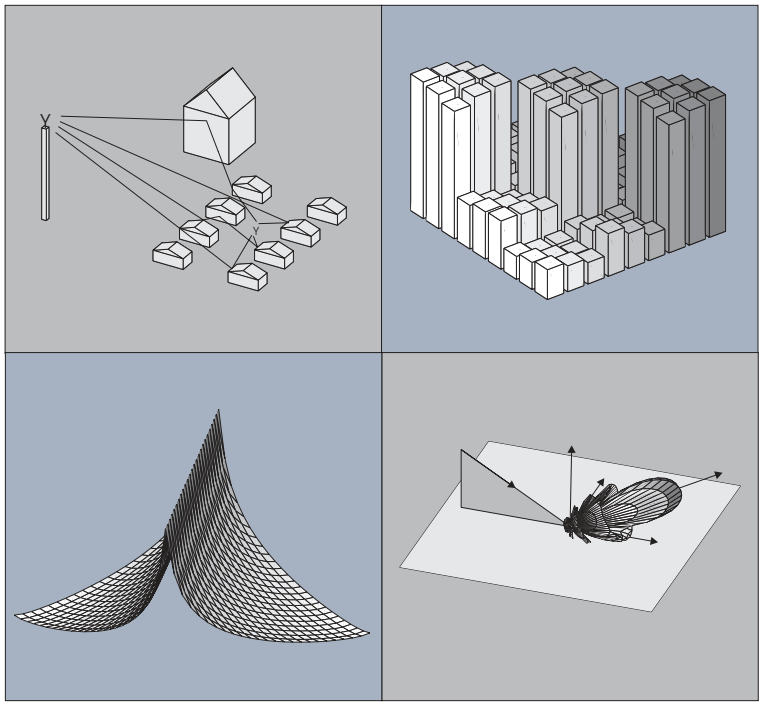
CHALMERS



Antennas and Propagation from a Signal Processing Perspective

Thomas Svantesson

Department of Signals and Systems
School of Electrical and Computer Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Sweden 2001

Thesis for the degree of Doctor of Philosophy

Technical Report No. 407

Antennas and Propagation from a Signal Processing Perspective

by

Thomas Svantesson

Department of Signals and Systems
School of Electrical and Computer Engineering
Chalmers University of Technology
S-412 96 Göteborg, Sweden

Göteborg 2001

Thomas Svantesson
Antennas and Propagation from
a Signal Processing Perspective.

This thesis has been prepared using LaTeX.

Doktorsavhandlingar vid Chalmers Tekniska Högskola
Ny serie nr 1703
ISBN 91-7291-019-4
ISSN 0346-718X

Technical Report No. 407
Signal Processing Group
Department of Signals and Systems
Chalmers University of Technology
SE-412 96 Göteborg, Sweden

Printed in Sweden by Chalmers Reproservice
Göteborg, May 2001

Abstract

This thesis studies several topics within the area of antennas and propagation from a signal processing perspective. However, theory and methods from electromagnetics and communications have also been used, contributing to an interdisciplinary character of the thesis. Several physical models are derived in the thesis that describes the wireless communications channel and several novel antennas. On the basis of these models, it is investigated how multiple antennas may be employed to increasing the capacity of future wireless communication systems. Furthermore, several aspects of using antenna arrays for finding the directions of arrival of electromagnetic waves, with applications in radar and acoustic sonar, are studied.

An often neglected issue when employing several antennas is that the elements affect each other through mutual coupling. By deriving expressions for the mutual coupling, the achievable direction finding performance is analyzed for this case. It is found that if the coupling is known, the effects on direction finding are small and can even increase performance in some cases. Similar results are also obtained when examining the capacity of communication systems employing multiple closely spaced antennas at both the transmitter and receiver.

To evaluate the performance of communication systems, a spatio-temporal channel model is proposed that is based on electromagnetic scattering and fundamental physics. By using a dyad notation and concepts from rough surface scattering, a compact formulation of the channel model is obtained. System models are then derived that employs multiple antennas at both the transmitter and receiver, so called Multi-Input Multi-Output (MIMO) systems. The polarization properties of the channel as well as those of the antennas are also included in the model, allowing for studies of different antenna arrangements.

Two novel antenna solutions are proposed that can be used as diversity receivers in MIMO systems or when performing high-resolution direction finding. By exciting higher order modes of biconical or microstrip antennas, several directionally dependent radiation patterns are obtained. Different patterns can also be obtained by employing parasitic elements. It is found that these antennas, the multimode and switched parasitic antenna, offer performance comparable to that of an antenna array.

Keywords: physical modeling, sensor array processing, mutual coupling, channel modeling, MIMO channels, diversity reception, channel capacity, higher order modes, switched parasitic antennas, parameter estimation.

Contents

Acknowledgments

Several individuals have played important roles in the course of writing this thesis. I would like to express my sincerest gratitude to the following persons.

First, I would like to thank my advisor Mats Viberg for all his help, encouraging words, and for contributing to the open-minded atmosphere in the department. In particular, he has encouraged me to pursue my own ideas (while critically scrutinizing them), for which I am very grateful.

Two other persons have meant a lot to me during these years, my room mates Dr. Tony Gustafsson and Lic. Magnus Lundberg. In the beginning, Tony really helped me get started and answered my many questions. Without you pushing, I probably never had submitted my first paper. Thank you! After Tony graduating, my new room mate Magnus, really took over where Tony left. He inspired me to achieve more than I ever thought possible. Thank you for all your excellent proof reading and sharing your expertise on deadlines.

Thanks also to all my former and present colleagues and friends at the department for making this a fun and interesting place to work at. I am also grateful to Lic. Fredrik Athley, Patrik Bohlin, Dr. Jonas Sjöberg, Dr. Tomas McKelvey, and Prof. Mats Viberg for their careful proof-reading of the manuscript.

I would also like to gratefully acknowledge the contributions of my co-authors, Lic. Mattias Wennström and Dr. Anders Ranheim (thanks for introducing me to Sunnmöre). The interactions have been fun, rewarding, and indeed very stimulating.

Prof. Anders Derneryd and Prof. Lars Josefsson at Ericsson Microwave Systems have also contributed to this thesis by sharing their deep knowledge in antenna and propagation issues, for which I am grateful. I am also grateful to Prof. Per-Simon Kildal for introducing me to the world of antennas and mutual coupling. Furthermore, I would like thank Dr. Christian Bergljung

and Dr. Peter Karlsson at Telia Research for providing me the opportunity to come to Telia Research in Malmö and perform channel measurements.

I also gratefully acknowledge the support from the Swedish Foundation for Strategic Research, under the Personal Computing and Communications (PCC) Program. These years within the PCC program have also been very rewarding as I have met a lot of good friends and perhaps future colleagues.

Last, but far from least, I would like to thank my parents and grandparents for always supporting me and doing their best to help me, no matter what.

Abbreviations and Acronyms

AIC	An Information theoretic Criterion
AP	Alternating Projection
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPR	Branch Power Ratio
CCDF	Complementary Cumulative Distribution Function
CDMA	Code Division Multiple Access
CMA	Constant Modulus Algorithm
CRB	Cramér-Rao lower Bound
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DML	Deterministic Maximum Likelihood
DOA	Direction Of Arrival
DOD	Direction Of Departure
EM	ElectroMagnetic
EMF	Electro-Motive Force
ESPRIT	Estimation of Signal Parameters via Rotational Invariance Techniques
FDD	Frequency Division Duplex
FEM	Finite Element Method
FIM	Fisher Information Matrix
GLRT	Generalized Likelihood Ratio Test
GTD	Geometrical Theory of Diffraction
HFSS	High Frequency Structure Simulator

LAN	Local Area Network
LOS	Line Of Sight
MEA	MultiElement Antenna
MB	Multiple-Bounce
MDL	Minimum Description Length
MDLC	Extended version of MDL that uses DML and AP
MFP	Matched Field Processing
MIMO	Multi-Input Multi-Output
ML	Maximum Likelihood
MODE	Method Of Direction Estimation
MOM	Method Of Moments
MRC	Maximum Ratio Combining
MSE	Mean Square Error
MUSIC	MUltiple SIgnal Classification
MVU	Minimum Variance Unbiased
NLOS	Non Line Of Sight
NSF	Noise Subspace Fitting
OTH	Over The Horizon
PAS	Power Azimuth Spectrum
PCC	Personal Computing and Communication
PE	Parasitic Element
PO	Physical Optics
RCS	Radar Cross Section
RMS	Root Mean Square
RMSE	Root Mean Square Error
SB	Single-Bounce
SIMO	Single-Input Multi-Output
SISO	Single-Input Single-Output
SML	Stochastic Maximum Likelihood
SNR	Signal to Noise Ratio
SPA	Switched Parasitic Antenna
SSF	Signal Subspace Fitting
STC	Space Time Coding
STBC	Space Time Block Coding

SVD	Singular Value Decomposition
TE	Transverse Electric
TEM	Transverse ElectroMagnetic
TM	Transverse Magnetic
UE	User Equipment
UMTS	Universal Mobile Telephone System
ULA	Uniform Linear Array
VHF	Very High Frequency
VOR	VHF Omni-directional Radio range
VS	Vector Sensor
WCDMA	Wideband Code Division Multiple Access
WLAN	Wireless Local Area Network
WSF	Weighted Subspace Fitting

Notation

In this thesis the following conventions are used. Vectors are written as boldface lower-case letters. Matrices are written as boldface upper-case letters. The meaning of the following symbols are, if nothing else is explicitly stated:

$\mathbb{R}^m,\ \mathbb{C}^m$	The set of real and complex-valued m-vectors, respectively.
$\mathcal{R}(\mathbf{A})$	The column space of the matrix $\mathbf{A}$.
$\mathsf{rk}\ (\mathbf{A})$	The rank of the matrix $\mathbf{A}$.
$\mathbf{A}^\dagger$	The Moore-Penrose pseudo-inverse of a $m \times n$ matrix $\mathbf{A}$. If $m \geq n$ and $\mathbf{A}$ is of full rank it is defined as $\mathbf{A}^\dagger = \left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H$.
$\mathbf{A}^H,\ \mathbf{A}^*,\ \mathbf{A}^T$	Complex conjugate transpose, Complex conjugate and Transpose operator, respectively.
$\mathbf{A}_{i,j}$	The $(i,j)^{th}$ element of $\mathbf{A}$.
$\mathbf{A}_{j,:}$	The j^{th} row of $\mathbf{A}$. Other submatrices of $\mathbf{A}$ are denoted with a similar MATLAB-like notation.
$\mathrm{Tr}(\mathbf{A})$	The trace operator.
$\|\mathbf{A}\|_F$	The Frobenius matrix norm $\|\mathbf{A}\|_F = \sqrt{\mathrm{Tr}(\mathbf{A}\mathbf{A}^H)}$.
$\mathbf{I}_m$	The $m \times m$ identity matrix. Subscript m is often omitted.

$\tilde{\mathbf{I}}$ The reverse permutation operator defined as

$$\tilde{\mathbf{I}} = \begin{bmatrix} 0 & & & 1 \\ & & \iddots & \\ & 1 & & \\ \iddots & & & \\ 1 & & & 0 \end{bmatrix}.$$

$\mathbf{0}_m$ An $m \times m$ matrix with zeros. Subscript m is often omitted.

$\mathbf{1}_k$ A vector of zeros except for a one in the k^{th} position. The total length of the vector is given by the context.

$\boldsymbol{\Upsilon}_k$ A banded symmetric Toeplitz matrix with zeros everywhere except on the k^{th} and the $-k^{th}$ subdiagonal which consists of ones

$$\boldsymbol{\Upsilon}_k = \begin{bmatrix} \mathbf{0}_k & 1 & & \mathbf{0}_{n-k-1} \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ \mathbf{0}_{n-k-1} & & 1 & \mathbf{0}_k \end{bmatrix}.$$

$\text{vec}(\mathbf{A})$ The vectorization operator. Stacks the columns of $\mathbf{A}$ into a vector, i.e. when $\mathbf{A}$ has n columns:

$$\text{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{A}_{:,1} \\ \vdots \\ \mathbf{A}_{:,n} \end{bmatrix}.$$

$\mathbf{A} > \mathbf{0}$ The matrix $\mathbf{A}$ is positive definite. Furthermore, $\mathbf{A} \geq \mathbf{0}$ means that $\mathbf{A}$ is positive semidefinite, and $\mathbf{A} - \mathbf{B} > \mathbf{0}$ means that the matrix difference $\mathbf{A} - \mathbf{B}$ is positive definite.

$\mathbf{A} \odot \mathbf{B}$ The Hadamard (or Schur) product, i.e. element-by-element multiplication.

$\mathbf{A} \otimes \mathbf{B}$ The Kronecker product. When $\mathbf{A}$ is $m \times n$ and $\mathbf{B}$ is $r \times s$, then

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} \mathbf{A}_{1,1}\mathbf{B} & \mathbf{A}_{1,2}\mathbf{B} & \cdots & \mathbf{A}_{1,n}\mathbf{B} \\ \mathbf{A}_{2,1}\mathbf{B} & \mathbf{A}_{2,2}\mathbf{B} & \cdots & \mathbf{A}_{2,n}\mathbf{B} \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_{m,1}\mathbf{B} & \mathbf{A}_{m,2}\mathbf{B} & \cdots & \mathbf{A}_{m,n}\mathbf{B} \end{bmatrix}$$

is of dimension $mr \times ns$.

$\mathbf{A}^{-T}, \mathbf{A}^{-H}$	$(\mathbf{A}^{-1})^T$ and $(\mathbf{A}^{-1})^H$.
δ_{ij}	The Kronecker delta function.
$\widehat{(\cdot)}$	An estimated quantity.
$\vec{(\cdot)}$	Geometrical vector, i.e. a vector of length 3. Typically, the spherical basis vectors are written as $\vec{r}$, $\vec{\phi}$, and $\vec{\theta}$.
$\lceil x \rceil$	Rounding off x to the nearest integer towards infinity.
$\arg\min_x f(x)$	The minimizing argument of the function $f(x)$.
$E\{\mathbf{x}\}$	Statistical expectation of a random vector $\mathbf{x}$.
ϵ	The permittivity.
μ	The permeability.
η	The intrinsic impedance is defined as $\eta = \sqrt{\mu/\epsilon}$. For free space the intrinsic impedance becomes $120\pi\ \Omega \approx 377\ \Omega$.
k	The wave number $k = \frac{2\pi}{\lambda} = \omega\sqrt{\mu\epsilon}$, where λ is the wavelength.
$\vec{k}$	The wave vector, i.e. the direction which the wave travels. Note that $\lvert\vec{k}\rvert = k$.
$\vec{\rho}$	The polarization vector.
$\vec{\mathcal{E}}, \vec{\mathcal{H}}$	The instantaneous values of the electric and magnetic field.
$\vec{E}, \vec{H}$	The complex form of the electric and magnetic field with the time harmonic variation suppressed.
$\vec{\mathbf{S}}$	The scattering dyad that is defined as the juxtaposition $\vec{ab}$ of the vectors $\vec{a}$ and $\vec{b}$, with no dot product or cross product between them. In matrix form this corresponds to the outer product of two vectors $\mathbf{a}$ and $\mathbf{b}$, i.e. $\mathbf{ab}^T$ (assuming column vectors).

Chapter 1

Introduction

This thesis deals with signal processing topics in the area of antennas and propagation. A number of problems within this diverse area are analyzed using theory and methods from signal processing but also concepts from antennas and propagation, communications and information theory. In fact, this interdisciplinary character is a major theme and contribution of the thesis, since most of the literature on the analyzed topics do not make this connection. Most of the chapters in this thesis are rather self-contained in order for the chapters to be read with some appreciation by readers with knowledge from different research fields. This introductory chapter will briefly introduce the different areas, that will be analyzed. Furthermore, a few guidelines in reading this thesis will be given together with an outline and a list of the publications by the author. A reader already familiar with the topics may skip this introductory chapter and proceed directly to the following chapters.

1.1 Background

Studying signal processing often means that a sensor of some kind is employed. For instance, in a bicycle computer, a magnetic sensor is employed to get a signal that is used to calculate the instantaneous speed, average speed, etc. In this calculation, a simple physical model, that describes how the received signal depends on the speed of the wheel, is needed. The algorithm that calculates the speed from the measured signal is then based upon this model. In this simple example, it is relatively easy to obtain a model of the measured signal

When devising signal processing schemes for signals that are transmitted

and received via radio, the situation is considerably more complex. The signal that is transmitted will interact with the surroundings of the transmitter and receiver, i.e. the channel, and the signal that finally reaches the receiver will typically be severely distorted. Often, the signal is reflected by tall buildings, lamp posts, and other objects. The received signal will consist of many contributions from different propagation paths, i.e. multipaths. The type of antenna will greatly affect this interaction with the surroundings. Hence, when devising signal processing schemes for radio, the properties of the antenna need to be considered. Furthermore, if the transmitted signal bears some special properties, as it does in wireless communications, this should to be accounted for by the signal processing schemes. However, also signals transmitted in radar and underwater acoustics are designed to exhibit special properties that are exploited in the receiver. Therefore, wireless communications as well as radar applications are of necessity interdisciplinary, since it involves signal processing, antenna and propagation, communications and information theory.

Unfortunately, much of the research in wireless communications and radar is performed within the respective research field. By combing the knowledge of the different fields, there is a large potential of devising systems with better overall performance. For instance, most signal processing algorithms for estimating the direction to aircrafts, do not explicitly take antenna issues into account. In a sense, this thesis attempts to bridge this gap by including ideas and concepts from electromagnetics, physics, communications, and signal processing.

Several different topics are studied in this thesis. First, the impact of mutual coupling between different antennas in direction finding applications, such as radar, is studied in several chapters. Therefore, the fundamental principles of array antennas and direction finding will be introduced in this chapter. Furthermore, two novel methods for direction finding employing higher order modes of microstrip and biconical antennas will be studied. The direction finding part of the thesis ends with an interesting antenna solution for high-resolution direction finding, a switched parasitic antenna, that offers several interesting advantages compared to traditional array antennas.

The remaining part of the thesis deals with wireless communications, and in particular modeling of the wireless channel. A spatio-temporal channel model is derived based on fundamental physics and several channel properties are highlighted. A very brief background on channel modeling and wireless communications is therefore also provided in this introductory chapter. Some focus is put on the recently introduced concept of employing multiple antennas at both the transmitter and the receiver to substantially increase the data rate. Thus, some introductory material on this subject will also be

given.

Some guidelines in reading the thesis and the organization of it are also provided. Finally, a list of the publications by the author is given.

1.2 Personal Computing and Communication

The Personal Computing and Communication (PCC) program was started in 1997, within the Swedish Foundation for Strategic Research (SSF). When forming the PCC, the aim was to perform research in areas of strategic importance for Swedish industry, while encouraging co-operation between different research disciplines. The activities of the PCC may be summarized in the PCC vision *Personal multimedia communication to all at the same cost as fixed telephony today* [Mol98a, Mol98b]. PCC is a cooperation between Chalmers University of Technology (CTH), Royal Institute of Technology (KTH) and Lund Institute of Technology (LTH). Other universities participating in the program are Uppsala University, and Luleå University of Technology.

The focus of bringing together researchers from different fields, has increased the awareness for the research problems in neighboring research areas. In particular, this thesis work has benefitted from this interdisciplinary nature of PCC with interaction of researchers from electromagnetics, communications, and signal processing. The PCC program consists of several research projects, of which this thesis is part of the project *Adaptive Antennas in Wideband Radio Access Networks*. The scope of the project includes several topics in the fields of wave propagation and channel modeling, analog and digital electronics, signal processing, communication theory, and system analysis. In fact, these topics relatively well summarize the thesis which deals with many of these subjects.

1.3 Physical Modeling of Array Antennas

The use of array antennas is an important theme of this thesis. Therefore, some background material on physical modeling of antennas and array antennas in particular, will be provided in this section. These models will be used throughout the thesis, although most frequently in the direction finding applications. Hence, the presentation of this material is made with this application in mind.

Direction finding using an array of sensors is of great importance in a variety of applications, such as radar, sonar, communications, and recently

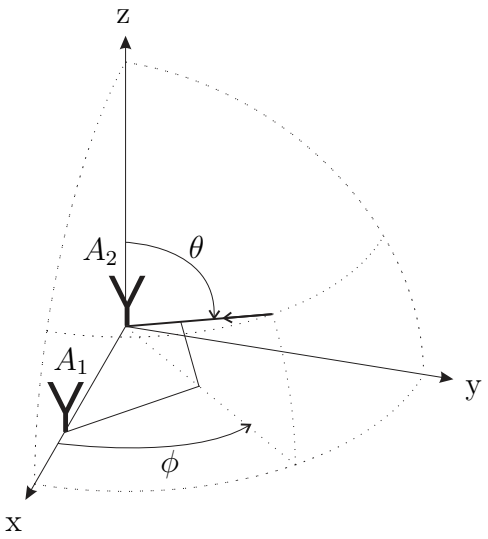


Figure 1.1: The geometry of the array and the relation to the coordinates.

also personal locating services. In these applications, the Direction Of Arrival (DOA) of a wave that is incident upon the array is calculated using a model for the measured voltages. Hence, obtaining an accurate model in these applications is critical. However, also when modeling the radio channel, the properties of the antenna are important. For instance, interaction of the different antenna elements and their radiation patterns greatly impacts the performance of most communication systems. Therefore, the fundamental properties of antennas are briefly reviewed in this section and a simple model based on fundamental electromagnetics is derived.

The principal function of an antenna is to convert an electromagnetic wave into an induced voltage or current that can be measured. If the antenna consists of several elements, a number of voltages or currents are measured. The physical principle that governs DOA estimation is that an incident wave will reach each antenna element at different time instants. A typical scenario is shown in Figure 1.1, where a wave is incident from the (θ, ϕ) direction upon an array of two elements. Here, the wave reaches antenna A1 some time before antenna A2. By exploiting a model for the measured signals at each antenna element it is possible to calculate the DOA from those signals.

Consider first a special case where the wave consist of a single sinusoidal, i.e., a time-harmonic field [Bal89]. For this case, Maxwell's equations can be written in a much simpler form, and therefore most of the antenna literature uses time-harmonic fields. Of course, any signal can be written as a sum

of sinusoidal signals, and thus any signal can be represented using time-harmonic fields. The time harmonic variation of the form $e^{j\omega t}$ is suppressed by using complex quantities and the instantaneous values can be found by taking the real part of the product of the quantity and $e^{j\omega t}$. For example, the instantaneous value of a field at point r_p becomes

$$\vec{\mathcal{E}}(r_p) = Re[\vec{E}(r_p)e^{j\omega t}]. \tag{1.1}$$

If the source of the incident wave is located sufficiently far away from the array, the wave front can be considered plane. This is referred to as the *far-field* case. The field at a point r_p of a plane electromagnetic wave propagating along the $\vec{k}$ direction[1] can then be written as

$$\vec{E}_{wave}(r_p) = \vec{\rho}se^{-j\vec{k}\cdot\vec{r}_p}, \tag{1.2}$$

where $\vec{\rho}$ denotes the polarization of the incident field and s the strength of the field. Note that $\vec{r}_p$ denotes the vector from the origin to the point r_p and thus the phase reference point is located at the center of the coordinate system. This type of wave is usually called a uniform plane wave, i.e. the amplitude, phase, and direction of the field is the same over a plane orthogonal to the propagation direction. For more details regarding wave propagation, see [Sau99] that treats propagation mechanisms and antenna fundamentals from a wireless communications perspective.

The action of the antenna is to convert the field into a voltage or current that can be measured. How different types of antennas does this conversion is one of the fundamental topics of antenna theory. For an in-depth discussion about the properties of different antenna types, see the following classical antenna textbooks [Bal82, Col85, Ell81, Kra88, ST98]. However, most antenna textbooks focus on transmitting properties and not on receiving modes. A recent compact summary that includes receiving properties can be found in [Mey00], and in [Kil99] equivalent circuits for antennas in receive mode are presented. In the communication literature, the impact of the antenna is usually neglected, however in [Jak74] there is a brief introduction to antennas and propagation that can be useful.

An expression for the induced voltage can easily be derived using the concept of reciprocity and knowledge of the far-field radiation function of the receiving antenna $\vec{G}(\vec{r})$, where $\vec{r}$ is the direction of radiation. Using the

[1]The direction of the wave vector $\vec{k}$ is the direction in which the wave travels and the magnitude $k = |\vec{k}|$ corresponds to the wave number $k = \frac{2\pi}{\lambda} = \omega\sqrt{\mu\epsilon}$, where λ is the wavelength. The symbols ϵ and μ denote the dielectric constant and the permeability of the medium respectively.

results from [Mey00, Kil99], the induced voltage becomes

$$V = c\ \vec{G}(-\vec{k}) \cdot \vec{E}_{inc}, \tag{1.3}$$

where c is a scalar defined as

$$c = -j\frac{4\pi}{k\eta I_{in}}\,. \tag{1.4}$$

Here, k is the wave number, η the intrinsic impedance[2] of the propagation medium, I_{in} the input current used when calculating $\vec{G}$, and $\vec{E}_{inc}$ the incident field. Note that here $-\vec{k}$ denotes the direction from which the wave is incident. The expression for the induced voltage is, thus, obtained once the radiation function of the antenna is obtained. For most common antenna types, the radiation function $\vec{G}$ (or approximations of it) can be found in antenna textbooks [Bal82, Col85, Ell81, Kra88, ST98].

By combining (1.2) and (1.3), the voltage can be written as

$$V = c\ \vec{G}(-\vec{k}) \cdot \vec{\rho}\ e^{-j\vec{k}\cdot\vec{r}_p} s = H(\theta,\phi)\ e^{-j\vec{k}\cdot\vec{r}_p} s, \tag{1.5}$$

where the directional properties of the antenna and the polarization of the incident wave are combined into the scalar $H(\theta,\phi)$. Note that $H(\theta,\phi)$ corresponds to the response of the antenna to a sinusoidal signal at the carrier frequency, i.e. the frequency domain version of the antenna impulse response at the carrier frequency. Remember that the above voltage is written using the complex signal representation. The instantaneous value of the induced voltage is obtained in a manner similar to (1.1) as

$$\mathcal{V}(t) = \mathrm{Re}\left[sH(\theta,\phi)\ e^{-j\vec{k}\cdot\vec{r}_p} e^{j\omega t}\right] = |sH(\theta,\phi)|\cos\left(\omega t - \omega\tau + \angle sH(\theta,\phi)\right), \tag{1.6}$$

where $\tau = \vec{k}\cdot\vec{r}_p/\omega$ is a time delay that corresponds to the time it takes for the wave to travel from the phase reference center, i.e. the origin, to the point r_p. Typically, the phase reference point is located at one antenna element and thus the time delay τ corresponds to the time needed by the wave to propagate over the array between the current element and the reference element.

It is important to remember here that the preceding analysis was based on a single harmonic, i.e. a single frequency. In a practical application, the signal will span a band of frequencies and the field strength will be time-varying. However, if the time delays τ are small compared to the reciprocal

[2]The intrinsic impedance is defined as $\eta = \sqrt{\mu/\epsilon}$. For free space the intrinsic impedance becomes $120\pi\ \Omega \approx 377\ \Omega$.

of the bandwidth of the signal, the instantaneous value of the induced voltage can be written as

$$\mathcal{V}(t) \approx |s(t)H(\theta,\phi)| \cos\left(\omega t - \omega\tau + \angle s(t)H(\theta,\phi)\right). \tag{1.7}$$

This approximation is usually called the narrowband assumption in array signal processing [KV96], and essentially means that the signal $s(t)$ does not vary over the array, i.e. $s(t) \approx s(t+\tau)$. In order to arrive at (1.7) it was also assumed that the response of the antenna element, $H(\theta,\phi)$, does not change significantly over the frequency band of interest. This is a reasonable assumption in almost all modern communication systems where the relative bandwidth usually is a fraction of a percent. Therefore, in the light of (1.7) and the narrowband assumption, it is possible to use the complex signal representation of the voltage also for the time-varying case

$$V(t) = H(\theta,\phi)\ e^{-j\vec{k}\cdot\vec{r}_p} s(t). \tag{1.8}$$

This complex representation greatly simplifies the analysis, and will be used throughout the thesis. A more detailed discussion on the above approximations may be found in [SM97].

When estimating the DOA using an array of n antenna elements, n different voltages will be measured. If these voltages are collected in a vector $\mathbf{x}(t)$, the resulting model becomes

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} H_1(\theta,\phi)\ e^{-j\vec{k}\cdot\vec{r}_1} \\ H_2(\theta,\phi)\ e^{-j\vec{k}\cdot\vec{r}_2} \\ \vdots \\ H_n(\theta,\phi)\ e^{-j\vec{k}\cdot\vec{r}_n} \end{bmatrix} s(t) = \mathbf{a}(\theta,\phi)s(t), \tag{1.9}$$

where $H_i(\theta,\phi)$ denotes the response of antenna element i and $\vec{r}_i$ the location vector from the phase reference center to antenna element i. The $n \times 1$ vector $\mathbf{a}(\theta,\phi)$ models the spatial response of the array due to an incident plane wave from the (θ,ϕ) direction. Therefore, the vector $\mathbf{a}(\theta,\phi)$ is usually called steering vector or array response vector in the array processing literature.

Example Uniform Linear Array of Dipoles

Consider a plane wave incident upon an array of 3 thin dipoles of length l as illustrated in Figure 1.2. The dipoles are oriented along the z axis and spaced equidistantly along the x axis with a separation distance of d. Such an arrangement is referred to as a Uniform Linear Array

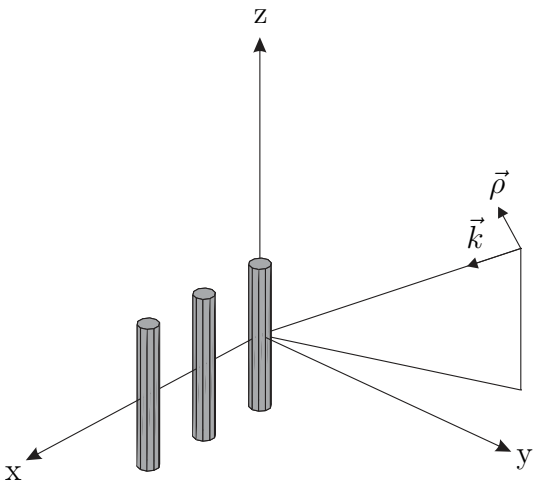


Figure 1.2: The geometry of the dipole array and a plane wave incident along $\vec{k}$ from the (θ, ϕ) direction.

(ULA). In this case, the wave will reach the element at the origin first and after some time the other elements. The time delay for antenna element i, where $i = 0, 1, 2$ and 0 corresponds to the origin, is easily calculated as

$$\tau_i = \frac{\vec{k} \cdot \vec{r}_i}{\omega} = \frac{k}{\omega} \left(\sin\theta \ (\vec{x}\cos\phi + \vec{y}\sin\phi) + \vec{z}\cos\theta\right) \cdot \vec{x}di = \frac{kdi}{\omega} \sin\theta\cos\phi. \tag{1.10}$$

Before the full expression for the induced voltages over the array can be obtained, the directional properties of the dipole element must be found. The far-field radiation function $\vec{G}$ of the dipole element is included in almost any antenna handbook [Bal82, Col85, Ell81, Kra88, ST98] as

$$\vec{G} = \vec{\theta} \frac{j\eta I_{in}}{2\pi \sin\left(\frac{kl}{2}\right)} \left[\frac{\cos\left(\frac{kl}{2}\cos\theta\right) - \cos\frac{kl}{2}}{\sin\theta}\right] \tag{1.11}$$

Now, the antenna response $H(\theta, \phi)$ can be found using (1.4) and (1.5) as

$$H(\theta,\phi) = c\vec{G}(-\vec{k}) \cdot \vec{\rho} = \frac{\lambda}{\pi \sin\frac{kl}{2}} \left[\frac{\cos\left(\frac{kl}{2}\cos\theta\right) - \cos\frac{kl}{2}}{\sin\theta}\right], \tag{1.12}$$

where $\vec{\rho} = -\vec{\theta}$ as indicated in Figure 1.2. The final array model is then obtained by using (1.5) and (1.12) as

$$\mathbf{x}(t) = \frac{\lambda}{\pi \sin\frac{kl}{2}} \left[\frac{\cos\left(\frac{kl}{2}\cos\theta\right) - \cos\frac{kl}{2}}{\sin\theta} \right] \begin{bmatrix} 1 \\ e^{-j\psi} \\ e^{-j2\psi} \end{bmatrix} s(t) = \mathbf{a}(\theta,\phi)s(t). \tag{1.13}$$

Note that the phase shifts $\psi = kd\sin\theta\cos\phi$ (often called electrical angle) only depend on the element positions and not on the type of antenna element. The impact of the antenna element is included in the antenna response $H(\theta,\phi)$. In this example, the elements are identical and the antenna response is simply a scalar that could be absorbed into the signal $s(t)$. If the antenna response of each element is different, it should be incorporated into the steering vector $\mathbf{a}(\theta,\phi)$, as implied by (1.9). The steering vector for a dipole array was calculated in this example, but the steering vector for most antenna elements can be found in a similar manner by using formulas of the radiation function $\vec{G}$ from antenna textbooks. A more detailed derivation of the array response of a dipole array, that also includes the effects of mutual coupling, may be found in [Sva99b] and in Chapter 2.

□

The derivation of the measured voltages considered only one incident wave. When several waves are incident upon the array, the superposition principle can be applied if the antenna elements and the receiver are linear. To simplify the notation in the following analysis, the waves are assumed to arrive in the x-y plane ($\theta = 90°$), see Figure 2.1. The model when p uniform plane waves are incident upon an array of n elements can thus be written as

$$\mathbf{x}(t) = \sum_{l=1}^{p} \mathbf{a}(\phi_l)s_l(t) = \mathbf{A}(\boldsymbol{\phi})\mathbf{s}(t), \tag{1.14}$$

where

$$\mathbf{A}(\boldsymbol{\phi}) = \begin{bmatrix} \mathbf{a}(\phi_1) & \mathbf{a}(\phi_2) & \cdots & \mathbf{a}(\phi_p) \end{bmatrix} \tag{1.15}$$

$$\mathbf{s}(t) = \begin{bmatrix} s_1(t) & s_2(t) & \cdots & s_p(t) \end{bmatrix}^T. \tag{1.16}$$

Note that the vector of measured voltages, at time t, $\mathbf{x}(t)$ is $n \times 1$, the steering matrix $\mathbf{A}(\boldsymbol{\phi})$ is $n \times p$, and the signal vector $\mathbf{s}(t)$ is $p \times 1$. The DOAs are contained in the $p \times 1$ parameter vector $\boldsymbol{\phi}$.

In all measurement situations, noise will inevitably appear and it is typically difficult to model. Essentially, noise represents everything that do not obey the assumed model and here it is included as an additive term as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\phi})\mathbf{s}(t) + \mathbf{n}(t). \tag{1.17}$$

There are many sources of noise [Sko62], for example environmental noise such as cosmic noise, atmospheric absorption noise, solar noise etc. It can further be man-made noise, such as jammers and power tools etc. The receiver will also generate some noise such as thermal noise, shot noise and flicker noise. Often, receiver noise is the dominating noise source; and then additive white Gaussian noise is a good model. In this thesis, the noise is, if not stated otherwise, assumed to be both spatially and temporally white and Gaussian distributed,

$$E\{\mathbf{n}(t)\} = 0 \quad E\{\mathbf{n}(t)\mathbf{n}^H(s)\} = \sigma^2\mathbf{I}\ \delta_{ts} \quad E\{\mathbf{n}(t)\mathbf{n}^T(s)\} = \mathbf{0}\ \forall\ t, s. \tag{1.18}$$

The motivation for this assumption is that if there are many sources of noise, the sum will be Gaussian distributed according to the central limit theorem [Pap91]. Also, the further analysis of direction finding performance is greatly simplified by assuming white Gaussian noise. Therefore, an overwhelming part of the literature of array signal processing assumes white Gaussian noise. For other noise models, such as colored noise or noise of other distributions, see [Hay95b].

The transmitted signal is in a similar manner often assumed to be Gaussian distributed with

$$E\{\mathbf{s}(t)\} = 0 \quad E\{\mathbf{s}(t)\mathbf{s}^H(s)\} = \mathbf{P}\ \delta_{ts} \quad E\{\mathbf{s}(t)\mathbf{s}^T(s)\} = \mathbf{0}\ \forall\ t, s. \tag{1.19}$$

Note that different source signals may be correlated, leading to a non-diagonal source covariance matrix $\mathbf{P}$. The assumption of temporally white and Gaussian signals is not critical. It is only used when deriving the maximum likelihood estimator and the Cramér-Rao lower bound. The actual performance of the methods under study is typically independent on the source distribution, but depends only on $\mathbf{P}$ [OVSN93], unless the distribution or the signal itself is known. For other choices of signal statistics, see [Hay95b]. A crucial assumption is, however, that the signal is uncorrelated with the noise.

Most DOA estimation schemes rely on the properties of the second order moment of the measured voltages $x(t)$, i.e. the spatial covariance matrix. Therefore, some characteristics of the spatial covariance matrix of the measured voltages, $\mathbf{x}(t)$, are presented next. The latter is obtained as

$$\mathbf{R} = E\{\mathbf{x}(t)\mathbf{x}(t)^H\} = \mathbf{A}E\{\mathbf{s}(t)\mathbf{s}^H(t)\}\mathbf{A}^H + E\{\mathbf{n}(t)\mathbf{n}^H(t)\} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \sigma^2\mathbf{I}. \tag{1.20}$$

Note that the cross-terms vanish since the signal and the noise are assumed to be uncorrelated and the noise is zero-mean.

Many of the DOA estimation schemes presented in the literature rely on the fact that the eigendecomposition of the covariance matrix $\mathbf{R}$ can be written as a sum of two parts [KV96]. One part consisting of eigenvectors corresponding to eigenvalues equal to the noise variance, and a second part that is related to the signal

$$\mathbf{R} = \sum_{k=1}^{n} \lambda_k \mathbf{e}_k \mathbf{e}_k^H = \mathbf{E}_s \mathbf{\Lambda}_s \mathbf{E}_s^H + \sigma^2 \mathbf{E}_n \mathbf{E}_n^H, \tag{1.21}$$

where $\mathbf{E}_s = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$ denotes the signal eigenvectors, $\mathbf{E}_n = [\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n]$ denotes the noise eigenvectors, and $\mathbf{\Lambda}_s = \text{diag}[\lambda_1, \ldots, \lambda_p]$ denotes the signal eigenvalues. This decomposition, which some times is referred to as spectral factorization, and is used extensively in subspace based DOA estimation [KV96, VS01].

Remarks

1. The model derived here for the case of an electromagnetic wave incident upon an antenna array is closely connected to the sonar problem where an acoustic wave is incident upon an hydrophone array [Lig73, Ows85]. Applications of similar models for medical imaging can be found in [FK92] and also for chemical sensor arrays in [NPP95].

2. If the antenna is located close to the source of radiation, the wave front is spherical and the plane wave assumptions is no longer valid. However, if the spherical nature of the wave is correctly modeled, the range as well as the DOA can be estimated [SN94].

3. Only direction finding using an antenna array is considered in this introduction, but other antenna arrangements are also used. For instance, a mechanically steered antenna [Sko62] has been used in radar for a long time. Other possibilities are to employ electromagnetic vector-sensors [NP97], switched parasitic elements [SW01], or higher order electromagnetic modes [Sva00e]. Higher order modes and switched parasitic antennas are analyzed in Chapter 7 and 8 of this thesis.

4. The linear array, that only can be used to estimate either azimuth or elevation, is by far the most analyzed type of array. However, two dimensional arrays have also been analyzed. See [Hay95b, He93] and the

references therein for results on two dimensional arrays, and in particular for circular arrays that are used to estimate both the elevation and the azimuth angles.

5. The expression for the steering vector was calculated analytically in this section. In general, these expressions will not correspond to a measured response of an antenna array. Typically, differences in temperature, aging of components, changes in the electromagnetic environment of the array will change the array response. Some type of calibration, of both the antenna system as well as the receiver, is therefore usually needed. Often the response of the array is measured and stored in a lookup table. For an example of the calibration procedures of a high performance digital beamforming antenna, see [PDS97]. An important topic here is mutual coupling that will affect the performance. This is discussed in detail in Chapters 2-6.

□

1.4 Direction Finding

The problem of estimating the Direction Of Arrival (DOA) of multiple signals arriving at an array of sensors has received considerable attention for several decades [KV96, VS01, VB88]. The underlying physical principle of all antenna array DOA algorithms is the fact that a transmitted signal will arrive at the different elements at different time instants. By exploiting the data model derived in previous section, the directions from which the signals arrive can be obtained. Traditionally, the most notable application has been source localization in radar and sonar. These applications have essentially driven the development of DOA algorithms since the first approaches of spatial filtering or beamforming. The conventional beamformer dates back to the second world-war, and is a mere application of Fourier-based spectral analysis to spatio-temporally sampled data. Later, adaptive beamformers and classical time-delay estimation techniques were applied to enhance one's ability to resolve closely spaced signal sources [App76, Cap69, Gab80]. However, these approaches suffer from the fact that the performance directly depends upon the physical size of the array (aperture). The limitations of the beamformer started a development of other algorithms that alleviated some of these drawbacks. Inspired by the maximum likelihood principles [Kay93, MS69, Sch68], a wealth of DOA estimation schemes based on parametric modeling have appeared during the last two decades. By more fully

exploiting the underlying physical principles (model), a much higher performance than that of the beamformer was now possible.

Estimating the DOA is also of interest in wireless communications. The angular properties of the mobile communications channel impacts many parts of the system design. For instance, the angular spread of the channel essentially determines the necessary antenna spacing in diversity applications. It has further been proposed to employ antenna arrays to reduce the co-channel interference by transmitting energy only in the direction to a specific user and essentially no energy in the directions of other users. In these types of systems, estimating the DOA forms an integral part of the system. The DOA parameters may also offer a more stable characterization of the channel than a direct filter description. Furthermore, as the mobile phone becomes more ubiquitous, the interest of employing the phone for personal locating services increases. Here, the DOA can be used for obtaining the location of the mobile phone. Thus, there are many reasons for employing DOA estimation also in wireless communications.

Most of these algorithms make the assumption that the signal essentially emanates from a single point. This point source assumption significantly simplifies the estimation problem, and the data model derived in the previous section is applicable. This is a reasonable assumption in radar and macrocellular communications, where the receiver antenna is located relatively far from the source. However, for some applications, such as micro- and pico-cell applications, the received energy can not be considered to emanate from a single direction [KRB00]. In these applications, the source is better modeled as a spread source [VS01]. Here, only the simpler case of point sources will be briefly reviewed. For more material, see [God97, He93, Hay95a, KV96, VS01] and further references therein.

1.4.1 Spectral-Based Techniques

As the name suggest, spectral based methods rely on calculating a spatial spectrum and finding the DOAs as the location of peaks in the spectrum. These methods were the first that were developed and are easy to apply. Probably the most widely used method of obtaining estimates of the DOAs is the beamforming method. As the name beamforming suggests, the received energy is focused to one direction (or beam) at a time. This can be expressed as

$$y(t) = \mathbf{w}^H \mathbf{x}(t), \tag{1.22}$$

where the weighting vector $\mathbf{w}$ can be seen as a spatial filter that emphasizes one particular direction. Given samples $y(1), y(2), \ldots, y(N)$, the output

power is measured by

$$P(\mathbf{w}) = \frac{1}{N}\sum_{t=1}^{N}|y(t)|^2 = \frac{1}{N}\sum_{t=1}^{N}\mathbf{w}^H\mathbf{x}(t)\mathbf{x}^H(t)\mathbf{w} = \mathbf{w}^H\hat{\mathbf{R}}\mathbf{w}, \tag{1.23}$$

where $\hat{\mathbf{R}}$ is obtained as

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{x}(t)\mathbf{x}^H(t) \tag{1.24}$$

which is an unstructured maximum likelihood estimate of the covariance matrix in (1.20). Many different choices of the weighting vector $\mathbf{w}$ can be made leading to different properties of the beamforming schemes [Sko62, VB88].

Conventional Beamforming If the weighting vector is chosen in order to maximize the received power in a certain direction ϕ [VB88], as

$$\mathbf{w}_{BF} = \frac{\mathbf{a}(\phi)}{\sqrt{\mathbf{a}^H(\phi)\mathbf{a}(\phi)}}, \tag{1.25}$$

the classical spatial spectrum is obtained

$$P_{BF}(\phi) = \frac{\mathbf{a}^H(\phi)\hat{\mathbf{R}}\mathbf{a}(\phi)}{\mathbf{a}^H(\phi)\mathbf{a}(\phi)}. \tag{1.26}$$

This spectrum is in array processing literature referred to as conventional (or Bartlett) beamformer, since this actually is a natural extension of the classical Fourier based spectral analysis with different window functions [Bar48, SM97]. In fact, if a Uniform Linear Array (ULA) of isotropic elements is used, the spatial spectrum in (1.26) is a spatial analog of the classical periodogram in time series analysis. Note that other types of arrays correspond to non-uniform sampling schemes in time-series analysis. As with the periodogram, the spatial spectrum has a resolution threshold. Waves arriving with electrical angle separation[3] less than $2\pi/n$ can not be resolved with this method. For example, using a five element ULA with an element separation of $d = \lambda/2$ results in a resolution threshold of 23°. Other choices of weighting vectors $\mathbf{w}$, that result in lower resolution thresholds, were therefore investigated.

[3]The electrical angle is defined as $kd\ cos\phi$.

Capon's Beamformer One of the most popular beamforming methods that to some extent alleviates the limitations of the conventional beamformer is the Capon's beamformer [Cap69, Lac71]. This beamformer attempts to minimize the power contributed by noise and any signals coming from other directions than the looking direction, while maintaining a fixed gain in the look direction. This type of beamformer is sometimes also referred to as Minimum Variance Distorsionless Response (MVDR) filter in acoustics literature. The Capon weighting vector is

$$\mathbf{w}_{CAP} = \frac{\hat{\mathbf{R}}^{-1}\mathbf{a}(\phi)}{\mathbf{a}^H(\phi)\hat{\mathbf{R}}^{-1}\mathbf{a}(\phi)} \tag{1.27}$$

and if inserted into (1.23), the MVDR spatial spectrum becomes

$$P_{CAP}(\phi) = \frac{1}{\mathbf{a}^H(\phi)\hat{\mathbf{R}}^{-1}\mathbf{a}(\phi)}. \tag{1.28}$$

Although more complex than the conventional beamformer, Capon's method offers significantly reduced resolution threshold. The lower resolution threshold (reduced spectral leakage) is achieved at the cost of reduced noise suppression capability [Cap69]. A formula for calculating the resolution threshold for the Capon's method assuming a perfect estimate of $\hat{\mathbf{R}}$ is given in [Ser95], see Table 1.1 on page 17.

A large number of alternative methods for beamforming have been presented in the literature, see [VB88] for an overview. An example of an application of conventional beamforming in a mobile communication scenario can be found in [MLEP+00] where the performance gain by exploiting conventional beamforming in a GSM network is investigated.

Still the resolution threshold for beamforming methods is quite high and that was one of the motivations for the interest in the so-called subspace methods that will be described next.

MUSIC Subspace based methods rely on observations regarding the eigendecomposition of the covariance matrix into a signal subspace and a noise subspace, as discussed in previous subsection. One of the most popular subspace methods, MUltiple SIgnal Classification (MUSIC), was introduced in [Sch79]. The method is based on the observation in (1.21), that the noise eigenvectors are perpendicular to the steering matrix or the signal subspace. The algorithm calculates the noise subspace using an eigendecomposition of the estimated covariance matrix in (1.24). Then, the estimates of the DOAs are taken as those ϕ that gives in the smallest value of $\mathbf{a}^H(\phi)\hat{\mathbf{E}}_n$, i.e. the

values that result in a steering vector furthest away from the noise subspace. Usually this is formulated as finding the p largest peaks in the "MUSIC spectrum"

$$P_{MU}(\phi) = \frac{1}{\mathbf{a}^H(\phi)\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\mathbf{a}(\phi)}. \tag{1.29}$$

Note that the eigenvectors $\hat{\mathbf{E}}_n$ are easily obtained by either an eigendecomposition of the sample covariance matrix, or a Singular Value Decomposition (SVD) of the data matrix, and that numerical reliable routines for eigendecomposition and SVD are included in most software packages.

The main motivation for introducing the subspace methods was to reduce the resolution threshold of the beamforming methods. The resolution threshold of MUSIC depends in a complicated manner on several parameters such as number of samples, number of elements, and the SNR. Therefore, the formula for the resolution threshold becomes more complicated than for the beamforming methods, see Table 1.1. First, the different resolution thresholds of conventional beamforming, Capon, and MUSIC will be examined in a simulation example.

Example Beamforming, Capon, and MUSIC

The beamforming (BF), Capon, and MUSIC spectra are shown in Figure 1.3 for the case when two waves are incident upon an ULA of 5 elements with half-wavelength spacing. The true directions are 85° and 95° and 100 samples of the induced voltages are taken, i.e. $N = 100$ in (1.24). Furthermore, the sources are assumed to be uncorrelated and of equal strength, i.e. $\mathbf{P} = 10^{(15/10)}\mathbf{I}$ with a noise power $\sigma^2 = 1$, resulting in an SNR of 15 dB. It is clear that the conventional beamforming method (1.26) fails to resolve the sources, since the angular separation in this example (10°) is less then the resolution threshold of conventional beamforming (23°). Capon's method just barely resolves the sources while the MUSIC method results in two well defined peaks at the true DOAs.

□

Resolution Threshold A critical property of the spectral-based methods is the ability to resolve closely spaced sources. For the conventional beamforming method a relatively simple expression for the minimum separation in electrical angle can easily be obtained as $\Delta = \frac{2\pi}{n}$. A slightly more complicated expression for the Capon's beamformer can be obtained for the lim-

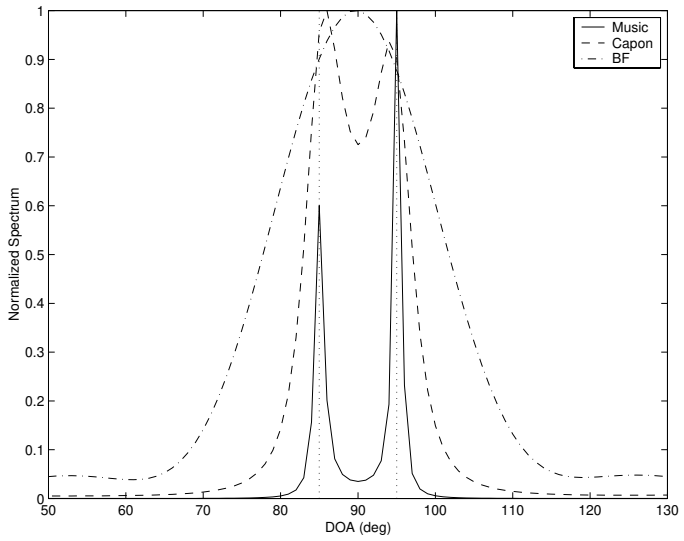


Figure 1.3: Normalized spectra of MUSIC (solid), Capon (dashed), and beamforming (dash-dotted) methods versus DOA. The true DOAs are indicated by dotted vertical lines. An ULA of 5 elements with half-wavelength spacing is used when $N = 100$ samples of the induced voltages are taken at an SNR of 15 dB.

iting case of a perfect covariance estimate and uncorrelated sources [Ser95], see Table 1.1. For correlated sources, see [Ser95]. Note that the resolution threshold of Capon's beamformer decreases with increased SNR (ξ), while the conventional beamformer is independent of the power. The resolution limit for the MUSIC method is more complex [KB86], and depends on the number of samples as well as SNR ξ and number of elements n. For more detailed descriptions on the assumptions on the resolution limits of Capon and MUSIC, see [KB86, Ser95].

BF	Capon	MUSIC
$\Delta = \frac{2\pi}{n}$	$\Delta = 8.71 \left[\frac{1}{n^5 \xi}\right]^{\frac{1}{4}}$	$\Delta = \left\{\Delta : \frac{2880(n-2)}{Nn^4\Delta^4}\left[1 + \sqrt{1 + \frac{Nn^2\Delta^2}{60(n-1)}}\right] = \xi\right\}$

Table 1.1: Resolution thresholds for beamforming (BF), Capon, and MUSIC methods. Here, n denotes the number of antennas, ξ the SNR, and N the number of snapshots.

Example Resolution Thresholds of BF, Capon, and MUSIC

The resolution threshold of the beamforming (BF), the Capon, and the MUSIC methods are shown in Figure 1.4 for the case when two

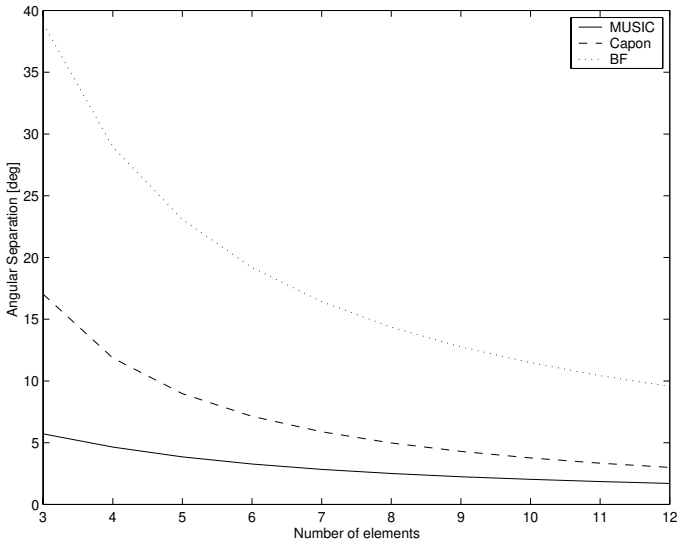


Figure 1.4: The resolution thresholds of MUSIC (solid), Capon (dashed), and beamforming (dash-dotted) methods versus the number of half-wavelength spaced antenna elements when 100 samples are taken at a SNR of 15 dB.

waves are incident upon an ULA with 3 to 12 elements with half-wavelength spacing. The waves are assumed to arrive symmetrically about the array normal and 100 samples of the induced voltages are taken, i.e. $N = 100$ in (1.24). Furthermore, the sources are assumed to be uncorrelated and of equal strength, i.e. $\mathbf{P} = 10^{(15/10)}\mathbf{I}$ with a noise power $\sigma^2 = 1$ resulting in a SNR=15 dB. It is clear that the conventional beamforming experiences the highest resolution threshold. The Capon's method has a lower resolution threshold and the MUSIC method a much lower threshold. Note that if the SNR is increased, the threshold of MUSIC and Capon will drop further while the beamforming method will not be affected. In the case of MUSIC, the threshold will also drop if more samples are taken.

□

The high resolution of MUSIC appears to make it the preferred method over beamforming and Capon at all times. Unfortunately, also the MUSIC method suffers from a few drawbacks that prevents it's application to certain problems. A problem in practical applications is that it is assumed that the number of sources is known in order to separate the eigenvectors into noise eigenvectors and signal eigenvectors. Furthermore, if the signals are

correlated the covariance matrix will not be full rank and the separation into signal and noise subspace becomes difficult. One typical example of correlated signals is in a multipath scenario, where several replicas of the signal arrive through different paths. These drawbacks were the prime motivation of introducing the parametric approaches that will be examined in the next section.

1.4.2 Parametric Methods

While the spectral-based methods presented in the previous section are computationally attractive and simple to apply, they do not always yield a sufficient accuracy. This is in particular the case for scenarios with highly correlated signals. An alternative is to employ so-called parametric array processing methods, that directly estimate the DOAs without first calculating a spectrum. These algorithms yield a higher performance in terms of accuracy and resolution by exploiting the underlying data model to a larger extent. The cost for this performance increase is a higher complexity and more computations, since typically a multi-dimensional search for the parameters are needed.

A new performance measure is now required since the DOA estimates are obtained without computing a spectrum, thus making beamwidth and resolution threshold less important. Instead two statistical properties of the DOA estimates are usually employed as performance measure.

- *Consistency:* An estimate is consistent if it converges to the true value when the number of data (or SNR) tends to infinity.
- *Statistical Efficiency:* An estimator is statistically efficient if it asymptotically attains the Cramér-Rao Bound (CRB), which is a lower bound on the covariance matrix of any unbiased estimator (see [Kay93]).

Of these, the statistical efficiency is the most important since almost all of the methods presented here are consistent. Formulae for calculating the lower bound (CRB) may be found in many textbooks on statistical signal processing.

Essentially all parametric DOA estimation methods may be formulated as

$$\hat{\boldsymbol{\phi}} = \arg\min_{\boldsymbol{\phi}} V(\phi, \mathbf{X}), \tag{1.30}$$

where $\mathbf{X} = [\mathbf{x}(t_1), \mathbf{x}(t_2), \ldots, \mathbf{x}(t_N)]$, N denotes the number of samples, and the DOA estimates are taken as the minimizing arguments of the criterion function V. Many different functions have been proposed in the literature

with different properties. The perhaps most well known and frequently used method is the Maximum Likelihood (ML) method. That method fits the above description and will be discussed in Chapter 3 in the context of mutual coupling. Several other parametric DOA estimation methods are discussed in Chapters 3-6 in the context of both known and unknown mutual coupling. A version of the ML method for an antenna employing higher order modes is also outlined in Chapter 7. For more material on parametric DOA estimation, see [God97, He93, Hay95a, KV96, VS01, VB88].

1.5 Modeling of the Wireless Communication Channel

An accurate knowledge of the mobile communications channel is of greatest importance when designing radio systems. The performance of a detector or receiver is highly dependent on the characteristics of the radio channel. If an accurate model of the channel is available, it is possible to design receiver algorithms that achieve a high performance by exploiting the properties of the channel. Thus, to accurately model the radio channel is an important problem that affects the performance of wireless communication systems. This thesis mainly deals with cellular mobile communication systems for which the term wireless communication systems also will be used.

Unfortunately, the propagation situation in a practical wireless communication channel is very complex. The signal that is transmitted from the mobile phone will reach the base station antenna through many different paths. At the base station antenna it will appear as if energy is arriving over an angular sector rather than from a distinct direction. Therefore, the model in the previous section that assumed a plane wave incident from a distinct direction needs to be modified. Numerous textbooks on radio channel characteristics have been written with applications in cellular networks [Ber00, Bla00, Par92].

Consider the propagation scenario in Figure 1.5, that contains a mobile antenna and a base station antenna as well as numerous scattering objects. Several local scattering objects (houses) are located in the vicinity of the mobile, and a large scattering object (large building) is located further away. The transmitted field from the mobile will experience electromagnetic scattering, reflection, refraction, and diffraction before it reaches the base station antenna where a voltage finally is induced. For an introduction to scattering and wave propagation, see [Bal89, Sau99] and the references therein.

By exploiting fundamental physical principles, it is possible to arrive at

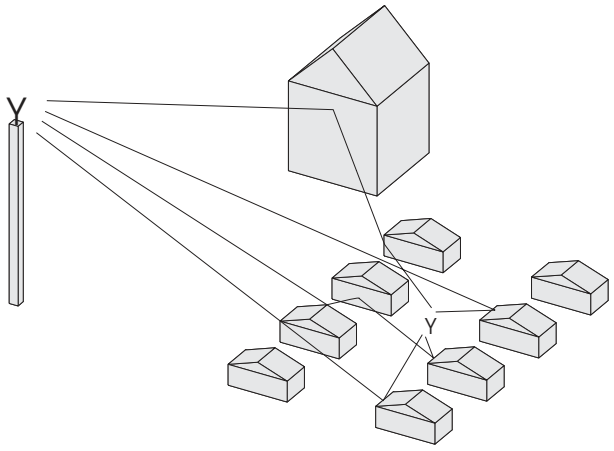


Figure 1.5: Typical suburban scatter environment with local scattering and a dominant scatterer.

a detailed channel model that includes most of the propagation phenomena encountered in practice. In fact, due to the recent developments in computational power, it is possible to approximately solve Maxwell's equations for the propagation scenario at hand. Through massive calculations, the channel impulse response of a practical scenario can be found if the geometry and electrical properties of the scenario are specified. Probably the most popular solution is based on a concept called ray-tracing, where the propagating field is viewed as a ray. Although computationally intensive, accurate results can be obtained [CPdAG98, DKW97, SI94]. These ray-tracing schemes can in some cases replace measurement campaigns by using them to simulate data if an accurate database of the scenario is available. However, for the purpose of evaluating different receiver designs, these models are way too complex to be useful. Instead, the primary use of ray-tracing schemes may be found in cell-planning tools.

Ray-tracing models have mainly been designed and used within the area of electromagnetics. In the signal processing and communications areas, much simpler models have been used. These models typically make an assumption regarding the statistical distribution of some properties of the channel. Often, very little physical knowledge is incorporated into these models. However, with the advent of antenna array techniques, a need for modeling the spatial dimension arose.

At the time antenna arrays were introduced, the base station was typically serving a large area (macrocellular application). Hence, the base station was placed rather high (elevated base station) on a mast, with negligible influence

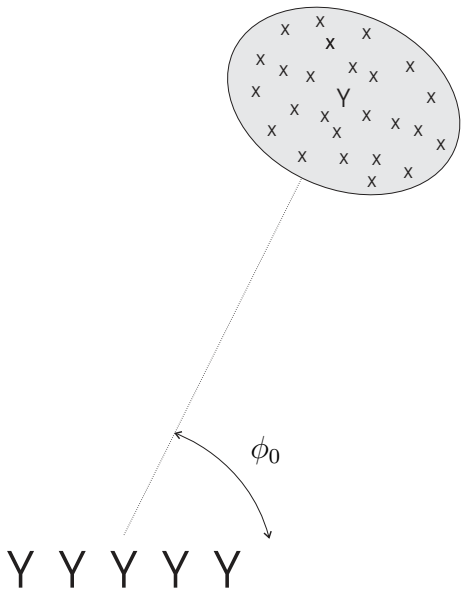


Figure 1.6: Local scattering model.

of local scatterers. The mobile, which often was located relatively far from the base, was on the other hand surrounded by local scatterers. For these types of scenarios, a common model, is the so-called local scattering model [Ast99, Zet97]. It is assumed that the transmitted signal is reflected by many scattering elements in the close vicinity of the mobile, see Figure 1.6. The direction to the center of the local cluster of scatterers around the mobile is typically referred to as the nominal DOA ϕ_0. At the mobile, however, the received signal appear as if coming from all directions. Typically, it is assumed that the time-delays of the different paths within the cluster can be modeled as a phase-shifts, i.e., a narrowband assumption is made. If, in each path, this phase-shift is incorporated into the complex amplitude ρ, the channel model can be written as

$$\mathbf{x}(t) = \mathbf{v}s(t) + \mathbf{n}(t) \ , \ t = 1, \ldots, N. \tag{1.31}$$

As before, $s(t)$ is the signal waveform and $\mathbf{n}(t)$ is the noise vector. The vector $\mathbf{v}$ is termed the *spatial signature*, and using the above assumptions it can be written as

$$\mathbf{v} = \sum_{l=1}^{p} \rho_l \mathbf{a}(\phi_l) \ , \tag{1.32}$$

where ρ_l is a complex amplitude and ϕ_l is the DOA of the lth signal path from the cluster. The number of paths in the cluster is denoted p. In case

of no multipath or negligible DOA spread, the model applies with $p = 1$ and essentially reduces to the expressions in Section 1.3.

With the above framework, the spatial dimension is included in the model. Hence, a number of different systems can be studied. For instance, studies of systems employing interference suppression via beamforming and other techniques have been performed [Lin00, Pel99, Ran99]. The impact of different receiver designs that employ array antennas has been studied in numerous publications during the last decade. Also, the success of cellular communication systems has resulted in a vast literature on radio channel modeling. Most of the literature on spatial signal processing employs the model of (1.31) or similar models. See [Ert98] for an excellent overview of spatial models in signal processing. The interested reader is referred to the following textbooks [Jak74, Lee89, Par92, Pro95, Sau99, Ste92] and further references therein. See also the following tutorials and overviews [ARY95, FL96, Has93] for further information.

The above model was essentially developed for a narrowband, single-polarized, macrocellular mobile communications scenarios. However, due to the enormous success of mobile communications, the focus has shifted from narrowband to broadband, single-polarization to multiple-polarization, and from large cells to pico-cells and indoor networks. Hence, a number of conditions have changed, and new properties and parameters of the channel need to be modeled. In particular, the recent proposal of using multiple antennas, i.e., arrays at both the transmitter and receiver, requires another type of channel model. Impressive data rates have been demonstrated using these Multi-Input Multi-Output (MIMO) systems [FG98, LP00, WFGV98]. Furthermore, it was recently found that exploiting all six possible Electro-Magnetic (EM) polarizations, i.e. the three components of the electric field and the magnetic field respectively [AMd01], might provide six uncorrelated transmission paths. Consequently, modeling the EM polarization properties of the MIMO channel is an important topic.

An interesting way of accurately modeling all these different properties is to base a channel model on fundamental physics, thereby including most of these properties. A general model that accounts for many of the physical properties encountered in practice, will be derived in this thesis (Chapter 9), by exploiting EM theory. The aim is to bridge the gap between the complicated wave propagation environment of radio signals and the idealized models previously used. This type of model is straightforwardly extended to modeling MIMO systems, which is the topic of Chapters 9-12.

The channel model is derived using fundamental results from EM scattering. A channel scenario is defined by a number of scattering objects positioned to resemble different channel environments, such as micro-cells or

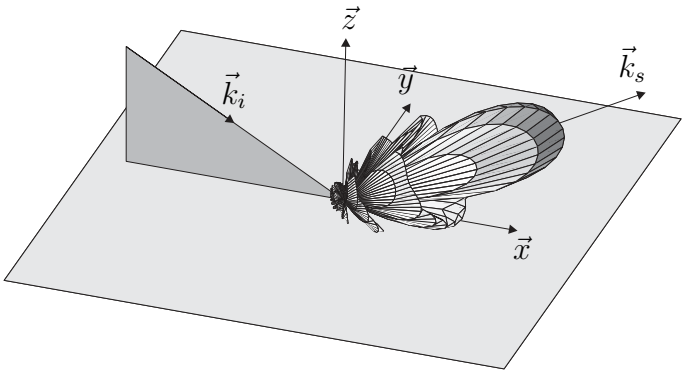


Figure 1.7: Scattering from a dielectric plate with oblique incidence.

indoor scenarios. The positions and properties of these scatterers can be of stochastic, deterministic, or semi-deterministic nature. Once the characteristics of all objects are determined, the field at the receiving antenna can be obtained using wave propagation results. The field incident upon each scatterer is calculated from the radiation pattern of the transmitting antenna, and then the field incident upon the receiving antenna, due to the scattering, is calculated.

The scattering process that includes both magnitude and polarization properties is accounted for using a scattering matrix $\vec{\mathbf{S}}$ that relates the scattered field to the incident field as

$$\vec{E}_s = \vec{\mathbf{S}} \cdot \vec{E}_i. \tag{1.33}$$

By using a scattering matrix formulation, a compact description of the channel is obtained. Scattering matrices are often used in electromagnetics, but seldom applied in signal processing models. A relatively simple expression of $\vec{\mathbf{S}}$ can be obtained by utilizing the theory of rough surface scattering and oblique scattering by dielectric plates. Here, properties such as width of scattering lobe, polarization, and size of the scatterer are included using relatively few parameters. An illustration of a scattering situation that can be described by using the scattering matrix is shown in Figure 1.7. In this case, a plane wave is scattered by a dielectric plate under oblique incidence. Here, $\vec{k}_i$ and $\vec{k}_s$ denote the normalized incidence and scatter wave vectors, and $\vec{x}, \vec{y}$ and $\vec{z}$ denote the axis in the rectangular system of coordinates. In this figure, the width of the scattering lobe is clearly visible. By combining this scattering matrix formulation with the channel scenario described above, a channel model that incorporates many of the physical mechanisms encountered in radio channel propagation is obtained. A detailed derivation and

discussion on physical channel modeling is given in Chapter 9. The channel model derived therein, is used in various forms in the following chapters to study different aspects of communication systems. In particular, the focus of these comparisons is on the border between antennas and signal processing.

The fundamental function of a channel model in signal processing and communications is to relate the transmitted signal to the received version of it. If a baseband signal $s(t)$ is transmitted, the received signal $x(t)$ can be expressed as the convolution of the channel impulse response $h(\tau, t)$ and $s(t)$ as

$$x(t) = \int_{-\infty}^{\infty} h(\tau, t)s(t - \tau)d\tau + n(t). \tag{1.34}$$

Here, the impulse response $h(\tau, t)$ is a function of both the delay τ, but is also a function of the time t that accounts for the time evolution of the impulse response. One example of time evolution effects is the Doppler shift that occurs due to movement of either the mobile, the base station, or the scatterer. Furthermore, an additive noise $n(t)$ is incorporated in (1.34). This is the by far most common assumption regarding noise. However, many other assumptions can of course be made, see [Jak74, Pro95, Sau99].

The above expression for a single transmit and receive antenna is straightforwardly extended to the case of multiple antennas. For a communication link with n_r receive antennas and n_t transmit antennas, the channel can be described by a $n_r \times n_t$ matrix $\mathbf{H}(\tau, t)$ of complex baseband impulse responses. Here, the element $\mathbf{H}_{ij}(\tau, t)$ denotes the impulse response from transmit antenna j to receive element i. Hence, $n_r n_t$ impulse responses are needed to characterize this type of MultiElement Antenna (MEA) or Multi-Input Multi-Output (MIMO) channel. In this case, the $n_r \times 1$ vector of received signals $\mathbf{x}(t)$ becomes

$$\mathbf{x}(t) = \int_{-\infty}^{\infty} \mathbf{H}(\tau, t)\mathbf{s}(t - \tau)d\tau + \mathbf{n}(t), \tag{1.35}$$

where $\mathbf{s}(t)$ denotes the $n_t \times 1$ vector of transmitted signals and $\mathbf{n}(t)$ is the noise vector of the same length. Here, it should be stressed that a continuous representation of the signals and impulse responses have been used. This appears natural when deriving the channel model from physics and electromagnetics. However, most of present and most likely also future wireless communication systems employ digital signal processing to a large extent. When devising receiver structures and detector algorithms for these systems, a discrete time representation is needed. If the received signal is sampled as $\mathbf{x}(n) = \mathbf{x}(nT)$, where T denotes the sampling period, the following relation

can be used

$$\mathbf{x}(n) = \sum_{k=-\infty}^{\infty} \mathbf{H}(k,n)\mathbf{s}(n-k) + \mathbf{n}(n). \tag{1.36}$$

Note that the channel $\mathbf{H}$ in (1.36) differs from the $\mathbf{H}$ in (1.35) and that the sampled versions of the transmitted signal and the noise are denoted $\mathbf{s}(n)$ and $\mathbf{n}(n)$. Although this formulation results in some abuse of notation, it is the normal notation used in most textbooks. For narrowband systems, where the main part of the received energy arrives at essentially the same time, the model may be further simplified to

$$\mathbf{x}(n) = \mathbf{H}(n)\mathbf{s}(n) + \mathbf{n}(n). \tag{1.37}$$

Here, the channel model reduces to complex matrices comprising complex scalars, that relate the received signals of each element to the corresponding transmitted signal from each antenna. Note that these narrowband channels are often called frequency-flat channels, since the frequency response of a channel, where the main part of the energy is received at one time instant, is approximately constant. Correspondingly, channels where the energy arrives at different time instants (as in (1.36)) are often referred to as frequency-selective channels. Both types of channels will be modeled and studied in this thesis. If the channel further is assumed to be time-invariant, the time dependency of the channel may be dropped, i.e. $\mathbf{H}$. Thus, leaving only a single matrix to characterize the channel. This representation is often used in Space Time Coding applications [NSC00].

It is important to stress that the explicit meaning of impulse response and the transmit and receive signals critically depend on the system that is modeled. In many systems, matched filtering is employed. In that case, the received signal $\mathbf{x}(n)$ is sampled after being filtered with a filter matched to the transmitted signal waveform. When modeling these systems, the pulse shaping filters of the communication system are often incorporated into the channel via a convolution of the "true physical" channel with the pulse shape of the system. Here, the filters of the transmitting and receiving equipment should also be accounted for. Hence, the signals $\mathbf{x}(n)$ and $\mathbf{s}(n)$ then denote the received and transmitted "symbols". However, the specific meaning needs to be determined for each system. For example, in many systems oversampling is used. Also, many present and future systems, such as WCDMA and CDMA2000, employ spectrum spreading. The sequence of information symbols are multiplied by a spreading sequence that changes values many times over the symbol interval. The elements of this spreading sequence are often called chips. The resulting signal has a wider bandwidth, hence the term spreading. By exploiting the properties of the spreading sequence,

different users can be separated while occupying the same frequency at the same time. For these systems, the signals are sampled at this higher rate (chip-rate), or even higher if oversampling is employed. Hence, the actual meaning of transmitted and received signal will depend on the system that is analyzed. In this thesis, the underlying physical modeling is the main objective. Therefore, the models derived herein need to be extended to incorporate properties specific to individual systems such as pulse-shaping function and oversampling.

1.6 Diversity and Channel Capacity

In order to assess the validity of a channel model, it is necessary to compute some properties of the obtained channel that can be compared to measurements or other models. Obviously, the preferred way is to perform extensive measurement campaigns in environments that resemble the model scenarios in order to address the quality of the model. Unfortunately, when writing this thesis, no extensive measurement data was available. However, some measurements were performed shortly before writing this thesis, and some preliminary ideas of how these may be used to improve the models are given in Chapter 13 as suggestions of future work. Instead of measurements, a number of overall performance measures were used to relate the output of the models to results obtained when employing real systems. Two different types of performance measures have been studied; diversity gain in different scenarios and calculation of the channel capacity under different conditions. These performance measures have also been used when evaluating two novel antenna solutions, namely multimode and parasitic antennas, in Chapters 10 and 11. The main ideas of diversity and capacity will therefore, be briefly reviewed in this section.

There are a number of characteristics of the wireless channel that makes it much more difficult to communicate over, compared to wired communication systems. Perhaps the most important difficulty, at least from a radio interface perspective, is the fading of the signal strength. The signal that is received will typically be the superposition of many propagation paths. These signals will sometimes add constructively and sometime destructively, hence creating a rapid fluctuation of the signal strength. An example of a fading signal is shown in Figure 1.8, where a y oriented dipole is used and the signal is normalized such that its maximum is of unit magnitude. Here, the rapid changes are clearly visible and that the signal power may drop 20dB in some fades. This envelope was calculated using the channel model derived in Chapter 9, where also other effects such as shadowing and other channel

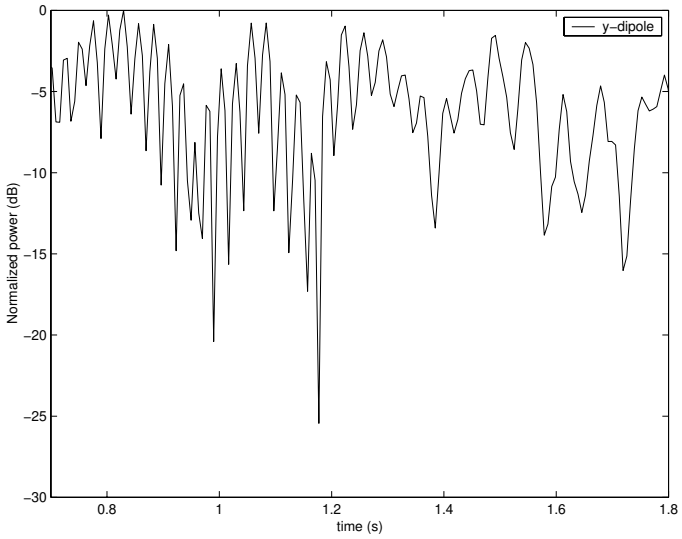


Figure 1.8: A fading signal.

characteristics are included. The rate of change depends on the speed of the mobile or other objects in the surroundings. This type of effects are often called Doppler effects and will be described in more detail in Chapter 9. An example of an everyday situation that illustrates the Doppler effect is the siren mounted on emergency vehicles and police cars. As the angle towards the siren changes, so does the appeared frequency. In wireless communications, there are many objects (or sirens) that contribute to the received signal. Hence, a very complex situation arises as seen in Figure 1.8.

Many different schemes of mitigating these severe drops of signal power have been designed throughout the history of wireless transmission and communications. One of the most important techniques for mitigating the fading of the channel is diversity combining. The main idea is to obtain several independent copies of the received signal, for example by using multiple antennas, that are combined in order to increase the received average power. An important property for the combining to be successful is that while one of the signals are exhibiting a deep fade, the others should not. Hence, the different branches should experience low mutual correlation in order for successful mitigating of deep fading dips. However, the received signals should also be of comparable average strengths for the combining to be effective. An example of diversity combining is shown in Figure 1.9, where a combination technique called Maximum Ratio Combining (MRC) [Jak74, Pro95, Ste92] is used. That technique is optimum in the sense that the received Signal to

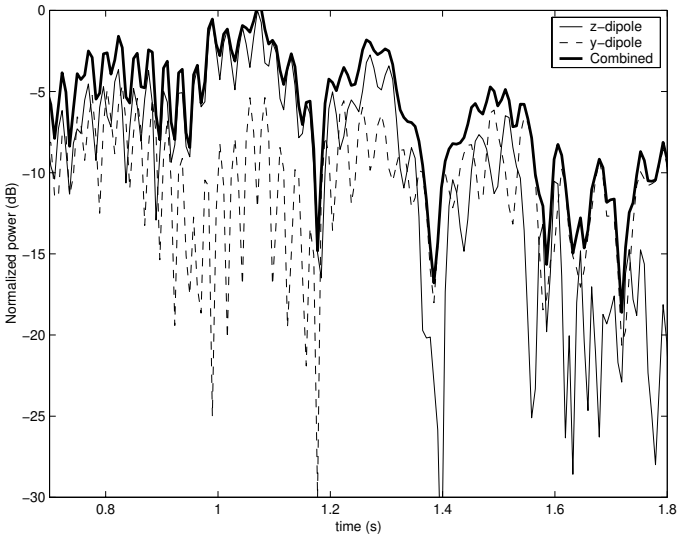


Figure 1.9: Diversity combining.

Noise Ratio (SNR) is maximized. Two dipoles are used, one z oriented and one y oriented, and the scenario is the same as in Figure 1.8. By employing diversity combining, many of the deep fades may be avoided. Hence, the performance of a communication system may be significantly improved. This is particularly important when a high performance is required, since then almost all the transmission errors will occur in the fades.

Many different ways of obtaining different branches or copies of the signal of interest have been proposed.

- **Space diversity:** Two or more antennas, that are separated in space, are used to yield uncorrelated versions of the received signal.

- **Time diversity:** By transmitting the same signal multiple times, several different versions of the channel are encountered.

- **Frequency diversity:** Transmitting the signal over different frequencies that fade independently also provides diversity.

- **Pattern diversity:** Receiving the signal with different radiation patterns is also a mean of obtaining different diversity branches.

- **Polarization diversity:** Several versions of the signal may also be obtained by employing antennas of different polarization.

Often, polarization and pattern diversity are considered to be a kind of space diversity, since all three depend on the spatial characteristics of the channel. Space diversity as well as polarization and pattern diversity will be examined in some detail in this thesis. In particular, the correlation between the different branches are studied since this will essentially (together with the magnitude of the branches) determine the potential gain in using diversity combining. The correlation between the received signal levels of spatially separated antennas will be studied in Chapter 9, as a mean of evaluating the properties of the channel model derived therein. Two novel diversity antennas are studied in Chapter 10 and 11, where the correlation properties are investigated.

Traditionally, only receive diversity has been implemented in mobile communications. Mostly, space diversity in the form of spatially separated antenna elements at the base station has been employed. Recently, the increased demand of high data rates and spectrum efficient communications, has raised an interest in employing also transmit diversity. Significant performance gains are possible by employing space diversity also in the transmit mode. By combining antenna signal processing with coding concepts from communication theory, i.e. Space Time Coding (STC), excellent performance may be obtained [NSC00]. During the last decade, an intense activity in the area of STC has lead to numerous publications. In fact, some of the STC ideas have already found their way to some of the current third-generation wireless system standards.

In this thesis, the performance improvements of using multiple antennas at both the transmitter and receiver, will be examined. Recently, it has been pointed out by a number of researchers that high data rates may be possible by employing these MIMO systems. For example, in [SOK00] data rates at 300Mbits/s, for a HiperLAN/2 type of systems, were discussed when employing three transmit and receive elements. Those discussions were based on measurements. However, most of the analysis of MIMO systems have been performed using rather idealistic channel models and conditions. In this thesis, the performance of these systems will be examined using a physically based channel model. However, instead of examining the performance of a specific STC scheme, an upper bound on the achievable data rate, i.e. the channel capacity C, is considered. The Shannon channel capacity is a measure of the maximum possible rate, that can be transported over a channel, with arbitrarily low Bit Error Rate (BER) probability. It can be achieved by coding of infinite delay. The seminal capacity formula due to Shannon [Sha48] has been extended to the case of multiple transmit and receive antennas by several researchers. In [FG98], a measure of the capacity for a quasi-static random channel has been derived. A more general framework

on capacity is presented in [Tel95], where capacity measures are derived for channels with and without fading. Here, the formulation for a quasi-static MIMO channel $\mathbf{H}$ [FG98] will be used

$$C = \log_2 \det \left[\mathbf{I} + \frac{\xi}{N} \mathbf{H}\mathbf{H}^H \right] \text{ bits/s/Hz}, \tag{1.38}$$

where ξ/N denotes the SNR at each receive antenna. Note that the channel matrix $\mathbf{H}$ is $M \times N$, where M, N denotes the number of antenna elements at the receiver and transmitter, respectively. Furthermore, the element $\mathbf{H}_{ij}$ represents the complex path gain from transmitter j to receiver i. It should be stressed that the above capacity formulation differs from the definition of Shannon. Here, the capacity is a random value that can be seen as an "instantaneous capacity". No matter what small desired rate, there is a non-zero probability that the channel realization is incapable of supporting it, no matter the code length. On the other hand, one can talk about a tradeoff between the outage probability, namely the probability of not receiving a desired rate. Hence, the capacity results are often presented in the form of a Complementary Cumulative Distribution Function (CCDF), or as the capacity for an certain probability of outage.

The channel capacity will be used to illustrate the properties of the channel model that is derived in Chapter 9. Furthermore, the capacity of two novel antenna solutions, the multimode and parasitic antennas, are examined in Chapter 10 and 11. The performance of a relatively simple STC scheme employing block-codes in the context of parasitic antennas is also examined in Chapter 11. Finally, the subject of mutual coupling is revisited in Chapter 12, where the impact of coupling upon the channel capacity is investigated. This ends the introductory material on antennas and propagation in radar and mobile communication applications. Next, the content of the different chapters will be briefly outlined.

1.7 Outline

The purpose of this section is to provide a chapter-wise description of the thesis, as well as providing an overview of the main contributions. It should also be mentioned that most parts of the thesis are modified versions of submitted journal manuscripts, a book chapter, technical reports, and conference papers. Often, only minor editing has been done. Therefore, with the exception of Chapters 2-6, the chapters are self-contained and can be read in any order.

Chapter 2

Estimating the DOA of a wave incident upon an array in the presence of a known coupling is considered. By using basic electromagnetic concepts, the mutual coupling in an array of dipoles is calculated. The direction finding accuracy in the presence of a known coupling is then investigated by calculating a lower bound on the variance of the estimated angles, the Cramér-Rao lower Bound. It is found that a known coupling does not significantly affect the performance. Parts of this chapter have appeared in [SV97, Sva98b].

Chapter 3

Methods for estimating the DOAs in the presence of a known coupling are discussed in this chapter. Many DOA algorithms, derived for the hypothetical coupling free case, can be extended to include a known mutual coupling. By first reviewing the published methods that do include a known coupling and then introducing several extensions of other DOA algorithms, the field of sensor array processing is introduced. The problem of detecting the number of incident waves is also briefly discussed. Modified versions of this chapter have been published in [Sva98a, SV98]

Chapter 4

Usually the mutual coupling can be obtained using calibration measurements, but in some cases, obtaining the coupling might not be possible. For instance, the environment might be changing too quickly to allow for calibration measurements. If the coupling is left uncompensated, it will drastically reduce the possibilities of performing direction finding. One way of mitigating these effects is to estimate the coupling along with the DOAs using signal processing methods. However, the electromagnetical coupling model derived in Chapter 2 generally contains too many parameters to be estimated.

In Chapter 4, a parameterization of the coupling with a feasible number of parameters, unlike a direct parameterization of the coupling matrix, is derived. Furthermore, the question of uniqueness of the parameter estimates is discussed. Also, a necessary condition for uniqueness is derived, that gives a limit on the number of coupling parameters that can be uniquely estimated. The main ideas of this chapter have also appeared in [Sva99d, Sva99e].

Chapter 5

The parameterization of the coupling is in this chapter used to estimate both an unknown coupling and DOAs. A ML method is derived. Based on the ML expressions, the Cramér-Rao lower Bound (CRB), for this model, is also derived. Then, several estimation methods are proposed along with a review of a previously published method. The properties of the different estimation methods are examined in a few computer experiments. The estimation methods derived in this section, have been published in [Sva99c, Sva00f].

Chapter 6

The impact of the model errors, introduced by using the reduced model to estimate the coupling, is examined in this chapter. The effects are studied using data generated using the full electromagnetic model and estimated with the reduced model derived in Chapter 4, using the estimation methods derived in Chapter 5. The effects of an unknown coupling on the estimation of the signal in the presence of a jammer is also examined. The results of this chapter have appeared in [Sva99a, SV99]. This chapter ends the study of mutual coupling effects on direction finding. With only minor changes, the Chapters 2-6 have previously been included in the author's Licentiate Thesis [Sva99b]. A heavily edited version of this material has also been submitted to *IEEE Trans. on Antennas and Propagation* [Sva00g].

Chapter 7

Direction finding by exploiting higher order modes of a modified biconical antenna is the topic of this chapter. The DOA of an incident wave is estimated by using the fact that the different modes have different directional properties. Since only one antenna element is used, the spatial requirements of an array of elements is avoided as well as the issue of mutual coupling between antenna elements. The direction finding performance is examined by calculating the CRB and the ML estimator for several different cases. It is found that the multimode antenna offers direction finding performance comparable to the traditional antenna array using only a single antenna element. This chapter is essentially based upon material previously published in [Sva00e]. An edited shorter version has also been submitted to *IEEE Signal Processing Letters* [Sva01d].

Chapter 8

In this chapter, the idea of exploiting the directional radiation patterns of an Switched Parasitic Antenna (SPA) for direction finding, is studied. By employing passive elements (parasites), which can be shorted to ground using pin diodes, directional radiation patterns can be obtained. The direction finding performance of the SPA is examined by calculating a lower bound on the direction finding accuracy, the CRB. It is found that the SPA offers a compact implementation with high-resolution direction finding performance using only a single radio receiver. Thus, exploiting SPAs for direction finding is an interesting alternative to traditional antenna arrays, offering compact and low-cost antenna implementations. This chapter will appear with minor changes in [SW01].

Chapter 9

A general framework for physical spatio-temporal channel modeling is derived in this chapter. Based on this framework, models for single antenna systems as well as Multi-Input Multi-Output (MIMO) models for multi-element antenna systems, are derived. The model is based on results from ElectroMagnetic (EM) scattering and wave propagation, thereby including many of the channel characteristics encountered in practice. A general description that captures the most important scattering properties of arbitrary objects is derived in a compact dyad formulation, by using results from rough surface scattering. The polarization properties of the channel and the antennas are thus included in the model, which allows for studies of the impact of different antenna arrangements and polarizations. Several channel environments are simulated, and properties such as spatial correlation, channel capacity, and time evolution of the channel are calculated. The proposed physical channel model is also suitable for other interesting application areas, such as long-range prediction of fading signals, feedback based transmit diversity schemes, and wideband MIMO systems where the temporal properties are important. This chapter represents only minor changes of [Sva01e]. A much shorter version comprising the main ideas of the chapter has been submitted to *IEEE Journal on Selected Areas in Communications* [Sva01b]. Parts of the results have also been published in the following conference proceedings [Sva00a, Sva00b, SV01]. A paper regarding multi-polarized MIMO systems that to some extent rely on results presented in this chapter has also been submitted to *IEEE Vehicular Technology Conference Fall 2001* [Sva01a].

