Parametric Characterization and Estimation of Dispersive Multi-Path Components with SAGE in Radio Propagation Channel

SIGNAL AND INFORMATION PROCESSING IN COMMUNICATIONS SYSTEMS

DEPARTMENT OF COMMUNICATION TECHNOLOGY

Group 1090

Lingfeng Liu and Daniel Klokmose Nielsen

June 6^{th} *2007*

AALBORG UNIVERSITY

AALBORG UNIVERSITY
Department of Communication Technology
Fredrik Bajers Vej 7 - 9220 Aalborg Øst - Phone: (+45) 9635 8080

TITLE:

Parametric Characterization and Estimation of Dispersive Multi-Path Components with SAGE in Radio Propagation Channel

PERIOD:
September 2006 - June 2007

PROJECT GROUP:
1090

GROUP MEMBERS:
Lingfeng Liu
Daniel Klokmose Nielsen

SUPERVISORS:
Xuefeng Yin
Troels Pedersen
Bernard Fleury

COPIES:
7

NUMBER OF PAGES:
116

ABSTRACT:

This project is concerned about development and estimation of a parametric model to describe nominal and dispersive values of DoA (Direction of Arrival) or DoD (Direction of Departure) and delay for a radio propagation channel. As parametric model a joint FB5 and truncated Gaussian is derived along with an approximated model using a multivariate truncated Gaussian. Both are maximum entropy models. The Maximum Likelihood Estimator is approximated by the Space-Alternating Generalized Maximization-Expectation algorithm, which requires an appropriate design of data spaces and proper initialization. Scenarios with time variant and invariant radio propagation channels are examined with synthetic and measurement data to assess the performance and validate the proposed parametric model. The SAGE estimator works and the results are consistent with other estimators.

ATTACHMENTS:
CD-ROM

- "A SAGE Algorithm for the Estimation of Direction Power Spectrum of Individual Path Components"
- "Characterization of the Azimuth-Elevation Power Spectrum of Individual Path Components"

Preface

This report has been prepared by group 1090 at the 9^{th} and 10^{th} semester of Signal and Information Processing in Communications M.Sc. at Aalborg University from September 2006 to June 2007.

Formulas, figures and tables are numbered in succession inside each chapter. References to formulas are notated as e.g. Eq. (1.1), references to figures are notated as e.g. Fig. (1.1). At the end of the report you can find the bibliography which entries are referred to as e.g. [1] in the report. A list of symbols and acronyms used in the report can be found at 7.

Whenever the word "channel" is mentioned in the report, it implicitly represents the radio propagation channel.

At the back cover the reader will find a CD-ROM attached containing the report in the pdf file format, Matlab source code and figures.

We would like to thank our supervisors Xuefeng Yin, Bernard H. Fleury and Troels Pedersen with their feedback on the project. Also, thanks to Nicolai Czink and Attaphongse Taparugssanagorn (Pong) for providing measurement data.

Aalborg University, June 7, 2007

Lingfeng Liu

Daniel Klokmose Nielsen

Contents

Wordlist

Symbols

T_i	Transmit antenna i
M_T	Number of transmit antennas
R_i	Receive antenna i
M_R	Number of transmit antennas
$s(t)$	Transmitted signal
$\boldsymbol{x}(t)$	Input signal of the receive antenna array
$\boldsymbol{y}(t)$	Output signal of the receive antenna array
$\boldsymbol{g}(t)$	Channel impulse response vector
$\boldsymbol{h}(\tau_k)$	Estimated channel impulse response vector at delay τ_k
$\boldsymbol{H}$	Vectorized estimated channel impulse response vector
$w(t)$	Noise variable
$\boldsymbol{w}(t)$	Noise vector
δ_w^2	Variance of the noise
$\boldsymbol{I}$	Identical matrix
$\boldsymbol{\Sigma}$	Covariance matrix of a random vector
$\boldsymbol{\mu}$	Mean of a random vector
$\boldsymbol{\Omega}$	Direction vector
τ	Delay variable
$a(\boldsymbol{\Omega}, \tau)$	Complex amplitude of the transmitted signal
$\boldsymbol{C}(\boldsymbol{\Omega})$	Antenna response vector
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	A Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\mathcal{S}$	Support space of direction vector
$\mathcal{T}$	Support space of delay variable
T_s	Sounding period
t_n	Start time of the nth sounding period
$\boldsymbol{p}$	Propagation vector
P	Power of one path component
θ	Elevation of the direction vector
ϕ	Azimuth of the direction vector
$\boldsymbol{\gamma}_1$ or $\bar{\boldsymbol{\Omega}}$	Mean direction of FB5
$\boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3$	Spread directions of FB5
α	Direction tilt angle of FB5
β	Delay tilt angle of joint direction and delay power distribution
$\boldsymbol{\alpha}_\theta$ and $\boldsymbol{\alpha}_\phi$	Basis for the orthogonal plane of $\boldsymbol{\gamma_1}$
κ	Spread factor of FB5
ζ	Angular power spread ratio in FB5