Chapter 10

Here, a novel way of exploiting higher modes of antennas as diversity branches in MIMO systems, is examined. Essentially, antennas employing multiple modes offer characteristics similar to an antenna array, through multiple modes and using only a single element. Analytical expressions for the correlation between different modes in a realistic environment is presented for a biconical and a circular microstrip antenna that employs higher order modes. It is found that the correlation is low enough to yield a diversity gain. Furthermore, the channel capacity of a MIMO system using a multimode antenna, i.e. an antenna employing multiple modes, is found to be comparable to the capacity of an array. This chapter has, with minor changes, been submitted to *IEEE Trans. on Vehicular Technology* [Sva01c]. The chapter has also resulted in the following conference papers [Sva00d, Sva00h, Sva00c].

Chapter 11

The potential of using the switched parasitic antenna (SPA) as a novel diversity technique in a MIMO systems, is investigated in this chapter. It is shown that the correlation between the received signal modes is sufficiently low to yield a diversity gain. The capacity limit using the SPA is investigated for different SPA configurations and it is found that the capacity is comparable with an array antenna configuration in certain situations. Finally, a space time block coding scheme is used to evaluate the bit error rate of a MIMO-SPA system. It is found that the SPA requires a slightly higher SNR than an antenna array solution to yield the same performance. However, the array antenna requires a radio transceiver for each antenna, as opposed to the SPA which uses only one transceiver. An slightly modified version of the chapter has been submitted to *The* 12^{th} *IEEE International Symposium on Personal, Indoor and Mobile Radio Communication (PIMRC)* [WS01].

Chapter 12

This chapter studies the capacity of multiple element antenna systems, with an emphasis on the effect that mutual coupling between the antenna elements has on the capacity. The results presented shows, contrary to some earlier claims, that correlation between different channel coefficients as a function of antenna spacing, can in fact decrease when the mutual coupling effect is accounted for. As a consequence, capacity also improves. An edited version of this chapter will be appear in [SR01].

Chapter 13

Finally, the thesis is concluded in Chapter 13 by commenting on the results derived and some proposals for future work are presented.

1.8 Contributions

Below, the publications by the author, including some publications not covered by this thesis, are listed.

List of Publications

[NS96] G. Nilsson and T. Svantesson. "Analysis of Radar Backscatter of Stones Using a Modified Extended Physical Optics Approximation". Master's thesis, Chalmers University of Technology, 1996.

[RSB00] A. Ranheim, T. Svantesson, and P. Bohlin. "Combining Spatial and Temporal Processing in a CDMA System Using an Array of Antennas". In *Proc. Nordiskt antennsymposium*, pages 137–142, Lund, Sweden, September 2000.

[SR01] T. Svantesson and A. Ranheim. "Mutual Coupling Effects on the Capacity of Multielement Antenna Systems". In *Proc. IEEE ICASSP 01*, Salt Lake City, Utah, May 2001. To appear.

[SV97] T. Svantesson and M. Viberg. "On the Direction Finding Accuracy of a Calibrated Array in the Presence of Mutual Coupling". In *Proc. Nordiskt antennsymposium*, pages 385–392, Göteborg, Sweden, May 1997.

[SV98] T. Svantesson and M. Viberg. "Mutual Coupling in Antenna Arrays: Effects and Cures". In *Proc. PCC Workshop*, pages 99–103, Stockholm, November 1998.

[SV99] T. Svantesson and M. Viberg. "Adaptive Antennas and Channel Modeling: The First Wild Years". In *Proc. PCC Workshop*, pages 74–77, Lund, November 1999.

[SV01] T. Svantesson and M. Viberg. "A Radio Channel Model for Multielement Antenna Systems". In *Proc. PCC Workshop*, Nynäshamn, April 2001.

[Sva98a] T. Svantesson. "The Effects of Mutual Coupling on the Direction Finding Accuracy of a Linear Array of Thin Dipoles of Finite Length". Technical report, Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, 1998.

[Sva98b] T. Svantesson. "The Effects of Mutual Coupling Using a Linear Array of Thin Dipoles of Finite Length". In *Proc. 8th IEEE Signal Processing Workshop On Statistical Signal and Array Processing*, Portland, USA, September 1998.

[Sva99a] T. Svantesson. "Antennas and Propagation from a Signal Processing Perspective". In *Proc. NRS "Antennen och dess omvärld"*, pages 35–38, Lund, November 1999.

[Sva99b] T. Svantesson. "Direction Finding in the Presence of Mutual Coupling". Licentiate thesis, Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, March 1999.

[Sva99c] T. Svantesson. "Methods of Mitigating the Effects of Mutual Coupling in Antenna Arrays: A Signal Processing Approach". In *Proc. RadioVetenskap och Kommunikation (RVK 99)*, pages 441–445, Karlskrona, Sweden, June 1999.

[Sva99d] T. Svantesson. "Modeling and Estimation of Mutual Coupling in a Uniform Linear Array of Dipoles". Technical report, Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, 1999.

[Sva99e] T. Svantesson. "Modeling and Estimation of Mutual Coupling in a Uniform Linear Array of Dipoles". In *Proc. ICASSP 99*, pages 2961–2964, Phoenix, USA, March 1999.

[Sva00a] T. Svantesson. "A Spatio-Temporal Channel Model Based on EM Scattering". In *Proc. IEEE AP-S*, pages 1110–1113, Salt Lake City, UT, July 2000.

[Sva00b] T. Svantesson. "A Study of Polarization Diversity Using an Electromagnetic Spatio-Temporal Channel Model". In *Proc. IEEE VTC 2000 Fall*, Boston, MA, September 2000.

[Sva00c] T. Svantesson. "An Antenna Solution for MIMO Channels: The Multimode Antenna". In *Proc. 34th Asilomar Conf. Sig., Syst., Comput.*, Pacific Grove, CA, October 2000.

[Sva00d] T. Svantesson. "Multimode Antenna Diversity". In *Proc. Nordiskt antennsymposium*, pages 97–102, Lund, Sweden, September 2000.

[Sva00e] T. Svantesson. "Multimode Based Direction Finding". In *Proc. 34th Asilomar Conf. Sig., Syst., Comput.*, Pacific Grove, CA, October 2000.

[Sva00f] T. Svantesson. "Mutual Coupling Compensation Using Subspace Fitting". In *Proc. IEEE SAM 2000 Workshop*, pages 494–498, Boston, MA, March 2000.

[Sva00g] T. Svantesson. "Mutual Coupling From a Signal Processing Perspective". Submitted to *IEEE Trans. on Antennas and Propagation*, 2000.

[Sva00h] T. Svantesson. "On the Potential of Multimode Antenna Diversity". In *Proc. IEEE VTC 2000 Fall*, Boston, MA, September 2000.

[Sva01a] T. Svantesson. "A Physical MIMO Radio Channel Model for Multi-Element Multi-Polarized Antenna Systems". Submitted to IEEE VTC 01 Fall, Atlantic City, NY, October, 2001.

[Sva01b] T. Svantesson. "A Physical Spatio-Temporal Channel Model for MIMO Systems". Submitted to *IEEE Journal on Selected Areas in Communications*, 2001.

[Sva01c] T. Svantesson. "Correlation and Channel Capacity of MIMO Systems Employing Multimode Antennas". Submitted to *IEEE Trans. on Vehicular Technology*, 2001.

[Sva01d] T. Svantesson. "High-Resolution Direction Finding Using the Higher Modes of a Biconical Antenna". Submitted to *IEEE Signal Processing Letters*, 2001.

[Sva01e] T. Svantesson. "Physical Channel Modeling of Multi-Element Antenna Systems". Technical report, Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, 2001.

[SW01] T. Svantesson and M. Wennström. "High-Resolution Direction Finding Using a Switched Parasitic Antenna". In *Proc. IEEE SSP 01*, Singapore, August 2001. To appear.

[VS01] M. Viberg and T. Svantesson. "Direction-Of-Arrival Estimation in Mobile Communication Scenarios". In L. Godara, editor, *The Handbook on Antennas in Wireless Communications*. CRC Press, 2001. To be published.

[WS01] M. Wennström and T. Svantesson. "An Antenna Solution for MIMO Channels: The Switched Parasitic Antenna". Submitted to PIMRC 2001, San Diego, CA, September, 2001.

Chapter 2

Coupling Effects on Direction Finding Accuracy

Estimating the DOA of a wave incident upon an array in the presence of a known coupling is considered. By using basic electromagnetic concepts, the mutual coupling in an array of dipoles is calculated. The direction finding accuracy in the presence of a known coupling is then investigated by calculating a lower bound on the variance of the estimated angles, the Cramér-Rao lower Bound. It is found that a known coupling does not significantly affect the performance.

2.1 Introduction

In the last decades, the area of sensor array processing has attracted considerable interest in the signal processing community [Hay95a, KV96]. The focus of this work has been on high resolution DOA estimation algorithms for detection and identification of aircrafts using radar systems. These algorithms can also be used in sonar applications where underwater arrays of acoustical sensors are used to locate and identify other vessels. However, in recent years the rapid development in the communication field has inspired some work on DOA estimation when using antenna arrays at base stations.

The key to obtaining high resolution in the DOA estimates is to use a parameterized model of the array measurements. The quality of the estimates then depends heavily on how well the model describes the data. In the array case this implies that the response of the array needs to be known. Also, the statistical assumptions regarding the signals and the noise must be correct for the quality of the estimates to obey the theoretical predictions

derived using the assumed model. However, in practice, the array response is never known exactly, and the statistical assumptions are only approximately correct even for large data sets. Temperature, pressure, humidity and mechanical vibrations are only some factors which affect the properties of the array resulting in a time-varying array response.

Much work has been made on different calibration schemes that try to overcome uncertainties in the array response. Those efforts have mainly been focused on obtaining estimation algorithms that are robust to modeling errors [Hay95a]. Essentially, two different approaches of minimizing the errors have been studied. Firstly, the model errors are treated as random perturbations from some nominal model, and robust estimation methods are derived based on the statistical assumptions regarding the perturbations. Those analyses typically lead to weighted estimation algorithms that minimize the variance of the estimates [SK92, SK93, VS94]. Secondly, some part of the array response is assumed unknown, but not random, and estimated together with the DOAs. Typically, the locations of the sensors are assumed unknown but close to some nominal array and estimated [FB00, LM87, RS87, WF89].

Independently of the approach used, one modeling error that frequently is considered in the literature is unknown gain and phase of the sensors used in the array. Many different algorithms that estimate the gain and phase, either together with the DOAs or using known calibration sources, have appeared in the literature [FW88, Hay95a, NS96a, PK85, SK92, WF97].

Most of these methods do not use any detailed physical reasoning or measurements. An important source of modeling error in practical antennas is that the different elements of the antenna affect each other through mutual coupling, and this effect can drastically reduce the performance of the direction finding algorithms. The subject of mutual coupling has not attracted much interest, as compared to the case of independent sensors, in the signal processing literature. Mutual coupling is on the other hand a well known problem for antenna designers, and in the electromagnetical literature mutual coupling is a well covered subject [Bal82, Ell81, Kil98].

Although very little work on mutual coupling has been published in the signal processing literature, some studies have been made. In [PK91] and [LDvdV99], the effects of an uncompensated mutual coupling on the estimation performance is studied using measurements. Another study of coupling effects on direction finding performance may be found in [DLX00], where the coupling is obtained using a computational electromagnetics solver. When using antenna arrays in digital communication applications it might be argued that the information retrieval is the main concern and not DOA estimation. But if the DOA estimation is severely affected due to mutual coupling, it is reasonable to believe that mutual coupling also affects the signal detection

problem when array antennas are used [BE96, LLSB98]. Since many communication systems now proposed include antennas arrays, the mutual coupling effects is an important problem that needs to be studied. Some initial experiments in [LSM00] indicate that in some cases, mutual coupling can in fact decrease the Bit-Error-Rate performance on a Nakagami fading channel. A similar result is presented in Chapter 12, where it is found that coupling can decrease the correlation between signals received by closely spaced antenna elements. Recently the effects of mutual coupling on the Constant Modulus Algorithm (CMA) was investigated in [YHZ98], and in [WKS+98] the effects on an aircraft navigation aid radio beacon facility (VOR) was studied.

If mutual coupling can pose a problem when estimating DOAs or signals, it is then of interest to reduce or otherwise mitigate the effects of mutual coupling. The most natural way of doing this is to design the antenna from the start in order to avoid high levels of mutual coupling and that is usually done. Reducing the mutual coupling is in fact one of the most important design problems in antenna design, and much work has been published. The effects of coupling on the radiation pattern for wire antennas has been thoroughly studied, see for instance [IN71, KP85, SH69]. Another way of reducing the effects of mutual coupling is to introduce extra antenna elements that are not used, i.e. dummy elements. The mutual coupling effects on DOA estimation by using dummy elements are investigated in [Lun96]. Furthermore, the coupling can also be compensated for by using analog low-loss networks [AR76, Spe96].

But it is not always best to design the antenna with the lowest possible mutual coupling. Instead, it is important to account for the coupling correctly. Actually, a known coupling can sometimes increase the estimation performance if compensated for correctly.

Adding additional hardware to combat mutual coupling is expensive and power consuming. An appealing alternative is to use signal processing methods instead, since the available processing power grows larger for every year due to the rapid development of the silicon industry. Therefore, this and the following chapters will focus on the effects of mutual coupling and methods of compensating for it using signal processing methods.

If the coupling is known, it is relatively straightforward to compensate for the mutual coupling. One natural idea is to multiply the data with a correction matrix and then apply a coupling free estimation method [Jos94, SH90]. Note that this method can be implemented either in hardware or using signal processing. However, the multiplication affects not only the signal part, but also the noise which changes color. Therefore, methods that rely on a white noise covariance are not, at least not directly, applicable using this method. Most DOA estimation schemes are possible to extend

to include a known mutual coupling by using an effective steering matrix. However, in the literature, few methods has been analyzed in the presence of a known coupling. The popular MUSIC method [Sch79] has been extended to the coupling case in [PF94, RW92, YLU89]. Furthermore, in [Sva99d], the Root-Music [Bar83] and the ESPRIT [Roy87] algorithms are modified to include a known coupling. The ESPRIT method has also been studied in the presence of coupling in [HW90]. In Chapter 3, these methods will be briefly discussed and several other existing coupling free estimation methods will also be extended to include a known coupling.

The coupling might not be known in all scenarios, and if left uncompensated it will reduce the possibilities of direction finding. Different compensation algorithms or calibration methods can now be applied. For instance, it can be assumed that the DOAs are known and only the coupling and perhaps additional calibration parameters are estimated [NS96a, PDS97, PK91, SAGA98, SDL00, SL01]. On the other hand, if the DOAs are also assumed to be unknown and estimated along with the coupling, so-called auto-calibration methods result. An iterative version of the MUSIC method, that estimates both coupling and DOAs, was developed in [FW91]. However, it was found in [PK91] to give nonunique estimates and in [PF94] to experience slow convergence. In Chapter 5, the method of [FW91] will be reviewed and several other methods of estimating unknown DOAs and coupling will be introduced.

Note that in the following, the coupling is assumed to be non-random and is actually calculated using basic electromagnetics. No knowledge about the coupling is assumed when the coupling is estimated. Of course, in a realistic scenario some knowledge about the coupling usually exists and should be exploited. For instance, methods that use calibration data together with some random perturbation, as well as unknown gain and phase of the sensor responses, should of course be considered. However, in order to get a tractable analysis, only the simplified case of a completely known coupling will be considered in this chapter. Extensions to an unknown coupling will be made in the following chapters.

2.2 Array Model and Coupling

An expression for the received voltage at each antenna element is necessary when analyzing the performance of the different direction finding schemes in the presence of mutual coupling. However, antennas are usually analyzed in transmit mode since the radiating properties are the same when transmitting and receiving (reciprocity). A recent compact summary that includes receiving properties can be found in [Mey00] and in [Kil99] equivalent cir-

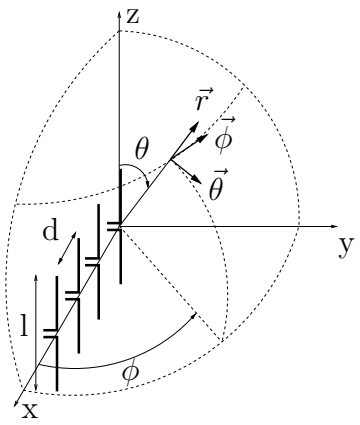


Figure 2.1: The geometry of the array of dipoles and the relation to the coordinates.

cuits for antennas in receive mode are presented. Therefore, expressions of the induced voltages will be derived for a linear array of n thin dipoles of finite length l, see Figure 2.1. The spacing between the dipoles is d and in Figure 2.1, the spherical basis vectors are introduced. The steps in the derivation are general and valid for any type of antenna, but the dipole array was chosen since the dipole is a widely used element and the expressions for the voltages simplifies considerably.

If a uniform plane wave impinges upon the array from the (θ,ϕ) direction, the wave will reach each element at different time instants. By measuring the induced voltages on the elements of the array, the DOA can be estimated using these time instants assuming a narrowband signal of known frequency. However, the received voltage on each element will induce a current on the element, which in turn radiates a field which affects the surrounding elements, i.e. mutual coupling. The induced voltage at element i, V_{ind_i}, consists of a direct voltage from the wave V_{wave_i} and a voltage induced from the neighboring elements

$$V_{\text{ind}_i} = V_{\text{wave}_i} - \sum_{j \neq i} I_j Z_{ij}. \tag{2.1}$$

Here, Z_{ij} is the mutual impedance between elements i and j, and I_j is the current at element j. The mutual impedance can be calculated using the method of induced electro-motive-force [Bal82]. Usually, these calculations are intractable, but for the case of thin dipoles of finite length the expressions can be found in most antenna handbooks [Bal82, Ell81]. For two identical finite length dipoles placed side-by-side, the expressions for the mutual impedance are simplified if the length l is a multiple of half a wavelength.

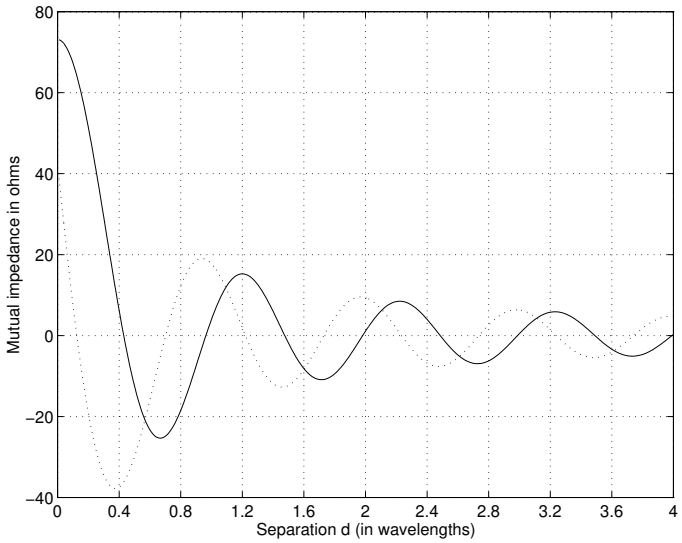


Figure 2.2: Mutual impedance of thin half-wavelength dipole oriented side-by-side (R solid, X dotted).

The mutual impedance between element m and n in the array then become

$$Z_{mn} = R + jX \tag{2.2a}$$

$$R = \frac{\eta}{4\pi \sin^2(kl/2)} \left[2C_i(u_0) - C_i(u_1) - C_i(u_2)\right] \tag{2.2b}$$

$$X = -\frac{\eta}{4\pi \sin^2(kl/2)} \left[2S_i(u_0) - S_i(u_1) - S_i(u_2)\right] \tag{2.2c}$$

$$u_0 = kd|m-n| \tag{2.2d}$$

$$u_1 = k\left(\sqrt{d^2|m-n|^2 + l^2} + l\right) \tag{2.2e}$$

$$u_2 = k\left(\sqrt{d^2|m-n|^2 + l^2} - l\right), \tag{2.2f}$$

where C_i and S_i are the cosine and sine integrals

$$C_i(x) = -\int_x^\infty \frac{\cos y}{y}\, dy \tag{2.3a}$$

$$S_i(x) = \int_0^x \frac{\sin y}{y}\, dy. \tag{2.3b}$$

In Figure 2.2 the mutual impedance of two thin half-wavelength dipoles oriented side-by-side is shown as a function of the element separation d. The mutual impedance approaches the self-impedance when the element separation diminishes. For larger element separations the mutual impedance

quickly vanishes. Hence, only elements that are close affect each other. More detailed calculations of the mutual impedance that takes the presence of open-circuited elements into account is possible by employing a Method Of Moments (MOM) approach [AS00].

The voltage induced directly from the wave V_{wave_i} can be calculated in a manner similar to when calculating the mutual impedance [Bal82] but using the incident field instead. The field at the antenna from the uniform plane wave can be written as

$$\vec{E}_{wave}(z') = \vec{\rho} E_o e^{-j\vec{k}\cdot\vec{z}} = \left(\frac{\rho_\theta \vec{\theta} + \rho_\phi \vec{\phi}}{\sqrt{\rho_\theta^2 + \rho_\phi^2}} \right) E_o e^{-jkz' \cos\theta}, \tag{2.4}$$

where the strength of the field at the surface is denoted E_0, the polarization of the field is included in the parameters ρ_θ and ρ_ϕ, and $\vec{k}$ is the wave vector. The induced voltage is then obtained as

$$V_{wave} = -\frac{1}{I_{in}} \int_{-l/2}^{l/2} \vec{E}_{wave} \cdot \vec{I}\, dz', \tag{2.5}$$

where I_{in} is the current at the input terminals. For a very thin dipole, the current distribution at $z = z'$ can be approximated by [Bal82, Ell81]

$$\vec{I}(z') = \vec{z}\ I_0 \sin\left[k(\frac{l}{2} - |z'|) \right] \quad -l/2 \le z' \le l/2 \ . \tag{2.6}$$

Note that this current distribution corresponds to a MOM approach employing a single basis function. Performing the integration in (2.5), the induced voltage becomes

$$V_{\text{wave}_i} = \frac{\rho E_{o_i}}{sin\left(\frac{kl}{2}\right)} \frac{\lambda}{\pi} \left[\frac{\cos\left(\frac{kl}{2}\cos\theta\right) - \cos\frac{kl}{2}}{\sin\theta} \right]. \tag{2.7}$$

The field strength at dipole i is denoted E_{oi}, λ is the wavelength, and $k = 2\pi/\lambda$ is the wave number. The parameter $\rho = \rho_\theta / \sqrt{\rho_\theta^2 + \rho_\phi^2}$ is a polarization mismatch factor. In the following analysis, the polarization mismatch will be neglected. Thus, ρ will be equal to unity.

Since the induced voltage at each element depends on the currents on the other elements, a circuit model is needed in order to solve for the received voltages. A simple circuit model of element i is shown in Figure 2.3. Here, Z_A is the antenna impedance that is calculated in the same manner as the mutual impedance, and Z_T is the terminating impedance that represents the

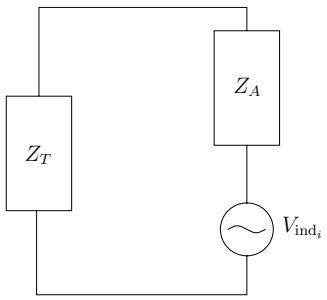


Figure 2.3: Circuit model of the i:th antenna element.

measurement equipment. Equations (2.1) and (2.7) along with the circuit model form a system of equations, from which the received or measured voltage over Z_T can be obtained as

$$\mathbf{V}_T = \mathbf{C}H(\theta) \underbrace{\begin{bmatrix} 1 \\ e^{-jkd\sin\theta\cos\phi} \\ \vdots \\ e^{-jkd(n-1)\sin\theta\cos\phi} \end{bmatrix}}_{geometrical\ \ array factor} \underbrace{E_0}_{s}, \tag{2.8}$$

where $H(\theta)$ is the normalized response of an individual element to the plane wave

$$H(\theta) = \frac{Z_T}{Z_T + Z_A} \frac{\lambda}{\pi \sin\frac{kl}{2}} \left[\frac{\cos\left(\frac{kl}{2}\cos\theta\right) - \cos\frac{kl}{2}}{\sin\theta} \right], \tag{2.9}$$

s is the strength of the source (the wave), and $\mathbf{C}$ is the coupling matrix

$$\mathbf{C} = (Z_T + Z_A)(\mathbf{Z} + Z_T\mathbf{I})^{-1}. \tag{2.10}$$

Note that $\mathbf{I}$ denotes the identity matrix and $\mathbf{Z}$ the mutual impedance matrix. The expression for the coupling is valid for any antenna if the corresponding impedances are used. However, for center-fed slots the expressions derived here can be used directly by exploiting Booker's relation [Ell81]. The terminating impedance is chosen as the complex conjugate of the antenna impedance in order to reduce the power loss. If a large terminating impedance is chosen, the matrix $\mathbf{C}$ will be more diagonal dominant. As a result, the effects of coupling are reduced at the cost of increased mismatch and loss of power [RW92].

The time-delay between each element becomes, due to the narrowband assumption, a simple phase shift and in (2.8) this corresponds to the geometrical array factor. Often the geometrical array factor and the response

are combined into a steering vector $\mathbf{a}(\theta, \phi)$, and the model for the mutual coupling in an array of finite length thin dipoles becomes

$$\mathbf{V}_T = \mathbf{C}\ \mathbf{a}(\theta, \phi)\ s. \tag{2.11}$$

This model differs from the one normally used in sensor array processing only in the introduction of a coupling matrix $\mathbf{C}$ that acts on the steering vector. Therefore, the effects of mutual coupling can be included by changing the steering vector into an effective steering vector $\mathbf{a}_e(\theta, \phi) = \mathbf{C}\ \mathbf{a}(\theta, \phi)$.

Expressions for the coupling matrix have been presented by several authors [GK83, LC95], however without including the expressions for the voltages which are essential for the direction finding. An interesting interpretation of the coupling is given in [LC95], where $\mathbf{C}$ is rewritten as

$$\frac{Z_T + Z_A}{Z_T}\left(\frac{\mathbf{Z}}{Z_T} + \mathbf{I}\right)^{-1} = \frac{Z_T + Z_A}{Z_T}\sum_{m=0}^{\infty}\left(-\frac{\mathbf{Z}}{Z_T}\right)^m. \tag{2.12}$$

The zeroth order contribution is the direct excitation from the wave and the first order term corresponds to the direct coupling between two dipoles. The second order term corresponds to the "two-trip" coupling and the higher order terms can be interpreted in a similar manner.

The derivation of the measured voltages considered only one incident wave and no noise. In all measurement situations, noise will inevitably appear and it is difficult to model. Often, receiver noise is the dominating noise source; and then additive white Gaussian noise is a good model. Furthermore, to simplify the notation in the following analysis, the waves are assumed to arrive in the x-y plane ($\theta = 90°$), see Figure 2.1. The model when p uniform plane waves or sources are incident then becomes

$$\mathbf{x}(t) = \mathbf{C}\mathbf{A}(\boldsymbol{\phi})\mathbf{s}(t) + \mathbf{n}(t), \tag{2.13}$$

where the vector of measured voltages, at time t, $\mathbf{x}(t)$ is $n \times 1$, the coupling matrix $\mathbf{C}$ is $n \times n$, the steering matrix $\mathbf{A}(\boldsymbol{\phi})$ is $n \times p$, the signal vector $\mathbf{s}(t)$ is $p \times 1$ and the noise vector $\mathbf{n}(t)$ is $n \times 1$. The DOAs are contained in the parameter vector $\boldsymbol{\phi}$. In order for the analysis in the following sections to be valid, some additional assumptions are needed:

- the coupling matrix has full rank, i.e., $\mathrm{rk}(\mathbf{C}) = n$
- $\mathbf{n}(t)$ is temporally white and circularly Gaussian distributed: $\mathbf{n}(t) \in \mathcal{N}(0, \sigma^2\mathbf{I})$
- $\mathbf{s}(t)$ is also temporally white and circularly Gaussian distributed: $\mathbf{s}(t) \in \mathcal{N}(0, \mathbf{P})$

The noise is both spatially and temporally white, while the signal is only temporally white. Furthermore, the signal is assumed to be uncorrelated with the noise. The covariance matrix of the measured voltages $\mathbf{x}(t)$ becomes [KV96]

$$\mathbf{R} = E\{\mathbf{x}(t)\mathbf{x}(t)^H\} = \mathbf{CAPA}^H\mathbf{C}^H + \sigma^2\mathbf{I}. \tag{2.14}$$

Note that the eigendecomposition of the covariance matrix $\mathbf{R}$ can be written as a sum of two parts [KV96]. One part consisting of eigenvectors corresponding to eigenvalues equal to the noise variance, and a second part that is related to the signal

$$\mathbf{R} = \sum_{k=1}^{n} \lambda_k \mathbf{e}_k \mathbf{e}_k^H = \mathbf{E}_s \mathbf{\Lambda}_s \mathbf{E}_s^H + \sigma^2 \mathbf{E}_n \mathbf{E}_n^H, \tag{2.15}$$

where $\mathbf{E}_s = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$ denotes the signal eigenvectors, $\mathbf{E}_n = [\mathbf{e}_{p+1}, \ldots, \mathbf{e}_n]$ denotes the noise eigenvectors, and $\mathbf{\Lambda}_s = \mathrm{diag}[\lambda_1, \ldots, \lambda_p]$ denotes the signal eigenvalues. This decomposition will be used extensively when dealing with subspace based estimation in later sections.

2.3 Direction Finding Accuracy

The model in the previous section will now be used to examine how mutual coupling between the elements of the array affects the DOA estimation accuracy. Since it is the effect of the coupling in the DOA estimation that is interesting and not a particular estimator, at least not in the first case, a measure of how good DOA estimates the model can deliver is needed. Here, the Cramér-Rao lower Bound is used which gives a lower bound on the variance of the estimated angles for an unbiased estimator.

Now, it is interesting to examine what happens to this bound when coupling is included. It is important to note that the coupling is assumed to be known and the signals are assumed to be non-random in the following calculation of the CRB for the DOA:s.

For sufficiently large number of samples N, the (asymptotic) Cramér-Rao lower Bound can be derived using direct differentiation [CTO89, Hay95b, SN89]. The introduction of a known coupling is accounted for by changing the steering vector in the derivation to an effective steering vector $\mathbf{CA}(\boldsymbol{\phi})$. This does not affect the derivation much, and the asymptotic CRB inequality becomes

$$E\{(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0)(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0)^T\} \geq \mathbf{B} \tag{2.16}$$

$$\mathbf{B} = \frac{\sigma^2}{2N}\left[Re\{(\mathbf{D}^H\mathbf{C}^H\mathbf{P}_{CA}^{\perp}\mathbf{CD}) \odot \mathbf{S}^T\}\right]^{-1}, \tag{2.17}$$

where

$$\mathbf{D} = \left[\frac{\partial \mathbf{a}(\phi)}{\partial \phi} \Big|_{\phi=\phi_1}, \cdots, \frac{\partial \mathbf{a}(\phi)}{\partial \phi} \Big|_{\phi=\phi_p} \right] \tag{2.18}$$

$$\mathbf{P}_{CA}^{\perp} = \mathbf{I} - \mathbf{CA}(\mathbf{A}^H \mathbf{C}^H \mathbf{CA})^{-1} \mathbf{A}^H \mathbf{C}^H \tag{2.19}$$

$$\mathbf{S} = \lim_{N \to \infty} \frac{1}{N} \sum_{t=t_1}^{t_N} \mathbf{s}(t) \mathbf{s}^H(t) \ . \tag{2.20}$$

Here, $\odot$ denotes the Hadamard (or Schur) product, i.e., element-wise multiplication and $\mathbf{P}_{CA}^{\perp}$ is the orthogonal projector on the null space of $(\mathbf{CA})^H$. The matrix $\mathbf{S}$ is the limiting signal sample covariance matrix.

The bound in (2.17) represents the lowest possible variance when using an unbiased estimator $\hat{\boldsymbol{\phi}}$. It applies to a thought situation where the experiment of obtaining N samples and estimating the angles is performed many times with the same signal vector $\mathbf{s}(t)$, but different noise realizations $\mathbf{n}(t)$.

Next, this bound is calculated for a few different scenarios and examined when coupling is included and not.

2.4 Computer Experiments

To examine the effect of mutual coupling, the CRB from the previous section is calculated for a few different scenarios. The CRB when coupling is included is compared with the CRB when coupling is excluded. This calculation is carried out for different locations of the source signals, correlated signals, and for different element spacings. In all of these examples the terminating impedance is chosen as the complex conjugate of the single antenna impedance.

2.4.1 Angular Dependence

Consider a five element half-wavelength dipole array with element separation of $\lambda/2$. Two plane waves are incident upon the array and the DOA:s (ϕ_1, ϕ_2) of these are to be estimated. The signal waveforms are uncorrelated and the signal power is equal to the noise power (SNR= 0 dB) and $N = 200$ snapshots are collected.

The logarithm of the CRB for ϕ_2 with coupling is shown in Figure 2.4. The fact that it is harder to estimate angles close together is obvious from the figure. Note that there is a corresponding figure for the other angle ϕ_1. For angles close to end-fire ($\phi = 0°$) the CRB is quite large, and this is due

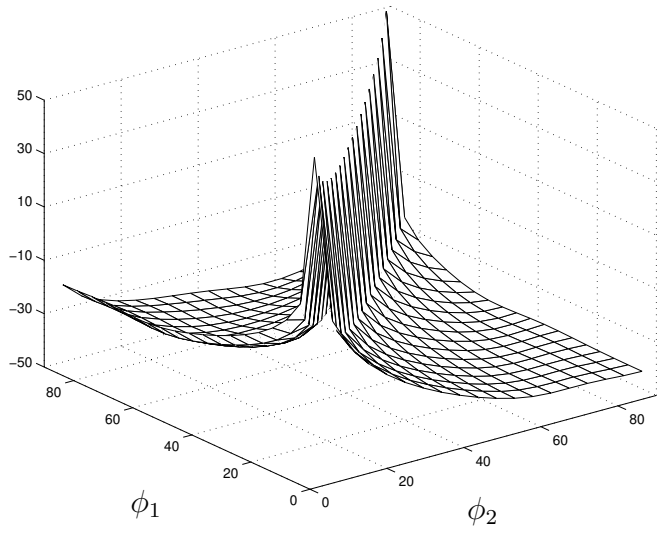


Figure 2.4: The logarithm of $CRB_{coupling}$ for ϕ_2 versus the angles for an array of 5 $\lambda/2$ dipoles spaced $\lambda/2$ apart.

to the fact that the CRB is inversely proportional to the derivative of the steering vector, and for $\phi_2 = 0°$ the derivative is zero.

Now it is interesting to compare the CRB with and without coupling. In Figure 2.5, the ratio of the standard deviation with and without coupling is shown. Here, it is found that the coupling does not significantly change the CRB bound if the coupling is known. This result depends on the geometrical configuration of the array and the array elements, but also on several parameters such as number of elements, correlation of the signals and element spacing. These dependencies are studied next and the influence of the number of elements is studied first.

In Figure 2.6, the CRB ratio for an array of 25 elements (otherwise the same as previous array) is shown. A larger difference is found here. The CRB with coupling gets lower as ϕ_2 gets larger. This can partly be explained by considering the received power which is maximum for $\phi_2 = 90°$, see [Sva99b]. This effect gets larger for a larger array, and that is why it did not show in the first figure for the array with few elements. If the change in received power is compensated for, the ratio will mostly reside at unity except for angles near end-fire ($\phi_2 = 0°$), where it is hard to estimate the DOA. This is shown in Figure 2.7.

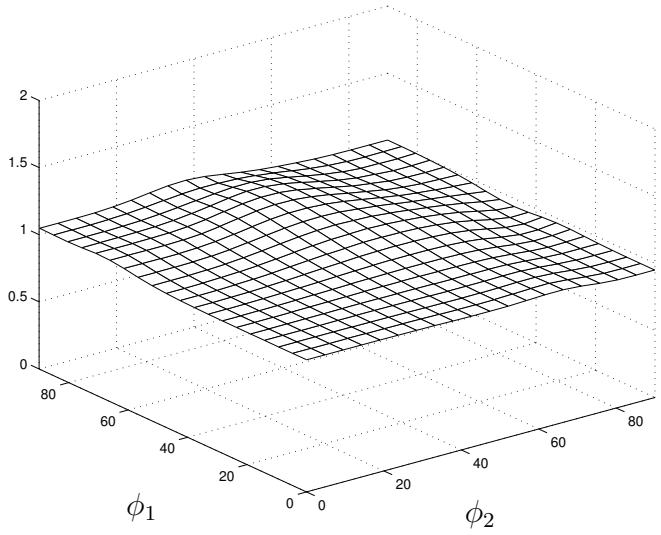


Figure 2.5: The ratio of the standard deviation for ϕ_2 with and without coupling versus the angles for an array of 5 $\lambda/2$ dipoles spaced $\lambda/2$ apart.

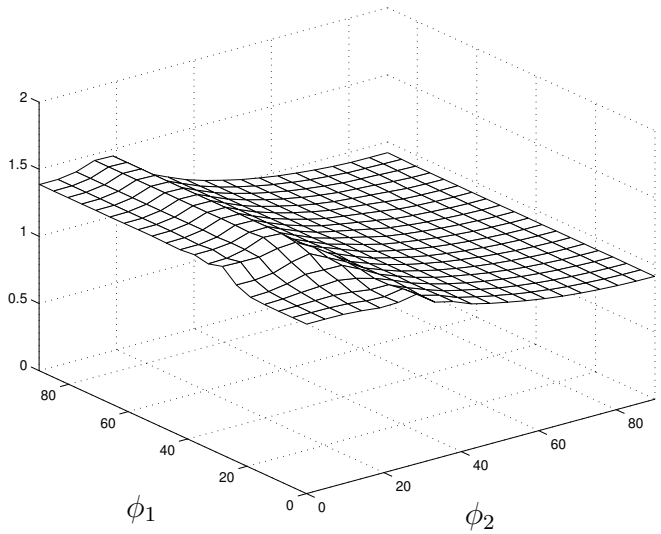


Figure 2.6: The ratio of the standard deviation for ϕ_2 with and without coupling versus the angles for an array of 25 $\lambda/2$ dipoles spaced $\lambda/2$ apart.

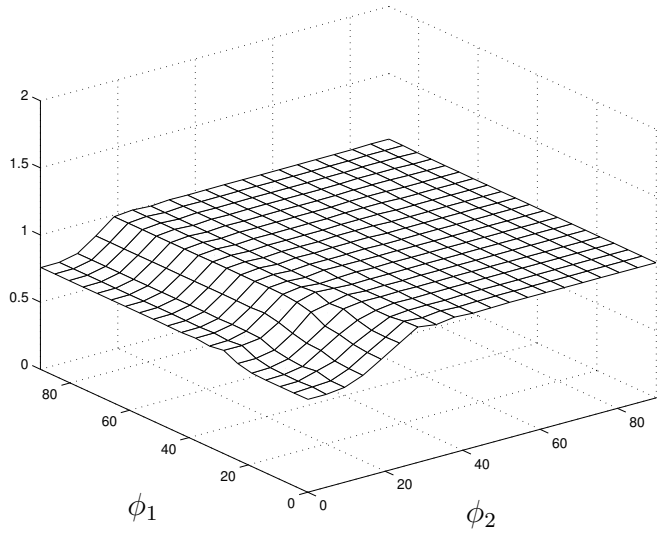


Figure 2.7: The power corrected ratio of the standard deviation for ϕ_2 with and without coupling versus the angles for an array of 25 $\lambda/2$ dipoles spaced $\lambda/2$ apart.

2.4.2 Correlated Signals

Consider a 25 element half-wavelength dipole array with element separation of $\lambda/2$. Two plane waves are incident upon the array and the DOA:s $(84°, 85°)$ are to be estimated. The signal to noise ratio (SNR) is -10dB and $N = 1000$ snapshots are collected. In this simulation, the correlation dependence of the CRB is examined. The signal covariance matrix is

$$\mathbf{S} = 10^{(\mathrm{SNR}/10)} \times \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \tag{2.21}$$

The CRB is now calculated for different values of ρ. In Figure 2.8, the standard deviation for the wave incident from 84° is calculated for different correlations. First of all it is noted that it becomes harder to estimate the DOA when the signal correlation increases. But the effects of coupling do not depend on the correlation. The ratio of the standard deviation with coupling (solid) and without (dashed) is essentially unchanged by the change in correlation. Here, coupling actually improved the estimation performance but that depends on the incidence angles as discussed previously, see Figure 2.6. The conclusion that the effects of coupling are almost independent of the correlation remains.

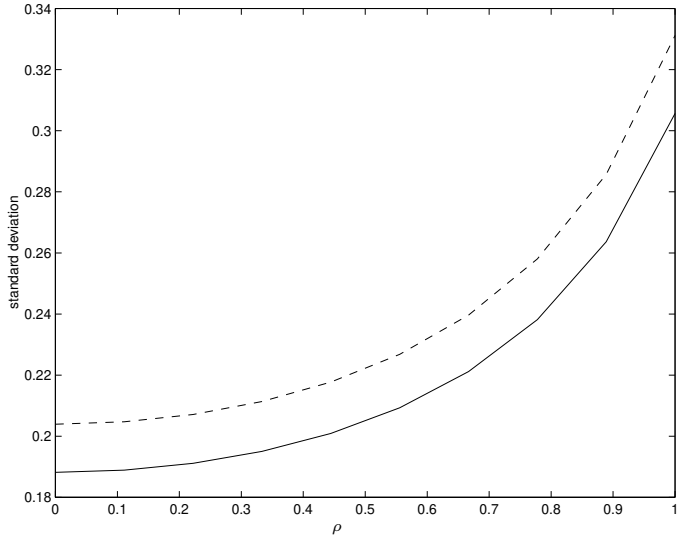


Figure 2.8: The standard deviation with coupling (solid) and without (dashed) for the wave incident from 84° upon an array of 25 $\lambda/2$ dipoles spaced $\lambda/2$ apart for different correlations.

2.4.3 Element Spacing

Consider a 25 element half-wavelength dipole array with element separation of $\lambda/2$. Two plane waves are incident upon the array and the DOA:s $(84°, 85°)$ are to be estimated. The signals are uncorrelated and the SNR is -10dB and $N = 1000$ snapshots are assumed.

In this scenario, the influence of the element spacing of the CRB is examined. The electrical size has a large effect on the prospect of estimating the DOA:s as well as the radiation pattern. With larger element spacing the electrical size of the antenna increases and the estimation performance is increased. The effects of element spacing on the ratio of the standard deviation with and without coupling is shown in Figure 2.9. Here it is found that the ratio approaches unity as the spacing gets larger. Another perhaps obvious result is that this ratio plot is similar to the corresponding plot of the mutual impedance, see Figure 2.2.

When the coupling is known, the impact on the CRB is not so large and for the DOA:s $(84°, 85°)$, and a moderate element spacing, the coupling only changes the CRB about 10-20%. But as discussed earlier, for some DOA:s close to end-fire the changes are larger.

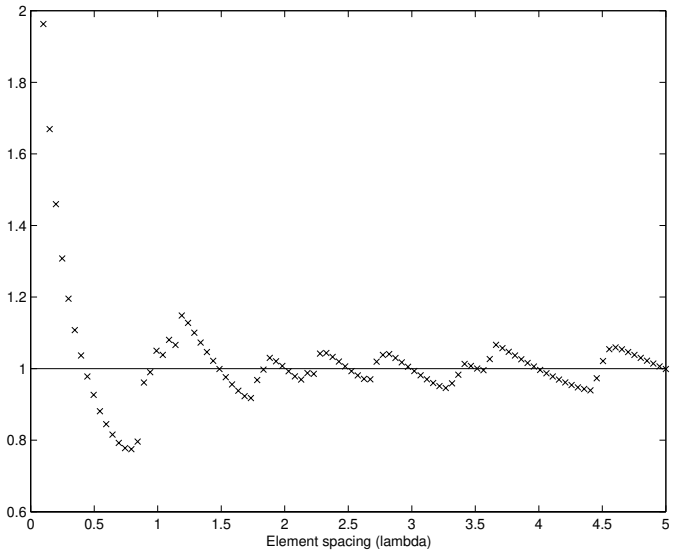


Figure 2.9: The ratio of the standard deviation with and without coupling for the wave incident from 84° upon an array of 25 $\lambda/2$ dipoles for different element spacings.

2.5 Conclusions

The problem of how a known coupling affects the prospects of estimating the directions of arrival using an array of elements was addressed. The direction finding accuracy, i.e. the Cramér-Rao lower Bound on the estimated DOA:s, was used as a measure of these effects. The induced voltages in a linear array of dipole elements, including mutual coupling, was derived. Hence, a model for the mutual coupling was obtained. Some stochastical assumptions were also introduced that were needed when calculating the lower bound on the variance of the DOA estimates.

The CRB with a known mutual coupling was presented and analyzed for a few different estimation scenarios. It was found that the coupling effects on the CRB were dependent on the DOA:s but the effects was relatively small. Furthermore, the correlations of the signals did not change the coupling effects, considering the CRB, appreciably. The element separation did affect the CRB. But for moderate element spacings, the effects were not large.

To sum up the chapter, the effects of a known mutual coupling in the scenarios studied here are not large (of course depending on the application). Depending on the geometrical configuration, the array elements, number of elements, and element spacing, the coupling increases or decreases the CRB.

In the analysis of the CRB, the coupling was introduced through the coupling matrix $\mathbf{C}$. Other effects than coupling can also be included in the $\mathbf{C}$ matrix. For example, effects such as the gain and phase of the individual sensors could be included. In that case, the analysis of the CRB in this chapter is still valid by simply changing the $\mathbf{C}$ matrix to include those effects. In the next chapter, some methods of estimating the DOAs in the presence of a known coupling will be introduced.

Chapter 3

Estimation With a Known Coupling

The problem of estimating the DOA, when a known coupling is present, will be considered in this chapter. The mutual coupling is calculated using the model that was derived in Section 2.2 for the induced voltages on a Uniform Linear Array (ULA) of thin and finite length dipoles. First, the model will be reviewed and some statistical assumptions regarding the signals and the noise will be introduced. In Chapter 2, the CRB was used to examine the estimation performance. Here, the corresponding bound is derived, using the statistical assumptions made in this chapter.

Several methods for estimating the DOA are then reviewed and thus also a small part of the area of sensor array processing. However, the methods are discussed in the context of mutual coupling and the algorithms that have appeared in the literature that include a known coupling are reviewed first. After that, some straightforward extensions of existing methods for the coupling free case are introduced. Finally, the problem of estimating the number of signals is briefly discussed.

3.1 Data Model

The direction of an incident wave can be estimated by measuring the received voltages at the different elements of the array and by assuming the wave to be plane, simple geometry then gives the direction. However, due to mutual coupling, the measured voltage at each element will depend not only on the incident field but also on the voltages on the other elements. The received voltage on each element will induce a current on the element which in turn radiates a field which affects the surrounding elements, i.e. mutual coupling. Here, mutual coupling in an array of n dipoles of finite length,

l, is considered. To simplify the analysis the dipoles are considered thin, i.e. the radius $a \ll l$. The dipoles are placed linearly side-by-side with the same element separation, d, resulting in a ULA. The received voltages from this array, when p far-field narrow-band sources are used, are derived using basic electromagnetic concepts in Section 2.2. The resulting model for the measured voltages at time t is

$$\mathbf{x}(t) = \mathbf{C}\mathbf{A}(\boldsymbol{\phi})\mathbf{s}(t) + \mathbf{n}(t), \tag{3.1}$$

where the vector of measured voltages $\mathbf{x}(t)$ is $n \times 1$, the coupling matrix $\mathbf{C}$ is $n \times n$, the steering matrix $\mathbf{A}(\boldsymbol{\phi})$ is $n \times p$, the signal vector $\mathbf{s}(t)$ is $p \times 1$ and the noise vector $\mathbf{n}(t)$ is $n \times 1$. The DOAs are contained in the parameter vector $\boldsymbol{\phi}$. The steering matrix of a ULA has a Vandermonde structure [Lüt96] with elements $[\mathbf{A}(\boldsymbol{\phi})]_{vw} = e^{-jkd(v-1)\cos\phi_w}$, where k is the wavenumber. The data model in (3.1) is identical to the usual data model used in array processing except for the coupling matrix

$$\mathbf{C} = (Z_A + Z_T)(\mathbf{Z} + Z_T\mathbf{I})^{-1}, \tag{3.2}$$

where Z_A is the antenna impedance, Z_T is the impedance of the measurement equipment at each element, and $\mathbf{Z}$ is the mutual impedance matrix. The impedance of the measurement equipment, Z_T, is often chosen as the complex conjugate of the dipole impedance in order to reduce the power loss. However, other choices of Z_T are possible. If Z_T is chosen to be large, the coupling matrix in (3.2) approaches the unity matrix [RW92]. In that case the effects of mutual coupling are mitigated at the cost of a large impedance mismatch and consequently a large power loss. For a further discussion about impedance matching, see [AR76, Spe96], where the impedance matching is addressed in a beamforming network context that includes mutual coupling. Here, the terminating impedance is chosen as the complex conjugate of the dipole impedance.

The model in (3.1) needs to be complemented with some additional assumptions that are used in the following sections.

- the coupling matrix has full rank, i.e. $\mathrm{rk}(\mathbf{C}) = n$, which implies that $\mathrm{rk}(\mathbf{C}\mathbf{A}(\boldsymbol{\phi})) = \mathrm{rk}\,(\mathbf{A}(\boldsymbol{\phi})) = p$ for distinct DOAs
- $\mathbf{n}(t)$ is circularly Gaussian distributed $E\{\mathbf{n}(t)\} = 0$, $E\{\mathbf{n}(t)\mathbf{n}^H(s)\} = \sigma^2\mathbf{I}\ \delta_{ts}$ and $E\{\mathbf{n}(t)\mathbf{n}^T(s)\} = \mathbf{0}\ \forall\ t, s$
- $\mathbf{s}(t)$ is also circularly Gaussian distributed $E\{\mathbf{s}(t)\} = 0$, $E\{\mathbf{s}(t)\mathbf{s}^H(s)\} = \mathbf{P}\ \delta_{ts}$ where $\mathbf{P}$ is full rank and $E\{\mathbf{s}(t)\mathbf{s}^T(s)\} = \mathbf{0}\ \forall\ t, s$

- $n > p$, i.e. the number of dipoles are larger than the number of incident signals

Note that δ_{ts} denotes the Kronecker delta and that both the signal and the noise are random. Specifically, they are both assumed to be Gaussian distributed and this model is sometimes called the stochastic or unconditional signal model. In Chapter 2, the signal was assumed to be non-random and that model is therefore sometimes called the deterministic or the conditional signal model. However, since the signal in that case is considered unknown and hence also estimated along with the DOAs, the number of unknowns increases as the number of samples N increases. Actually, it is shown in [OVK90, SN90a] that the asymptotic (large N) covariance of the DOAs estimates using the deterministic signal model is larger than the asymptotic covariance for a stochastic signal model when $n < \infty$. Furthermore, and most importantly, the Cramér-Rao lower Bound derived using the deterministic signal model cannot be reached and therefore is not asymptotically tight, in contrast to the bound based on the stochastic signal model that is asymptotically tight.

Since both the signal and the noise are assumed to be Gaussian, all information about the DOAs and the signals are contained in the first and second order moments. Many estimation schemes have appeared in the literature that exploit the properties of the second moment to estimate the DOAs. Some of these schemes will be discussed in the following sections. But first, in order to assess the quality of the different estimation methods, a lower bound for the variance of the DOA estimates will be derived.

3.2 Direction Finding Accuracy

Most proposed estimators are asymptotically unbiased and the variance of the estimates is therefore a good figure of merit for these estimators. A lower bound on the variance of DOA estimates is therefore very useful when addressing the quality of a estimation schemes. Probably the most widely used bound is the Cramér-Rao lower Bound, which is a lower bound on the variance of DOA estimates, assuming that the estimator is unbiased. There are unbiased estimators that, at least asymptotically, attains the bound and hence the bound is called asymptotically tight and the estimators are called efficient. The CRB therefore serves as a benchmark when evaluating different estimation methods. In Chapter 2, the effects of mutual coupling on the DOA estimation was examined by calculating the CRB using the model in (3.1) and the deterministic signal model. Here, the CRB will be derived for a stochastic

signal model, since the bound in this case becomes asymptotically tight and thus more appropriate when considering different estimation schemes.

Compact forms for the CRB for the stochastic signal model were derived for the coupling free case in [OWVK89, SN90b]. Coupling can be included in those derivations by simply changing the steering matrix $\mathbf{A}$ to $\mathbf{CA}$. Note that each column of $\mathbf{CA}$ now contains contributions from several different elements. The Cramér-Rao lower Bound when coupling is included can be written as

$$E\{(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0)(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0)^T\} \geq \mathbf{B}_{STO} \tag{3.3}$$

$$\mathbf{B}_{STO} = \frac{\sigma^2}{2N}\left[Re\{(\mathbf{D}^H\mathbf{C}^H\mathbf{P}_{CA}^{\perp}\mathbf{CD}) \odot (\mathbf{PA}^H\mathbf{C}^H\mathbf{R}^{-1}\mathbf{CAP})^T\}\right]^{-1} \quad , \tag{3.4}$$

where

$$\mathbf{D} = \left[\frac{\partial \mathbf{a}(\phi)}{\partial \phi}\Big|_{\phi=\phi_1}, \cdots, \frac{\partial \mathbf{a}(\phi)}{\partial \phi}\Big|_{\phi=\phi_P}\right] \tag{3.5}$$

$$\mathbf{P}_{CA}^{\perp} = \mathbf{I} - \mathbf{P}_{CA} = \mathbf{I} - \mathbf{CA}(\mathbf{CA})^{\dagger} \tag{3.6}$$

Here, $\odot$ denotes the Hadamard (or Schur) product, i.e., element-wise multiplication, $\mathbf{P}_{CA}^{\perp}$ is the orthogonal projector on the null space of $(\mathbf{CA})^H$. The orthogonal projector on the column space of $(\mathbf{CA})^H$ is denoted $\mathbf{P}_{CA}$ and $\mathbf{M}^{\dagger}$ is the Moore-Penrose pseudo inverse of $\mathbf{M}$, which here can be written as $\mathbf{M}^{\dagger} = (\mathbf{M}^H\mathbf{M})^{-1}\mathbf{M}^H$. The matrix $\mathbf{R}$ is the covariance matrix of the measured voltages $\mathbf{x}(t)$.

$$\mathbf{R} = \mathbf{CAPA}^H\mathbf{C}^H + \sigma^2\mathbf{I}. \tag{3.7}$$

Next, it is examined how the mutual coupling affects the estimation performance. The CRB is calculated both with a known mutual coupling and without. In Figure 3.1, the ratio of the standard deviation with and without coupling for ϕ_2, when two waves from (ϕ_1, ϕ_2) impinge on an array of 5 $\lambda/2$ dipoles spaced $\lambda/2$ apart, is shown. The ratio mostly resides close to unity and thus in this respect does mutual coupling not affect the estimation accuracy much.

This is in agreement with the results in Chapter 2, where the CRB was calculated using the deterministic signal model in a few different scenarios. Comparing with Figure 2.5 further supports this conclusion and no fundamental differences between the stochastic and deterministic CRB concerning mutual coupling is found in this case.

However, in the previous section it was mentioned that the CRB is larger for the stochastic than for the deterministic signal model. In Figure 3.2, the logarithm of the ratio of the standard deviation for stochastic and deterministic signal model for ϕ_2 versus the angles for the same array as in

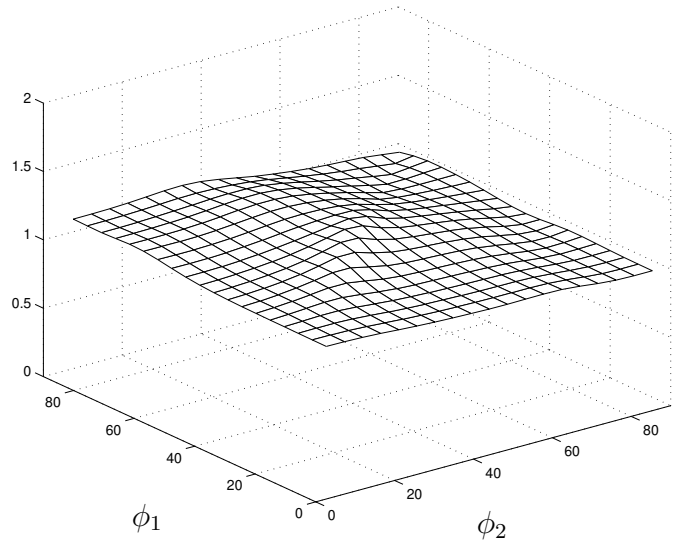


Figure 3.1: The ratio of the standard deviation with and without coupling for ϕ_2 versus the angles for an array of 5 $\lambda/2$ dipoles spaced $\lambda/2$ apart.

Figure 3.1 is shown. The logarithm is positive for all angles, i.e. the CRB for the stochastic signal model is larger than for the deterministic model as stated previously. The difference is not very large for angles with sufficient angular separation. Also, the large ridge through the plot corresponds to estimating DOAs for two signals arriving from the same direction. In this case, the estimation procedure breaks down and it is hard to draw any conclusions regarding these angles.

The effects of mutual coupling on the CRB using the stochastic signal model are thus more or less the same as on the CRB calculated for the deterministic signal model. Now, a measure for the quality of the proposed DOA estimation schemes is derived. In the next section, several methods for estimating the DOAs in the presence of a known coupling will be studied and compared to the CRB.

3.3 Estimation Methods

The problem of estimating the DOA in the presence of a known coupling is not very different from the usual DOA estimation problem. Most algorithms which have appeared in the last decades can be modified to include a known mutual coupling. Essentially, it is just to change the steering matrix $\mathbf{A}$ to an

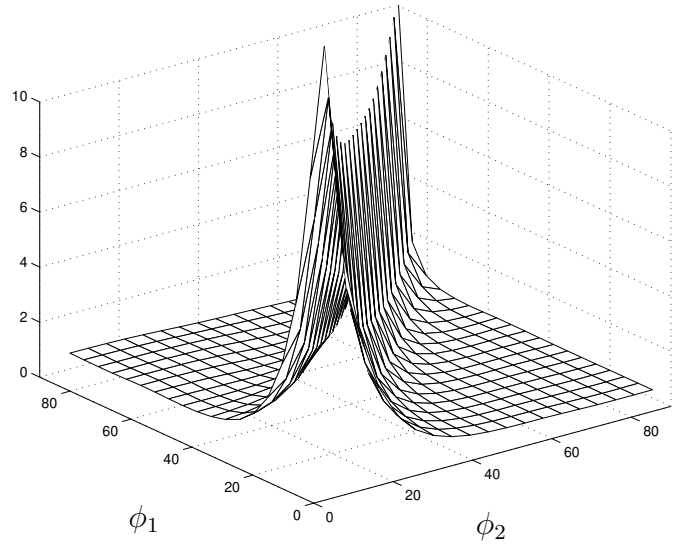


Figure 3.2: The logarithm of the ratio of the standard deviation for stochastic and deterministic signal model for ϕ_2 versus the angles for an array of 5 $\lambda/2$ dipoles spaced $\lambda/2$ apart.

effective steering matrix $\mathbf{CA}$. However, if the algorithm uses some special structure of the steering matrix, it can not, at least not directly, be used when mutual coupling is present. Many algorithms exploit the second order properties of the covariance matrix of the measured voltages. Therefore, the properties of the covariance matrix will be examined first. Then, algorithms that have appeared in the literature, that include a known coupling, are reviewed and this part also serves as an introduction to the area of sensor array processing. Also, some straightforward extensions of other coupling free algorithms are presented.

3.3.1 The Structure of the Covariance Matrix

Using the model for the measured voltages in (3.1), (3.2), and the statistical assumptions, the covariance matrix of the measured voltages becomes

$$\begin{aligned}\mathbf{R} &= E\{\mathbf{x}(t)\mathbf{x}(t)^H\} = \mathbf{A}_e E\{\mathbf{s}(t)\mathbf{s}(t)^H\}\mathbf{A}_e^H + E\{\mathbf{n}\mathbf{n}^H\} \\ &= \mathbf{A}_e\mathbf{P}\mathbf{A}_e^H + \sigma^2\mathbf{I},\end{aligned} \tag{3.8}$$

where $\mathbf{A}_e = \mathbf{CA}(\phi)$. Note that the cross-terms vanish since the signal and the noise are assumed to be uncorrelated. Assume that there is one vector $\mathbf{y}$

that is perpendicular to the steering matrix $\mathbf{A}_e$, or

$$\mathbf{A}_e^H \mathbf{y} = \mathbf{0}. \tag{3.9}$$

Then $\mathbf{R}\mathbf{y} = \sigma^2\mathbf{y}$, i.e. $\mathbf{y}$ is an eigenvector of $\mathbf{R}$ with corresponding eigenvalue σ^2. Since the steering matrix $\mathbf{A}_e$ is full rank with dimensions $n \times p$ and $\mathbf{P}$ also is full rank, there are $n - p$ independent vectors $\mathbf{y}$ that are perpendicular to $\mathbf{A}_e$. Actually, it is possible to write the eigendecomposition of the covariance matrix as a sum of two parts. One part consisting of eigenvectors corresponding to eigenvalues equal to the noise variance, and a second part that is related to the signal

$$\mathbf{R} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^H = \mathbf{E}_s\mathbf{\Lambda}_s\mathbf{E}_s^H + \sigma^2\mathbf{E}_n\mathbf{E}_n^H. \tag{3.10}$$

Furthermore, by exploiting the orthonormality of the eigenvectors

$$\mathbf{E}_s\mathbf{E}_s^H + \mathbf{E}_n\mathbf{E}_n^H = \mathbf{I}, \tag{3.11}$$

the covariance matrix can be written as

$$\mathbf{R} = \mathbf{E}_s(\mathbf{\Lambda}_s - \sigma^2\mathbf{I})\mathbf{E}_s^H + \sigma^2\mathbf{I} = \mathbf{E}_s\widetilde{\mathbf{\Lambda}}\mathbf{E}_s^H + \sigma^2\mathbf{I}. \tag{3.12}$$

Comparing (3.12) with (3.8) reveals that the subspace spanned by the columns of $\mathbf{E}_s$ must be equal to the subspace spanned by the columns of $\mathbf{A}_e\mathbf{P}$ or

$$\mathcal{R}\{\mathbf{E}_s\} = \mathcal{R}\{\mathbf{A}_e\mathbf{P}\} \subseteq \mathcal{R}\{\mathbf{A}_e\} \tag{3.13}$$

with equality if the signal covariance matrix $\mathbf{P}$ is of full rank. The subspace spanned by $\mathbf{A}_e$ (or $\mathbf{E}_s$) is usually called the signal subspace and the corresponding space spanned by $\mathbf{E}_n$ is therefore called the noise subspace. Note that due to the orthogonality of the eigenvectors, these spaces are orthogonal.

A related entity is the array manifold $\mathcal{A}$, that is defined as the collection of all steering vectors within the DOA range, or

$$\mathcal{A} = \{\mathbf{a}(\phi_i) \mid \phi_i \in \Phi\}. \tag{3.14}$$

The array manifold is a p-dimensional subset of the n-dimensional space and this observation is important in the derivation of the subspace based methods presented later.

Almost all methods discussed in the following sections will exploit the properties of the covariance matrix of the measured voltages $\mathbf{R}$. However, it is the induced voltages $\mathbf{x}(t)$ that are measured and not the covariance matrix. Usually, $\mathbf{R}$ is estimated as

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{x}(t)\mathbf{x}^H(t) \tag{3.15}$$

which is an unstructured maximum likelihood estimate of the covariance matrix in (3.8).

The structure in (3.8) and (3.12) has been exploited in numerous DOA estimation schemes, presented during the latest decades. For a review of some of these methods, see [Hay95a, KV96, VS01]. Many of those methods are possible to use also in the presence of mutual coupling with slight changes, but in the signal processing literature only a few methods, that include a known coupling, have been presented. Those methods will be reviewed in the next subsection, and then some additional methods that include a known coupling are proposed.

3.3.2 Estimation Methods: Beamforming and MUSIC

Probably, the most used method of obtaining estimates of the DOAs is beamforming [Sko62, VB88]. As the name beamforming suggests, the received energy is focused to one direction (or beam) at a time. This can be expressed as

$$y(\phi) = \mathbf{a}^H(\phi)\mathbf{x}. \tag{3.16}$$

A spatial spectrum is obtained by calculating $|y(\phi)|^2$ for $\phi \in \Phi$. The presence of mutual coupling can be taken care of by changing the multiplication in (3.16) from $\mathbf{a}^H(\phi)$ to $\mathbf{a}(\phi)^H\mathbf{C}^{-1}$, i.e. premultiply the measured voltages $\mathbf{x}(t)$ with $\mathbf{C}^{-1}$ as suggested in [Jos94, SH90]. The noise covariance matrix then changes from $\sigma^2\mathbf{I}$ to $\sigma^2\mathbf{C}^{-1}\mathbf{C}^{-H}$. If the method relies on that the noise covariance is white ($\sigma^2\mathbf{I}$), the premultiplication method does not work. Note that the beamforming method ignores the noise and premultiplication can thus be applied. Similar methods, but more oriented towards pattern design, can be found in [IN71, KP85, SH69].

The spatial spectrum in (3.16) is a spatial analogue of the classical periodogram in time series analysis. As with the periodogram, the spatial spectrum has a resolution threshold. Waves arriving with electrical angle separation[1] less than $2\pi/n$ can not be resolved with this method. That was one of the motivations for the interest in the so-called subspace methods that will be described next.

Subspace based methods rely on observations regarding the eigendecomposition of the covariance matrix into a signal subspace and a noise subspace, as discussed in previous subsection. One of the most popular subspace methods, MUSIC, was introduced in [Sch79]. The method is based on the observation in (3.9), that the noise eigenvectors are perpendicular to the steering matrix or the signal subspace. The algorithm calculates the noise subspace

[1]The electrical angle is defined as $kd\ cos\phi$.

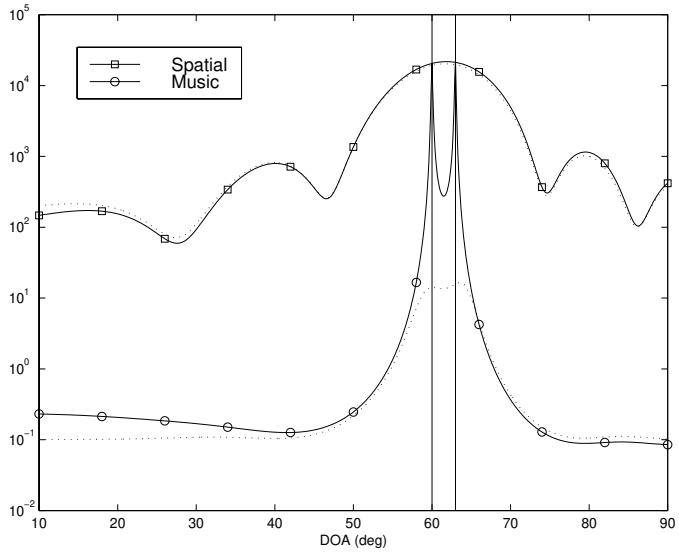


Figure 3.3: The spatial and MUSIC spectrum when coupling is compensated for (solid) and not (dotted). An array of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart is used and two waves are incident from $(60°, 63°)$, and 1000 snapshots with SNR = 20dB.

using an eigendecomposition of the estimated covariance matrix in (3.15). Then, the estimates of the DOAs are taken as those ϕ that gives in the smallest value of $\mathbf{a}^H(\phi)\hat{\mathbf{E}}_n$. Usually this is formulated as finding the p largest peaks in the "MUSIC spectrum"

$$P_{MU}(\phi) = \frac{1}{\mathbf{a}^H(\phi)\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\mathbf{a}(\phi)}. \tag{3.17}$$

This algorithm is straightforwardly adopted to the case of a mutual coupling by simply exchanging the steering vector $\mathbf{a}(\phi)$ to $\mathbf{Ca}(\phi)$. Another way of handling the coupling is to employ the method of premultiplication as described above. But as mentioned previously, that changes the noise covariance matrix and a generalized eigendecomposition has to be used in order to obtain the correct noise eigenvectors. These two methods have been investigated in [LZ90, PF94, YLU89]. Here, the steering vector is changed and in Figure 3.3, the MUSIC spectrum of (3.17) is shown when two sources are present at $(60°, 63°)$, using an array with 10 $\lambda/2$ elements spaced $\lambda/2$ apart. The solid line is when the coupling is compensated for, i.e. using the correct coupling matrix, and the dotted line is when the coupling matrix is unknown and replaced by the unity matrix. With a known coupling, the method works

well and two peaks at the correct angles appear. However, if the coupling is unknown the method does not work well. Also included in Figure 3.3 is the corresponding spatial spectra for the beamforming method in (3.16). The beamforming method fails to resolve the sources, since in this case the resolution limit is more than 10°. The resolution threshold of MUSIC is much lower, and therefore does that method succeed in resolving the DOAs. This, of course, is the main reason for using subspace based methods like MUSIC instead of beamforming. However, the difference between when the coupling is known or not is small for the beamforming method, i.e. the beamforming method is robust. The MUSIC method, as all subspace based methods, is more sensitive to errors in the steering vector as seen in the figure.

An interesting application of the MUSIC method for Over The Horizon (OTH) radar arrays is found in [SAGA98], where ionized meteor trails are used to estimate the sensor positions and mutual coupling in a calibration scheme.

As mentioned above, several other algorithms are easily extended to include a known coupling. In the next section some straightforward extensions will be proposed.

3.3.3 Other Methods

Most DOA estimation algorithms can be extended to include a known coupling. In this subsection, some straightforward extensions of other DOA algorithms will be proposed. Some of these algorithms will be used later when the DOAs are estimated along with an unknown coupling. First two computationally simple algorithms, that exploit a special structure of the antenna array, will be considered.

The popular MUSIC method discussed in the previous subsection can be applied to any type of antenna array. However, if the antenna is a ULA, the estimation procedure can be simplified. In this case, the steering vector has elements $[\mathbf{a}(\phi)]_v = e^{-jkd(v-1)\cos\phi}$ and the expression in (3.17) can be viewed as a polynomial in $z = e^{-jkd\cos\phi}$. Therefore, the search for the DOAs in the MUSIC method described above can be avoided. Instead, find the roots of the polynomial

$$f(z) = \mathbf{a}^T(z^{-1})\mathbf{C}^H\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\mathbf{C}\mathbf{a}(z). \tag{3.18}$$

Pick the p roots closest to the unit circle, $\hat{z}_i$ and calculate the angle and solve for ϕ_i in $\angle\hat{z}_i = \angle e^{-jkd\cos\phi_i}$. This method is known in the literature as the Root-MUSIC method and was presented in [Bar83]. Not only does this method avoid the numerical search, but also improves the resolution threshold of the spectral MUSIC discussed in previous subsection. A performance

analysis in [RH89] shows that the improvement in resolution threshold is due to the fact that a small error in the radius of the root does not effect the estimation performance. Thus, only errors in the angle of the root effect the estimation performance and this decreases the resolution threshold. The presence of mutual coupling is taken care of by simply introducing the coupling matrix, $\mathbf{C}$, as in (3.18).

Another subspace-based method that also relies on a special structure of the array is Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [Roy87]. ESPRIT exploits a shift structure of a special steering matrix constructed as

$$\mathbf{A}_s = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_1\mathbf{\Phi} \end{bmatrix}, \tag{3.19}$$

where $\mathbf{A}_1$ is a $l \times p$ steering matrix and $\mathbf{\Phi}$ is $p \times p$ diagonal matrix with elements $\mathbf{\Phi}_{ii} = e^{-jk\Delta\cos\phi_i}$. Of course, not all arrays have this shift structure, but for example a ULA can be divided into two parts which have this structure. Several different strategies for forming these two parts are discussed in the literature [OVK91], but here maximum overlapping structures will be used, i.e. the first part consists of the first $n-1$ elements and the second part of the last $n-1$ elements, that is, $l = n-1$ and $\Delta = d$. ESPRIT exploits this shift structure and the fact that for the coupling free case $\mathcal{R}\{\mathbf{E}_s\} = \mathcal{R}\{\mathbf{A}(\boldsymbol{\phi})\}$ (assuming $\mathbf{P}$ full rank). When coupling is present $\mathcal{R}\{\mathbf{E}_s\} = \mathcal{R}\{\mathbf{C}\mathbf{A}(\boldsymbol{\phi})\}$ and the shift structure has vanished. However, $\mathcal{R}\{\mathbf{C}^{-1}\mathbf{E}_s\} = \mathcal{R}\{\mathbf{A}(\boldsymbol{\phi})\}$ and the shift structure is obtained. Another way of obtaining this result is to premultiply the data with $\mathbf{C}^{-1}$ and the same result is obtained by calculating the generalized eigenvalue factorization. Thus, there exists a nonsingular $\mathbf{T}$ such that

$$\mathbf{J}\mathbf{C}^{-1}\mathbf{E}_s = \mathbf{A}_s(\boldsymbol{\phi})\mathbf{T}, \tag{3.20}$$

where $\mathbf{J}$ is a $2(n-1) \times n$ selection matrix. The basic idea is to find those ϕ:s that best fit this relation. This fit can be made in a least squares sense or in a total least square sense [OVK91, Roy87]. Here, the total least squares method is used. In Figure 3.4, the Root Mean Square Error (RMSE) of the modified MUSIC and ESPRIT methods are shown. The Cramér-Rao lower Bound is also shown for the case with coupling and without. The MUSIC method is close to the CRB, but ESPRIT performs worse, since it only exploits the invariance structure and not the full steering matrix. The same relation is is found also in the coupling free case [OVK91]. However, it is interesting to observe that the CRB when a known coupling is present is slightly lower than the coupling free case. This is not generally true, but the difference is usually small, see Figure 3.1. A slightly different study of the

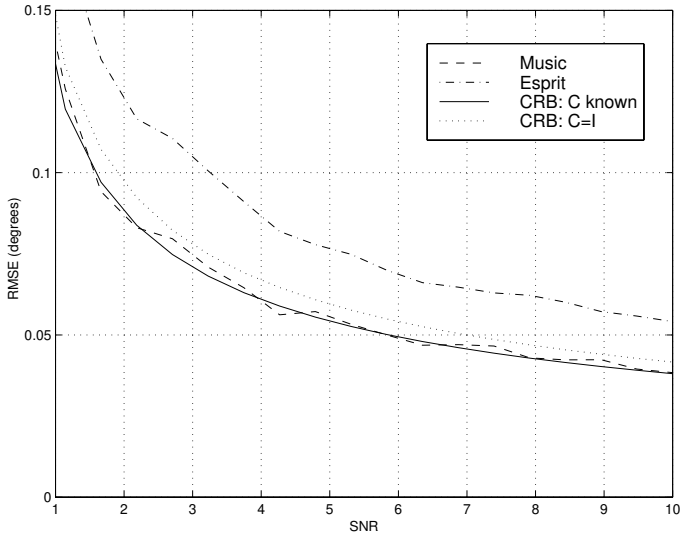


Figure 3.4: CRB and RMSE of Root-Music and ESPRIT for ϕ_1 using an array of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart when two waves are incident from $(80°, 85°)$ and 1000 snapshots with SNR = 20dB.

ESPRIT algorithm in the presence of mutual coupling has been presented in [HW90].

Perhaps the most well known and frequently used model-based estimation technique in signal processing is the ML method. The ML method assumes a model for the signals and also a statistical framework for the data generation. Two different assumptions regarding the signals have appeared in the signal processing literature and here the stochastic signal model is used, as discussed in Section 3.1. The method is therefore called the Stochastic Maximum Likelihood (SML) method. Using the model in (3.1) and the statistical assumptions in Section 3.1, the negative log likelihood function to be minimized becomes (with the parameter independent terms ignored)

$$l_{SML}(\boldsymbol{\phi}, \mathbf{P}, \sigma^2) = \log|\mathbf{R}| + \mathrm{Tr}\{\mathbf{R}^{-1}\hat{\mathbf{R}}\}, \tag{3.21}$$

where $\mathbf{R}$ and $\hat{\mathbf{R}}$ are defined in (3.8) and (3.15). This expression can be put in a concentrated form [Böh86, Jaf88, SN95]. For fixed $\boldsymbol{\phi}$, the minimizing value of $\mathbf{P}$ and σ^2 becomes

$$\hat{\sigma}^2(\boldsymbol{\phi}) = \frac{1}{n-p}\mathrm{Tr}\{\mathbf{P}_A^{\perp}\hat{\mathbf{R}}\} \tag{3.22}$$

$$\hat{\mathbf{P}}(\boldsymbol{\phi}) = \mathbf{A}^{\dagger}\left(\hat{\mathbf{R}} - \hat{\sigma}^2(\boldsymbol{\phi})\mathbf{I}\right)\mathbf{A}^{\dagger H}, \tag{3.23}$$

where $\mathbf{A}^\dagger$ denotes the Moore-Penrose pseudo inverse of $\mathbf{A}$. Inserting these expressions in (3.21), the concentrated negative log likelihood function is obtained

$$\hat{\boldsymbol{\phi}} = \arg\min_{\boldsymbol{\phi}} \left\{ \log |\mathbf{A}\hat{\mathbf{P}}(\boldsymbol{\phi})\mathbf{A}^H + \hat{\sigma}^2(\boldsymbol{\phi})\mathbf{I}| \right\} \tag{3.24}$$

and the DOA estimates are taken as the minimizing arguments of it. A known coupling is easily taken care of by simply exchanging the steering matrix $\mathbf{A}$ to the effective steering matrix $\mathbf{A}_e = \mathbf{C}\mathbf{A}$. The SML method is efficient since the variance of the DOA estimates attains the Cramér-Rao lower Bound. The ESPRIT method discussed previously does not attain the CRB, and the MUSIC method only attains the bound if the signals are uncorrelated and both the number of snapshots and the source signal power are large. The function in (3.22) is unfortunately a non-linear function of the DOAs, and therefore a numerical search is necessary. This search can be computationally heavy and that is why subspace-based methods like MUSIC and ESPRIT, that are less computationally intensive but not efficient, are attractive. The advantage of a lower computational load is thus traded for a higher variance of the estimates.

Another type of methods that have attracted considerable attention are the "Subspace Fitting" methods [SS90b, VO91, VOK91]. With a specific weighting of the criterion, those methods are efficient and thus have properties similar to the SML method. Also, connections to spectral-based methods such as MUSIC and beamforming as well as connections to ESPRIT, can be established. The Subspace Fitting methods are, as all subspace methods, based on the properties of the covariance matrix in (3.10). Two versions of Subspace Fitting have appeared; one based on the signal subspace (SSF) and one based on noise signal subspace (NSF). The Signal Subspace Fitting (SSF) version is based on the observation that the signal eigenvectors are equal to a linear combination of steering vectors. This is essentially the same observation that is made in the ESPRIT method, see (3.20). Here, the criterion is formulated as

$$\{\hat{\boldsymbol{\phi}}, \hat{\mathbf{T}}\} = \arg\min_{\boldsymbol{\phi},\mathbf{T}} \|\hat{\mathbf{E}}_s - \mathbf{A}\mathbf{T}\|_W^2, \tag{3.25}$$

where $\|\mathbf{A}\|_W^2$ denotes $\mathrm{Tr}\{\mathbf{A}\mathbf{W}\mathbf{A}^H\}$ and $\mathbf{W}$ is a positive definite weighting matrix. This function can be concentrated in the same manner as the SML method and the concentrated criterion function becomes

$$\hat{\boldsymbol{\phi}}_{SSF} = \arg\min_{\boldsymbol{\phi}} \mathrm{Tr}\left\{ \mathbf{P}_A^\perp \hat{\mathbf{E}}_s \mathbf{W} \hat{\mathbf{E}}_s^H \right\}. \tag{3.26}$$

The Noise Subspace Fitting (NSF) formulation of the Subspace Fitting criterion is based on the same observation as in MUSIC, namely that the columns

of $\mathbf{A}$ are orthogonal to the noise subspace or $\mathbf{E}_n^H \mathbf{A} = 0$. In this case, an estimate of the DOA is obtained by minimizing the following criterion

$$\hat{\boldsymbol{\phi}}_{NSF} = \arg\min_{\boldsymbol{\phi}} \|\hat{\mathbf{E}}_n^H \mathbf{A}\|_U^2, \tag{3.27}$$

where $\mathbf{U}$ is a $p \times p$ positive definite weighting matrix. Note that the steering matrix enters the criterion function quadratically and thus estimates of parameters that enter linearly in the steering matrix can be found analytically. Different weightings give different asymptotic properties, and by choosing a specific weighting it can be shown that the estimates calculated using (3.26) and (3.27) are asymptotically equivalent to the SML method that is efficient [OVSN93]. Furthermore, if the weighting matrix $\mathbf{U} = \mathbf{I}$, the NSF method reduces to MUSIC. Both of the subspace fitting methods are readily adapted to the presence of a known coupling by exchanging the steering matrix into an effective steering matrix as before.

The weighted subspace methods in (3.26) and (3.27) require a numerical search, and thus the computational load is higher than methods like MUSIC and ESPRIT. One optimization method suitable for these problems is the damped Newton method [DS83]. This method can also be used to compute the estimate when the SML method is used.

If the array is a ULA, the numerical search can be avoided by exploiting the additional structure provided by the ULA. The basic idea is to parameterize the nullspace of $\mathbf{A}^H$ directly instead of parameterizing $\mathbf{A}$, and then use $\mathbf{P}_A^{\perp}$. Several different papers using this parameterization in different contexts have appeared. In [BM86a], the parameterization is applied in combination with the Deterministic Maximum Likelihood (DML) method and in [SS90b] the parameterization is used in the Method Of Direction Estimation (MODE) context (similar to a Root-SSF formulation). The computations involved are then only eigendecompositions and polynomial rootings, in the same manner as in Root-Music, instead of the search otherwise performed. How to include the presence of a known mutual coupling in these methods is not straightforward and is still an open question.

Now, several methods for estimating the DOAs when a known mutual coupling is present have been presented. Most of these represent straightforward extensions of existing algorithms for the coupling free case, but few have been proposed in the literature. The exposition of algorithms also introduced the field of sensor array processing as well as the effects of mutual coupling. However, the survey was of course far from complete and for further algorithms and methods see [Hay95a, He93, KV96].

All of the methods discussed in this chapter rely on a parameterization (parametric methods) except for beamforming, that is a spectral based

method. All of the parametric methods discussed herein assume that the number of sources or signals is known. If the number of sources is not known, it must be estimated along with the DOAs. This will be briefly touched upon in the next section.

3.4 Signal Detection

The parametric methods presented in the previous sections all require the knowledge of the number of sources or number of incident waves. Estimating the number of signals is usually called detection, and here that subject will be briefly discussed. The division into a signal subspace and a noise subspace in Section 3.3.1 requires the knowledge of the dimension of the signal subspace p'. When the signal covariance $\mathbf{P}$ is of full rank, the dimension of the signal subspace is equal to the number of signals, $p' = p$. In that case the problem reduces to testing the multiplicity of the smallest eigenvalue (σ^2). Two methods based on information theory have been proposed in [WK85]. The first method is Akaike's An Information theoretic Criterion (AIC), which essentially is the log-likelihood adjusted with a penalty term for choosing too many signals or too large model order. The second method is Rissanen's and Schwartz's Minimum Description Length (MDL) method, which basically is the same as AIC but using a slightly different penalty term.

However, if the signal covariance $\mathbf{P}$ is not of full rank, i.e. there are coherent signals, the dimension of the signal subspace p' will differ from the number of signals p and the methods in [WK85] will not work. Methods that simultaneously estimate both the number of signals and the DOAs have been proposed instead. The MDL method was extended to simultaneously estimate the number of signals and the DOAs in [WZ89a]. The estimation technique used for the DOA estimation was the DML method and a suitable method of initialization is the Alternating Projection (AP) method in [ZW88]. That method (MDLC) estimates the number of signals even when the signal covariance is rank deficient.

Another way of handling rank deficient signal covariances is the method of Generalized Likelihood Ratio Test (GLRT) [OVSN93, Poo94, Sil70], where the hypothesis that the data are described by the model in (3.8) with p signals is tested against the hypothesis that the covariance matrix of the measurements has an arbitrary structure

$$H_0 : \mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \sigma^2\mathbf{I}$$
$$H_1 : \mathbf{R} \text{ is arbitrary.}$$

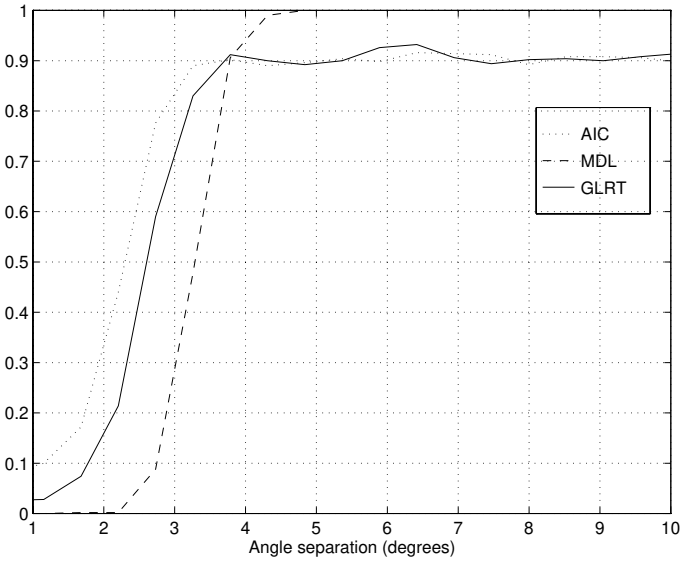


Figure 3.5: The probability of correct detection as a function of the angular separation for the methods GLRT (95%), AIC, and MDL when the signals are non-coherent and the first signal is incident from $\phi_1 = 90°$.

The estimated number of signals is the smallest value, p, of which the hypothesis H_0 not is rejected.

A method that also simultaneously estimates the DOAs and the number of signals is described in [VOK91], where the Weighted Subspace Fitting (WSF) method is used for estimating the DOAs. The detection scheme is based on the observation that a large distance between the signal subspace and the array manifold indicates that there might be coherent sources. This observation is formalized and results in a chi-square test. Furthermore, that method needs knowledge of the dimension of the signal subspace, and the methods described above could be used to estimate this. As in the case of the GLRT, a multi-dimensional search results, and a good method of obtaining initialization values is the method of Alternating Projections [ZW88]. Here, the methods AIC, MDL, MDLC and GLRT are investigated in a few simulations.

In Figure 3.5, the probability of correct detection, when an array of 6 $\lambda/2$ dipoles spaced $\lambda/2$ apart is used and two signals are incident upon the array, is shown. The first signal is incident from $\phi_1 = 90°$ and the other incidence angle is varied from 90° to 80°, forming an angular separation in the interval

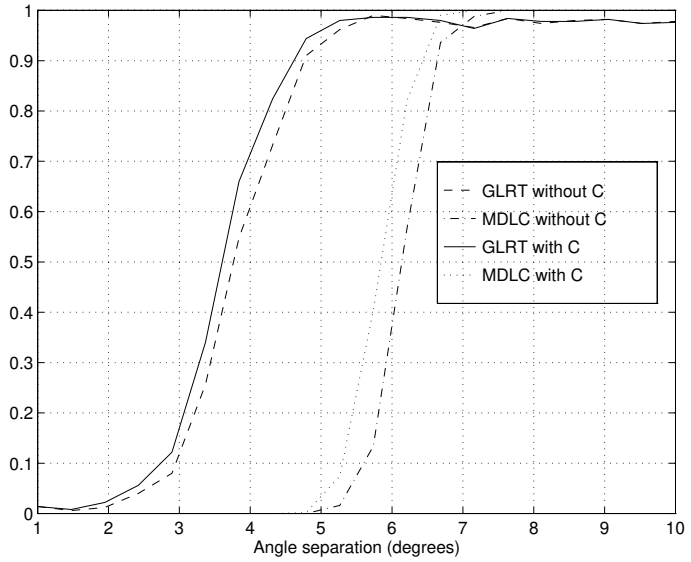


Figure 3.6: The probability of correct detection as a function of the angular separation for the methods GLRT (99%) and MDLC for coherent signals calculated with and without coupling. The first signal is incident from $\phi_1 = 90°$.

$(0°, 10°)$. The signal covariance matrix is

$$\mathbf{P} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \tag{3.28}$$

and the number of samples is $N = 100$. The GLRT method, using 95% confidence level as threshold, works reasonably good already at an angular separation of 3°. Since the signals are non-coherent, the simpler methods AIC and MDL also work. The AIC method has better performance than the MDL method for small angular separations. For larger angular separations the MDL method works much better with almost 100% detection. In [WK85], it was shown that the MDL method is consistent whereas the AIC is not, and the results in this simulation example are in agreement with those results. In the simulations above, the mutual coupling was assumed known and the methods above were extended to include the coupling by simply changing the steering matrix to the effective one as before. However, it is interesting to examine the effects of coupling on the detection performance. In Figure 3.6, the GLRT and the MDLC methods are used for both the case with mutual coupling present and the coupling free case ($\mathbf{C} = \mathbf{I}$). Now, the signals are

correlated and the signal covariance is

$$\mathbf{P} = \begin{bmatrix} 4 & 2.8 \\ 2.8 & 2 \end{bmatrix} \tag{3.29}$$

and the threshold for the GLRT method is calculated based on a confidence level of 99% The GLRT method works better than the Extended version of MDL that uses DML and AP (MDLC) method that was based on the MDL method and the deterministic signal model. Also, it is interesting that both methods work better with mutual coupling present than in the coupling free case. This is in agreement with the observations made when estimating only the directions. However, as in the estimation case these effects are angle dependent and not generally true. In Figure 3.7, the same simulations are performed using $\phi_1 = 50^\circ$ instead of $\phi_1 = 90^\circ$. Here, the effects of coupling are negligible if the coupling is compensated for.

If the coupling is left uncompensated, the GLRT method collapse and detection fails. The MDLC methods appears to be more robust, but that method requires, as found in the examples, a larger angular separation of the sources to work than the GLRT method. This is one important effect of mutual coupling. Since in the case of DOA estimation, uncompensated coupling only introduces some errors and reduces the resolution, but the effects on the detection are more severe since signals are lost and false signals appear. Of course, the simulations presented here are by no means exhaustive and no general conclusions can be drawn. But, as in the case of estimation of the DOAs alone, it has been demonstrated that a compensated known coupling actually can increase the performance in some cases.

3.5 Conclusions

The effects of a known coupling was examined. First, by studying the estimation performance by calculating the CRB for the stochastic signal model. Then, the area of sensor array processing was presented by reviewing the most popular methods of DOA estimation. In most cases, the methods were extended to include the presence of a known coupling. Particularly, the MUSIC, Root-Music, SML, and the Subspace fitting methods were extended to include a known coupling. These methods will be further extended in Chapter 5 to estimate an unknown coupling along with the DOAs.

Furthermore, the problem of signal detection was discussed. If the covariance matrix of the signals is nonsingular, the methods AIC and MDL can be used. Those methods were found to work even in the presence of coupling, known or unknown, without any modifications. However, if the covariance

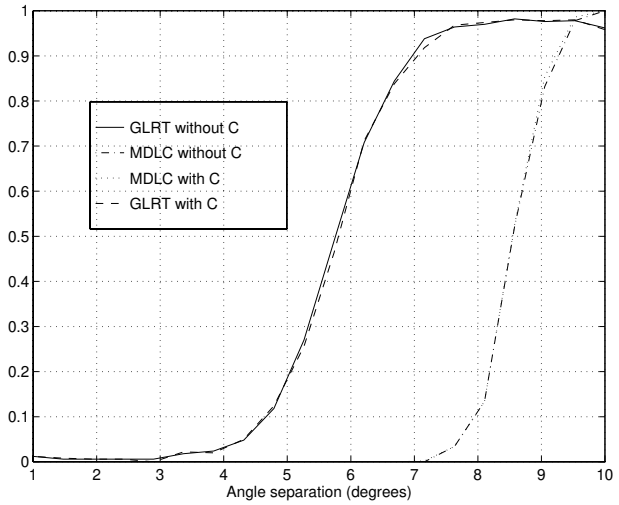


Figure 3.7: The probability of correct detection as a function of the angular separation for the methods GLRT (99%) and MDLC for coherent signals calculated with and without coupling. The first signal is incident from $\phi_1 = 50°$.

matrix was singular, the MDLC and GLRT methods were extended to include the presence of a known coupling. The coupling was found to sometimes increase the detection performance and sometimes decrease. Thus, the effects of mutual coupling on the detection performance was essentially the same as the effects on the estimation performance.

The DOA estimation methods considered in this chapter compensated for a known coupling by introducing the coupling matrix $\mathbf{C}$. However, other effects such as the gain and phase of the individual elements could also be included in the $\mathbf{C}$ matrix. These effects can then also be compensated for by using the algorithms derived in this chapter. In the same manner can the detection schemes, discussed in this chapter, also compensate for other effects by modifying the $\mathbf{C}$ matrix.

Chapter 4

Modeling of an Unknown Coupling

The methods discussed in the previous chapter compensate for the coupling by adjusting the signal processing methods to account for the coupling. However, to do that the coupling has to be known. In some rare occasions the coupling might be analytically obtained, as in the case of a ULA of dipoles in free space. But in most cases, calibration measurements are necessary. In some cases, obtaining the coupling might not be possible due to time aspects, the environment is changing to quickly, or that calibration measurements with known sources is simply not possible in the actual environment. If the coupling is left uncompensated, it will drastically reduce the possibilities of performing direction finding.

One way of mitigating these effects is to estimate the coupling along with the DOAs using signal processing methods. Since this is an introductory study on the effects of mutual coupling and the improvements by estimating the coupling, a simple coupling scenario is considered. The coupling is considered to be unknown and deterministic, and no assumptions regarding the coupling are made. If some knowledge of the coupling exists, it should of course be included in a more realistic scenario, as well as gain and phase uncertainties in the array response. Furthermore, the nominal array response could be known and some random deviation from this response considered. But, as a first study, a simple case is considered in order to get some first results regarding the possible improvements by estimating the coupling as well as the DOAs.

Directly estimating the coupling using the full electromagnetic coupling matrix, derived in Section 2.2, usually requires too many parameters, and therefore a reduced model will be derived in this section. That model will be used in the following sections to compensate for the coupling by estimating it along with the DOAs. Furthermore, the uniqueness of the estimates

obtained using that model is discussed. The section is concluded by deriving a necessary condition for uniqueness, that gives a limit on the number of coupling parameters that are possible to estimate.

4.1 Reduced Coupling Model

The model for the coupling derived in Section 2.2 needs to be simplified in order to be used to estimate the mutual coupling. A direct estimation of the full coupling matrix is usually not feasible, since it involves the estimation of a relatively large number of parameters. The coupling matrix contains n^2 complex entries, and if no constraints on the matrix are imposed, n^2 complex parameters need to be estimated. However, by examining the structure of the coupling matrix $\mathbf{C}$, it is found that the number of parameters can be significantly reduced and a model with few parameters can be found.

First, it is illuminating to study the magnitude of the coupling matrix elements. In Figure 4.1, the logarithm of the magnitude of the elements of the coupling matrix are shown for the case of an array of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart. From this figure it is clear that the magnitude of the coupling decreases quite rapidly with the distance. Thus, only the coupling with the closest elements need to be considered. This corresponds to only parameterizing the elements on the first subdiagonals[1] of the coupling matrix, since the elements far from the main diagonal are relatively small.

Furthermore, the matrix entries along each subdiagonal are almost the same. To see this, the normalized magnitude of the elements on a few subdiagonals are shown in Figure 4.2 for the same array as in Figure 4.1. The magnitude of the coupling is approximately the same along the subdiagonals, and thus it can be modeled by using only one parameter for each subdiagonal resulting in a coupling matrix of Toeplitz structure. A more complex model that better approximates the coupling could be formed by adding additional parameters that model the shape of the magnitude change along the subdiagonals. Here, the objective is to derive a simple model and thus only one parameter for each subdiagonal is used, which also gives a linear model in the coupling parameters.

Also, noting that the coupling from element i to k is the same as the coupling from k to i results in a symmetric (not Hermitian!) Toeplitz matrix.

[1]The i^{th} subdiagonal of a $n \times n$ matrix $\mathbf{C}$ is the $n-i$ length vector with the k^{th} element $\mathbf{C}_{k,k+i}$.

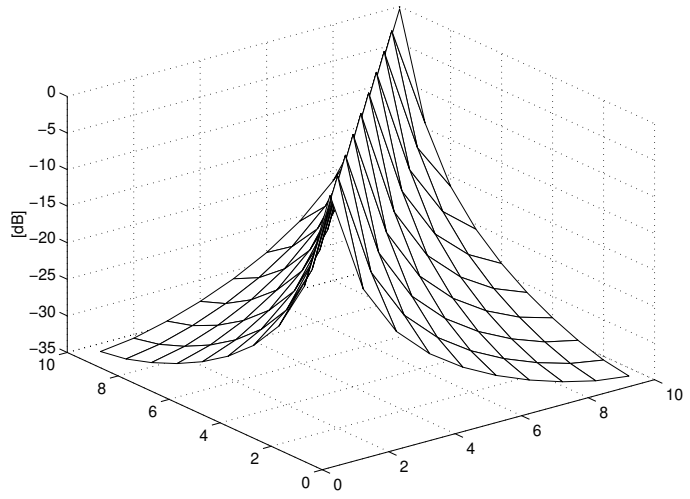


Figure 4.1: The magnitude of the elements of the coupling matrix of an array of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart.

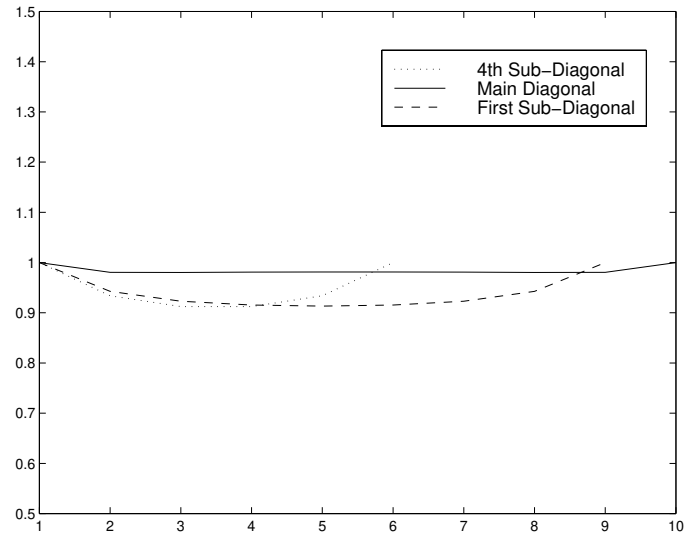


Figure 4.2: The normalized magnitude of the elements on the main diagonal, the first, and 4th subdiagonals of the coupling matrix of an array of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart.

Based on these observations, a reduced coupling model can be formulated as

$$\mathbf{C} = \begin{bmatrix} 1 & c_1 & \cdots & c_q & 0 & \cdots & 0 \\ c_1 & 1 & & & \ddots & & \vdots \\ \vdots & & \ddots & & & \ddots & 0 \\ c_q & & & \ddots & & & c_q \\ 0 & \ddots & & & \ddots & & \vdots \\ \vdots & & \ddots & & & \ddots & c_1 \\ 0 & \cdots & 0 & c_q & \cdots & c_1 & 1 \end{bmatrix}, \tag{4.1}$$

where $c_i \in \mathbb{C}$ and $0 < |c_{i+1}| < |c_i| < 1$. The coupling parameters that are estimated are then

$$\mathbf{c}^T = \begin{bmatrix} c_1 & c_2 & \cdots & c_q \end{bmatrix}. \tag{4.2}$$

Note that the number of parameters to be estimated has decreased from n^2, when estimating all the elements in the $\mathbf{C}$ matrix, to $q < n$. The reduced model of the coupling matrix then becomes a banded symmetric Toeplitz matrix. How many coupling parameters to be include depends on the element separation distance since the coupling is reduced with increased separation distance, see Section 2.4.3. It also depends on how many elements there are and the length of the dipole. Also, note that the dipole experiences a strong mutual coupling and other types of antenna elements often experiences less coupling and thus also less parameters are needed. The coupling model in (4.1) was also suggested in [FW91], however, without any detailed physical motivation. Here, the mutual coupling for a dipole array was derived and it was found that the above model relatively well describes the effects of mutual coupling. Thus, a detailed physical motivation for using the above coupling model has been presented.

This reduced model is used to estimate the coupling in the following chapters. Before estimation based on the reduced coupling model is considered, some conditions for obtaining unique estimates of the mutual coupling and the DOAs will be discussed.

4.2 Uniqueness

A good model should of course give unique estimates of directions and coupling parameters. This problem is decoupled from the estimation problem, and thus the noise can be ignored in the following analysis. The problem of uniqueness has been analyzed for the coupling free case in [WZ89b], where two sufficient conditions for uniqueness are derived. One condition

for uniqueness for every batch of signals is derived using subspace concepts, and another condition for almost every batch of signals is derived based on results from topology. The uniqueness problem has also been analyzed in [BM86b, NSS91, Swi92, Wax92]. These results seem difficult to extend to the case when an unknown coupling is present, since in that case the effective steering matrix is $\mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\phi})$, and thus the unknowns are multiplied with each other.

The problem of uniqueness is the question of under which conditions the unknown parameters of the model can be determined uniquely. The obtained data are the measured voltages, and the following model is used

$$\mathbf{x}(t) = \mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\lambda})\mathbf{s}(t) + \mathbf{n}(t), \tag{4.3}$$

where $\boldsymbol{\lambda} = e^{jkd\cos\boldsymbol{\phi}}$. Now, the received vector $\mathbf{x}(t)$ is sampled at N time instants and stacked in a matrix $\mathbf{X}$ and the following model results

$$\mathbf{X} = \mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\lambda})\mathbf{S} + \mathbf{N}, \tag{4.4}$$

where $\mathbf{X}$ and $\mathbf{N}$ are $n \times N$ matrices and $\mathbf{S}$ is a $p \times N$ matrix. The problem of uniqueness is based on the following noiseless relation

$$\mathbf{X} = \mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\lambda})\mathbf{S} = \mathbf{C}(\mathbf{c}')\mathbf{A}(\boldsymbol{\lambda}')\mathbf{S}' \quad \Rightarrow \quad \{\mathbf{c}, \boldsymbol{\lambda}, \mathbf{S}\} = \{\mathbf{c}', \boldsymbol{\lambda}', \mathbf{S}'\}. \tag{4.5}$$

If equation (4.5) holds for every set of measured voltages $\mathbf{X}$, only one set of parameters gives $\mathbf{X}$, thus the solution $(\mathbf{c},\ \boldsymbol{\lambda},\ \mathbf{S})$ is unique. This is a rather difficult problem already without coupling, and with an unknown coupling it seems difficult to obtain a general result. One step towards the general result is to consider the parameterization of the array. If

$$\mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\lambda}) = \mathbf{C}(\mathbf{c}')\mathbf{A}(\boldsymbol{\lambda}') \quad \Rightarrow \quad \{\mathbf{c}, \boldsymbol{\lambda}\} = \{\mathbf{c}', \boldsymbol{\lambda}'\} \tag{4.6}$$

holds, the array is said to be unambiguously parameterized. However, if the signal is known and of full row rank, it can be removed from (4.5) by multiplying with the pseudo-inverse of $\mathbf{S}$. Therefore, an unambiguously parameterized array implies uniqueness if the signal is known and of full row rank. A known signal could correspond to a training sequence in communication applications, and in that case the solution is unique if (4.6) holds. To examine the uniqueness and ambiguity problem for the parameterization of the reduced coupling model, the condition in (4.6) is rewritten as

$$f(\boldsymbol{\psi}, \mathbf{y}) = \mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\lambda}) - \mathbf{C}(\mathbf{c}')\mathbf{A}(\boldsymbol{\lambda}') = \mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\lambda}) - \mathbf{Y} = 0, \tag{4.7}$$

where $\boldsymbol{\psi}^T = [\mathbf{c}^T\ \boldsymbol{\lambda}^T]$ and $\mathbf{Y} \in \mathbb{C}^{n\times p}$. Of course, one solution is $\mathbf{c} = \mathbf{c}'$ and $\boldsymbol{\lambda} = \boldsymbol{\lambda}'$. If this is the only solution, the array is unambiguously parameterized.

Also, note that if the equation in (4.7) holds for one source or one column of the steering matrix $\mathbf{A}$, it holds for all columns of $\mathbf{A}$. Thus the derivation is performed for only one source. The steering matrix $\mathbf{A}(\boldsymbol{\lambda})$ then reduces to a vector $\mathbf{a}(\lambda)$ and $\mathbf{Y}$ reduces to the vector $\mathbf{y}$. Furthermore, if there are q coupling parameters and only one column is considered, then $\boldsymbol{\psi} \in \mathbb{C}^{q+1}$.

Now, (4.7) represents a system of equations in $\boldsymbol{\psi}$, and the implicit function theorem [Rud76] can be applied to obtain conditions for the existence of a unique solution of the system of equations. However, the parameter vector, $\boldsymbol{\psi}$, is in general complex and therefore an analytic form of the implicit function theorem is used [Ber77].

Theorem 1 The Implicit Function Theorem
Let $\mathbf{f}$ *be continuously differentiable in an open set* $E \subset \mathbb{C}^{n+q+1}$*, with values in* $\mathbb{C}^n$*. Suppose* $(\mathbf{a}, \mathbf{b}) \in E$, $\mathbf{f}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$*, and*

$$\det \mathbf{J}(\mathbf{a}, \mathbf{b}) \neq 0, \tag{4.8}$$

where $\mathbf{J}_{ij} = \frac{\partial \mathbf{f}_i}{\partial \psi_j}$ *is the Jacobian.*

Then there exists a neighborhood B *of* $\mathbf{b}$ *(*$B \subset \mathbb{C}^n$*) and a unique function* $\mathbf{g}$ *which is continuously differentiable in* B*, with values in* $\mathbb{C}^{q+1}$*, such that* $\mathbf{g}(\mathbf{b}) = \mathbf{a}$*, and*

$$\mathbf{f}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \mathbf{0} \qquad (\mathbf{y} \in B). \tag{4.9}$$

Actually, if $q = n - 1$, the analytic implicit function theorem reduces to the analytic inverse function theorem.

Before the question of general uniqueness of the array parameterization is studied, two special cases will be considered. First, the uniqueness when estimating the DOAs in the presence of a known coupling is examined. Secondly, the calibration case where the DOAs are known and the coupling parameters are estimated is studied.

When the coupling is known, the DOAs in (4.7) are simply obtained by multiplication of the inverse of the coupling matrix.

$$\mathbf{A}(\boldsymbol{\lambda}) = \mathbf{C}^{-1}\mathbf{y} \tag{4.10}$$

Remember that the coupling matrix is assumed to be of full rank in all the cases considered here. Furthermore, in this case the results of [BM86b, NSS91, Swi92, WZ89b, Wax92] can then be applied since the introduction of a known full rank coupling matrix can easily be eliminated by multiplication by the inverse. Thus, the uniqueness properties do not change when a known coupling is introduced.

Next, consider the case when the DOAs are known and the coupling parameters are estimated. This corresponds to a calibration scenario using

known sources at known locations. The system of equations in (4.7) now becomes

$$f(\boldsymbol{\psi}, \mathbf{y}) = \mathbf{C}(\mathbf{c})\mathbf{a}(\lambda) - \mathbf{y} = 0, \tag{4.11}$$

where only one source is considered as mentioned previously. However, since the unknown coupling parameters enter the expressions linearly, the equations can be rewritten as

$$f(\boldsymbol{\psi}, \mathbf{y}) = \mathbf{M}(\lambda)\mathbf{c} - \mathbf{y}' = 0. \tag{4.12}$$

If the Jacobian is non-zero or the $n \times n$ matrix $\mathbf{M}$ is non-singular, then according to Theorem 1 the solution is unique. Thus, the problem reduces to examine the rank of the matrix $\mathbf{M}$. Before a general formula for the $\mathbf{M}$ matrix is presented, first consider a special case. For an array of 7 elements, the relation $\mathbf{C}(\mathbf{c})\mathbf{a}(\lambda)$ can be written as

$$\begin{bmatrix} 1 & c_1 & c_2 & c_3 & c_4 & c_5 & c_6 \\ c_1 & 1 & c_1 & c_2 & c_3 & c_4 & c_5 \\ c_2 & c_1 & 1 & c_1 & c_2 & c_3 & c_4 \\ c_3 & c_2 & c_1 & 1 & c_1 & c_2 & c_3 \\ c_4 & c_3 & c_2 & c_1 & 1 & c_1 & c_2 \\ c_5 & c_4 & c_3 & c_2 & c_1 & 1 & c_1 \\ c_6 & c_5 & c_4 & c_3 & c_2 & c_1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \\ \lambda^3 \\ \lambda^4 \\ \lambda^5 \\ \lambda^6 \end{bmatrix}. \tag{4.13}$$

The equations in (4.13) are linear in the coupling parameters and can be rewritten as

$$\begin{bmatrix} 1 & \lambda & \lambda^2 & \lambda^3 & \lambda^4 & \lambda^5 & \lambda^6 \\ \lambda & 1+\lambda^2 & \lambda^3 & \lambda^4 & \lambda^5 & \lambda^6 & 0 \\ \lambda^2 & \lambda+\lambda^3 & 1+\lambda^4 & \lambda^5 & \lambda^6 & 0 & 0 \\ \lambda^3 & \lambda^2+\lambda^4 & \lambda+\lambda^5 & 1+\lambda^6 & 0 & 0 & 0 \\ \lambda^4 & \lambda^3+\lambda^5 & \lambda^2+\lambda^6 & \lambda & 1 & 0 & 0 \\ \lambda^5 & \lambda^4+\lambda^6 & \lambda^3 & \lambda^2 & \lambda & 1 & 0 \\ \lambda^6 & \lambda^5 & \lambda^4 & \lambda^3 & \lambda^2 & \lambda & 1 \end{bmatrix} \begin{bmatrix} 1 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix}. \tag{4.14}$$

The $\mathbf{M}$ matrix is obtained by omitting the first column of the matrix in (4.14). If $\mathbf{M}$ is of full rank the coupling parameters are obtained by simply multiplying with the left-inverse of $\mathbf{M}$. To examine the rank of the $\mathbf{M}$ matrix, a few steps of Gaussian elimination are performed. First, subtract from each

row, except the first one, λ times the preceding row

$$\text{rk } \mathbf{M} = \text{rk} \begin{bmatrix} \lambda & \lambda^2 & \lambda^3 & \lambda^4 & \lambda^5 & \lambda^6 \\ 1 & 0 & 0 & 0 & 0 & -\lambda^7 \\ 0 & 1 & 0 & 0 & -\lambda^7 & 0 \\ 0 & 0 & 1 & -\lambda^7 & 0 & 0 \\ 0 & 0 & -\lambda^7 & 1 & 0 & 0 \\ 0 & -\lambda^7 & 0 & 0 & 1 & 0 \\ -\lambda^7 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.15}$$

Thus if $\lambda^7 \neq \pm 1$, the columns of $\mathbf{M}$ are independent and $\mathbf{M}$ is of full column rank and the left-inverse exists. For the DOAs, this corresponds to

$$\lambda^7 \neq \pm 1 \Rightarrow e^{j14kd\cos\phi} \neq e^{j2\pi} \Rightarrow \cos\phi \neq \frac{l\pi}{7kd} \quad l = 0, \ldots, 6. \tag{4.16}$$

Therefore, if the maximum number of coupling parameters are estimated, the solution is unique for all angles except at 7 specific angles where the matrix becomes rank deficient. However, if 3 coupling parameters (c_1, c_2, c_3) are estimated it works for all angles, since the first three columns in (4.15) are independent.

By studying the structure in (4.15) and (4.14) a general formula for $\mathbf{M}$ can be found. The rows of the upper half become

$$\mathbf{M}_{i,:} = [\lambda^i, \lambda^{i-1} + \lambda^{i+1}, \ldots, 1 + \lambda^{2(i-1)}, \lambda^{2(i-1)+1}, \ldots, \lambda^{n-1}, 0, \ldots] \tag{4.17}$$

and rows of the lower half are given by

$$\mathbf{M}_{i,:} = [\lambda^i, \lambda^{i-1} + \lambda^{i+1}, \ldots, \lambda^{2i+1-n} + \lambda^{n-1}, \lambda^{2i-n}, \ldots, 1, 0, \ldots]. \tag{4.18}$$

If the number of sensors is odd, the middle row becomes

$$\mathbf{M}_{i,:} = [\lambda^i, \lambda^{i-1} + \lambda^{i+1}, \ldots, 1 + \lambda^{2(i-1)}, 0, \ldots] \tag{4.19}$$

as was the case in the above example with 7 elements, see (4.14). The rank of $\mathbf{M}$ is examined by performing a few row-reductions just as in the above example. Using the expressions for the rows of $\mathbf{M}$ in (4.17)-(4.19), it is found that the structure will be similar to the one in the above example. The row-reduced matrix has ones on the main diagonal and $-\lambda^n$ on the anti-diagonal after row-reduction, and thus the $\mathbf{M}$ is of full rank if $\lambda^n \neq \pm 1$ or

$$\cos\phi \neq \frac{l\pi}{nkd} \quad l = 0, \ldots, n-1. \tag{4.20}$$

Therefore, if the DOA of a single source is known, $n-1$ coupling parameters defined like in (4.1) can be determined uniquely if the DOA does not

coincide with the n locations defined by (4.20). However, if the number of coupling parameters are reduced, the solution can be guaranteed to be unique. Considering only the left half of $\mathbf{M}$, those columns are guaranteed to be independent, see (4.15). Note that the number of estimated coupling parameters equals the number of columns of $\mathbf{M}$. Therefore, if only the left half of $\mathbf{M}$ is included, which corresponds to using[2]

$$q = q_0 = \lceil n/2 \rceil - 1 \tag{4.21}$$

columns, q_0 coupling parameters can be determined uniquely.

Therefore, in a calibration scenario, all $n-1$ coupling parameters can be uniquely determined if the signal and the DOA are known and not equal to the ambiguity angles of (4.20). To avoid the DOA ambiguity, the number of estimated parameters can be reduced to q_0 in order to obtain guaranteed uniqueness. However, at least for small arrays, it seems rather unlikely that the ambiguity angles can not be avoided in a calibration scenario. For larger arrays and more coupling parameters, the number of ambiguity angles increases. Since the problem experiences poor conditioning close to the ambiguity angles, the full number of estimated coupling parameters should not be estimated for larger arrays.

Now, the special cases of known DOAs or known coupling have been discussed. The general case when solving for both DOA and coupling can be analyzed by calculating the Jacobian of the system of equations in (4.7) and invoking Theorem 1. Here, only one unknown DOA is considered, since if the solution is unique for one column it will be unique for all columns, as discussed previously. Start once again by considering a special case with 5 elements and 4 coupling parameters. In that case the Jacobian becomes

$$\begin{bmatrix} \lambda & \lambda^2 & \lambda^3 & \lambda^4 & c_1 + 2c_2\lambda + 3c_3\lambda^2 + 4c_4\lambda^3 \\ 1+\lambda^2 & \lambda^3 & \lambda^4 & 0 & 1 + 2c_1\lambda + 3c_2\lambda^2 + 4c_3\lambda^3 \\ \lambda+\lambda^3 & 1+\lambda^4 & 0 & 0 & c_1 + 2\lambda + 3c_1\lambda^2 + 4c_2\lambda^3 \\ \lambda^2+\lambda^4 & \lambda & 1 & 0 & c_2 + 2c_1\lambda + 3\lambda^2 + 4c_1\lambda^3 \\ \lambda^3 & \lambda^2 & \lambda & 1 & c_3 + 2c_2\lambda + 3c_1\lambda^2 + 4\lambda^3 \end{bmatrix}. \tag{4.22}$$

Here, the left part of the Jacobian ($n-1$ left columns) is recognized as the matrix that appeared when assuming the DOA to be known and solving for the coupling parameters. That matrix was possible to simplify using row-operations and conditions for full rank was derived. The last column corresponds to the estimation of the DOA and for the solution of both the coupling parameters and the DOA to be unique, the Jacobian must be of

[2]The notation $\lceil x \rceil$ denotes rounding off x to the nearest integer towards infinity.

full rank according to Theorem 1. Therefore, by performing the same row-operations as previously, the Jacobian can be written as

$$\begin{bmatrix} \lambda & \lambda^2 & \lambda^3 & \lambda^4 & c_1 + 2c_2\lambda + 3c_3\lambda^2 + 4c_4\lambda^3 \\ 1 & 0 & 0 & -\lambda^5 & 1 + c_1\lambda + c_2\lambda^2 + c_3\lambda^3 - 4c_4\lambda^4 \\ 0 & 1 & -\lambda^5 & 0 & c_1 + \lambda + c_1\lambda^2 + c_2\lambda^3 - 4c_3\lambda^4 \\ 0 & -\lambda^5 & 1 & 0 & c_2 + c_1\lambda + \lambda^2 + c_1\lambda^3 - 4c_2\lambda^4 \\ -\lambda^5 & 0 & 0 & 1 & c_3 + c_2\lambda + c_1\lambda^2 + \lambda^3 - 4c_1\lambda^4 \end{bmatrix} \tag{4.23}$$

To show that the determinant of the matrix in (4.23) is non-zero, i.e. full rank, seems hard. An even harder problem is to generalize to n sensors and q coupling parameters. That is an open problem which demands more work, but for one simple special case the solution can be shown to be unique and thus the parameterization of the array in that case unambiguous. Consider the same 5 element array, but let only one coupling parameter be used. In that case, it suffices to show that the first column (corresponding to the coupling parameter) and the last column (corresponding to the DOA) are independent

$$\begin{bmatrix} \lambda & c_1 \\ 1 & 1 + c_1\lambda \\ 0 & c_1 + \lambda + c_1\lambda^2 \\ 0 & \lambda(c_1 + \lambda + c_1\lambda^2) \\ -\lambda^5 & c_1\lambda^2 + \lambda^3 - 4c_1\lambda^4 \end{bmatrix}. \tag{4.24}$$

Subtracting λ times the second row from the first and changing order gives

$$\begin{bmatrix} 1 & 1 + c_1\lambda \\ 0 & c_1 - \lambda - c_1\lambda^2 \\ 0 & c_1 + \lambda + c_1\lambda^2 \\ 0 & \lambda(c_1 + \lambda + c_1\lambda^2) \\ -\lambda^5 & c_1\lambda^2 + \lambda^3 - 4c_1\lambda^4 \end{bmatrix}. \tag{4.25}$$

Furthermore, adding the third row to the second shows that the columns are independent. Adding more sensors only results in taller columns but the first rows remain unchanged. Therefore, the parameterization using one coupling parameter is locally unambiguous for arrays containing 5 elements or more. This analysis started with a five element array and more work is needed to prove that a smaller number of elements also results in an unambiguous parameterization using one coupling parameter. Also, extending to more coupling parameters requires further work. However, the more elements that are included the taller the vectors become and it becomes more likely that the set of columns in question become independent.

The preceding analysis examined if the parameterization of the array was unique. To prove that the parameterization of the general case in equation (4.5) with unknown signals, angles, and coupling is unique seems difficult. Another interesting question is how many coupling parameters that can be uniquely estimated using an n element array when p signals are incident from unknown directions. In that case the analysis becomes much more complicated. Therefore, a necessary condition is derived next that gives a bound on how many coupling parameters that at best can be estimated.

A necessary condition for uniqueness, which is valid for all types of arrays (not only ULAs), can also be derived using the analytical implicit function theorem. According to the discussion in Section 3.3.1 all information is contained in the second order moments, and since the uniqueness problem is decoupled from the estimation problem, the noise can be eliminated. All information is then contained in the signal eigenvectors $\mathbf{E}_s$. Assume that the signal matrix $\mathbf{S}$ is of rank p', then the uniqueness relation in (4.5) can be written as

$$\mathbf{E}_s = \mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\phi})\mathbf{T}, \tag{4.26}$$

where the signals are accounted for through the $p \times p'$ matrix $\mathbf{T}$ of rank p'. A necessary condition for obtaining a unique solution of $(\mathbf{c}, \boldsymbol{\phi}, \mathbf{T})$ is of course to have at least an equal number of unknowns and independent equations, see for instance [Ber77, Rud76]. The number of real equations are $2np'$ since there are n sensors and p' independent complex signals. The number of real unknowns are $2q + p + 2pp'$ corresponding to q complex coupling parameters, p real DOAs, and pp' complex elements in $\mathbf{T}$. The necessary condition for a unique solution is that the number of equations is larger than the number of unknowns

$$2np' \geq 2q + p + 2pp'. \tag{4.27}$$

A useful formulation of the above relation is to express the number of elements needed to estimate p signals with a total rank of p' as

$$n \geq p + \frac{p}{2p'} + \frac{q}{p'}. \tag{4.28}$$

The estimation of the coupling parameters results in an increase of the number of antenna elements needed to obtain unique estimates. How large this cost becomes is easily seen in an example. Assume that the signal is full rank, i.e. $p' = p$, and study three different cases: 1) No coupling parameters are estimated, 2) The reduced coupling model with $n-1$ parameters, and 3) The reduced coupling model with only 2 complex parameters. A calculation of the number of elements n needed to estimate p signals based on (4.28) is shown in Figure 4.3. The cost of introducing the coupling parameters

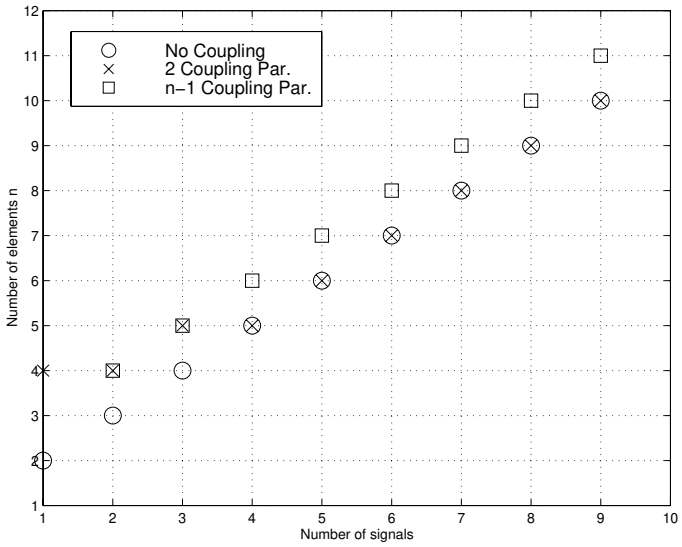


Figure 4.3: The number of elements necessary for the existence of a unique solution when p signals are incident upon the array.

is rather small. Usually only one additional element is needed to obtain unique estimates. Note that estimating a single DOA and the full number of coupling parameters $(n-1)$ is not possible if the estimates are to be unique. It is important to remember that the figures presented here only represent a necessary condition for the solution to be unique. Furthermore, the result in (4.28) reduces to the result in [BM86b, WZ89b] when no coupling is present, but it differs slightly from the results in [FW91], which also includes coupling and uncertainties in gain and phase of the elements.

4.3 Conclusions

If it is not possible to obtain the coupling matrix through calibration, one way of reducing the effects of an unknown coupling is to estimate the coupling along with the DOAs. However, in that case a simple model with relatively few parameters is needed since it must be estimated using the same data. In this chapter, a reduced coupling model with less parameters than the full coupling model derived using electromagnetics, was introduced. This model, a banded symmetric Toeplitz matrix, coincided with the model conjectured in [FW91], but here it was based on physical reasoning. Note that although the reduced model captures the properties of the full electromagnetical model of the coupling, it of course deteriorates the direction finding. Therefore, a

calibration scheme where the full coupling matrix is estimated should be applied if possible. But if not possible, estimating the coupling reduces the effects of an unknown coupling.

When estimating the coupling along with the DOAs new problems arise. One problem is the question of uniqueness of the estimates of the parameters of the reduced coupling model. This was analyzed in this chapter and although no general results regarding the uniqueness were obtained, it was shown that the parameterization of the array, in a calibration scenario, is locally unambiguous. If the number of coupling parameters are equal to or less than $\lceil n/2 \rceil - 1$, the parameterization is locally unambiguous for all DOAs. Otherwise the parameterization becomes ambiguous at n different angles in that calibration scenario. For the case of both unknown DOA and coupling, no general results were obtained, but it was found that arrays with more than four elements are unambiguously parameterized (locally) if one coupling parameter is used.

Furthermore, a necessary condition for uniqueness was presented. In the next chapter, this reduced model is used to estimate both the coupling and the DOAs using different methods.

Chapter 5

Estimation With an Unknown Coupling

If it is not possible to obtain the coupling matrix through calibration or some other means, the effects of an unknown coupling can be mitigated by estimating both the coupling and the DOAs. However, the number of unknown parameters generally becomes too large if a direct estimation of the elements of the coupling matrix is performed together with a DOA estimation. Using the reduced coupling model, derived in the previous chapter, the number of parameters becomes feasible. Here, the reduced coupling model will be used throughout this chapter in the simple scenario of a completely unknown coupling. In practice, some knowledge of the coupling usually exists, which of course should be used.

First, the model and the stochastic assumptions will be reviewed. Then, the Stochastic Maximum Likelihood (SML) estimator for the present data model, is derived. The Cramér-Rao lower Bound on the variance for the case of an unknown coupling is derived based on the SML expressions.

After establishing the lower bound, several methods of estimation are discussed. Estimation of an unknown mutual coupling has received very little interest in the signal processing literature. An iterative version of the MUSIC method, that estimates both coupling and DOAs, was developed in [FW91]. Therefore, that iterative MUSIC algorithm is discussed first and some related methods are proposed. Furthermore, the NSF method is extended to estimate both an unknown coupling and the DOAs. The properties of the different estimation methods are then examined and compared to the CRB in a few computer experiments. The chapter is concluded by briefly considering the problem of signal detection in the presence of an unknown coupling.

5.1 Data Model

The input data to the estimation algorithms in the following sections are the measured voltages of the antenna elements. As before, the antenna is a uniform linear array of n dipoles and p waves are incident upon the array. One important difference to the case of a known coupling is that the reduced coupling model is used here. Sampling the voltages $\mathbf{x}(t)$ at N time instants $t = t_1, \ldots, t_N$ (taking N "snapshots") gives the following model

$$\mathbf{X} = \mathbf{C}\mathbf{A}(\boldsymbol{\phi})\mathbf{S} + \mathbf{N}, \tag{5.1}$$

where

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} \mathbf{x}(t_1) & \cdots & \mathbf{x}(t_N) \end{bmatrix} && (5.2) \\ \mathbf{S} &= \begin{bmatrix} \mathbf{s}(t_1) & \cdots \mathbf{s}(t_N) \end{bmatrix} && (5.3) \\ \mathbf{N} &= \begin{bmatrix} \mathbf{n}(t_1) & \cdots & \mathbf{n}(t_N) \end{bmatrix}. && (5.4) \end{aligned}$$

The matrix of measured voltages $\mathbf{X}$ is $n \times N$, the signal matrix $\mathbf{S}$ is $p \times N$, the noise matrix $\mathbf{N}$ is $n \times N$, and the coupling matrix $\mathbf{C}$ is $n \times n$ with elements $\mathbf{C}_{ij} = c_{|i-j|}$. Note that $c_0 = 1$ and $|c_i| < 1 \; \forall \; 1 \leq i \leq q$. The steering matrix $\mathbf{A}(\boldsymbol{\phi})$ is $n \times p$, and for a ULA it has a Vandermonde structure with elements $[\mathbf{A}(\boldsymbol{\phi})]_{vw} = e^{-jkd(v-1)\cos\phi_w}$, where k is the wavenumber and d is the element separation distance. The DOAs are contained in the parameter vector $\boldsymbol{\phi}$. Furthermore, the coupling matrix is assumed to be of full rank and both the signal $\mathbf{s}(t)$ and the noise $\mathbf{n}(t)$ are circularly Gaussian distributed as in Section 3.1. Also, the number of dipoles are chosen as to fulfill the necessary condition derived in previous section

$$n \geq p + \frac{p}{2p'} + \frac{q}{p'}, \tag{5.5}$$

where p' is the rank of the signal matrix $\mathbf{S}$ and $q \leq n - 1$ is the number of coupling parameters used.

Based on the data model in (5.1) and the statistical assumptions above, the maximum likelihood estimator for an unknown coupling and unknown DOAs will be derived in the next section.

5.2 Maximum Likelihood

The method of Maximum Likelihood estimation is one of the most popular estimation method used in signal processing. One of the reasons is that

for large enough data records the ML estimator is approximately the Minimum Variance Unbiased (MVU) estimator, and thus it is asymptotically efficient. Here, the maximum likelihood estimator for the data model presented in the previous section will be derived. Since not only the noise, but also the signals, are assumed to be stochastic, the resulting maximum likelihood method is usually called the Stochastic Maximum Likelihood (SML) method, see Section 3.3.3. For the coupling-free case the SML has been derived in a concentrated form [Böh86, Jaf88, SN95]. The derivation of the SML, when an unknown coupling is present, follows basically the same lines as the coupling-free case. However, some properties do change and of course the number of unknowns increase, and therefore the important steps of the derivation are included in some detail in Appendix 5A.

The SML estimation method is based on maximizing the likelihood by choosing the values of the unknown parameters that give the maximum. The unknown parameters are the DOAs $\boldsymbol{\phi}$, the coupling $\mathbf{c}$, and the free parameters of the signal and noise covariances. Note that both the coupling parameters c_i and the signal covariance $\mathbf{P}$ are in general complex. When the objective is to estimate the DOAs, the signal and noise covariances are nuisance parameters and the expression that is to be maximized can be concentrated into the following form

$$\hat{\boldsymbol{\Theta}} = \{\hat{\boldsymbol{\phi}}, \hat{\mathbf{c}}\} = \arg\min_{\phi,c} V_{SML}(\boldsymbol{\phi}, \mathbf{c}) = \arg\min_{\phi,c} \log\det\left[\mathbf{C}\mathbf{A}\hat{\mathbf{P}}\mathbf{A}^H\mathbf{C}^H + \hat{\sigma}^2\mathbf{I}\right], \tag{5.6}$$

where

$$\hat{\mathbf{P}} = (\mathbf{C}\mathbf{A})^{\dagger}\left[\hat{\mathbf{R}} - \hat{\sigma}\mathbf{I}\right](\mathbf{C}\mathbf{A})^{\dagger H} \tag{5.7}$$

$$\hat{\sigma}^2 = \frac{1}{n-p}\mathrm{Tr}\left\{\mathbf{P}^{\perp}_{CA}\hat{\mathbf{R}}\right\}. \tag{5.8}$$

Note that the dependence of $\mathbf{C}(\mathbf{c})$, $\mathbf{A}(\boldsymbol{\phi})$, $\hat{\mathbf{P}}(\boldsymbol{\phi}, \mathbf{c})$, and $\hat{\sigma}^2(\boldsymbol{\phi}, \mathbf{c})$ on the parameters have been omitted in order to reduce the notation complexity. For a more detailed discussion of deriving the likelihood function, see Appendix 5A. To obtain a closed form solution for the coupling or the DOAs is, in general, not possible. Therefore, the parameter values are obtained through a numerical search. Many different optimization techniques have been proposed in the signal processing literature. Here, a damped Newton method [DS83] will be used. The estimate is calculated iteratively as

$$\boldsymbol{\Theta}^{k+1} = \boldsymbol{\Theta}^k - \mu_k\mathbf{H}^{-1}\mathbf{V}', \tag{5.9}$$

where $\boldsymbol{\Theta}^k$ is the estimate at iteration k, μ_k is the step length, $\mathbf{H}$ denotes the Hessian matrix of the criterion function, and $\mathbf{V}'$ denotes the gradient. The

Hessian and the gradient are evaluated at $\mathbf{\Theta}^k$. The step length is chosen using a simple scheme, starting with a value $\mu \leq 1$ and then decreasing the value until the value of the criterion function is less than in previous iteration. If the criterion not decreases, the iteration is terminated. Of course there is no guarantee that the search will end in the global minimum. However, if the search is initiated "sufficiently close", the global minimum will be found.

When implementing the above search method, two problems appear. First, for $-\mathbf{H}^{-1}\mathbf{V}'$ to be a descent direction the Hessian must be positive definite, which not necessarily have to be the case far from the true optimum. Also, the evaluation of the Hessian is computationally cumbersome, and therefore the Hessian is replaced by the Fisher Information Matrix (FIM), which is more easily evaluated and also guaranteed to be positive definite if the estimation problem is well posed. This leads to a method that in the statistical literature is known as the scoring method [OVSN93]. Furthermore, it will be shown in the next section that the asymptotic Hessian essentially coincides with the asymptotic covariance matrix of the estimates. Since, the ML method is statistically efficient, an expression for the FIM and the Cramér-Rao lower bound are thus also found.

In order to implement the search, the gradient and the asymptotic Hessian of the criterion function in (5.6) must be calculated. The steps involved in this calculation are similar to those in [SN90b], but the introduction of the coupling matrix results in slight changes. The calculations are included in Appendix 5A, where it also is shown that the gradient becomes

$$\mathbf{V}'(\mathbf{\Theta}) = \begin{bmatrix} 2\mathrm{ReTr}\left[\hat{\mathbf{G}}\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{CA}\frac{\partial(\mathbf{CA})}{\partial\Theta_1}\right] \\ \vdots \\ 2\mathrm{ReTr}\left[\hat{\mathbf{G}}\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{CA}\frac{\partial(\mathbf{CA})}{\partial\Theta_{p+2q}}\right] \end{bmatrix}, \tag{5.10}$$

where

$$\hat{\mathbf{G}} = (\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{C}\mathbf{A})^{-1} - \frac{1}{\hat{\sigma}^2}(\mathbf{A}^H\mathbf{C}^H\mathbf{C}\mathbf{A})^{-1}, \tag{5.11}$$

and $\hat{\sigma}^2$ is defined in (5.8). The estimated covariance matrix, $\hat{\mathbf{R}}$, is defined in (3.15). The expression for the asymptotic Hessian is also derived in Appendix 5A

$$\mathbf{H} = \mathbf{V}''_a(\mathbf{\Theta}) = \begin{bmatrix} \mathbf{H}_{\phi\phi} & \mathbf{H}_{\phi c} \\ \mathbf{H}^T_{\phi c} & \mathbf{H}_{cc} \end{bmatrix}. \tag{5.12}$$

The block matrices are as follows

$$\mathbf{H}_{\phi\phi} = \frac{2}{\sigma^2}\mathrm{Re}\left\{\left(\mathbf{P}\mathbf{A}^H\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}\mathbf{A}\mathbf{P}\right)^T \odot \left(\mathbf{D}^H\mathbf{C}^H\mathbf{P}^{\perp}_{CA}\mathbf{C}\mathbf{D}\right)\right\}, \tag{5.13}$$

where

$$\mathbf{D} = \left[\frac{\partial \mathbf{a}(\phi)}{\partial \phi} \Big|_{\phi=\phi_1}, \cdots, \frac{\partial \mathbf{a}(\phi)}{\partial \phi} \Big|_{\phi=\phi_p} \right]. \tag{5.14}$$

Also,

$$\mathbf{H}_{\phi c} = \begin{bmatrix} \frac{\partial^2 \mathbf{V}_a}{\partial \phi \partial c_r} & \frac{\partial^2 \mathbf{V}_a}{\partial \phi \partial c_i} \end{bmatrix} = \mathbf{H}_{\phi c}^T, \tag{5.15}$$

where

$$\left[\frac{\partial^2 \mathbf{V}_a}{\partial \phi \partial c_r} \right]_{ij} = \tfrac{2}{\sigma^2} \mathrm{ReTr} \left[\mathbf{P} \mathbf{A}^H \mathbf{C}^H \mathbf{R}^{-1} \mathbf{C} \mathbf{A} \mathbf{P} \mathbf{A}^H \boldsymbol{\Upsilon}_j \mathbf{P}_{CA}^{\perp} \mathbf{C} \tfrac{\partial \mathbf{A}}{\partial \phi_i} \right] \tag{5.16}$$

$$\left[\frac{\partial^2 \mathbf{V}_a}{\partial \phi \partial c_i} \right]_{ij} = -\tfrac{2}{\sigma^2} \mathrm{ImTr} \left[\mathbf{P} \mathbf{A}^H \mathbf{C}^H \mathbf{R}^{-1} \mathbf{C} \mathbf{A} \mathbf{P} \mathbf{A}^H \boldsymbol{\Upsilon}_j \mathbf{P}_{CA}^{\perp} \mathbf{C} \tfrac{\partial \mathbf{A}}{\partial \phi_i} \right] \tag{5.17}$$

and $\boldsymbol{\Upsilon}_k$ is banded symmetric Toeplitz matrix with zeros everywhere except on the k^{th} and the $-k^{th}$ subdiagonal which consists of ones. Finally,

$$\mathbf{H}_{cc} = \begin{bmatrix} \mathbf{H}_{c_r c_r} & \mathbf{H}_{c_r c_i} \\ \mathbf{H}_{c_r c_i}^T & \mathbf{H}_{c_i c_i} \end{bmatrix}, \tag{5.18}$$

where

$$[\mathbf{H}_{c_r c_r}]_{ij} = \tfrac{2}{\sigma^2} \mathrm{ReTr} \left[\mathbf{P} \mathbf{A}^H \mathbf{C}^H \mathbf{R}^{-1} \mathbf{C} \mathbf{A} \mathbf{P} \mathbf{A}^H \boldsymbol{\Upsilon}_j \mathbf{P}_{CA}^{\perp} \boldsymbol{\Upsilon}_i \mathbf{A} \right] \tag{5.19}$$

$$[\mathbf{H}_{c_r c_i}]_{ij} = -\tfrac{2}{\sigma^2} \mathrm{ImTr} \left[\mathbf{P} \mathbf{A}^H \mathbf{C}^H \mathbf{R}^{-1} \mathbf{C} \mathbf{A} \mathbf{P} \mathbf{A}^H \boldsymbol{\Upsilon}_j \mathbf{P}_{CA}^{\perp} \boldsymbol{\Upsilon}_i \mathbf{A} \right] \tag{5.20}$$

$$[\mathbf{H}_{c_i c_i}]_{ij} = -\tfrac{2}{\sigma^2} \mathrm{ReTr} \left[\mathbf{P} \mathbf{A}^H \mathbf{C}^H \mathbf{R}^{-1} \mathbf{C} \mathbf{A} \mathbf{P} \mathbf{A}^H \boldsymbol{\Upsilon}_j \mathbf{P}_{CA}^{\perp} \boldsymbol{\Upsilon}_i \mathbf{A} \right]. \tag{5.21}$$

Now, using the scoring method, i.e. equation (5.9) with the gradient and Hessian in (5.10) and (5.12), the implementation of the SML method is straightforward.

The relation between the SML method and the Cramér-Rao bound will be used in the next section to derive a formula for the CRB.

5.3 Direction Finding Accuracy

When comparing different estimation schemes, it is important to known the fundamental performance limit of the estimation problem at hand. The Cramér-Rao lower bound gives the lowest possible variance of a unbiased estimator and thus serves as a fundamental performance limit that is useful when comparing different estimation schemes. In Section 3.2, the CRB was calculated for the case of a known mutual coupling. Here, the bound will be calculated when an unknown mutual coupling is present.

The Cramér-Rao lower bound for a parameter vector $\mathbf{\Theta}$ is defined as

$$E\{(\hat{\mathbf{\Theta}} - \mathbf{\Theta}_0)(\hat{\mathbf{\Theta}} - \mathbf{\Theta}_0)^T\} \geq \mathbf{I}^{-1}(\mathbf{\Theta}), \tag{5.22}$$

where $\mathbf{I}(\mathbf{\Theta})$ is the FIM with elements

$$\mathbf{I}(\mathbf{\Theta})_{ij} = -\mathrm{E}\left[\frac{\partial^2 \log p(\mathbf{x};\mathbf{\Theta})}{\partial\Theta_i \partial\Theta_j}\right]. \tag{5.23}$$

Note that $p(\mathbf{x};\mathbf{\Theta})$ is the probability density function for the measurements parameterized by the unknown parameters $\mathbf{\Theta}$ that are to be estimated. When viewed as a function of the unknown parameters $\mathbf{\Theta}$, $p(\mathbf{x};\mathbf{\Theta})$ is termed the likelihood function.

The CRB is usually evaluated using (5.22)-(5.23) and for the coupling free case, formulas have appeared in [Ban71, Sch81]. However, when estimating the DOAs, the signal and noise covariances can often be regarded as nuisance parameters and thus the CRB for these parameters are not important. A more compact expression for the CRB on the DOAs is possible to obtain by observing the fact that the likelihood function can be concentrated regarding the signal and the noise covariance as derived in the previous section. Compact expressions for the CRB for the coupling free case have been derived in [OWVK89, SN90b].

The CRB for the case when an unknown mutual coupling is present has been derived in the literature [FW91], where also the phase and gain of the antenna elements are considered unknown. In that derivation the definition in (5.22) was used, resulting in a large Fisher matrix to be inverted. Here, compact expressions for the CRB on the DOAs in the presence of an unknown mutual coupling will be derived based on the concentrated likelihood function.

An expression for the CRB on the DOAs can be derived by exploiting the relation between the ML method and the CRB. From general statistical theory it follows that if there is a unique solution, the ML estimator converges in probability to the true parameter vector under very general conditions [KS61]. Furthermore, the ML estimator is found to be asymptotically efficient [KS61], i.e. the covariance matrix for the ML estimates coincides with the CRB for large number of samples ($N \gg 1$). Once the asymptotic expression for the covariance matrix for the ML estimates is obtained, the general CRB formula is also obtained. The number of samples enters only as a scaling, since the likelihood expression is proportional to the number of samples, and therefore is the CRB expression valid for all N [SN90b].

An expression for the asymptotic covariance matrix for the SML estimates is calculated in Appendix 5B and the general CRB with an unknown coupling

becomes

$$E\{(\hat{\boldsymbol{\Theta}}-\boldsymbol{\Theta}_0)(\hat{\boldsymbol{\Theta}}-\boldsymbol{\Theta}_0)^T\} \geq \mathbf{B}_{STO_c} = \frac{1}{N}[\mathbf{V}_a''(\boldsymbol{\Theta})]^{-1}, \tag{5.24}$$

where the asymptotic Hessian is defined in (5.12).

However, in most cases the variance of the DOA estimates is important and it is of interest to obtain an explicit expression for the these. In (5.12) the asymptotic Hessian is written in a block matrix form and a partitioned inverse can be obtained as

$$\mathbf{B}_{STO_c}^{\phi\phi} = \frac{1}{N}\mathbf{H}_{\phi\phi}^{-1} + \frac{1}{N}\mathbf{H}_{\phi\phi}^{-1}\mathbf{H}_{\phi c}\left(\mathbf{H}_{cc} - \mathbf{H}_{\phi c}^T\mathbf{H}_{\phi\phi}^{-1}\mathbf{H}_{\phi c}\right)^{-1}\mathbf{H}_{\phi c}^T\mathbf{H}_{\phi\phi}^{-1}. \tag{5.25}$$

Note that since the Hessian is evaluated at the true parameters it is positive definite and the inverse exists, if the problem is well posed. The first term in (5.25) is the CRB in the case of a known coupling. The second term is positive definite, since the second term is a quadratic form of the Schur complement which is positive definite if the Hessian is [SS89]. The CRB, when an unknown coupling is present, is therefore always larger than when the coupling is known

$$\mathbf{B}_{\text{unknown}}^{\phi\phi} > \mathbf{B}_{\text{known}}^{\phi\phi}. \tag{5.26}$$

The above result is very intuitive and of course expected. It is interesting to examine how much larger the bound becomes when the DOAs are estimated together with an unknown coupling compared to the case when the coupling is known. Also, it is interesting to compare with the hypothetical coupling free case.

First note that in this chapter, an unknown coupling is estimated using the data model in (5.1), where the reduced coupling model is used. That model is also used when calculating the CRB. If the data are generated using a different model, a modeling error will occur and in that case the formula for the CRB should be modified to include a biased estimator, see [SS89]. However, in this case the analysis of the model errors seems complicated. Thus, the different estimation schemes are examined using data generated from the reduced coupling model and no model error is present. The calculated CRB for a known coupling in this section will therefore differ from the CRB for a known coupling calculated in Section 3.2 since different coupling matrices are used. The difference will however be rather small, since the reduced coupling model relatively well describes the coupling. The coupling parameter values are here chosen from the first column of the coupling matrix calculated using electromagnetics. Usually, only a few coupling parameters are included when estimating and that also give rise to a small difference. In the following, the increase in the square root of the CRB (the standard deviation) when the

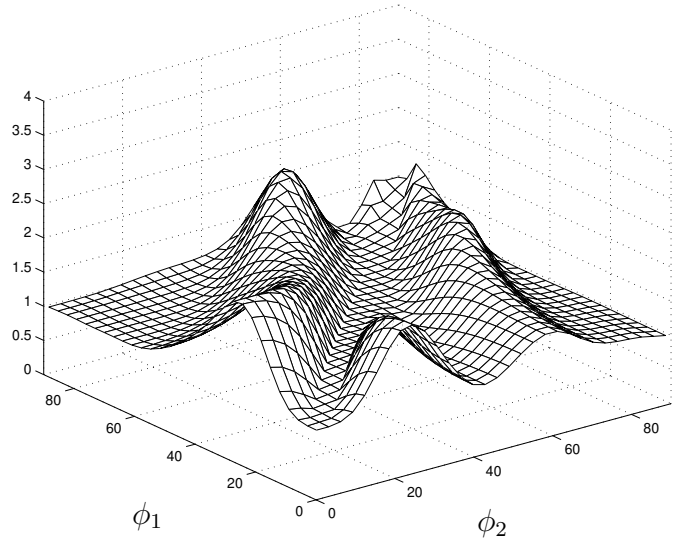


Figure 5.1: The ratio of the standard deviation for unknown coupling and known coupling for ϕ_2 versus the angles for an array of 5 $\lambda/2$ dipoles spaced $\lambda/2$ apart and $q = 1$.

coupling needs to be estimated (using the reduced model), is compared to when the coupling is known. Although these comparisons of the CRB for different cases gives some insight on how well it is possible to estimate the DOAs, it is only meaningful for reasonably large SNRs. For high SNR scenarios the model error will dominate in a realistic scenario. Here, however, the data are generated using the same model as the estimators are based on and, thus the estimators considered here will be asymptotically unbiased.

In Figure 5.1, the ratio of the standard deviation for an unknown coupling and a known coupling for ϕ_2 versus the angles is shown. An array of 5 $\lambda/2$ dipoles spaced $\lambda/2$ apart is used, and only one coupling parameter is estimated ($q = 1$). The ratio is larger than unity, but the increase is not dramatic. Also, the ratio is above unity everywhere due to the fact that the CRB for an unknown coupling is always larger than the CRB for a known coupling as found in (5.26). It is also important to note that the main diagonal corresponds to two waves arriving from the same angle and thus impossible to discriminate. The Hessian becomes rank deficient and the CRB becomes infinite. Therefore, it is hard to draw any conclusions regarding the valley that appear along the main diagonal. A similar figure for ϕ_1 also exists, but is not shown since it is only a mirrored version.

Increasing the number of coupling parameters that are estimated also

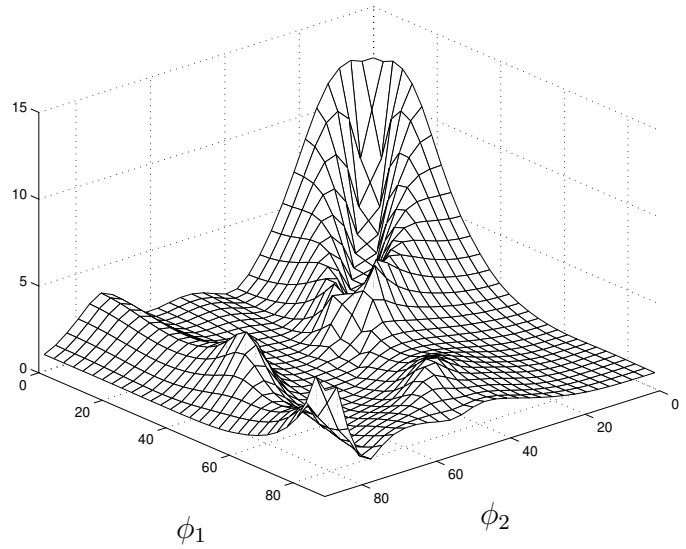


Figure 5.2: The ratio of the standard deviation for unknown coupling and known coupling for ϕ_2 versus the angles for an array of 5 $\lambda/2$ dipoles spaced $\lambda/2$ apart and $q = 2$.

increases the CRB. In Figure 5.2, the same array as in Figure 5.1 is used but the number of estimated coupling parameters, q, is two. Here, the increase of the standard deviation when the coupling needs to be estimated is larger than when only one parameter was estimated. For angles close to end-fire, where it is hard also in the coupling free case to estimate the DOAs, the increase is especially large.

Estimating more coupling parameters further increases the CRB, and the numerical evaluation becomes more critical. When a full coupling matrix is used, i.e. four coupling parameters for a 5 element array, then the smooth surface in the previous figures disappears, indicating that the problem is close to singular. Estimating a large number of coupling parameters, although below the necessary condition derived in Section 4.2, results in a large variance and often the gain in a smaller bias is not worth the large variance of the estimates.

Increasing the number of elements allows for a larger number of coupling parameters to be estimated. In Figure 5.3, two coupling parameters are estimated using a 10 element array. Comparing with Figure 5.2 where also two coupling parameters are estimated but with a five element array it is found that the CRB ratio is much lower for the larger array. Thus, more coupling parameters can be estimated using a larger array, and a better

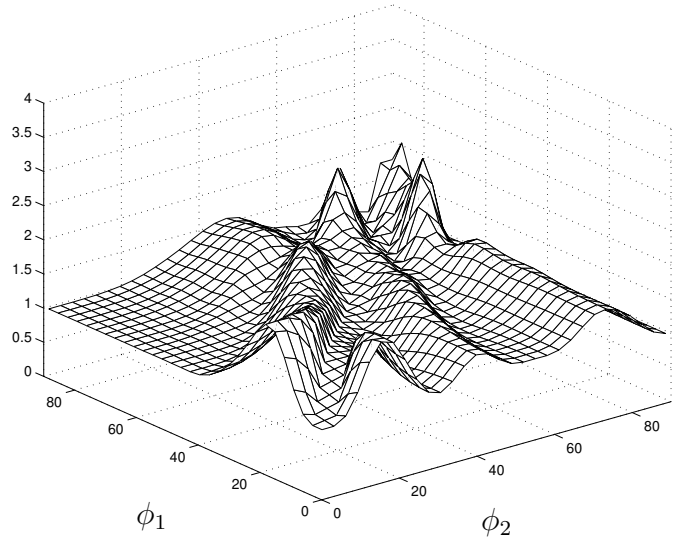


Figure 5.3: The ratio of the standard deviation for unknown coupling and known coupling for ϕ_2 versus the angles for an array of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart and $q = 2$.

model fit can therefore be obtained. The same array, as in in Figure 5.3, but with almost a full coupling matrix is shown in Figure 5.4. Six out of nine parameters are used, and since the coupling decreases rapidly (see Figure 3.2), this many parameters are usually not needed.

All the CRB ratios presented in the above figures were calculated as the ratio of the CRB for an unknown coupling to a known coupling, since that clearly illustrates the increase in variance when the coupling is unknown. However, the ratio of the CRB for an unknown coupling to the hypothetical coupling free case is essentially the same. In Figure 5.5, the ratio for a known coupling to the hypothetical coupling free case is shown for the 10 element array with six coupling parameters used in Figure 5.4. The ratio is close to unity for most angles, and thus using the CRB for a known coupling or the CRB for the hypothetical coupling free case to calculate the ratio essentially gives the same result.

Now, a lower bound on the variance (CRB) has been obtained that can be used to address the quality of different estimation schemes. In the following sections a few methods of estimating both an unknown coupling and DOAs are discussed. Then, some simulations are performed where these methods are compared to each other and to the CRB that was derived in this section.

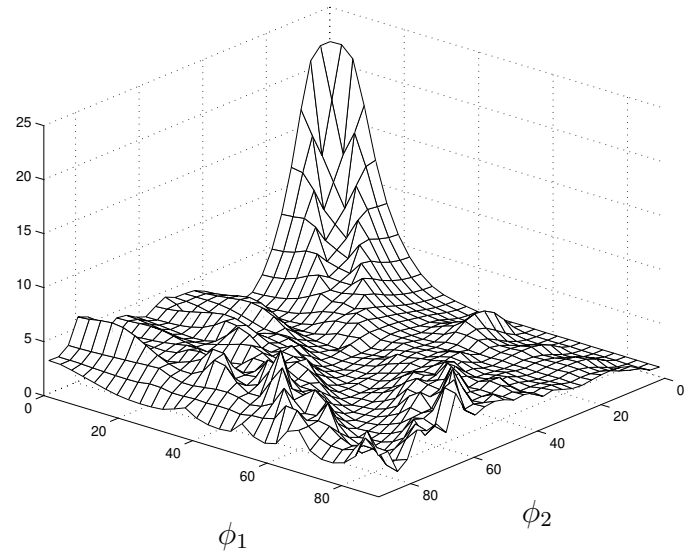


Figure 5.4: The ratio of the standard deviation for unknown coupling and known coupling for ϕ_2 versus the angles for an array of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart and $q = 6$.

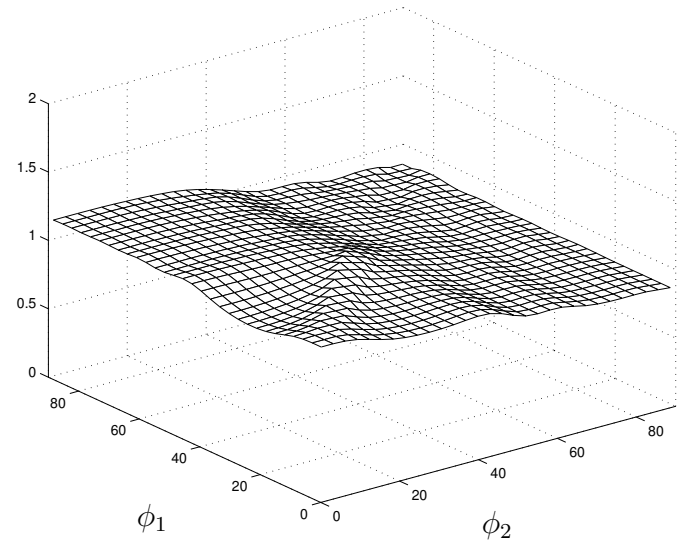


Figure 5.5: The ratio of the standard deviation for a known coupling and without coupling for ϕ_2 versus the angles for an array of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart and $q = 6$.

5.4 Iterative MUSIC

The SML method requires a multidimensional search, which is computationally demanding already when only DOAs are estimated. The search becomes even more demanding when q coupling parameters are added. Also, the initialization of the search becomes more important the more parameters that are estimated, and the possibilities of ending up in a local minimum increases. Therefore, it is of great interest to eliminate the full parameter search, or at least reduce the number of parameters that are needed to search over. In [FW91], an iterative version of the MUSIC algorithm that accounts for an unknown coupling as well as gain and phase uncertainties is presented. The only search performed is the DOA search, i.e. search for peaks in the MUSIC spectrum just as in the coupling free case. The price that has to be paid for this reduction in computational complexity is that the estimation scheme must be performed several times, i.e. iterated. Unfortunately, it was found in [PK91] that the iterative MUSIC method of [FW91] gives nonunique estimates and in [PF94] that is experiences slow convergence. The main reason for this is the fact that both an mutual coupling and unknown gain and phase are estimated in [FW91], leading to an ambiguous array parameterization. However, for the case of only an unknown mutual coupling, the analysis in Chapter 4 is valid and non-uniqueness can be avoided.

Using the notation in the previous sections the criterion that is minimized becomes [FW91]

$$J_c = \|\hat{\mathbf{E}}_n^H \mathbf{C}(\mathbf{c})\mathbf{\Gamma}\mathbf{A}(\boldsymbol{\phi})\|_F^2 = \sum_{k=1}^{p} \|\hat{\mathbf{E}}_n^H \mathbf{C}(\mathbf{c})\mathbf{\Gamma}\mathbf{a}(\phi_k)\|_F^2, \tag{5.27}$$

where F denotes the Frobenius norm, $\hat{\mathbf{E}}_n$ the noise eigenvectors (see (3.10)), and $\mathbf{\Gamma}$ contains the gain and phase. Here, only the mutual coupling will be studied and thus $\mathbf{\Gamma} = \mathbf{I}$ will be assumed in the following. Furthermore, in [FW91] the coupling model is the same as the reduced coupling model derived in previous sections. In [FW91], however, no detailed physical motivation was presented. A similar model for a circular array was also conjectured in [FW91].

Now, when estimating the DOAs and the unknown mutual coupling the estimation scheme becomes an iterative two step procedure.

Step 1: Estimating the DOAs

Search for the p highest peaks in the MUSIC spectrum

$$P_{MU}(\phi) = \|\hat{\mathbf{E}}_n^H \mathbf{C}(\mathbf{c})\mathbf{a}(\phi)\|^{-2}. \tag{5.28}$$

The DOAs estimates are then taken as the angles where these peaks appear. Note that this is the same step as in the MUSIC algorithm for a known

coupling. In the first iteration, the identity matrix is used as coupling matrix.
Step 2: Estimating the Mutual Coupling
Insert the estimated DOAs from previous step in the criterion function J_c. The following lemma [FW91] will then prove useful.

Lemma 1 *For any $M \times 1$ complex vector $\mathbf{x}$ and any $M \times M$ banded complex symmetric Toeplitz matrix $\mathbf{B}$*

$$\mathbf{Bx} = \mathbf{Q}(\mathbf{x}) \cdot \mathbf{b} \tag{5.29}$$

where the $L \times 1$ vector $\mathbf{b}$ is given by

$$\mathbf{b}_i = \mathbf{B}_{1i}, \qquad i = 1, 2, \cdots, L \tag{5.30}$$

and L is the highest subdiagonal that is different from zero. The $M \times L$ matrix $\mathbf{Q}(\mathbf{x})$ is given by the sum of the following matrices

$$\begin{aligned} [\mathbf{W}_1]_{kl} &= \begin{cases} \mathbf{x}_{k+l-1}, & k+l \leq M+1 \\ 0, & \textit{otherwise} \end{cases} \\ [\mathbf{W}_2]_{kl} &= \begin{cases} \mathbf{x}_{k-l-1}, & k \geq l \geq 2 \\ 0, & \textit{otherwise} \end{cases} \end{aligned} \tag{5.31}$$

Using Lemma 1 the criterion function can be written as

$$J_c = \bar{\mathbf{c}}^H \left\{ \sum_{k=1}^{p} \mathbf{Q}^H(k) \hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Q}(k) \right\} \bar{\mathbf{c}}, \tag{5.32}$$

where

$$\mathbf{Q}(k) = \mathbf{Q}(\mathbf{a}(\phi_k)) \tag{5.33}$$

$$\bar{\mathbf{c}} = [1 \quad \mathbf{c}^T]^T \tag{5.34}$$

$$\mathbf{c}_k = \mathbf{C}_{1k}, \qquad k = 1, 2, \cdots, L \tag{5.35}$$

and the vector $\mathbf{c}$ contains the unknown coupling parameters. Now, (5.32) represents a quadratic minimization problem under a linear constraint, since the coupling matrix $\mathbf{C}$ is assumed to be unity on the main diagonal. Solving the constrained minimization problem using Lagrangian multipliers

$$\hat{\bar{\mathbf{c}}} = \mathbf{G}^{-1}\mathbf{1}_1(\mathbf{1}_1^T\mathbf{G}^{-1}\mathbf{1}_1)^{-1}, \tag{5.36}$$

where

$$\mathbf{G} = \sum_{k=1}^{p} \mathbf{Q}^H(k) \hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Q}(k). \tag{5.37}$$

Note that the vector $\mathbf{1}_1$ has a one in the first position and the rest is zero.

The iterative MUSIC algorithm thus estimates the DOAs in *Step 1* and estimates the coupling in *Step 2*, and the final estimates are obtained by iterating between these two steps. At each iteration the value of the criterion function decreases, and thus convergence is guaranteed [FW91] which is very important. Another important property of an iterative algorithm is how many iterations that are usually needed. The convergence rate of the iterative MUSIC method and the iterative methods introduced in the next section will be examined in some simulation examples in Section 5.7

5.5 Other Iterative Methods

The idea to iteratively estimate the DOAs and the coupling parameters is very intuitive, and can of course be applied to other DOA estimation algorithms. Many methods for estimating the DOAs when a known coupling is present can be formulated iteratively when an unknown coupling is present. The main reason for formulating an iterative procedure is to reduce the computational cost by avoiding an extensive numerical search, and thus also avoiding initialization and convergence problems. A closed form solution for the coupling parameter should therefore exist in order to avoid a numerical search for the coupling parameters. First, a straightforward extension of the Root-Music to the case of unknown coupling will be discussed. The Root-Music method can the formulated as two steps:

Step 1: Estimating the DOAs

When the array is a ULA, the search in the MUSIC method for the DOAs can be avoided by viewing the MUSIC function as a polynomial in $z = e^{jkd\cos\phi}$, exactly like the known coupling case, see Section 3.3.3. The DOA estimates are then obtained by picking the p roots of the polynomial (3.18)

$$f(z) = \mathbf{a}^T(z^{-1})\mathbf{C}^H\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\mathbf{C}\mathbf{a}(z). \tag{5.38}$$

closest to the unit circle, $\hat{z}_i$, and calculate the angle and solve for ϕ_i in $\angle\hat{z}_i = \angle e^{-jkd\cos\phi_i}$.

Step 2: Estimating the Mutual Coupling

Return to the MUSIC criterion and insert the DOA estimates into the criterion

$$J_c = \|\hat{\mathbf{E}}_n^H\mathbf{C}(\mathbf{c})\mathbf{A}(\hat{\boldsymbol{\phi}})\|_F^2, \tag{5.39}$$

where the estimates of the coupling parameters $\hat{\mathbf{c}}$ are taken as the values of $\mathbf{c}$ that minimize the criterion function. The same formulas result as in the

iterative MUSIC method in the previous section

$$\hat{\tilde{\mathbf{c}}} = \mathbf{G}^{-1}\mathbf{1}_1(\mathbf{1}_1^T\mathbf{G}^{-1}\mathbf{1}_1)^{-1}, \tag{5.40}$$

where

$$\mathbf{G} = \sum_{k=1}^{p} \mathbf{Q}^H(k)\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\mathbf{Q}(k). \tag{5.41}$$

However, at the true parameters the noise eigenvectors are orthogonal to the steering matrix or $\mathbf{E}_n \perp \mathbf{A}_e$, and the criterion in (5.39) becomes zero. Since the criterion also can be written as a quadratic form of $\mathbf{G}$, this means that $\mathbf{G}$ becomes singular. Close to the minimum, $\mathbf{G}$ therefore becomes ill conditioned, and the expression for coupling estimates that includes $\mathbf{G}^{-1}$ needs to be modified. By applying the formula of a partitioned inverse, [Lüt96], the coupling estimates can be written as

$$\hat{\mathbf{c}} = -\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H, \tag{5.42}$$

where $\mathbf{G}$ is partitioned as

$$\mathbf{G} = \begin{bmatrix} G_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{12}^H & \mathbf{G}_{22} \end{bmatrix}. \tag{5.43}$$

Note that G_{11} is a scalar and that the matrix $\mathbf{G}$ is Hermitian, i.e. $\mathbf{G} = \mathbf{G}^H$. This expression for the coupling estimates is better conditioned and can be used close to the minimum, i.e. close to the true parameters. This modification of the coupling estimate is of course suitable to make also in the iterative MUSIC method described in the previous section.

The iterative Root-Music algorithm estimates the DOAs in *Step 1* and estimates the coupling in *Step 2*, and the final estimates are obtained by iterating between these two steps just as in the iterative MUSIC method. The drawback of this method is that convergence is not guaranteed, since the roots of the polynomial (3.18) are not exactly the minima of the criterion function in (5.39). The minima for the criterion function lie by definition on the unit circle, but the roots of polynomial lie close to the unit circle, but not exactly on the unit circle. Usually this difference is small and the method works well, but due to this fact convergence can not be guaranteed.

Another method that would be interesting to obtain an iterative form of is the ESPRIT method. However, it appears difficult to obtain an iterative method with guaranteed convergence without resorting to a numerical search. Nevertheless, it is straightforward to extend the ESPRIT algorithm to a two step iterative version for the case of an unknown coupling.

Step 1: Estimating the DOAs
The DOAs are obtained as those ϕ:s that best fit the relation

$$\mathbf{J}\mathbf{C}^{-1}\mathbf{E}_s = \mathbf{A}_s(\boldsymbol{\phi})\mathbf{T}. \tag{5.44}$$

The total least squares solution to this fitting problem can be put in a general subspace fitting context as [VO91]

$$\min_{\boldsymbol{\phi},\mathbf{T}} \|\mathbf{J}\mathbf{C}^{-1}\hat{\mathbf{E}}_s - \mathbf{A}_s\mathbf{T}\|_F^2, \tag{5.45}$$

where $\mathbf{A}_s$ is the steering matrix with the shift structure defined in (3.19). Solving this minimization problem gives an estimate of the DOAs.
Step 2: Estimating the Mutual Coupling
It seems difficult to obtain an estimate of the coupling parameters from the above criterion without resorting to a numerical search. In that case it is probably better to use a multi-dimensional estimation scheme as SML from the start. However, a closed form solution for the coupling parameters is possible to obtain by changing of the criterion that is to be minimized to

$$\min_{\mathbf{c}} \|\mathbf{J}\hat{\mathbf{E}}_s - \mathbf{J}\mathbf{C}(\mathbf{c})\mathbf{A}\mathbf{T}\|_F^2, \tag{5.46}$$

where the values of $\mathbf{A}$ and $\mathbf{T}$ are calculated using the DOA estimates from *Step 1*. That, of course, eliminates any possibilities of guaranteed convergence, but gives a computationally cheap solution with reasonably good properties if the criteria are sufficiently similar.

The iterative ESPRIT algorithm thus estimates the DOAs in *Step 1* and estimates the coupling in *Step 2*, and the final estimates are obtained by iterating between these two steps.

In the iterative versions of Root-Music and ESPRIT, a large numerical search was avoided by jumping between two different criteria, that had analytical solutions, at the expense of giving up any convergence guarantees. The main motivation for this was that the computational demands of these methods were kept small. Similar iterative formulations of multi-dimensional search methods like WSF and MODE seem to be of little interest, since a search for the DOAs still has to be performed. A better solution is probably to include the coupling parameters in the search instead, even though this increases the computational complexity. However, for one multi-dimensional method, it is possible to concentrate the criterion function regarding the coupling parameters in the same manner as the SML method was concentrated with respect to the noise and signal covariances. That method will be discussed in the next section.

5.6 Noise Subspace Fitting

Most DOA estimation algorithms are possible to adopt to an unknown coupling, but solving for the unknown parameters usually requires a multi-dimensional search. One method, the Noise Subspace Fitting method (see Section 3.3.3), is however possible to concentrate in the same manner as the SML method. Therefore, only a search for the DOAs are needed just as in the coupling free case and the introduction of the unknown coupling parameters only marginally increases the computational complexity. The NSF criterion function can be formulated as

$$\hat{\mathbf{\Theta}} = \left\{ \hat{\boldsymbol{\phi}}, \hat{\mathbf{c}} \right\} = \arg\min_{\Theta} V_{NSF}(\mathbf{\Theta}) = \arg\min_{\Theta} \|\hat{\mathbf{E}}_n^H \mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\phi})\|_U^2, \tag{5.47}$$

where $\mathbf{U}$ is a $p \times p$ positive definite weighting matrix. That function can be concentrated regarding the coupling parameters, see Appendix 5C

$$V_{NSF}(\boldsymbol{\phi}) = G_{11} - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H, \tag{5.48}$$

where $\mathbf{G}$ is partitioned as in (5.43) and calculated as

$$\mathbf{G} = \mathbf{M}^H \left[\hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \otimes \mathbf{A}^* \mathbf{U}^* \mathbf{A}^T \right] \mathbf{M}, \tag{5.49}$$

where $\mathbf{M}$ is a $n^2 \times n$ selection matrix

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \\ \vdots \\ \mathbf{M}_n \end{bmatrix} \quad \text{where} \quad \mathbf{M}_k = \begin{cases} \mathbf{I} & k = 1 \\ \begin{bmatrix} \tilde{\mathbf{I}}\ \mathbf{I}_{2:k,:} \\ \mathbf{I}_{1:n-k+1,:} \end{bmatrix} & k > 1 \end{cases} \tag{5.50}$$

and $\mathbf{I}_{k:m,:}$ denotes the k^{th} to m^{th} rows of the unity matrix and $\tilde{\mathbf{I}}$ is the reverse permutation operator defined as

$$\tilde{\mathbf{I}} = \begin{bmatrix} 0 & & & 1 \\ & & \iddots & \\ & 1 & & \\ \iddots & & & \\ 1 & & & 0 \end{bmatrix}. \tag{5.51}$$

Alternatively, the matrix $\mathbf{G}$ can be written as

$$\mathbf{G} = \sum_{k=1}^{p} \mathbf{Q}^H(\mathbf{A}\mathbf{U}_{k,:}^H)\hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Q}(\mathbf{A}_{:,k}), \tag{5.52}$$

where the function $\mathbf{Q}(\mathbf{x})$ is defined in Lemma 1. The criterion is concentrated with respect to the q coupling parameters, and only a p dimensional search for the DOAs are needed just as in the coupling free case. The DOA parameters enter the expression non-linearly and a damped Newton method is used, similar to the SML method in Section 5.2. In the SML method large simplifications were possible by using an asymptotic Hessian in the search process leading to the scoring method. Here, the true Hessian is used in a damped Newton method. The gradient and Hessian, needed in the implementation of the search method, are derived in Appendix 5C.

The weighting $\mathbf{U}$ in (5.47) can be chosen as to minimize the variance of the DOA estimates. The derivation in the coupling free case is valid also in this case since $\mathbf{A}_e(\boldsymbol{\Theta}) = \mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\phi})$ can be seen as an effective steering matrix with parameters $\mathbf{c}$ and $\boldsymbol{\phi}$. Choose the weighting as [OVSN93]

$$\mathbf{U}_{opt} = \mathbf{A}_e^{\dagger}(\boldsymbol{\Theta}_0)\hat{\mathbf{E}}_s\widetilde{\boldsymbol{\Lambda}}^2{\boldsymbol{\Lambda}_s}^{-1}\hat{\mathbf{E}}_s^H\mathbf{A}_e^{\dagger H}(\boldsymbol{\Theta}_0), \tag{5.53}$$

where $\boldsymbol{\Theta}_0$ are the true parameters, $\widetilde{\boldsymbol{\Lambda}}$ is defined in (3.12), and $\boldsymbol{\Lambda}_s$ is defined in (3.10). Also, note that the replacement of $\boldsymbol{\Theta}_0$ by a consistent estimate $\hat{\boldsymbol{\Theta}}$ does not affect the asymptotic properties if $p' = p$ [OVSN93], and with this weighting the variances of the estimates are minimized and achieves the CRB. Furthermore, it can be shown that the estimates (not only in the distribution) of the NSF method are asymptotically equivalent with those of the SML method [OWVK89]. Thus, when an unknown mutual coupling is estimated, the SML method that requires a search also for the coupling parameters is asymptotically equivalent to the NSF method that only requires a search for the DOAs. The NSF method is therefore a good candidate for DOA estimation if the signal covariance is of full rank ($p' = p$).

Now, several methods of estimating both DOAs and coupling have been discussed and in the next section the properties of these algorithms will be examined in some simulation scenarios.

5.7 Computer Experiments

Usually it is hard to analytically compare the performance of different estimation schemes, and here the different algorithms discussed in previous sections will be evaluated by calculating the estimates in a few different scenarios numerically. First, the convergence rate of the iterative formulations of the algorithms will be studied, and then variance of the different estimates are compared to the CRB. Throughout this section the different methods of estimation are compared using data generated by the reduced model and

thus the bias of the estimation methods is small, since there is no model error. That allows for comparisons with the CRB calculated in previous sections. Further experiments are needed to examine how well the reduced model approximates the full electromagnetic model, see Chapter 6.

The convergence rate of an iterative algorithm is of course very important, and here the convergence of the DOA estimation schemes discussed previously will be addressed. Consider a 10 element half-wavelength dipole ULA with an element separation of $\lambda/2$. Two plane waves are incident upon the array, and the DOAs (ϕ_1, ϕ_2) of these are estimated. Since it is the convergence rate that is studied, the signal power is chosen large (SNR=50dB) in order to avoid fluctuations due to the stochastic nature which is studied later when the variance is studied. The signal waveforms are assumed uncorrelated and the model and assumptions in Section 5.1 are used. Furthermore, the number of snapshots is large ($N = 10000$) to ensure that the covariance estimate $\hat{\mathbf{R}}$ is close to the true covariance.

In the first example, the waves are incident from $(76°, 78°)$ and the number of coupling parameters is $q = 9$, i.e. the maximum number for a 10 element array. The DOA estimates versus the number of iterations are shown in Figure 5.6 for the spectrally based iterative MUSIC, Root-Music, and for the ESPRIT method. The iterative MUSIC experiences the best convergence properties, followed by the Root-Music method. The ESPRIT method converges very slowly, probably due to that the criteria in (5.45) and (5.46) differ too much and the convergence properties deteriorate. That the Root-Music performs worse than the iterative MUSIC is also due to jumping between two different criteria, but the criteria are reasonably similar, in this case, since the Root-Music method converges for a larger number of iterations. Also, remember that the difference between the criteria is that the roots of Root-Music do not lie exactly on the unit circle. However, calculating the MUSIC spectra and searching for peaks is computationally demanding compared to the rooting operation in Root-Music (especially when a large resolution is wanted). Therefore, the Root-Music can use more iterations and still be competitive. The extended ESPRIT method, on the other hand, does not converge at all when the maximum number of coupling parameters was estimated and is thus not a good alternative.

Next, it is interesting to compare the convergence rate of the iterative MUSIC method with that of the SML and NSF methods that were implemented using the damped Newton method. The initial values needed for the search in the SML and NSF methods were obtained by inserting the result from the first iteration of the Root-Music method, which requires a modest amount of computations. The maximum number of coupling parameters was estimated in the previous example resulting in the most difficult

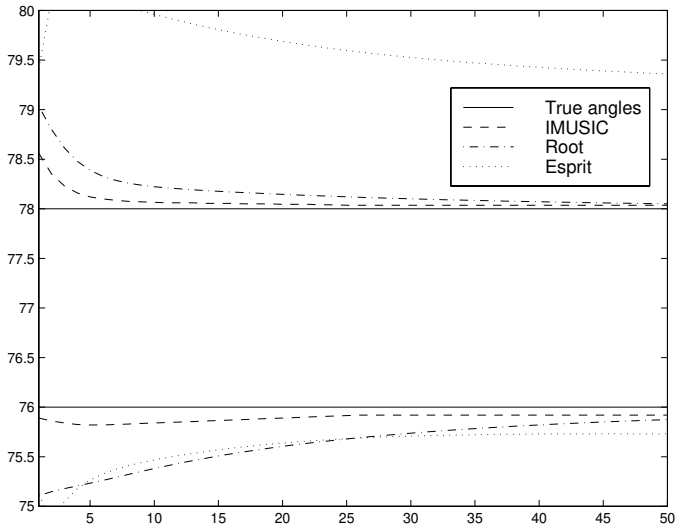


Figure 5.6: The DOA estimates of the iterative MUSIC, Root-Music, and ESPRIT method versus the number of iterations using a ULA of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart when two waves are incident from $(76°, 78°)$ and 10000 snapshots with SNR=50dB and $q = 9$.

scenario. Usually only a few coupling parameters are estimated since the magnitude falls off rapidly, see Figure 4.1. Therefore, in the next example only 3 coupling parameters will be estimated.

In Figure 5.7, two plane waves are incident from $(41°, 49°)$ and 3 coupling parameters are used. Although the number of coupling parameters is less, this is still a harder estimation scenario since the angles are closer to end-fire. Here, the convergence properties of the iterative MUSIC method is worse than for the SML and NSF methods. Furthermore, the iterative MUSIC method does not converge to the true DOA even after 25 iterations. The resolution in the spectrum calculation was 0.005 which probably is the reason why it does not converge to the true value. The rate of convergence is, however, much slower for the iterative MUSIC method than for the Newton search methods SML and NSF.