η	Delay power spread coefficient in the joint direction and delay power distribution
$\boldsymbol{c}$	Delay spread direction in the joint direction and delay power distribution
B	Delay spread coefficient
$\mathbb{R}^N$	N dimensional real space
L	Number of path components
K	Number of delay samples
N	Number of sounding periods
$\boldsymbol{S}$	Signal part of a hidden data
Rss(n)	Autocorrelation function of the transmitted signal, $s(t)$, at delay n
$\boldsymbol{b}_i$	Eigenvector i
$\boldsymbol{\lambda}_i$	The i'th eigenvalue
$\bar{(\cdot)}$	A mean value
$\hat{(\cdot)}$	An observation
$\tilde{(\cdot)}$	An estimate
$(\cdot)_i^j$	Estimation for the ith path component at the jth iteration

Operators

$\star$	Convolution operator
$(\cdot)^*$	Conjugate operator
$\|\cdot\|$	Norm operator
$(\cdot)^T$	Transpose operator
$(\cdot)^H$	Hermitian transpose operator

Acronyms

CDM	Code Division Multiplexing
DoA	Direction of Arrival
DoD	Direction of Departure
EM	Expectation-Maximization
FB5	The Fisher-Bingham distribution with 5 parameters.
FDM	Frequency Division Multiplexing
ISI	Inter Symbol Interference
LOS	Line Of Sight
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MUSIC	MUltiple SIgnal Classification
NLOS	Non Line Of Sight
OFDM	Orthogonal Frequency Division Multiplexing
PN	Pseudo Noise
Rx	Receiver
SAGE	Space-Alternating Generalized Expectation-maximization
SIMO	Single-Input-Single-Output
ST	Space Time
TDM	Time Division Multiplexing
Tx	Transmitter
WSSUS	White Sense Stationary Uncorrelated Scattering
ZMCSCG	Zero-Mean Circularly Symmetric Complex Gaussian

1 Introduction

1.1 Background

Since the late 1990's MIMO (Multiple-Input Multiple-Output) antenna systems have been the subject of enormous interest among engineers and researchers. The reason is that the theoretical capacity of a MIMO antenna system is capable to far exceed the performance of a Single-Input Single-Output (SISO) system [1][2]. The extra degrees of freedom can be used to increase reliability through space time diversity techniques [3] and/or to increase the data rate with space time multiplexing techniques [4].

It is the trend that 4G mobile systems[1] and beyond employ multi antenna transceivers to improve the spectral efficiency and link quality of the radio propagation channel. A combination of Orthogonal Frequency Division Multiplexing (OFDM) and MIMO antenna technology is one such implementation which exploits the spatial benefits of MIMO antenna arrays [5].

Channel modeling and estimation are important topics. They are introduced in order to capture appropriate information about the radio channel and serves as a way to simulate the channel environment. Channel estimation is a necessity to maintain the performance of spatial multiplexing [6], e.g. in MIMO-OFDM.

Radio propagation channel modeling and estimation provides an insight of how electromagnetic waves propagate and interact in real radio channels. With a proper model the influence of the antenna array can be removed from the observations [7], which enables a comparison of various MIMO transceiver structures.

The radio propagation channel can be viewed as a superposition of multiple propagation paths denoted as path components, illustrated in Fig. 1.1. Description of the propagation paths includes the Direction of Arrival (DoD), Direction of Departure (DoD), delay, polarization and Doppler frequency [8].

Radio propagation channel models may be divided into two categories: Specular propagation paths [9] and dispersive propagation paths [10]. Research on the characterization of dispersive propagation path and estimation is quite active in recent years. Betlehem et al. [11] showed that the shape and orientation of the joint power distribution contours have a major impact on capacity and diversity of non-line-of-sight (NLOS) channels. "The capacity is dependent upon the orientation of the distribution contour while the diversity order is proportional to the area of distribution contours".