Although the convergence rate is important, it is also of interest to examine how good estimates the estimator delivers on the average. The Root Mean Square Error (RMSE) is calculated based on the true DOA and thus also the bias is included. By comparing with the CRB, the quality of the estimator is examined. Note that if the estimator is unbiased, the RMSE equals the standard deviation of the DOA estimates. Furthermore, if the method also is efficient, the RMSE coincides with the square-root of the CRB.

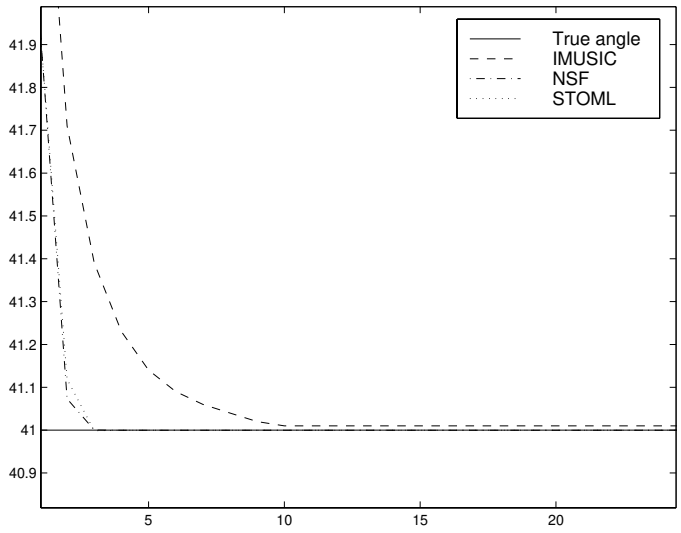


Figure 5.7: The DOA estimate of ϕ_1 for the iterative MUSIC, SML, and NSF method versus the number of iterations using a ULA of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart when two waves are incident from $(41°, 49°)$ and 10000 snapshots with SNR=50dB and $q = 3$.

In the first example, a 10 element half-wavelength dipole ULA with an element separation of $\lambda/2$ is used. Two plane waves are incident from $(75°, 80°)$, 100 snapshots are taken, and $q = 3$ coupling parameters are used and estimated. The signals are assumed to be uncorrelated and the SNR is varied from 0 dB to 30 dB. The ESPRIT method is excluded due the poor convergence properties, and the Root-Music is excluded since it essentially is the same method as the iterative MUSIC method, which performs slightly better. In Figure 5.8, the RMSE of the DOA estimate of ϕ_1 using the iterative MUSIC, SML and NSF methods is shown for different SNR. The RMSE of the SML method essentially follows the CRB as it should, and the NSF method requires slightly larger SNR to coincide with the CRB. The numerical search in SML and NSF require initial values for the DOA, and these values are obtained as the estimates from the Root-Music method after a few iterations. Remember that the NSF method is only asymptotically equivalent to the SML method, and for a small number of samples and low SNR the methods will differ. The iterative MUSIC requires larger SNR than the NSF method to work well. Also, by a close inspection of Figure 5.8 it is found that the RMSE of the iterative MUSIC method is higher than that of the SML and NSF methods that reach the CRB, see [Sva99e].

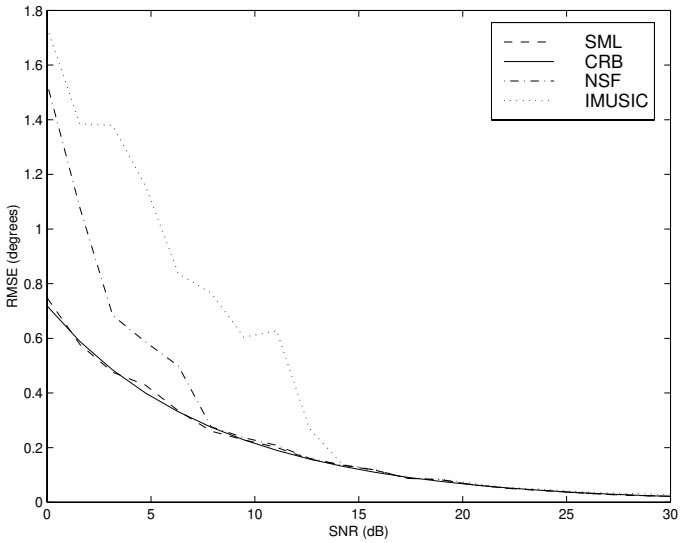


Figure 5.8: The RMSE of the DOA estimate of ϕ_1 for the iterative MUSIC, SML, and NSF method versus SNR using a ULA of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart using 200 Monte-Carlo simulations. Two waves are incident from $(75°, 80°)$ and 100 snapshots are used with $q = 3$.

Next, a more difficult scenario is considered where the DOAs are further away from broad-side $(41°, 49°)$ and two more coupling parameters ($q = 5$) are estimated. In Figure 5.9 the RMSE for the iterative MUSIC, SML, and NSF are shown together with the CRB. The SML and NSF methods follow relatively closely the CRB, but the iterative MUSIC method has a larger RMSE for high SNR. Note that the spectral resolution was $0.01°$, which should be sufficient to give a lower RMSE than that obtained in Figure 5.9. The high RMSE is due to slow convergence, and when the maximum number of iterations is reduced to 10, the RMSE further increases [Sva99e]. But for the SML and NSF methods the reduction in iterations does not affect the RMSE.

The strength of the iterative MUSIC method is that convergence is guaranteed, unlike the Newton search employed in the SML and NSF methods that can end up in a local minimum. However, in the examples studied here the Newton search finds the true minimum. The initial values taken from the first estimates from the Root-Music are close enough to the true minimum for the search to converge at the global minimum. One important property of an iterative method is the number of iterations needed to converge to the minimum, and here the Newton methods require sometimes substantially

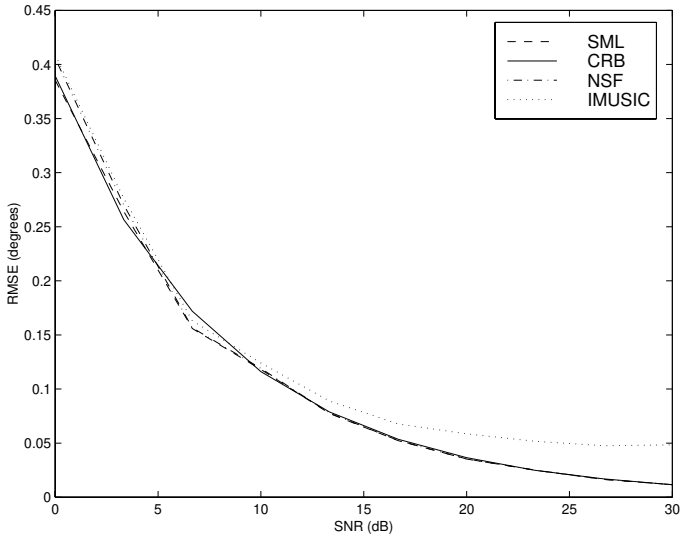


Figure 5.9: The RMSE of the DOA estimate of ϕ_1 for the iterative MUSIC, SML, and NSF method versus SNR using a ULA of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart using 200 Monte-Carlo simulations. Two waves are incident from $(41°, 49°)$ and 1000 snapshots with $q = 5$ and an iteration limit of 20.

less computations than the iterative-MUSIC method. Also, the calculation of the MUSIC spectrum and searching for peaks is computationally demanding when many iterations are needed. The implementation is of course very important, but the iterative MUSIC method is not less demanding than the NSF and SML methods, and for the coupling free case the main motivation for using MUSIC is that it is computationally cheap.

An important drawback of the iterative MUSIC method is that its threshold for resolving closely spaced signal sources is substantially higher than for the Root-Music, SML, and NSF methods. This is especially important when estimating DOAs in the presence of an unknown coupling, since the effect of the coupling is to mix different steering vectors and thus blurring the MUSIC spectrum. In Figure 5.10, the MUSIC spectrum for a known compensated coupling and an uncompensated coupling when two waves are incident from $(30°, 33°)$ is shown. Two clearly discernible peaks appear when the coupling is compensated for, but for the unknown coupling only one broad peak appears. This affects the iterative MUSIC method, since it estimates the angles as a first step and then estimates the coupling. Obtaining two DOA estimates appears difficult, and the method is not really applicable in this case.

However, resolving closely spaced sources can be of great importance and

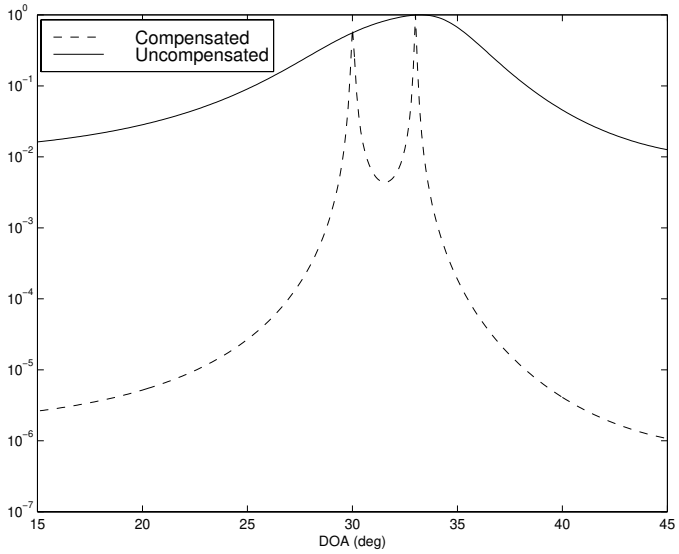


Figure 5.10: The MUSIC spectrum for a compensated coupling and an uncompensated coupling using a ULA of 15 $\lambda/2$ dipoles spaced $\lambda/2$ apart when two waves are incident from $(30°, 33°)$ and 100 snapshots with SNR = 40dB.

a lot of work have been done in this field. Here, the RMSE for closely spaced sources will be examined for the SML and NSF method. Also included in this example is the Root-Music algorithm, since the array is a ULA and the Root-Music can be applied. The initial values to the search performed for the NSF and SML methods are, as previously, obtained from Root-Music. In Figure 5.11, the RMSE of the DOA estimate of ϕ_1 for the iterative MUSIC, SML, and NSF method versus the angular separation is shown. The SML method works best and achieves the CRB for relatively small angular separations and the NSF method reaches the CRB for slightly larger separation. As expected, the Root-Music method works much worse, since two different criteria are used in that method.

The NSF method only requires a search over the DOAs, unlike the SML method that also searches for the coupling parameters resulting in a larger search. In most of the examples considered in this section. the NSF method either coincided with the SML method or was quite close. Therefore, it seems that the NSF method would be preferable in many instances, but it has one drawback since it does not work for coherent signals. In Figure 5.12, the RMSE of the DOA estimate of ϕ_1 for the iterative MUSIC, SML, and NSF method is shown for different correlations when two waves are incident from $(41°, 49°)$. Note that the signal covariance matrix and the correlation

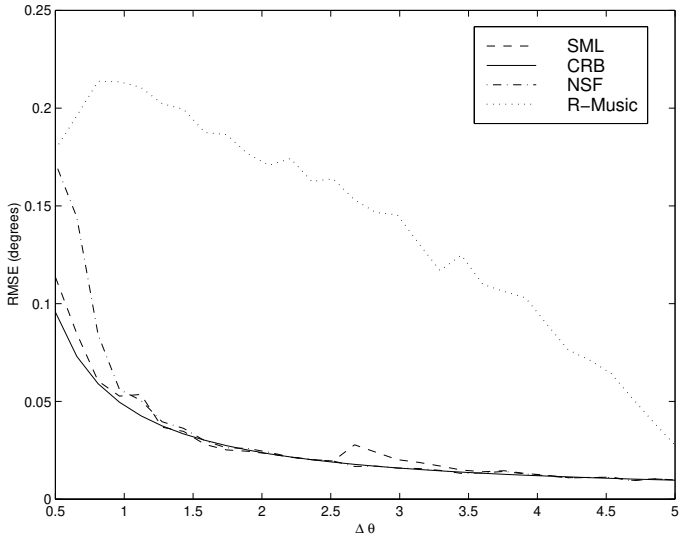


Figure 5.11: The RMSE of the DOA estimate of ϕ_1 for the iterative MUSIC, SML, and NSF method versus angular separation using a ULA of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart using 300 Monte-Carlo simulations. One wave is incident from 45° and the other is varied from 45.5° to 50°. The number of snapshots is 1000 and $q = 5$ coupling parameters are estimated with an iteration limit of 40 and SNR = 35dB.

parameter ρ are defined as

$$\mathbf{P} = SNR \times \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}. \tag{5.54}$$

The SML method follows the CRB closely as it should, and for low correlations, the NSF method coincides with the SML method. Examining the figure closely, reveals that the NSF method starts to differ from the SML at a correlation of $\rho = 0.9$. The RMSE for the NSF method turns much larger for correlations above $\rho = 0.95$. The iterative MUSIC method has a larger RMSE value for all correlations, but start to perform much worse around $\rho = 0.9$ where the RMSE value becomes several times larger than the CRB. Thus, for coherent signals the noise subspace methods do not work well and other methods should be used for highly correlated signals.

The numerical examples in this section indicate that the extension of the ESPRIT method derived herein is not a good candidate for DOA estimation due to its convergence properties. A method that also uses two different criteria is the Root-Music method, but the criteria were similar enough and the method converged. Faster convergence was obtained by using the iterative

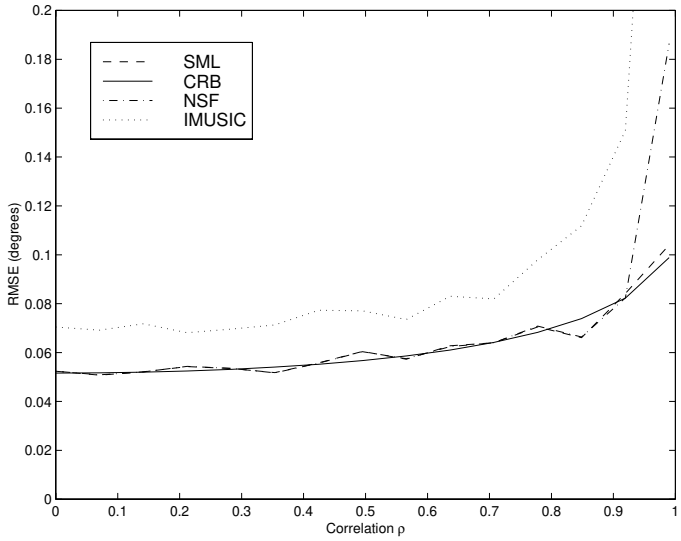


Figure 5.12: The RMSE of the DOA estimate of ϕ_1 for the iterative MUSIC, SML, and NSF method versus correlation using a ULA of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart using 200 Monte-Carlo simulations. Two waves are incident from $(41°, 49°)$ and 500 snapshots with $q = 5$ and an iteration limit of 20 and SNR = 20dB.

MUSIC method that experienced rapid convergence in the easier scenarios. However, for the more difficult scenarios also the iterative MUSIC method experienced slow convergence.

The fastest convergence was obtained for the SML and NSF methods that usually only needed a few iterations. When comparing the average error of the estimates similar results were obtained. The Root-Music method was slightly worse than the iterative MUSIC method that worked well in the easier scenarios, but for the more difficult scenarios the convergence rate became a problem. Also, the iterative MUSIC method experienced a higher RMSE than the SML and NSF method that were asymptotically efficient, i.e. reached the CRB.

One major drawback for the iterative MUSIC method is that it has a higher resolution threshold than SML and NSF. This was found to be especially important when the coupling is unknown and needs to be estimated. If not compensated for, the coupling will blur the MUSIC spectrum and further reduce the resolution threshold. Thus, the iterative MUSIC method is not suited for resolving closely spaced sources where SML and NSF probably should be used.

Finally, it was found that both the MUSIC method and the NSF method

do not work for highly correlated signals just as in the coupling free case. Apart from the correlated case, the NSF method closely followed the SML method. Therefore, the NSF method seems preferable since it only requires a p dimensional search unlike the SML method that requires a $p+q$ dimensional search.

5.8 Signal Detection

All parametric methods in this section require that the number of incident signals is known. If the number of incident waves is unknown it can be estimated and that is usually called detection. However, when the coupling also is unknown and estimated another problem appears. The number of coupling parameters to use is also unknown, and two types of detection results, signal detection and coupling detection.

The number of incident waves or signals is usually important since that could correspond to the number of aircrafts or submarines. If the signal covariance matrix is of full rank, the signal detection part essentially reduces to testing the rank of the steering matrix or the multiplicity of the smallest eigenvalue. In that case, methods like AIC and MDL can be used and no correction for the coupling is necessary. However, when the signal covariance matrix is rank deficient, i.e. there are coherent signals, other detection schemes are needed.

When the coupling is known, methods that simultaneously estimate the DOAs and the number of signals are used, see Section 3.4. If the coupling is not known and the number of coupling parameters also is unknown, those methods needs to be modified. That will not be done here, and through out this report the number of coupling parameters is assumed known. Also, in a more realistic scenario the number of coupling parameters is simply chosen as the minimum number parameters that reasonably well describe the coupling.

5.9 Conclusions

As a first approach of investigating the effects of mutual coupling, a simplified scenario of a completely unknown and deterministic mutual coupling was considered. Of course, in a more realistic scenario also a random coupling error and gain and phase uncertainties should be studied. Here, estimating an unknown coupling in this simplified scenario was investigated. First, the SML estimator was derived and using those expressions, the CRB for the unknown coupling case was also derived. If a feasible number of coupling

parameters was used, it was found that the increase in the CRB was not large.

Furthermore, a number of methods for estimating the DOAs in the presence of an unknown coupling was discussed. First, the iterative spectral MUSIC method, was reviewed and then several extensions of coupling free methods were introduced. The Root-Music method was extended to include the presence of an unknown coupling. Also, the NSF method was extended to include an unknown coupling, and it was shown that the criterion could be concentrated with respect to the coupling and thus only a search for the DOAs was needed.

These methods were also analyzed in a few numerical examples where it was found that the convergence rate of the iterative MUSIC and the Root-Music was slower than for the NSF and SML methods. The main argument of using MUSIC in the coupling free case is that it is computationally cheap. But with coupling those methods require many iterations and therefore do not require less computational power than multidimensional search methods like SML and NSF. One problem with multidimensional search methods is to find initial values for the search. Here it was found that the Root-Music method gave initial values accurate enough. Therefore, the SML and NSF methods were found to be good candidates for finding the DOA. The NSF method was found to have similar performance as the efficient SML method in most cases and therefore the NSF was usually to prefer since it only requires a search over the DOAs. But for highly correlated signals the NSF method does not work well (together with the MUSIC methods) and the SML method should be used instead.

5A The Maximum Likelihood Method

The Likelihood Function

The measured voltages $\mathbf{x}(t)$ are, using the stochastic assumptions and the data model in Section 5.1, Gaussian distributed. However, since the voltages in general are complex, they belong to a multivariate complex Gaussian distribution. Especially, since both the signal and the noise are circularly Gaussian, the likelihood function can be written in a compact form [Kay93] as

$$p(\mathbf{X};\mathbf{\Theta}) = \frac{1}{\pi^{Nn}|\mathbf{R}|^N} \exp - \left[\sum_{k=1}^{N} \mathbf{x}^H(t_k)\mathbf{R}^{-1}\mathbf{x}(t_k)\right], \tag{5A.1}$$

where

$$\mathbf{R} = \mathbf{CAPA}^H\mathbf{C}^H + \sigma^2\mathbf{I} \tag{5A.2}$$

is the covariance matrix of the voltages. Note that both the signal and the noise are assumed to be zero mean processes. Taking the logarithm of the likelihood function results in

$$\log p(\mathbf{X};\boldsymbol{\Theta}) = -N\left(n\log(\pi) + \log|\mathbf{R}| + \frac{1}{N}\sum_{k=1}^{N}\mathbf{x}^H(t_k)\mathbf{R}^{-1}\mathbf{x}(t_k)\right) \quad (5A.3)$$

Using that $Tr(\mathbf{ABC}) = Tr(\mathbf{BCA})$, the logarithm of the likelihood function can be written as

$$\log p(\mathbf{X};\boldsymbol{\Theta}) = -N\left(n\log(\pi) + \log|\mathbf{R}| + \mathrm{Tr}(\mathbf{R}^{-1}\hat{\mathbf{R}})\right), \quad (5A.4)$$

where

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{k=1}^{N}\mathbf{x}(t_k)\mathbf{x}^H(t_k). \quad (5A.5)$$

Now it is possible to maximize this expression with respect to the signal and noise covariances, see [SN95], and (5.6) results.

The Gradient and Hessian

In general, the coupling parameters are complex and instead of defining a complex differentiation, here the complex parameters are divided into a complex and real part. This results in a parameter vector $\boldsymbol{\Theta} = \{\boldsymbol{\phi}, \mathbf{c}_r, \mathbf{c}_i\}$ and the SML criterion function becomes

$$V_{SML}(\boldsymbol{\Theta}) = \log\det\left[\mathbf{CA}\hat{\mathbf{P}}\mathbf{A}^H\mathbf{C}^H + \hat{\sigma}^2\mathbf{I}\right], \quad (5A.6)$$

and inserting the expression for $\hat{\mathbf{P}}$ in (5.7) gives

$$\begin{aligned} V_{SML}(\boldsymbol{\Theta}) &= \log\det\left[\mathbf{CA}(\mathbf{CA})^\dagger\left[\hat{\mathbf{R}} - \hat{\sigma}\mathbf{I}\right](\mathbf{CA})^{\dagger H}\mathbf{A}^H\mathbf{C}^H + \hat{\sigma}^2\mathbf{I}\right] \\ &= \log\det\left[\mathbf{P}_{CA}\hat{\mathbf{R}}\mathbf{P}_{CA} + \hat{\sigma}^2\mathbf{P}_{CA}^{\perp}\right]. \end{aligned} \quad (5A.7)$$

The gradient and the Hessian are defined as

$$\mathbf{V}' = \frac{\partial V_{SML}}{\partial\boldsymbol{\Theta}} \quad (5A.8)$$

$$\mathbf{H}_{ij} = \frac{\partial^2 V_{SML}}{\partial\Theta_i\partial\Theta_j}. \quad (5A.9)$$

Using the differentiation result [Kay93]

$$\frac{\partial\log\det\mathbf{A}(x)}{\partial x} = \mathrm{Tr}\left[\mathbf{A}^{-1}(x)\frac{\partial\mathbf{A}^H(x)}{\partial x}\right], \quad (5A.10)$$

and the simpler notation $\mathbf{A}_e = \mathbf{CA}$, the gradient can be written as

$$\mathbf{V}'_k = \mathrm{Tr}\left[\left(\mathbf{P}_{A_e}\hat{\mathbf{R}}\mathbf{P}_{A_e} + \hat{\sigma}^2\mathbf{P}^{\perp}_{A_e}\right)^{-1}\frac{\partial}{\partial\Theta_k}\left(\mathbf{P}_{A_e}\hat{\mathbf{R}}\mathbf{P}_{A_e} + \hat{\sigma}^2\mathbf{P}^{\perp}_{A_e}\right)\right]. \quad \text{(5A.11)}$$

Also,

$$\frac{\partial\mathbf{P}_{A_e}}{\partial\Theta_k} = \left(\mathbf{P}_{A_e}\frac{\partial\mathbf{P}_{A_e}}{\partial\Theta_k}\right) + \left(\mathbf{P}_{A_e}\frac{\partial\mathbf{P}_{A_e}}{\partial\Theta_k}\right)^H \quad \text{(5A.12)}$$

and $\mathrm{Tr}\mathbf{B}(\mathbf{C}+\mathbf{C}^H) = 2\mathrm{ReTr}[\mathbf{BC}]$ gives

$$\begin{aligned}\mathbf{V}'_k = 2\mathrm{ReTr}&\left[\left(\mathbf{P}_{A_e}\hat{\mathbf{R}}\mathbf{P}_{A_e} + \hat{\sigma}^2\mathbf{P}^{\perp}_{A_e}\right)^{-1}\mathbf{P}_{A_e}\hat{\mathbf{R}}\frac{\partial\mathbf{P}_{A_e}}{\partial\Theta_k} + \hat{\sigma}^2\mathbf{P}^{\perp}_{A_e}\frac{\partial\mathbf{P}^{\perp}_{A_e}}{\partial\Theta_k}\right] \\ &+ \mathrm{Tr}\left[\left(\mathbf{P}_{A_e}\hat{\mathbf{R}}\mathbf{P}_{A_e} + \hat{\sigma}^2\mathbf{P}^{\perp}_{A_e}\right)^{-1}\frac{\partial\hat{\sigma}^2}{\partial\Theta_k}\mathbf{P}^{\perp}_{A_e}\right].\end{aligned} \quad \text{(5A.13)}$$

From (5.8) it follows that

$$\frac{\partial\hat{\sigma}^2}{\partial\Theta_k} = -\frac{1}{n-d}\mathrm{Tr}\left[\frac{\partial\mathbf{P}_{A_e}}{\partial\Theta_k}\hat{\mathbf{R}}\right]. \quad \text{(5A.14)}$$

Now, invoke the matrix inversion lemma

$$\begin{aligned}\left(\mathbf{P}_{A_e}\hat{\mathbf{R}}\mathbf{P}_{A_e} + \hat{\sigma}^2\mathbf{P}^{\perp}_{A_e}\right)^{-1} &= \left(\mathbf{CA}\hat{\mathbf{P}}\mathbf{A}^H\mathbf{C}^H + \hat{\sigma}^2\mathbf{I}\right)^{-1} \\ &= \frac{1}{\hat{\sigma}^2}\left[\mathbf{I} - \mathbf{A}_e\left(\hat{\mathbf{P}}\mathbf{A}_e^H\mathbf{A}_e + \hat{\sigma}^2\mathbf{I}\right)^{-1}\hat{\mathbf{P}}\mathbf{A}_e^H\right]\end{aligned} \quad \text{(5A.15)}$$

and insert in (5A.13)

$$\begin{aligned}\mathbf{V}'_k = 2\mathrm{ReTr}&\left[\frac{1}{\hat{\sigma}^2}\left(\mathbf{I} - \mathbf{A}_e\left(\hat{\mathbf{P}}\mathbf{A}_e^H\mathbf{A}_e + \hat{\sigma}^2\mathbf{I}\right)\hat{\mathbf{P}}\mathbf{A}_e^H\right)\right. \\ &\left.\left(\mathbf{P}_{A_e}\hat{\mathbf{R}}\frac{\partial\mathbf{P}_{A_e}}{\partial\Theta_k} + \hat{\sigma}^2\mathbf{P}^{\perp}_{A_e}\frac{\partial\mathbf{P}^{\perp}_{A_e}}{\partial\Theta_k}\right)\right] \\ &- \mathrm{Tr}\left[\frac{1}{\hat{\sigma}^2}\left(\mathbf{I} - \mathbf{A}_e\left(\hat{\mathbf{P}}\mathbf{A}_e^H\mathbf{A}_e + \hat{\sigma}^2\mathbf{I}\right)\hat{\mathbf{P}}\mathbf{A}_e^H\right)\frac{\mathbf{P}^{\perp}_{A_e}}{n-p}\mathrm{Tr}\left(\frac{\partial\mathbf{P}_{A_e}}{\partial\Theta_k}\hat{\mathbf{R}}\right)\right].\end{aligned} \quad \text{(5A.16)}$$

However,

$$\mathrm{Tr}\frac{\mathbf{P}^{\perp}_{A_e}}{n-p}\mathrm{Tr}\left(\frac{\partial\mathbf{P}_{A_e}}{\partial\Theta_k}\hat{\mathbf{R}}\right) = 2\mathrm{ReTr}\left[\mathbf{P}_{A_e}\hat{\mathbf{R}}\frac{\partial\mathbf{P}_{A_e}}{\partial\Theta_k}\right] \quad \text{(5A.17)}$$

and $\mathbf{A}_e^H \mathbf{P}_{A_e}^{\perp} = 0$ gives

$$\mathbf{V}_k' = 2\mathrm{ReTr}\left[\mathbf{P}_{A_e}^{\perp}\frac{\partial \mathbf{P}_{A_e}^{\perp}}{\partial \Theta_k} - \frac{1}{\hat{\sigma}^2}\mathbf{A}_e\left(\hat{\mathbf{P}}\mathbf{A}_e^H\mathbf{A}_e + \hat{\sigma}^2\mathbf{I}\right)^{-1}\hat{\mathbf{P}}\mathbf{A}_e^H\mathbf{P}_{A_e}\hat{\mathbf{R}}\frac{\partial \mathbf{P}_{A_e}}{\partial \Theta_k}\right]. \tag{5A.18}$$

Next, note that

$$\mathrm{Tr}\left(\mathbf{P}_{A_e}^{\perp}\frac{\partial \mathbf{P}_{A_e}^{\perp}}{\partial \Theta_k}\right) = 0 \tag{5A.19}$$

and

$$\frac{\partial \mathbf{P}_{A_e}}{\partial \Theta_k}\mathbf{A}_e = \mathbf{P}_{A_e}^{\perp}\frac{\partial \mathbf{A}_e}{\partial \Theta_k}(\mathbf{A}_e)^{\dagger}\mathbf{A}_e + (\mathbf{A}_e)^{\dagger H}\frac{\partial \mathbf{A}^H\mathbf{C}^H}{\partial \Theta_k}\mathbf{P}_{A_e}^{\perp}\mathbf{A}_e = \mathbf{P}_{A_e}^{\perp}\frac{\partial \mathbf{A}_e}{\partial \Theta_k}. \tag{5A.20}$$

Furthermore,

$$\begin{aligned}\hat{\mathbf{P}}(\mathbf{A}_e)^H\mathbf{A}_e + \hat{\sigma}^2 =& (\mathbf{A}_e)^{\dagger}\left[\hat{\mathbf{R}} - \hat{\sigma}^2\mathbf{I}\right](\mathbf{A}_e)^{\dagger H}\mathbf{A}_e^H\mathbf{A}_e + \hat{\sigma}^2 \\ =& (\mathbf{A}_e^H\mathbf{A}_e)^{-1}\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{A}_e\end{aligned} \tag{5A.21}$$

which reduces the gradient expression to

$$\begin{aligned}\mathbf{V}_k' =& -2\mathrm{ReTr}\left[\frac{1}{\hat{\sigma}^2}(\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{A}_e)^{-1}\mathbf{A}_e^H\mathbf{A}_e\left(\hat{\mathbf{P}}\mathbf{A}_e^H\mathbf{P}_{A_e}\hat{\mathbf{R}}\mathbf{P}_{A_e}^{\perp}\frac{\partial(\mathbf{A}_e)}{\partial \Theta_k}\right)\right] \\ =& 2\mathrm{ReTr}\left[\left\{(\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{A}_e)^{-1} - \frac{1}{\hat{\sigma}^2}(\mathbf{A}_e^H\mathbf{A}_e)^{-1}\right\}\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{P}_{A_e}^{\perp}\frac{\partial(\mathbf{A}_e)}{\partial \Theta_k}\right].\end{aligned} \tag{5A.22}$$

Next, introduce the matrix

$$\hat{\mathbf{G}} = (\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{A}_e)^{-1} - \frac{1}{\hat{\sigma}^2}(\mathbf{A}_e^H\mathbf{A}_e)^{-1}. \tag{5A.23}$$

Inserting $\mathbf{A}_e = \mathbf{C}\mathbf{A}$, the gradient can be written

$$\mathbf{V}'(\mathbf{\Theta}) = \begin{bmatrix} 2\mathrm{ReTr}\left[\hat{\mathbf{G}}\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}_{CA}^{\perp}\frac{\partial(\mathbf{C}\mathbf{A})}{\partial \Theta_1}\right] \\ \vdots \\ 2\mathrm{ReTr}\left[\hat{\mathbf{G}}\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}_{CA}^{\perp}\frac{\partial(\mathbf{C}\mathbf{A})}{\partial \Theta_{p+2q}}\right] \end{bmatrix}. \tag{5A.24}$$

Here, it is important to note that the expression for the differentiation with respect to the DOA can be simplified by noting that

$$\frac{\partial(\mathbf{C}\mathbf{A})}{\partial \phi_k} = \mathbf{C}\frac{\partial \mathbf{A}}{\partial \phi_k}, \tag{5A.25}$$

where the derivative has only one non-zero column, $\mathbf{a}_k$, and thus can the expressions for the p derivatives with respect to the DOAs be reduced to

$$\frac{\partial \mathbf{V}}{\partial \phi_k} = 2\mathrm{Re}\left[\hat{\mathbf{G}}_{k,:}\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{\overline{CA}}\mathbf{C}\mathbf{a}_k\right], \tag{5A.26}$$

where $\mathbf{G}_{k,:}$ denotes the k^{th} row of the matrix $\mathbf{G}$. For the derivatives with respect to the coupling there are in general no non-zero columns and therefore the expressions become

$$\frac{\partial \mathbf{V}}{\partial c_{rk}} = 2\mathrm{ReTr}\left[\hat{\mathbf{G}}\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{\overline{CA}}\boldsymbol{\Upsilon}_k\mathbf{A}\right] \tag{5A.27}$$

$$\frac{\partial \mathbf{V}}{\partial c_{ik}} = -2\mathrm{ImTr}\left[\hat{\mathbf{G}}\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{\overline{CA}}\boldsymbol{\Upsilon}_k\mathbf{A}\right], \tag{5A.28}$$

where $\boldsymbol{\Upsilon}_k$ is a banded symmetric Toeplitz matrix with zeros everywhere except on the k^{th} and the $-k^{th}$ subdiagonal which consists of ones.

The Hessian, using the chain-rule, becomes

$$\begin{aligned}\frac{\partial^2 \mathbf{V}}{\partial\Theta_i\partial\Theta_j} = &2\mathrm{ReTr}\left[\frac{\partial}{\partial\Theta_j}\left(\frac{\partial(\mathbf{CA})}{\partial\Theta_i}\hat{\mathbf{G}}\right)\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{\overline{CA}}\right] \\ &+ 2\mathrm{ReTr}\left[\frac{\partial(\mathbf{CA})}{\partial\Theta_i}\hat{\mathbf{G}}\frac{\partial}{\partial\Theta_j}\left(\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{\overline{CA}}\right)\right].\end{aligned} \tag{5A.29}$$

Evaluating this expression exactly leads to rather cumbersome expressions, but by considering the asymptotic Hessian, i.e. a large number of samples $N \to \infty$, the expression can be reduced considerably. Since the estimates of σ and $\mathbf{P}$ are known to be asymptotically statistically consistent and efficient [Böh86] and both are functions of $\hat{\mathbf{R}}$ which also is consistent and thus the asymptotic Hessian is found by simply replacing $\hat{\mathbf{R}}$ with $\mathbf{R}$. Also, the steering matrix $\mathbf{A}(\phi)$ is evaluated at the true DOAs ϕ_0. Using the reduced notation $\mathbf{A}_e = \mathbf{CA}$ once again and observing that

$$\mathbf{A}_e^H\mathbf{R}\mathbf{P}^{\perp}_{A_e} = \mathbf{A}_e^H\left(\mathbf{A}_e\mathbf{P}\mathbf{A}_e^H + \sigma^2\mathbf{I}\right)\left(\mathbf{I} - \mathbf{A}_e(\mathbf{A}_e^H\mathbf{A}_e)^{-1}\mathbf{A}_e^H\right) = 0, \tag{5A.30}$$

the asymptotic Hessian expression becomes

$$\frac{\partial^2 \mathbf{V}_a}{\partial\Theta_i\partial\Theta_j} = 2\mathrm{ReTr}\left[\frac{\partial\mathbf{A}_e}{\partial\Theta_i}\hat{\mathbf{G}}\left(\frac{\partial\mathbf{A}_e}{\partial\Theta_j}\mathbf{R}\mathbf{P}^{\perp}_{A_e} + \mathbf{A}_e^H\mathbf{R}\frac{\partial\mathbf{P}^{\perp}_{A_e}}{\partial\Theta_j}\right)\right]. \tag{5A.31}$$

However,

$$\mathbf{A}_e^H\mathbf{R}\frac{\partial\mathbf{P}^{\perp}_{A_e}}{\partial\Theta_j} = -\mathbf{A}_e^H\mathbf{R}\mathbf{A}_e(\mathbf{A}_e^H\mathbf{A}_e)^{-1}\frac{\partial\mathbf{A}_e^H}{\partial\Theta_j}\mathbf{P}^{\perp}_{A_e} \tag{5A.32}$$

and

$$\mathbf{R}\mathbf{P}_{A_e}^{\perp} = \sigma^2 \mathbf{P}_{A_e}^{\perp}. \tag{5A.33}$$

Thus,

$$\frac{\partial^2 \mathbf{V}_a}{\partial\Theta_i \partial\Theta_j} = 2\mathrm{ReTr}\left[\hat{\mathbf{G}}\left(\sigma^2\mathbf{I} - \mathbf{A}_e^H\mathbf{R}\mathbf{A}_e(\mathbf{A}_e^H\mathbf{A}_e)^{-1}\right)\frac{\partial \mathbf{A}_e^H}{\partial\Theta_j}\mathbf{P}_{A_e}^{\perp}\frac{\partial \mathbf{A}_e}{\partial\Theta_i}\right] \tag{5A.34}$$

and by applying the matrix inversion lemma

$$\begin{aligned}\left(\mathbf{A}_e^H\mathbf{R}\mathbf{A}_e\right)^{-1} =& (\mathbf{A}_e^H\mathbf{A}_e)^{-1}\left(\mathbf{A}_e^H\mathbf{A}_e\mathbf{P} + \sigma^2\mathbf{I}\right)^{-1} \\ =& \frac{1}{\sigma^2}(\mathbf{A}_e^H\mathbf{A}_e)^{-1}\left[\mathbf{I} - \mathbf{P}\mathbf{A}_e\mathbf{A}_e\left(\mathbf{P}\mathbf{A}_e^H\mathbf{A}_e + \sigma^2\mathbf{I}\right)^{-1}\right]\end{aligned} \tag{5A.35}$$

the asymptotic Hessian reduces to

$$\frac{\partial^2 \mathbf{V}_a}{\partial\Theta_i \partial\Theta_j} = \frac{2}{\sigma^2}\mathrm{ReTr}\left[\mathbf{P}\mathbf{A}_e^H\mathbf{R}^{-1}\mathbf{A}_e\mathbf{P}\frac{\partial \mathbf{A}_e^H}{\partial\Theta_j}\mathbf{P}_{A_e}^{\perp}\frac{\partial \mathbf{A}_e}{\partial\Theta_i}\right]. \tag{5A.36}$$

Furthermore, the asymptotic Hessian can be written as a block matrix

$$\mathbf{H} = \mathbf{V}_a''(\mathbf{\Theta}) = \frac{\partial^2 \mathbf{V}_a}{\partial \mathbf{\Theta}^2} = \begin{bmatrix}\mathbf{H}_{\phi\phi} & \mathbf{H}_{\phi c} \\ \mathbf{H}_{c\phi} & \mathbf{H}_{cc}\end{bmatrix}. \tag{5A.37}$$

Since the derivative with respect to the DOA results in a matrix with only one non-zero column, $\mathbf{H}_{\phi\phi}$ can be written as

$$\mathbf{H}_{\phi\phi} = \frac{2}{\sigma^2}\mathrm{Re}\left\{\left(\mathbf{P}\mathbf{A}^H\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}\mathbf{A}\mathbf{P}\right)^T \odot \left(\mathbf{D}^H\mathbf{C}^H\mathbf{P}_{CA}^{\perp}\mathbf{C}\mathbf{D}\right)\right\}, \tag{5A.38}$$

where

$$\mathbf{D} = \left[\frac{\partial \mathbf{a}(\phi)}{\partial\phi}\Big|_{\phi=\phi_1}, \cdots, \frac{\partial \mathbf{a}(\phi)}{\partial\phi}\Big|_{\phi=\phi_p}\right]. \tag{5A.39}$$

Also,

$$\mathbf{H}_{\phi c} = \begin{bmatrix}\dfrac{\partial^2 \mathbf{V}_a}{\partial\phi\partial c_r} & \dfrac{\partial^2 \mathbf{V}_a}{\partial\phi\partial c_i}\end{bmatrix} = \mathbf{H}_{\phi c}^T, \tag{5A.40}$$

where

$$\left[\frac{\partial^2 \mathbf{V}_a}{\partial\phi\partial c_r}\right]_{ij} = \frac{2}{\sigma^2}\mathrm{ReTr}\left[\mathbf{P}\mathbf{A}^H\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}\mathbf{A}\mathbf{P}\mathbf{A}^H\mathbf{\Upsilon}_j\mathbf{P}_{CA}^{\perp}\mathbf{C}\frac{\partial \mathbf{A}}{\partial\phi_i}\right] \tag{5A.41}$$

$$\left[\frac{\partial^2 \mathbf{V}_a}{\partial\phi\partial c_i}\right]_{ij} = -\frac{2}{\sigma^2}\mathrm{ImTr}\left[\mathbf{P}\mathbf{A}^H\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}\mathbf{A}\mathbf{P}\mathbf{A}^H\mathbf{\Upsilon}_j\mathbf{P}_{CA}^{\perp}\mathbf{C}\frac{\partial \mathbf{A}}{\partial\phi_i}\right]. \tag{5A.42}$$

Finally,

$$\mathbf{H}_{cc} = \begin{bmatrix}\mathbf{H}_{c_rc_r} & \mathbf{H}_{c_rc_i} \\ \mathbf{H}_{c_rc_i}^T & \mathbf{H}_{c_ic_i}\end{bmatrix}, \tag{5A.43}$$

where

$$[\mathbf{H}_{c_r c_r}]_{ij} = \frac{2}{\sigma^2}\mathrm{ReTr}\left[\mathbf{P}\mathbf{A}^H\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}\mathbf{A}\mathbf{P}\mathbf{A}^H\boldsymbol{\Upsilon}_j\mathbf{P}_{CA}^{\perp}\boldsymbol{\Upsilon}_i\mathbf{A}\right] \quad (5A.44)$$

$$[\mathbf{H}_{c_r c_i}]_{ij} = -\frac{2}{\sigma^2}\mathrm{ImTr}\left[\mathbf{P}\mathbf{A}^H\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}\mathbf{A}\mathbf{P}\mathbf{A}^H\boldsymbol{\Upsilon}_j\mathbf{P}_{CA}^{\perp}\boldsymbol{\Upsilon}_i\mathbf{A}\right] \quad (5A.45)$$

$$[\mathbf{H}_{c_i c_i}]_{ij} = -\frac{2}{\sigma^2}\mathrm{ReTr}\left[\mathbf{P}\mathbf{A}^H\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}\mathbf{A}\mathbf{P}\mathbf{A}^H\boldsymbol{\Upsilon}_j\mathbf{P}_{CA}^{\perp}\boldsymbol{\Upsilon}_i\mathbf{A}\right]. \quad (5A.46)$$

5B The Asymptotic Covariance of SML

The SML estimates $\hat{\boldsymbol{\Theta}}$ are obtained as

$$\hat{\boldsymbol{\Theta}} = \{\hat{\boldsymbol{\phi}}, \hat{\mathbf{c}}\} = \arg\min_{\Theta} V_{SML}(\boldsymbol{\Theta}), \quad (5B.1)$$

where $V_{SML}(\boldsymbol{\Theta})$ is defined in (5A.6). Introducing the simplified notation $V_{SML}(\boldsymbol{\Theta}) = V(\boldsymbol{\Theta})$, the gradient of V with respect to $\boldsymbol{\Theta}$ which is derived in Appendix 5A becomes

$$\mathbf{V}'(\hat{\boldsymbol{\Theta}}) = \frac{\partial \mathbf{V}(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}}\Big|_{\boldsymbol{\Theta}=\hat{\boldsymbol{\Theta}}} = 0. \quad (5B.2)$$

A Taylor series expansion of (5B.2) around the true parameter vector $\boldsymbol{\Theta}_0$ gives

$$0 = \mathbf{V}'(\boldsymbol{\Theta}_0) + \mathbf{V}''(\boldsymbol{\Theta}_0)(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) + \cdots, \quad (5B.3)$$

where $\mathbf{V}''(\boldsymbol{\Theta}_0)$ denotes the Hessian matrix. The higher order terms can be neglected since $\hat{\boldsymbol{\Theta}}$ converges to $\boldsymbol{\Theta}_0$ as $N \to \infty$ under very general conditions [KS61]. Furthermore, the Hessian can be replaced by the asymptotic Hessian

$$\mathbf{V}_a''(\boldsymbol{\Theta}_0) = \lim_{N\to\infty} \mathbf{V}''(\boldsymbol{\Theta}_0), \quad (5B.4)$$

since all terms that tend to zero faster than $(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)$ can be neglected in the asymptotic analysis. The asymptotic Hessian were derived in Appendix 5A. Thus,

$$\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0 = -\left[\mathbf{V}_a''(\boldsymbol{\Theta}_0)\right]^{-1}\mathbf{V}'(\boldsymbol{\Theta}_0) \quad (5B.5)$$

assuming that the inverse exists which is true if the problem is well posed, see Section 5.3. The covariance matrix for the asymptotic SML estimates then becomes

$$\mathbf{B}_{SML} = \left[\mathbf{V}_a''(\boldsymbol{\Theta}_0)\right]^{-1}\mathbf{B}\left[\mathbf{V}_a''(\boldsymbol{\Theta}_0)\right]^{-1}, \quad (5B.6)$$

where

$$\mathbf{B} = \lim_{N\to\infty} \mathrm{E}\left[\mathbf{V}'(\boldsymbol{\Theta}_0)\right]\left[\mathbf{V}'(\boldsymbol{\Theta}_0)\right]^T \quad (5B.7)$$

Since the asymptotic Hessian already is derived in previous appendix, it is left to evaluate $\mathbf{B}$. The gradient has elements

$$\mathbf{V}'(\mathbf{\Theta}_i) = 2\mathrm{ReTr}\left[\hat{\mathbf{G}}\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{CA}\frac{\partial(\mathbf{CA})}{\partial\Theta_i}\right], \tag{5B.8}$$

where $\hat{\mathbf{G}}$ is defined in (5A.23). Since $\mathbf{A}^H\mathbf{C}^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{CA}$ tends to zero as $N \to \infty$, $\hat{\mathbf{G}}$ can be replaced with $\mathbf{G}$, and $\hat{\mathbf{R}}$ and $\hat{\sigma}$ are replaced by $\mathbf{R}$ and σ without affecting the asymptotic covariance matrix $\mathbf{B}$ [SN90b]. Using the simpler notation $\mathbf{A}_e = \mathbf{CA}$, the elements become

$$\begin{aligned}
\mathbf{B}_{ij} &= 4\lim_{N\to\infty}\mathrm{E}\left\{\mathrm{Re}\ \mathrm{Tr}\left(\mathbf{G}\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_e)}{\partial\Theta_i}\right)\cdot\mathrm{Re}\ \mathrm{Tr}\left(\mathbf{G}\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_e)}{\partial\Theta_j}\right)\right\} \\
&= 2\lim_{N\to\infty}\mathrm{E}\left\{\mathrm{Re}\left[\mathrm{Tr}\left(\mathbf{G}\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_e)}{\partial\Theta_i}\right)\cdot\mathrm{Tr}\left(\mathbf{G}\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_e)}{\partial\Theta_j}\right)\right]\right\} \\
&+ 2\lim_{N\to\infty}\mathrm{E}\left\{\mathrm{Re}\left[\mathrm{Tr}\left(\mathbf{G}\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_e)}{\partial\Theta_i}\right)\cdot\mathrm{Tr}\left(\frac{\partial(\mathbf{A}_e^H)}{\partial\Theta_j}\mathbf{P}^{\perp}_{A_e}\hat{\mathbf{R}}\mathbf{A}_e\mathbf{G}\right)\right]\right\}.
\end{aligned} \tag{5B.9}$$

Now, using $\mathrm{Tr}\mathbf{A}\cdot\mathrm{Tr}\mathbf{B} = \mathrm{Tr}\mathbf{A}\otimes\mathbf{B}$ [Lüt96] the first term in (5B.9) becomes

$$\begin{aligned}
&\mathrm{Re}\left\{\mathrm{E}\left[\mathrm{Tr}\left[\left(\mathbf{G}\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_e)}{\partial\Theta_i}\right)\otimes\left(\mathbf{G}\mathbf{A}_e^H\hat{\mathbf{R}}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_e)}{\partial\Theta_j}\right)\right]\right]\right\} \\
&= \mathrm{Re}\left\{\mathrm{E}\left[\mathrm{Tr}\left[\frac{1}{N}\sum_{m=1}^{N}\left(\mathbf{G}\mathbf{A}_e^H\mathbf{x}(t_m)\mathbf{x}^H(t_m)\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_e)}{\partial\Theta_i}\right)\right.\right.\right. \\
&\left.\left.\left.\otimes\frac{1}{N}\sum_{n=1}^{N}\left(\mathbf{G}\mathbf{A}_e^H\mathbf{x}(t_n)\mathbf{x}^H(t_n)\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_e)}{\partial\Theta_j}\right)\right]\right]\right\},
\end{aligned} \tag{5B.10}$$

where the expression for $\hat{\mathbf{R}}$ (5A.5) has been inserted. Furthermore, since

$$\mathbf{x}^H(t_n)\mathbf{P}^{\perp}_{A_e} = \mathbf{n}^H(t_n)\mathbf{P}^{\perp}_{A_e} \tag{5B.11}$$

the first term can be written as

$$\begin{aligned}
\frac{1}{N^2}\mathrm{E}\Bigg[\sum_{l=1}^{p}\sum_{k=1}^{p}\sum_{m=1}^{N}\sum_{n=1}^{N}&\underbrace{\mathbf{G}_{k,:}\mathbf{A}_e^H\mathbf{x}(t_m)}_{\mathbf{A}}\underbrace{\mathbf{n}^H(t_m)\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i}}_{\mathbf{B}} \\
&\times\underbrace{\mathbf{G}_{l,:}\mathbf{A}_e^H\mathbf{x}(t_n)}_{\mathbf{C}}\underbrace{\mathbf{n}^H(t_n)\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_{e_l})}{\partial\Theta_j}}_{\mathbf{D}}\Bigg].
\end{aligned} \tag{5B.12}$$

Here, $\mathbf{G}_{k,:}$ denotes the k^{th} row of $\mathbf{G}$ and $\frac{\partial(\mathbf{A}_{e_l})}{\partial\Theta_j}$ denotes the l^{th} column. The expectation of the product of four Gaussian random variables can be expressed as [JS88]

$$\begin{aligned}\mathrm{E}[\mathbf{ABCD}] =& \mathrm{E}[\mathbf{AB}]\cdot\mathrm{E}[\mathbf{CD}] \\ &+ \sum_{k=1}^{b_2}\mathrm{E}[(\mathbf{1}_k^T\mathbf{C})\otimes\mathbf{A}]\cdot\mathrm{E}[\mathbf{D}\otimes(\mathbf{B}\mathbf{1}_k)] \\ &+ \mathrm{E}[\mathbf{A}(\mathrm{E}[\mathbf{BC}])\mathbf{D}] - 2\mathrm{E}[\mathbf{A}]\mathrm{E}[\mathbf{B}]\mathrm{E}[\mathbf{C}]\mathrm{E}[\mathbf{D}],\end{aligned} \tag{5B.13}$$

where $\mathbf{1}_k$ denotes the vector having a "1" at the k^{th} position and zero elsewhere. Also, $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}$ are jointly Gaussian random variables of dimensions $(a_1\times a_2),(b_1\times b_2),(c_1\times c_2),(d_1\times d_2)$. In the present case, $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}$ are scalars. Since the noise and the signal are assumed uncorrelated and $\mathbf{A}_e^H\mathbf{P}_{A_e}^{\perp}=0$

$$\mathrm{E}[\mathbf{AB}] = \mathrm{E}[\mathbf{BC}] = 0. \tag{5B.14}$$

Next, note that $\mathrm{E}[\mathbf{n}^*(t_m)\mathbf{n}^H(t_n)]=0\ \forall\ m,n$ giving

$$\mathrm{E}[\mathbf{D}\otimes\mathbf{B}\mathbf{1}_k] = \mathrm{E}[\mathbf{DB}] = \mathrm{E}[\mathbf{B}^T\mathbf{D}] = 0. \tag{5B.15}$$

Finally, since both the signal and the noise are assumed to be zero mean $\mathrm{E}[\mathbf{B}]=0$ and all the terms in (5B.12) become zero, and so also the first term in (5B.9).

Proceeding with the second term in (5B.9), it can also be rewritten along the same lines as

$$\begin{aligned}\frac{1}{N^2}\mathrm{E}\Bigg[\sum_{l=1}^{p}\sum_{k=1}^{p}\sum_{m=1}^{N}\sum_{n=1}^{N}&\underbrace{\mathbf{G}_{k,:}\mathbf{A}_e^H\mathbf{x}(t_m)}_{\mathbf{A}}\underbrace{\mathbf{n}^H(t_m)\mathbf{P}_{A_e}^{\perp}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i}}_{\mathbf{B}} \\ &\times\underbrace{\frac{\partial(\mathbf{A}_{e_l}^H)}{\partial\Theta_j}\mathbf{P}_{A_e}^{\perp}\mathbf{n}(t_n)}_{\mathbf{C}}\underbrace{\mathbf{x}^H(t_n)\mathbf{A}_e\mathbf{G}_{l,:}^H}_{\mathbf{D}}\Bigg].\end{aligned} \tag{5B.16}$$

The product formula in (5B.13) will be used also for the second term in (5B.9), but first the different expectations need to be evaluated. Since the matrices $\mathbf{A}$ and $\mathbf{B}$ are the same as in previous case $\mathrm{E}[\mathbf{AB}]$=0. Next, note that

$$\begin{aligned}\mathrm{E}[\mathbf{DB}] &= \mathrm{E}\left[\mathbf{x}^H(t_n)\mathbf{A}_e\mathbf{G}_{l,:}^H\mathbf{n}^H(t_m)\mathbf{P}_{A_e}^{\perp}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i}\right] \\ &= \frac{\partial(\mathbf{A}_{e_k}^T)}{\partial\Theta_i}\mathbf{P}_{A_e}^{\perp\ T}\mathrm{E}[\mathbf{n}^*(t_m)\mathbf{n}^H(t_n)]\mathbf{A}_e\mathbf{G}_{l,:}^H = 0\end{aligned} \tag{5B.17}$$

and also that

$$\begin{aligned}\mathrm{E}]\mathbf{BC}] &= \mathrm{E}\left[\mathbf{n}^H(t_n)\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i}\frac{\partial(\mathbf{A}^H_{e_l})}{\partial\Theta_j}\mathbf{P}^{\perp}_{A_e}\mathbf{n}(t_m)\right] \\ &= \frac{\partial(\mathbf{A}^H_{e_l})}{\partial\Theta_j}\mathbf{P}^{\perp}_{A_e}\mathrm{E}[\mathbf{n}(t_m)\mathbf{n}^H(t_n)]\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i} = \sigma^2\frac{\partial(\mathbf{A}^H_{e_l})}{\partial\Theta_j}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i}\delta_{mn},\end{aligned} \tag{5B.18}$$

where δ_{mn} denotes the Kronecker delta. Furthermore,

$$\begin{aligned}\mathrm{E}[\mathbf{A}\left(\mathrm{E}[\mathbf{BC}]\right)\mathbf{D}] = \mathrm{E}[\mathbf{BC}]\mathrm{E}[\mathbf{AD}] &= \mathrm{E}[\mathbf{BC}]\mathbf{G}_{k,:}\mathbf{A}_e^H\mathrm{E}[\mathbf{x}(t_n)\mathbf{x}^H(t_m)]\mathbf{A}_e\mathbf{G}^H_{l,:} \\ &= \sigma^2\frac{\partial(\mathbf{A}^H_{e_l})}{\partial\Theta_j}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i}\mathbf{G}_{k,:}\mathbf{A}_e^H\mathbf{R}\mathbf{A}_e\mathbf{G}^H_{l,:}\delta_{mn}\end{aligned} \tag{5B.19}$$

and since $\mathrm{E}[\mathbf{B}] = 0$ the second term in (5B.9) becomes

$$\begin{aligned}&\mathrm{Re}\left[\frac{1}{N^2}\sum_{l=1}^{p}\sum_{k=1}^{p}\sum_{m=1}^{N}\sum_{n=1}^{N}\sigma^2\frac{\partial(\mathbf{A}^H_{e_l})}{\partial\Theta_j}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i}\mathbf{G}_{k,:}\mathbf{A}_e^H\mathbf{R}\mathbf{A}_e\mathbf{G}^H_{l,:}\delta_{mn}\right] \\ &= \frac{\sigma^2}{N^2}\mathrm{Re}\left[\sum_{k=1}^{N}\sum_{l=1}^{N}\frac{\partial(\mathbf{A}^H_{e_l})}{\partial\Theta_j}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i}\mathbf{G}_{k,:}\mathbf{A}_e^H\mathbf{R}\mathbf{A}_e\mathbf{G}^H_{l,:}\right] \\ &= \frac{\sigma^2}{N}\mathrm{ReTr}\left[\frac{\partial(\mathbf{A}^H_{e_l})}{\partial\Theta_j}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i}\mathbf{G}\mathbf{A}_e^H\mathbf{R}\mathbf{A}_e\mathbf{G}^H\right].\end{aligned} \tag{5B.20}$$

Thus, the elements of the asymptotic covariance matrix $\mathbf{B}$ becomes

$$\mathbf{B}_{ij} = \frac{2\sigma^2}{N}\mathrm{ReTr}\left[\frac{\partial(\mathbf{A}^H_{e_l})}{\partial\Theta_j}\mathbf{P}^{\perp}_{A_e}\frac{\partial(\mathbf{A}_{e_k})}{\partial\Theta_i}\mathbf{G}\mathbf{A}_e^H\mathbf{R}\mathbf{A}_e\mathbf{G}^H\right]. \tag{5B.21}$$

However, that is only a scaled version of the asymptotic Hessian in (5A.36) and therefore

$$\mathbf{B} = \frac{1}{N}\mathbf{V}_a''(\mathbf{\Theta}). \tag{5B.22}$$

Thus, the covariance matrix of the gradient of the concentrated log likelihood coincides with the expected value of the Hessian of the concentrated log likelihood apart from a scale factor N. For the full log likelihood function, this is a general result [MKB79]; and it is not surprising that it is valid also for the concentrated likelihood function. Finally, the asymptotic covariance matrix for the SML estimates, using (5B.6), becomes

$$\mathbf{B}_{SML} = \frac{1}{N}[\mathbf{V}_a''(\mathbf{\Theta})]^{-1}. \tag{5B.23}$$

5C The Noise Subspace Fitting Method

The NSF estimates are obtained as

$$\hat{\boldsymbol{\Theta}} = \left\{\hat{\boldsymbol{\phi}}, \hat{\mathbf{c}}\right\} = \arg\min_{\Theta} V_{NSF}(\boldsymbol{\Theta}) = \arg\min_{\Theta} \|\hat{\mathbf{E}}_n^H \mathbf{C}(\mathbf{c})\mathbf{A}(\phi)\|_U^2, \tag{5C.1}$$

where $\mathbf{U}$ is a $p \times p$ positive definite weighting matrix. The criterion can also be written as

$$\begin{aligned} V_{NSF}(\boldsymbol{\phi}, \mathbf{c}) = \mathrm{Tr}\left[\mathbf{U}\mathbf{A}^H(\boldsymbol{\phi})\mathbf{C}^H(\mathbf{c})\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\mathbf{C}(\mathbf{c})\mathbf{A}(\boldsymbol{\phi})\right] = \\ \mathrm{Tr}\Big[\underbrace{\mathbf{A}(\boldsymbol{\phi})\mathbf{U}\mathbf{A}^H(\boldsymbol{\phi})}_{\mathbf{A}}\underbrace{\mathbf{C}^H(\mathbf{c})}_{\mathbf{B}}\underbrace{\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H}_{\mathbf{C}}\underbrace{\mathbf{C}(\mathbf{c})}_{\mathbf{D}}\Big] \end{aligned} \tag{5C.2}$$

Using the following rule for the Kronecker product [Lüt96]

$$\mathrm{Tr}\{\mathbf{ABCD}\} = \mathrm{vec}(\mathbf{D}^{\mathrm{T}})^{\mathrm{T}}(\mathbf{C}^{\mathrm{T}} \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}) \tag{5C.3}$$

gives

$$V_{NSF} = \mathrm{vec}\left(\mathbf{C}^{\mathrm{T}}\right)^{\mathrm{T}}\left[\hat{\mathbf{E}}_{\mathrm{n}}^{*}\hat{\mathbf{E}}_{\mathrm{n}}^{\mathrm{T}} \otimes \mathbf{A}\mathbf{U}\mathbf{A}^{\mathrm{H}}\right]\mathrm{vec}\left(\mathbf{C}^{\mathrm{H}}\right). \tag{5C.4}$$

Next, the vectorized coupling matrix can also be written as

$$\mathrm{vec}(\mathbf{C}) = \mathbf{M}\bar{\mathbf{c}}, \tag{5C.5}$$

where $\bar{\mathbf{c}}^T = [1\ \mathbf{c}^T]$ and $\mathbf{M}$ is a $n^2 \times n$ selection matrix

$$\mathbf{M} = \begin{bmatrix}\mathbf{M}_1\\ \mathbf{M}_2\\ \vdots\\ \mathbf{M}_n\end{bmatrix} \quad \text{where} \quad \mathbf{M}_k = \begin{cases}\mathbf{I} & k = 1\\ \begin{bmatrix}\tilde{\mathbf{I}}\ \mathbf{I}_{2:k,:}\\ \mathbf{I}_{1:n-k+1,:}\end{bmatrix} & k > 1\end{cases} \tag{5C.6}$$

and $\mathbf{I}_{k:m}$ denotes the k^{th} to m^{th} rows of the unity matrix and $\tilde{\mathbf{I}}$ is the reverse permutation operator defined in (5.51). Now, the criterion becomes

$$\bar{\mathbf{c}}^T\mathbf{M}^T\left[\hat{\mathbf{E}}_n^*\hat{\mathbf{E}}_n^T \otimes \mathbf{A}\mathbf{U}\mathbf{A}^H\right]\mathbf{M}\bar{\mathbf{c}}^*. \tag{5C.7}$$

However, since the criterion is real it can be conjugated

$$V_{NSF} = \bar{\mathbf{c}}^H\mathbf{G}\bar{\mathbf{c}}, \tag{5C.8}$$

where

$$\mathbf{G} = \mathbf{M}^H\mathbf{V}\mathbf{M} \tag{5C.9}$$

$$\mathbf{V} = \hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H \otimes \mathbf{A}^*\mathbf{U}^*\mathbf{A}^T. \tag{5C.10}$$

Note that since the norm in (5C.1) is nonnegative, the matrix $\mathbf{G}$ is a positive semidefinite matrix and due to Sylvester's law of inertia [Lüt96] so is the matrix $\mathbf{V}$. Furthermore, the dimensions of the matrices $\mathbf{G}$ and $\mathbf{V}$ are important. The matrix $\mathbf{G}$ is $q+1 \times q+1$, but the dimension of $\mathbf{V}$ is $n^2 \times n^2$ which can be large if the array is large. Thus, the above vectorization operation may not be feasible for large arrays.

However, it is possible to use Lemma 1 on page 103 to rewrite the criterion in (5C.1) to the same form as in (5C.8) with the corresponding $\mathbf{G}$ matrix defined as

$$\mathbf{G} = \sum_{k=1}^{p} \mathbf{Q}^H(\mathbf{A}\mathbf{U}_{k,:}^H)\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\mathbf{Q}(\mathbf{a}_k), \tag{5C.11}$$

where $\mathbf{U}_{k,:}$ denotes the k^{th} row of $\mathbf{U}$ and $\mathbf{a}_k$ denotes the k^{th} column of $\mathbf{A}$. This formulation avoids the large matrix $\mathbf{V}$ at the cost of introducing a sum of p terms. Depending on the scenario either (5C.11) or (5C.9) may be the best expression for evaluating the criterion. The Kronecker expression in (5C.9) is valid also for other linear parameterizations of the coupling matrix that not result in a banded complex symmetric Toeplitz matrix. However, Lemma 1 can also be generalized to other parameterizations, and which formulation to choose depends on the explicit parameterization and the size of the array.

Now, in order to concentrate the criterion regarding the coupling parameters $\mathbf{c}$ use the partitioned form of the matrix $\mathbf{G}$

$$\mathbf{G} = \begin{bmatrix} G_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{12}^H & \mathbf{G}_{22} \end{bmatrix} \tag{5C.12}$$

which follows since $\mathbf{G}$ is Hermitian and

$$\begin{aligned} V_{NSF} =& G_{11} + \mathbf{G}_{12}\mathbf{c} + \mathbf{c}^H\mathbf{G}_{12}^H + \mathbf{c}^H\mathbf{G}_{22}\mathbf{c} = \\ & G_{11} + \left(\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H + \mathbf{c}\right)^H \mathbf{G}_{22}\left(\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H + \mathbf{c}\right) - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H. \end{aligned} \tag{5C.13}$$

The coupling parameters that minimize the criterion become

$$\hat{\mathbf{c}} = -\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H. \tag{5C.14}$$

Inserting (5C.14) into (5C.13) gives the concentrated criterion function

$$V_{NSF}(\boldsymbol{\phi}) = G_{11} - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H. \tag{5C.15}$$

The criterion is concentrated with respect to the q coupling parameters and only a p dimensional search for the DOAs are needed, just as in the coupling free case. The DOA parameters enter the expression non-linearly, and a damped Newton method is used, similar to the SML method in Section 5.2.

But in the SML method, large simplifications were possible by using an asymptotic Hessian in the search process leading to the scoring method. In the NSF method, the FIM could be used instead of the true Hessian if the weighting matrix $\mathbf{U}$ is chosen as to minimize the variance of the estimates. That is possible since the NSF method achieves the CRB. Here, however, the true Hessian is used in a damped Newton method.

To implement the search, it is necessary to find expressions for the gradient and Hessian of the NSF criterion and that is done next.

The k^{th} element of the gradient becomes

$$\mathbf{V}'_k = \frac{\partial V_{NSF}}{\partial \phi_k} = \frac{\partial G_{11}}{\partial \phi_k} - 2\mathrm{Re}\left[\frac{\partial \mathbf{G}_{12}}{\partial \phi_k}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H\right] + \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\frac{\partial \mathbf{G}_{22}}{\partial \phi_k}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H. \tag{5C.16}$$

The expression for the Hessian becomes

$$\begin{aligned}\mathbf{H}_{km} = \frac{\partial^2 V_{NSF}}{\partial \phi_k \partial \phi_m} = \frac{\partial^2 G_{11}}{\partial \phi_k \partial \phi_m} + 2\mathrm{Re}\Bigg[&\frac{\partial \mathbf{G}_{12}}{\partial \phi_k}\mathbf{G}_{22}^{-1}\frac{\partial \mathbf{G}_{22}}{\partial \phi_m}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H + \\ &\frac{\partial \mathbf{G}_{12}}{\partial \phi_m}\mathbf{G}_{22}^{-1}\frac{\partial \mathbf{G}_{22}}{\partial \phi_k}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H - \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\frac{\partial \mathbf{G}_{22}}{\partial \phi_k}\mathbf{G}_{22}^{-1}\frac{\partial \mathbf{G}_{22}}{\partial \phi_m}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H - \\ &\frac{\partial^2 \mathbf{G}_{12}}{\partial \phi_k \partial \phi_m}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H - \frac{\partial \mathbf{G}_{12}}{\partial \phi_k}\mathbf{G}_{22}^{-1}\frac{\partial \mathbf{G}_{12}^H}{\partial \phi_k}\Bigg] + \mathbf{G}_{12}\mathbf{G}_{22}^{-1}\frac{\partial^2 \mathbf{G}_{22}}{\partial \phi_k \partial \phi_m}\mathbf{G}_{22}^{-1}\mathbf{G}_{12}^H.\end{aligned} \tag{5C.17}$$

The derivatives of $\mathbf{G}$ depends on which of the formulations (5C.9) or (5C.11) is used. First, the expressions for the Kronecker formulation in (5C.9) will be given and then the expressions for the summation form in (5C.11). Since neither the selection matrix nor the noise eigenvalues depend on the DOAs, the derivative becomes

$$\frac{\partial \mathbf{G}}{\partial \phi_k} = \mathbf{M}^H\left[\left(\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\right) \otimes \left(\frac{\partial \mathbf{A}^*\mathbf{U}^*\mathbf{A}^T}{\partial \phi_k}\right)\right]\mathbf{M}, \tag{5C.18}$$

where

$$\frac{\partial \mathbf{A}^*\mathbf{U}^*\mathbf{A}^T}{\partial \phi_k} = \left(\frac{\partial \mathbf{a}_k^*}{\partial \phi_k}\mathbf{U}_{k,:}^*\mathbf{A}^T\right) + \left(\frac{\partial \mathbf{a}_k^*}{\partial \phi_k}\mathbf{U}_{k,:}^*\mathbf{A}^T\right)^H. \tag{5C.19}$$

Furthermore,

$$\frac{\partial^2 \mathbf{G}}{\partial \phi_k \partial \phi_m} = \mathbf{M}^H\left[\left(\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\right) \otimes \left(\frac{\partial^2 \mathbf{A}^*\mathbf{U}^*\mathbf{A}^T}{\partial \phi_k \partial \phi_m}\right)\right]\mathbf{M}, \tag{5C.20}$$

where

$$\frac{\partial^2 \mathbf{A}^*\mathbf{U}^*\mathbf{A}^T}{\partial \phi_k \partial \phi_m} = \left(\mathbf{U}_{km}^*\frac{\partial \mathbf{a}_k^*}{\partial \phi_k}\frac{\partial \mathbf{a}_m^T}{\partial \phi_m}\right) + \left(\mathbf{U}_{km}^*\frac{\partial \mathbf{a}_k^*}{\partial \phi_k}\frac{\partial \mathbf{a}_m^T}{\partial \phi_m}\right)^H \tag{5C.21}$$

if $k \neq m$ and

$$\left(\frac{\partial^2 \mathbf{a}_k^*}{\partial \phi_k^2}\mathbf{U}_{k,:}^* \mathbf{A}^T + \mathbf{U}_{kk}^* \frac{\partial \mathbf{a}_k^*}{\partial \phi_k}\frac{\partial \mathbf{a}_k^T}{\partial \phi_k}\right) + \left(\frac{\partial^2 \mathbf{a}_k^*}{\partial \phi_k^2}\mathbf{U}_{k,:}^* \mathbf{A}^T + \mathbf{U}_{kk}^* \frac{\partial \mathbf{a}_k^*}{\partial \phi_k}\frac{\partial \mathbf{a}_k^T}{\partial \phi_k}\right)^H \quad \text{(5C.22)}$$

for $k = m$. For the summation form of $\mathbf{G}$ in (5C.11), the derivative becomes

$$\frac{\partial \mathbf{G}}{\partial \phi_k} = \sum_{l=1}^{p} \mathbf{U}_{lk}^* \mathbf{Q}^H(\frac{\partial \mathbf{a}_k}{\partial \phi_k})\hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Q}(\mathbf{a}_l) + \mathbf{Q}^H(\mathbf{A}\mathbf{U}_{k,:}^H)\hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Q}(\frac{\partial \mathbf{a}_k}{\partial \phi_k}). \quad \text{(5C.23)}$$

The second derivatives become

$$\frac{\partial^2 \mathbf{G}}{\partial \phi_k \partial \phi_m} = \mathbf{U}_{mk}^* \mathbf{Q}^H(\frac{\partial \mathbf{a}_k}{\partial \phi_k})\hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Q}(\frac{\partial \mathbf{a}_m}{\partial \phi_m}) + \mathbf{U}_{km}^* \mathbf{Q}^H(\frac{\partial \mathbf{a}_m}{\partial \phi_m})\hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Q}(\frac{\partial \mathbf{a}_k}{\partial \phi_k}) \quad \text{(5C.24)}$$

if $k \neq m$ and

$$\begin{aligned}\sum_{l=1}^{p} \mathbf{U}_{lk}^* \mathbf{Q}^H(\frac{\partial^2 \mathbf{a}_k}{\partial \phi_k^2})\hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Q}(\mathbf{a}_l) + 2\mathbf{U}_{kk}^* \mathbf{Q}^H(\frac{\partial \mathbf{a}_k}{\partial \phi_k})\hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Q}(\frac{\partial \mathbf{a}_k}{\partial \phi_k}) + \\ \mathbf{Q}^H(\mathbf{A}\mathbf{U}_{k,:}^H)\hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{Q}(\frac{\partial^2 \mathbf{a}_k}{\partial \phi_k^2})\end{aligned} \quad \text{(5C.25)}$$

if $k = m$. Finally, inserting the expressions for the derivatives of $\mathbf{G}$ into (5C.16) and (5C.17) gives the gradient and the Hessian.

Chapter 6

Effects of Model Errors

In the preceding chapter, the effects of an unknown coupling were mitigated by estimating the coupling along with the DOAs. To reduce the number of estimated coupling parameters, the reduced coupling model derived in Chapter 4, was used. For this coupling model, the CRB was calculated and several estimation methods were derived. The quality of these estimation methods was then addressed by comparing the empirical variance of the estimates in a few Monte-Carlo simulations with the CRB. However, the data used in those comparisons were generated using the reduced coupling model, in order for the comparison to the CRB to be valid. Usually when examining different estimation schemes the correct data model is used (as in Chapter 5), since when designing estimation schemes the focus is on finding the best way of utilizing the model at hand. But for the methods to work well in practice, it is of course important to examine how well the model actually describes the measured data. This will be examined in this chapter, but instead of measured data, the expressions for the induced voltages, derived in Section 2.2 using fundamental electromagnetics, will be used. The DOAs and the coupling parameters will then be estimated using the reduced coupling model introduced in Chapter 4 and the estimation schemes from Chapter 5.

How the estimation of both the coupling and the DOAs work is examined in a few scenarios. Several methods are used to estimate the parameters and the result is presented using the MUSIC spectra to gain some insight of the problems. Then, the average RMSE error is calculated for a few cases using the different estimation methods. Finally, since the effect of coupling essentially changes the response of the array and in some way blurs the response, it is interesting to examine the effects on the signal estimates when a jammer is present.

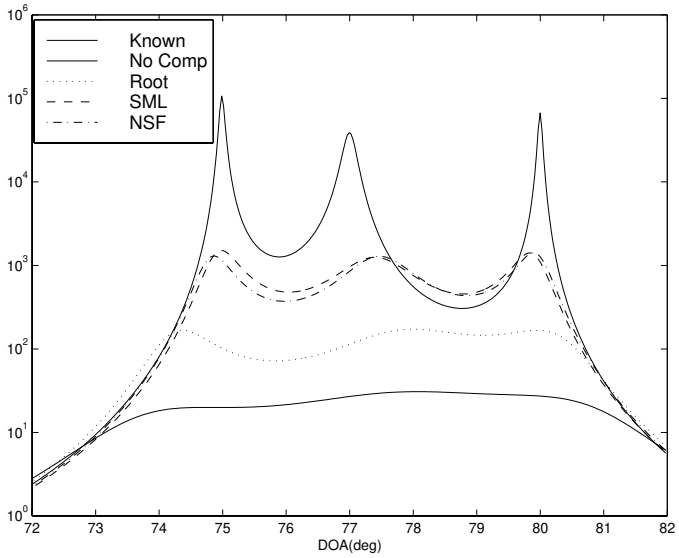


Figure 6.1: The MUSIC spectrum calculated from the estimates of the Root-Music (with known and estimated $\mathbf{C}$), SML, and NSF with 8 coupling parameters using an array of 15 $\lambda/2$ dipoles spaced $\lambda/2$ apart. Three waves are incident from $(75°, 77°, 80°)$ and 1000 snapshots with SNR = 30dB are used.

6.1 Impact of Model Errors

How the joint estimation of both the coupling and DOAs works on data generated using the electromagnetic model and how large impact the model errors have on the DOA estimation is examined in this section. The focus is on the effects on the DOA estimation, since usually the coupling parameters are only nuisance parameters. The effects will be examined by calculating the estimates using the Root-Music, SML, and the NSF methods. Also, the DOA estimates are calculated without any compensation for the coupling using the Root-Music method. The Root-Music method is used instead of the iterative MUSIC, due to the the lower resolution threshold of the latter method. The results of the calculations are illustrated by displaying the MUSIC-spectra using the estimates from the different methods.

First, consider a 15 element half-wavelength dipole array with an element separation of $\lambda/2$. Three plane waves are incident upon the array and the DOA:s $(75°, 77°, 80°)$ are to be estimated. The signal to noise ratio (SNR) is 30dB and $N = 1000$ snapshots are collected. If the coupling is left uncompensated, the MUSIC spectrum in Figure 6.1 fails to resolve the sources, and therefore the iterative MUSIC method can not be used. Here, the Root-

Music method, SML method, and the NSF method are used with 8 coupling parameters, and the resulting estimates are illustrated by calculating the MUSIC spectrum using the estimated coupling. When the coupling is considered unknown and estimated along with the DOAs, peaks in the spectra appear for the different methods.

Estimating the coupling along with the DOAs results in a larger model that better describes the data, and therefore the peaks become sharper. But the estimation methods are affected differently from the fact that the reduced coupling model only partly captures the properties of the full electromagnetic model. The Root-Music method just barely resolves three sources. The SML and NSF methods produce similar spectra and three sharper peaks appear. Remember that NSF is an asymptotic version of the SML method. But the peaks do not appear at the true DOAs, due to the modeling error that results from using the simpler coupling model. How large the error in the DOA estimates becomes depends on the angles.

Furthermore, since the different methods are affected in different ways by the model error, any general conclusions are hard to find. Also included in the figure is the MUSIC spectrum for the case of a known coupling, and in that case three sharp peaks appear. In that case, the only difference to the usual MUSIC method is that the steering matrix has changed into an effective steering matrix, as discussed in Section 3.3.2. The spectrum is essentially the same as in the hypothetical coupling free case.

In the next example, two waves are incident from $(80°, 84°)$ upon a smaller array of 8 elements and 4 coupling parameters are used, see Figure 6.2. The uncompensated MUSIC method almost fails to resolve the sources, whereas MUSIC with known coupling gives sharp peaks. When the coupling is estimated, the Root-Music method gives two peaks, but rather far from the true DOAs. Slightly higher peaks are obtained using the SML method, that also appear closer to the true DOAs. The best method in this example is the NSF method, that gives the sharpest peaks closest to the true DOAs. Here, the difference between the NSF and SML methods indicate that a large portion of the error is due to the model error, since otherwise these methods would give almost identical estimates. The effect of the model error is heavily dependent on the incidence angles, thus making comparisons for specific angles of less interest. Therefore, in the next section, an average measure of the model error will be introduced.

Another parameter that affects the estimation performance is how many coupling parameters that are estimated. If too few parameters are used, the model does not capture the properties of the coupling. If too many parameters are used, the variance is increased as analyzed in Section 5.3. In Figure 6.3, both the DOAs and the coupling parameters are estimated using

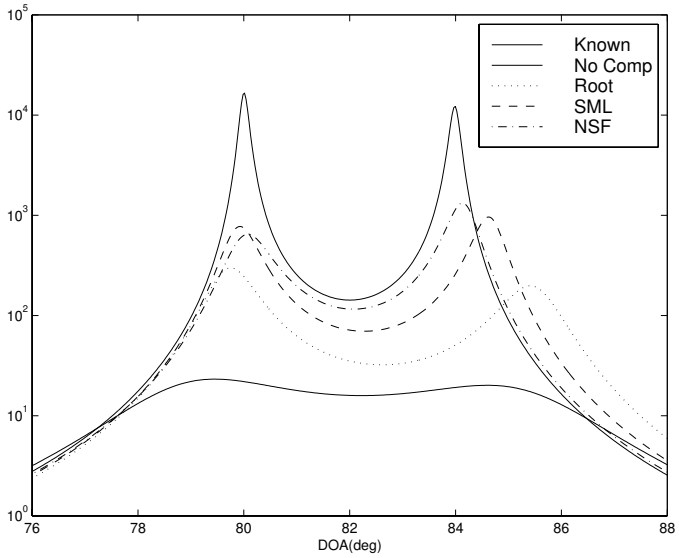


Figure 6.2: The MUSIC spectrum calculated from the estimates of the Root-Music (with known and estimated $\mathbf{C}$), SML, and NSF with 4 coupling parameters using an array of 8 $\lambda/2$ dipoles spaced $\lambda/2$ apart when two waves are incident from $(80°, 84°)$, and 1000 snapshots with SNR = 20dB.

the SML method with an array of 10 $\lambda/2$ dipoles spaced $\lambda/2$ apart when two waves are incident from $(80°, 85°)$. The RMSE for $\phi = 80°$ is calculated for different SNRs and numbers of coupling parameters. The number of snapshots used was 1000, and the initialization values in the search were taken as the estimates obtained from the Root-Music method after 5 iterations.

Increasing the SNR does not help if the coupling is left uncompensated, since the dominating error source in that case is the model error. However, already when using only one single coupling parameter, the situation improves significantly. The RMSE decreases when the SNR is increased, indicating that the model works better now. The best performance is obtained using two or three coupling parameters depending on the SNR value. Thus, the performance can be increased by estimating only a few coupling parameters indicating that coupling can be accounted for.

Now, the MUSIC spectrum for several different cases and also the influence of the number of coupling parameters on the estimation performance have been examined. The model errors were found to dominate the error in the DOA estimation in many cases, indicating that if the full coupling matrix can be found by a calibration scheme it should be used. However, that is not possible in all cases and by using a few coupling parameters the

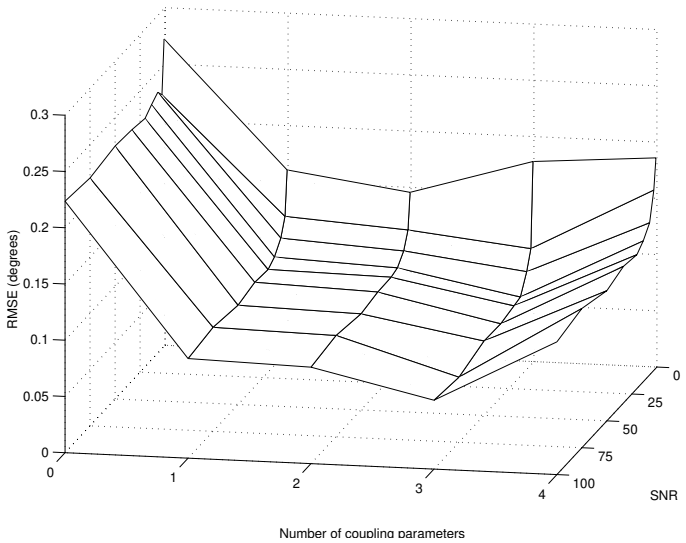


Figure 6.3: The RMSE for $\phi = 80$ when estimating both DOAs and coupling parameters using an array of $10\lambda/2$ dipoles spaced $\lambda/2$ apart when two waves are incident from $(80°, 85°)$. The RMSE is calculated for different SNR and number of coupling parameters where 1000 snapshots were used.

estimation performance can be improved. Furthermore, the effects of the model error was highly dependent on the DOAs, and also the different methods worked quite differently for different incidence angles, and any general conclusions were hard to find. Therefore, in the next section an average error is calculated that makes more general conclusions possible.

6.2 Average Error

The impact of the model errors is heavily dependent on the incidence angles, and therefore it is hard to draw any general conclusions. To obtain more knowledge of how well the reduced model and estimation methods work, an extensive simulation study or even better a large scale measurement study is needed. But since both of these options are time- and resource-consuming, an averaging operation is used to get some overall performance measure, although without any claims of being exhaustive. Since the effects of coupling are heavily dependent on the incidence angles, a natural measure is to average the RMSE over all possible incidence angles. Assuming that the DOAs ($\boldsymbol{\phi}$) are uniformly distributed in the area of interest, the average RMSE can be

defined as

$$\tilde{R} = \sum_{k=1}^{p} E\ [\text{RMSE}(\phi_k)]. \tag{6.1}$$

However, calculating the expectation is difficult, and therefore an approximative measure can be defined as

$$\bar{R} = \sum_{k=1}^{p} \frac{1}{N_{av}} \sum_{n=1}^{N_{av}} \text{RMSE}_n(\phi_k), \tag{6.2}$$

where the RMSE value is averaged over N_{av} Monte-Carlo simulations using DOAs that are uniformly distributed over the area of interest. Using this average RMSE value, the effect of the modeling errors can be evaluated and the different algorithms compared. Such a comparison, of course, only tells part of the story, but at least it is better than calculating the estimates for specific scenarios.

First, a 5 element half-wavelength dipole array with element separation of $\lambda/2$ will be considered, since the coupling effects are larger for a small array. Two plane waves are incident upon the array and the DOAs of these are assumed to be uniformly distributed in the interval $20° - 90°$. Waves incident closer to end-fire are excluded, since the estimation performance there is severely reduced. The signal to noise ratio (SNR) is chosen large (50dB), in order to let the part of the RMSE due to the model errors dominate over the variance part. Also, $N = 1000$ snapshots are collected and $N_{av} = 5000$ Monte-Carlo simulations are averaged to calculate the average RMSE value $\bar{R}$ in Figure 6.4.

To avoid that a single realization dominates the RMSE value, i.e. so-called outliers, estimates that are more than $10°$ off are excluded from the averaging operation. The uncompensated Root-Music method results in an average RMSE value close to $4°$. However, by modeling the coupling and estimating only one coupling parameter the average RMSE can be significantly reduced. The averaging operation gives only a very rough picture of the effects of model errors and the different estimation methods behave similarly in this case. To examine the quality of the different estimation methods a more detailed analysis is needed.

Adding more coupling parameters does not further reduce the average RMSE error significantly. The RMSE error is still quite large, thus the reduced coupling model considered here only captures some of the coupling properties. But that the RMSE decreases significantly when only estimating one coupling parameter, indicates that the model is reasonably good.

For larger arrays, more coupling parameters can be estimated and next, a larger array of 15 elements is considered. The average RMSE is calculated by

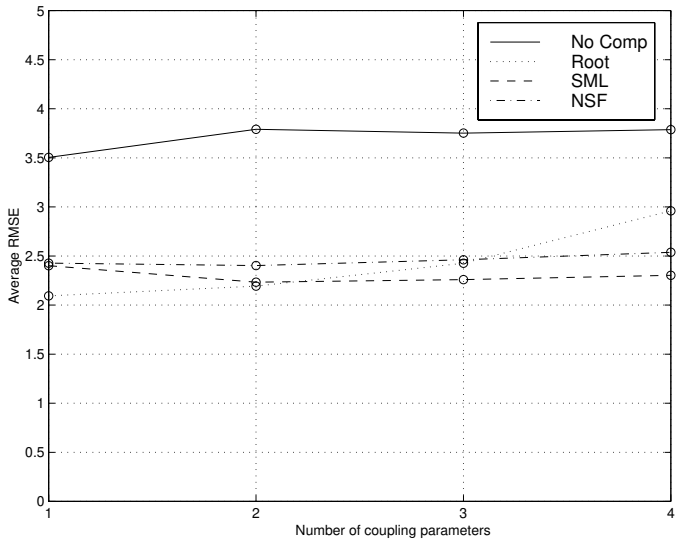


Figure 6.4: The average RMSE $\bar{R}$ for the Root-Music (with uncompensated and estimated $\mathbf{C}$), SML, and NSF methods versus the number of coupling parameters. An array of 5 $\lambda/2$ dipoles spaced $\lambda/2$ apart is used. Two waves are incident from uniformly distributed angles in $(20°, 90°)$. The number of snapshots is $N = 1000$, SNR = 50dB, and $N_{av} = 5000$.

averaging the results from 5000 simulations. Here, the average RMSE error is smaller due to the larger array, and therefore estimates that deviates more than 3° are considered to be outliers and are excluded from the averaging operation. Another implication of a larger array is that it allows for more parameters to be estimated with just a small increase in variance, as discussed in Section 5.3. The RMSE value when estimating 5 coupling parameters is decreased to less than half the value when no compensation is applied. The different estimation methods give similar results, also in this case.

The contribution to the average RMSE from model errors is quite large and for larger SNR values, the model error part will dominate. In Figure 6.6, the average RMSE is shown for an array of 8 elements for different SNRs. For small SNRs, the variance of the estimates dominate $\bar{R}$ and estimating coupling parameters further increases the RMSE. For larger SNRs, the model error part dominate, and here the RMSE can be decreased by estimating a few coupling parameters.

The RMSE values in these simulations are relatively large indicating that better models than the reduced coupling model can be found. Of course, the full coupling matrix should be estimated if possible, but if the DOAs are

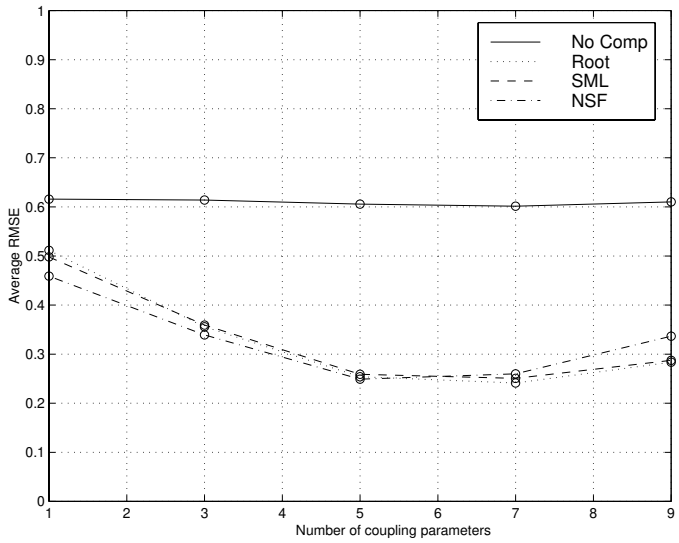


Figure 6.5: The average RMSE $\bar{R}$ for the Root-Music (with uncompensated and estimated $\mathbf{C}$), SML, and NSF methods versus the number of coupling parameters. An array of 15 $\lambda/2$ dipoles spaced $\lambda/2$ apart is used. Two waves are incident from uniformly distributed angles in $(20°, 90°)$. The number of snapshots is $N = 1000$, SNR $= 50$dB, and $N_{av} = 5000$.

unknown only a few coupling parameters can be estimated. In that case, only part of the properties of the coupling can be obtained. But the estimation problem considered here with no knowledge about the coupling is probably not a realistic one. In reality some knowledge about the coupling usually exists and should be incorporated and some deviation from a nominal model could be estimated using the reduced coupling model considered here. Also, the mutual coupling between dipoles is strong and other antenna elements usually experience less coupling, which probably makes it easier to model the coupling in that case. Therefore, other models that better describe the coupling, can probably be found and if the coupling parameters enter the coupling matrix linearly, the methods considered here are easily modified to this model. For example, in the NSF method this corresponds to simply changing the $\mathbf{M}$ matrix.

Now, the effects of the model errors introduced by using the reduced model for estimating the coupling on the average RMSE value of the DOAs have been studied. Next, the effects on signal estimation in the presence of jammers is examined.

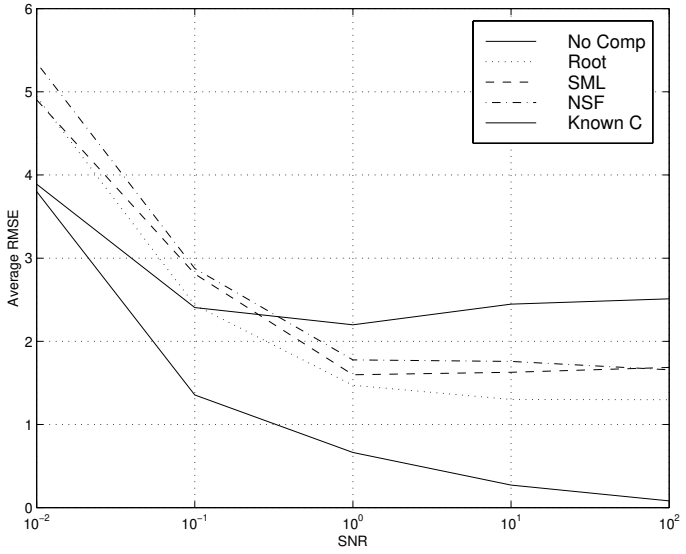


Figure 6.6: The average RMSE $\bar{R}$ for the Root-Music (with uncompensated and estimated $\mathbf{C}$), SML, and NSF methods versus SNR using an array of 8 $\lambda/2$ dipoles spaced $\lambda/2$ apart when two waves are incident from uniformly distributed angles in $(20°, 90°)$. The number of snapshots is $N = 500$, $q = 3$, and $N_{av} = 5000$.

6.3 Interference Rejection

One problem when estimating the signals that are incident on the array is the interference from other signals that also induce voltages on the antenna elements. In particular, if the interfering signals are of significantly higher power, the estimation problem becomes difficult. Interfering signals of large power are usually called jammers and a rich literature on the subject of interference suppression exists [Far92, He93, VB88]. Here, the influence of an unknown coupling will be examined in a very simple manner. Only one signal is transmitted and there is only one jammer present. First, the DOAs of these waves are estimated and then the signal values of the interesting signal is estimated as

$$\hat{\mathbf{s}} = (\mathbf{C}(\hat{\mathbf{c}})\mathbf{A}(\hat{\phi}))^{\dagger}\mathbf{X}, \tag{6.3}$$

where $\mathbf{X}$ is the measured data and $(\hat{\mathbf{c}}, \hat{\boldsymbol{\phi}})$ are the estimated parameters. This is of course a very simple method to obtain an estimate of the signal, but here the effects of mutual coupling will be examined using this method. In [Adv97, AS96], a MOM solution of the antenna is used to account for the mutual coupling by considering also the reflections from the open circuited

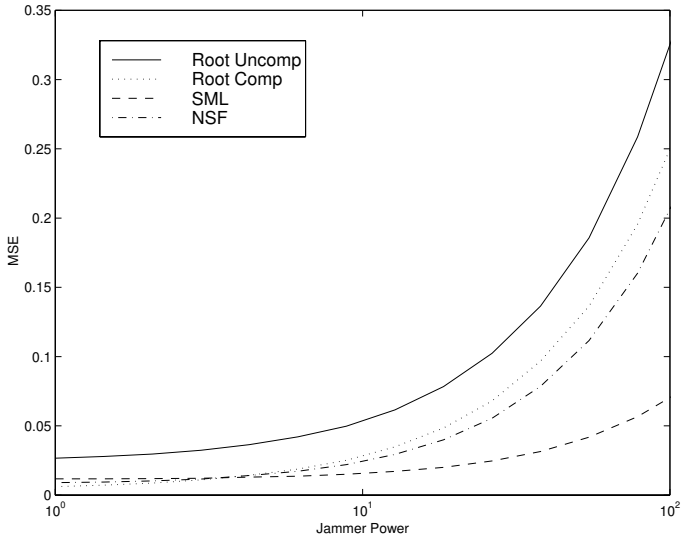


Figure 6.7: The MSE of the estimated signal versus jammer power using the Root-Music (with compensated and uncompensated $\mathbf{C}$), SML, and NSF with 3 coupling parameters using an array of 8 $\lambda/2$ dipoles spaced $\lambda/2$ apart. The jammer signal is located 10° off the interesting signal. The result is averaged over the incidence angle of the interesting signal in the interval $(50°, 90°)$.

neighboring elements and thus obtaining close to complete jammer cancellation. Using the simple method suggested here, the cancellation will be far from complete, but on the other hand no explicit knowledge of the actual antenna setup and no MOM solution is necessary. Also, this simple method will work on most antenna arrays, but of course with various result. Furthermore, if the NSF method is used, the numerical search will be of the same dimension as if the coupling was known.

Consider an array of 8 $\lambda/2$ dipoles spaced $\lambda/2$ apart and the DOAs and 3 coupling parameters are estimated with the Root-Music, SML, and NSF methods respectively. The noise power is -30dB and 1000 snapshots are taken. The jammer signal is 10° off the interesting signal in this simulation. However, since the effects of mutual coupling are highly dependent on the DOA, the result is averaged over the incidence angle of the interesting signal in the interval $(50°, 90°)$. In Figure 6.7, the Mean Square Error (MSE) of the estimated signal using the different methods is shown.

If the coupling is left uncompensated, the signal estimate deviates significantly already for equal signal powers, and increases rapidly with increasing jammer power. The Root-Music method reduces the error in signal power

for low jammer power, but for high jammer power the reduction is small. The NSF method offers more error reduction or interference suppression for higher jammer power. The best method is the SML method, that offers the best interference rejection especially for high jammer power. Thus, the interference rejection can be increased significantly by estimating a few coupling parameters along with the DOAs. For better performance, a more complex model is needed. However, as discussed in previous section usually some knowledge about the coupling exists and should be used to formulate another coupling model.

6.4 Conclusions

If it is not possible to obtain the coupling matrix by calibration measurements, an unknown uncompensated coupling will deteriorate the DOA estimation performance. However, the effects of the unknown coupling can be mitigated by estimating the unknown coupling and that was examined in this chapter. Data was generated using the full electromagnetical coupling model, and the coupling was estimated using the reduced coupling model. That introduces model errors and the impact of these errors was examined.

First, the effects of an unknown coupling was investigated by estimating the DOAs, and the coupling parameters using the methods derived in Chapter 5. It was found that the effects were angle dependent, but the estimation performance could be improved by estimating only a few coupling parameters.

To get some general result of the improvement by estimating the coupling, an average RMSE value was calculated and it was found that the model error could be substantially reduced by estimating only a few coupling parameters. But for high SNRs, the model error dominated the RMSE value. The value was reduced compared to the uncompensated case, although much larger than the RMSE value for the known coupling case. But, if it is not possible to obtain the coupling matrix by some calibration scheme, the effects can thus be reduced by estimating a few coupling parameters.

Finally, the effects of an unknown mutual coupling were examined in a scenario where an interesting signal was received in the presence of a jammer signal. If coupling was not compensated for, the estimated signal deviated significantly from the true signal. However, by estimating the coupling along with the DOAs, the deviation was decreased substantially.

Chapter 7

Multimode Based Direction Finding

Direction finding by exploiting higher order modes of a modified biconical antenna is considered. The DOA of an incident wave is estimated by using the fact that the different modes have different directional properties. Since only one antenna element is used, several problems encountered when using an array of elements are avoided such as problems with mutual coupling between antenna elements. The direction finding performance is examined by calculating the CRB and the ML estimator for several different cases. It is found that an antenna employing multiple modes, i.e. an multimode antenna, offers direction finding performance comparable to the traditional antenna array using only a single physical antenna element.

7.1 Introduction

Direction finding is of great importance in a variety of applications, such as radar, sonar, communications, and recently also personal locating services. In the last two decades, direction finding and sensor array processing has attracted considerable interest in the signal processing community. The focus of this work has been on high resolution DOA estimation algorithms [KV96]. These algorithms exploit the fact that an electromagnetic wave that is received by an array of antenna elements reaches each element at different time instants. A different approach was presented in [NP94], where the DOA is estimated from the cross-product of the electric and magnetic field that are measured using an EM Vector Sensor (VS). Of these, the antenna array has received by far the most attention in the signal processing literature.

Recently, a modified biconical antenna was introduced in [DW98]. It offers interesting possibilities of forming directional beams by exiting higher

order modes. The biconical antenna is known as an easy to build broadband antenna with omnidirectional radiation. If higher order modes are excited, directional radiation patterns can be achieved [DW98]. By exciting the different modes separately, a number of antenna ports are obtained while only using a single antenna element. In a sense, the biconical antenna offers characteristics similar to an antenna array through multiple modes, without requiring many expensive antenna elements. The possibilities of forming directional beams, using a multimode antenna, was considered in [DW98], but no direction finding analysis was undertaken. Another application of a biconical antenna can be found in [ZHM+00], where several directional beams are used when studying angular properties at both transmitter and receiver, i.e. the Direction Of Departure (DOD) and DOA, in an indoor scenario. The direction finding performance of the biconical antenna and its potential use in high resolution direction finding will be analyzed in this paper.

7.2 Biconical Antenna

Many different types of antenna elements support excitation of higher modes. Here, a biconical antenna [DW98] will be analyzed. An example of how a multimode microstrip antenna can be used as a diversity receiver, is given in Chapter 10. A biconical antenna consists of two conical horns facing opposite directions as illustrated in Figure 7.1, where α denotes the flare angle and l the length of the horn. Note that the flare angle of the upper and lower cones are identical, i.e. the antenna is symmetric.

The concept of exciting higher order modes is rather unknown in the signal processing society, but is well known by antenna designers. Essentially, different modes in electromagnetics represent different solutions to Maxwell's equations that fulfill the boundary conditions for the geometry at hand. In fact, it is possible to separately excite several modes at the same frequency on one antenna structure and regard these as different antenna ports. A multimode antenna is thus an antenna where several modes are excited separately on the same antenna structure.

In the traditional analysis of the biconical antenna, a coaxial waveguide is used to to excite the antenna. By enlarging the diameter of the coaxial feed, more solutions or modes are possible. The enlargement of the coaxial waveguide allows for higher order circular modes in the feed. These circular waves are then transformed into spherical waves that propagate on the biconical structure that in turn radiates a field. It is important to note that it is possible to design the overmoded coaxial feed in such a way that no modal coupling between modes occurs [DW98].

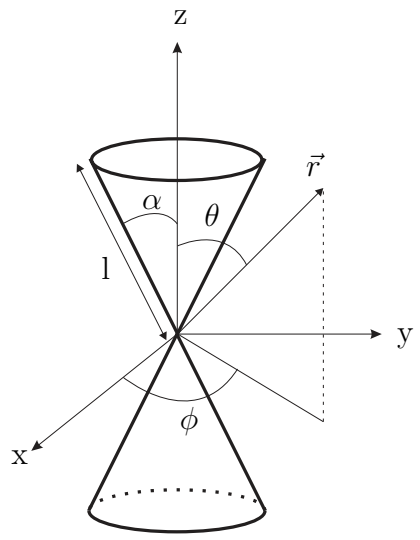


Figure 7.1: The geometry of the biconical antenna and the relation to the coordinates.

In principle, the far-field is derived by integrating the fields from an infinite biconical transmission line over the biconical aperture [Bal89, Sva00h]. The lowest order mode, the dominant or fundamental mode, is a so-called Transverse ElectroMagnetic (TEM) mode, where both the electric and magnetic field are perpendicular to the propagation direction. The higher order modes result in solutions of Transverse Electric (TE) type, i.e. where only the electric field is perpendicular to the propagation direction. Note that these modes are simply solutions to Maxwell's equations that fulfill the boundary conditions of the biconical structure for specific values of l and α. The far-field expression for the TEM mode can be written as

$$\vec{E}_0 = \vec{\theta}\,\frac{e^{-jkr}}{r} G_0(\theta, l, \alpha), \tag{7.1}$$

where k denotes the wave number[1], r the radial distance from the antenna, and $G_0(\theta, l, \alpha)$ denotes the element pattern. The expression for the element pattern is quite complicated, see Appendix 10A and [DW98, Sva00h].

The corresponding expression of the far-field of the TE_{mn} mode can also be written as in (7.1). Here, the first index refers to the azimuthal mode order and the second to the elevation order [Bal89, DW98]. There is, however, one

[1]The wave number $k = \frac{2\pi}{\lambda} = \omega\sqrt{\mu\epsilon}$, where λ is the wavelength. The symbols ϵ and μ denote the dielectric constant and the permeability of the medium respectively.

important difference to the TEM case. While the TEM field actually is independent of ϕ, the TE field does have an azimuth dependency

$$G_m(\theta, \phi, l, \alpha) = \cos(m\phi)\ G'_m(\theta, l, \alpha) \quad m \geq 1 \quad , \tag{7.2}$$

where $G'_m(\theta, \phi, l, \alpha)$ is the remaining azimuth independent term of the radiation pattern, see Appendix 10A and [DW98, Sva00h].

Since the radiation intensity pattern provides some insight into the properties of the biconical antenna with higher order modes, it will be calculated next. The radiation intensity $U(\theta, \phi)$ can be written as

$$U(\theta, \phi) = \frac{r^2}{2\eta}\left|\vec{E}_{tot}\right|^2, \tag{7.3}$$

where η is the intrinsic impedance of the propagation medium and the total field radiated from the biconical antenna $\vec{E}_{tot}$ is the sum of the modal fields as

$$\vec{E}_{tot} = \sum_{m=1}^{M} A_m \vec{E}_m(\theta, \phi), \tag{7.4}$$

where A_m is the amplitude of mode m. Only the azimuth radiation pattern will be considered here. Also, to reduce the notational complexity the patterns are normalized to $G'_m(\theta, l, \alpha) = 1$. The radiation pattern of the multimode biconical antenna can then be written as

$$U(90^\circ, \phi) = \frac{1}{2\eta}\left|\sum_{m=1}^{M} A_m \cos(m\phi)\right|^2. \tag{7.5}$$

The radiation pattern using $A_m = 1\ \forall\ m = 1, \ldots, M$ for a biconical antenna with $M = 4$ modes is shown in Figure 7.2. The width of the main lobe is actually smaller than that of the antenna array for this case when $\phi = 0^\circ$ (end-fire). However, if the beam is steered towards broadside ($\phi = 90^\circ$) the opposite result is obtained. To steer the beam towards angle ϕ_0, the weights should be chosen as $A_m = \cos m\phi_0$, which is related to the Discrete Cosine Transform (DCT) in the same manner as the uniformly spaced linear array is related to the Discrete Fourier Transform (DFT).

In order to analyze the direction finding properties, an expression for the antenna in receive mode is needed. Unfortunately, no explicit expressions are included in antenna textbooks, since antennas are usually analyzed in transmit mode and not in receive mode. However, expressions for the received voltages are easily derived [Mey00, Kil99]. The analysis is based on reciprocity and knowledge of the far-field radiation function of the antenna

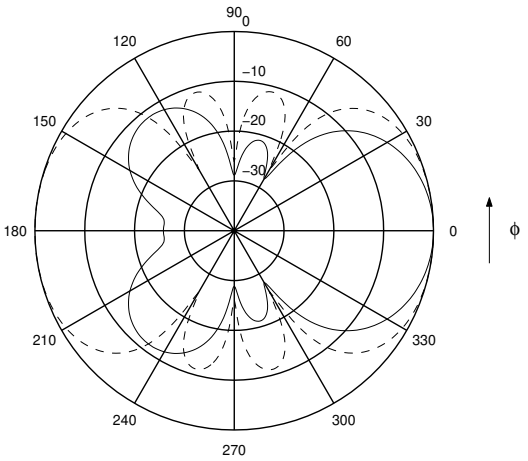


Figure 7.2: The radiation pattern using $A_m = 1 \ \forall \ m = 0, \ldots, M$ for a biconical antenna with $M = 4$ modes (solid) and the corresponding pattern of a four element array with $\lambda/2$ element separation (dashed).

$\vec{G} = \vec{g}G(\theta, \phi)$, where $\vec{g}$ is the orientation of the far-field. Using the results from [Kil99], the induced voltage becomes

$$v = -j\frac{4\pi}{k\eta}\left(\frac{1}{I_{in}}\vec{G} \cdot \vec{E}_i\right), \tag{7.6}$$

where k is the wave number, I_{in} the input current used when calculating $\vec{G}$, and $\vec{E}_i$ denotes the incident field. In the following, only waves incident in the $x-y$ plane, i.e. $(\theta = 90°)$ will be considered. Assuming an incident field from the ϕ direction of strength E_s and polarization $\vec{\rho}$

$$\vec{E}_i = \vec{\rho}E_s, \tag{7.7}$$

the received voltage of mode m can be written, using (7.1), (7.2), and (7.6) as

$$x_m(t) = \cos(m\phi)E_s\vec{\rho} \cdot \vec{\theta}, \tag{7.8}$$

if the normalization $G'_m(\theta, l, \alpha) = 1$ is used. It is further assumed that no polarization losses occur, i.e. $\vec{\rho} \cdot \vec{\theta} = 1$. If p waves are incident upon an antenna with M modes, the received voltages can be written in matrix form as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\phi})\mathbf{s}(t) + \mathbf{n}(t), \tag{7.9}$$

where the vector of measured voltages $\mathbf{x}(t)$ is $M \times 1$. The matrix $\mathbf{A}(\boldsymbol{\phi})$ $(M \times p)$ corresponds to the response of the different modes and has elements

$[\mathbf{A}(\boldsymbol{\phi})]_{pq} = \cos p\phi_q$. This matrix is typically called steering matrix in the sensor array processing literature. The signal vector $\mathbf{s}(t)$ is $p \times 1$ and contains the strength of the received fields. Finally, the noise vector $\mathbf{n}(t)$ is $M \times 1$.

7.3 Direction Finding Performance

The data model (7.9) is identical to the usual data model used in sensor array processing [KV96], except for a new steering matrix. This will of course change the direction finding properties. Before the properties of a specific DOA estimation scheme is studied, a lower bound, the Cramér-Rao lower Bound (CRB), on the variance of the DOA estimates will be analyzed. Note that it is possible to asymptotically achieve this bound with methods that are described in the next section. Expressions for the CRB was derived for an array of antenna elements in [OWVK89]; and can also be applied to the multimode antenna by changing the steering matrix.

$$E\{(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0)(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0)^T\} \geq \mathbf{B} \tag{7.10}$$

$$\mathbf{B} = \frac{\sigma^2}{2N}\left[Re\{(\mathbf{D}^H\mathbf{P}_A^{\perp}\mathbf{D}) \odot (\mathbf{S}\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A}\mathbf{S})^T\}\right]^{-1} ,$$

where

$$\mathbf{D} = \left[\frac{\partial \mathbf{a}(\phi)}{\partial \phi}\Big|_{\phi=\phi_1}, \cdots, \frac{\partial \mathbf{a}(\phi)}{\partial \phi}\Big|_{\phi=\phi_p}\right] \tag{7.11}$$

Here, $\odot$ denotes the Hadamard (or Schur) product, i.e., element-wise multiplication and $\mathbf{P}_A^{\perp} = \mathbf{I} - \mathbf{P}_A = \mathbf{I} - \mathbf{A}\mathbf{A}^{\dagger}$ [2] is the orthogonal projector onto the null space of $(\mathbf{A})^H$. The matrix $\mathbf{R} = \mathbf{A}\mathbf{S}\mathbf{A}^H + \sigma^2\mathbf{I}$ is the covariance matrix of the measured voltages $\mathbf{x}(t)$ and N denotes the number of time samples.

The square root of the CRB, i.e. the standard deviation, is shown in Figure 7.3 for different antenna configurations as two waves are incident from $(30°, 30° + \Delta)$. Only the CRB for the first DOA, i.e. the wave arriving from 30°, is shown since the CRB for the second DOA will behave similarly. Therefore, in the following CRB comparisons only the first DOA will be considered. The standard deviation for a uniform linear array of three elements spaced $\lambda/2$ apart is compared to a biconical antenna employing three to five modes. As expected, the performance increases with the number of modes. Also, note that the three element array behaves similar to the biconical antenna employing four modes. However, these results depend on the

[2] $\mathbf{M}^{\dagger}$ is the Moore-Penrose pseudo inverse of $\mathbf{M}$.

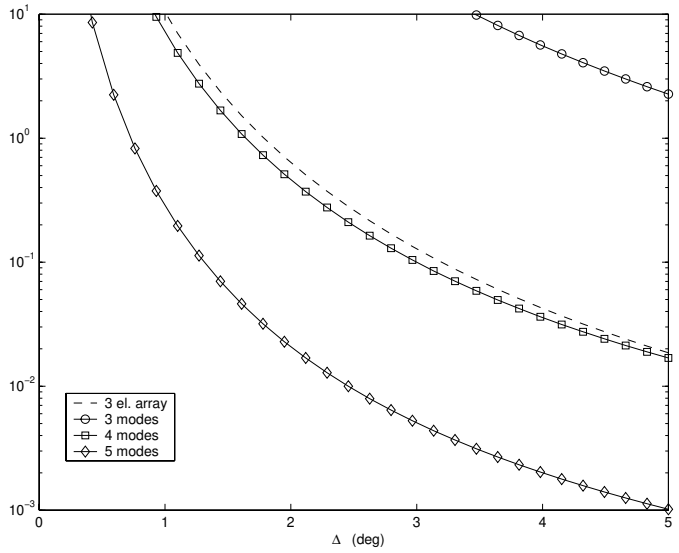


Figure 7.3: The square root of the CRB for several different configurations when two waves are incident from $(30°, 30° + \Delta)$ with SNR=10dB and 1000 samples.

incidence angles, since the array will work best for broadside and worst for end-fire incidence.

In Figure 7.4, the standard deviation is shown for the same antenna configurations as in Figure 7.3 when two waves are incident from $(\phi_0, \phi_0 + 5°)$. It is found that the multimode antenna performance is worse for angles close to end-fire $\phi_0 = 0°$ (similar to the traditional array). Also, the manner in which the performance increases when moving away from end-fire, is almost identical for the array and the multimode antenna. However, since the received power of the biconical antenna depends on the DOA, due to the real sinusoidal terms, the performance will vary with the DOA. This dependency will be different for each mode since the sinusoidal terms are different, see (7.2).

Thus, the multimode antenna offers DOA estimation performance similar to that of an array. In the next section, two methods for DOA estimation with a multimode antenna are evaluated.

In order for the analysis in the following sections to be valid, some additional assumptions are needed:

- the steering matrix has full rank, i.e., $\mathrm{rk}(\mathbf{A}) = p$

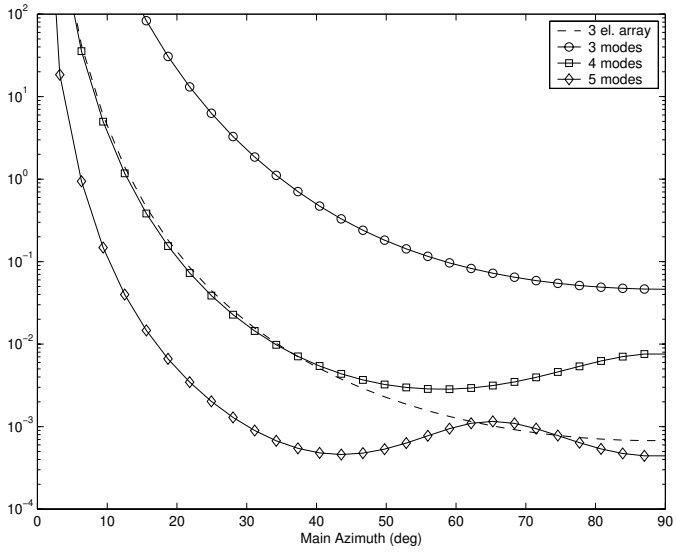


Figure 7.4: The square root of the CRB for several different configurations when two waves are incident from $(\phi_0, \phi_0 + 5°)$ with SNR=10dB and 1000 samples.

- $\mathbf{n}(t)$ is temporally white and circularly Gaussian distributed: $\mathbf{n}(t) \in \mathcal{N}(0, \sigma^2\mathbf{I})$
- $\mathbf{s}(t)$ is also temporally white and circularly Gaussian distributed: $\mathbf{s}(t) \in \mathcal{N}(0, \mathbf{S})$

The noise is both spatially and temporally white, while the signal is only assumed to be temporally white. Furthermore, the signal is assumed to be uncorrelated with the noise.

7.4 Estimation Methods

The analysis in the previous section was based on the CRB on the estimation error. In this section, algorithms that approximately achieve this lower bound will be discussed. In principle, all DOA estimation schemes derived for a general antenna array can also be applied to the multimode antenna by inserting the new steering matrix. For an overview of DOA estimation methods, see [KV96, VS01]. Here, the DOA estimation performance for two of the most well known algorithms will be examined and compared to the CRB.

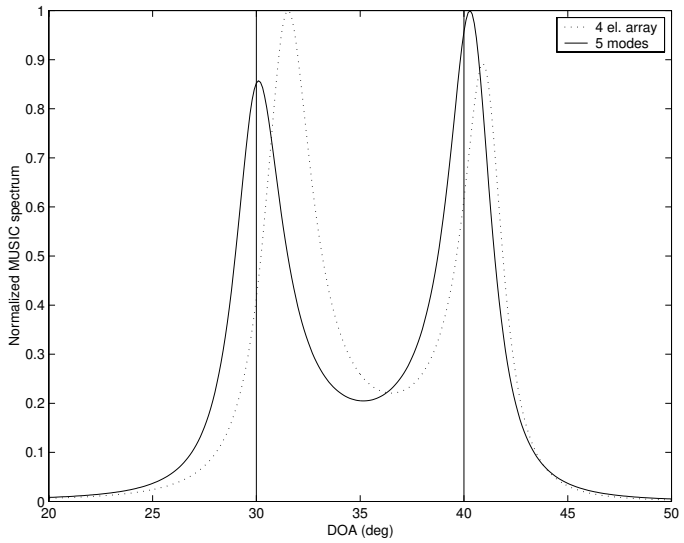


Figure 7.5: The normalized MUSIC spectrum when two waves are incident from 30° and 40° upon a five mode multimode antenna and a four element array with SNR=10dB and 1000 samples.

In [Sch79], a very popular high resolution DOA estimation method, MUSIC, was introduced where the DOA estimates are taken as those ϕ that maximizes the MUSIC criterion function

$$\hat{\boldsymbol{\phi}} = \arg\max_{\boldsymbol{\phi}} \frac{\mathbf{a}^H(\phi)\mathbf{a}(\phi)}{\mathbf{a}^H(\phi)\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\mathbf{a}(\phi)}. \tag{7.12}$$

Usually this is formulated as finding the p largest peaks in the "MUSIC spectrum". Here, $\hat{\mathbf{E}}_n$ denotes the $M-p$ eigenvectors corresponding to the $M-p$ smallest eigenvalues of the estimated covariance matrix $\hat{\mathbf{R}}$. A typical example of a MUSIC spectrum is shown in Figure 7.5, where two waves are incident from 30° and 40° upon a five mode multimode antenna and a 4 element array with SNR=10dB and 1000 samples. This figure indicates that the multimode antenna, in this case, offers a performance similar to the antenna array. Furthermore, when the incident signals are uncorrelated, the MUSIC method is unbiased and approximately achieves the CRB (i.e. efficient). Unfortunately, the method breaks down for correlated signals, which are common in communication applications due to multipath. Thus, in these cases another method is required.

The Maximum Likelihood (ML) method is the chief systematic approach of model-based estimation. In [Böh86], the ML DOA estimator for an array

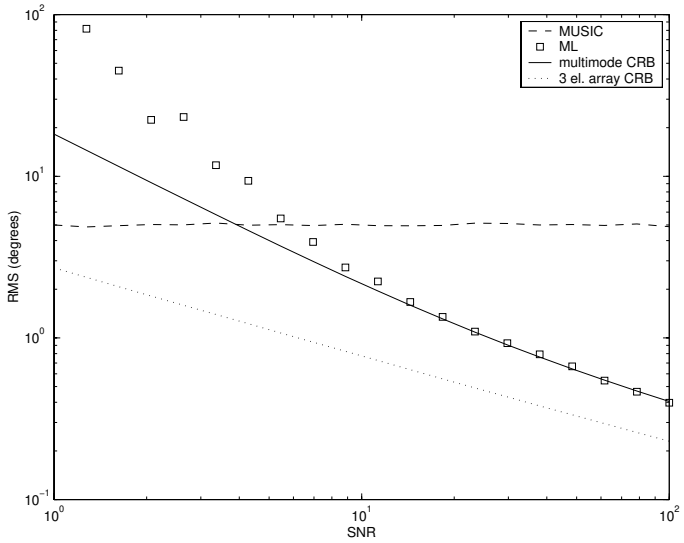


Figure 7.6: The RMSE (first DOA) for ML and MUSIC when two waves are incident from 40° and 60° upon a four mode antenna versus SNR with 1000 samples and correlated signals (0.9).

of antenna elements was introduced. The ML estimates are unbiased and efficient also for correlated signals. Furthermore, the ML method is easily modified to the multimode antenna by inserting the new steering matrix. Using the model in (7.9) and the statistical assumptions in Section 7.2, the negative log likelihood function to be minimized becomes

$$\hat{\boldsymbol{\phi}} = \arg\min_{\boldsymbol{\phi}} \ \log\left|\mathbf{P}_A\hat{\mathbf{R}}\mathbf{P}_A + \frac{\mathrm{Tr}[\mathbf{P}_A^{\perp}\hat{\mathbf{R}}]}{M-p}\mathbf{P}_A^{\perp}\right| \tag{7.13}$$

and the DOA estimates are taken as the minimizing arguments of it. In Figure 7.6, the RMSE for ML and MUSIC when two waves are incident from 40° and 60° upon a four mode antenna, is shown for different SNRs. Two highly correlated signals $\mathbf{S}_{1,2} = \mathbf{S}_{2,1} = 0.9$ are used. As expected, the MUSIC method fails, while the ML method achieves the CRB already at quite low SNRs. The simulation examples clearly indicates that the biconical antenna offers DOA estimation performance comparable to that of a traditional antenna array, although with one additional mode than elements.

7.5 Conclusions

By exciting the biconical antenna with higher order modes, directional radiation patterns are obtained that successfully can be used to estimate DOAs. Since the multimode antenna only requires one antenna element it avoids many of the problems encountered when employing an array of elements. Another advantage of the biconical antenna is that it can be manufactured at a very low cost [DW98]. A data model for the multimode antenna was derived and the direction finding performance was examined by calculating the CRB and the ML estimator for several different cases. It was found that the multimode antenna offers direction finding performance comparable to the traditional antenna array using only a single antenna. Thus, exploiting multimodes for direction finding is a new and interesting alternative that offers several advantages over traditional arrays.

However, it should be stressed that only the theoretical possibility of using the higher order modes of antennas for direction finding was examined. More work on the antenna design is needed to find suitable positions of the feeds, appropriate dimensions, and a matching network that minimizes the losses in the antenna. It is also important to note that in this chapter, only the case of exploiting the individual modes directly is examined. Many other possibilities exists. For instance, is also possible to use several modes to form narrow beams, that can be used for direction finding applications. Hence, exploiting higher order modes for direction finding offers many interesting possibilities.

Chapter 8

Direction Finding Using a Switched Parasitic Antenna

Direction finding by exploiting the directional radiation patterns of a Switched Parasitic Antenna (SPA) is considered. By employing passive elements (parasites), which can be shorted to ground using pin diodes, directional radiation patterns can be obtained. The direction finding performance of the SPA is examined by calculating a lower bound on the direction finding accuracy, i.e. the CRB. It is found that the SPA offers a compact implementation with high-resolution direction finding performance using only a single radio receiver. Thus, exploiting SPAs for direction finding is an interesting alternative to traditional antenna arrays offering compact and low-cost antenna implementations.

8.1 Introduction

Direction finding is of great importance in a variety of applications, such as radar, sonar, communications, and recently also personal locating services. In the last two decades, direction finding and sensor array processing has attracted considerable interest in the signal processing community. The focus of this work has been on high resolution, i.e. a resolution higher than the width of the main lobe, DOA estimation algorithms [KV96]. These algorithms exploit the fact that an electromagnetic wave that is received by an array of antenna elements reaches each element at different time instants. Although the performance of these systems is excellent, an unfortunate aspect is the high costs of employing a radio receiver for each antenna element. Furthermore, it is expensive to calibrate and maintain antenna arrays with

many antenna elements.

Recently, it was proposed to employ an SPA for direction finding [PTS+98, PTL+97] that only uses a single active radio receiver, thereby significantly reducing the cost. The SPA offers characteristics similar to an array antenna with several beams by using passive antenna elements that serve as reflectors when shorted to ground. Different directional patterns can be achieved by switching the short-circuits of the passive elements using pin diodes. The possibilities of exploiting these patterns for high-resolution DOA estimation will be examined in this paper, since no attempt to employ high-resolution DOA methods was undertaken in [PTS+98, PTL+97].

8.2 Switched Parasitic Antenna

Switched Parasitic Antennas offering directional patterns dates back to the early work of Yagi and Uda in the 1930's [Bal82]. The concept is to use a single active antenna element, connected to a radio transceiver, in a structure with one or several passive antenna elements, operating near resonance. These passive elements are called Parasitic Elements (PEs) and act together with the active element to form an array, as in the well known Yagi-Uda array [Bal82]. To alter the radiation pattern, the termination impedances of the PEs are switchable, to change the current flowing in those elements. The PEs become reflectors when shorted to the ground plane using pin diodes [STLO00] and when not shorted, the PEs have little effect on the antenna characteristics. The receiver is always connected to the center antenna element so there are no switches in the RF direct signal path.

An interesting possibility to obtain directional information is to sample the received signal with several different radiation patterns, since the switching time of a pin diode is only of the order of a few nanoseconds. This technique of oversampling the received signal is common in many communication systems, but here the oversampling is performed in both time and space, i.e. spatio-temporal oversampling. If the increased sampling rate (or bandwidth) poses a problem, a bandpass sampling strategy could also be employed. In this chapter, the potential in using the different radiation patterns of an SPA for direction finding will be examined. However, further work is needed on the practical aspects of the antenna design as well as sampling strategies.

In the literature, it has been proposed to use monopoles on a ground plane [STLO00] or patch antennas [PT97] as SPAs. In this paper, a monopole on a ground plane is used because of its omnidirectional properties. A four Direction Symmetric (4-DS) monopole parasitic antenna is shown in Figure

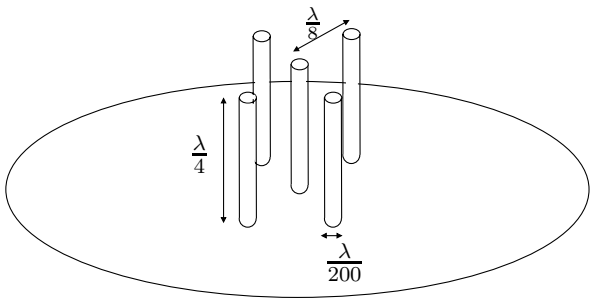


Figure 8.1: A five element monopole SPA. The center element is active and connected to the transceiver. The four passive antenna elements can be switched in or out of resonance using appropriately biased pin diodes.

8.1 and a three Direction Symmetric (3-DS) antenna is shown in Figure 8.2. The antenna in Figure 8.2 has an additional circle of parasitic elements that always are shorted to ground. The effect of this arrangement is an increased directivity as their lengths are shorter than the corresponding resonant length ($\approx \lambda/4$) and will lead the induced Electro-Motive Force (EMF) [Bal82]. The lengths and distances displayed in Figure 8.1 and 8.2 are not optimal in any way. Note that the resulting antennas are very compact ($\lambda/4$ and $\lambda/2$) compared to corresponding linear arrays with $\lambda/2$ separation distance (2λ and $3\lambda/2$).

The antennas in Figures 8.1 and 8.2 were simulated using the HFSS (High Frequency Structure Simulator) from Agilent Technologies Inc. which is a 3D simulator using the Finite Element Method (FEM), to solve for the electromagnetic field. The software was used to calculate the far-field radiation pattern of the antenna for different settings of the switched parasitics. The monopole elements were cylindrical with a length to radius ratio $l/r = 100$ which yields a first resonance at approximately 0.24λ [Bal82].

The far-field power radiation pattern in the azimuth plane $F(\phi)$ for the 4-DS SPA is shown in Figure 8.3. The corresponding pattern for the 3-DS SPA is similar and not shown. The directivity of the two antennas are 9.9 dB and 10.0 dB respectively. Thereby, only a small gain in directivity was achieved by the extra ring of shorted parasites.

Once the far-field radiation properties are found, it is straightforward to derive a model for the received voltages [Kil99, Mey00]. If p waves are incident upon an antenna with M symmetry directions, the received voltages can be written in matrix form as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\phi})\mathbf{s}(t) + \mathbf{n}(t), \tag{8.1}$$

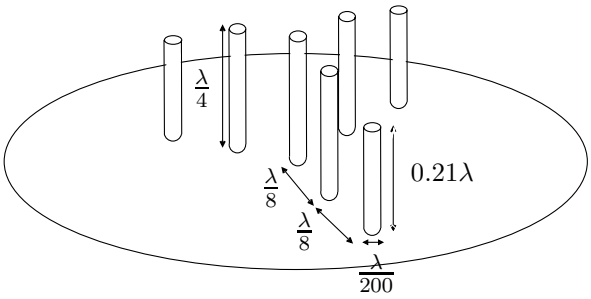


Figure 8.2: A seven element monopole SPA. The center element is active and connected to the transceiver. The three passive antenna elements closest to the active can be switched in or out of resonance using appropriately biased pin diodes. The three outermost monopoles are hardwired to ground.

where the vector of measured voltages $\mathbf{x}(t)$ is $M \times 1$. The matrix $\mathbf{A}(\boldsymbol{\phi})$ ($M \times p$) corresponds to the response of the different symmetry directions and has elements $[\mathbf{A}(\boldsymbol{\phi})]_{qr} = F(\phi_r + 2q\pi/M)$. This matrix is typically called steering matrix in the sensor array processing literature. The signal vector $\mathbf{s}(t)$ is $p \times 1$ and contains the strength of the received fields. Finally, the noise vector $\mathbf{n}(t)$ is $M \times 1$.

In order for the analysis in the following sections to be valid, some additional assumptions are needed:

- the steering matrix has full rank, i.e., $\mathrm{rk}(\mathbf{A}) = p$
- $\mathbf{n}(t)$ is temporally white and circularly Gaussian distributed: $\mathbf{n}(t) \in \mathcal{N}(0, \sigma^2\mathbf{I})$
- $\mathbf{s}(t)$ is also temporally white and circularly Gaussian distributed: $\mathbf{s}(t) \in \mathcal{N}(0, \mathbf{S})$

The noise is both spatially and temporally white, while the signal is only assumed to be temporally white. Furthermore, the signal is assumed to be uncorrelated with the noise.

8.3 Direction Finding Performance

The data model (3.1) is identical to the usual data model used in sensor array processing [KV96], except for a new steering matrix. This will of course change the direction finding properties. Before the properties of a specific DOA estimation scheme is studied, a lower bound, the Cramér-Rao lower Bound (CRB), on the variance of the DOA estimates will be analyzed. Note

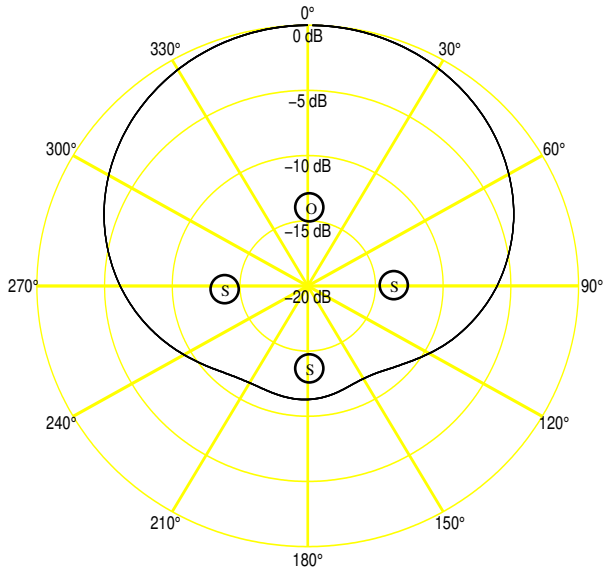


Figure 8.3: Power radiation pattern of the five element monopole antenna shown in Figure 8.1 with three parasitics shorted (S) to ground and one open (O).

that it is possible to asymptotically achieve this bound with many methods in the literature [KV96].

Expressions for the CRB was derived for an array of antenna elements in [OWVK89]; and can also be applied to the parasitic antenna by changing the steering matrix.

$$E\{(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0)(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0)^T\} \geq \mathbf{B} \tag{8.2}$$

$$\mathbf{B} = \frac{\sigma^2}{2N}\left[Re\{(\mathbf{D}^H\mathbf{P}_A^{\perp}\mathbf{D}) \odot (\mathbf{S}\mathbf{A}^H\mathbf{R}^{-1}\mathbf{A}\mathbf{S})^T\}\right]^{-1} \ ,$$

where the elements of $\mathbf{D}_{qr} = \frac{\partial F(\phi+2q\pi/M)}{\partial \phi}\Big|_{\phi=\phi_r}$. Furthermore, $\odot$ denotes the Hadamard (or Schur) product, i.e., element-wise multiplication and $\mathbf{P}_A^{\perp} = \mathbf{I} - \mathbf{P}_A = \mathbf{I} - \mathbf{A}\mathbf{A}^{\dagger}$ [1] is the orthogonal projector onto the null space of $(\mathbf{A})^H$. The matrix $\mathbf{R} = \mathbf{A}\mathbf{S}\mathbf{A}^H + \sigma^2\mathbf{I}$ is the covariance matrix of the measured voltages $\mathbf{x}(t)$ and N denotes the number of time samples.

The square root of the CRB, i.e. the standard deviation, is shown in Figure 8.4 for the antenna configurations in Figure 8.1 and 8.2 as two waves are incident from $(30°, 30° + \Delta)$. Only the CRB for the first DOA, i.e.

[1] $\mathbf{M}^{\dagger}$ is the Moore-Penrose pseudo inverse of $\mathbf{M}$.

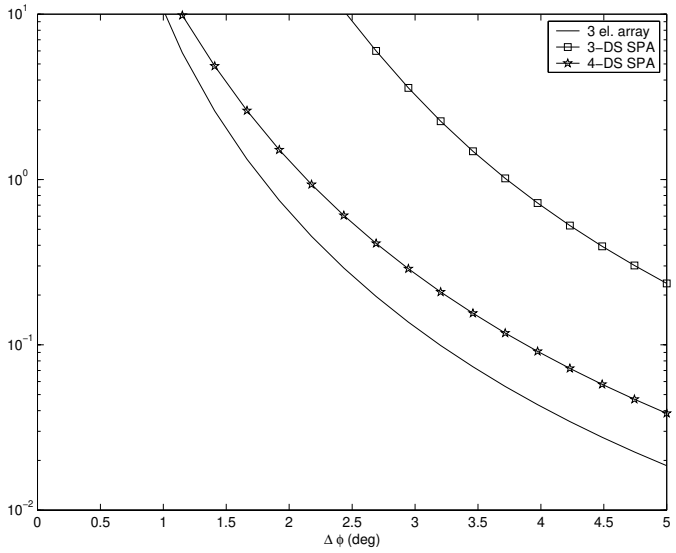


Figure 8.4: The square root of the CRB for the configurations in Figure 8.1 and 8.2 when two waves are incident from $(30°, 30° + \Delta)$ with SNR=10dB and 1000 samples.

the wave arriving from 30°, is shown since the CRB for the second DOA will behave similarly. The standard deviation for a uniform linear array of three elements spaced $\lambda/2$ apart is compared to the 4-DS and 3-DS SPAs. As expected, the performance is better when using four rather than the three symmetry directions. Also, note that the three element array performs slightly better than the 4-DS SPA. However, these results depend on the incidence angles, since the array will work best for broadside and worst for end-fire incidence.

In Figure 8.5, the standard deviation is shown for the same antenna configurations as in Figure 8.4 when two waves are incident from $(\phi_0, \phi_0 + 5°)$. The parasitic antenna, due to its symmetrical properties, offers similar direction finding performance properties for all incidence angles. The linear array performs worse than the parasitic antenna at end-fire incidence, while performing much better at broad-side incidence. However, for many direction finding applications, the direction finding performance of the parasitic antenna is sufficient and the cost reduction of using only a single radio receiver outweighs the loss in performance for broad-side angles. It should also be stressed that the antenna designs in Figure 8.1 and 8.2 are by no means optimal and better DOA properties may be obtained by a proper optimization.

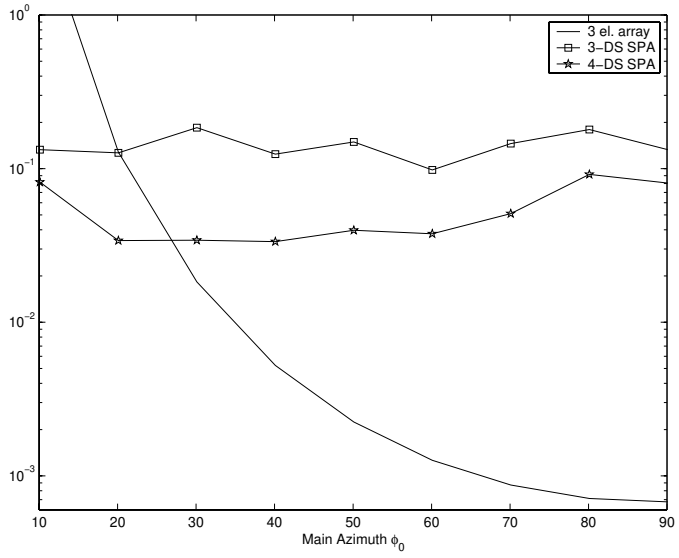


Figure 8.5: The square root of the CRB for the configurations in Figure 8.1 and 8.2 when two waves are incident from $(\phi_0, \phi_0 + 5°)$ with SNR=10dB and 1000 samples.

8.4 Estimation Methods

The analysis in the previous section was based on the CRB on the estimation error. In this section, algorithms that approximately achieve this lower bound will be discussed. In principle, all DOA estimation schemes derived for a general antenna array can also be applied to a parasitic antenna by inserting a new steering matrix. For an overview of DOA estimation methods, see [KV96, VS01].

In [Sch79], a popular high resolution DOA estimation method, MUSIC, was introduced where the DOA estimates are taken as those ϕ that maximizes the MUSIC criterion function

$$\hat{\boldsymbol{\phi}} = \arg\max_{\boldsymbol{\phi}} \frac{\mathbf{a}^H(\phi)\mathbf{a}(\phi)}{\mathbf{a}^H(\phi)\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^H\mathbf{a}(\phi)}, \tag{8.3}$$

where the steering vector has elements $\mathbf{a}_q(\phi) = F(\phi + 2q\pi/M)$, $q = 1, \ldots, M$. Usually this is formulated as finding the p largest peaks in the "MUSIC spectrum". Here, $\hat{\mathbf{E}}_n$ denotes the $M - p$ eigenvectors corresponding to the $M - p$ smallest eigenvalues of the estimated covariance matrix $\hat{\mathbf{R}}$. A typical example of a MUSIC spectrum is shown in Figure 7.5, where two waves are incident from 25° and 45° upon a 4-DS SPA and a three element array with

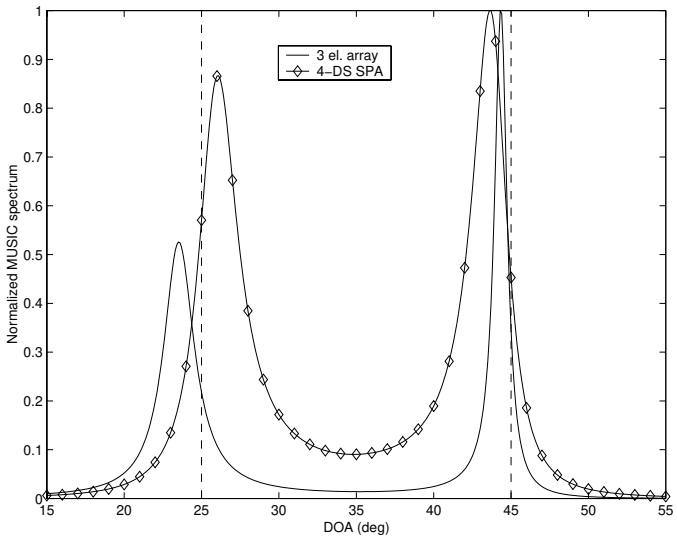


Figure 8.6: The normalized MUSIC spectrum when two waves are incident from 25° and 45° upon a 4-DS parasitic antenna and a three element array with SNR=10dB and 1000 samples.

SNR=10dB and 1000 samples. This figure indicates that the SPA, in this case, offers a high-resolution direction finding performance similar to that of an antenna array without the cost of many radio receivers. Most other DOA estimation schemes [KV96, VS01] can also be applied to SPAs with similar results. For instance, the SML algorithm [OWVK89] for this type of antenna was implemented. The RMSE of the ML estimator achieved the CRB bound from Section 7.3, as expected.

8.5 Conclusions

The potential use of an Switched Parasitic Antenna for high-resolution direction finding was investigated. By employing passive elements, which can be shorted to ground using pin diodes, directional radiation patterns are obtained that can be used successfully to estimate DOAs. The main advantage with this concept is that only one radio receiver is needed, thereby reducing the costs significantly compared to traditional antenna arrays where one radio receiver per element typically is employed. Another advantage of the SPA is that a very compact implementation of the antenna is possible.

A data model for the SPA was presented and the direction finding performance was examined by calculating the CRB and the MUSIC estimator. It

was found that the SPA offers a compact implementation with high-resolution direction finding performance using only a single radio receiver. Thus, exploiting SPAs for direction finding is an interesting alternative that offers several advantages over traditional arrays.

Chapter 9

Physical Channel Modeling of MIMO Systems

A general framework for physical spatio-temporal channel modeling is presented. Based on this framework, models for single antenna systems as well as Multi-Input Multi-Output (MIMO) models for multi-element antenna systems, are derived. The model is based on results from ElectroMagnetic (EM) scattering and wave propagation, thereby including many of the channel characteristics encountered in practice. A general description that captures the most important scattering properties of arbitrary objects is derived in a compact dyad formulation by using results from rough surface scattering. The polarization properties of the channel and the antennas are thus included in the model, which allows for studies of the impact of different antenna arrangements and polarizations. Several channel environments are simulated, and properties such as spatial correlation, channel capacity, and time evolution of the channel are calculated. The proposed physical channel model is also suitable for other interesting application areas, such as long-range prediction of fading signals, feedback based transmit diversity schemes, and wideband MIMO systems where the temporal properties are important.

9.1 Introduction

An accurate knowledge of the mobile communications channel is of greatest importance when designing radio systems. The performance of a detector or receiver is highly dependent on the characteristics of the radio channel. If an accurate model of the channel is available, it is possible to design receiver

algorithms that achieve a high performance by exploiting the properties of the channel. Thus, to accurately model the radio channel is an important problem that affects the performance of wireless communication systems. Consequently, radio channel modeling has been a subject of intense research for a long time.

The type of channel model that is desired depends critically on the type of system that is considered. For instance, different types of channel models are needed for narrowband, broadband, single antenna systems, and multiple antenna systems. Early channel modeling work studied the behavior of the received envelope of a single antenna system. This lead to statistical models of the fading of the received signal [Aul79, Cla68, Jak74].

With the advent of antenna array systems, new types of channel characteristics appeared that needed to be modeled. The correlation between different antenna elements was now modeled using different assumptions regarding the physical properties of the channel [Jak74, Lee73, Lee89]. These models were typically developed for narrowband systems, where only the spatial domain of the channel was considered while the temporal domain was largely neglected. As the need for higher data rates was increased, larger bandwidths was necessary. To accurately model broadband systems, the temporal domain must be considered and properties such as time-of-arrival and delay spread must be modeled. The success of cellular communication systems has resulted in a vast literature on radio channel modeling. The interested reader is referred to the following textbooks [Jak74, Lee89, Par92, Pro95, Sau99, Ste92] and further references therein. See also the tutorials and overviews [ARY95, Ert98, FL96, Has93] for further information.

The need to model many different types of radio channels has resulted in a large number of different modeling approaches that can be found in the literature. One reason for the abundance of approaches is that the propagation of a transmitted radio signal is a very complicated process, and modeling all the phenomena in a physically correct manner is virtually impossible. The transmitted signal will usually arrive at the destination along several paths, i.e. multipaths, where the signal is reflected, refracted or diffracted. Thus, many different types of simplifications and approximations are necessary in order to obtain a simple yet accurate model of the wireless radio communications channel. The different approaches to channel modeling may roughly be divided in the following categories:

1. Ray-Tracing Models
   The advances in computing power has made it possible to accurately predict the radio channel if a detailed map of the area is available. By employing optical concepts, i.e. treating the waves as rays, the

field strength can be calculated. Hence, these approaches are usually called ray-tracing approaches [AN00, Ami92, CPdAG98, DKW97, SI94]. However, ray-tracing methods are still very computer intensive and their application is limited to modeling deterministic scenarios where a good map is available. The main application of ray-tracing in channel prediction is cell-planning where the average received power over an area is needed.

2. Empirical Models
   A number empirical channel models have been developed that also rely on optical concepts [Ber00, Bla00, Zha97], where two of the the most well known examples are the Okumura [Oku68] and Hata [Hat80] models. In this case, no detailed modeling of the area is performed. Instead, simplified versions of ray-tracing are fitted to measured data through a number of parameters. Typical examples of models that are fitted to data are diffraction via roof-tops and propagation over smooth/irregular terrain. These models are also mainly used in cell-planning.

3. Statistical Models
   These models typically use a highly simplified physical reasoning to arrive at an statistical description of the radio channel [Aul79, Cla68, Ert98, SV87, SJJS00, TCJ+72]. Although in some cases, these statistical models are obtained by fitting statistical models to measurement data. A more elaborate physical reasoning may be found in [BD91, DR93]. A stochastic spatial channel model that can be employed for studies of different antenna arrangements, is proposed in [ZFDW00]. In that work, the statistical properties are estimated from huge ray-tracing data sets. Statistical models are used both for simulating channel data but also serves as a starting point when designing optimal receivers.

4. Semi-Deterministic Models
   Here, a more accurate physical modeling is performed. A typical approach is to place scattering objects in the area to be modeled and then calculate the properties of this deterministic scenario using fundamental physics [BJ98, Ert98, FMB98, LR96, LLL97, NA98, PRR96, RDNP95, ZFW98]. By assuming some statistical distribution of the location of these objects and their properties, a semi-deterministic radio channel model is obtained. Similar to the purely statistical models, these models may be used to simulate data and designing receiver algorithms. Their main benefit is, however, that they provide some

physical insight of the channel properties while avoiding the intensive computations of ray-tracing schemes.

Note that the ray-tracing channel models are often called propagation models in the literature, while models of more statistical nature are called channel models. The term propagation models is mostly used for models that calculate the received signal from a deterministic environment employing theory from EM wave propagation. The term channel model is typically used in communication and signal processing literature for statistical based models that do not rely on detailed results from wave propagation. Here, the term channel model is used throughout the report for all these model types.

Recently, it has been proposed to use multiple antennas at both the transmitter and receiver, resulting in MIMO channels. Impressive capacity gains using these MIMO systems have been demonstrated for rich scattering environments [FG98, Tel95]. The success of MIMO systems is due to the fact that they more fully exploit the spatial dimension of the available radio spectrum than previous systems. These MIMO systems require a new set of channel models that accurately models the correlation between the antenna elements, not only on the receiver side, but also on the transmitter side. Furthermore, the high data rates that these systems will support also requires that the temporal domain is accurately modeled. Thus, MIMO channels puts a set of new demands on the channel model. The literature on MIMO channel models is relatively scarce [Bur00, GBGP00, PAKM00, SJBF00]. This report aims to derive a more general model of the MIMO channel that is applicable to arbitrary antennas (including dual-polarized antennas) and also suitable for channel prediction studies.

To obtain a realistic model for the MIMO channel results from fundamental electromagnetic, scattering theory as well as linear algebra and mathematical statistics will be used. Thus, the proposed model may fall under the semi-deterministic category above. In principle, a simple and realistic MIMO channel model for simulation and receiver design may be of either statistical or semi-deterministic nature. Here, a semi-deterministic model will be derived, since it provides some physical insight of the new phenomena that are introduced by using antenna arrays at both transmitter and receiver. Furthermore, a semi-deterministic model may more easily incorporate the effect of different antenna arrangements that will form an integral part of most MIMO systems. It also makes it possible to relate different channel capacity figures to environmental factors [BFHS00] such as angular spread, dual-polarized antennas, and Doppler-spread. Finally, using a model that incorporates physical dimensions, the time-variation of the channel due to motion may be accurately modeled [HHDH99]. Thus, it may be used

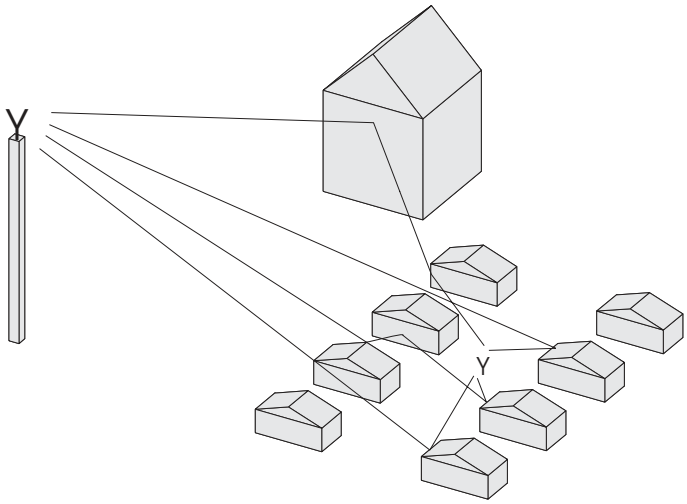


Figure 9.1: Typical suburban scatter environment with local scattering and a dominant scatterer.

for studying channel prediction schemes that have gained increased interest in the literature [DHHH00, EK99]. Channel prediction is of particular importance in transmit diversity applications and MIMO systems, where the amount of signaling may be significantly reduced by using prediction schemes.

The outline of this report is as follows: Since the MIMO radio channel model exploits some results from electromagnetic wave propagation and scattering, some topics from these areas are reviewed in Section 9.2 and Section 9.3. Based on these results, a model of the impulse response for the single antenna case is derived in Section 9.4. This model is then extended to a Single-Input Multi-Output (SIMO) model for the case of a multi-element antenna, i.e. array, at the receiver in Section 9.5. The general case of arrays at both receiver and transmitter, i.e. a MIMO channel, is then treated in Section 9.6. Finally, some conclusions are presented in Section 9.7.

9.2 Radio Wave Propagation

The propagation situation in a practical wireless communication channel is very complex. The signal that is transmitted from the mobile phone will reach the base station antenna through many different paths. Consider the propagation scenario in Figure 9.1 that contains a mobile antenna and a base station antenna as well as numerous scattering objects. Several local scattering objects (houses) are located in the vicinity of the mobile and a

large scattering object (large house) is located further away. The transmitted field from the mobile will experience electromagnetic scattering, reflection, refraction, and diffraction before it reaches the base station antenna where a voltage finally is induced.

In principle, it is possible to quite accurately calculate the received signal at the base station by employing ray-tracing concepts. Of course, a detailed map and the electrical properties of the materials are required in that case. However, that type of calculation is only valid in that particular scenario and does not provide much insight into the general characteristics of the channel. Another alternative is to conduct extensive measurements and fit a statistical model to the measured data. Although the measurements provide much physical insight, the resulting statistical model is only valid for the channel conditions under which the measurements were conducted. In the case of a MIMO system relatively few extensive measurement campaigns have been conducted and many parameters are needed to describe the relations between the transmit- and receive-elements. Thus, a physics based channel model will be derived in this report by employing results from radio wave propagation, antenna theory, and electromagnetic scattering. However, measurement data may be used also here to find the distributions of the channel parameters that best describe different types of channels.

Consider the scenario in Figure 9.2 that contains a transmit antenna (mobile) and a receive antenna (base station) as well as a scattering object (scatterer). The vector $\vec{r}_{ms}$ is defined as the vector from the mobile to the scatterer and the vector $\vec{r}_{sb}$ as the vector from the scatterer to the base station. In the uplink scenario, the mobile transmits a wave along the $\vec{r}_{ms}$ direction. This wave reaches the scatterer after some time and gives rise to a scattered wave. The scattered wave later impinges on the base station antenna from the $\vec{r}_{sb}$ direction and induces a voltage which is measured. If only one scattering object is present, as in Figure 9.2, it is rather straightforward to find an expression for the received field at the base station using fundamental wave propagation and scattering theory. However, if many objects are present, things become much more involved since the interaction between the different objects must be accounted for. For instance, the signal in one of the paths in Figure 9.1 is first reflected by two building walls before being received by the base station antenna. A very complex model would be obtained if these types of paths, i.e. multiple scattering, are accurately modeled. In fact, that model would resemble the computer intensive ray-tracing schemes previously alluded to.

By placing a scattering object at the last reflection and neglecting the interaction between scatterers, some of the properties of multiple scattering will be retained but a much simpler model is obtained. This would in Figure

9.1 correspond to placing a scatterer at some of the houses. Using this approach, the total field at the receiving antenna at the base station is obtained by summing the contributions from all scatterers. The following analysis will use some results from wave propagation. For an introduction to scattering and wave propagation, see [Bal89, Sau99]. Further material on scattering may be found in [BS63, BSea87, FM94, dH57, Wai55] and the references therein. Numerous textbooks on radio channel characteristics have been written with applications in cellular networks [Ber00, Bla00, Jak74, Par92].

The different parts of the propagation path in Figure 9.2 will now be examined in detail and expressions for the transmitted field, the scattered field and the received voltages will be derived.

The transmitted field is of course determined by what type of antenna that is used. Antenna theory is a mature research area where analytical expressions exists for many different types of antennas. Most of the analytical results were obtained rather early, before the computer revolution occurred. For results on most common antenna types refer to the following textbooks and further references therein [Bal82, Bal89, Kra50, Ell81, KH69, SF52, ST98, Wee68].

For all antennas it is possible to express the transmitted field at a point $\vec{r}_{ms}$, with the antenna positioned at the origin, in the form

$$\vec{E}_t(\vec{r}_{ms}) = \frac{e^{-jk|\vec{r}_{ms}|}}{|\vec{r}_{ms}|} G_t(\vec{r}_{ms}) \vec{g}_t(\vec{r}_{ms}), \tag{9.1}$$

where the first term includes a phase shift due to the propagation time and the divergence of the transmitted field, k denotes the wave number[1], $G_t(\vec{r}_{ms})$ is the element pattern and $\vec{g}_t(\vec{r}_{ms})$ is the orientation of the transmitted field.

Example Radiation from a Dipole Antenna

For example, the corresponding expressions for a $\vec{z}$ oriented dipole antenna of length l, is

$$G_t = \frac{j\eta I_0}{2\pi} \left[\frac{\cos\left(\frac{kl}{2}\cos\theta\right) - \cos\left(\frac{kl}{2}\right)}{\sin\theta} \right] \tag{9.2}$$

$$\vec{g}_t = \vec{\theta} = \vec{x}\cos\theta\cos\phi + \vec{y}\cos\theta\sin\phi - \vec{z}\sin\theta, \tag{9.3}$$

[1]The wave number $k = \frac{2\pi}{\lambda} = \omega\sqrt{\mu\epsilon}$, where λ is the wavelength. The symbols ϵ and μ denote the dielectric constant and the permeability of the medium respectively.

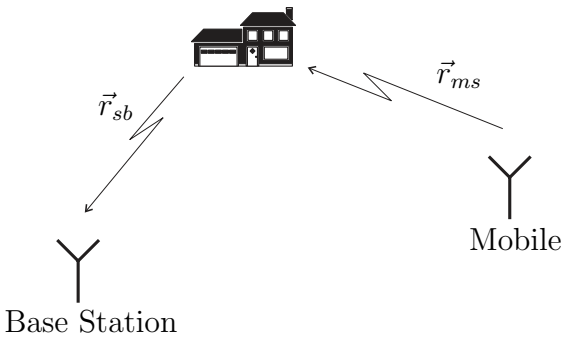


Figure 9.2: Example of a propagation path where a mobile transmits a signal that is received by a base station antenna via a scattering object.

where $\vec{x}, \vec{y}, \vec{z}$ denotes the basis vectors in the rectangular coordinate system, η is the intrinsic impedance of the medium[2], I_0 is the input current, and (θ, ϕ) are the respective angles of the spherical coordinate system which are calculated from the direction of $\vec{r}_{ms}$.

□

After some time, the transmitted wave will reach the scatterer in Figure 9.2. Scattering occurs and a new field $\vec{E}_s$ is radiated, that is received at the base station. The field at the base station can be expressed in a manner similar to (9.1) as

$$\vec{E}_s(\vec{r}_{sb}) = \frac{e^{-jk|\vec{r}_{sb}|}}{|\vec{r}_{sb}|} E_0 \vec{\alpha} \tag{9.4}$$

where E_0 is the amplitude of the field incident upon the scatterer and $\vec{\alpha}$ represents the scattering process and contains both direction and amplitude. The amplitude of the field, E_0 may be expressed using (9.1) and the following expression for the scattered field is obtained

$$\vec{E}_s(\vec{r}_{sb}) = \frac{e^{-jk(|\vec{r}_{ms}|+|\vec{r}_{sb}|)}}{|\vec{r}_{ms}||\vec{r}_{sb}|} G_t(\vec{r}_{ms})\vec{\alpha}. \tag{9.5}$$

The scattering function $\vec{\alpha}$ depends on the direction of incidence, the direction of scattering, and the orientation of the incident field ($\vec{g}_t$) as well as the properties of the scatterer. For objects of simple shapes, it is possible to derive analytic expressions for $\vec{\alpha}$ that can be used to study the underlying

[2]The intrinsic impedance is defined as $\eta = \sqrt{\mu/\epsilon}$. For free space the intrinsic impedance becomes $120\pi\ \Omega \approx 377\ \Omega$.

physics [Sva00a]. However, for channel simulation purposes it is usually sufficient with much simpler forms that are based on statistical electromagnetic scattering theory. Section 9.3 will treat the choice of scattering function in more detail.

An expression for the received voltage can be obtained once the scattered field is known. However, antennas are typically analyzed in transmission mode and no expressions for the induced voltages, when using the antenna in receive mode, are included in the antenna textbooks. Recently, expressions of the antenna in receive mode was presented in [Kil99, Mey00, Sva98b], where the induced voltage from a plane wave was derived. Using those results, the induced voltage at the base station antenna due to the scattered field $\vec{E}_s$, is obtained as

$$v = CG_r(-\vec{r}_{sb})\vec{g}_r(-\vec{r}_{sb}) \cdot \vec{E}_s, \tag{9.6}$$

where G_r and $\vec{g}_r$ are the element pattern and polarization of the receiving antenna. The scalar C is defined as

$$C = -j4\pi/k\eta I_0, \tag{9.7}$$

where I_0 is the input current used in the calculation of the radiation function. Note that the evaluation of the voltage in (9.6) involves the scalar product between the vector $\vec{g}_r$ and the field $\vec{E}_s$. Using (9.1),(9.5), and (9.6), the received voltage can also be written as

$$v = C\frac{e^{-jk(|\vec{r}_{ms}|+|\vec{r}_{sb}|)}}{(|\vec{r}_{ms}| \star |\vec{r}_{sb}|)^{\gamma}} G_t(\vec{r}_{ms})G_r(-\vec{r}_{sb})\vec{g}_r(-\vec{r}_{sb}) \cdot \vec{\alpha}. \tag{9.8}$$

Note that the wave travels along $\vec{r}_{sb}$ but is incident along the $-\vec{r}_{sb}$ direction, hence the minus sign in the element pattern and directional dependency of the receive antenna. A more general formulation of the attenuation of the wave is employed in (9.8), where the symbol $|\vec{r}_{ms}| \star |\vec{r}_{sb}|$ denotes either $|\vec{r}_{ms}||\vec{r}_{sb}|$ or $|\vec{r}_{ms}| + |\vec{r}_{sb}|$. The first is used if scattering is the dominant process and the second is used when specular reflection dominates. Although deviating slightly from the strict physical motivation, both may be used for channel modeling, depending on what channel properties that are desired [FMB98]. Furthermore, the factor γ determines the attenuation over distance of the wave. Here, a value of $\gamma = 1$, corresponds to free space propagation.

The expression in (9.8) represents the voltage due to the scattered field from one object. Here, the field resulting from the direct path from mobile to base station is not included. If the direct path is needed, the expression in (9.8) should be combined with an additional term containing the field transmitted from the mobile that induces the voltage at the base station. By

combining (9.1) and (9.6), the voltage from the direct path becomes

$$v_d = C\frac{e^{-jk|\vec{r}_{mb}|}}{|\vec{r}_{mb}|^{\gamma}}G_t(\vec{r}_{mb})G_r(-\vec{r}_{mb})\vec{g}_r(-\vec{r}_{mb})\cdot\vec{g}_t(\vec{r}_{mb}), \tag{9.9}$$

where $\vec{r}_{mb}$ is the vector from the mobile to the base station. The contribution of v_d to the total voltage can be large if the scattering parameter $\vec{\alpha}$ is small. In the following, the direct wave will not be included.

A shift in the frequency of the received signal, i.e. the Doppler shift, will occur if the mobile is moving [Jak74, Sau99]. This effect is implicitly included in both (9.8) and (9.9) which can be illustrated by considering a moving mobile. The movement of the mobile will result in a change in the vector $\vec{r}_{ms}$. If we assume a reasonably large distance between the scatterer and the mobile, the vector can be approximated as [BD91]

$$k|\vec{r}_{ms}| = k|\vec{r}_s - \vec{r}_m| \approx k|\vec{r}_s - \vec{r}_{m_o}| + \nu_d t, \tag{9.10}$$

where $\nu_d = kv_s\cos\zeta$ is the Doppler frequency, ζ is the angle between the velocity of the mobile and the vector $\vec{r}_{ms}$. Note also that the speed and the original position of the mobile are denoted v_s and $\vec{r}_{m_o}$. Thus, the expressions in (9.8) and (9.9) includes fundamental properties such as the Doppler shift. In the following the implicit form of (9.8) and the explicit form of (9.10) will be used alternatively.

The total received field of a typical radio channel that has many scattering objects is obtained, as previously alluded to, by summing the contributions from each individual propagation path. Using the results above, it is now possible to form a frequency domain transfer function for a scenario with several scatterers, one mobile antenna, and a receiving antenna at the base station. Using (9.8) gives the following transfer function when N_s scatterers are included

$$h(\omega,t) = \sum_{q=1}^{N_s} C\frac{e^{-jk(|\vec{r}_{ms_q}|+|\vec{r}_{s_qb}|)}}{(|\vec{r}_{ms_q}|\star|\vec{r}_{s_qb}|)^{\gamma}}G_t(\vec{r}_{ms_q})G_r(-\vec{r}_{s_qb})\vec{g}_r(-\vec{r}_{s_qb})\cdot\vec{\alpha}_q, \tag{9.11}$$

Note that the vectors and the distances in (9.11) depend on time t and that all antenna parameters, k, and the scattering dyad depend implicitly on the frequency ω. Although the expression for the voltage in (9.8) is somewhat complicated, it consists of relatively simple functions of the antennas, the propagation distance, and the properties of the scattering objects. However, the fact that the direction of propagation as well as the directions of the fields are included, results in a more realistic model of the spatial domain than just working with scalars. Also, the impact of polarization is correctly accounted

for by calculating both the direction and the amplitude of the fields. Hence, using the above expression, it is possible to examine different types of antenna arrangements and polarization effects. This is particularly desirable in MIMO channels where some physical insight is useful when designing systems and evaluating receiver algorithms.

Note that the channel model is time-variant if the mobile, base station, or any of the scatterers are moving. To reduce the notational complexity only movement of the mobile will be considered here. Furthermore, the Doppler shift is included in (9.11) in its implicit form to further simplify the notation.

The expression in (9.11) was derived for a single transmit and receive antenna. It is straightforward to extend the scalar expression above to vector and matrix expressions for the more general cases of antenna arrays at either the base station or the mobile or at both. In Section 9.5, an expression for a SIMO channel is presented that takes into account an array at the base station (or mobile). The general case of multiple antennas at both transmitter and receiver is treated in Section 9.6. First, the scattering function $\vec{\alpha}$ will be examined in some detail in the next section.

9.3 Electromagnetic Scattering

A simple radio channel model based on EM wave propagation was introduced in the previous section. The received voltage was obtained by adding the contribution from many propagation paths, where each path involved a scattered wave. However, the actual properties of the scattering objects was not discussed. In this Section, the scattering function $\vec{\alpha}$ will be analyzed in detail. First, two objects of simple shape will be analyzed which will give some insight into the scattering process. This insight is then used to formulate a general scattering function that captures the most important scattering properties of scattering objects.

9.3.1 Objects of Simple Shapes

For objects of simple shapes, it is possible to solve the scattering problem and obtain an analytical expression of the scattered field. Many of these analytical results were developed rather early when no computers were available and the value of analytical results was greater. In the beginning of the last century, analytical results for many objects of simple shapes were derived. In the following sections, the field scattered by two simple objects, the sphere and the cylinder, will be examined.

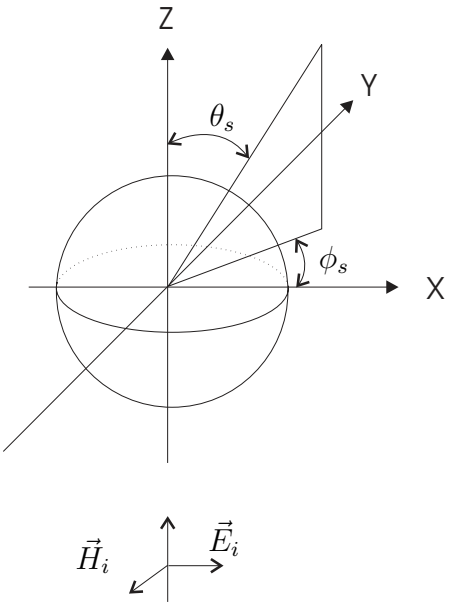


Figure 9.3: The geometry of the sphere and the orientation of the fields.

Scattering by Spheres

The scattering of a plane wave incident upon a dielectric sphere was one of the first the scattering problems that was solved. For detailed derivations of the scattering calculations, see [Bal89, BSea87, KCH71, Mie08]. Consider a sphere of radius a located at the origin, see Figure 9.3. A uniform plane wave polarized in the $\vec{x}$ direction, traveling along the $\vec{z}$ axis, is incident upon the sphere. The incident electric field can be written as

$$\vec{E}^i = \vec{x} E_0 e^{-jkz}, \tag{9.12}$$

where k denotes the wave number. For objects with a geometry coinciding with a coordinate system, the scattered field can be found using modal techniques. In the case of a sphere, the incident plane wave is expressed in spherical wave functions. The scattered field is expressed in similar functions and the fields are then matched at the surface of the sphere using boundary conditions. The scattered field at a point (r, θ_s, ϕ_s) that is sufficiently far from the sphere, i.e. the far-field expressions, then becomes

$$\vec{E}^s(r, \theta_s, \phi_s) = E_0 \frac{e^{-jkr}}{r} \vec{\alpha}. \tag{9.13}$$

The expression for the scattering function $\vec{\alpha}$ is quite complicated and is given in Appendix 9A.

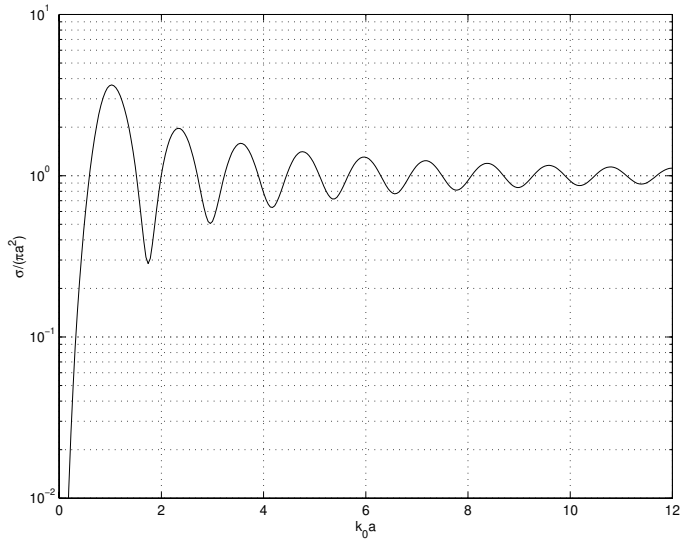


Figure 9.4: Normalized radar cross section of a perfectly conducting sphere.

The scattering function $\vec{\alpha}$ is closely related to another quantity that commonly is used in scattering analysis, namely the Radar Cross Section (RCS) [Ruc70]. The RCS simply relates the scattered energy to the incident energy and can actually be expressed using $\vec{\alpha}$ as

$$\sigma = 4\pi|\vec{\alpha}|^2. \tag{9.14}$$

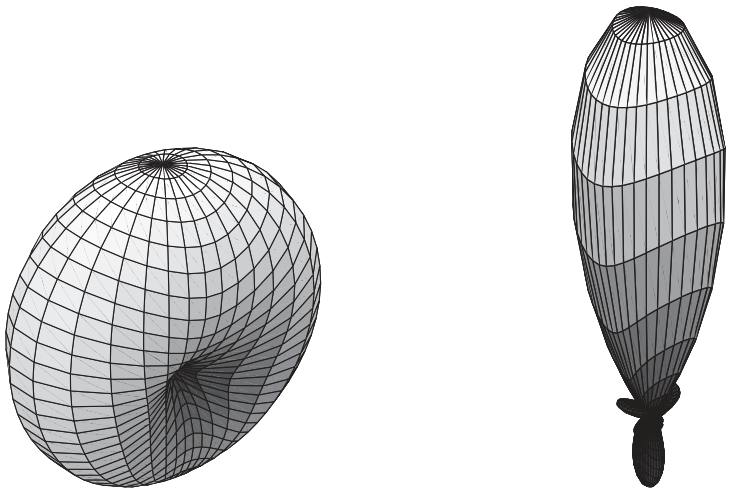
For a thorough study of the RCS of the sphere and many other objects, see [BSea87, Ruc70, CS68]. Also, the RCSs of spheroids and crushed rocks are studied in [NS96b]. It is interesting to examine the frequency behavior of $\vec{\alpha}$ or σ since these functions will have a large impact on the channel model. Remember that the expressions for the received voltage in Section 9.2 were derived in the frequency domain, and that an impulse response can easily be obtained via the inverse Fourier transform of these expressions. Thus, the frequency behavior of the scattering function $\vec{\alpha}$ will partly determine the impulse response.

The normalized RCS of a perfectly conducting sphere is shown in Figure 9.4 as a function of frequency. Note that the frequency axis is normalized as $k_0 a = 2\pi f a/c$ where c denotes the speed of light and a denotes the radius of the sphere. For a sphere of radius 0.1m, the frequency range in Figure 9.4 is 0-5.7 GHz. Three different regions are clearly visible in Figure 9.4. For low frequencies, the RCS rapidly increases ($\sim f^4$) with increasing frequency. This region is usually called the Rayleigh region. In this region, the dimension

of the object is small compared to the wavelength. At higher frequencies, the wavelength is comparable to the dimensions of the object which gives rise to some resonances in the RCS. Therefore, this region is often called the resonance region or the Mie region. As the frequency increases even further, the dimensions of the object become much larger than the wavelength and the resonances decay with increased frequency. For objects with dimensions much larger than the wavelength, it is possible to use optical theory. Thus, this region is often called the optical region. Note that the material of the scattering object will of course affect the RCS, but the three different regions will still be distinct. In Figure 9.4, the RCS of a perfectly conducting sphere was shown. If a dielectric sphere would have been examined, the results would have been similar but with other amplitude levels and a more noisy shape of the curve.

Future wireless communication systems will be able to support high data rates and require significantly more bandwidth than present systems. Hence, these systems are often called broadband systems. However, in terms of the physical bandwidth, i.e. the bandwidth relative to the carrier frequency, these systems can still be regarded as narrowband systems. For example, consider a system such as HiperLAN/2 [KJSWW99] which operate on the 5GHz band with a channel bandwidth of 20MHz. The difference in scattering properties for the sphere considered in Figure 9.4, in this frequency band, is negligible since the above bandwidth corresponds to values of k_0a in the range 10.47-10.51. In this range, the changes in backscattered energy are very small. Thus, in this respect the bandwidth is quite narrow which will prove useful in Section 9.4 where the impulse response of the channel model is derived.

Of course, the radius of the sphere does not only affect the amount of scattering but also the directional properties of the scattered field. In Figure 9.5, the magnitude of the scattered field as a function of the scattering direction is shown. Here, a plane wave is incident upon a sphere with a dielectric constant (relative permittivity) $\epsilon_r = \epsilon_1/\epsilon_0 = 4$. Two different cases are considered; Rayleigh scattering with a sphere of size $k_0a = 0.42$ and Mie scattering with a sphere of size $k_0a = 4.2$. For the case of the smaller sphere, the wavelength is large compared to the radius of the sphere. In this case, the sphere scatters almost isotropically and thus acts as a point source. However, in the second case, the radius of the sphere is comparable to the wavelength and the scattering is more confined to the forward direction. In the optical region, the beamwidth of the scattered field becomes even narrower. Thus, as the size of the scattering object is increased, the scattered energy comes more focused to one scattering direction. This property will be used later in Section 9.3.2, where a general scattering function will be derived.



(a) $k_0a = 0.42$ (b) $k_0a = 4.2$

Figure 9.5: Scattering from a dielectric sphere with $\epsilon_r = 4$ and varying size parameter k_0a when a wave is incident from below.

Another interesting observation from Figure 9.5 is that the forward scattering is much more pronounced than the backscattering. In fact, in most cases, the forward scattering term dominates for dielectric spheres. However, for a perfectly conducting sphere the backscattered energy dominates. For more details on scattering from spheres, see [BSea87, CS68, KCH71, Kri99, dH57]. Next, the scattering properties of finite cylinders will be examined.

Scattering by Finite Cylinders

The scattering of a plane wave incident obliquely upon a dielectric circular cylinder results in more complicated expressions than those for the sphere. Consequently, expressions for the scattered field appeared in the literature significantly later than for the sphere [Wai55]. Due to the complexity of the exact solution and the fact that it is only valid for infinite cylinders, a number of approximative solutions have appeared in the literature. Earlier work may be found in [PK83, Ric65] that gives numerical results for various objects. More recent work includes dyad and tensor approximative formulations [SS90a, SS96]. Scattering from dielectric and conducting cylinders is also treated in some textbooks on scattering and radar cross section analysis [BSea87, Ruc70, CS68].

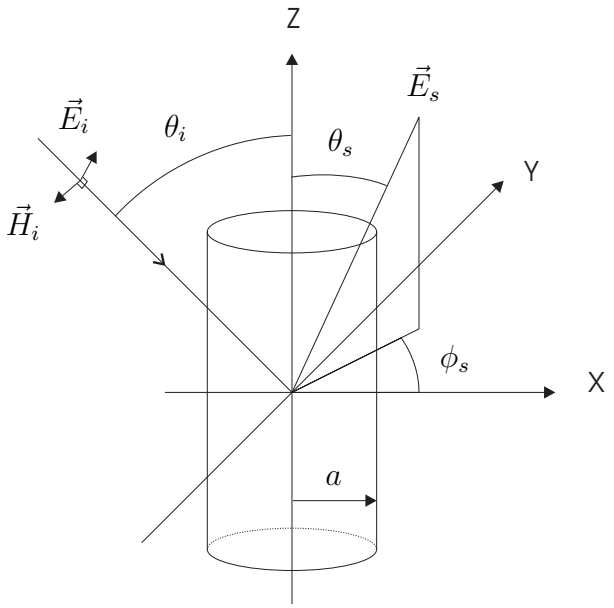


Figure 9.6: Cylinder Geometry

Consider a uniform plane wave incident obliquely upon a cylinder, see Figure 9.6. Here, θ_i denotes the elevation incidence angle while θ_s and ϕ_s denotes the desired scattering direction. The length of the cylinder is denoted l and the radius is denoted a. The scattered field is found using similar techniques as for the sphere. The solution is more complicated since the cylinder is of more complicated geometry. Furthermore, the final solution is only valid for infinite cylinders. However, an approximation is possible that is sufficiently accurate for channel modeling. Using those approximations for the finite cylinder along with the original solutions then gives the far-field expressions for the scattered field in a point (r, θ_s, ϕ_s), see Figure 9.6, as

$$\vec{E}^s(r, \theta_s, \phi_s) = E_0 \frac{e^{-jkr}}{r} \vec{\alpha}. \tag{9.15}$$

The expression for the scattering function $\vec{\alpha}$ is even more complicated that for the sphere and is given in Appendix 9B.