[1] 4G is not just one defined technology or standard, but rather a collection of technologies and protocols

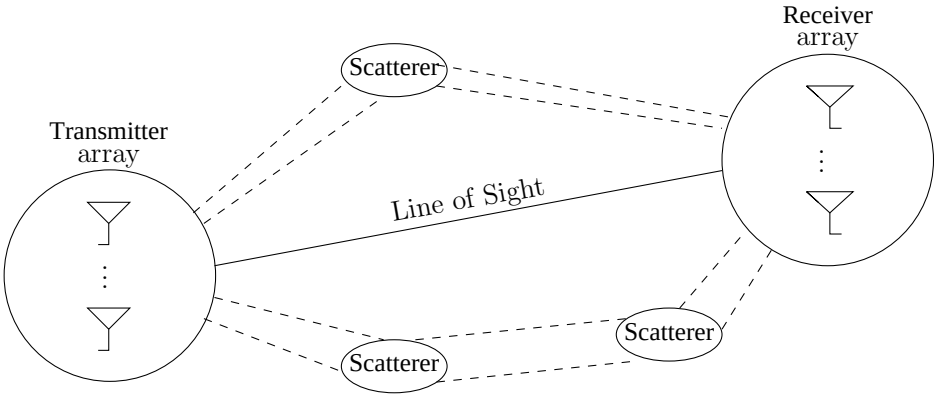


Figure 1.1: Illustration of a radio propagation channel. The signal departs from the Tx array and interacts with different scatterers changing its properties(e.g. amplitude, phase, delay and direction). Also a line of sight path is illustrated which may provide a strong signal with small dispersion in delay and direction.

1.1.1 High resolution parametric estimation methods

High resolution estimation parametric methods, formally based on specular path propagation are effective approaches to exploit the propagation characteristics of e.g. MIMO system. The high resolution parametric description of the multi antenna system can also be applied into navigation, locating and radar tracing and sensing as well.

Beam-forming

Beam-forming techniques are a subclass of spectral analysis techniques and they are commonly used with electromagnetic and audio signals. The techniques work by "steering" the sensitivity of a sensor array in a specific direction and are typically used as a first analysis for the power spectrum of the propagation paths. The Bartlett and Capon beam-former family are among the most popular beam-forming techniques [12].

MUSIC

The MUltiple SIgnal Classification (MUSIC) algorithm published in [13] is another high resolution spectral analysis technique. The idea behind the MUSIC is to divide the signal space into a noise subspace and signal plus noise subspace. The estimator will exhibit peaks in the vicinity of the DoA, which is where the steering vector will have a large distance to the orthogonal projector onto the noise subspace [12]. The MUSIC algorithm is very sensitive to estimation of the rank of the subspaces.

ESPRIT

The ESPRIT [14] algorithm also belongs to high resolution spectral analysis. The underlying antenna array is divided into sub-arrays. Each sub-array consists of the same number of

elements and the spacing between the those are known.

Compared to the MUSIC algorithm, the ESPRIT algorithm does not require any knowledge of the radiation pattern. However, the ESPRIT algorithm needs multiple sensor doublets. The elements in each doublet must have identical radiation patterns and are separated by a known constant distance. Except for the identity requirement for the radiation patterns of the elements in the same doublet, the identical radiation pattern may have arbitrary gain, phase and polarization.

EM

The Expectation-Maximum (EM) algorithm [15] is an approximation to the Maximum Likelihood Estimator (MLE) of the joint specular path components. The EM is designed on the expectation of the hidden data space for each specular path propagation and the performance of it is quite dependent on the algorithm design.

SAGE

The Space-Alternating Generalized Expectation-Maximization (SAGE) algorithm is introduced in [9]. SAGE is a further improvement for the EM with faster convergence rate and less complexity, but the design of hidden data spaces is crucial. The realization of SAGE depends on the expectation of hidden data for the individual path component.

Review of dispersive path modeling and estimation

The dispersive path components model is believed to be a more accurate model of the radio propagation channel compared to specular path components. Dispersion of the path components include direction, delay and frequency and it is described by a continuous function. In [10], a dispersive path model was proposed by only considering the horizontal angular spread. The bi-azimuth (DoA and DoD) dispersion was modeled by a von-Mises distribution and estimation was conducted with the SAGE.

Another dispersive path model of one single side (receiver or transmitter) was proposed in [16][17] where both the horizontal and vertical angular spread were considered. The joint horizontal and vertical power angular spread was modeled by the Fisher-Bingham-5(FB5) distribution and an investigation was conducted to estimate multi-path components using the SAGE in narrow band scenarios.

1.2 Problem statement

This project is concerned about how to extend the dispersive path component model into more dimensions specifically in delay. In [16], the authors focused on a single side of the radio propagation channel in a narrow band scenario, where the latter assumption means that only dispersion of path components in direction were considered. An extension into a wide-band scenario requires modeling of path component dispersion in both direction and delay.

Furthermore, the project will also consider the multi-path components estimation problem. In this matter the SAGE algorithm will be introduced in order to obtain a practical estimator of the multi-path components.

We only consider SIMO and MISO systems since the results may be extended to MIMO systems. Also doppler frequency and polarization will not be investigated. Our interests is to model and estimate the power spectrum of dispersive path components in both time variant and time invariant radio propagation channel conditions. The phase distribution of path component will be analyzed for the time invariant scenario.

In short, the problem of the project can be described in the following.

1. To find an appropriate model of the radio propagation channel in SIMO and MISO systems with dispersive path components in both direction and time.
2. To design a well structured estimation algorithm with preferably fast convergence.

1.3 Report overview

This report is divided into 3 parts: A general part about the channel model; channel propagation characterization and estimation for time variant scenario; channel propagation characterization and estimation for the time invariant scenario.

Part I

Chapter 2 describes signal models of SIMO and MISO systems. Channel sounding techniques are introduced. The radio channel propagation is modeled under the far-field assumption to simplify the radio channel propagation model. A dispersive path component model is introduced to describe a group of correlated path components.

Part II

In Chapter 3 the power distribution model of the dispersive path components is analyzed. The maximum entropy modeling principle is used. The FB5 distribution is suggested as the power distribution in direction and the truncated Gaussian distribution is suggested as the power distribution in delay. A joint power distribution in direction and delay is derived based on maximum entropy principle and also approximated by a multivariate Gaussian distribution.

Chapter 4 analyzes the estimation method in the time variant scenario for multi-path components. The SAGE is chosen as a solution for the estimation. Principles and the general structure of the SAGE are discussed.

Chapter 5 provides the specific algorithm design for the channel propagation estimation in a time variant scenario. A detailed initialization procedure design for SAGE is given.

Chapter 6 documents the estimation results from both synthetic and real measurement data. Observation for the measurement data is described. The analysis of the model mismatch

between the model and the measurement is carried out and a solution to minimize the model mismatch is proposed.

Part III

Chapter 7 suggests a simplified distribution of the path components for the time invariant scenario by using the power distribution model in Chapter 3 and an identical phase distribution.

Chapter 8 modifies the SAGE for the time invariant scenario.

Chapter 9 modifies the algorithm to the time invariant scenario.

Chapter 10 provides estimation results of the measurement data in the time invariant scenario. Improvements to minimize the model mismatch is again discussed.

Chapter 11 concludes the project. Suggestions to improve the radio channel model are presented and further research topics in this field of research are provided as well.

2 System Modeling

This chapter introduces the basic structure of MIMO, SIMO and MISO systems [6]. The channel sounding techniques help to extract a SIMO or MISO system from a MIMO system. The analysis indicates that by using channel sounding techniques the SIMO and MISO system will have a reciprocal relation in signal model and radio propagation patterns. The concept of continuous distributed path components is introduced as well as several assumptions for the radio propagation channel.

2.1 General structure of MIMO, SIMO and MISO systems

Considering a general MIMO system with M_T transmit antennas (Tx) and M_R receive antennas (Rx) working simultaneously, the radio channel of the MIMO system can be viewed as a coexistence of many sub-channels formed by each pair of transmit and receive antenna. The composition can be represented by the following sub-channel matrix.

$$\begin{bmatrix} (T_1,R_1) & (T_1,R_2) & \dots & & (T_1,R_{M_R}) \\ (T_2,R_1) & (T_2,R_2) & & & \\ & & \ddots & & \\ \vdots & & & (T_i,R_j) & \vdots \\ & & & & \ddots \\ (T_{M_T},R_1) & & & & (T_{M_T},R_{M_R}) \end{bmatrix},$$

where each entry represents one sub-channel formed by the T_i transmit antenna and R_j receive antenna.