The scattering parameter for the dielectric cylinder is related to the RCS in the same manner as the sphere, see (9.14). The normalized radar cross section of a dielectric cylinder as a function of the size parameter $k_0 a$ is shown in Figure 9.7. An axial ratio $l/a \gg 1$ is used in order for the approximative solution to be valid. Furthermore, a cylinder of lossy dielectric ($\epsilon_r = 2.56 - i0.1024$) is considered in this example. The three different re-

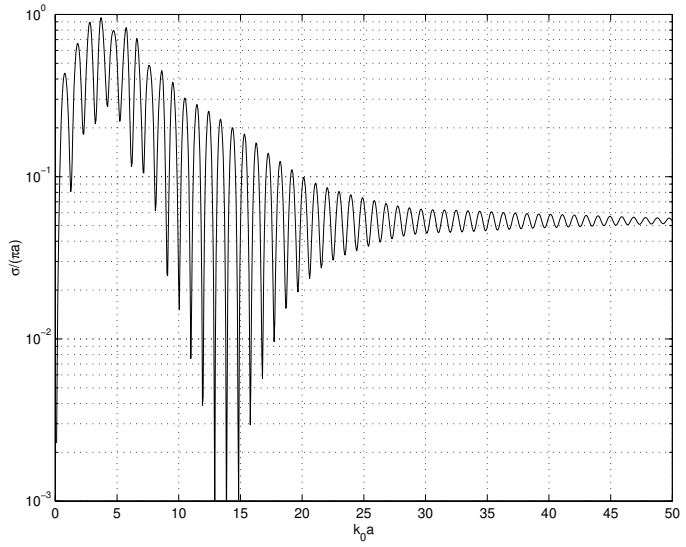
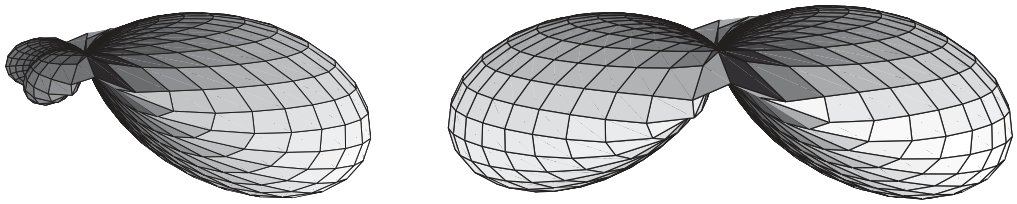


Figure 9.7: Normalized radar cross section of a dielectric cylinder with $\epsilon_r = 2.56 - i0.1024$ (TMz).

gions, the Rayleigh, Mie and optical region are again clearly visible in Figure 9.7. Note that this RCS figure exhibits more rapid variations than the corresponding figure for the sphere, see Figure 9.4. This is partly due to the more complicated geometry and partly due to the fact that the sphere is perfectly conducting while the cylinder is of lossy dielectric material.

It is also interesting to note that the scattering from a cylinder, in a manner similar to the sphere, may be considered as narrowband. Returning to the HiperLAN/2 example and a cylinder radius of 0.1m, the frequency range in Figure 9.7 is 0-23.9GHz. Thus, a HiperLAN/2 channel corresponds to the range 10.47-10.51, as for the sphere, in which the RCS remains almost unchanged. In this respect, the scattering from a cylinder can be considered to be of narrowband type.

The directional properties of the scattered field of a dielectric cylinder are illustrated in Figure 9.8. Here, the axial ratio is $l/a = 10$, the size parameter is $k_0 a = 0.1$, and two values of the relative permittivity $\epsilon = 4$ and $\epsilon = 4000$ are considered. Since the size of the cylinder is of comparable size as the wavelength, the width of the scattering lobe is rather broad. If the cylinder would have been smaller, close to isotropical radiation would have been obtained. On the other hand, if the cylinder would have been much larger than the wavelength, a very narrow scattering lobe would have been obtained. Thus, the scattering characteristics are, in this respect, similar to



(a) $\epsilon_r = 4$ (b) $\epsilon_r = 4000$

Figure 9.8: Scattering from a dielectric cylinder with size parameter $k_0 a = 0.1$ and varying relative permittivity ϵ_r when a wave is incident from the left.

those of the sphere. In fact, the same is true for most scattering objects.

Furthermore, for the case of $\epsilon_r = 4$ forward scattering dominates while two scattering lobes are obtained for the case of $\epsilon_r = 4000$. These type of results have been verified by several authors [CS68, PK83, Ruc70]. Note that scattering objects with very high ϵ_r values show similar scattering properties as perfectly conducting objects. Thus, also in this respect do the cylinder and the sphere exhibit similarities. The width of the scattering lobes in elevation is due to the fact that the cylinder is finite. For an infinite cylinder, the energy is scattered along a conical surface with half-angle θ_i. As the length of the cylinder becomes smaller, the energy will be spread around this cone, as seen in Figure 9.8.

Another interesting aspect of the scattering properties of the cylinder is that the dielectric cylinder depolarizes the signal [Bal89]. Depolarization means that the scattering introduces additional components in the scattered field as compared to the incident field. This depolarization occurs only for dielectric cylinders and not for smooth perfectly conducting cylinders. The symmetry of the scattering object as well shape of the surface determines the polarization properties of the object [Bal89, BS63, Kri99]. For instance, the symmetry of the sphere assures no depolarization of the scattered field.

The polarization of the scattered wave is, along with the amplitude and directional dependency, the main important properties of the scattering process. The next section will introduce a general scattering model that is based on the typical scattering properties found in this section.

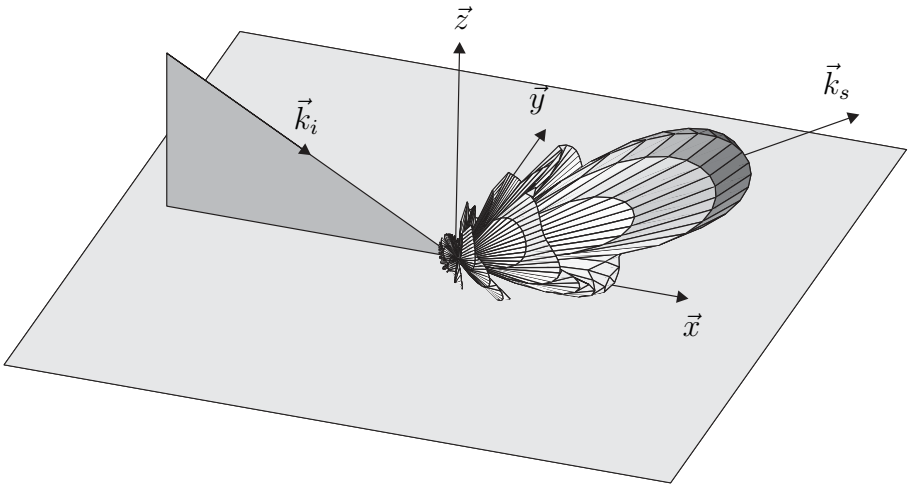


Figure 9.9: Scattering from a dielectric plate with oblique incidence.

9.3.2 General Scattering Model

In the previous section, the scattering properties of objects of simple shapes such as the sphere and the cylinder were examined. The most important characteristics were the width of the scattering lobe, the polarization, and the frequency dependency. However, the amount of scattered energy is also of great importance. The backscattered energy is mainly dependent on the size of the scattering object. For instance, a large building wall will obviously contribute more to the total scattering environment than a small lamp post. Even in an indoor environment the scattering from the walls of the building plays an important part, although not as dominant as in the outdoor scenario. Therefore, is is interesting to examine the scattering from walls of finite extent. Also, it will show that the scattering from walls depend quite strongly on the shape of the surface. In particular, walls with a non-smooth surface, i.e. a rough surface, will show quite a different scattering behavior than that of a smooth wall or plane.

Scattering from a wall may accurately be described by considering the classical electromagnetic problem of plane wave scattered by a flat rectangular plate. Explicit solutions of this problem may be found in many textbooks [Bal89, Kri99]. These solutions rely on approximations such as Physical Optics (PO) or Geometrical Theory of Diffraction (GTD). For the purpose of obtaining a simple channel model, these approximations are satisfactory.

An example of scattering by a dielectric plate with oblique incidence is shown in Figure 9.9. Here, $\vec{k}_i$ and $\vec{k}_s$ denote the normalized incidence and

scatter wave vectors ($|\vec{k}_i| = |\vec{k}_s| = 1$) and $\vec{x}, \vec{y}$ and $\vec{z}$ denote the axis in the rectangular system of coordinates. In fact, throughout this section, the wave vectors will be assumed to be normalized. Note also that the plane of incidence, i.e. the plane spanned by $\vec{k}_i$ and the surface normal vector $\vec{n} = \vec{z}$, is shaded in dark gray. The width of the scattering lobe depends strongly on the size of the plate. For a plate of infinite extent, the width would be infinitely small. If on the other hand, the plate is very small, the energy will scatter in all directions and appear as a point source. Thus, the size of the plate essentially determines the width of the scattering lobe. This is in agreement with the results from the scattering analysis of spheres and cylinders. However, note that the size of the plate, as well as the size of the spheres and cylinders, also determines the magnitude of the scattered field.

Based on the above analysis, it is easy to draw the conclusion that large buildings would only scatter in a very small angle and therefore contribute only for a very narrow spatial section. However, there are two other factors that affects the total scattering. Since the building is large, it is typically possible to find a point where the directions of the incident wave and the scattered or reflected wave fulfills Snell's law [Bal89] and specular reflection will dominate. Furthermore, the surface of a building will not be perfectly smooth which affects the scattering properties significantly. An often used criterion for a surface to be smooth is that the height difference Δh between two points on the surface, should obey [BS63]

$$\Delta h < \frac{\lambda}{8 \cos \theta_i}, \tag{9.16}$$

where λ denotes the wavelength and θ_i the incidence angle relative the surface normal. This criterion for a surface to be smooth or rough is usually called the Rayleigh criterion. If the surface is rough, the energy in the specular direction will decrease while increasing the scattered energy in other directions. The reason is that a rough surface will have a large number of scattering points on the surface with different surface normals that will scatter the energy in many different directions. Thus, the scattered lobe in Figure 9.9 could also be interpreted as scattering from an infinite plate with a rough surface.

An interesting property of rough surface scattering is that the depolarization phenomenon is determined by the incidence and scatter directions. In fact, it can be shown [BS63, Bla00] that the field scattered by a rough surface is depolarized in the same way as the field reflected by a plane inclined in such a way that the specular direction coincides with the scatter direction $\vec{k}_s$. For more details regarding the depolarization of rough scattering, see [BS63, Bla00].

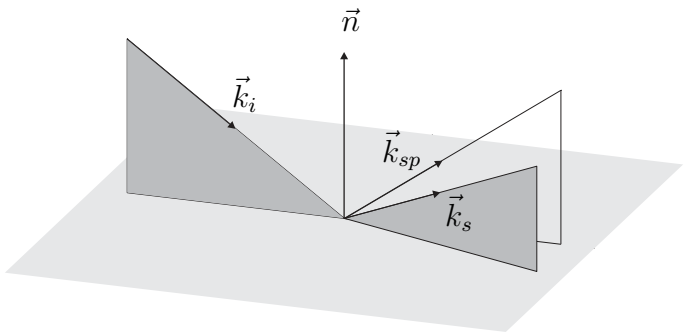


Figure 9.10: A wave incident upon the surface of a scatterer when the desired scattering direction does not coincide with the specular scattering direction.

A simple model of scattering by arbitrary objects can be obtained by employing the results from rough surface scattering and the scattering by a dielectric plate. The polarization of the scattered field is obtained by employing the depolarization result from rough surface scattering while the width of the scattering lobe is obtained by relating the size of the scattering object to a dielectric plate. The resulting model is then based on fundamental electromagnetic theory while still being relatively straightforward to apply.

Consider the scattering scenario in Figure 9.10, where a wave is incident along the $\vec{k}_i$ direction upon an object with surface normal $\vec{n}$. The specular direction that fulfills Snell's law is indicated as $\vec{k}_{sp}$, and the desired scattering direction is denotes $\vec{k}_s$ as usual. In Figure 9.10, the incidence plane and the scattering plane are shaded using a slightly darker gray than the surface of the scattering object. Based on the above observations, the scattered field is modeled as

$$\vec{E}_s = \rho \vec{\mathbf{S}}_p \cdot \vec{E}_i, \tag{9.17}$$

where $\vec{\mathbf{S}}$ denotes a scattering dyad[3] or a scattering matrix [Bal89, FM94]. In (9.17), the depolarization is determined by the dyad $\vec{\mathbf{S}}$ while the shape and the width of the scattering lobe is determined by ρ. Note that both these quantities are functions of the incidence and scattering directions as well as the orientation of the scattering objects.

An expression for the scattering dyad $\vec{\mathbf{S}}$ is easily obtained by employing the result from rough surface scattering that states that the polarization of

[3]A dyad is defined by the juxtaposition $\vec{a}\vec{b}$ of the vectors $\vec{a}$ and $\vec{b}$, with no dot product or cross product between them. In matrix form this corresponds to the outer product of two vectors $\mathbf{a}$ and $\mathbf{b}$, i.e. $\mathbf{a}\mathbf{b}^T$ (assuming column vectors). Thus, a dyad is simply a linear transform [Bal89].

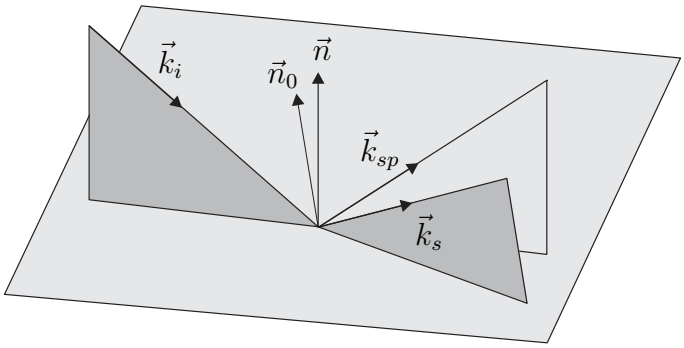


Figure 9.11: Illustration of scattering geometry and the orientation of the vectors.

the scattered field can be obtained from the polarization of a wave incident upon a plane inclined in such a way that the specular direction of that plane coincides with the desired scattering direction. In Figure 9.11, a simple illustration of the inclination of such a plane is shown. Here, $\vec{n}$ denotes the original surface normal and $\vec{n}_0$ denotes the surface normal of the inclined plane. Note that $\vec{k}_{sp}$ denotes the specular direction of the original plane and that $\vec{k}_s$ now denotes both the desired scattering direction and the specular reflection direction of the inclined plane. The new surface normal is easily obtained as

$$\vec{n}_0 = \frac{\vec{k}_s - \vec{k}_i}{||\vec{k}_s - \vec{k}_i||}, \tag{9.18}$$

where the denominator assures a unit length surface normal. Furthermore, the incidence angle (and now also the scattering angle) becomes

$$\vec{k}_s \cdot \vec{n}_0 = -\vec{k}_i \cdot \vec{n}_0 = \cos\theta_0. \tag{9.19}$$

It is now straightforward to obtain an expression for the scattered field for the scenario illustrated in Figure 9.11, and formulas may be found in many textbooks [Bal89, Bla00, FM94]. Unfortunately, these textbooks do not employ dyad notation. A more convenient expression may be found in [Foc65], where a coordinate independent formulation of the scattering problem of a plane wave incident obliquely upon an infinite dielectric plate was obtained using a dyad notation. Using the notation in Figure 9.11 and (9.17)-(9.19), the expression for the scattering dyad becomes

$$\vec{\mathbf{S}}_p = \frac{-1}{\sin^2\theta_0}\left[R_1(\vec{n}_0\cos 2\theta_0 + \vec{k}_i\cos\theta_0)\vec{n}_0 - R_2(\vec{n}_0 \times \vec{k}_i)\vec{n}_0 \times \vec{k}_i\right], \tag{9.20}$$

where the coefficients R_1 and R_2 depend on the material of the dielectric as

$$\begin{aligned} R_1 &= \frac{\cos\theta_0 - \sqrt{\frac{\epsilon_1}{\epsilon 2}}\sqrt{1 - \frac{\epsilon_1}{\epsilon_2}\sin^2\theta_0}}{\cos\theta_0 + \sqrt{\frac{\epsilon_1}{\epsilon 2}}\sqrt{1 - \frac{\epsilon_1}{\epsilon_2}\sin^2\theta_0}} \\ R_2 &= \frac{\cos\theta_0 - \sqrt{\frac{\epsilon_2}{\epsilon 1}}\sqrt{1 - \frac{\epsilon_1}{\epsilon_2}\sin^2\theta_0}}{\cos\theta_0 + \sqrt{\frac{\epsilon_2}{\epsilon 1}}\sqrt{1 - \frac{\epsilon_1}{\epsilon_2}\sin^2\theta_0}} \end{aligned}. \tag{9.21}$$

Here, ϵ_1 and ϵ_2 denote the permittivity of the surrounding medium and the dielectric of the plate, respectively. These coefficients are usually called reflection coefficients in the literature. An interesting special case is a perfectly conducting plane, where the coefficients become $R_1 = 1$ and $R_2 = -1$. The scattering dyad in (9.20) can also be expressed in matrix form as

$$\mathbf{S}_p = \frac{-1}{\sin^2\theta_0}\left[R_1(\mathbf{n}_0\cos 2\theta_0 + \mathbf{k}_i\cos\theta_0)\mathbf{n}_0^T - R_2\mathbf{M}\mathrm{vec}(\mathbf{n}_0\mathbf{k}_\mathrm{i}^\mathrm{T})\mathrm{vec}^\mathrm{T}(\mathbf{n}_0\mathbf{k}_\mathrm{i}^\mathrm{T})\mathbf{M}^\mathrm{T}\right], \tag{9.22}$$

where a matrix formulation of the vector cross product $\vec{v} = \vec{a} \times \vec{b}$ is used

$$\mathbf{v} = \mathbf{M}\ \mathrm{vec}(\mathbf{a}\mathbf{b}^\mathrm{T}). \tag{9.23}$$

The selection matrix $\mathbf{M}$ is defined as

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{9.24}$$

Now, two equivalent expressions for the depolarization have been given in (9.20) and (9.22). The dyad formulation is frequently used in the propagation literature, but the matrix formulation may appear more intuitive to researchers in the fields of signal processing and wireless communications. Next, an expression for the model parameter ρ in (9.17) that determines the shape and width of the scattered field will be presented.

Obviously, the expression for ρ depends strongly on the shape of the scatterer. Expressions are readily available in the literature for objects such as flat rectangular plates or cylindrical plates [Bal89, BSea87, BS63]. However, the aim is to design a single function that is able of modeling most scatterers. However, most scatterers exhibits a maximum in the specular direction and decrease smoothly as the desired scattering direction diverges from the specular direction, in the manner observed in Figures 9.5, 9.8, and 9.9. A simple function that produces this type of behavior is

$$\rho = \mathrm{sinc}(a||\vec{k}_s - \vec{k}_{sp}||), \tag{9.25}$$

where a denotes a size related parameter that determines the rate of decrease as the scattering direction diverges from the specular direction and $\text{sinc}(x) = \sin(x)/x$. By using the relation

$$\vec{k}_{sp} = \vec{k}_i - 2(\vec{k}_i \cdot \vec{n})\vec{n} \tag{9.26}$$

equation (9.25) can also be written as

$$\rho = \text{sinc}(a||\vec{k}_s - \vec{k}_i + 2(\vec{k}_i \cdot \vec{n})\vec{n}||), \tag{9.27}$$

where $\vec{n}$ as previously denotes the surface normal of the scattering object.

The model for scattering by arbitrary objects can, using the above results, finally be formulated as

$$\vec{\mathbf{S}} = A\rho\vec{\mathbf{S}}_p, \tag{9.28}$$

where ρ is given in (9.25) and either of the two formulations (9.20) and (9.22) may be used. The parameter A is included to account for the size of different objects or the amount of scattered energy. The scattering dyad in (9.28) depends on the parameters $\vec{k}_i$, $\vec{k}_s$, $\vec{n}$, ϵ_2, a, and A. To reduce the number of model parameters, the special case of a perfectly conducting object will be used in the following. Thereby, the number of parameters that models the scattering dyad are reduced to three. In Section 9.4, the impact of different choices of these parameters on the channel impulse response will be discussed. These parameters may be modeled as stochastic or deterministic, depending on the scenario. Here, measurements may be used to decide upon the distributions of the parameters.

9.4 SISO Channel Modeling

The impulse response of a channel with a single transmitting and a single receiving antenna, i.e. an Single-Input Single-Output (SISO), will be derived in this section using the results from Section 9.2 and Section 9.3. The spatio-temporal properties of the channel model are then examined and discussed in a few channel scenarios.

9.4.1 Impulse Response

The frequency domain transfer function for a scenario with several scatterers was derived in Section 9.2. The form of this transfer function may be simplified by using the general scattering model introduced in Section 9.3. In particular, it is possible to relate the scattering function $\vec{\alpha}$ to the transmitted field in more direct way. Since, the scattering function $\vec{\alpha}$ denotes the

polarization of the scattered field it can be rewritten (with a slight abuse of notation) using (9.17) as

$$\vec{\alpha} = \vec{E}_s = \vec{\mathbf{S}}\vec{E}_i = \vec{\mathbf{S}}\vec{g}_t, \tag{9.29}$$

where the polarization of the incident field $\vec{E}_i$ is expressed using the polarization of the transmitted field from the antenna $\vec{g}_t$. Note that $\vec{g}_t$ depends on the direction of radiation as well as the antenna type. Furthermore, the relationship to the radar cross section can now be written as

$$\sigma = 4\pi|\vec{\mathbf{S}}\vec{g}_t|. \tag{9.30}$$

A more compact form of the frequency domain transfer function may be written using (9.29)

$$h(\omega, t) = \sum_{q=1}^{N_s} C\frac{e^{-jk(|\vec{r}_{ms_q}|+|\vec{r}_{s_qb}|)}}{(|\vec{r}_{ms_q}| \star |\vec{r}_{s_qb}|)^\gamma} G_t(\vec{r}_{ms_q})G_r(-\vec{r}_{s_qb})\vec{g}_r(-\vec{r}_{s_qb})\vec{\mathbf{S}}_q\vec{g}_t(\vec{r}_{ms_q}), \tag{9.31}$$

where C is given in (9.7), $\vec{\mathbf{S}}$ in (9.28), and the antenna parameters $G_t, G_r, \vec{g}_t$, and $\vec{g}_r$ are defined in Section 9.2. By selecting the antennas and determining the properties of the scatterers and their positions, the frequency domain transfer function is easily calculated using (9.31). Throughout this section, $\star$ denotes a multiplication, and an attenuation factor of $\gamma = 1$ will be used.

A time domain transfer function is typically employed when analyzing and simulating the wireless channel. Fortunately, it is straightforward to obtain a complex baseband channel impulse response from the frequency domain transfer function (9.11) by performing an inverse Fourier transform. First, several things should be noticed. Even if a wideband system is considered, the bandwidth compared to the carrier frequency can typically be considered narrow. For instance, the HiperLAN/2 standard, that supports high data rates, has a relative bandwidth of less than one percent. Thus, the antenna parameters $C, G_t, G_r, \vec{g}_t$, and $\vec{g}_r$ are almost constant over the signal bandwidth. Furthermore, the scattering dyad $\vec{\mathbf{S}}$ was in Section 9.3 also found to be almost constant over the bandwidth. The remaining term that depends on the frequency is the phase shift that corresponds to the time delay. Thus, for a system with bandwidth ω_b, the complex baseband impulse response can be written as

$$h(\tau, t) = C' \sum_{q=1}^{N_s} \frac{\text{sinc}\left[\omega_b(\tau - \tau_q)\right]}{(|\vec{r}_{ms_q}| \star |\vec{r}_{s_qb}|)^\gamma} G_t(\vec{r}_{ms_q})G_r(-\vec{r}_{s_qb})\vec{g}_r(-\vec{r}_{s_qb})\vec{\mathbf{S}}_q\vec{g}_t(\vec{r}_{ms_q}), \tag{9.32}$$

where $\tau_q = (|\vec{r}_{ms_q}| + |\vec{r}_{s_qb}|)/c$, c denotes the speed of light in the surrounding medium, and

$$C' = -j\frac{4\omega_b}{k_c \eta I_0}, \tag{9.33}$$

where k_c denotes the wave number of the carrier frequency. The impulse response in (9.32) is a function of both the delay τ and the time t, that accounts for the time evolution of the impulse response. One example of time evolution effects is the Doppler shift that occurs due to movement of either the mobile, the base station, or the scatterer. Note that the Doppler is included implicitly in (9.32), but explicit expressions were given in Section 9.2. Furthermore, the impulse responses given above is the baseband impulse response where the carrier frequency is suppressed.

The functions given in (9.31) and (9.32) are two characterizations of the wireless channel which are related through the Fourier transform with respect to the delay τ. Another set of functions can be obtained by Fourier transforming with respect to the time t. Altogether, four different functions can be obtained by Fourier transforming with respect to both time and delay. This characterization is known as the family of Bello functions, after their originator [Bel63]. For example, by taking the Fourier transform of (9.32) with respect to the time t results in the delay Doppler-spread function that is widely used to determine the Doppler spread. This function will be calculated for a scattering scenario in Section 9.4.2.

The constant C' can of course be omitted when simulating a particular type of system, since it is just a constant. However, when studying the frequency dependency it should be included, and care must be taken not to violate the narrowband assumption used to arrive at (9.32).

The expression for the impulse response may be simplified even further, by evoking another narrowband assumption. If the bandwidth ω_b of the system is small compared to the spread in arrival times $\Delta\tau$, i.e. $\omega_b \ll 2\pi/\Delta\tau$, the different time delays corresponding to each scatterer (or path) can then be treated as simple phase shifts as

$$h_{nb}(t) = \sum_{q=1}^{N_s} \frac{e^{-j\omega_c\tau_q}}{(|\vec{r}_{ms_q}| \star |\vec{r}_{s_qb}|)^\gamma} G_t(\vec{r}_{ms_q}) G_r(-\vec{r}_{s_qb}) \vec{g}_r(-\vec{r}_{s_qb}) \vec{\mathbf{S}}_q \vec{g}_t(\vec{r}_{ms_q}), \tag{9.34}$$

where the constant C' has been omitted and ω_c denotes the carrier angular frequency. Thus, it is important to separate the two different narrowband assumptions. The first narrowband assumptions is regarding the bandwidth of the system relative the carrier frequency, while the second narrowband assumption regards the system bandwidth relative the delay spread of the

channel. Thus, a system may be narrowband in the first sense, but not in the second sense.

It is possible to relate the second narrowband assumption to a spread in path length Δl instead of delay time $\omega_b \ll 2\pi/\Delta\tau = c/\Delta l$. For example, a system of bandwidth 20MHz at a carrier frequency of 5.2GHz can be considered narrowband in the second sense if the difference in path length is less than 15m, while it is still narrowband in the first sense. For a system with a bandwidth of 1MHz, the corresponding distance becomes 300m. However, for all reasonable communication systems it is always possible to apply (9.32).

In (9.32) and (9.34), a continuous time representation has been used. However, the major part of the signal processing and communication theory is performed using a discrete time representation of the channel. Fortunately, a discrete impulse response can be obtained by sampling the functions given above [Pro95]. An example of how this can be performed in a MIMO setting is given in [SJBF00], where also the pulse shaping filter response of the communication system is included. Next, some of the most important properties of the channel model will be illustrated in a few simulation examples that provide some insight into the different physical phenomena.

9.4.2 Spatio-Temporal Properties

The spatio-temporal properties of the channel model will be examined in a few examples in this section. First, a very simple example of a single scatterer will be considered to illustrate the general properties of the model and the scattering dyad $\vec{\mathbf{S}}$. Next, a more complicated example with many scattering elements with different polarization characteristics will be examined.

Consider the channel scenario illustrated in Figure 9.12, where only a single scatterer is present. The scatterer is oriented as $\vec{n} = -\vec{y}$, and the mobile travels along a path in the $\vec{x}$ direction with a speed of 10m/s. The base station is situated beneath the scatterer and the mobile trajectory. The scatterer is assumed to be perfectly conducting, and the scattering amplitude is $A = 3$. At both the base station and the mobile, $\vec{z}$ oriented dipole antennas are used (see (9.2) and (9.3)). Since all objects are situated in the $x-y$ plane, all wave propagation is horizontal and no polarization losses occur. Several fundamental channel properties can now be illustrated using the channel models introduced above and also explained since the geometry of the channel is simple. Since only one scatterer is present, the delay spread will be zero and the narrowband model in (9.34) may be used.

First, the received power as the mobile moves along the trajectory will be examined. The squared magnitude of the received voltage $|h_{nb}(t)|^2$ ver-

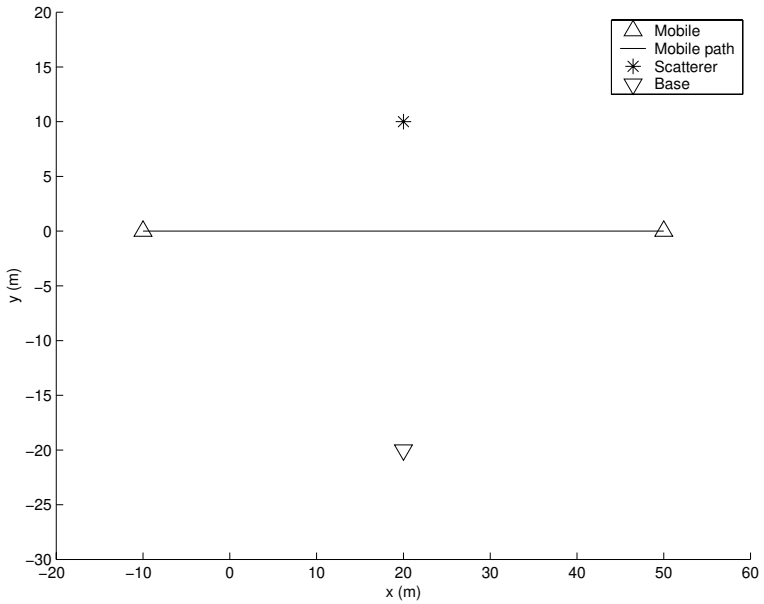


Figure 9.12: Geometry of a simple channel scenario example with only one scatterer oriented as $\vec{n} = -\vec{y}$.

sus time t is shown in Figure 9.13 for three different scatterer sizes. As is intuitively expected, the maximum received power occurs when the mobile is the closest to the scatterer. Also, the impact of different scattering object sizes is examined by using several values for the size parameter a, see (9.25). A large value of the scatterer size parameter a results in a narrow scattering lobe, while a lower value gives a broader area of scattering as expected. This figure may be useful when determining suitable values for a when simulating different channels.

Although the scattered power reveals much of the scattering properties, the Doppler effects are not included since the phase is neglected. In Figure 9.14, the real part of the complex baseband response is shown for the same scenario as previously. Note that this corresponds to the output of a filter matched to a single sinusoid of carrier frequency. Here, the Doppler effect is clearly visible. As the mobile approaches the scatterer, the angle ζ between the velocity of the mobile ($\vec{x}$) and the direction towards the scatterer approaches 90°. Using the results obtained in Section 9.2, the Doppler frequency is $\nu_d = kv_s \cos\zeta$. Thus, the Doppler frequency should be zero at the scatterer, as found in Figure 9.14, since the angle ζ then becomes 90°. The maximum Doppler frequency is consequently expected at the beginning and the end of the mobile path in Figure 9.12. In the simulation of Figure 9.14,

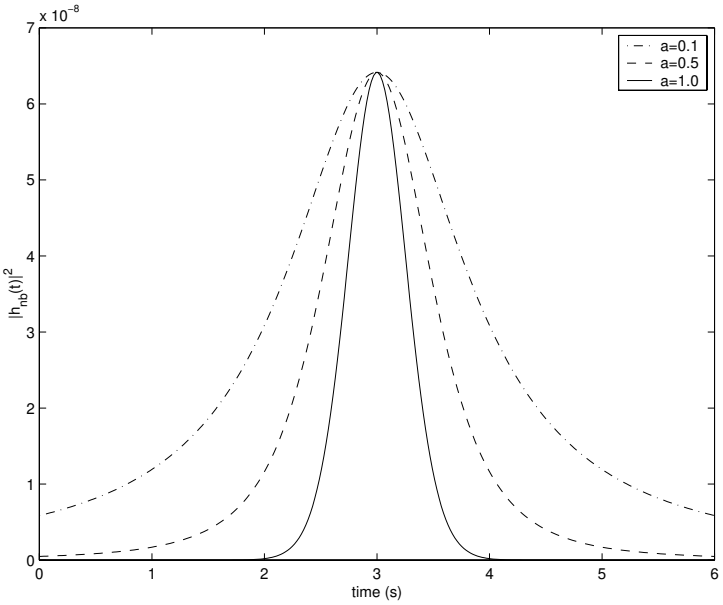


Figure 9.13: Received energy versus time for the scenario in Figure 9.12, and for three different scatterer sizes.

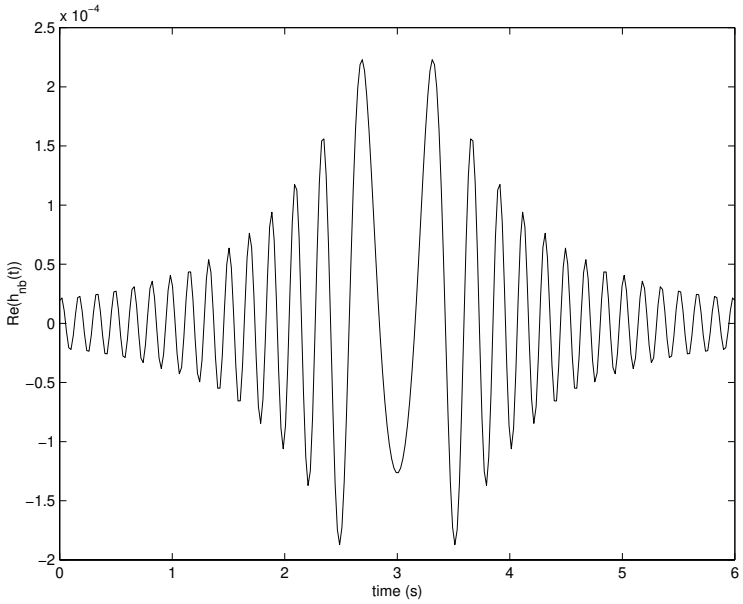


Figure 9.14: The real part of the complex response, $\mathrm{Re}(h_n(t))$ versus time, for the scenario in Figure 9.12 when $a = 0.5$ and $f_c = 200$MHz.

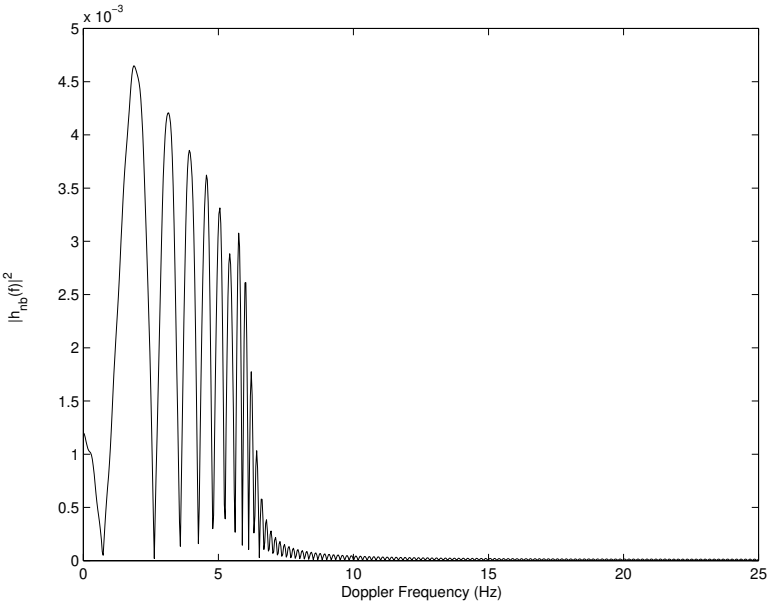


Figure 9.15: The Doppler spectrum for the scenario in Figure 9.12 when $a = 0.5$ and $f_c = 200$MHz resulting in a maximum theoretical Doppler frequency of 6.7Hz.

a carrier frequency of 200MHz was used, which gives a maximum Doppler frequency of 6.7Hz. This agrees reasonably well with Figure 9.14. However, by calculating the Doppler spectrum, i.e. taking the Fourier transform of the signal in Figure 9.14, the Doppler effects are illustrated more explicitly.

The Doppler spectrum for the scenario in Figure 9.12 is shown in Figure 9.15. It is clearly visible that the Doppler frequencies are in the range $0 - 7$Hz, which agrees well with the expected maximum Doppler of 6.7Hz. Thus, the channel model is capable of reproducing the most fundamental characteristics of the wireless channel, such as the Doppler and the time evolution of the channel. However, since only one scattering object was used, no multipath effects were present. Furthermore, no polarization effects were included since all wave propagation was horizontal, and both antennas were vertically polarized. Next, a more complicated scenario that contains some of these aspects will be studied.

Consider the channel geometry in Figure 9.16, where a more complicated scattering scenario is shown. Here, the base station is located at $(0, 0, 20)$ while the mobile travels between $(98, 0, 0)$ and $(102, 0, 0)$, where all values are are given in meters. The mobile moves with a velocity of 2m/s, i.e. the total travel time is 2s. The mobile dipole antenna is oriented as $\vec{a} = \vec{x} + \vec{y} + \vec{z}$

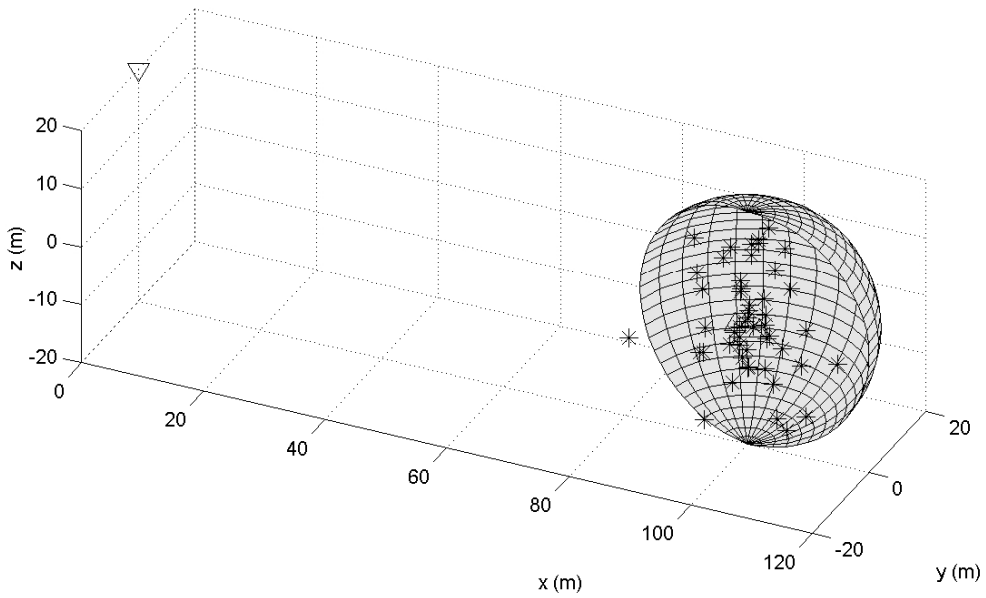


Figure 9.16: The geometry of a more complicated channel scenario with 50 scatterers clustered around the mobile. The base station and mobile antennas are denoted as in Figure 9.12.

and the expressions in Section 9.2 can be used after an appropriate rotation of the antenna coordinate system. For a brief review on rotating antennas using Euler rotation angles, see [Mil99]. Two different antenna arrangements at the base station will be considered, one $\vec{z}$ oriented and one $\vec{y}$ oriented. A carrier frequency $f_c = 5$GHz will be used, and both a narrowband system and a wideband system with a bandwidth of 100MHz will be examined. The position of each scatterer $\vec{r}_s$ is obtained as

$$\vec{r}_s = \vec{x}\ (100 + r\sin\theta\cos\phi) + \vec{y}\ r\sin\theta\sin\phi + \vec{z}\ \cos\theta, \tag{9.35}$$

where the radius of the scattering cluster r is uniformly distributed $r \in U(1, 20)$, and the angles θ, ϕ are also uniformly distributed as $\phi \in U[0, 2\pi)$ and $\theta \in U[0, \pi)$. Furthermore, the scattering properties of each scattering object, as defined in Section 9.3 equation (9.28), are distributed as

$$\begin{aligned} a &\in U(0.01, 0.5) \\ \epsilon_2 &= \infty \\ A &\in N(0, 1) \\ \vec{n} &\in U(0, 1) \end{aligned} \quad . \tag{9.36}$$

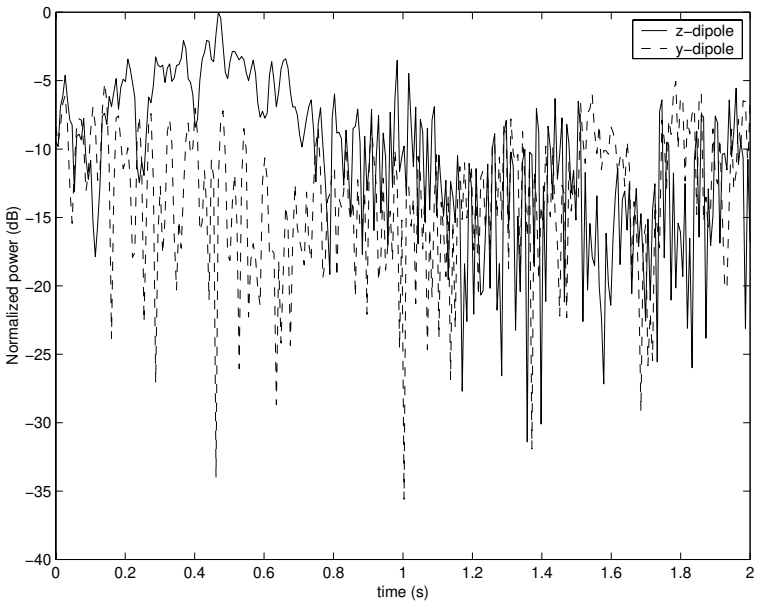


Figure 9.17: The received power from both the $\vec{z}$ oriented and the $\vec{y}$ oriented base station dipoles for the channel scenario in Figure 9.16.

Note that the surface normal $\vec{n}$, with the above definition, will be equally probable in all directions. Further, the size parameter a is chosen as to obtain a large scattering lobe and thus energy contributions from most objects. The scattering amplitude is simply modeled using a complex normal distribution while all objects are assumed to be perfectly conducting.

The received power from both the $\vec{z}$ oriented and the $\vec{y}$ oriented base station dipoles are shown in Figure 9.17 for a narrowband system employing (9.34). The wideband case will be treated later. Note that the powers are normalized to the maximum received power of the $\vec{z}$ dipole. The rapid variations of the received power is due to the interaction of several scatterers and propagation paths, i.e. multipath effects. This is an example of a fading signal, where deep fades are encountered quite frequently. Note that some slow fading properties due to the movement also is visible in Figure 9.17. It is important to stress that Figure 9.17 may be compared to Figure 9.13, where the corresponding curve is shown for a single scatterer. The corresponding curve for the real part of the complex response for the channel scenario in Figure 9.16 exhibits similar behavior as the power plot in Figure 9.17 and is therefore not plotted here.

The received power is almost the same for the $\vec{z}$ dipole as for the $\vec{y}$ dipole, i.e. the vertical polarization is of similar strength as the horizontal, in this

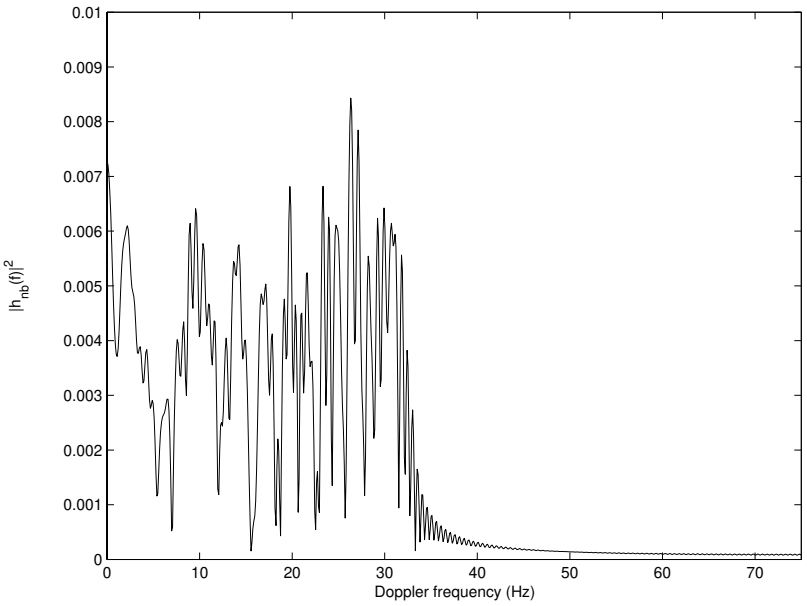


Figure 9.18: The Doppler spectrum for the scenario in Figure 9.16 with $f_c = 5$GHz and a velocity of 2m/s resulting in a theoretical maximum Doppler frequency of 33.3Hz.

case. For other simulation environments, such as larger separation distances, vertical polarization will be slightly larger as found in several measurement studies [Sor98, KTS84, Vau90]. However, it is important to stress that the geometry of the channel determines the polarization properties. Here, it can be useful to use measurement results to determine suitable values and distributions for the different parameters involved in the model [PMF99].

In Figure 9.18, the Doppler spectrum for the scenario in Figure 9.16 when $f_c = 5$GHz, is shown. Using the formulas given in Section 9.2 and the fact that the mobile is moving with a speed of 2m/s, the maximum theoretical Doppler frequency becomes 33.3Hz. This theoretical result agrees very well with the simulations presented in Figure 9.18. However, the actual shape of the Doppler spectra obviously differs significantly from the theoretical curves usually presented in textbooks [Jak74, Pro95, Yac93], that assume uniformly distributed scatterers around the receiver.

The arrival times of the individual waves and the magnitude of the amplitudes are shown in Figure 9.19 for the $\vec{z}$ oriented base station dipole. Also shown is the magnitude of the wideband impulse response as given in (9.32), for a system with a bandwidth of 100MHz. First, it is noted that the arrival times correspond well with the distance over the speed of light for the

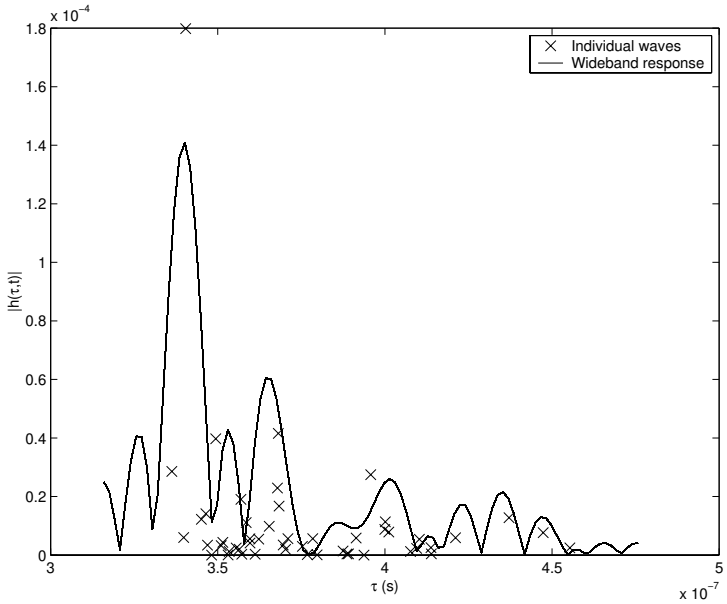


Figure 9.19: The magnitude of the wideband impulse response, the arrival times, and the magnitude of the amplitudes of the individual waves for the channel illustrated in Figure 9.16.

channel scenario. A number of parameters may be used to characterize the wideband impulse response [Jak74, Pro95, Sau99, Yac93]. For example, the mean delay

$$\tau_o = \frac{1}{P_T} \sum_{q=1}^{N_s} P_q \tau_q, \tag{9.37}$$

where $P_T = \sum_{q=1}^{N_s} P_q$, and the Root Mean Square (RMS) delay spread which is defined as

$$\tau_{RMS} = \sqrt{\frac{1}{P_T} \sum_{q=1}^{N_s} P_q \tau_q^2 - \tau_o^2}. \tag{9.38}$$

Here, P_q denotes the received power from path q and τ_q denotes the delay of path q. For the simulation in Figure 9.19, the mean delay is 356ns and the RMS delay spread 16ns. The RMS value is rather low for this type of channel and indicates that a few additional scatterers at other locations typically contribute to the received signal. For instance, several clusters of scatterers may be considered. This would also increase the depolarization of the signals and result in a quite small difference between the received power for different polarizations. This is frequently encountered in practice.

Furthermore, the amplitudes decrease quite rapidly as the delay increases, which agrees quite well with previous measurements and models [SV87, Sau99, TCJ+72]. In those models, the envelope of the impulse response was modeled as exponentially decaying, which also could be used as a model for the response in Figure 9.19. However, it needs to be re-iterated that all results depend on the chosen channel scenario, the position, and the characteristics of the scatterers. The possibilities of calculating different channel properties are virtually endless, and in the next section a few alternatives will be indicated. Also, some relationships to existing channel models will be established.

9.4.3 Discussion

An important contribution of the proposed channel model is that the polarization properties and the antenna properties are included in the model. For example, it is easy to examine the impact of different antenna solutions by simply inserting different antenna functions. Here, both analytical and numerical results can be employed. This is of particular interest, since smart antennas in communication systems are rapidly becoming a key element in the design of future high data rate systems. Also, since the full electromagnetic field is obtained at the receiver, it is possible to examine the potential of employing different electromagnetic polarizations as a diversity receiver [Sva00b]. Already, many communication systems employ dual polarized antennas as two diversity branches. However, recent results indicate that up to six different diversity branches may be available by exploiting the full potential of electromagnetic polarization [AMd01]. All this is possible to model with the proposed channel model, and some results will be presented for the case of MIMO channels in Section 9.6.

Another interesting channel characteristic which is captured by the channel model is the time evolution of the impulse response. As the mobile moves, the directions to the scattering objects will change gradually, and thus affect the characteristics of the channel in a physically sound manner. This is typically not modeled in the purely statistical models often suggested in the signal processing and communications literature. However, with the advent of transmit diversity and MIMO systems, the time evolution of the channel is becoming increasingly more important. In particular, it is important to be able to predict the behavior of the channel in order to more fully exploit the wireless channel [DHHH00, Ekm00, HHDH99]. Hence, the channel model proposed here may be used to simulate mobile radio channels also for this purpose. These properties are even more relevant for multiple antenna systems and in Section 9.6 the topic of channel prediction will be further

discussed.

Furthermore, by using the scattering dyad formulation of the channel model, compact expressions are obtained that are physically intuitive. It is possible to relate effects in the impulse response to features in the channel scenario. Furthermore, it is also possible to relate effects in different receiver algorithms in a similar manner, which is not possible in a measurement campaign. The model also covers both narrowband and wideband systems. Thus, it can be used to simulate many types of radio systems since it does not rely on a specific frequency or time resolution.

By varying the properties and the locations of the scattering objects, many different channel types are possible to simulate. These choices can, of course either be made by physical reasoning, or by calculating these parameters from measurement data. In fact, almost any type of channel can be obtained by an appropriate selection of scattering objects and positions. By making some specific assumptions, it is possible to relate the new channel model to existing ones.

If only horizontal propagation is considered and all antennas are vertically polarized, the scattering dyad may be replaced by a scalar parameter. In that case, the proposed channel model essentially reduces to a model proposed in [FMB98]. That model is also based on positioning scattering objects and calculating the received signal, but without explicitly including any electromagnetic wave propagation. Several channel environments are proposed in [FMB98] that can be adopted to the proposed model by including electromagnetic polarization effects. A similar model was proposed in [BKM96] that also exploits the concept of positioning scattering objects and calculating the received signal, however, without including polarization and 3-D wave propagation. Interestingly, the channel parameters were extracted using measurements and a number of interesting results were obtained. A similar scheme may be applied also to the present channel model, if suitable channel measurements that includes polarization are available.

Finally, it is possible to relate the channel model to some of the purely statistically based channel models. For instance, by placing the scattering objects in a circle in a plane surrounding the receiver, a uniform distribution of the incidence angles at the receiver would result. Thus, in this case the proposed channel model would reduce to the Clark's model [Cla68]. If the scatterers instead were placed on a circular segment within an elevation sector, the extension of Clark's model proposed in [Aul79] would result. Furthermore, if the scatterers are positioned along a circle circumventing the mobile, the well known Jake's model [Jak74] is obtained. A number of other relationships to well known models may also be established. However, many of these assume that an array antenna is used which will be discussed in the

following sections.

9.5 SIMO Channel Modeling

The impulse response of a channel with a single transmitting antenna and an array of receiving antennas, i.e. an Single-Input Multi-Output (SIMO), will be derived in this section using the results from previous Section. The spatio-temporal properties of the channel model are then examined and discussed.

9.5.1 Impulse Response

The transfer function or impulse response for a SIMO channel is easily obtained by calculating the SISO channel impulse response for each individual element at the array and combining the final response as

$$\mathbf{h}(\tau, t) = \begin{bmatrix} h_1(\tau, t) \\ h_2(\tau, t) \\ \vdots \\ h_n(\tau, t) \end{bmatrix}, \tag{9.39}$$

where n denotes the number of antenna elements and $h_i(\tau, t)$ the time domain impulse response of antenna i. The impulse response of each antenna is calculated using (9.32) for the wideband channel and using (9.34) for the narrowband channel.

A simpler expression of the SIMO channel may be obtain if the reciprocal of the bandwidth of the signal is large in comparison with the time needed for the wave to propagate across the array aperture. This assumption is usually called the narrowband assumption in sensor array signal processing literature. Note that this narrowband assumption is different from the narrowband assumptions discussed previously. Here, the bandwidth of the signal is related to the array aperture while previously narrowband related to either the bandwidth of the signal or the delay spread. Thus, a signal may be narrowband related to the array aperture while being wideband when related to the delay spread.

For a signal, narrowband in the aperture sense, the arrival time at each antenna element may be modeled using simple phase shifts which eliminates many of the calculations involved in evaluating (9.39) The induced voltages for an array of n elements when a wave is incident along $\vec{k}_q$ can then be written as the following $n \times 1$ vector

$$\mathbf{v}^T = v\mathbf{a}^T(\vec{k}_q) = v\left[e^{-j\vec{l}_1\cdot\vec{k}_q}, e^{-j\vec{l}_2\cdot\vec{k}_q}, \ldots, e^{-j\vec{l}_n\cdot\vec{k}_q}\right], \tag{9.40}$$

where $\vec{l}_i$ is the location vector of element i, $\vec{k}_q$ is the wave vector of the incident wave, and v is the induced voltage already given in (9.8). Often the first element is chosen as the phase reference center or origin, thereby measuring all the other time delays or distances are relative to the first element. Using (9.40), the wideband SIMO impulse response may be written as

$$\mathbf{h}(\tau,t) = C' \sum_{q=1}^{N_s} \frac{\mathbf{a}(\vec{k}_q)\text{sinc}\left[\omega_b(\tau-\tau_q)\right]}{(|\vec{r}_{ms_q}| \star |\vec{r}_{s_qb}|)^\gamma} G_t(\vec{r}_{ms_q}) G_r(-\vec{r}_{s_qb}) \vec{g}_r(-\vec{r}_{s_qb}) \vec{\mathbf{S}}_q \vec{g}_t(\vec{r}_{ms_q}), \tag{9.41}$$

and the narrowband SIMO impulse response as

$$\mathbf{h}_{nb}(t) = \sum_{q=1}^{N_s} \frac{\mathbf{a}(\vec{k}_q)e^{-j\omega_c\tau_q}}{(|\vec{r}_{ms_q}| \star |\vec{r}_{s_qb}|)^\gamma} G_t(\vec{r}_{ms_q}) G_r(-\vec{r}_{s_qb}) \vec{g}_r(-\vec{r}_{s_qb}) \vec{\mathbf{S}}_q \vec{g}_t(\vec{r}_{ms_q}), \tag{9.42}$$

where $\mathbf{a}(\vec{k}_q)$ is calculated for q^{th} incident wave using (9.40). An implicit assumption in the above simplified expression is that the wave fronts are planar. If the array is close to the source or the array aperture is very large, the expressions need to be corrected for non-planar (spherical) wavefronts [HB91]. This is not necessary using the first expression (9.39), where the full path length is used to calculate the response of each element.

The introduction of several antenna elements gives rise to several new channel characteristics that will be examined in the next section.

9.5.2 Spatio-Temporal Properties

The temporal properties of the SIMO model are similar to the SISO model, since the SIMO response essentially consist of a number of SISO impulse responses. However, there is one important new property, namely the relation or correlation between the signals received by the different array elements. The correlation properties will be examined for the same scattering scenario as described in Section 9.4.2 and illustrated in Figure 9.16. Each individual impulse response will exhibit similar Doppler shift, fading, and RMS delay as found in the SISO analysis and will not be further examined here. Instead, angular properties such as the DOA and the angular spread will be examined in conjunction with the correlation properties. All channel parameters will be the same as in Figure 9.16 except for the attenuation of the waves and the orientation of the mobile antenna $\vec{a} = \vec{z}$. Here, $|\vec{r}_{ms_q}| \star |\vec{r}_{s_qb}|$ will denote $|\vec{r}_{ms_q}| + |\vec{r}_{s_qb}|$, while the attenuation factor γ still is $\gamma = 1$. This change will result in a less rapid attenuation of the scattering around the mobile

and thus increase the delay spread and the angular spread. An alternative way of obtaining this effect is to use a smaller attenuation factor γ, which also would result in a less rapid attenuation. What parameter values to use depends critically on the type of scenario that is desired.

The correlation properties of the received signals from the array depend strongly on the angular properties of the channel. In particular, the angular spread and the mean Direction Of Arrival essentially determine the correlation levels. The mean DOA is calculated in a manner similar to the mean delay as

$$\phi_o = \frac{1}{P_T} \sum_{q=1}^{N_s} P_q \phi_q, \tag{9.43}$$

where $P_T = \sum_{q=1}^{N_s} P_q$ and the RMS angular spread becomes

$$\phi_{RMS} = \sqrt{\frac{1}{P_T} \sum_{q=1}^{N_s} P_q \phi_q^2 - \phi_o^2}. \tag{9.44}$$

Here, P_q denotes the received power, ϕ_q denotes the DOA, and τ_q denotes the delay of path q. In Figure 9.20, the angular spread ϕ_{RMS} is shown for the scenario in Figure 9.16. The average RMS angular spread in this case is 3.9°. Also, the variations between different realizations of the channel are not very large. The angular spread value may appear low. However, with the chosen distribution of scatterers the concentration of scatterers will be denser close to the mobile. Also, the scatterers further from the mobile will often experience longer path delays and consequently also a larger attenuation, and will therefore contribute less to the total received power. An RMS angular spread of 3.9° is, therefore, reasonable.

A very important property of a wireless system employing multiple antennas is the correlation between the signals received by the different antennas. In principle, the correlation properties determine the performance of diversity combining and essentially also the channel capacity. The basic idea of diversity combining is that if the received signal in one antenna experiences a deep fade (low signal SNR), the signals received by the other elements may not, provided that the signal levels exhibit low correlation. The correlation between the envelopes of the received signals, i.e. the envelope correlation, is obtained using [PS60] as

$$\rho_e \approx \frac{|E[h_{nb}^1 {h_{nb}^2}^*]|^2}{E[|h_{nb}^1|^2]E[|h_{nb}^2|^2]}. \tag{9.45}$$

The envelope correlation between two $\vec{z}$ oriented base station antenna elements, as a function of the element separation distance, is shown in Figure

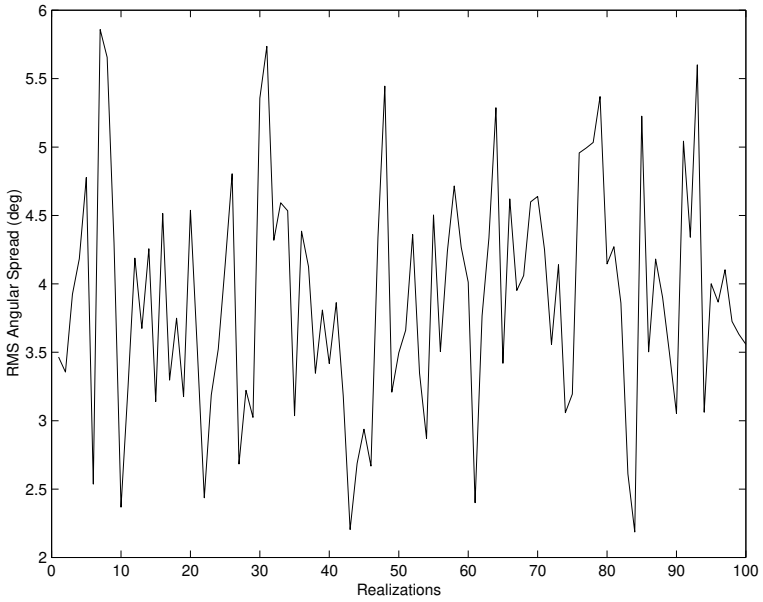


Figure 9.20: The angular spread ϕ_{RMS} for the channel illustrated in Figure 9.16.

9.21. Also shown is the correlation when assuming a Gaussian or Uniform distribution of the DOA with the same angular spread [Ast96]. First, it is noted that the correlation using the Gaussian assumption follows the actual correlation behavior of the scenario under consideration fairly well. A slightly higher correlation is obtained using the Uniform assumption. It is also observed that the ripple behavior is not accounted for by using the Gaussian assumption. The above properties of the correlation as a function of the separation distance have been confirmed both using measurements and other channel models in numerous papers. An excellent overview of spatial channel models can be found in [Ert98]. Since the correlation properties are of great importance when designing receiver algorithms, much work has been performed on deriving accurate channel models for the correlations properties. However, these models only account for the spatial separation and not the polarization of the antennas. Most future communication systems will most likely exploit polarization diversity, since it offers relatively compact antenna solutions. Therefore, it is of great interest to model also the correlation properties between different polarization branches.

In Figure 9.22, the square root of the envelope correlation between two $\vec{z}$ oriented antennas, two $\vec{y}$ oriented and the correlation between one $\vec{z}$ oriented and one $\vec{y}$ oriented antennas are shown vs separation distance. Here, 100

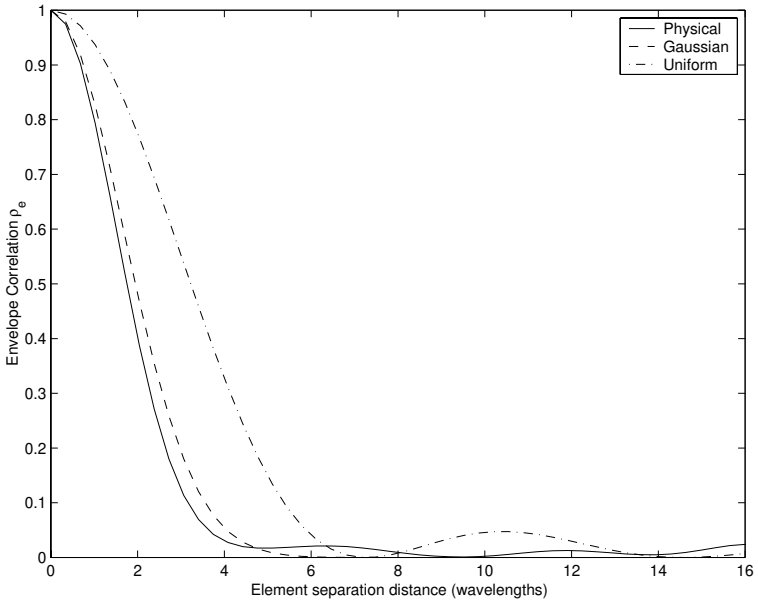


Figure 9.21: The envelope correlation between two $\vec{z}$ oriented base station antenna elements as a function of the element separation distance. Also shown is the correlation when assuming a Gaussian or Uniform distribution of the DOA with the same angular spread.

channel realizations were used to calculate the correlations. Note that the $\vec{z}$ oriented case corresponds to the curve shown in Figure 9.21. The correlation for the $\vec{y}$ oriented case is relatively similar to the vertical case as might be expected. One possible explanation for the slightly slower decorrelation in the horizontal plane is that the wave propagation is predominantly horizontal, leaving more degrees of freedom in the vertical plane that can contribute to the decorrelation. Also, it is interesting to note that the received signals from an $\vec{x}$ and an $\vec{y}$ oriented dipoles are almost uncorrelated in this scenario. Obviously, the polarization correlation properties depends on the choice of the channel parameters, and care should be taken to not draw to many conclusions regarding the levels of correlation. However, if polarization dependent measurements are available, the values of the channel parameters may be tuned to produce similar results as the measurements. The resulting channel model may then be used to simulate different systems and also to relate physical parameters to effects in the signal processing schemes. In the next section, the possibilities of the SIMO model as well as some relationships to existing channel models will be discussed.

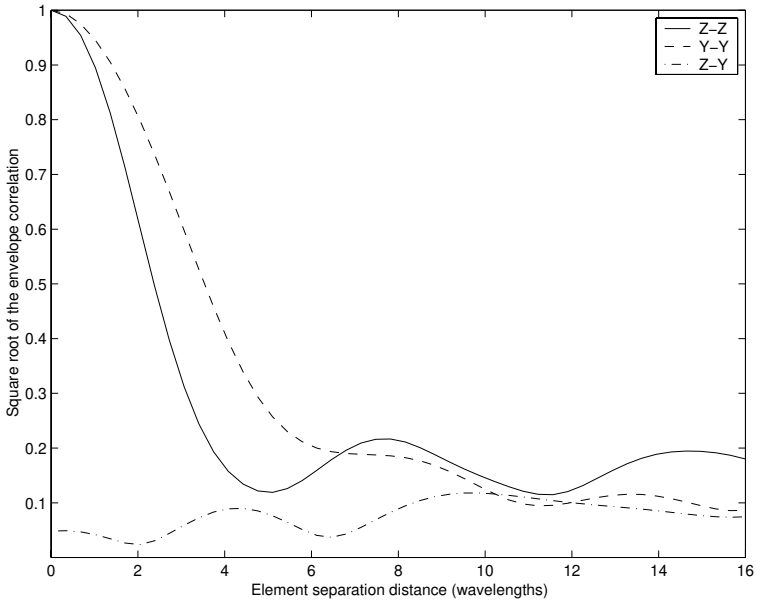


Figure 9.22: The square root of the envelope correlation between two $\vec{z}$ oriented, two $\vec{y}$ oriented, and between one $\vec{z}$ oriented and one $\vec{y}$ oriented antennas vs separation distance.

9.5.3 Discussion

The introduction of an array antenna at the base station offers many advantages compared to the single antenna case, such as beamforming and diversity combining. Therefore, there has been a significant interest in modeling the channel with a single mobile and a base station antenna array, i.e. the SIMO channel. The focus of this research has been on obtaining models for the correlation between the elements and how the correlation depends on environmental parameters such as element spacing. Excellent overviews of spatial channel models may be found in [Ert98, LR99], where references are given to numerous papers on SIMO channel models. Analytical expressions for the correlation under different assumptions can, for instance, be found in [Ast96, Jak74, Sau99].

Most of the SIMO channel models presented in the literature are possible to obtain as a special case of the channel model proposed in this report, by assuming an appropriate distribution of the scatterers and their respective properties. In particular, if the polarization properties are neglected and a scalar scattering coefficient is used, the proposed channel model reduces to models proposed in [BKM96, FMB98]. In those works, suggestions for real-

istic values of the channel parameters, such as number of scatterers, number of scatter centers, attenuation factors, radius of scattering centers, and typical distances between base and mobile can be found. Essentially, the SIMO model presented in this section is only a combination of several SISO models. This is one advantage of physical modeling compared to statistical modeling, where a whole set of new correlation functions and distributions need to be determined.

If it is difficult to accurately model the correlation between the different antennas at the base station, it is even more difficult to model systems with multiple antennas at both transmitter and receiver. Thus, models based on fundamental physics are well suited for this modeling task which is the topic of the next section.

9.6 MIMO Channel Modeling

The impulse response of a channel with multiple transmit and receive antennas, i.e. an Multi-Input Multi-Output (MIMO), will be derived in this section using the results from previous sections. The spatio-temporal properties of the channel model are then examined and discussed.

9.6.1 Impulse Response

Essentially, the extension from a single transmit antenna and single receive antenna to a MIMO system with n_t transmit and n_r receive antennas only requires the calculation of $n_t \times n_r$ impulse responses or channel coefficients. Using the proposed physical channel model, this simply means evaluating (9.32) and (9.34) for many different transmit and receive positions.

Most of the analysis of MIMO systems, so far, has been regarding narrowband systems [CJT98, FG98, SFGK00, NSC00]. In that case, the MIMO channel can be expressed as an $n_r \times n_t$ channel matrix $\mathbf{H}(t)$ with elements $\mathbf{H}_{ij}(t) = h_{nb}^{ij}(t)$, where $h_{nb}^{ij}(t)$ denotes the channel coefficient between transmit antenna j and receive antenna i, calculated using (9.34). Different normalizations of this narrowband MIMO channel matrix have been used in the literature, where $||\mathbf{H}||_F^2 = n_r n_t$ appears to be the most common normalization [Bur00].

For wideband MIMO channels, numerous possibilities of representing the channel exist, such as tensor notation or concatenated matrices [LP00, SJBF00]. However, the set of channel impulse responses that define the MIMO channel, $\{h^{ij}(t,\tau); i = 1,2,\ldots,n_r, j = 1,2,\ldots,n_t\}$, is easily evaluated using (9.32), with the corresponding antenna positions for each impulse

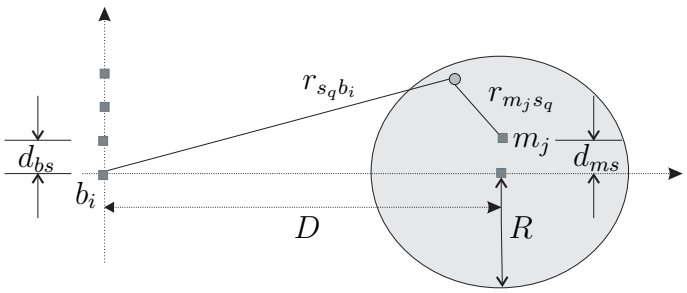


Figure 9.23: Microcell channel environment.

response. Although most of the MIMO analysis has been performed for narrowband systems, future systems will be wideband where it is critical that the temporal properties of the MIMO channel is accounted for [LP00].

Both narrowband and wideband MIMO systems will be simulated and discussed in the next section.

9.6.2 Spatio-Temporal Properties

Some of the properties of MIMO channels will be examined in this section. In particular, the correlation properties as seen from the transmitter and the receiver will be examined along with the achievable data rate (channel capacity). The potential of using EM polarization as diversity branches will also be briefly addressed. Two examples of channel environments will be examined where suitable values of the channel parameters are given. First, a microcell propagation environment will be presented that allows for comparisons of spatial correlations with previous SIMO models [Ert98, Sau99]. A picocell or indoor model will also be presented, that models areas where the scattering is dense and high channel capacities have been measured [KSMP00, SOK00].

In Figure 9.23, a channel environment is shown that may be characterized as a microcell environment. The scatterers are uniformly distributed within a sphere of radius R, centered at the mobile that is located a distance D from the base station. The antenna element separation is denoted d_{ms} and d_{bs} at the mobile and base respectively. Values for the number of scatterers N_s and the carrier frequency f_c together with the distances D and R are given in Table 9.1. These values correspond to similar environments given in [BJ98, FMB98], that were found to yield realistic channel characteristics. The polarization properties are determined by the parameters $\vec{n}$, a, and A. Here, it is assumed that the surface normals are uniformly distributed, resulting in strong depolarization. Other choices can of course be made such

Microcell Environment	Picocell/Indoor Environment
$D = 5$km	$D = 30$m
$R = 0.5$km	$R = 100$m
$N_s = 50$	$N_s = 50$
$f_c = 5$GHz	$f_c = 5$GHz
$\vec{n} \in U[0,1]$, $\|\vec{n}\| = 1$	$\vec{n} \in U[0,1]$, $\|\vec{n}\| = 1$
$a \in U[0.01, 0.5]$	$a \in U[0.01, 0.5]$
$A \in \mathcal{CN}(0,1)$	$A \in \mathcal{CN}(0,1)$

Table 9.1: Channel parameters for two channel environments.

as a distribution biased towards horizontal normals, to account for the fact that buildings and walls typically have vertical surfaces. A uniform distribution also models the width of the scattering lobe that is of sufficient range to include both broad and narrow scattering lobes. Finally, the scattering coefficient is modeled as a zero mean complex Gaussian process, see Table 9.1.

A channel environment with picocell and indoor characteristics can be obtained using almost the same modeling approach as above. In this case, a scattering sphere is located between the base and the mobile that has a scattering radius large enough to enclose both mobile and base. The corresponding channel parameters are given in Table 9.1. Both the microcell and indoor environments use $\gamma = 1$ and multiplicative attenuation, i.e. $|\vec{r}_{ms}| \star |\vec{r}_{sb}|$ denotes $|\vec{r}_{ms}||\vec{r}_{sb}|$.

Since the overwhelming part of the analysis of channel capacity and space-time processing has been conducted for narrowband systems, the correlation properties of narrowband MIMO systems will be examined first. Wideband characteristics and polarization properties are discussed after that.

A very important property of the channel that essentially determines the channel capacity is the correlation between different antennas or channel coefficients. The related case of a single mobile antenna and a base station array has been thoroughly studied in numerous theoretical and practical investigations, as discussed in the previous section. However, the MIMO case is less understood, where essentially only theoretical results and conjectures have appeared in the literature [CRFLL00, SFGK00, PAKM00].

In Figure 9.24, the envelope correlation ρ_e [Jak74, Sau99] of a signal received by two base station antennas separated by a distance d_{bs} is shown for various separation distances. Here, 500 realizations of the microcell scenario given in Figure 9.23 and Table 9.1 is used, and the antennas are $\vec{z}$ oriented

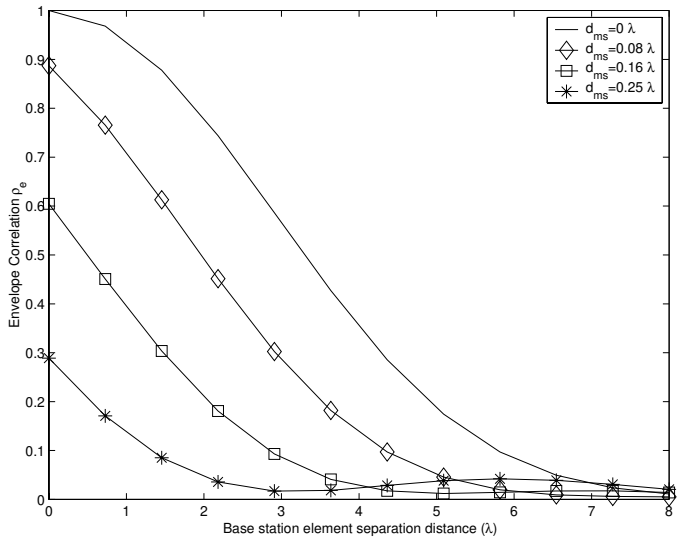


Figure 9.24: The envelope correlation $\rho_{n_2 m_2}^{n_1 m_1}$, as seen from the base station, for various element separation distances.

dipole elements. The solid curve is directly comparable to the previous analysis on SIMO systems [Ert98, Jak74, Sau99], and is well modeled using a Gaussian distribution for the angular properties. Also included in the plot is the corresponding envelope correlation when the signal is emanating from two mobile antennas separated by a distance d_{ms}, i.e.

$$\rho_{n_2 m_2}^{n_1 m_1} \approx \frac{E^2[\mathbf{H}_{n_1 m_1} \mathbf{H}_{n_2 m_2}^*]}{E[|\mathbf{H}_{n_1 m_1}|^2] E[|\mathbf{H}_{n_2 m_2}|^2]}, \tag{9.46}$$

where the channel matrix $\mathbf{H}$ is defined in Section 9.6.1. The base station antennas n_1 and n_2 are separated a distance d_{bs} while the mobile antennas m_1 and m_2 are separated a distance d_{ms}.

It has been conjectured [CRFLL00, PAKM00] that the correlation properties at the transmitter and receiver are independent from one another, and that the overall correlation properties may be obtained through multiplication. This is supported by the simulations presented in Figure 9.24, where the correlation decreases in a multiplicative manner as the mobile correlation decreases with increasing d_{ms}.

However, this is not the case in Figure 9.25, where instead the correlation properties are shown versus mobile separation distance d_{ms}. Here, the general shape of the correlation curve changes as d_{bs} is changed, thus indicating that the correlation properties are not independent. Further measurements are needed to determine if the predictions of the model are in agreement with practical channel environments.

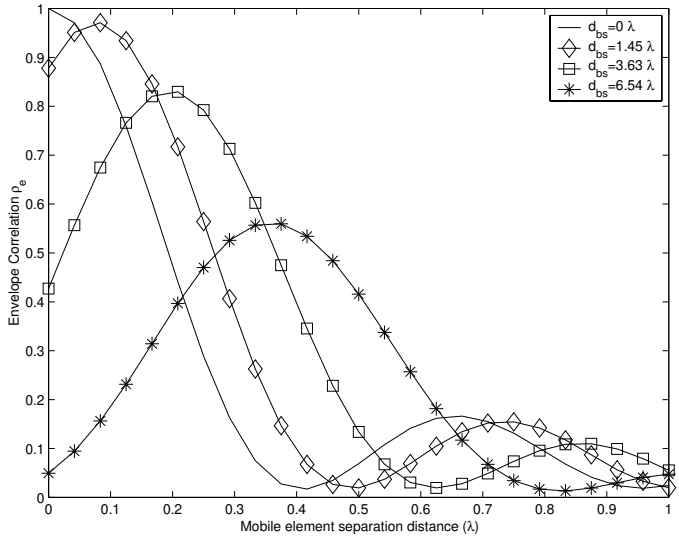


Figure 9.25: The envelope correlation $\rho_{n_2 m_2}^{n_1 m_1}$, as seen from the mobile, for various element separation distances.

For systems with a high degree of correlation in one end and relatively low correlation at the other, the multiplicative conjecture may provide a reasonable first order approximation. The magnitude of the envelope correlation matrix, i.e. a normalization of the correlation matrix $\mathbf{R} = E[\mathrm{vec}(\mathbf{H})\mathrm{vec}^{\mathrm{H}}(\mathbf{H})]$, is shown in Figure 9.26 for a 3×3 microcell MIMO system with $d_{ms} = 0.5\lambda$ and $d_{bs} = 0.75\lambda$. In this case, the correlation is well approximated by $\mathbf{R} = \mathbf{R}_{bs} \otimes \mathbf{R}_{ms}$ (see Figure 9.27), where $\otimes$ denotes the Kronecker product and the correlation properties at the base station and mobile are denoted $\mathbf{R}_{bs}$ and $\mathbf{R}_{ms}$, respectively. However, for systems with intermediate correlation, the simulations indicate that the correlation properties cannot accurately be described by multiplication. As mentioned above, further measurements are needed to investigate these properties in detail.

Many future MIMO systems are designed to offer high data rates in indoor environments, where most likely multiple polarizations will be used to obtain compact implementations suitable for hand-held devices. The channel capacity of a 2×2 MIMO system using co-located 45° slanted dipoles [Bal82, Sau99], will be simulated next. The channel capacity is calculated using formulas given in [FG98] and the indoor channel parameters given in Table 9.1.

In Figure 9.28, the channel capacity is shown for the dual-polarized MIMO system along a path of length 10λ (0.6m), with an average SNR of 10dB (path-average). For comparison, the corresponding capacities of a SISO

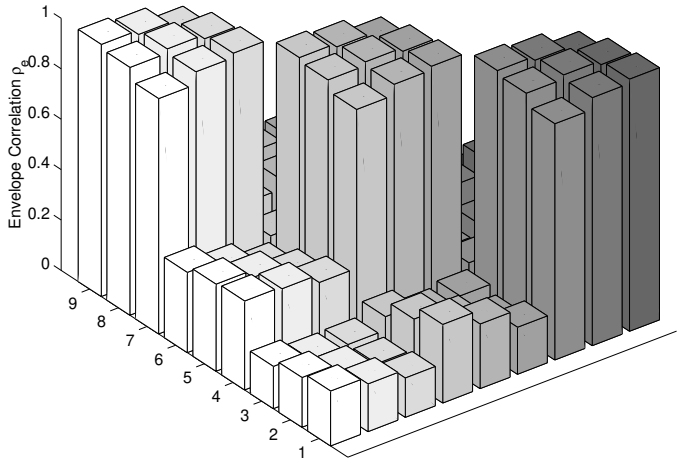


Figure 9.26: The magnitude of the envelope correlation matrix for a 3×3 MIMO system with $d_{ms} = 0.5\lambda$ and $d_{bs} = 0.75\lambda$.

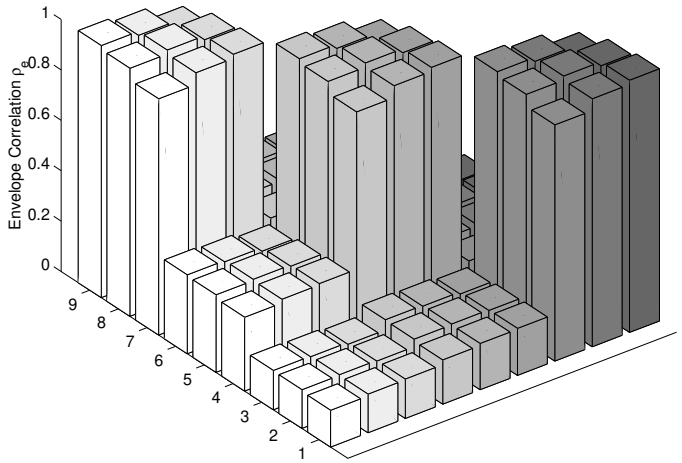


Figure 9.27: The magnitude of the Kronecker envelope correlation product matrix $\mathbf{R} = \mathbf{R}_{bs} \otimes \mathbf{R}_{ms}$ for a 3×3 MIMO system with $d_{ms} = 0.5\lambda$ and $d_{bs} = 0.75\lambda$.

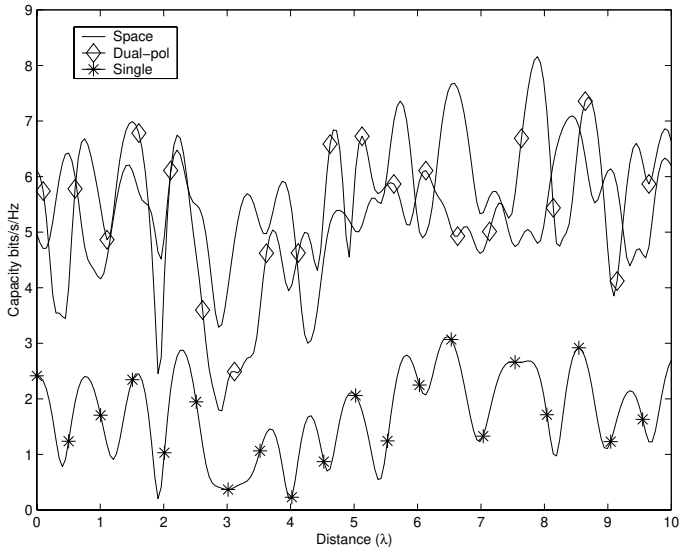


Figure 9.28: The channel capacity of a dual-polarized MIMO system along a path of length 10λ (0.6m), with an average SNR of 10dB (path-average). Also, shown is the corresponding capacities of a SISO system and a 2×2 spatially separated MIMO system.

system and a 2×2 MIMO system of $\vec{z}$ oriented dipole elements separated as $d_{ms} = 0.5\lambda$ and $d_{ms} = 2\lambda$, are also shown. In this case (a single channel realization), the dual-polarized and the spatially separated systems offer similar capacities that are about three times the corresponding capacity of a single antenna system. Here, the normalization $||\mathbf{H}||_F^2 = n_r n_t = 4$ is employed, thereby including both antenna gain and diversity gain [And00, Bur00], which can be observed in Figure 9.28 where the deeper capacity dips are avoided in the 2×2 systems.

It was recently proposed that all six possible EM polarizations can be used as diversity branches [AMd01] to increase capacity. In Figure 9.29, the CCDF of the magnitude of the eigenvalues of $\mathbf{H}$ are shown for an indoor MIMO system employing 6 co-located elements, i.e. three electrical dipoles and three ring antennas (magnetic dipoles). The magnitude of the eigenvalues in Figure 9.29 are similar enough to provide relative large capacity gains over a single antenna system. This is clearly indicated in Figure 9.30, where the CCDF of the channel capacity for the same scenario is shown. The channel capacity of a corresponding SISO system and a hypothetical independent and identically distributed (i.i.d) system is also shown. In fact, the capacity of the multipolarized system is close to the i.i.d system that serves as an upper bound on the capacity [FG98]. Thus, employing all EM polarizations

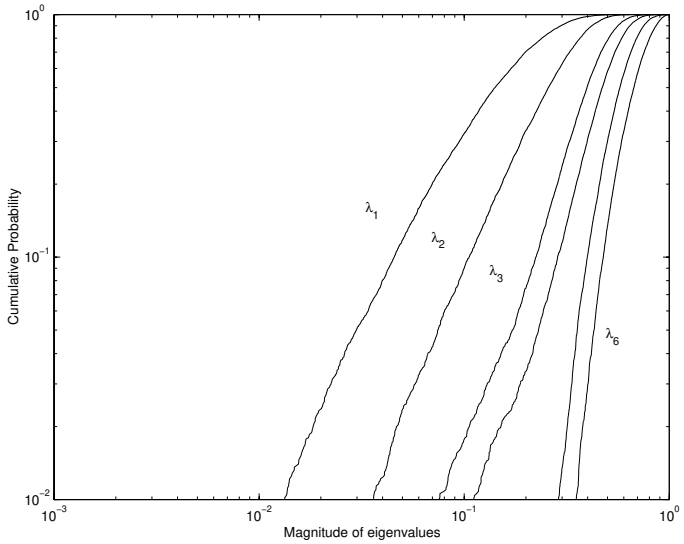


Figure 9.29: The CCDF of the magnitude of the eigenvalues of an indoor MIMO system employing all six EM polarizations.

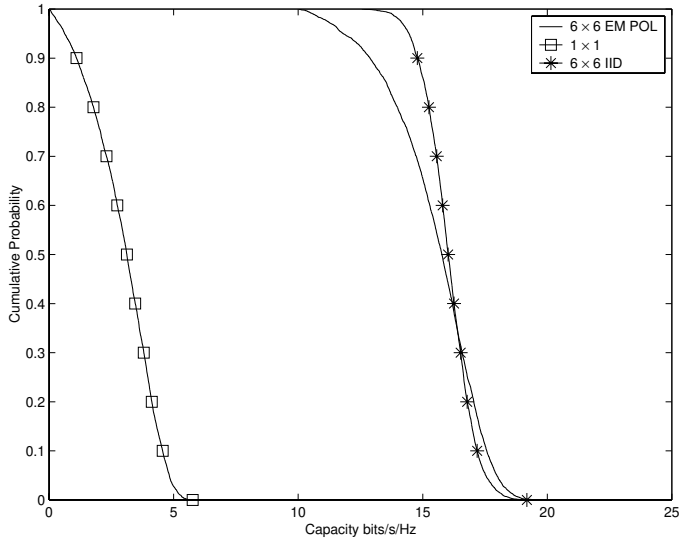


Figure 9.30: The CCDF of the channel capacity in an indoor MIMO system employing all six EM polarizations.

offers interesting new possibilities of achieving high capacities, where novel antenna solutions may offer very compact implementations.

As mentioned in the beginning of this section, most of the MIMO analysis has been of narrowband character. However, future system will most likely be wideband. Thus, the temporal domain also needs to be modeled. Using a physical channel model, this is easily included in a similar manner as in the SISO case, see Figure 9.16.

A number of numerical examples of channel environments have been presented in this section. However, many other types of channels can be obtained by adjusting the channel parameters, and it is an ongoing work to determine parameter sets which based on measurements match reality.

9.6.3 Discussion

Essentially, the extension from a single transmit antenna and single receive antenna to a MIMO system only requires that several impulse responses need to be calculated. Using the proposed physical channel model, this simply resulted in evaluating the SISO expressions for many different transmit and receive positions. This is one of the advantages with employing a physical channel model compared to a purely stochastic model, where a whole set of new correlation functions and distributions needs to be determined. Furthermore, in a manner similar to the SISO and SIMO cases, it is straightforward to simulate a system employing different polarizations. By positioning the scatterers and using different characteristics, most types of channels may be simulated. One important benefit of the proposed physical model is that it provides physical insight, which is especially interesting in the less measured and understood MIMO systems.