A SIMO system can be viewed as a special MIMO system where only one transmit antenna works. A selection of all the sub-channels in a row of the sub-channel matrix leads to a SIMO system with one transmit antenna T_i and all M_R receiver antennas. Similarly a selection of a column in the sub-channel matrix leads to a MISO system with one transmit antenna R_i and all M_T receiver antennas.

For either MIMO, SIMO or MISO system, each sub-channel can be described by its channel impulse response. The channel impulse response describes how the radio channel changes the transmitted signal in both amplitude and phase. In the interest of research, we only discuss the channel impulse responses for SIMO and MISO systems.

SIMO signal model

Consider a SIMO system with one transmit antenna and M_R receive antennas, where the signal $s(t)$ is transmitted. The output signal from the receive antenna array is a $M_R \times 1$ vector denoted as $\boldsymbol{y}(t)$.

$$\boldsymbol{y}(t) = \boldsymbol{g}(t) \star s(t) + \boldsymbol{w}(t) \tag{2.1}$$

or equivalently:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \\ \vdots \\ y_{M_R}(t) \end{bmatrix} = \begin{bmatrix} g_1(t) \\ g_2(t) \\ g_3(t) \\ \vdots \\ g_{M_R}(t) \end{bmatrix} \star s(t) + \begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \\ \vdots \\ w_{M_R}(t) \end{bmatrix},$$

where $\star$ is the convolution symbol, $g_i(\tau, t)$ is the sub-channel impulse response and $\boldsymbol{w}(t)$ is assumed to be additive zero-mean spatially and temporally white Gaussian noise:

$$\boldsymbol{w}(t) \sim \mathcal{N}(0, \delta_w^2 \boldsymbol{I}),$$

where $\boldsymbol{I}$ is a $M_R \times M_R$ identical matrix and δ_w^2 is the variance of the noise.

MISO signal model

In a MISO system with M_T transmit antennas and one receive antenna, the transmitted signal is a $M_T \times 1$ vector, $\boldsymbol{s}(t)$. The received signal is:

$$\boldsymbol{y}(t) = \boldsymbol{g}(t)^T \star \boldsymbol{s}(t) + w(t) \tag{2.2}$$

or equivalently

$$y(t) = \begin{bmatrix} g_1(t) & g_2(t) & g_3(t) & \dots & g_{M_R}(t) \end{bmatrix} \star \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \\ \vdots \\ s_{M_T}(t) \end{bmatrix} + w(t),$$

where $g_i(t)$ is the sub-channel response and $w(t)$ is assumed to be white Gaussian noise with variance δ_w^2:

$$w(t) \sim \mathcal{N}(0, \delta_w^2).$$

The signal models of SIMO and MISO systems share some similar properties. For both systems, there are one side with a single antenna and another side with multi-antennas

array. Both system works under Gaussian noise as assumption. These imply a way to unify the signal models for both systems. The following section in fact will apply the channel sounding technique to improve the signal models.

2.2 Channel sounding techniques

To investigate the radio environment of a multi-antenna system, the properties of each sub-channel in the system needs to be examined. Channel sounding techniques are used to decompose the mixed radio channel into orthogonal sub-channels. The time and spatial overlapping sub-channels are allocated within different divisions so that interference between the sub-channels is avoided. Channel sounding is also an efficient way to reduce the cost of the estimation for the channel impulse response.

There are three basic techniques [18]:

1. Time Division Multiplexing (TDM) by allocating different sub-channels at different time slots.
2. Frequency Division Multiplexing (FDM) by allocating different sub-channels at different frequencies.
3. Code Division Multiplexing (CDM) by using distinguishable codewords for different sub-channels.

The channel sounding techniques can be further developed as mixtures such as joint TDM and CDM.

Wyzocki et al [19] provide a practical CDM channel sounding technique by using Walsh-chirp sequence and Prof. Bernard Fleury has described a TDM channel sounding techniques in [9].

In this project, the TDM channel sounding technique is applied.

Consider the MIMO system with M_T transmit antennas and M_R receive antennas. To divide the $M_T \times M_R$ parallel sub-channels in time, a switch is equipped at both the Tx and Rx array, see Fig. 2.1. The switch is only able to select one antenna at one time. The switching scheme divides the time period into time slots of equal length. During any time slot, only one transmit antenna and one receive antenna are selected. In other words, during any time slot, only one sub-channel works in the radio channel. The switching scheme is depicted in Fig. 2.2.