The framework presented in this report has been used to simulate MIMO channels employing novel antenna solutions such as multimode antennas [Sva00c] and parasitic antennas [WS01]. See also Chapter 10 and 11 for more details. Moreover, the impact of mutual coupling on the channel capacity [SR01] has been studied using a version of the channel model presented in this section, see Chapter 12. In this section, the recently proposed idea of exploiting all six EM polarizations in MIMO systems [AMd01] was briefly addressed. The distribution of the eigenvalues of the channel matrix was calculated for such a system in an indoor scenario. It was found that the capacity, in that case, was close to the theoretical capacity. Thus, exploiting EM polarization offers essentially the same capacity as a traditional spatially separated antenna array. Since very compact antenna solutions may be possible for systems employing EM polarizations, this is an interesting future topic. The effect of antenna separation on the correlation properties was

also examined in this section. The possibility of modeling the correlation at each end as independent, was also studied. These issues are important when designing MIMO systems, and they have therefore been addressed in several recent publications [Bur00, CRFLL00, PAKM00]. Although a number of measurements have been presented [KSMP00, MWS00, SOK00], more empirical results are needed in order to confirm the simulation results that has been presented. Finding parameter sets for the proposed channel model that satisfactorily matches reality is an important ongoing work.

Since the model naturally incorporates temporal channel properties, the proposed physical model also appears to be suitable when analyzing the performance of wideband MIMO systems [LP00]. Another interesting area of physical channel modeling is long-range prediction of fading signals, where realistic channel models are desired to evaluate the performance of adaptive transmission techniques [DHHH00]. Realistic channel models based on physics may also prove useful when designing efficient feedback strategies for systems employing transmit diversity [NLTW98].

9.7 Conclusions

A spatio-temporal channel model based on EM scattering and fundamental physics was presented. By studying the scattering properties of objects of simple shapes, such as spheres and cylinders, a simple function that captures the most important scattering properties was derived. A compact formulation was obtained by using a dyad notation and concepts from rough surface scattering. A simple, yet detailed, channel model was then obtained by employing results from EM wave propagation theory and elementary antenna theory. In a sense, the model bridges the gap between the complicated wave propagation environment of radio signals and the idealized models previously used when analyzing MIMO systems.

These results were then used to formulate SISO, SIMO, and MIMO channel models. Essentially, the extension from a single transmit antenna and single receive antenna to multiple antennas only requires that several impulse responses or channel coefficients need to be evaluated. Hence, a physical channel model can be modified to different antenna arrangements rather easily as compared to a purely stochastic model where a whole set of new correlation functions and distributions need to be determined. For each type of channel model, important channel characteristics were simulated and discussed. For instance, properties such as antenna correlation, Doppler spread, channel evolution, channel capacity, fading, and polarization diversity were analyzed.

In principle, most channel types can be obtained by positioning the scatterers and using different object characteristics. However, it is important to find parameter sets that yield channel properties that correspond to real channel environments. Here, measurements are needed and it is an ongoing work to determine parameter values which match reality. An interesting method of doing that is Matched Field Processing (MFP), which has been employed for estimation of physical channel parameters in CDMA systems [AG00].

One important contribution of the proposed physical model is that it provides physical insight, which is especially interesting in the less measured and understood MIMO systems. It also provides the possibility of evaluating the potential of different polarization schemes, such as dual-polarized antennas and sixth-fold polarization diversity that recently was proposed [AMd01]. The proposed model also appears suitable for wideband analysis [LP00] and for evaluating channel prediction schemes [DHHH00]. Other applications where the time evolution of the channel is important is transmit diversity schemes with feedback [NLTW98]. Thus, the proposed channel model has many interesting applications. The framework of this report has already been used to analyze the impact of several different antenna configurations [Sva00a, Sva00b, Sva00c, SR01, WS01].

9A Scattering by a Sphere

Expressions of $\vec{\alpha}$, that were introduced in Section 9.2, for scattering of a plane wave by a sphere will be presented in this appendix. Consider a sphere of radius a located at the origin, see Figure 9.3. A uniform plane wave polarized in the $\vec{x}$ direction traveling along the $\vec{z}$ axis is incident upon the sphere. The incident electric field can be written as

$$\vec{E}^i = \vec{x} E_0 e^{-jkz}, \tag{9A.1}$$

where k denotes the wave number. For objects with a geometry coinciding with a coordinate system, the scattered field can be found using modal techniques. In the case of a sphere, the incident plane wave is expressed in spherical wave functions. The scattered field is expressed in similar functions and the fields are then matched at the surface of the sphere using boundary conditions. The scattered field at a point (r, θ_s, ϕ_s) that is sufficiently far from the sphere, i.e. the far-field expressions, then becomes

$$\vec{E}^s(r, \theta_s, \phi_s) = E_0 \frac{e^{-jkr}}{r} \left[\vec{\theta} E^s_\theta + \vec{\phi} E^s_\phi \right], \tag{9A.2}$$

where

$$E_\theta^s = \frac{j\cos(\phi_s)}{k}\sum_{n=1}^{\infty} j^n \frac{2n+1}{n(n+1)}\left[a_n u_1(\theta_s) - b_n u_2(\theta_s)\right] \tag{9A.3}$$

$$E_\phi^s = \frac{j\sin(\phi_s)}{k}\sum_{n=1}^{\infty} j^n \frac{2n+1}{n(n+1)}\left[a_n u_2(\theta_s) - b_n u_1(\theta_s)\right]. \tag{9A.4}$$

Here, $u_1(\theta_s) = \sin\theta_s P_n^{1\prime}(\cos\theta_s)$ and $u_2(\theta_s) = \frac{P_n^1(\cos\theta_s)}{\sin\theta_s}$, where P_n^m is the associated Legendre function [Bal89] and the prime denotes differentiation with regards to the function argument. The coefficients a_n and b_n depend on the electrical size of the sphere $s = ka$ and $s_1 = k_1 a$. It should also be noted that k_1 is the wave number for the sphere and k for the surrounding medium. Thus, the dielectric properties of the sphere enter the expressions through the wave number. Assuming that the permeability of the sphere is the same as the surrounding medium, i.e. $\mu_1 = \mu$. The coefficients can then be written as [Bal89, BSea87, KCH71]

$$a_n = -\frac{k^2 j_n(s)[s_1 j_n(s_1)]' - k_1^2 j_n(s_1)[s j_n(s)]'}{k^2 h_n^{(2)}(s)[s_1 j_n(s_1)]' - k_1^2 j_n(s_1)\left[s h_n^{(2)}(s)\right]'} \tag{9A.5}$$

$$b_n = -\frac{j_n(s)[s_1 j_n(s_1)]' - j_n(s_1)[s j_n(s)]'}{h_n^{(2)}(s)[s_1 j_n(s_1)]' - j_n(s_1)\left[s h_n^{(2)}(s)\right]'}, \tag{9A.6}$$

where $j_n(x)$ is the spherical Bessel function [Bal89] of order n, $h_n^2(x)$ is the spherical Hankel function of the second kind [Bal89] of order n, and the prime denotes differentiation with regards to the function argument.

At the time these solutions were derived, evaluation of the expressions was almost impossible. However, with access to a computer the above expressions are readily evaluated. Furthermore, only a few terms of the sum need to be included to achieve the accuracy needed in this application. To obtain more accurate values, a rule of thumb of how many terms that should be evaluated is [Wis80]

$$n_c = s + 4.05 s^{1/3} + 2. \tag{9A.7}$$

The solution given in (9A.2)-(9A.6) is the general solution valid for all frequencies, where the special case of a perfectly conducting sphere is also included by choosing a complex value for the dielectric constant. However, the solution is based on a $\vec{x}$ oriented field incident along the $\vec{z}$ axis. These conditions will rarely be fulfilled in the channel model derived in Section 9.2, since the same coordinate system is used for all scatterers. By employing a local coordinate system for each object, the above solution can be applied. Thus, a coordinate transformation is necessary where the local coordinate system should be oriented such a way that $\vec{x'}$ coincides with the direction of $\vec{g}_t$ and that $\vec{z'}$ coincides with the direction of $\vec{r}_{ms}$. The final basis vector

is then obtained as $\vec{y}' = \vec{z}' \times \vec{x}'$, where $\times$ denotes the vector cross product [Bal89]. The transformation from local to global coordinates becomes

$$\mathbf{T}_{lg} = \left[\frac{\vec{g}_t}{|\vec{g}_t|}, \frac{\vec{r}_{ms} \times \vec{g}_t}{|\vec{r}_{ms}||\vec{g}_t|}, \frac{\vec{r}_{ms}}{|\vec{r}_{ms}|} \right]. \tag{9A.8}$$

The scattering angles (θ_s, ϕ_s) are easily found from the scattering direction expressed in local coordinates as $\vec{r}'_{sb} = \mathbf{T}_{lg}^{-1} \vec{r}_{sb}$. Unfortunately, the scattered field is also expressed in local coordinates. To get the scattered field back into the global coordinate system, the vectors $\vec{\theta}$ and $\vec{\phi}$ in (9A.2) need to be transformed to the global rectangular coordinate system. Combining these steps, the scattering parameter $\vec{\alpha}$ can finally be written as

$$\vec{\alpha} = \mathbf{T}_{lg} \mathbf{T}_{sr} \vec{\theta}' E_\theta^S + \mathbf{T}_{lg} \mathbf{T}_{sr} \vec{\phi}' E_\phi^S, \tag{9A.9}$$

where $\mathbf{T}_{sr}$ denotes the transformation from spherical to rectangular coordinates [Bal89], E_θ^S is given in (9A.3), and E_θ^S is given in (9A.4). Thus, $\vec{\alpha}$ depends on $\vec{r}_{ms}, \vec{r}_{sb}$, and $\vec{g}_t$ as well as on the size and the electric properties of the sphere.

9B Scattering by a Finite Cylinder

Expressions of $\vec{\alpha}$, that were introduced in Section 9.2, for scattering of a plane wave by a cylinder will be presented in this appendix. Consider a uniform plane wave incident obliquely upon the cylinder, see Figure 9.6. The incident electric field can be written as

$$\vec{E}^i = E_0 e^{jk(z\cos\theta_i - x\sin\theta_i)} \vec{e} = E_0 e^{jk(z\cos\theta_i - x\sin\theta_i)} \left[\alpha^{TM} \left(\vec{x}\cos\theta_i + \vec{z}\sin\theta_i \right) + \alpha^{TE} \vec{y} \right], \tag{9B.1}$$

where the incident field has been separated in a Transverse Magnetic (TM) and Transverse Electric (TE) part. That is, one part where the electric field is perpendicular to $\vec{z}$ (TE) and one part where the magnetic field is perpendicular to $\vec{z}$ (TM). The reason for this is that the solution will be different for the TE and TM parts. The scattered field is found in a manner similar to the sphere case using modal techniques. However, in this case the plane wave is expressed in cylindrical functions. An exact solution is unfortunately only available for an infinite cylinder [Wai55]. Approximate solutions exist [Bal89, Ruc70] that employ concepts from PO, and thus give good results for scattering directions close to specular reflection. However, for this application of channel modeling, the accuracy is acceptable. Using

those approximations for the finite cylinder along with the original solutions then gives the far-field expressions for the scattered field in a point (r, θ_s, ϕ_s), see Figure 9.6, as

$$\vec{E}^s(r, \theta_s, \phi_s) = E_0 \frac{e^{-jkr}}{r} \left[\vec{\theta} E_\theta^s + \vec{\phi} E_\phi^s \right], \tag{9B.2}$$

where

$$E_\theta^s = \tfrac{jl}{\pi} \tfrac{\sin U}{U} \tfrac{\sin \theta_s}{\sqrt{\sin \theta_i}} \textstyle\sum_{n=-\infty}^{\infty} \left(\alpha^{TE} f_{2,n} - \alpha^{TM} f_{1,n} \right) e^{jn\phi} \tag{9B.3}$$

$$E_\phi^s = \tfrac{jl}{\pi} \tfrac{\sin U}{U} \tfrac{\sin \theta_s}{\sqrt{\sin \theta_i}} \textstyle\sum_{n=-\infty}^{\infty} \left(\alpha^{TE} f_{1,n} + \alpha^{TM} f_{2,n} \right) e^{jn\phi}, \tag{9B.4}$$

where $U = \frac{kl}{2}(\cos\theta_i + \cos\theta_s)$ and l is the length of the cylinder. The coefficients $f_{1,n}$ and $f_{2,n}$ are defined as

$$\begin{aligned} f_{1,n} &= -\frac{J_n(v)}{H_n^{(2)}(v)} - \frac{2j \frac{H_n^{(2)\prime}(v)}{v H_n^{(2)}(v)} - \frac{K J_n'(u)}{u J_n(u)}}{\pi v^2 \left[H_n^{(2)}(v) \right]^2 D} \\ f_{2,n} &= -\frac{2}{\pi v^2} \left(\frac{1}{u^2} - \frac{1}{v^2} \right) \frac{n \cos\theta_i}{\left[H_n^{(2)}(v) \right]^2 D}, \end{aligned} \tag{9B.5}$$

where

$$\begin{aligned} D = \left[\frac{H_n^{(2)\prime}(v)}{v H_n^{(2)}(v)} - \frac{K J_n'(u)}{u J_n(u)} \right] \left[\frac{H_n^{(2)\prime}(v)}{v H_n^{(2)}(v)} - \frac{N^2 J_n'(u)}{K u J_n(u)} \right] \\ - \left[\frac{1}{v^2} - \frac{1}{u^2} \right] n^2 \cos^2\theta_i \end{aligned} \tag{9B.6}$$

and $K = \mu_1/\mu$, $v = ka\sin\theta_i$, $u = \sqrt{k_1^2 - k^2\cos^2\theta_i}a$, and $N = k_1^2/k^2$. Also, the cylindrical Bessel function [Bal89] of order n is denoted $J_n(x)$ and the Hankel function of the second kind [Bal89] of order n is denoted $H_n^{(2)}$. The prime denotes differentiation with regards to the function argument. The rule of thumb given in (9A.7) for spheres can be applied also here.

Similar to the sphere analysis, a coordinate transformation is needed in order to apply the above results in the channel model in Section 9.2. In this case, the $\vec{z}'$ axis of the local coordinate system should coincide with the cylinder axis $\vec{r}_{\text{cyl}}$. Also, the field should be incident along the $\vec{x}'$ axis (see Figure 9.6). This is ensured by choosing $\vec{x}' = \vec{r}_{ms} - (\vec{r}_{ms} \cdot \vec{r}_{\text{cyl}})\vec{r}_{\text{cyl}}$. Thus, the transformation from local to global coordinates can be expressed as

$$\mathbf{T}_{lg} = \left[\frac{\vec{r}_{ms} - (\vec{r}_{ms} \cdot \vec{r}_{\text{cyl}})\vec{r}_{\text{cyl}}}{|\vec{r}_{ms} - \vec{r}_{ms} \cdot \vec{r}_{\text{cyl}})\vec{r}_{\text{cyl}}|}, \frac{\vec{r}_{\text{cyl}} \times (\vec{r}_{ms} - (\vec{r}_{ms} \cdot \vec{r}_{\text{cyl}})\vec{r}_{\text{cyl}})}{|\vec{r}_{\text{cyl}} \times (\vec{r}_{ms} - (\vec{r}_{ms} \cdot \vec{r}_{\text{cyl}})\vec{r}_{\text{cyl}})|}, \vec{r}_{\text{cyl}} \right]. \tag{9B.7}$$

Using $\mathbf{T}_{lg}$ the incident field in local coordinates is easily found as $\vec{E}^{i\prime} = \mathbf{T}_{lg}^{-1}\vec{E}^{i}$, from which the TE and TM parts of the field $(\alpha^{TE}, \alpha^{TM})$ are calculated . Furthermore, the angles θ_i' and θ_s' are obtained from the transformed vectors $\vec{r}'_{ms} = \mathbf{T}^{-1}\vec{r}_{ms}$ and $\vec{r}'_{sb} = \mathbf{T}^{-1}\vec{r}_{sb}$.

Finally, the scattering parameter $\vec{\alpha}$ can be written as

$$\vec{\alpha} = \mathbf{T}_{lg}\mathbf{T}_{sr}\vec{\theta}'E_\theta^S + \mathbf{T}_{lg}\mathbf{T}_{sr}\vec{\phi}'E_\phi^S, \tag{9B.8}$$

where $\mathbf{T}_{sr}$ denotes the transformation from spherical to rectangular coordinates [Bal89], E_θ^S is given in (9B.3), and E_θ^S is given in (9B.4). Note that the scattering parameter for the cylinder is related to the RCS in the same manner as for the sphere, see (9.14).

Chapter 10

Exploiting Multimode Diversity in MIMO Systems

A novel way of exploiting higher modes of antennas as diversity branches in MIMO systems is introduced. Essentially, antennas employing multiple modes offer characteristics similar to an antenna array, through multiple modes and using only a single element. Analytical expressions for the correlation between different modes in a realistic environment is presented for a biconical and a circular microstrip antenna that employs higher order modes. It is found that the correlation is low enough to yield a diversity gain. Furthermore, the channel capacity of a MIMO system using a multimode antenna, i.e. an antenna employing multiple modes, is found to be comparable with the capacity of an array. Since only one element is needed, the multimode antenna offers several advantages over traditional arrays, and is an interesting antenna solution for future high capacity MIMO systems.

10.1 Introduction

The outstanding growth of wireless communication during the last years has resulted in an enormous demand for higher data rates. Recently, it has been proposed to use multiple antennas, i.e. arrays, at both the transmitter and the receiver to substantially increase the attainable data rate, i.e. the channel capacity. Impressive capacity gains using these MIMO systems have been shown for rich scattering environments, both theoretically and experimentally [FG98, LP00, WFGV98].

Unfortunately, several practical aspects pose problems for achieving the

predicted high data rates. Introducing antenna arrays at both the transmitter and receiver will undoubtedly increase the cost of producing the terminals, since it is expensive to manufacture, calibrate, and maintain antenna arrays with several elements. Furthermore, fitting several elements onto a small hand-set introduces several new and interesting challenges in antenna design. Practical problems with the feed and size requirements need to be overcome, while still providing an aesthetically pleasing design. Hence, the deployment of MIMO systems introduces several new challenges in antenna design that also encourages the development of new antenna solutions.

An interesting antenna solution, which appears well suited for MIMO systems is the multimode antenna. In a sense, the multimode antenna offers characteristics similar to those of an antenna array through multiple modes, but using only a single antenna element. Therefore, this type of antenna appears to be an interesting candidate for MIMO systems that are thought to operate in environments where employing many antenna elements may pose several problems related to size requirements, low cost implementations, mutual coupling, and design issues.

Employing the higher order modes of antennas has not received much attention in the literature. In [Vau88], a circular microstrip antenna employing higher order modes, was designed based on an element figure of merit that was introduced in [Vau86]. Recently, a modified biconical antenna that employs several modes to form directional beams was introduced in [DW98, ZHM+00]. The microstrip and biconical elements appear to be the only antennas that have been used to excite multiple modes in the context of beamforming and diversity. Hence, these antennas will be used in this chapter to examine the channel capacity using an antenna employing multiple modes, i.e. a multimode antenna, in a realistic propagation environment.

The correlation between the signals received by different modes will also be calculated since this is one of the most important properties when using several antennas for diversity or as a part of a MIMO system. If the received signals exhibit low correlation, it is likely that at least one mode or antenna yields a strong signal, while others may experience a deep fade. By properly combining the received signals from the different modes or antennas, a higher average SNR can be achieved. The correlation properties also essentially determine the success of MIMO systems, since these need several independent transmission paths to be able to support the desired high data rates. Hence, the correlation properties of multimode antennas will be studied and an analytical expression of the modal correlation will be derived. Based on the insight of the correlation analysis, the channel capacity of a MIMO system employing multimode antennas is then calculated in a typical channel scenario using a spatial channel model. First, however, the concept

of higher order modes will be introduced, and the properties of two antennas employing higher order modes will be briefly reviewed.

10.2 Multimode Antennas

A multimode antenna is an antenna where several modes are excited separately on the same antenna structure. The concept of modes is well known in the antenna community, but relatively unknown in the communications and signal processing communities. Essentially, different modes in electromagnetics represent different solutions to Maxwell's equations that fulfill the boundary conditions for the geometry at hand. In fact, it is possible to excite several modes at the same temporal frequency on one antenna structure separately, and regard these as separate antenna ports. Here, two different antenna structures, namely the biconical and microstrip, will be briefly described. These structures are then used in the following sections, where the correlation properties and the channel capacity of the resulting channel are studied.

10.2.1 Multimode Biconical Antenna

Many different types of antenna elements support excitation of several modes, but recently a multimode antenna based on a biconical element was proposed in [DW98]. A biconical antenna consists of two conical horns facing opposite directions as illustrated in Figure 10.1. Here, α denotes the flare angle and l the length of the horn. Note that the flare angle of the upper and lower cones are identical, i.e. the antenna is symmetric.

In the traditional analysis of the biconical antenna, a coaxial waveguide is used to excite the antenna. By enlarging the diameter of the coaxial feed, more solutions or modes are possible. The enlargement of the coaxial waveguide allows for higher order circular modes in the feed. These circular waves are then transformed into spherical waves that propagate on the biconical structure, that in turn radiates a field. It is important to note here that it is possible to design the overmoded coaxial feed in such a way that no modal coupling between modes occurs [DW98]. The radiated far-field from the higher order modes was in [DW98] found to be angularly dependent. This will show to be the key to achieve diversity by using multimode antennas.

The expressions for the transmitted far-field is derived by integrating the fields of an infinite biconical transmission line over the biconical aperture [Bal89, Dem98]. The far-field expression for mode m can be written [DW98,

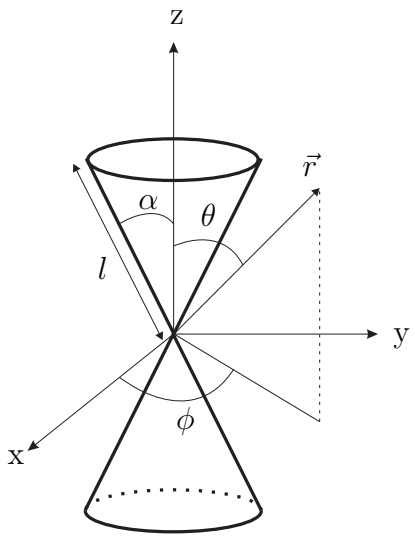


Figure 10.1: The geometry of the biconical antenna and the relation to the coordinates.

Sva00h] as

$$\vec{E}_m = \vec{\theta}\,\frac{e^{-jkr}}{r} G_m(\theta, \phi, l, \alpha), \tag{10.1}$$

where k denotes the wave number[1], r the radial distance from the antenna, and $G_m(\theta, \phi, l, \alpha)$ denotes the element pattern of mode m. Note that the fields of the different modes are simply the solutions of Maxwell's equations that fulfill the boundary conditions of the biconical structure for specific values of l and α. The full expression for $G_m(\theta, \phi, l, \alpha)$ is quite complicated and is given in Appendix 10A. However, the azimuthal dependency of the radiation pattern is easily factored out as

$$G_m(\theta, \phi, l, \alpha) = \cos m\phi \; G'_m(\theta, l, \alpha), \tag{10.2}$$

where $G'_m(\theta, l, \alpha)$ is the remaining azimuth independent term of the radiation pattern.

10.2.2 Multimode Circular Microstrip Antenna

There is a large volume of literature devoted to microstrip antennas. Nearly all of the literature deal with the fundamental mode. However, [Der79,

[1]The wave number $k = \frac{2\pi}{\lambda} = \omega\sqrt{\mu\epsilon}$, where λ is the wavelength. The symbols ϵ and μ denote the dielectric constant and the permeability of the medium respectively.

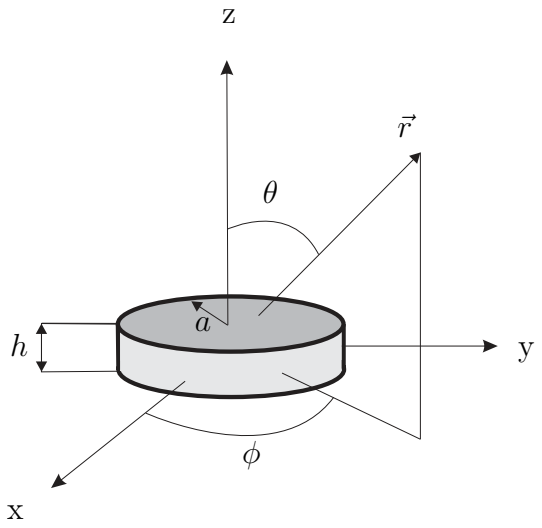


Figure 10.2: The geometry of the circular microstrip antenna and the relation to the coordinates.

How75, Hua84, LSR79] allude to some higher order modes. Furthermore, in [Vau88] a two-port higher order mode circular microstrip antenna, based on a figure of merit calculation [Vau86], is proposed.

An illustration of a circular microstrip antenna is shown in Figure 10.2, where a denotes the radius and h denotes the thickness. The upper circular conductor (shaded dark-gray) is separated from the underlying ground plane by a dielectric substrate (shaded light-gray). The main principle behind the antenna is that the upper conductor and the ground plane together form a circular slot that radiates. Typically, this type of antenna is excited using a co-axial feed or a microstrip feed. Details regarding the design of microstrip antennas can be found in antenna textbooks. Here, only a simple model for the radiation of this type of antenna is needed. For the purposes of this paper, an adequate model of the radiation of mode m is to view the antenna as a magnetic ring current with an edge voltage $V_m = V_m^0 \cos m\phi$ [Der79, JHW81, Vau88]. Using this simple model, the radiated far-field may be written as

$$\vec{E}_m = \frac{e^{-jkr}}{r} E_{m,\theta}\vec{\theta} + E_{m,\phi}\vec{\phi}, \tag{10.3}$$

where

$$\begin{aligned} E_{m,\theta} &= \frac{j^m V_m^0 ka}{2} \cdot (J_{m+1}(ka\sin\theta) - J_{m-1}(ka\sin\theta)) \cdot \cos m\phi \\ E_{m,\phi} &= \frac{j^m V_m^0 ka}{2} \cdot (J_{m+1}(ka\sin\theta) + J_{m-1}(ka\sin\theta)) \cdot \cos\theta \sin m\phi \ . \end{aligned} \tag{10.4}$$

Here, $J_n(x)$ denotes the Bessel function of the first kind and of order n. More involved models and more accurate solutions can be found in antenna textbooks. For the purposes here, the above solution is sufficient. Interestingly, it is found that the expressions for the multimode microstrip antenna are very similar those of the multimode biconical antenna. If only the horizontal plane ($\theta = 90°$) is considered, the $\vec{\phi}$ component of the field vanishes, and only a $\vec{\theta}$ component remains. Thus, the far-field expression of the biconical antenna in (10.1) is also valid for the microstrip antenna. In fact, the ϕ dependency of the microstrip antenna reduces to $\cos m\phi$, so also the expression for the radiation pattern of the bicone in (10.2) is applicable. The same analysis regarding correlation and capacity is thereby valid for both the biconical and the microstrip antennas with higher order modes. In the following, both of these antennas will be referred to as a multimode antenna.

In a multipath scenario, the radiation patterns of (10.2) for the different modes are different enough to yield significantly different received signals. For example, one mode may be in a deep fade while the others still receive a strong signal. Thus, the signals received by the different modes may be combined to increase the average SNR, i.e. yield a diversity gain. However, the potential diversity gain of employing a multimode antenna depends strongly on the correlation between the different modes. Hence, an expression for the correlation between different modes will be derived in the next section for a typical channel scenario encountered in practice.

10.3 Modal Correlation

The correlation between the different modes ultimately determines the diversity gain, but also the potential increase in capacity that can be achieved by employing the multimode antenna as a MIMO receiver. Therefore, an expression for the correlation between two modes will be derived and examined in different channel scenarios. The correlation between signals received by two spatially separated antennas has been investigated in numerous works, but the modal correlation has not been treated in any detail.

The level of correlation depends on the angular distribution of the received power, i.e. the Power Azimuth Spectrum (PAS). Several different

PAS models have been proposed in the literature. However, recent measurement campaigns have shown that for both urban and rural areas, a Laplacian distribution offers the best fit [PMF98]. Therefore, the modal correlation calculations will use a truncated Laplacian model of the PAS $P(\phi)$

$$P(\phi) = ce^{-\sqrt{2}\frac{|\phi-\bar{\phi}|}{\sigma_\phi}} \quad \phi \in [-\pi, \pi), \tag{10.5}$$

where the value of the constant c is chosen so as to get a probability weight of one, i.e. $\int_{-\pi}^{\pi} P(\phi)d\phi = 1$. Note that $\bar{\phi}$ denotes the main azimuth direction, i.e. $\bar{\phi} = \int_{-\pi}^{\pi} \phi P(\phi)d\phi$ and that the parameter σ_ϕ controls the angular spread.

The received signals by mode m and n, assuming L multipaths, can be written as

$$s_m = \sum_{l=1}^{L} \alpha_l e^{j\beta_l} \cos(m\phi_l) \tag{10.6}$$

$$s_n = \sum_{l=1}^{L} \alpha_l e^{j\beta_l} \cos(n\phi_l), \tag{10.7}$$

where α_l is a random scattering amplitude and β_l is a random phase of path l. It is further assumed that the phase is uniformly distributed $\beta \in U[-\pi, \pi)$ and that both β and ϕ are independent stochastic quantities. Here, it is also assumed that α and β are independent.

Now, an approximative expression for the envelope correlation $\rho_e(m, n)$ between the signals s_m and s_n can be obtained from the correlation coefficient ρ_{mn} as $\rho_e(m, n) \approx |\rho_{mn}|^2$ [PS60]. Tedious, but straightforward, calculations then give the envelope correlation coefficient $\rho_e(m, n)$ as

$$\rho_e(m, n) \approx |\rho_{mn}|^2 = \frac{|E[s_m s_n^*]|^2}{E[|s_m|^2]E[|s_n|^2]} = \frac{\kappa^2(m, n)}{\kappa(m, m)\kappa(n, n)}, \tag{10.8}$$

where

$$\begin{aligned}\kappa(m, n) = &\frac{\sqrt{2}\sigma_\phi \cos(m-n)\bar{\phi}}{2 + \sigma_\phi^2(m-n)^2} + \frac{\sqrt{2}\sigma_\phi \cos(m+n)\bar{\phi}}{2 + \sigma_\phi^2(m+n)^2} - \\ &- \sqrt{2}\sigma_\phi e^{-\frac{\sqrt{2}\pi}{\sigma_\phi}} \left(e^{\frac{\sqrt{2}\bar{\phi}}{\sigma_\phi}} + e^{-\frac{\sqrt{2}\bar{\phi}}{\sigma_\phi}} \right) \left[\frac{(-1)^{m-n}}{2 + \sigma_\phi^2(m-n)^2} + \frac{(-1)^{m+n}}{2 + \sigma_\phi^2(m+n)^2} \right].\end{aligned} \tag{10.9}$$

The full derivation of the envelope correlation coefficient is given in Appendix 10B. In Figure 10.3, the envelope correlation between two modal combinations $(1, 2)$ and $(2, 5)$ is evaluated for different azimuth angles $\bar{\phi}$, and compared to the correlation between two spatially separated antennas [PMF98]. Note that the azimuth spread in this comparison is 20°, which

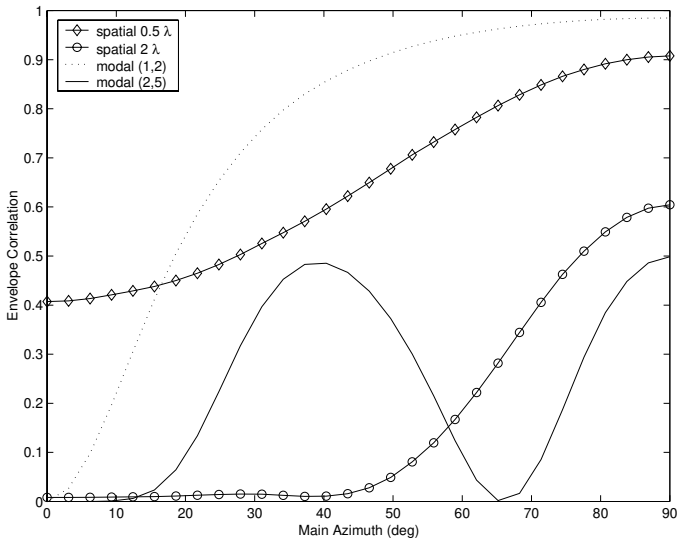


Figure 10.3: The envelope correlation of two spatially separated elements and two modal correlations $\rho_e(1,2)$ and $\rho_e(2,5)$ versus main azimuth $\bar{\phi}$ with an angular spread σ_ϕ of 20°.

could constitute a dense urban environment [PMF98]. The correlation between modes $(1,2)$ is quite high, but the correlation between $(2,5)$ is well below 0.7, that usually is taken as a rule of thumb limit for obtaining "uncorrelated" channels [PS60]. Hence, the modal correlation is low enough to give a diversity gain [Sva00h], and thereby also has the potential of increasing the channel capacity. Since the correlation is strongly dependent on the main azimuth angle, a natural performance measure is therefore to calculate the correlation averaged over azimuth. The average correlation between the modes from the scenario in Figure 10.3 becomes in matrix form

$$\bar{\mathbf{R}} = \begin{bmatrix} 1.00 & 0.73 & 0.54 & 0.36 & 0.23 \\ 0.73 & 1.00 & 0.37 & 0.45 & 0.22 \\ 0.54 & 0.37 & 1.00 & 0.46 & 0.27 \\ 0.36 & 0.45 & 0.46 & 1.00 & 0.45 \\ 0.23 & 0.22 & 0.27 & 0.45 & 1.00 \end{bmatrix}. \tag{10.10}$$

Most of the modal correlations are below 0.7, and thus can contribute to a diversity and capacity gain. In fact, the two correlations shown in Figure 10.3, correspond to the highest and lowest modal correlations.

How the modal correlation depends on the azimuth spread is shown in Figure 10.4, where again the highest and lowest modal correlations are shown.

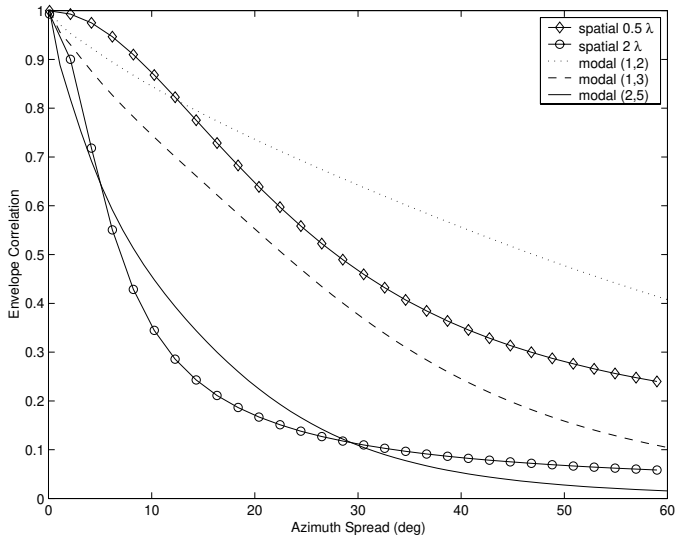


Figure 10.4: The average envelope correlation of two spatially separated elements and two modal correlations $\rho_e(1,2)$, $\rho_e(1,3)$, and $\rho_e(2,5)$ versus the angular spread.

Also shown is the correlation between modes (1,3) that represents a typical correlation level. The figure clearly indicates that the modal correlation decreases with increased azimuth spread, as in the case of spatially separated antennas. The conclusion from the correlation analysis is thus that the modal correlations are generally slightly higher than the spatial correlations for well separated elements, but still low enough to give a diversity and capacity gain. Of course, the level of correlation for the spatially separated antennas is strongly dependent on the element separation distance. However, it is important to remember that the multimode diversity branches are obtained using only one antenna element but several modes.

10.4 Multimode Channel Capacity

Recently, a number of researchers have pointed out the remarkable channel capacity gains available using multi-element antennas at both the transmitter and receiver. Typically, the analysis of these systems, often called Multi-Input Multi-Output (MIMO) systems, have used the rather ideal assumptions of independently fading channels and no mutual coupling. Correlated fading at the antenna elements could pose a serious problem, typically at the mobile hand-set where the elements are closely spaced. Thus, the multimode

antenna appears to be suitable for this type of application, since only a single antenna element is needed and there is no coupling between modes.

The instantaneous channel capacity of a MIMO channel $\mathbf{H}$ can be written as [FG98]

$$C = \log_2 \det \left[\mathbf{I} + \frac{\xi}{N} \mathbf{H}\mathbf{H}^H \right], \tag{10.11}$$

where ξ/N denotes the SNR at each receive antenna. Note that the channel matrix $\mathbf{H}$ is $M \times N$, where M, N denotes the number of antenna elements at the receiver and transmitter, respectively. Furthermore, the element $\mathbf{H}_{ij}$ represents the complex path gain from transmitter j to receiver i. A simple, yet detailed, channel model that includes the spatial dimension is to place a circular disc of uniformly distributed scatterers around the mobile. Although the full model presented in Chapter 9 can be applied also here, the polarization properties and the attenuation of the waves are not important in this application. Thus, a reduced version that still includes the radiation patterns of the antennas is better suited for this channel scenario. Hence, the following model can be used

$$\mathbf{H}_{ij} = \sum_{l=1}^{L} \alpha_l e^{jk(r_{t_j s_l} + r_{s_l r_i})} g_{t_j}(\phi_l^T) g_{r_i}(\phi_l^R), \tag{10.12}$$

where the number of scatterers is denoted L, and α_l is a complex gaussian distributed reflection coefficient with zero mean and unit variance. Furthermore, $r_{t_j s_l}$ and $r_{s_l r_i}$ denote the distances from mobile antenna j to scatterer l and from scatterer l to base antenna i respectively. Finally, $g_{t_j}(\phi_l^T)$ and $g_{r_i}(\phi_l^R)$ denote the radiation patterns of the antennas. The channel is normalized in the sense that $||\mathbf{H}||_2 = 1$. It should be noted that the normalization of the channel (or SNR) removes the signal strength and focuses on the richness of the scattering. In practice, both will of course affect the capacity. In fact, it has been found in [MBKF00] that a high capacity is still obtained in LOS scenarios with high correlations due to a high SNR value.

With the above channel model, the channel capacity is a random value, and in Figure 10.5, the Complementary Cumulative Distribution Function (CCDF) of the capacity for a scenario with $L = 20$ scatterers and $\xi/N =$ 10dB is shown. The capacity is calculated when a multimode antenna is used at the mobile station (MS), and when a multimode antenna is used at both the MS and the base station (BS). These capacity figures are then compared to the corresponding figures when using an array at both MS and BS. To simplify the comparison, the calculations are performed for an equal number of antenna ports at the MS and BS, $M = N = 4$. Also included in the comparison is the capacity for an independently fading identically distributed

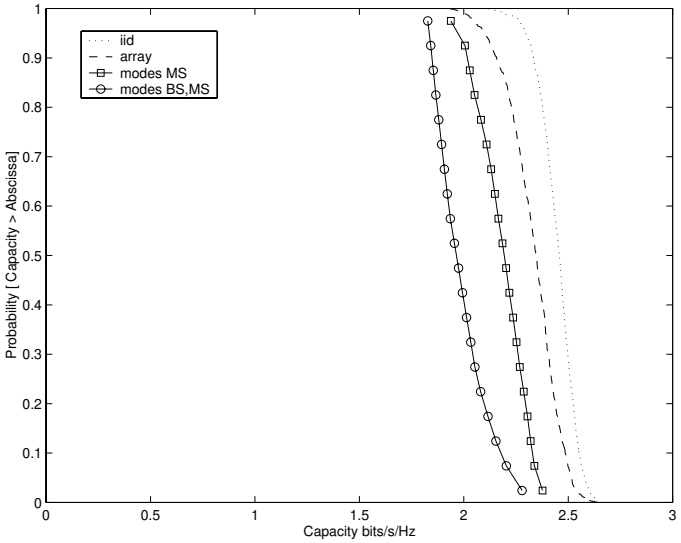


Figure 10.5: The CCDF of C with $M = N = 4$ for an array at both MS and BS, multimode at MS, multimode at both MS and BS, and independent Rayleigh channel. For the array, the element separation is $d_{bs} = 2\lambda$ and $d_{ms} = 0.5\lambda$. The scatterers are distributed on a disc of radius $R = 35\lambda$, the MS-BS distance is $D = 50\lambda$, and $\bar{\phi} = 45°$.

Rayleigh channel [FG98]. For the array, the element separation is $d_{bs} = 2\lambda$ at the BS and $d_{ms} = 0.5\lambda$ at the MS. Furthermore, the radius of the scatter disc is $R = 35\lambda$, the distance between the MS and BS is $D = 50\lambda$, and the main azimuth towards to mobile is $\bar{\phi} = 45°$. The scattering environment is rich enough in order for the array to be close to the theoretical curve. Using a multimode antenna at the mobile results in only a slightly lower capacity. Using a multimode antenna also at the BS results in lower capacity, since seen from the BS, the angular spread is much lower than at the MS. A lower angular spread gives a lower capacity, since it will result in a higher correlation, see Figure 10.4.

The correlation between the different modes also strongly depends on the main azimuth angle, as found in Figure 10.3. How the capacity depends on the main azimuth is shown in Figure 10.6. The dependency is found to be much less pronounced than in the correlation study. Note that in Figure 10.6, the capacity obtained at 90% of the time (i.e. 10% outage) is shown. The antenna array of course performs worse at $\bar{\phi} = 90°$ (end-fire), but the differences are in general small. Thus, the capacity is similar for most azimuth orientations for both the array and the multimode antenna.

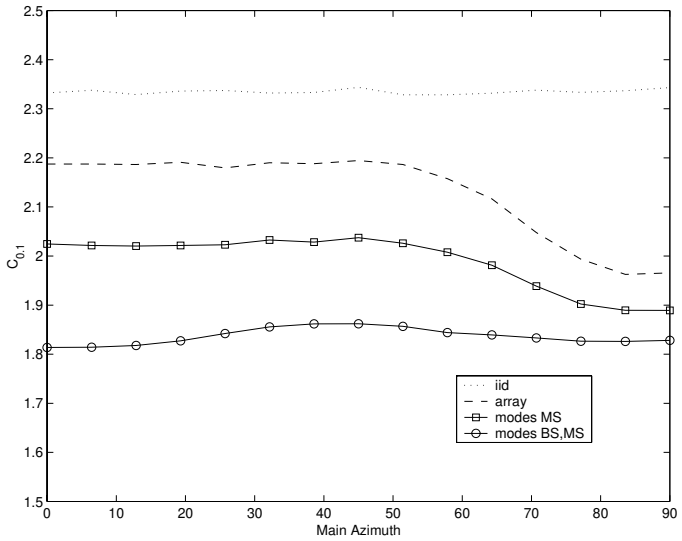


Figure 10.6: The channel capacity with 10% outage probability versus main azimuth for the same scenario as in Figure 10.5.

The capacity with 10% outage versus the radius of the scatter disc is shown in Figure 10.7. A richer scattering environment is obtained with increased radius, and thus the capacity increases as expected. This is in agreement with Figure 10.4, where it was found that the correlation decreases with increased azimuth spread. The increase in capacity is smaller for the multimode antenna than for the array, but the differences in capacity are not large. If the number of scatterers is increased to a large value, the capacity of the array attains the capacity of the Rayleigh channel, while the multimode antenna does not. However, here it is important to note that only one antenna element is used for the multimode receiver, while the antenna array employs four elements.

10.5 Conclusions

A novel way of exploiting diversity by employing the higher order modes of antennas was introduced. Two different types of multimode antennas were studied, namely a biconical and a circular microstrip antenna. Essentially, multimode antennas offer characteristics similar to an antenna array through multiple modes using just a single element. An analytical expression for the correlation between different modes in a realistic environment was presented, and it was found that the correlation is low enough to yield a diversity gain.

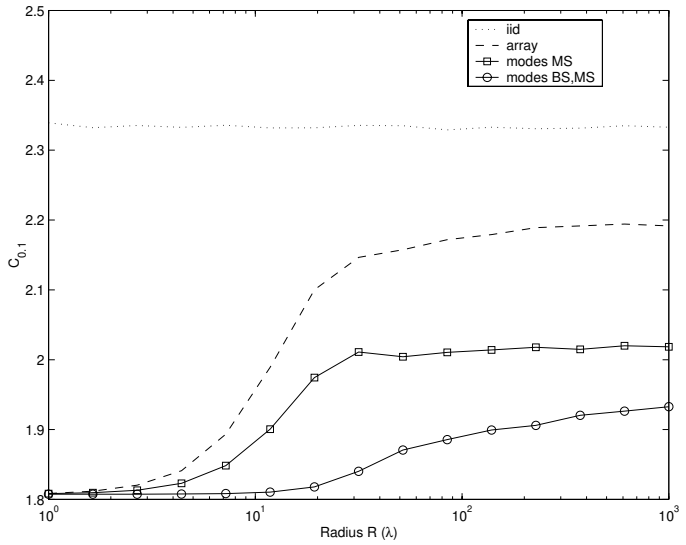


Figure 10.7: The channel capacity with 10% outage probability versus the radius of the scatter disc for the same scenario as in Figure 10.5.

Furthermore, the channel capacity of a MIMO system employing multiple antennas or modes at the MS and BS was simulated using a spatial channel model.

Surprisingly, it was found that the multimode antenna offers capacity gains similar to that of the array antenna for realistic MIMO scattering environments. Since the multimode antenna only requires a single antenna element, it avoids some of the problems of an array such as calibration problems and mutual coupling. Furthermore, the fact that only a single antenna structure is needed, a compact antenna can be obtained that also provides new possibilities in designing aesthetically pleasing antennas. One benefit of the biconical antenna is that it can be manufactured at very low cost, since it contains no expensive parts or materials [DW98]. However, the microstrip antenna may be more interesting when considering mobile hand-set implementations, due to its shape and "paste on" features. Thus, employing multimode antennas is an interesting antenna solution for future high capacity MIMO systems that offers several advantages over traditional arrays.

The concept of multimode antennas does not end with the antenna shapes considered here. In principle, most antennas can support higher order modes. Hence, it is an interesting antenna design topic to find and evaluate other structures with perhaps better characteristics than the microstrip and biconical antenna. More work on the antenna design is needed to find suitable

positions of the feeds, appropriate dimensions, and a matching network that minimizes the losses in the antenna. It is also important to note that in this chapter, only the potential of exploiting the individual modes are examined. Many other possibilities exists. For instance, a combination of spatial, modal, and polarization diversity is an interesting alternative that may offer very good performance. Furthermore, only the possibility of using the individual modes as diversity branches was studied in this chapter. It is also possible to use several modes to form narrow beams, that each could be used as a diversity branch. In that approach as well as the approach analyzed in this chapter, the diversity is based on the fact that the different signals are obtained through different radiation patterns. Hence, these diversity schemes are often called pattern or angular diversity, which for example has been studied in [Vau98], where the necessary beam separation of multibeam antennas in a uniformly distributed PAS has been treated. Thus, there are many interesting possibilities of using higher order modes of antennas in future high capacity MIMO systems.

10A Radiation Pattern of Biconical Antenna

The element pattern for the TEM mode can be written as [Dem98]

$$G_{TEM}(\theta, l, \alpha) = \frac{j\omega\mu I_0 l}{4\sqrt{\pi \ln \cot \frac{\alpha}{2}}} \cdot \int_{\theta'} \left[(u_1+1) j J_1(klu_2) + u_2 J_0(klu_2)\right] e^{jklu_1} d\theta', \quad \text{(10A.1)}$$

where I_0 is the input current, $u_1 = \cos\theta\cos\theta'$, $u_2 = \sin\theta\sin\theta'$, and J_k denotes the cylindrical Bessel function of order k [AS70, Bal89]. Note that the expression is independent of ϕ.

The element pattern for the TE_{mn} mode can be written as [Dem98]

$$G_{TE}(\theta, \phi, l, \alpha) = \cos m\phi \; G'_{TE}(\theta, l, \alpha), \quad \text{(10A.2)}$$

where

$$G'_{TE}(\theta,l,\alpha) = \frac{-j\omega\mu I_0 l j^{m-1} m}{2}\cdot \tag{10A.3}$$
$$\left[\int_{\theta'}\left(\frac{(u_1+1)m}{klu_2}+ju_2\right)L(\theta')e^{jklu_1}J_m(klu_2)d\theta'\right.$$
$$-\int_{\theta'}(u_1+1)L(\theta')e^{jklu_1}J_{m+1}(klu_2)d\theta'$$
$$\left.+\int_{\theta'}\frac{(\cos\theta'-\cos\theta)}{kl\sin\theta}\frac{\partial L(\theta')}{\partial\theta'}e^{jklu_1}J_m(klu_2)d\theta'\right]$$

and

$$L(\theta) \quad = \quad P_n^m(\cos\theta)\frac{\partial P_n^m(-\cos\alpha)}{\partial\theta} \quad - \quad P_n^m(-\cos\theta)\frac{\partial P_n^m(\cos\alpha)}{\partial\theta}. \tag{10A.4}$$

Here, $P_n^m(x)$ denotes the associated Legendre function of the first kind [AS70, Bal89]. The far-field expressions (10A.1)-(10A.3), although quite complicated, are easily evaluated numerically using a computer.

10B Correlation Analysis

The signals received by antenna 1 and 2, assuming L multipaths, can be written as

$$s_1 = \sum_{l=1}^{L}\alpha_l e^{j\beta_l}g_1(\phi_l) \tag{10B.1}$$

$$s_2 = \sum_{l=1}^{L}\alpha_l e^{j\beta_l}e^{j\Delta}g_2(\phi_l), \tag{10B.2}$$

where α_l is a scattering amplitude and β_l is a random phase of path l. Furthermore, $\Delta = \frac{2\pi d}{\lambda}\sin\phi_l$, where d is the separation distance between the antennas and ϕ_l is the azimuth angle of arrival of path l. The radiation pattern of antenna i for path l is denoted $g_i(\phi_l)$ and the azimuth angle is measured relative the normal of the array axis. Furthermore, the phase is assumed to be uniform $\beta \in U[-\pi,\pi)$, and both α and ϕ are stochastic quantities. It is also assumed that β and ϕ are independent. Finally, it assumed that α and β are independent. Using this formulation for the signals s_1 and s_2, both multimode and array antennas are included in the analysis.

Now, the envelope correlation ρ_e between the signals s_1 and s_2 can be calculated, approximately, from the correlation coefficient ρ_{12} as [PS60]

$$\rho_e \approx |\rho_{12}|^2. \tag{10B.3}$$

Due to the independence assumption and that the phase is uniformly distributed, the mean values of the signals are zero, and thus the expression for the correlation coefficient becomes

$$\rho_{12} = \frac{E[s_1 s_2^*]}{\sqrt{E[|s_1|^2]E[|s_2|^2}}. \tag{10B.4}$$

By introducing the real and imaginary part of the signals, $s = x + jy$, the expression for the numerator can be written as

$$\begin{aligned} E[s_1 s_2^*] &= E[(x_1 + jy_1)(x_2 + jy_2)^*] = \\ &= E[x_1 x_2] + E[y_1 y_2] + j\left(E[x_2 y_1] - E[x_1 y_2]\right). \end{aligned} \tag{10B.5}$$

Note that the expectation is taken over α, β, and ϕ. The first term in (10B.5) can be written as

$$\begin{aligned} E[x_1 x_2] &= E\Big[\left(\sum_{l=1}^{L} \alpha_l \cos(\beta_l) g_1(\phi_l)\right) \times \left(\sum_{l=1}^{L} \alpha_l \cos(\beta_l + \Delta(\phi_l)) g_2(\phi_l)\right)\Big] = \\ &= LE\left[E\left[\cos(\beta)\cos(\beta + \Delta(\phi))|\phi\right] \alpha^2 g_1(\phi) g_2(\phi)\right]. \end{aligned} \tag{10B.6}$$

Since

$$E\left[\cos(\beta)\cos(\beta + \Delta(\phi))|\phi\right] = \frac{\cos\Delta}{2}, \tag{10B.7}$$

the expression in (10B.6) becomes

$$\begin{aligned} E[x_1 x_2] &= \frac{L}{2} E\left[E\left[\alpha^2|\phi\right]\cos(\Delta) g_1(\phi) g_2(\phi)\right] = \\ &= \frac{L}{2}\int_{-\pi}^{\pi} \cos(\Delta) g_1(\phi) g_2(\phi) E\left[\alpha^2|\phi\right] p_\Phi(\phi) d\phi, \end{aligned} \tag{10B.8}$$

where $p_\Phi(\phi)$ denotes the probability density function for the azimuth angle. However, this distribution is related to the PAS [Ped00] as

$$P(\phi) = LE\left[\alpha^2|\phi\right] p_\Phi(\phi). \tag{10B.9}$$

Therefore,

$$E[x_1 x_2] = \frac{1}{2}\int_{-\pi}^{\pi} \cos(\Delta) g_1(\phi) g_2(\phi) P(\phi) d\phi. \tag{10B.10}$$

In a similar manner,

$$\begin{aligned} E[y_1 y_2] &= LE\left[E\left[\sin(\beta)\sin(\beta + \Delta(\phi))|\phi\right] \alpha^2 g_1(\phi) g_2(\phi)\right] = \\ &= \frac{L}{2} E\left[E\left[\alpha^2|\phi\right]\cos(\Delta) g_1(\phi) g_2(\phi)\right] = E[x_1 x_2], \end{aligned} \tag{10B.11}$$

since

$$E\left[\sin(\beta)\sin(\beta+\Delta(\phi))|\phi\right]=\frac{\cos\Delta}{2}. \tag{10B.12}$$

Furthermore,

$$\begin{aligned}E[x_2y_1]&=LE\left[E\left[\sin(\beta)\cos(\beta+\Delta(\phi))|\phi\right]\alpha^2g_1(\phi)g_2(\phi)\right]=\\&=-\frac{L}{2}E\left[E\left[\alpha^2|\phi\right]\sin(\Delta)g_1(\phi)g_2(\phi)\right]=-\frac{1}{2}\int_{-\pi}^{\pi}\sin(\Delta)g_1(\phi)g_2(\phi)P(\phi)d\phi,\end{aligned} \tag{10B.13}$$

since

$$E\left[\sin(\beta)\cos(\beta+\Delta(\phi))|\phi\right]=-\frac{\sin\Delta}{2}. \tag{10B.14}$$

Also,

$$\begin{aligned}E[x_1y_2]&=LE\left[E\left[\cos(\beta)\sin(\beta+\Delta(\phi))|\phi\right]\alpha^2g_1(\phi)g_2(\phi)\right]=\\&=\frac{L}{2}E\left[E\left[\alpha^2|\phi\right]\sin(\Delta)g_1(\phi)g_2(\phi)\right]=-E[x_2y_1],\end{aligned} \tag{10B.15}$$

since

$$E\left[\cos(\beta)\sin(\beta+\Delta(\phi))|\phi\right]=\frac{\sin\Delta}{2}. \tag{10B.16}$$

Finally,

$$\begin{aligned}E[|s_1|^2]=E[x_1^2+y_1^2]&=E\Big[\left(\sum_{l=1}^{L}\alpha_l\cos(\beta_l)g_1(\phi_l)\right)^2\times\\&\qquad\left(\sum_{l=1}^{L}\alpha_l\sin(\beta_l+\Delta(\phi_l))g_1(\phi_l)\right)^2\Big]=\\&=LE\left[\alpha^2\cos^2\beta g_1^2(\phi)+\alpha^2\sin^2\beta g_1^2(\phi)\right]=LE[\alpha^2g_1^2(\phi)]\end{aligned} \tag{10B.17}$$

and

$$E[|s_2|^2]=LE[\alpha^2g_2^2(\phi)]. \tag{10B.18}$$

Employing (10B.4), (10B.5), (10B.11), and (10B.15)-(10B.18), the correlation coefficient can be written as

$$\rho_{12}=\frac{2E[x_1x_2]+2jE[x_2y_1]}{\sqrt{L^2E[\alpha^2g_1^2(\phi)]E[\alpha^2g_2^2(\phi)]}}, \tag{10B.19}$$

where $E[x_1x_2]$ is given in (10B.10) and $E[x_2y_1]$ is given in (10B.13). Finally, using (10B.3), the envelope correlation coefficient becomes

$$\rho_e=\frac{\rho_{rr}^2+\rho_{ri}^2}{\rho_1\rho_2}, \tag{10B.20}$$

where

$$\begin{aligned}\rho_{rr} &= \int_{-\pi}^{\pi} \cos\left(\frac{2\pi d}{\lambda}\sin\phi\right) g_1(\phi)g_2(\phi)P(\phi)d\phi \\ \rho_{ri} &= \int_{-\pi}^{\pi} \sin\left(\frac{2\pi d}{\lambda}\sin\phi\right) g_1(\phi)g_2(\phi)P(\phi)d\phi \\ \rho_{1,2} &= \int_{-\pi}^{\pi} g_{1,2}^2(\phi)P(\phi)d\phi\end{aligned} \tag{10B.21}$$

The expressions for the envelope correlation coefficient for the antenna array and the multimode antenna are now readily obtained by inserting the corresponding radiation functions and locations. Here, the correlation evaluations will be performed using a truncated Laplacian model of the PAS

$$P(\phi) = ce^{-\sqrt{2}\frac{|\phi-\bar{\phi}|}{\sigma_\phi}} \quad \phi \in [-\pi, \pi), \tag{10B.22}$$

where the value of the constant c is chosen so as to get a probability weight of one, i.e. $\int_{-\pi}^{\pi} P(\phi)d\phi = 1$. Note that $\bar{\phi}$ denotes the main azimuth direction, i.e. $\bar{\phi} = \int_{-\pi}^{\pi} \phi P(\phi)d\phi$ and that the parameter σ_ϕ controls the angular spread. The Laplacian PAS has been experimentally verified in several measurement campaigns [PMF98].

Spatial correlation: The Antenna Array Assuming omnidirectional elements, i.e. $g_1(\phi) = g_2(\phi) = 1$ results in the following expressions

$$\begin{aligned}\rho_e &= \rho_{rr}^2 + \rho_{ri}^2 \\ \rho_{rr} &= \int_{-\pi}^{\pi} \cos\left(\frac{2\pi d}{\lambda}\sin\phi\right) e^{-\sqrt{2}\frac{|\phi-\bar{\phi}|}{\sigma_\phi}} d\phi \\ \rho_{ri} &= \int_{-\pi}^{\pi} \sin\left(\frac{2\pi d}{\lambda}\sin\phi\right) e^{-\sqrt{2}\frac{|\phi-\bar{\phi}|}{\sigma_\phi}} d\phi,\end{aligned} \tag{10B.23}$$

since $\rho_1 = \rho_2 = 1$ due to the omnidirectionality. The expressions in (10B.23) were also given in [PMF98], however without a detailed derivation.

Modal correlation: The multimode antenna Assuming that one multimode antenna is used, the separation distance will vanish (i.e. $d = 0$) and the radiation patterns of antenna port 1 and 2, i.e. mode m and n, are

$$\begin{aligned}g_1(\phi) &= cos(m\phi) \\ g_2(\phi) &= cos(n\phi).\end{aligned} \tag{10B.24}$$

Inserting (10B.24) into the expression for the envelope correlation coefficient (10B.20) and (10B.21) gives

$$\rho_e = \frac{\left(\int_{-\pi}^{\pi} \cos(m\phi)\cos(n\phi) e^{-\sqrt{2}\frac{|\phi-\bar{\phi}|}{\sigma_\phi}} d\phi\right)^2}{\int_{-\pi}^{\pi} \cos^2(m\phi) e^{-\sqrt{2}\frac{|\phi-\bar{\phi}|}{\sigma_\phi}} d\phi \int_{-\pi}^{\pi} \cos^2(n\phi) e^{-\sqrt{2}\frac{|\phi-\bar{\phi}|}{\sigma_\phi}} d\phi} \tag{10B.25}$$

Evaluation of the integrals give

$$\rho_e(m,n) = \frac{\kappa^2(m,n)}{\kappa(m,m)\kappa(n,n)}, \tag{10B.26}$$

where

$$\begin{aligned}\kappa(m,n) = \frac{\sqrt{2}\sigma_\phi \cos(m-n)\bar{\phi}}{2+\sigma_\phi^2(m-n)^2} + \frac{\sqrt{2}\sigma_\phi \cos(m+n)\bar{\phi}}{2+\sigma_\phi^2(m+n)^2} - \\ - \sqrt{2}\sigma_\phi e^{-\frac{\sqrt{2}\pi}{\sigma_\phi}} \left(e^{\frac{\sqrt{2}\bar{\phi}}{\sigma_\phi}} + e^{-\frac{\sqrt{2}\bar{\phi}}{\sigma_\phi}}\right) \times \\ \left[\frac{(-1)^{m-n}}{2+\sigma_\phi^2(m-n)^2} + \frac{(-1)^{m+n}}{2+\sigma_\phi^2(m+n)^2}\right].\end{aligned} \tag{10B.27}$$

Chapter 11

Employing Switched Parasitic Antennas in MIMO systems

A novel technique for obtaining diversity in MIMO systems via electronically directing the radiation pattern of a Switched Parasitic Antenna (SPA), is studied. The correlation between the received signal modes are shown to be sufficiently low to yield a diversity gain. The capacity limit using the SPA is investigated for different SPA configurations and it is found that the capacity is comparable with an array antenna configuration in certain situations. Finally, a space time block coding scheme is used to evaluate the bit error rate of a MIMO-SPA system. It is found that the SPA requires a 5 dB higher SNR than an antenna array solution to achieve a BER=10^{-2}. However, the array antenna requires a radio transceiver for every antenna, as opposed to the SPA which uses only one transceiver.

11.1 Introduction

Recent information theory results have demonstrated an enormous capacity potential of wireless systems with multiple antennas at both transmitter and receiver, so called Multi-Input Multi-Output (MIMO) systems [FG98]. An unfortunate aspect of MIMO systems is the high cost of multiple radio transceivers at the access point and at the user terminal. Furthermore, it is expensive to calibrate and maintain antenna arrays with many antenna elements. Recently, Switched Parasitic Antennas (SPAs) have been subject to an increased interest in the literature, for improving capacity in indoor Local Area Networks LANs [AKT00], as a diversity antenna [SLTV99, Vau99] and for tracking of base-stations [PTS+98].

In a sense, SPA offers characteristics similar to an array antenna with several fixed beams, but is more compact in size, and might be more suitable on certain mobile equipments. This chapter will examine the use of SPAs in MIMO systems, where the SPA is primarily used at the User Equipment (UE), in a realistic flat fading environment. It has been observed that when the fades of the MIMO receive channels are correlated, the channel capacity can be significantly smaller than when the fades are independent and identically distributed (i.i.d.) [SFGK00]. Hence, the capacity of the SPA in this type of scenarios is studied and compared to the capacity of a completely uncorrelated scenario. Furthermore, Space Time Block Coding (STBC) techniques are applied to examine how this correlation affects the BER of a SPA system.

11.2 Switched Parasitic Antennas

SPAs offering directional patterns dates back to the early work of Yagi and Uda in the 1930's. For mobile communications, Vaughan [Vau99] gave some examples of this technology to provide angle diversity by employing two parasitics on a mobile phone handset. The concept is to use a single active antenna element, connected to a radio transceiver, in a structure with one or several passive antenna elements, operating near resonance. The passive elements are called Parasitic Elements (PEs) and act together with the active element to form an array, as in the well known Yagi-Uda array [Bal82]. To alter the radiation pattern, the termination impedances of the PEs are switchable, to change the current flowing in those elements. The PE become reflectors when shorted to the ground plane using *pin* diodes [STLO00] and when not shorted, the PE have little effect on the antenna characteristics. The receiver is always connected to the center antenna element so there are no switches in the RF direct signal path.

The parasitic antennas can be designed using monopoles on a ground plane [STLO00, Vau99] or as parasitic patch antennas [PTL$^+$97]. Here, a monopole on the ground plane will be analyzed due to its omnidirectional properties. Examples of parasitic antennas are shown in Figure 11.1 for a 4-direction symmetry and in Figure 11.2 for a 3-direction symmetric antenna. The antenna in Figure 11.2 have an additional circle of parasitic elements that always are shorted to ground. The effect of this arrangement is an increased directivity as their length are shorter than the corresponding resonant length ($\approx \lambda/4$) and will lead the induced emf [Bal82].

The lengths and distances displayed in Figure 11.1 and 11.2 are not optimal in any way. The lengths can be adjusted to give the antenna certain

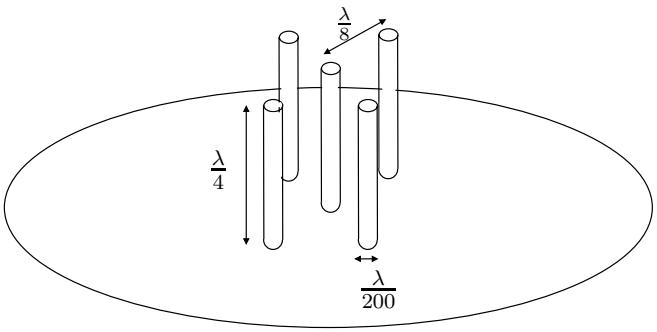


Figure 11.1: A five element monopole SPA. The center element is active and connected to the transceiver. The four passive antenna elements can be switched in or out of resonance using appropriately biased pin diodes.

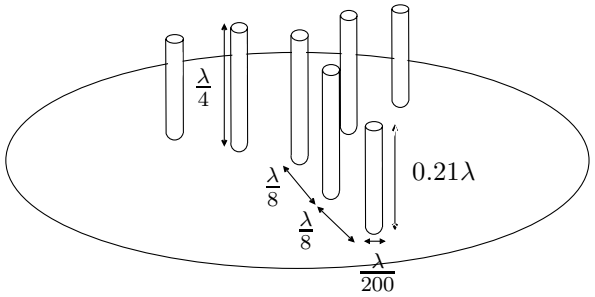


Figure 11.2: A seven element monopole SPA. The center element is active and connected to the transceiver. The three passive antenna elements closest to the active can be switched in or out of resonance using appropriately biased pin diodes. The three outermost monopoles are hardwired to ground.

characteristics, such as directivity and/or dual band tuning, as demonstrated in [STLO00], where a genetic algorithm approach was taken to optimize a six element switched beam antenna. If the parasitics are moved closer to the active element, the mutual coupling increases and the change in the radiation pattern when switching is greater. However, the antenna impedance changes also more dramatically, which makes the antenna matching difficult. Hence, a too large mutual coupling renders an inefficient antenna. The trade-off is thereby between compactness and high directivity on one hand, and antenna efficiency on the other.

The antennas in Figures 11.1 and 11.2 were simulated using High Frequency Structure Simulator (HFSS) from Agilent Technologies Inc., which is a 3D simulator using the finite element method to solve for the electromag-

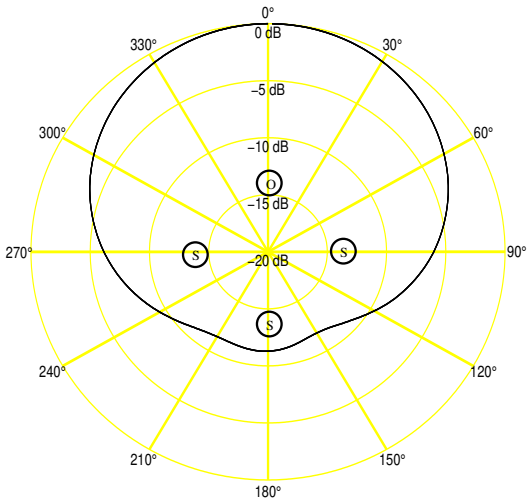


Figure 11.3: Power radiation pattern of the five element monopole antenna shown in Figure 11.1 with three parasitics shorted (S) to ground and one open (O).

netic field. The software was used to calculate the far-field radiation pattern of the antenna for different settings of the switched parasitics. The monopole elements were cylindrical with an length to radius ratio $l/r = 100$ which have a first resonance at approximately 0.24λ [Bal82]. The monopoles were assumed to be perfect conductors, which forces the electrical field to be normal to the surface. Furthermore, to shorten the simulation time, the ground plane was assumed to be of infinite extent. The chosen frequency was 2.15 GHz, suitable for the downlink in Universal Mobile Telephone System (UMTS) Frequency Division Duplex (FDD) mode. Hence, the wavelength is λ=0.1395 meter.

The far-field power radiation pattern for three shorted parasitics and one open for parasitic antenna 11.1 is shown in Figure 11.3 and the corresponding plot for Figure 11.2 is shown in Figure 11.4 for two shorted and one open parasitic. The directivity of the two antennas are 9.9 dB and 10.0 dB respectively.

11.3 MIMO Channel Capacity and Diversity Gain

To achieve a high capacity in MIMO systems or a large diversity gain, the signals received by different settings of the parasitics, called the M modes,

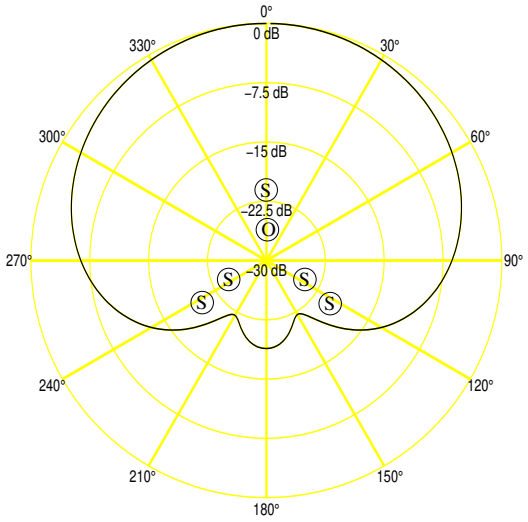


Figure 11.4: Power radiation pattern of the seven element monopole antenna shown in Figure 11.2 with two parasitics shorted (S) to ground and one open (O) plus three hardwired to ground (S).

must have low correlation (ideally zero). Hence, the correlation coefficient of the signal voltages received by two patterns are defined as [VA87]

$$\rho_{12} = \int_0^{2\pi} \int_0^{\pi} S(\theta,\phi) E_1(\theta,\phi) \cdot E_2^*(\theta,\phi) \sin\theta d\theta d\phi \tag{11.1}$$

where the two far field patterns $E_1(\theta,\phi)$ and $E_2(\theta,\phi)$ are normalized as

$$\int_0^{2\pi} \int_0^{\pi} S(\theta,\phi) |E_i(\theta,\phi)|^2 \sin\theta d\theta d\phi = 1 \tag{11.2}$$

for $i = 1, 2$. Above, $S(\theta,\phi)$ is the pdf of the incident waves.

The correlation coefficients for the complex voltage patterns corresponding to the power radiation patterns in Figure 11.3 and 11.4 are calculated assuming the Clark scenario [Cla68] with pdf $S(\theta,\phi) = \delta(\theta - \pi/2)/(2\pi \sin\theta)$ to model a ring of dense sources on a horizon about the receiving antenna. Here, it is assumed that three (two) parasitics are always shorted, to get four (three) different directions with 90° (120°) separation. This gives the correlation coefficient for adjacent patterns, or modes, for the parasitic antenna in Figure 11.3 as $|\rho_{i,i+1}|^2 = 0.1157$, and for opposite patterns $|\rho_{i,i+2}|^2 = 0.0120$. For the parasitic antenna in Figure 11.4, a value of $|\rho_{i,i+1}|^2 = 0.1002$ is obtained. The envelope correlation coefficients for the signals received from the modes are taken as $\rho_{eij} \approx |\rho_{ij}|^2$ [PS60]. A well known rule of thumb is

that "uncorrelated" signals in diversity branches corresponds to an envelope correlation lower than 0.7 [VA87]. Hence, the SPAs presented here would achieve a diversity gain in Clarke's scenario.

11.3.1 Channel capacity

The MIMO channel capacity is calculated using the SPA at the Base Station (BS) and/or at the User Equipment (UE). Here, a single user system is considered that employs point to point communication over a flat-fading channel with N transmit modes (or array antenna elements) and M receive modes (or array antenna elements) and no co-channel interference. The channel output corresponding to an input block spanning T symbol times is

$$\mathbf{Y} = \mathbf{HX} + \mathbf{V} \tag{11.3}$$

where the received signal $\mathbf{Y}$ is $M \times T$, the fading channel $\mathbf{H}$ is $M \times N$, the codeword matrix $\mathbf{X}$ is $N \times T$ and the receiver noise $\mathbf{V}$ is a $M \times T$ matrix. The entries of the noise matrix are *i.i.d.* Complex Gaussian with zero mean. It is also assumed that the channel is quasi-static, i.e. constant over the block of length T symbols.

With n parasitic elements, there are 2^n different modes, or settings of the switchable diodes. The transmit and receive modes of the parasitic antennas are chosen among these to minimize the envelope correlation. Here, it is assumed that the receiver switches through and samples the chosen modes during one symbol interval. This should be possible for many systems, since the switching time of a pin diode is only of the order of a few nanoseconds. The technique of oversampling the received signal is common in many communication systems, but here the oversampling is performed in both time and space, i.e. spatio-temporal oversampling. If the increased sampling rate (or bandwidth) poses a problem, a bandpass sampling strategy could also be employed. In this chapter, the potential in using the different radiation patterns of an SPA for diversity and capacity improvements are examined. However, further work is needed on the practical aspects of the antenna design as well as sampling strategies.

As the underlying concept here is angle diversity, a simple, yet detailed channel model that includes the spatial dimension is used. Since no polarization or time dependency is studied, a simpler version of the full model of Chapter 9, is used. A model similar to that used in the previous chapter will be used, where a circular disc of uniformly distributed scatterers is placed around the UE. Since the model is still based on fundamental physics, the correlations between the different antenna arrangements, which are of great

importance in this type of study, are physically motivated. Different channel scenarios can easily be obtained by varying the radius of the scattering disc as well as the location of the BS relative the disc, see also the previous chapter and [Sva00c].

The channel capacity of the following cases are investigated; BS antenna array-UE antenna array, BS antenna array-UE SPA, BS SPA-UE SPA. When the array is used, omnidirectional antenna elements are assumed with $d_{BS} = 2\lambda$ spacing at the BS, and $d_{UE} = 0.5\lambda$ spacing at the UE. Furthermore, it is assumed that the channel response is flat over frequency, an approximation valid if the communication bandwidth W is much less than the channel coherence bandwidth.

Shannon's capacity formula for an N input, M output MIMO channel, assuming equal power radiated from each transmitting antenna and $\mathbf{H}$ unknown at the transmitter can be written as

$$C = \log_2 \det \left[\mathbf{I}_M + \frac{\xi}{N} \mathbf{H}\mathbf{H}^H \right] \tag{11.4}$$

where ξ/N is the signal to noise ratio (SNR) at each receive antenna. The matrix elements $\mathbf{H}_{i,j}$ represents the complex path gain from transmitter j to receiver i.

Using the channel model [PRR96, Sva00c], the path gain can be written as

$$\mathbf{H}_{i,j} = \sum_{l=1}^{L} \alpha_l \exp\left\{ jk(r_{t_j \to s_l} + r_{s_l \to r_i}) \right\} g_{t_j}(\phi_l^T) g_{r_i}(\phi_l^R), \tag{11.5}$$

where the number of scatterers are L and α_l is the complex Gaussian distributed reflection coefficient with zero mean and unit variance. Furthermore, $r_{t_j \to s_l}$ and $r_{s_l \to r_i}$, denotes the distance from UE antenna j to scatterer l and scatterer l to BS antenna i respectively. Note that the wave number is denoted $k = 2\pi/\lambda$. Finally $g_{t_j}(\phi_l^R)$ and $g_{r_i}(\phi_l^R)$ are the complex voltage radiation patterns of the antennas. The channel matrix in (11.4) is normalized in the sense that $\|\mathbf{H}\|_F = \sqrt{M}$, where F denotes the Frobenius norm.

As the channel matrix $\mathbf{H}$ is a function of the random position and reflection coefficients of the scatterers, the capacity C in (11.4) is a random variable. With the SNR set to 4 dB, a Monte Carlo simulation with 4000 trials was performed to calculate the Complementary Cumulative Distribution Function (CCDF) for a scenario with $L = 20$ scatterers. The calculations are performed for an equal number of modes at transmit and receive (N=M=4). The distance between the BS and UE is $D = 50\lambda$. The CCDF of the capacity is shown in Figure 11.5 for the scattering disc radius $R = 50\lambda$. The curves are compared to the channel matrix $\mathbf{H}$ with *i.i.d.* elements with a

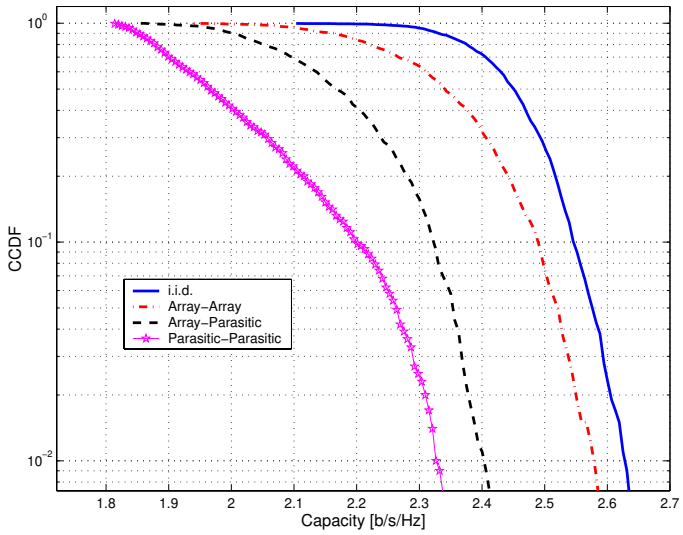


Figure 11.5: The complementary cumulative distribution function of the MIMO channel capacity for the N=M=4 case. The SNR is 4 dB and the scattering disc radius is 50λ. The parasitic antenna is shown in Figure 11.1.

complex Gaussian distribution. Using the array antenna at the UE results in a slightly higher capacity than the array-parasitic configuration, however at the expense of more hardware due to the use of four transceivers instead of one. If the SPA is used also at the BS, the capacity is further decreased, as the signals from the modes becomes correlated due to the small angular spread as seen at the BS.

The capacity at 10% outage is presented in Figures 11.6 and 11.7 for the two types of antenna configurations respectively. A large disc correspond to an indoor scenario, where both BS and UE are surrounded by scatterers. The other extreme, with a small scattering disc centered at the UE, as in an outdoor to indoor channel, results in a smaller capacity, due to the reduced angular spread, and hence, lower angle diversity gain. This can especially be seen in the case where the BS and the UE are both equipped with parasitic antennas. The capacity increases when the radius of the scattering disc exceeds the BS to UE distance, where full angular diversity also is possible at the BS.

At high bit rates, it might not be possible to switch through several modes during a symbol interval. Therefore, a system with only two modes was investigated. The parasitic antenna in Figure 11.1 was used, but only two modes with lowest signal correlation (opposite in direction) were exploited. The capacity at 10% outage is shown in Figure 11.8. Comparing Figure 11.6

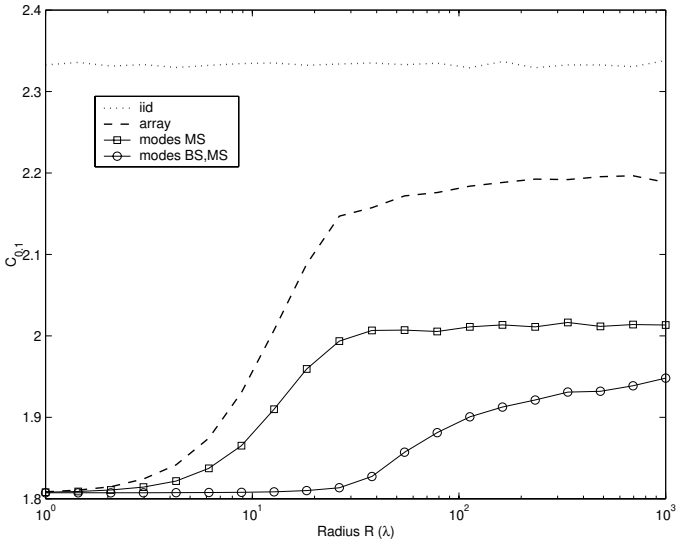


Figure 11.6: The channel capacity at 10% outage versus the radius of the scattering disc for the N=M=4 case. The parasitic antenna is shown in Figure 11.1. SNR=4 dB.

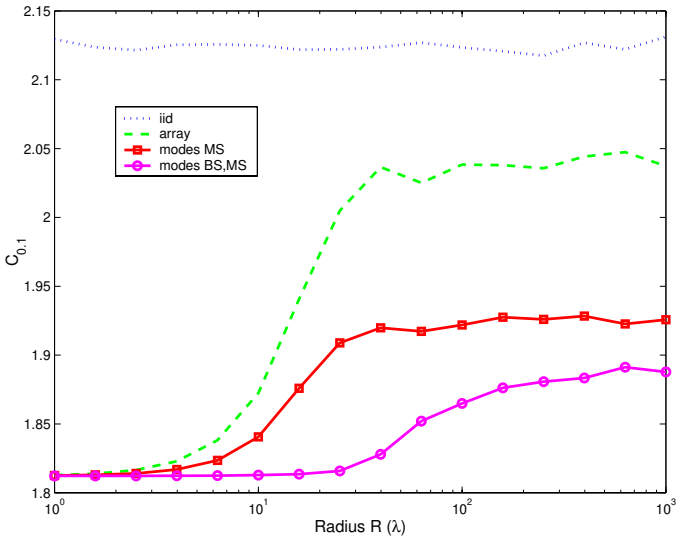


Figure 11.7: The channel capacity at 10% outage versus the radius of the scattering disc for the N=M=3 case. The parasitic antenna is shown in Figure 11.2.

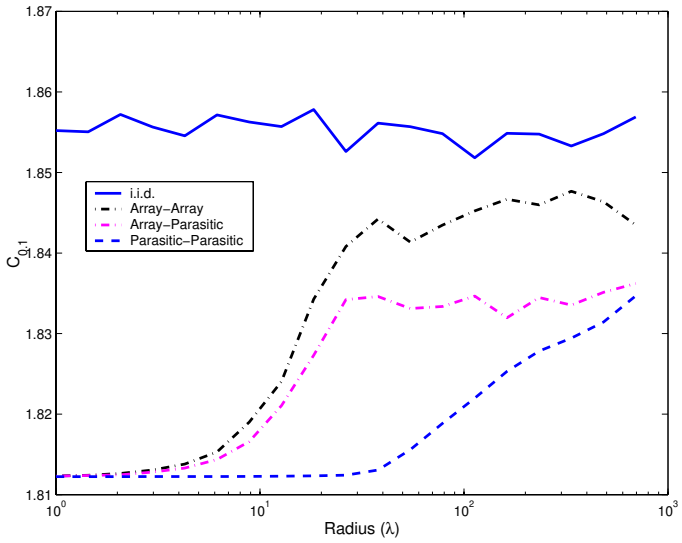


Figure 11.8: The channel capacity at 10% outage versus the radius of the scattering disc. Only N=M=2 of the 4 modes in the parasitic antenna in Figure 11.2 are used. SNR=4 dB.

and 11.8, it is observed that the overall capacity is lower, but the difference between the array and the SPA is slightly smaller at scattering disc radius around 10λ.

11.4 Evaluating BER Using STBC

In this section, Space Time Block Coding (STBC) [TJC99] will be used to exploit the available channel capacity discussed in the previous section. Here, it is assumed that the transmitter has no knowledge of the channel state information, and uses a very simple maximum likelihood detector based on linear processing at the receiver. The coding consists of mapping P consecutive symbols $x_1, \ldots x_P$ onto the transmission matrix $\mathbf{X}$. The entries of $\mathbf{X}$ are linear combinations of $x_1, \ldots x_P$ and their conjugates. For example, with $N = 2$ transmit antennas the Alamouti's scheme [Ala98] can be used

$$\mathbf{X} = \begin{pmatrix} x_1 & -x_2^\star \\ x_2 & x_1^\star \end{pmatrix} \quad . \tag{11.6}$$

Since P time slots are used to transmit P symbols, the rate of the code is one. For complex symbol constellations, rate one STBC only exists for

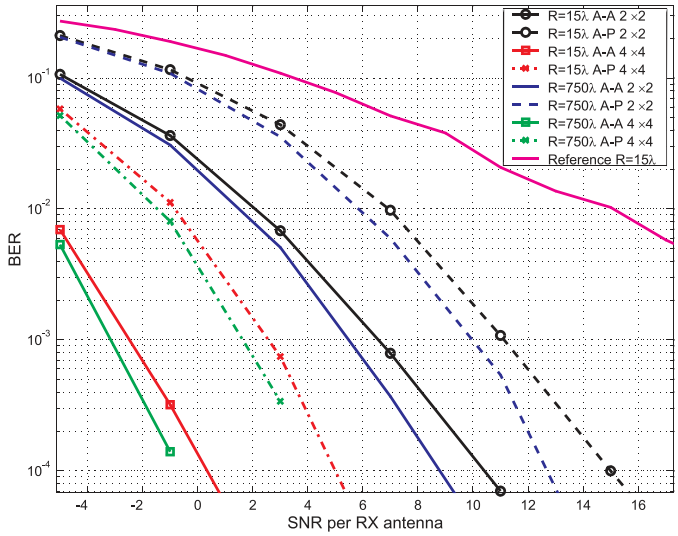


Figure 11.9: BER using STBC with N=M=2 and N=M=4 modes. The scattering disc radius is 750λ and 15λ. The BER curves a compared to the $N = M = 1$ reference antenna.

$N = 2$ [TJC99], however for real constellations there exists orthogonal, delay optimal rate one codes for $N = 2, 4, 8$, as for example

$$\mathbf{X} = \begin{pmatrix} x_1 & -x_2 & -x_3 & -x_4 \\ x_2 & x_1 & x_4 & -x_3 \\ x_3 & -x_4 & x_1 & x_2 \\ x_4 & x_3 & -x_2 & x_1 \end{pmatrix} . \tag{11.7}$$

To study the performance in terms of BER using the SPA, the systems are simulated employing the STBCs (11.6) and (11.7), using the 4-direction SPA in Figure 11.3 and comparing the results with the array antenna solution. When the Alamouti's STBC scheme is used, two opposite (180° separation direction) modes are used as the two receiving/transmitting modes for the SPA. The BER for these configurations is compared in Figure 11.9. Here, BPSK modulation is used in both cases, hence the bit rate 1 bit/s/Hz. It is observed that at the bit error rate of 10^{-2}, the 2-mode SPA gives 8 dB gain over the uncoded system with one antenna, and the 4-mode gives about 16 dB gain in SNR. The large and small scattering radius gives a difference in SNR gain less than 1 dB. The array-array configuration is about 5 dB better than the array-SPA configuration.

11.5 Conclusions

A MIMO system using a switched parasitic antenna has been analyzed, in terms of capacity and BER, assuming orthogonal space time block codes. The channel capacity was simulated using a spatial channel model. It was found that the SPA offered capacities close to those provided by an array antenna, in realistic MIMO scattering environments. Since the SPA only requires one receiver, it could be an attractive low cost solution to future user terminals using space time coding to increase data rates. The examples of SPA:s presented here are only for demonstrative purposes, the design of the SPA depends on the dimensions of the user equipment.

Chapter 12

Mutual Coupling Effects on the MIMO Capacity

A study of the capacity of multiple element antenna systems is presented, with particular emphasis on the effect that mutual coupling between the antenna elements has on the capacity. The results presented here shows, contrary to some earlier claims, that correlation between different channel coefficients as a function of antenna spacing, can in fact decrease when the mutual coupling effect is accounted for. As a consequence, capacity also improves. A realistic channel model is used to perform simulations to support these claims.

12.1 Introduction

The topic of Multi-Input Multi-Output (MIMO) communications systems have received considerable attention in recent years, and in particular the study of the capacity of such systems. Several authors have shown that the capacity gains resulting from the use of MultiElement Antenna (MEA) systems are potentially very significant [FG98, SFGK00]. A common assumption in the study of such systems is that the fading coefficients between different pairs of transmit-receive antennas are independent and identically distributed (i.i.d.). However, in practice the signals received by different antennas will be correlated which will reduce the capacity.

Measurement campaigns have been reported [KSMP00], that show a substantial capacity increase when using MEA systems, as long as elements are placed sufficiently far apart. Since the capacity of MEA systems is strongly dependent upon the number of transmit and receive elements available, it is

highly desirable to use as many antennas as possible. On the other hand, one typically has a limited amount of space/volume to distribute the antenna elements over. Unfortunately, closely spaced elements increase the correlation and thus decrease the capacity.

The effects on the correlation and thus also the capacity gain when using a MIMO system with a small inter-element antenna spacing will be investigated in this paper. Previously published studies in this area [Bur00, SFGK00] have ignored the fact that small inter-element antenna spacing will cause mutual coupling between elements. Mutual coupling is well known in the antenna community, but rather unknown in signal processing circles. In principle, the received voltage on each element will depend not only on the incident field, but also on the voltages on the other elements. This effect becomes significant at inter-element spacings of less than half a wavelength, and thus needs to be included in a correlation/capacity study when closely spaced elements are employed.

The general belief is that mutual coupling will deteriorate the channel, increase the correlation and reduce the achievable capacity. For instance, it was stated in [FG98] that since mutual coupling increases with reduced antenna spacing, it will also cause problems for achieving high capacity. In this paper, it is found that mutual coupling, actually can increase the channel capacity for scenarios with closely spaced antennas. It will be shown that for a typical realistic scattering model, mutual coupling can in fact have a decorrelating effect on the channel coefficients, and thereby also improve the capacity.

12.2 System Model

Consider a communication link with n_T transmit antennas and n_R receive antennas. Some important assumptions used throughout are:

- There is only a single user transmitting at any given time, so the received signal is corrupted by additive white Gaussian noise only.

- The communication is carried out in frames/packets of finite time-span, and the coherence-time of the channel is longer than the packet duration.

- The bandwidth of the transmitted signal is less than the coherence-bandwidth of the channel, i.e. the fading is frequency-flat.

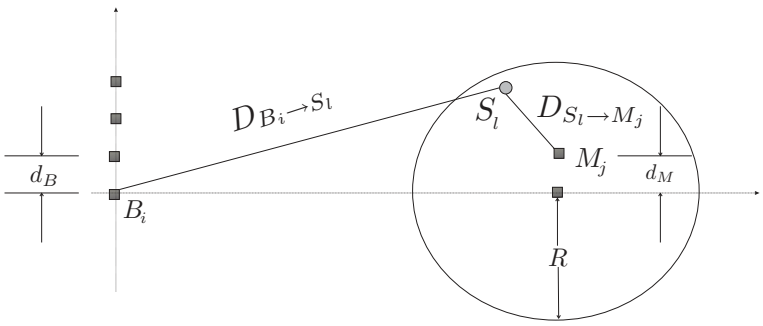


Figure 12.1: Geometry of channel

The following discrete-time vector/matrix model for the relation between input signal $\mathbf{s}_t$ and received signal $\mathbf{r}_t$ can then be formulated as

$$\mathbf{r}_t = \mathbf{H}\mathbf{s}_t + \mathbf{v}_t, \tag{12.1}$$

where $\mathbf{r}_t = [r_t^{(1)}, \ldots, r_t^{(n_R)}]^T$, $\mathbf{s}_t = [s_t^{(1)}, \ldots, s_t^{(n_T)}]^T$, $\mathbf{v}_t$ is spatially and temporally additive white Gaussian noise (AWGN) with unit variance, and t is a discrete-time index. The $n_R \times n_T$ channel matrix $\mathbf{H}$ is made up of elements $h_{i,j}$ as follows

$$\mathbf{H} = \begin{pmatrix} h_{1,1} & \ldots & h_{1,n_T} \\ \vdots & \ddots & \vdots \\ h_{n_R,1} & \ldots & h_{n_R,n_T} \end{pmatrix}, \tag{12.2}$$

where $h_{i,j}$ denotes the channel coefficient between the j:th transmit antenna and the i:th receiver element. In principle, any channel model that accurately includes the spatial dimension could be used to investigate the correlation properties of two spatially separated antennas. For an excellent overview, see [Ert98].

Here, the simple, yet detailed, channel model that has been used in several of the previous chapters will be employed. As previously, a circular disc of uniformly distributed scatterers is placed around the mobile. In Figure 12.1, a simple illustration of the scatter disc and the orientation of the mobile and base station is shown. Based on this model, the elements of the channel matrix in (12.2) are generated as follows. Assume there are L scatterers S_l, $l = 1, \ldots, L$, uniformly distributed on a disc of radius R, centered around the mobile. The channel parameter $h_{i,j}$ connecting transmit element j and receive element i is thus

$$h_{i,j} = \sum_{l=1}^{L} \alpha_l \exp\left(-j\frac{2\pi}{\lambda} \cdot (D_{B_i \to S_l} + D_{S_l \to M_j})\right) \tag{12.3}$$

where $D_{B_i \to S_l}$ and $D_{S_l \to M_j}$ are the distances from base station antenna i to scatterer l, and scatterer l to mobile antenna j, respectively, as shown in Figure 12.1. Also, α_l is the scattering coefficient from scatterer l, and is modeled as a normal complex random variable, with zero mean and unit variance. The channel matrix is finally normalized such that $\|\mathbf{H}\|_F^2 = n_T$. Thus, the increased antenna gain due to the use of multiple antennas is not included [Bur00].

12.3 Mutual Coupling

The principal function of an antenna is to convert an electromagnetic field into an induced voltage or current to be measured. However, the measured voltage at each antenna element will depend not only on the incident field, but also on the voltages on the other elements. Essentially, the received voltage on each element will induce a current on the element which in turn radiates a field which affects the surrounding element, i.e. the elements are said to be mutually coupled.

Mutual coupling is well known in the antenna community, since coupling between antenna elements is one of the most important properties to consider in antenna design. However, this phenomenon is is rarely accounted for or studied in the signal processing or communications literature. It is a simple matter to include the coupling effect in the model for the received voltage, by inserting a mutual coupling matrix, modifying (12.1) to

$$\mathbf{r}_t = \mathbf{CHs}_t + \mathbf{v}_t. \tag{12.4}$$

It is then natural to include the coupling effects into the channel by combining the two terms into a new channel matrix $\mathbf{H}'$. Note that (12.4) only includes coupling at the receiving antenna elements. In the scenario depicted in Figure 12.1, several closely spaced elements at the transmitter (mobile) will also experience mutual coupling. Thus, including this effect at both the transmitter and receiver, the expression for the channel becomes

$$\mathbf{H}' = \mathbf{C}_b\mathbf{HC}_m, \tag{12.5}$$

where the coupling matrix at the base $\mathbf{C}_b$ is $n_R \times n_R$ and the corresponding matrix at the mobile $\mathbf{C}_m$ is $n_T \times n_T$. Using fundamental electromagnetics and circuit theory, the coupling matrix of an array antenna can be written as [GK83]

$$\mathbf{C} = (Z_A + Z_T)(\mathbf{Z} + Z_T\mathbf{I})^{-1}, \tag{12.6}$$

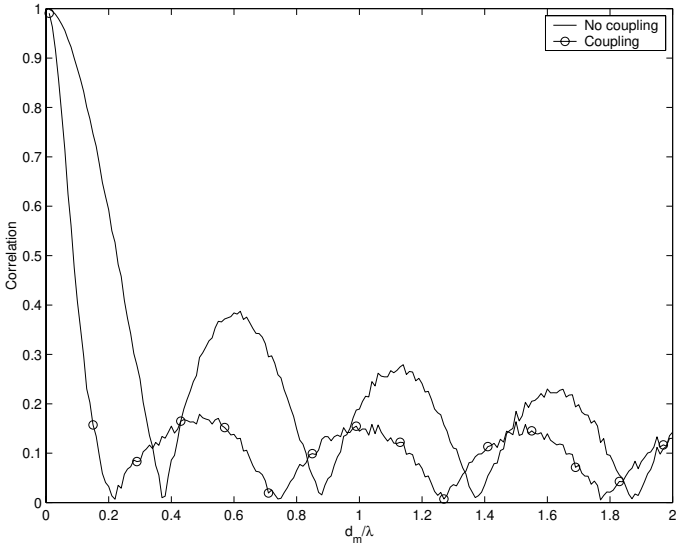


Figure 12.2: The effect of the coupling on the correlation between paths $E[h_{1,1}, h_{1,2}]$ versus different element spacings d_m at the mobile. The base station separation is $d_b = 0.5\lambda$, and the scatter disc radius $R = 200\lambda$.

where Z_A is the antenna impedance, Z_T is the impedance of the measurement equipment at each element, and $\mathbf{Z}$ is the mutual impedance matrix. This expression can be used for any array antenna, but for many types of elements an analytical expression for the mutual impedance matrix and the antenna impedance is difficult to obtain. However, one noticeable exception is the case of dipoles, which will be used as antenna elements in the communications scenarios investigated here. A more detailed derivation of the received voltages of an array of thin finite dipoles where coupling is included can be found in [Sva99b]. See also, Chapter 2 for further details. In the following, the impedance of the measurement equipment Z_T is chosen as the complex conjugate of the dipole impedance in order to reduce the power loss. The effect of the coupling on the correlation between paths $E[h_{1,1}, h_{1,2}]$ is shown in Figure 12.2 for different element spacings d_m at the mobile. Interestingly, the correlation between these two channel coefficients decreases faster when coupling is included in the calculations. In fact, it is possible to cut the element separation in half due to the coupling at the 0.1 correlation level (0.4λ to 0.2λ). Thus, the coupling phenomenon actually decorrelates the signals by acting as an additional "channel". Interestingly, it was recently reported in [LSM00] that coupling can in fact decrease the Bit-Error-Rate performance on a Nakagami fading channel. In summary, the mutual coupling may, contrary to common belief, actually decrease the correlation level

between channel coefficients and thereby also increase the channel capacity.

12.4 Channel Capacity

Considering a (n_T, n_R) MEA system, with channel matrix $\mathbf{H}$, the channel capacity is given by the expression

$$C = \log_2(|\mathbf{I}_{n_R} + \rho\frac{\mathbf{H}\mathbf{H}^*}{n}|) = \sum_{k=1}^{n} \log_2(1 + \rho\frac{\lambda_k}{n}), \tag{12.7}$$

where $\mathbf{H}$ is the $n_R \times n_T$ channel matrix, $n = \min(n_T, n_R)$, and λ_k are the eigenvalues of $\mathbf{H}\mathbf{H}^*$. Note that this expression assumes that the available transmit power ρ is uniformly allocated to the n_T transmit elements, which is the practical approach when the transmitter has no knowledge of the channel.

It is easily realized from this expression that a large capacity hinges on the presence of a rich scattering environment, being directly related to the rank of the channel matrix. Conversely, little or no scattering will result in a channel matrix of unit rank and thus low capacity.

In the event that there is coupling between the antenna elements of the mobile, the channel matrix should be modified to $\mathbf{H}_{cm} = \mathbf{H}\mathbf{C}$, where $\mathbf{C}$ is given in (12.6). The capacity in this case thus becomes

$$C = \log_2(|\mathbf{I}_{n_R} + \rho\frac{\mathbf{H}_{cm}\mathbf{H}_{cm}^*}{n}|). \tag{12.8}$$

Likewise, modifications for the case of coupling at the base or both at base-and-mobile, are similarly straightforward, as discussed in Section 12.3. In what follows, the capacity is computed for a large number of channel realizations, each with a random location of scattering elements within the disc. Hence, the capacity becomes a random variable, and its Complementary Cumulative Distribution Function (CCDF) as well as the outage capacity, will be computed for different parameter settings.

A scenario where the mobile has $n_T = 2$ antenna elements will be considered, where the elements are placed at broadside relative to the $n_R = 4$ element base-array. Furthermore, the distance between the base and mobile is $D = 300\lambda$. For each channel realization, a total of 100 scatterers are placed randomly and uniformly distributed on a disc of radius $R = 200\lambda$, centered on the mobile. The mobile has a total power $\rho = 10$, and 1000 channel realizations were generated for each value of d_m to compute the required statistics.

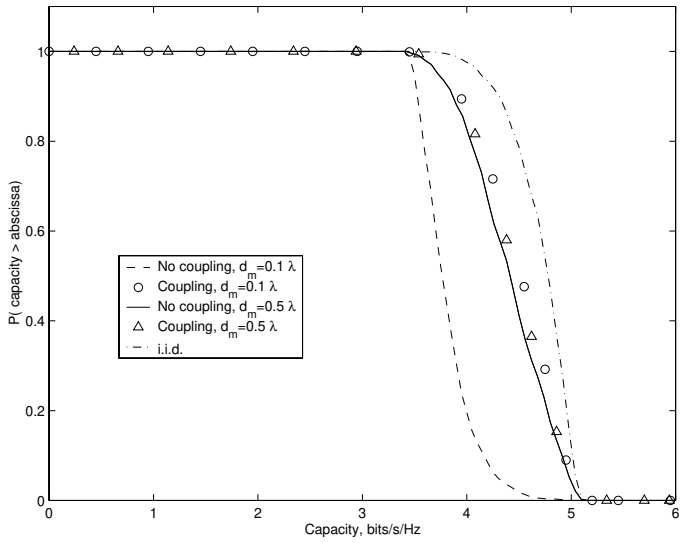


Figure 12.3: Capacity with and without coupling at mobile, for different values of mobile antenna separation, d_m.

Figure 12.3 shows the CCDF:s, both with and without mutual coupling between the mobile antenna elements, for two different values of mobile antenna separation, d_m. The separation between base-elements was held constant at $d_b = 0.5\lambda$. It also shows the capacity for the case of idealized i.i.d. channel coefficients [FG98]. If one ignores the effect of mutual coupling, it can be seen that one suffers a significant capacity loss by placing the mobile antennas closer together, i.e. from $d_m = 0.5\lambda$ to $d_m = 0.1\lambda$, as would be expected. The more interesting result is that when the coupling effect is accounted for, the difference in capacity is essentially identical for these two values of mobile antenna separation. In short, the advantage that is lost from the decreased antenna separation is more than compensated for when also the coupling between the same elements is taken into account. It should be mentioned that these observations hold true in the event that also the coupling between the base antenna elements, is included. Figure 12.4 shows the 10 % outage capacity $C_{0.1}$, i.e. there is a probability of 0.10 that the capacity is less than what is seen in the plot. First, notice how the capacity when coupling is included is higher for small separation distances, as compared to the no coupling case. For large separation distances, the strength of mutual coupling diminishes, and the capacities coincide. These results should be seen in conjunction with the correlation plot in Figure 12.2; The largest disparity between the two outage capacity curves occur at values of d_m for which the corresponding correlation curves also have a large disparity,

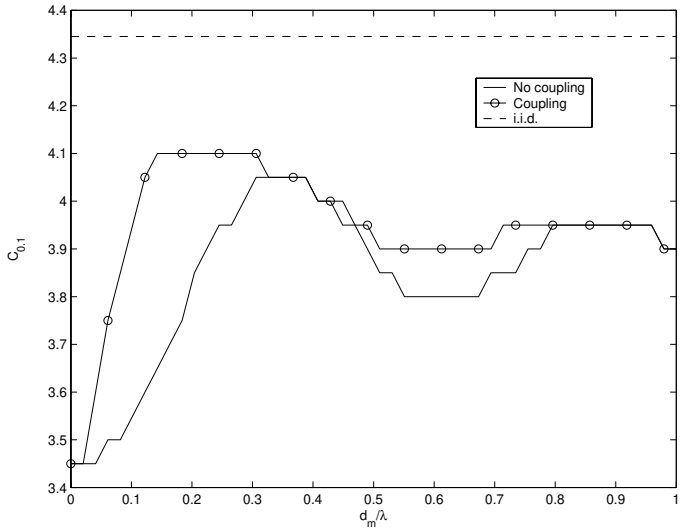


Figure 12.4: 10 % Outage capacity as a function of antenna spacing

in favor of the system which includes the coupling effect.

Other scenarios were also simulated to support the generality of the above conclusions to different propagation conditions. In summary, in the case of very localized scattering, the elements of the channel matrix will have a high degree of correlation regardless of whether coupling is included or not. Hence, there is little or nothing to be gained from an MEA system in terms of capacity improvements in such scenarios. For the other extreme, when the scattering environment is sufficiently rich to approach the i.i.d. assumption, coupling between antenna elements will clearly only degrade the performance. The results presented here have focused on more realistic scenarios that fall between these two extremes.

12.5 Conclusions

The capacity of multiple element antenna systems was studied, focusing on the effect of antenna spacing at the mobile terminal. In particular, the effect of mutual coupling between antenna elements was considered, and its effect on the correlation between channel coefficients and thereby on the capacity of such systems. Contrary to earlier claims regarding the effect of mutual coupling on capacity of MEA systems, results were presented that support the conclusion that coupling can in fact have a decorrelating effect on the channel coefficients and thereby also increase the capacity.

Chapter 13

Conclusions and Future Work

The main results from the different chapters together with some general conclusions are presented in this chapter. Finally, some interesting future work is proposed in last section.

13.1 Conclusions

This thesis has studied several aspects of topics within the area of antennas and propagation from a signal processing perspective. However, theory and methods from electromagnetics and communications have also been used. This interdisciplinary character has been a major theme of the thesis, that has resulted in several novel studies of antennas as well as spatio-temporal signal processing channel models. It should also be mentioned that when trying to incorporate physical properties into signal processing schemes, new data models are often required. Hence, an important part of this thesis is physical modeling, where the theories of electromagnetics and wave propagation meet signal processing. On the other hand, the borderline between signal processing and communication theory has also been studied in this thesis. This is a very interesting research area with an intense activity in MIMO and STC communication systems, of which both were briefly considered in connection with different antenna solutions. This thesis has thus studied a few problems along these interesting borderlines, and in the following the main results and some general conclusions from these studies will be presented.

The first chapters considered direction finding in the presence of mutual coupling between the elements of an antenna array. The effects of mutual coupling were studied from a signal processing perspective by analyzing sig-

nal processing schemes for direction finding. Electromagnetic aspects were also included by deriving the necessary expressions for the received voltages at the antenna elements in the presence of mutual coupling. Using these expressions, the direction finding performance that is possible to achieve with coupling was analyzed, unlike previously published work that analyzed specific estimation methods. A lower bound for the variance of any unbiased estimator was calculated. It was found that if the coupling is known, the effects on direction finding are small. In fact, in some cases a small increase in performance results due to the coupling.

If the coupling is unknown, there will be a loss of performance. The effects of an unknown mutual coupling can, however, be mitigated by estimating the unknown coupling along with the DOAs. A reduced coupling model suitable for estimation was derived, and its uniqueness properties were analyzed.

A number of DOA estimation schemes were extended to the case of both known and unknown coupling. In particular, the Noise Subspace Fitting method was shown to be a computationally attractive solution, while providing high estimation performance. Furthermore, the potential gain of estimating the unknown mutual coupling was examined by generating data, calculated using the EM derivation of the coupling, and employing the proposed algorithms. It was found that the effects of an unknown mutual coupling were substantially reduced by estimating the coupling along with the DOAs.

The direction finding part of the thesis ended with an analysis of two novel antenna arrangements for high resolution direction finding, namely the multimode and the switched parasitic antenna. Here, it was found that it is possible to obtain high resolution without employing an array of physically separated elements. Instead, the higher order modes were used to provide different copies of the received signal, that can be used for direction finding. Alternatively, a set of different radiation patterns may be obtained by employing a switched parasitic antenna. By calculating the CRB for these types of antennas, it was shown that high-resolution direction finding is possible. Some DOA estimation methods were also extended to these antennas, indicating good performance. However, there are many other interesting possibilities in using these types of antennas, of which some will be outlined in Section 13.2.

A large part of the thesis was also devoted to channel modeling and diversity studies. In particular, a spatio-temporal channel model based on EM scattering and fundamental physics was presented. By studying the scattering properties of objects of simple shapes, such as spheres and cylinders, a simple function that captures the most important scattering properties was derived. A compact formulation was obtained by using a dyad notation and

concepts from rough surface scattering. Hence, a simple, yet detailed, channel model was obtained by employing results from EM wave propagation theory and elementary antenna theory. These results were then used to formulate SISO, SIMO, and MIMO channel models. Essentially, the extension from a single transmit antenna and single receive antenna to multiple antennas only required that several impulse responses or channel coefficients need to be evaluated. For each type of channel model, important channel characteristics were simulated and discussed. The polarization properties of the channel as well as those of the antennas are included in the model, thus allowing for studies of different antenna arrangements. This is one of the main benefits of the model, and several different types of antennas were analyzed in the following chapters. Another important property of the model is that it is possible to relate various effects in the impulse responses to physical entities, since the model is derived from fundamental electromagnetics and wave propagation. Several performance measures, that can be found in the literature, were calculated in order to find suitable values for the channel parameters. However, measurements would of course be of great value in this task. In the next section, a few suggestions of how measurements may be used to find reasonable parameter values are given.

The potential use of the multimode and the switched parasitic antennas in mobile communications was also examined using the above channel model. Here, the use of these antennas as diversity receivers in MIMO systems was investigated. It was found that the performance was comparable to that of an antenna array of reasonable element separation. In particular, the multimode microstrip antenna offers interesting "paste on" features, while the switched parasitic provides a compact and cheap (only one RF-chain) solution. However, the studies presented here only considered some aspects of the use of these antennas. Some further possibilities and studies are suggested in the next section.

Finally, the mutual coupling issue was revisited from a slightly different point of view. In this study, the impact of coupling upon the channel capacity of a MIMO system consisting of closely spaced antenna elements, was examined. Surprisingly, it was found that mutual coupling actually can increase the capacity for certain channels. A plausible explanation for this effect is that the mutual coupling acts as an additional "channel" that may decrease the correlation by creating some additional multipath. In the light of the results from the DOA analysis in the presence of mutual coupling, it is not obvious that mutual coupling of necessity severely degrades systems and thereby should be avoided at all costs. However, further analysis is necessary and there will obviously be cases where coupling significantly does degrade the performance.

Although some results were obtained, there are still many open problems regarding the above issues that needs to be settled. It is the author's belief that most of the above topics essentially only are touched upon, and there are many interesting and important problems that need to be addressed. Some of these are indicated in the next section.

13.2 Future Work

In this thesis, the effects of mutual coupling on DOA estimation were analyzed via theoretical calculation of the mutual impedance between dipoles. It would of course be interesting to extend this analysis to other elements, such as patch elements. Furthermore, only the case of an equidistantly spaced linear array, was analyzed. Other array geometries would of course be interesting to examine. In particular, planar arrays such as circular arrays should be examined, since these offer 2-D DOA estimation. Since these also are known to experience problems with mutual coupling, it is an interesting topic for future research. Obviously, measurements would be interesting to examine, in order to determine the potential performance gain by estimating an unknown coupling together with the DOAs. Another topic for future work is to study other models of the coupling that better describes the coupling properties.

Two different multimode antennas, the biconical and microstrip antenna, were analyzed in the thesis. These antennas were studied in a DOA estimation context, and also as diversity receivers in MIMO systems. However, many interesting possibilities and problems in exploiting multiple modes still exist. In principle, most antennas can support higher order modes. Hence, it is an interesting antenna design topic to find and evaluate other structures with perhaps better characteristics than the microstrip and biconical antenna. More work on the antenna design is needed to find suitable positions of the feeds, appropriate dimensions, and a matching network that minimizes the losses in the antenna. It is also important to note that in this thesis, only the potential of exploiting the individual modes were examined. Many other possibilities exists. For instance, a combination of spatial, polarization, and modal diversity is an interesting alternative that may offer very good performance. Furthermore, only the possibility of using the individual modes as diversity branches was studied. It is also possible to use several modes to form narrow beams. Here, there may be interesting connections to high-resolution direction finding schemes for traditional mechanically rotating narrow beam antennas, that may be exploited. Thus, there are many interesting possibilities of using higher order modes of antennas, both for

direction finding and in future high capacity MIMO systems.

There are also many exciting directions for parasitic antennas that can be explored. For instance, the concept is not limited to monopoles. An even better solution may be a switched parasitic patch antenna that may be very well suited for deployment at the mobile. Furthermore, many other options of how these different radiation patterns can be used need to be explored. More work is also needed in designing the oversampling in space and time. This may prove especially critical for high data rates, where instead a selection scheme could be explored for MIMO systems and STC applications.

Another topic that may be studied is joint optimization of both the receiver, antenna, signal processing, and communication schemes. For example, simulating the full system of receiver network, antenna, channel, and receiver algorithms is useful in optimizing the overall performance of systems. With the rapid development of software in the electromagnetic and applied electronics area, this may not be unrealistic. Even if it should prove intractable to simulate the full system, it should be possible to analyze signal processing properties of different antenna arrangements via electromagnetic computation tools. This is something that most likely will benefit both the signal processing schemes and the antenna design process. An example of an interesting problem that could be analyzed is the relation between the communication quality and the radiation of mobile phones into human tissue. By an appropriate design of the antenna radiation pattern, the communication quality may increase while reducing the microwave exposure. In this type of study, models of both the antenna, head and channel may be used.

A large part of the thesis was spent on physical channel modeling. However, further work is needed in both validation and determining suitable parameter values. Here, measurements is an important tool. As previously mentioned, some measurements were performed shortly before writing this thesis. Some preliminary results of how measurements may be used to develop the channel model will therefore be given in this section. On the other hand, ray-tracing is also an interesting possibility of obtaining data that can be used to derive suitable parameter values. In this case, the surrounding environment is under complete control. Hence, it is possible to relate effects to physical entities in a more direct manner than using measurements that contains more uncertainty.

In the presented simulation results, the direct wave was neglected. However, for systems with LOS conditions, the direct wave will obviously affect the result. The normalized impulse response of an indoor measurement with LOS in an corridor, is shown in Figure 13.1. Also shown is an impulse response collected in NLOS conditions from an adjacent room. These measurements were collected at the office of Telia Research in Malmö, Sweden.

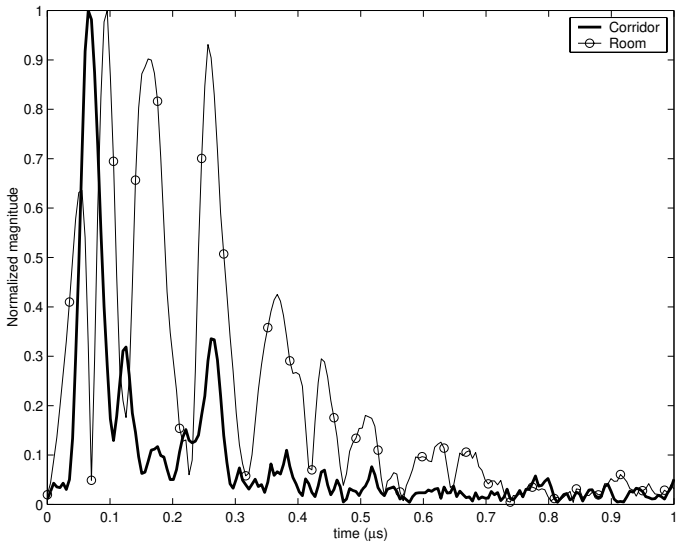


Figure 13.1: Wideband impulse responses from a typical office environment.

The measurements were conducted at a carrier frequency of 3.1GHz and a bandwidth of 200MHz. This equipment has also been used to collect synthetic MIMO measurements [Str01]. Since the aim of this section is only to indicate how measurements can be used to improve the model, the description of the system is brief. For details regarding the measurement equipment, see [Bör00].

Returning to Figure 13.1, it is obvious that the direct wave in the corridor needs to be included. Also, it is necessary to determine suitable values for the attenuation of the waves over distance to accurately model the decay of the impulse response. The geometrical distribution of the scatterers must also be determined to obtain an appropriate length of the impulse response in the obstructed scenario of the room. Another important implication of the location of the scatterers is the time evolution of the channel. In Figure 13.2, the impulse responses over a distance of four wavelengths is shown. These data were collected when moving along a corridor at the Telia office. Here, it is clear that the number of significant peaks (or filter taps) are changing as the location is changed. Hence, a suitable distribution for the scatterers needs to be determined. In particular, a good performance measure for fitting parameter distributions to the measured data is needed.

In most practical measurements, it has been observed that multipath components generally arrive in clusters. Considering the underlying physics, where the scatterers have a spatial extent and there are multiple small reflec-

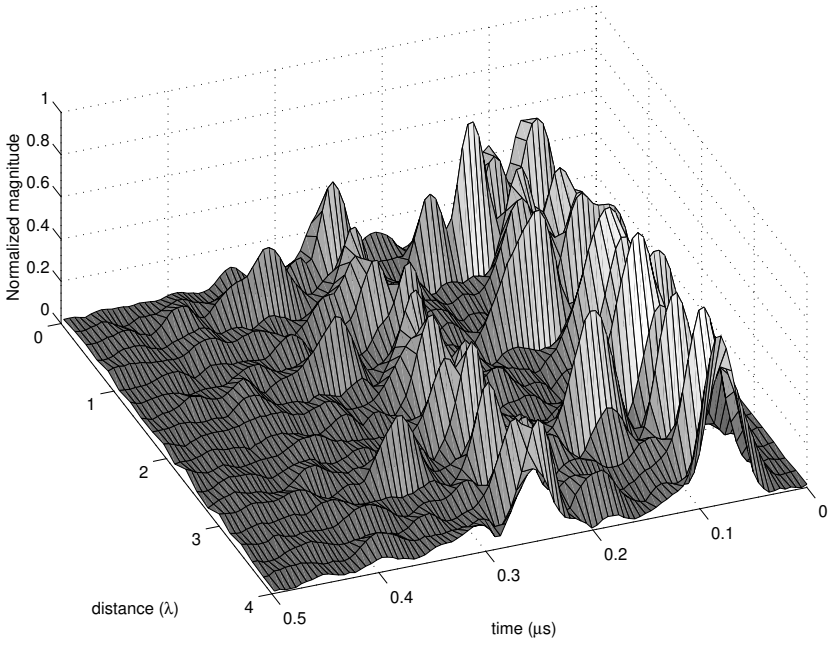


Figure 13.2: The evolution of the impulse response over distance (along a corridor).

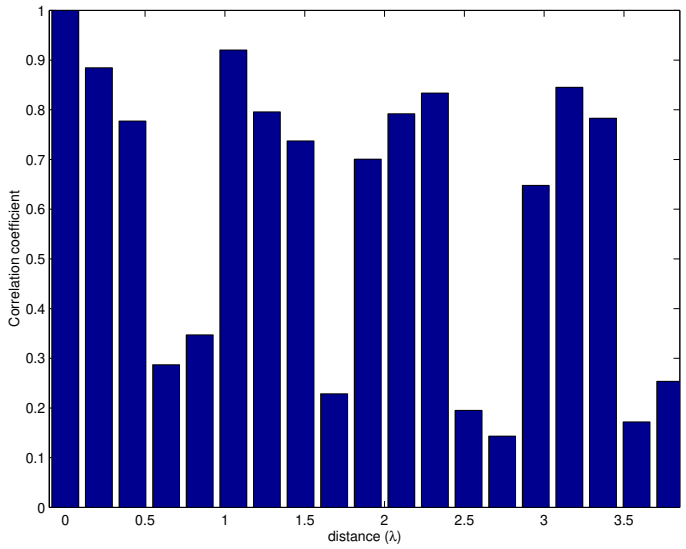


Figure 13.3: The narrowband correlation coefficient versus separation distance (along a corridor).

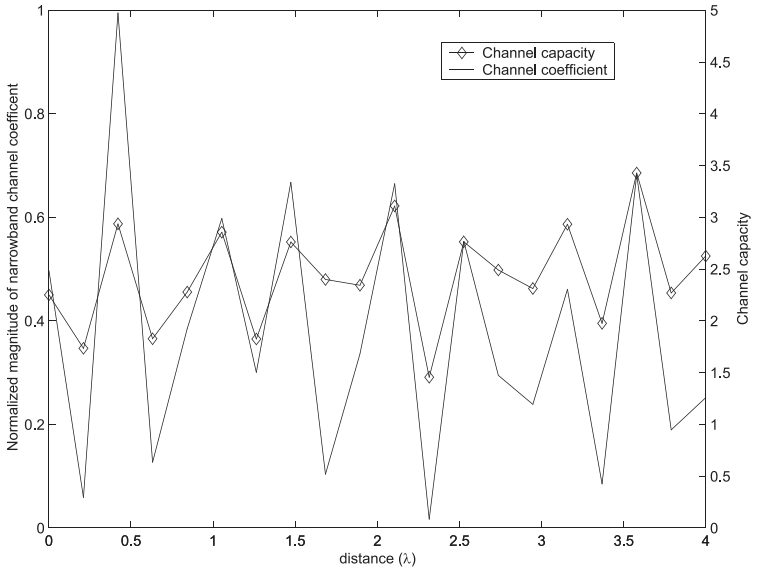


Figure 13.4: The magnitude of a single narrowband channel coefficient versus separation distance (along a corridor). The channel capacity of a 4×4 MIMO system for the same corridor is also shown.

tion points that contribute to the total signal, this seems reasonable. This clustering affects, for instance, the correlation between signals received by spatially separated antennas. Hence, these clusters should also be incorporated into the model. Here, some of the work on clusters of scatterers proposed in [BJ98, FMB98] could be useful. An example of the narrowband correlation between two signals received by spatially separated antennas is shown in Figure 13.3 versus separation distance. Here, the correlation decays with distance as it should. However, the decay is of oscillating character. By an appropriate positioning of the scatterers and clustering, this should be possible to model.

Furthermore, the rapid variation of the MIMO channel should be studied in more detail. In Figure 13.4, the variation of the magnitude of one channel coefficient of a 4×4 MIMO system is shown together with the total calculated channel capacity of the system. Here, it is found that although the individual coefficient experiences severe fading dips, the overall capacity appears much more stable. How this best is exploited in STC schemes and the relationship to different antenna solutions is also an interesting topic for future research. For instance, can the multiple polarization available through an electromagnetic vector sensor [NP97] be exploited here ?

Bibliography

[Adv97] R.S. Adve. "Accounting for the Effects of Mutual Coupling in Adaptive Antennas". In *Proc. 1997 IEEE National Radar Conference*, pages 361–366, Syracuse, New York, May 1997.

[AG00] D.A. Abraham and D.F. Gingras. "Matched Field Processing for CDMA Communication Systems in a Multipath Channel". In *Proc. IEEE SAM 2000 Workshop*, pages 17–21, Boston, MA, March 2000.

[AKT00] A.A. Almhdie, V. Kezys, and T.D. Todd. "Improved Capacity in TDMA/SDMA Using Switched Parasitic Antennas". In *Proc. of IEEE Personal, Indoor and Mobile Radio Communications*, pages 363–367, London,UK, 2000.

[Ala98] S.M. Alamouti. "A Simple Transmit Diversity Technique for Wireless Communications". *IEEE Journal on Selected Areas in Communications*, 16:1451–1458, 1998.

[AMd01] M.R. Andrews, P. P. Mitra, and R. deCarvalho. "Tripling the Capacity of Wireless Communications Using Electromagnetic Polarization". *Nature*, 409:316–318, January 2001.

[Ami92] N. Amitay. "Modeling and Computer Simulation of Wave Propagation in Lineal Line-of-Sight Microcells". *IEEE Trans. on Vehicular Technology*, 41(4):337–342, November 1992.

[AN00] G.E. Athanasiadou and A.R. Nix. "A Novel 3-D Indoor Ray-Tracing Propagation Model: The Path Generator and Evaluations of Narrow-Band and Wide-Band Predictions". *IEEE Trans. on Vehicular Technology*, 49(4):1152–1168, July 2000.

[And00] J.B. Andersen. "Array Gain and Capacity for Known Random Channels with Multiple Element Arrays at Both Ends". *IEEE Journal on Selected Areas in Communications*, 18(11):2172–2178, November 2000.

[App76] S.P. Applebaum. "Adaptive Arrays (reprinted from technical report, 1966)". *IEEE Trans. on AP*, AP-24:650–662, 1976.

[AR76] J.B. Andersen and H.H. Rasmussen. "Decoupling and Descattering Networks for Antennas". *IEEE Trans. on Antennas and Propagation*, AP-24:841–846, December 1976.

[ARY95] J.B. Andersen, T.S. Rappaport, and S. Yoshida. "Propagation Measurements and Models for Wireless Communications Channels". *IEEE Communications Magazine*, pages 42–49, January 1995.

[AS70] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. National Bureau of Standards:Applied Mathematics Series 55, Washington DC, 1970.

[AS96] R.S. Adve and T.K. Sarkar. "Elimination of the Effects of Mutual Coupling in an Adaptive Nulling System with a Look Direction Constraint". In *Proc. IEEE Antennas and Propagation Society International Symposium*, volume 2, pages 1164–1167, Baltimore, Maryland, July 1996.

[AS00] R.S. Adve and T.K. Sarkar. "Compensation for the Effects of Mutual Coupling on Direct Data Domain Adaptive Algorithms". *IEEE Trans. on Antennas and Propagation*, 48(1):86–94, January 2000.

[Ast96] D. Astély. "On Antenna Arrays in Mobile Communication Systems: Fast Fading and GSM Base Station Receiver Algorithms". Master's thesis, Royal Inst. Technology, Sweden, March 1996.

[Ast99] D. Astély. *"Spatial and Spatio-Temporal Processing with Antenna Arrays in Wireless Systems"*. PhD thesis, Royal Inst. Technology, Stockholm, Sweden, 1999.

[Aul79] T. Aulin. "A Modified Model for the Fading Signal at a Mobile Radio Channel". *IEEE Trans. on Vehicular Technology*, 28(3):182–203, August 1979.

[Bal82] C. Balanis. *Antenna Theory: Analysis and Design*. John Wiley and sons, Inc., New York, 1982.

[Bal89] C. Balanis. *Advanced Engineering Electromagnetics*. John Wiley and sons, Inc., New York, 1989.

[Ban71] W. J. Bangs. *Array Processing with Generalized Beamformers*. PhD thesis, Yale University, New Haven, CT, 1971.

[Bar48] M.S. Bartlett. "Smoothing Periodograms from Time Series with Continuous Spectra". *Nature*, 161:686–687, 1948.

[Bar83] A.J. Barabell. "Improving the Resolution Performance of Eigenstructure Based Direction-Finding Algorithms". In *Proc. IEEE ICASSP 83*, pages 336–339, Boston, MA, 1983.

[BD91] W.R. Braun and U. Dersch. "A Physical Mobile Radio Channel Model". *IEEE Trans. on Vehicular Technology*, 40(2):472–482, May 1991.

[BE96] C. Beckman and J. Eriksson. "Mutual Coupling Effects on Direction of Arrival Estimation". In *Proc. RadioVetenskap och Kommunikation (RVK 96)*, pages 410–414, Kiruna, Sweden, June 1996.

[Bel63] P.A. Bello. "Characterization of Randomly Time-Variant Linear Channels". *IEEE Trans. Commun. Systems*, CS-11:360–393, December 1963.

[Ber77] M.S. Berger. *Nonlinearity and Functional Analysis*. Academic Press, New York, 1977.

[Ber00] H.L. Bertoni. *Radio Propagation for Modern Wireless Applications*. Prentice-Hall, Upper Saddle River, New Jersey, 2000.

[BFHS00] D.W. Bliss, K.W. Forsythe, A.O. Hero, and A.L. Swindlehurst. "MIMO Environmental Capacity Sensitivity". In *Proc. 34th Asilomar Conf. Sig., Syst., Comput.*, pages 764–768, Pacific Grove, CA, October 2000.

[BJ98] J.J. Blanz and P. Jung. "A Flexibly Configurable Spatial Model for Mobile Radio Channels". *IEEE Trans. on Communications*, 46(3):367–371, March 1998.

[BKM96] J.J. Blanz, A. Klein, and W. Mohr. "Measurement-Based Parameter Adaptation of Wideband Spatial Mobile Radio Channel Models". In *Proc. IEEE 4th Int. Symp. on Spread Spectrum Tech. and Appl.*, pages 91–97, 1996.

[Bla00] N. Blaunstein. *Radio Propagation in Cellular Networks*. Artech House Publishers, 2000.

[BM86a] Y. Bresler and A. Macovski. "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise". *IEEE Trans. ASSP*, ASSP-34:1081–1089, October 1986.

[BM86b] Y. Bresler and A. Macovski. "On the Number of Signals Resolvable by a Uniform Linear Array". *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 34(6):1361–1375, December 1986.

[Böh86] J.F Böhme. "Estimation of Spectral Parameters of Correlated Signals in Wavefields". *Signal Processing*, 10:329–337, 1986.

[Bör00] H. Börjeson. *"Radio Wave Propagation in Confined Environments"*. PhD thesis, Department of Applied Electronics, Lund University, Lund, Sweden, 2000.

[BS63] P. Beckmann and A. Spizzichino. *The Scattering of Electromagnetic Waves from Rough Surfaces*. Pergamon Press, first edition, 1963.

[BSea87] J.J. Bowman, T.B.A. Senior, and P.L.E. Uslenghi et al. *Electromagnetic and Acoustic Scattering by Simple Shapes*. Hemisphere Publishing Corporation (SUMMA), 1987.

[Bur00] A.G. Burr. "Channel Capacity Evaluation of Multi-Element Antenna Systems Using a Spatial Channel Model". In *Proc. of AP2000*, Davos, Switzerland, April 2000.

[Cap69] J. Capon. "High Resolution Frequency Wave Number Spectrum Analysis". *Proc. IEEE*, 57:1408–1418, 1969.

[CJT98] C-N. Chuah, Kahn J, M, and D. Tse. "Capacity of Multi-Antenna Array Systems in Indoor Wireless Environment". In *Proc. IEEE Globecom 98*, volume 4, pages 1894–1899, 1998.

[Cla68] R.H. Clarke. "A Statistical Theory of Mobile Radio Reception". *Bell Syst. Tech. J.*, 47:957–1000, July 1968.

[Col85] R.E. Collin. *Antennas and Radio Wave Propagation*. McGraw-Hill, New York, 1985.

[CPdAG98] M.F. Catedra, J. Perez, F. Saez de Adana, and O. Gutierrez. "Efficient Ray-Tracing Techniques for Three-Dimensional Analyses of Propagation in Mobile Communications: Applications to Picocell and Microcell Scenarios". *IEEE Antennas and Propagation Magazine*, 40(2):15–28, April 1998.

[CRFLL00] D. Chizhik, F. Rashid-Farrokhi, J. Ling, and A. Lozano. "Effect of Antenna Separation on the Capacity of BLAST in Correlated Channels". *IEEE Communications Letters*, 4(11):337–339, November 2000.

[CS68] J.W. Crispin, Jr. and K.M. Siegel, editors. *Methods of Radar Cross-Section Analysis*. Academic Press, 1968.

[CTO89] H. Clergeot, S. Tressens, and A. Ouamri. "Performance of High Resolution Frequencies Estimation Methods Compared to the Cramer-Rao Bounds". *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 37(11):1703–1720, November 1989.

[Dem98] F. Demmerle. *"Bikonische Antenne mit mehrmodiger Anregung für den räumlichen Mehrfachzugriff (SDMA)"*. Phd Thesis (in german), Institut für Höchstfrequenztechnik und Elektronik der Universität Karlsruhe, Germany, 1998.

[Der79] A. Derneryd. "Analysis of the Microstrip Disc Element". *IEEE Trans. on Antennas and Propagation*, 27(5):660–664, 1979.

[dH57] H.C. Van de Hulst. *Light Scattering by Small Particles*. John Wiley and sons, Inc., 1957.

[DHHH00] A. Duel-Hallen, S. Hu, and H. Hallen. "Long-Range Prediction of Fading Signals". *IEEE Signal Processing Magazine*, 17(3):62–75, May 2000.

[DKW97] M. Döttling, F. Künchen, and W. Wiesbeck. "Deterministic Modeling of the Street Canyon Effect in Urban Micro and Pico Cells". In *Proc. IEEE ICC'97*, pages 36–40, Montreal, 1997.

[DLX00] K.R. Dandekar, H. Ling, and G. Xu. "Effect of Mutual Coupling on Direction Finding in Smart Antenna Applications". *Electronics Letters*, 36(22):1889–1891, October 2000.

[DR93] U. Dersch and R.J. Rüegg. "Simulations of the Time and Frequency Selective Outdoor Mobile Radio Channel". *IEEE Trans. on Vehicular Technology*, 42(3):338–344, August 1993.

[DS83] J.E. Dennis and R.B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice Hall, Englewood Cliffs, NJ, 1983.

[DW98] F. Demmerle and W. Wiesbeck. "A Biconical Multibeam Antenna for Space-Division Multiple Access". *IEEE Trans. on Antennas and Propagation*, 46(6):782–787, June 1998.

[EK99] T. Ekman and G. Kubin. "Nonlinear Prediction of Mobile Radio Channels: Measurements and MARS Model Designs". In *Proc. ICASSP 99*, Phoenix, AZ, March 1999.

[Ekm00] T. Ekman. "Prediction of Mobile Radio Channels". Licentiate thesis, Signals and Systems, Uppsala University, November 2000.

[Ell81] R. S. Elliot. *Antenna Theory and Design*. Prentice-Hall, Englewood Cliffs NJ, 1981.

[Ert98] R.B. Ertel et al. "Overview of Spatial Channel Models for Antenna Array Communication Systems". *IEEE Personal Communications Magazine*, 5(1):10–23, February 1998.

[Far92] A. Farina. *"Antenna-Based Signal Processing Techniques for Radar Systems"*. Artech House, Boston, 1992.

[FB00] B.P. Flanagan and K.L. Bell. "Improved Array Self Calibration with Large Sensor Position Errors for Closely Spaced Sources". In *Proc. of the 2000 IEEE Sensor Array and Multichannel Signal Processing Workshop*, pages 484–488, Boston, MA, 2000.

[FG98] G.J. Foschini and M.J. Gans. "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas". *Wireless Personal Communications*, 6:311–335, March 1998.

[FK92] A.J. Fenn and G.A. King. "Adaptive Nulling in the Hyperthermia Treatment of Cancer". *The Lincoln Laboratory Journal*, 5(2):223–240, 1992.

[FL96] B.H. Fleury and P.E. Leuthold. "Radiowave Propagation in Mobile Communications: An Overview of European Research". *IEEE Communications Magazine*, pages 70–81, February 1996.

[FM94] L.B. Felsen and N. Marcuvitz. *Radiation and Scattering of Waves*. Series on Electromagnetic Waves. IEEE Press, 1994. Reprint. Originally Published 1972.

[FMB98] J. Fuhl, A.F. Molisch, and E. Bonek. "Unified Channel Model for Mobile Radio Systems with Smart Antennas". *IEE Proc.-Radar, Sonar Navig.*, 145(1):32–41, February 1998.

[Foc65] V.A. Fock. *Electromagnetic Diffraction and Propagation Problems*. Pergamon Press, New York, 1965.

[FW88] B. Friedlander and A.J. Weiss. "Eigenstructure Methods for Direction Finding with Sensor Gain and Phase Uncertainties". In *Proc. IEEE ICASSP 88*, pages 2681–2684, New York City, April 1988.

[FW91] B. Friedlander and A.J. Weiss. "Direction Finding in the Presence of Mutual Coupling". *IEEE Trans. on Antennas and Propagation*, 39(3):273–284, March 1991.

[Gab80] W.F. Gabriel. "Spectral Analysis and Adaptive Array Superresolution Techniques". *Proceedings of the IEEE*, 68(6):654–666, June 1980.

[GBGP00] D. Gesbert, H. Bölcskei, D.A. Gore, and A.J. Paulraj. "Outdoor MIMO Wireless Channels: Models and Performance Prediction". In *Proc. 34th Asilomar Conf. Sig., Syst., Comput.*, Pacific Grove, CA, October 2000.

[GK83] I.J. Gupta and A.K. Ksienski. "Effect of Mutual Coupling on the Performance of Adaptive Arrays". *IEEE Trans. on Antennas and Propagation*, 31(5):785–791, September 1983.

[God97] L.C. Godara. "Applications of Antenna Arrays to Mobile Communications: Part II-Beamforming and Direction-of-Arrival Considerations". *Proceedings of the IEEE*, 85(8):1195–1245, August 1997.

[Has93] H. Hashemi. "The Indoor Radio Propagation Channel". *Proccedings of the IEEE*, 81(7):943–968, July 1993.

[Hat80] M. Hata. "Empirical Formula For Propagation Loss in Land Mobile Radio Services". *IEEE Trans. on Vehicular Technology*, 29(3):317–825, 1980.

[Hay95a] S. Haykin, editor. *Advances in Spectrum Analysis and Array Processing*, volume I-III. Prentice-Hall, 1991-1995.

[Hay95b] S. Haykin, editor. *Advances in Spectrum Analysis and Array Processing*, volume III. Prentice-Hall, 1995.

[HB91] Y-D Huang and M. Barkat. "Near-Field Multiple Source Localization by Passive Sensor Array". *IEEE Trans. on Antennas and Propagation*, 39(7):968–975, July 1991.

[He93] S. Haykin and J. Litva T.J. Shepherd (eds). *Radar Array Processing*. Springer-Verlag, Berlin, 1993.

[HHDH99] S. Hu, H. Hallen, and A. Duel-Hallen. "Physical Channel Modeling, Adaptive Prediction and Transmitter Diversity for Flat Fading Mobile Channels". In *Proc. IEEE Workshop Signal Processing Advances in Wireless Communications SPAWC'99*, pages 387–390, May 1999.

[How75] J.Q. Howell. "Microstrip Antennas". *IEEE Trans. on Antennas and Propagation*, 23:90–93, 1975.

[Hua84] J. Huang. "Circularly Polarized Conical Patterns from Circular Microstrip Antennas". *IEEE Trans. on Antennas and Propagation*, 32:991–994, 1984.

[HW90] B. Himed and D.D. Weiner. "Compensation for Mutual Coupling Effects in Direction Finding". In *Proc. IEEE ICASSP 90*, volume 5, pages 2631–2634, 1990.

[IN71] N. Inagaki and K. Nagai. "Exact Design of an Array of Dipole Antennas Giving the Prescribed Radiation Patterns". *IEEE Trans. on Antennas and Propagation*, 18:128–129, January 1971.

[Jaf88] A.G. Jaffer. "Maximum Likelihood Direction Finding of Stochastic Sources: A Separable Solution". In *ICASSP 88*, volume 5, pages 2893–2896, New York, April 1988.

[Jak74] W. Jakes. *Microwave Mobile Communications*. Wiley-Interscience, New York, 1974.

[JHW81] J.R. James, R.S. Hall, and C. Wood. *Microstrip Antenna Theory and Design*. Peregrinus, 1981.

[Jos94] L. Josefsson. "Mutual Coupling Effects on the Performance of Finite Radar Antenna Arrays". In *Proc. Radar 94*, pages 713–716, Paris, 3-6 May 1994.

[JS88] P.H. Janssen and P. Stoica. "On the Expectation of the Product of Four Matrix-Valued Gaussian Random Variables". *IEEE Trans. on Automatic Control*, 33(9):867–870, September 1988.

[Kay93] S. M. Kay. *Fundamentals of Statistical Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1993.

[KB86] M. Kaveh and A. J. Barabell. "The Statistical Performance of the MUSIC and the Minimum-Norm Algorithms in Resolving Plane Waves in Noise". *IEEE Trans. on ASSP*, ASSP-34:331–341, April 1986.

[KCH71] R.W.P. King and Jr. C.W. Harrison. "Scattering by Imperfectly Conducting Spheres". *IEEE Trans. on Antennas and Propagation*, AP-21(2):191–207, March 1971.

[KH69] R.W.P. King and C.W. Harrison. *Antennas and Waves: A Modern Approach*. The M.I.T. Press, 1969.

[Kil98] P-S. Kildal. *"Foundations of Antennas: A Unified Approach"*. Compendium, Gothenburg, 1998.

[Kil99] P-S. Kildal. "Equivalent Circuits of Receive Antennas in Signal Processing Arrays". *Microwave and Optical Technology Letters*, 21(4):244–246, May 1999.

[KJSWW99] J. Khun-Jush, P. Schramm, U. Wachsmann, and F. Wenger. "Structure and Performance of the HiperLAN2 Physical Layer". In *Proc. IEEE VTC 1999 Fall*, pages 2667 – 2671, September 1999.

[KP85] Y-W. Kang and D.M. Pozar. "Correction of Error in Reduced Sidelobe Synthesis due to Mutual Coupling". *IEEE Trans. on Antennas and Propagation*, AP-33:1025–1028, September 1985.

[Kra50] J. D. Kraus. *Antennas*. McGraw-Hill, New York, 1950.

[Kra88] J. D. Kraus. *Antennas*. McGraw-Hill, New York, second edition, 1988.

[KRB00] A. Kuchar, J-P. Rossi, and E. Bonek. "Directional Macro-Cell Channel Characterization from Urban Measurements". *IEEE Trans. on Antennas and Propagation*, 48(2):137–146, February 2000.

[Kri99] G. Kristensson. *Spridningsteori med antenntillämpningar*. Studentlitteratur, 1999. (in swedish).

[KS61] M.G. Kendall and A. Stuart. *The Advanced Theory of Statistics*, volume 2. Charles Griffin & Company Limited, London, 1961.

[KSMP00] J.P. Kermoal, L. Schumacher, P.E. Mogensen, and K.I. Pedersen. "Experimental Investigation of Correlation Properties of MIMO Radio Channels for Indoor Picocell Scenarios". In *Proc. IEEE VTC 2000 Fall*, Boston, MA, September 2000.

[KTS84] S. Kozono, T. Tsuruhara, and M. Sakamoto. "Base Station Polarization Diversity Reception for Mobile Radio". *IEEE Trans. on Vehicular Technology*, 33(4):301–306, November 1984.

[KV96] H. Krim and M. Viberg. "Two Decades of Array Signal Processing Research: The Parametric Approach". *IEEE Signal Processing Magazine*, 13(4):67–94, July 1996.

[Lac71] R.T. Lacoss. "Data Adaptive Spectral Analysis Methods". *Geophysics*, 36:661–675, 71.

[LC95] K.C. Lee and T.H. Chu. "A Circuit Model for Antenna Array Mutual Coupling Effects". In *Proc. IEEE APS Int. Symp.*, volume 2, pages 946–949, New York, NY, 1995.

[LDvdV99] A.N. Lemma, E.F. Deprettere, and A.-J van der Veen. "Experimental Analysis of Antenna Coupling for High-Resolution DOA Estimation Algorithms". In *Proc. IEEE SPAWC*, pages 362–365, 1999.

[Lee73] W.C. Lee. "Effects on Correlation Between Two Mobile Radio Base-Station Antennas". *IEEE Trans. on Communications*, COM-21(11):1214–1224, November 1973.

[Lee89] W.C. Lee. *Mobile Cellular Telecommunications Systems*. McGraw-Hill, 1989.

[Lig73] W.S. Liggett. *Passive Sonar: Fitting Models to Multiple Time Series*. Academic Press, New York, j.w.r. griffiths and p.l. stocklin and c. vam schoonefeld edition, 1973.

[Lin00] B. Lindmark. *"Analysis and Design of Base Station Antennas for Mobile Communications"*. PhD thesis, Chalmers University of Technology, Göteborg, Sweden, 2000.

[LLL97] M. Lu, T. Lo, and J. Litva. "A Physical Spatio-Temporal Model of Multipath Propagation Channels". In *Proc. IEEE VTC 97*, pages 810–814, Phoenix, AZ, May 1997.

[LLSB98] B. Lindmark, S. Lundgren, J.R. Sanford, and C. Beckman. "Dual-Polarized Array for Signal-Processing Applications in Wireless Communications". *IEEE Trans. on Antennas and Propagation*, 46(6):758–763, June 1998.

[LM87] J.T. Lo and S.L. Marple. "Eigenstructure Methods for Array Sensor Localization". In *Proc. IEEE ICASSP 87*, volume 4, pages 2260–2264, Dallas, USA, 1987.

[LP00] A. Lozano and C. Papadias. "Space-Time Receiver for Wideband BLAST in Rich-Scattering Wireless Channels". In *Proc. IEEE VTC 2000 Spring*, Tokyo, Japan, May 2000.

[LR96] J.C. Liberti and T.S. Rappaport. "A Geometrically Based Model for Line-Of-Sight Multipath Radio Channels". In *Proc. of IEEE VTC 96*, pages 844–848, April, 1996.

[LR99] J.C. Liberti, Jr. and T.S. Rappaport. *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*. Prentice-Hall, 1999.

[LSM00] J. Luo, J. Seidler, and S. McLaughlin. "Sensitivity Analysis of Compact Antenna Arrays in Correlated Nakagami Fading Channels". In *Proc. IEEE VTC 2000 Fall*, Boston, MA, September 2000.

[LSR79] Y.T. Lo, D. Solomon, and W.F. Richards. "Theory and Experiment in Microstrip Antennas". *IEEE Trans. on Antennas and Propagation*, 27:137–145, 1979.

[Lun96] S. Lundgren. "A Study of Mutual Coupling Effects on the Direction Finding Performance of ESPRIT With a Linear Microstrip Patch Array Using the Method of Moments". In *Proc. IEEE AP-S Symposium*, pages 1372–1375, Baltimore, July 1996.

[Lüt96] H. Lütkepohl. *Handbook of Matrices*. John Wiley & Sons, West Sussex, England, 1996.

[LZ90] J. Litva and M. Zeytinoglu. "Application of High-Resolution Direction Finding Algorithms to Circular Arrays with Mutual Coupling Present". Technical report, Part II prepared for the Defense Research Establishment Ottawa, McMaster University, Ontario, Canada, July 1990.

[MBKF00] D.P. McNamara, M.A. Beach, P. Karlsson, and P.N. Fletcher. "Initial Characterisation of Multiple-Input Multiple-Output (MIMO) Channels for Space-Time Communication". In *Proc. IEEE VTC 2000 Fall*, pages 1193–1197, Boston, MA, September 2000.

[Mey00] R.P. Meys. "A Summary of the Transmitting and Receiving Properties of Antennas". *IEEE Antennas and Propagation Magazine*, 42(3):49–53, June 2000.

[Mie08] G. Mie. "Beitrage zur Optik truber Meiden, speziell Kolloidaler Metallosungen". *Ann. Physik*, 25:377, 1908.

[Mil99] T. Milligan. "More Applications of Euler Rotation Angles". *IEEE Antennas and Propagation Magazine*, 41(4):78–83, August 1999.

[MKB79] K.V. Mardia, J.T. Kent, and J.M. Bibby. *Multivariate Analysis*. Academic Press, London, 1979.

[MLEP+00] P. Mogensen, P. Leth-Espensen, K.I. Pedersen, F. Fredriksen, and P. Zetterberg. "Performance of Adaptive Antennas in FH-GSM Using Conventional Beamforming". *Wireless-Personal-Communications*, 14(3):255–274, September 2000.

[Mol98a] B-A Molin. "PCC- A Distributed Multidisciplinary Research Program in Personal Computing and Communication". In *Proc. IEEE Globecom 98*, volume 6, pages 3159–3164, 1998.

[Mol98b] B-A Molin. "Personal Computing and Communication- A Swedish Strategic Research Program". In *Proc. PIMRC 1998*, volume 3, pages 1570–1574, Boston, MA, September 1998.

[MS69] V.H. MacDonald and P.M. Schultheiss. "Optimum Passive Bearing Estimation in a Spatially Incoherent Noise Environment". *J. Acoust. Soc. Am.*, 46(1):37–43, 1969.

[MWS00] C.C. Martin, J.H. Winters, and N.R. Sollenberger. "Multiple-Input-Multiple-Output (MIMO) Radio Channel Measurements". In *Proc. IEEE VTC 2000 Fall*, Boston, MA, September 2000.

[NA98] O. Nørklit and J.B. Andersen. "Diffuse Channel Model and Experimental Results for Array Antennas in Mobile Communications". *IEEE Trans. on Antennas and Propagation*, 46(6):834–840, June 1998.

[NLTW98] A. Narula, M.J. Lopez, M.D. Trott, and G.W. Wornell. "Efficient Use of Side Information in Multiple-Antenna Data Transmission over Fading Channels". *IEEE Journal on Selected Areas in Communications*, 16(8):1423–1436, October 1998.

[NP94] A. Nehorai and E. Paldi. "Vector-Sensor Array Processing for Electromagnetic Source Localization". *IEEE Trans. on Signal Processing*, 42:376–398, February 1994.

[NP97] A. Nehorai and E. Paldi. "Electromagnetic Vector-Sensor Array Processing". In V.K. Madisetti and D.B. Williams, editors, *The Digital Signal Processing Handbook*. CRC Press, 1997.

[NPP95] A. Nehorai, B. Porat, and E. Paldi. "Detection and Localization of Vapor-Emitting Sources". *IEEE Trans. on SP*, 43:243–253, Jan. 1995.

[NS96a] B.C. Ng and C.M.S. See. "Sensor-Array Calibration Using a Maximum-Likelihood Approach". *IEEE Trans. on Antennas and Propagation*, 44(6):827–835, June 1996.

[NS96b] G. Nilsson and T. Svantesson. "Analysis of Radar Backscatter of Stones Using a Modified Extended Physical Optics Approximation". Master's thesis, Chalmers University of Technology, 1996.

[NSC00] A.F. Naguib, N. Seshadri, and A.R. Calderbank. "Increasing Data Rate Over Wireless Channels". *IEEE Signal Processing Magazine*, 17(3):76–92, May 2000.

[NSS91] A. Nehorai, D. Starer, and P. Stoica. "Direction of Arrival Estimation in Applications with Multipath and Few Snapshots". *Circuits, Systems, and Signal Processing*, 10(3):327–342, 1991.

[Oku68] Y. Okumura et al. "Field Strength and Its Variability in the VHF and UHF Land Mobile Radio Service". *Review Elec. Commun. Lab.*, 16:825–843, 1968.

[OVK90] B. Ottersten, M. Viberg, and T. Kailath. "Asymptotic Robustness of Sensor Array Processing Methods". In *Proc. ICASSP 90*, pages 2635–2638, Albuquerque, NM, April 1990.

[OVK91] B. Ottersten, M. Viberg, and T. Kailath. "Performance Analysis of the Total Least Squares ESPRIT Algorithm". *IEEE Trans. on Signal Processing*, 39(5):1122–1135, May 1991.

[OVSN93] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai. "Exact and Large Sample ML Techniques for Parameter Estimation and Detection in Array Processing". In Haykin, Litva, and Shepherd, editors, *Radar Array Processing*, pages 99–151. Springer-Verlag, Berlin, 1993.

[Ows85] N.L. Owsley. "Sonar Array Processing". In S. Haykin, editor, *Array Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1985.

[OWVK89] B. Ottersten, B. Wahlberg, M. Viberg, and T. Kailath. "Stochastic Maximum Likelihood Estimation in Sensor Arrays by Weighted Subspace Fitting". In *Proc.* 23^{rd} *Asilomar Conf. Sig., Syst.,Comput.*, pages 599–603, Monterey, CA, November 1989.

[PAKM00] K.I. Pedersen, J.B. Andersen, J.P. Kermoal, and P. Mogensen. "A Stochastic Multiple-Input-Multiple-Output Radio Channel Model for Evaluation of Space-Time Coding Algorithms". In *Proc. IEEE VTC 2000 Fall*, Boston, MA, 2000.

[Pap91] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, third edition, 1991.

[Par92] D. Parsons. *The Mobile Radio Propagation Channel*. Pentech Press, New York-Toronto, 1992.

[PDS97] L. Pettersson, M. Danestig, and U. Sjöström. "An Experimental S-Band Digital Beamforming Antenna". *IEEE Aerospace and Electronics Systems Magazine*, May 1997.

[Ped00] K.I. Pedersen. *"Antenna Arrays in Mobile Communications"*. PhD thesis, Center for PersonKommunikation, Aalborg University, Denmark, January 2000.

[Pel99] P. Pelin. *"Space-time Algorithms for Mobile Communications"*. PhD thesis, Chalmers University of Technology, Göteborg, Sweden, 1999.

[PF94] K.M. Pasala and E.M. Friel. "Mutual Coupling Effects and Their Reduction in Wideband Direction of Arrival Estimation". *IEEE Trans. on Aerospace and Electronic Systems*, 30(4):1116–1122, October 1994.

[PK83] A.G. Papayiannakis and E.E. Kriezis. "Scattering from a Dielectric Cylinder of Finite Length". *IEEE Trans. on Antennas and Propagation*, 31(5):725–731, September 1983.

[PK85] A. Paulraj and T. Kailath. "Direction of Arrival Estimation by Eigenstructure Methods with Unknown Sensor Gain and Phase". In *Proc. IEEE ICASSP*, pages 17.7.1–17.7.4, Tampa, FL, March 1985.

[PK91] J. Pierre and M. Kaveh. "Experimental Performance of Calibration and Direction-Finding Algorithms". In *Proc. IEEE ICASSP*, volume 2, pages 1365–1368, Toronto, Canada, May 1991.

[PMF98] K.I. Pedersen, P.E. Mogensen, and B.H. Fleury. "Spatial Channel Characteristics in Outdoor Environments and their Impact on BS Antenna System Performance". In *Proc. 48th IEEE VTC*, volume 2, pages 719–723, May 1998.

[PMF99] K.I. Pedersen, P.E. Mogensen, and B.H. Fleury. "Dual-Polarized Model of Outdoor Propagation Environments for Adaptive Antennas". In *Proc. 49th IEEE VTC*, volume 2, pages 990–995, July 1999.

[Poo94] H.V. Poor. *An Introduction to Signal Detection and Estimation*. Springer-Verlag, Berlin, second edition, 1994.

[Pro95] J.G. Proakis. *Digital Communications*. McGraw-Hill, 1995.

[PRR96] P. Petrus, J.H. Reed, and T.S. Rappaport. "Geometrically Based Statistical Channel Model for Macrocellular Mobile Environments". In *Proc. IEEE Global Telecommunications Conference*, volume 2, pages 1197–1201, November 1996.

[PS60] J.N. Pierce and S. Stein. "Multiple Diversity with Nonindependent Fading". *Proc. of the IRE*, 48:89–104, January 1960.

[PT97] S.L. Preston and D.V. Thiel. "Direction Finding Using a Switched Parasitic Antenna Array". In *Proc. IEEE 1997 AP-S*, pages 1024–1027, Montreal, Canada, July 1997.

[PTL+97] S.L. Preston, D.V. Thiel, J.W. Lu, S.G. O'Keefe, and T.S. Bird. "Electronic Beam Steering Using Switched Parasitic Patch Elements". *Electronic letters*, 33:7–8, 1997.

[PTS+98] S.L. Preston, D.V. Thiel, T.A. Smith, S.G. O'Keefe, and J.W. Lu. "Base-Station Tracking in Mobile Communications Using a Switched Parasitic Antenna Array". *IEEE Trans. on Antennas and Propagation*, 46(6):841–844, June 1998.

[Ran99] A. Ranheim. *"Interference Rejection in Wireless Communication Systems"*. PhD thesis, Chalmers University of Technology, Göteborg, Sweden, 1999.

[RDNP95] G. Raleigh, S.N. Diggavi, A.F. Naguib, and A Paulraj. "Characterization of Fast Fading Vector Channels for Multi-Antenna Communication Systems". In *Proc.* 28^{rd} *Asilomar Conf. Sig., Syst., Comput.*, pages 853–857, Pacific Grove, CA, 1995.

[RH89] B.D. Rao and K.V.S. Hari. "Performance Analysis of Root-Music". *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 37(12):1939–1949, December 1989.

[Ric65] J.H. Richmond. "Scattering by a Dielectric Cylinder of Arbitrary Cross Section Shape". *IEEE Trans. on Antennas and Propagation*, 13:334–341, May 1965.

[Roy87] R. H. Roy. *"ESPRIT, Estimation of Signal Parameters via Rotational Invariance Techniques"*. PhD thesis, Stanford University, CA, August 1987.

[RS87] Y. Rockah and P.M. Schultheiss. "Array Shape Calibration Using Sources in Unknown Locations-Part I: Far-Field Sources". *IEEE Trans. on Antennas and Propagation*, 35(3):286–299, March 1987.

[RSB00] A. Ranheim, T. Svantesson, and P. Bohlin. "Combining Spatial and Temporal Processing in a CDMA System Using an Array of Antennas". In *Proc. Nordiskt antennsymposium*, pages 137–142, Lund, Sweden, September 2000.

[Ruc70] G.T. Ruck. *Radar Cross Section Handbook*, volume 2. Plenum Press, 1970.

[Rud76] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, 1976.

[RW92] C. Roller and W. Wasylkiwskyj. "Effects of Mutual Coupling on Super-Resolution DF in Linear Arrays". In *Proc. IEEE ICASSP*, pages V257–V260, San Francisco, CA, 1992.

[SAGA98] I.S.D. Solomon, Y.I. Abramovich, D.A. Gray, and S.J. Anderson. "Performance of OTH Radar Array Calibration". In *Proc. IEEE ICASSP 98*, volume 4, pages 2025–2028, Seattle, Washington, USA, May 1998.

[Sau99] S.R. Saunders. *Antennas and Propagation for Wireless Communication Systems*. John Wiley and sons, Inc., 1999.

[Sch68] F.C. Schweppe. "Sensor Array Data Processing for Multiple Signal Sources". *IEEE Trans. on IT*, IT-14:294–305, 1968.

[Sch79] R.O. Schmidt. "Multiple Emitter Location and Signal Parameter Estimation". In *Proc. RADC Spectrum Estimation Workshop*, pages 243–258, Rome, NY, 1979.

[Sch81] R. O. Schmidt. *A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation*. PhD thesis, Stanford Univ., Stanford, CA, November 1981.

[SDL00] T. Su, K. Dandekar, and H. Ling. "Simulation of Mutual Coupling Effect in Circular Arrays for Direction-Finding Applications". *Microwave and Optical Technology Letters*, 26(5):331–336, September 2000.

[Ser95] G.V. Serebryakov. "Direction-of-Arrival Estimation of Correlated Sources by Adaptive Beamforming". *IEEE Trans. on Signal Processing*, 43(11):2782–2787, November 1995.

[SF52] S.A. Schelkunoff and H.T. Friis. *Antennas: Theory and Practice*. Applied Mathematics Series. John Wiley and sons, Inc., 1952.

[SFGK00] D-S Shiu, G.J. Foschini, M.J. Gans, and J.M. Kahn. "Fading Correlation and Its Effect on the Capacity of Multielement Antenna systems". *IEEE Trans. on Communications*, 48(3):502–513, March 2000.

[SH69] B.J. Strait and K. Hirasawa. "Array Design for a Specified Pattern by Matrix Methods". *IEEE Trans. on Antennas and Propagation*, AP-17:237–239, March 1969.

[SH90] H. Steyskal and J. S. Herd. "Mutual Coupling Compensation in Small Array Antennas". *IEEE Trans. on Antennas and Propagation*, 38(12):1971–1975, December 1990.

[Sha48] C.E. Shannon. "A Mathematical Theory of Communication". *Bell. Syst. Tech. J.*, 27:623–656, October 1948.

[SI94] K.R. Schaubach and N.J Davis IV. "Microcellular Radio-Channel Propagation Prediction". *IEEE Antennas and Propagation Magazine*, 36(4):25–34, August 1994.

[Sil70] S.D. Silvey. *Statistical Inference*. Penguin, London, 1970.

[SJBF00] M. Stege, J. Jelitto, M. Bronzel, and G. Fettweis. "A Multiple Input-Multiple Output Channel Model for Simulation of TX- and RX-Diversity Wireless Systems". In *Proc. IEEE VTC 2000 Fall*, Boston, MA, September 2000.

[SJJS00] Q.H. Spencer, B.D. Jeffs, M.A. Jensen, and A.L. Swindlehurst. "Modeling the Statistical Time and Angle of Arrival Characteristics of an Indoor Multipath Channel". *IEEE Journal on Selected Areas in Communications*, 18(3):347–360, March 2000.

[SK92] A. Swindlehurst and T. Kailath. "A Performance Analysis of Subspace-Based Methods in the Presence of Model Errors, Part I: The MUSIC Algorithm". *IEEE Trans. on Signal Processing*, 40(7):1758–1774, July 1992.

[SK93] A. Swindlehurst and T. Kailath. "A Performance Analysis of Subspace-Based Methods in the Presence of Model Errors-Part II: Multidimensional Algorithms". *IEEE Trans. on Signal Processing*, 41(9):1277–1308, September 1993.

[Sko62] M. I. Skolnik. *Introduction to Radar System*. McGraw-Hill, New York, 1962.

[SL01] T. Su and H. Ling. "On Modeling Mutual Coupling in Antenna Arrays Using the Coupling Matrix". *Microwave and Optical Technology Letters*, 28(4):231–237, February 2001.

[SLTV99] N.L. Scott, M.O. Leonard-Taylor, and R.G. Vaughan. "Diversity Gain from a Single-Port Adaptive Antenna Using Switched Parasitic Elements Illustrated with a Wire and Monolopole Prototype". *Academic Press*, 47:1066–1070, 1999.

[SM97] P. Stoica and R. Moses. *Introduction to Spectral Analysis*. Prentice Hall, Upper Saddle River, NJ, 1997.

[SN89] P. Stoica and A. Nehorai. "MUSIC, Maximum Likelihood and Cramér-Rao Bound". *IEEE Trans. ASSP*, ASSP-37(5):720–741, May 1989.

[SN90a] P. Stoica and A. Nehorai. "MODE, Maximum Likelihood and Cramér-Rao Bound: Conditional and Unconditional Results". In *Proc. ICASSP 90 Conf*, pages 2715–2718, Albuquerque, NM, April 1990.

[SN90b] P. Stoica and A. Nehorai. "Performance Study of Conditional and Unconditional Direction-of-Arrival Estimation". *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 38(10):1783–1795, October 1990.

[SN94] D. Starer and A. Nehorai. "Path-Following Algorithm for Passive Localization of Near-Field Sources". *IEEE Trans. on Signal Processing*, 42(3):677–680, March 1994.

[SN95] P. Stoica and A. Nehorai. "On the Concentrated Stochastic Likelihood Function in Array Signal Processing". *IEEE Trans. on Circuits and Systems*, 14(5):669–674, 1995.

[SOK00] R. Stridh, B. Ottersten, and P. Karlsson. "MIMO Channel Capacity on a Measured Indoor Radio Channel at 5.8 GHz". In *Proc. 34th Asilomar Conf. Sig., Syst., Comput.*, Pacific Grove, CA, October 2000.

[Sor98] T.B. Sorensen et al. "Performance of Two-Branch Polarisation Antenna Diversity in an Operational GSM Network". In *Proc. 48th IEEE VTC*, volume 2, pages 741–746, 1998.

[Spe96] R.A. Speciale. "Advanced Design of Phased-Array Beam-Forming Networks". *IEEE Antennas and Propagation Magazine*, 38(4):22–34, August 1996.

[SR01] T. Svantesson and A. Ranheim. "Mutual Coupling Effects on the Capacity of Multielement Antenna Systems". In *Proc. IEEE ICASSP 01*, Salt Lake City, Utah, May 2001. To appear.

[SS89] T. Söderström and P. Stoica. *System Identification*. Prentice-Hall, London, U.K., 1989.

[SS90a] K. Sarabandi and T.B.A. Senior. "Low-Frequency Scattering From Cylindrical Structures at Oblique Incidence". *IEEE Trans. Geosci. Remote Sensing*, 28(5):879–885, September 1990.

[SS90b] P. Stoica and K.C. Sharman. "Novel Eigenanalysis Method for Direction Estimation". *Proc. IEE*, F:19–26, February 1990.

[SS96] J.M. Stiles and K. Sarabandi. "A Scattering Model for Thin Dielectric Cylinders of Arbitrary Cross Section and Electrical Length". *IEEE Trans. on Antennas and Propagation*, 44(2):260–266, February 1996.

[ST98] W.L. Stutzman and G.A. Thiele. *Antenna Theory and Design*. John Wiley and sons, Inc., New York, second edition, 1998.

[Ste92] R. Steele. *Mobile Radio Communications*. IEEE Press, New York, 1992.

[STLO00] R. Schlub, D.V. Thiel, J.W. Lu, and S.G. O'Keefe. "Dual-Band Six-Element Switched Parasitic Array For Smart Antenna Cellular Communications Systems". *Electronic Letters*, 36:1342–1343, 2000.

[Str01] R. Stridh. "Smart Antennas in Fourth Generation Wireless Systems: System Issues and Capacity". Licentiate thesis, Department of Signals, Sensors, and Systems, Royal Institute of Technology, Stockholm, Sweden, 2001.

[SV87] A.A.M Saleh and R.A. Valenzuela. "A Statistical Model for Indoor Multipath Propagation". *IEEE Journal on Selected Areas in Communications*, 5:128–137, February 1987.

[SV97] T. Svantesson and M. Viberg. "On the Direction Finding Accuracy of a Calibrated Array in the Presence of Mutual Coupling". In *Proc. Nordiskt antennsymposium*, pages 385–392, Göteborg, Sweden, May 1997.

[SV98] T. Svantesson and M. Viberg. "Mutual Coupling in Antenna Arrays: Effects and Cures". In *Proc. PCC Workshop*, pages 99–103, Stockholm, November 1998.

[SV99] T. Svantesson and M. Viberg. "Adaptive Antennas and Channel Modeling: The First Wild Years". In *Proc. PCC Workshop*, pages 74–77, Lund, November 1999.

[SV01] T. Svantesson and M. Viberg. "A Radio Channel Model for Multielement Antenna Systems". In *Proc. PCC Workshop*, Nynäshamn, April 2001.

[Sva98a] T. Svantesson. "The Effects of Mutual Coupling on the Direction Finding Accuracy of a Linear Array of Thin Dipoles of Finite Length". Technical report, Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, 1998.

[Sva98b] T. Svantesson. "The Effects of Mutual Coupling Using a Linear Array of Thin Dipoles of Finite Length". In *Proc. 8th IEEE Signal Processing Workshop On Statistical Signal and Array Processing*, Portland, USA, September 1998.

[Sva99a] T. Svantesson. "Antennas and Propagation from a Signal Processing Perspective". In *Proc. NRS "Antennen och dess omvärld"*, pages 35–38, Lund, November 1999.

[Sva99b] T. Svantesson. "Direction Finding in the Presence of Mutual Coupling". Licentiate thesis, Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, March 1999.

[Sva99c] T. Svantesson. "Methods of Mitigating the Effects of Mutual Coupling in Antenna Arrays: A Signal Processing Approach". In *Proc. RadioVetenskap och Kommunikation (RVK 99)*, pages 441–445, Karlskrona, Sweden, June 1999.

[Sva99d] T. Svantesson. "Modeling and Estimation of Mutual Coupling in a Uniform Linear Array of Dipoles". In *Proc. ICASSP 99*, pages 2961–2964, Phoenix, USA, March 1999.

[Sva99e] T. Svantesson. "Modeling and Estimation of Mutual Coupling in a Uniform Linear Array of Dipoles". Technical report, Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, 1999.

[Sva00a] T. Svantesson. "A Spatio-Temporal Channel Model Based on EM Scattering". In *Proc. IEEE AP-S*, pages 1110–1113, Salt Lake City, UT, July 2000.

[Sva00b] T. Svantesson. "A Study of Polarization Diversity Using an Electromagnetic Spatio-Temporal Channel Model". In *Proc. IEEE VTC 2000 Fall*, Boston, MA, September 2000.

[Sva00c] T. Svantesson. "An Antenna Solution for MIMO Channels: The Multimode Antenna". In *Proc. 34th Asilomar Conf. Sig., Syst., Comput.*, Pacific Grove, CA, October 2000.

[Sva00d] T. Svantesson. "Multimode Antenna Diversity". In *Proc. Nordiskt antennsymposium*, pages 97–102, Lund, Sweden, September 2000.

[Sva00e] T. Svantesson. "Multimode Based Direction Finding". In *Proc. 34th Asilomar Conf. Sig., Syst., Comput.*, Pacific Grove, CA, October 2000.

[Sva00f] T. Svantesson. "Mutual Coupling Compensation Using Subspace Fitting". In *Proc. IEEE SAM 2000 Workshop*, pages 494–498, Boston, MA, March 2000.

[Sva00g] T. Svantesson. "Mutual Coupling From a Signal Processing Perspective". Submitted to *IEEE Trans. on Antennas and Propagation*, 2000.

[Sva00h] T. Svantesson. "On the Potential of Multimode Antenna Diversity". In *Proc. IEEE VTC 2000 Fall*, Boston, MA, September 2000.

[Sva01a] T. Svantesson. "A Physical MIMO Radio Channel Model for Multi-Element Multi-Polarized Antenna Systems". Submitted to IEEE VTC 01 Fall, Atlantic City, NY, October, 2001.

[Sva01b] T. Svantesson. "A Physical Spatio-Temporal Channel Model for MIMO Systems". Submitted to *IEEE Journal on Selected Areas in Communications*, 2001.

[Sva01c] T. Svantesson. "Correlation and Channel Capacity of MIMO Systems Employing Multimode Antennas". Submitted to *IEEE Trans. on Vehicular Technology*, 2001.

[Sva01d] T. Svantesson. "High-Resolution Direction Finding Using the Higher Modes of a Biconical Antenna". Submitted to *IEEE Signal Processing Letters*, 2001.

[Sva01e] T. Svantesson. "Physical Channel Modeling of Multi-Element Antenna Systems". Technical report, Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, 2001.

[SW01] T. Svantesson and M. Wennström. "High-Resolution Direction Finding Using a Switched Parasitic Antenna". In *Proc. IEEE SSP 01*, Singapore, August 2001. To appear.

[Swi92] A. Swindlehurst. "DOA Identifiability for Rotationally Invariant Arrays". *IEEE Trans. on Signal Processing*, 40(7):1825–1828, July 1992.

[TCJ+72] G.L. Turin, F.D. Clapp, T.L. Johnston, S.B. Fine, and D. Lavry. "A Statistical Model of Urban Multipath Propagation". *IEEE Trans. on Vehicular Technology*, 21:1–9, February 1972.

[Tel95] E. Telatar. "Capacity of Multi-Antenna Gaussian Channels". Technical report, AT&T Bell Labs, June 1995.

[TJC99] V. Tarokh, H. Jafarkhani, and A.R. Calderbank. "Space-Time Block Codes from Orthogonal Design". *IEEE Trans. on Communications*, 45:1456–1467, 1999.

[VA87] R.G. Vaughan and J.B. Andersen. "Antenna Diversity in Mobile Communications". *IEEE Trans. on Vehicular Technology*, 36:149–172, 1987.

[Vau86] R.G. Vaughan. "Signals in Mobile Communications". *IEEE Trans. on Vehicular Technology*, 35(4):133–145, 1986.

[Vau88] R.G. Vaughan. "Two-Port Higher Mode Circular Microstrip Antennas". *IEEE Trans. on Antennas and Propagation*, 36(3):309–321, March 1988.

[Vau90] R.G. Vaughan. "Polarization Diversity in Mobile Communications". *IEEE Trans. on Vehicular Technology*, 39(3):177–186, August 1990.

[Vau98] R.G. Vaughan. "Pattern Translation and Rotation in Uncorrelated Source Distributions for Multiple Beam Antenna Design". *IEEE Trans. on Antennas and Propagation*, 46(7):982–990, July 1998.

[Vau99] R.G. Vaughan. "Switched Parasitic Elements for Antenna Diversity". *IEEE Trans. on Antennas and Propagation*, 47:399–405, 1999.

[VB88] B.D. Van Veen and K.M. Buckley. "Beamforming: A Versatile Approach to Spatial Filtering". *Signal Proc. Magazine*, pages 4–24, April 1988.

[VO91] M. Viberg and B. Ottersten. "Sensor Array Processing Based on Subspace Fitting". *IEEE Trans. SP*, SP-39(5):1110–1121, May 1991.

[VOK91] M. Viberg, B. Ottersten, and T. Kailath. "Detection and Estimation in Sensor Arrays Using Weighted Subspace Fitting". *IEEE Trans. SP*, SP-39(11):2436–2449, November 1991.

[VS94] M. Viberg and A.L. Swindlehurst. "A Bayesian Approach to Auto-Calibration for Parametric Array Signal Processing". *IEEE Trans. on Signal Processing*, 42(12):3495–3506, December 1994.

[VS01] M. Viberg and T. Svantesson. "Direction-Of-Arrival Estimation in Mobile Communication Scenarios". In L. Godara, editor, *The Handbook on Antennas in Wireless Communications*. CRC Press, 2001. To be published.

[Wai55] J.R. Wait. "Scattering of a Plane Wave from a Circular Dielectric Cylinder at Oblique Incidence". *Can J. Phys.*, 1955.

[Wax92] M. Wax. "On Unique Localization of Constrained-Signals Sources". *IEEE Trans. on ASSP*, ASSP-40(6):1542–1587, June 1992.

[Wee68] W.L. Weeks. *Antenna Engineering*. McGraw-Hill, 1968.

[WF89] A.J. Weiss and B. Friedlander. "Array Shape Calibration Using Sources in Uknown Locations- A Maximum Likelihood Approach". *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 37(12):1958–1966, December 1989.

[WF97] A.J. Weiss and B. Friedlander. "Comparison of Signal Estimation Using Calibrated and Uncalibrated Arrays". *IEEE Trans. on Aerospace and Electronic Syst.*, 33(1):241–249, January 1997.

[WFGV98] P.W. Wolniansky, G.J. Foschini, G.D. Golden, and R.A. Valenzuela. "V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel". In *Proc. 1998 URSI Int. Symp. on Sign., Syst., and Elect.*, pages 295–300, 1998.

[Wis80] W.J. Wiscombe. "Improved Mie Scattering Algorithms". *Appl. Opt.*, 19:1505–1509, 1980.

[WK85] M. Wax and T. Kailath. "Detection of Signals by Information Theoretic Criteria". *IEEE Trans. on ASSP*, ASSP-33(2):387–392, April 1985.

[WKS+98] R. Wakabayashi, H. Kawakami, G. Sato, T. Amano, and Y. Suzuki. "Analysis of Course Errors on CVOR Antennas (Including Effects of Mutual Coupling Between Elements". *IEEE Trans. on Vehicular Technology*, 47(2):392–405, May 1998.

[WS01] M. Wennström and T. Svantesson. "An Antenna Solution for MIMO Channels: The Switched Parasitic Antenna". Submitted to PIMRC 2001, San Diego, CA, September, 2001.

[WZ89a] M. Wax and I. Ziskind. "Detection of the Number of Coherent Signals by the MDL Principle". *IEEE Trans. on ASSP*, ASSP-37(8):1190–1196, August 1989.

[WZ89b] M. Wax and I. Ziskind. "On Unique Localization of Multiple Sources by Passive Sensor Arrays". *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 37(7):996–1000, July 1989.

[Yac93] M.D. Yacoub. *Foundations of Mobile Radio Engineering*. CRC Press, 1993.

[YHZ98] H. Yuan, K. Hirasawa, and Y. Zhang. "The Mutual Coupling and Diffraction Effects on the Performance of a CMA Adaptive Array". *IEEE Trans. Vehicular Tech.*, 47(3):728–736, August 1998.

[YLU89] C. Yeh, M. Leou, and D.R. Ucci. "Bearing Estimations with Mutual Coupling Present". *IEEE Trans. on Antennas and Propagation*, 37(10):1332–1335, October 1989.

[Zet97] P. Zetterberg. *"Mobile Cellular Communications with Base Station Antenna Arrays: Spectrum Efficiency, Algorithms and Propagation Models"*. PhD thesis, Royal Institute of Technology, Stockholm,Sweden, 1997.

[ZFDW00] T. Zwick, C. Fischer, D. Didascalou, and W. Wiesbeck. "A Stochastic Spatial Channel Model Based on Wave-Propagation Modeling". *IEEE Journal on Selected Areas in Communications*, 18(1):6–15, January 2000.

[ZFW98] T. Zwick, C. Fischer, and W. Wiesbeck. "A Statistical Model for Indoor Environments Including Angle of Arrival". In *Proc. IEEE VTC 98*, pages 615–619, 1998.

[Zha97] W. Zhang. "Wide-Band Propagation Model for Cellular Mobile Radio Communications". *IEEE Trans. on Antennas and Propagation*, 45(11):1669–1678, November 1997.

[ZHM+00] T. Zwick, D. Hampicke, J. Mauer, A. Richter, G. Sommerkorn, R. Thomä, and W. Wiesbeck. "Results of Double-Directional Channel Sounding Measurements". In *Proc. IEEE VTC 2000 Spring*, pages 2497–2501, 2000.

[ZW88] I. Ziskind and M. Wax. "Maximum Likelihood Localization of Multiple Sources by Alternating Projection". *IEEE Trans. on ASSP*, ASSP-36:1553–1560, October 1988.