The time period for a thorough traversal of all sub-channels is called one sounding period, denoted as T_s in Fig. 2.2. The switching scheme is repeated in each sounding period. For the TDM channel sounding, the sounding period is selected short enough to maintain the short term stationary of the radio propagation channel. This is to guarantee that characteristics of the radio propagation channel will not vary within one sounding period.

The training sequence, $s(t)$, can be reused for each sub-channel because the TDM mode guarantees separation of sub-channels. Guard intervals (T_g) are inserted between each time slot to avoid Inter Symbol Interference (ISI).

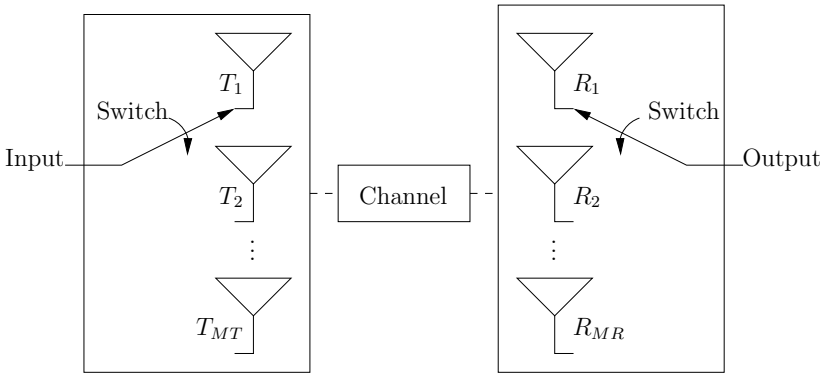


Figure 2.1: A TDM channel sounding technique where only one sub-channel is active at any time.

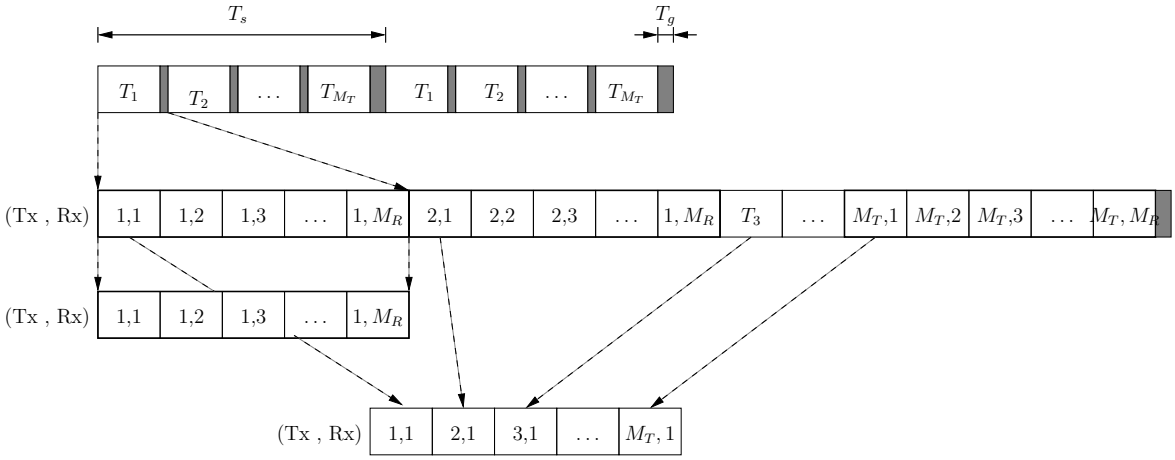


Figure 2.2: Realized switching pattern. The first row shows how T_s and T_g are defined. The switching pattern with T_1 and other transmitters are expanded in the second row. The two depicted selections of sub-channels form a SIMO and a MISO systems respectively.

By using the channel sounding techniques, one can easily extract a SIMO or MISO system from a MIMO system by picking up specific time slots for different sub-channel combinations. Fig. 2.2 gives examples of extracting a SIMO system and a MISO system from the same MIMO system. Since the sub-channels are well separated for both SIMO and MISO system with the help of channel sounding techniques, the signal model of the two systems can be represented by the same formation as given in Eq. (2.1).

The signal structure analogy for the SIMO and MISO system under channel sounding techniques also indicates that the channel characterization is reciprocal between the SIMO and MISO system. Therefore the same estimation method is possible for both systems.

2.3 Radio channel propagation model

Fig. 2.3 provides a general radio channel propagation plot of a SIMO system.

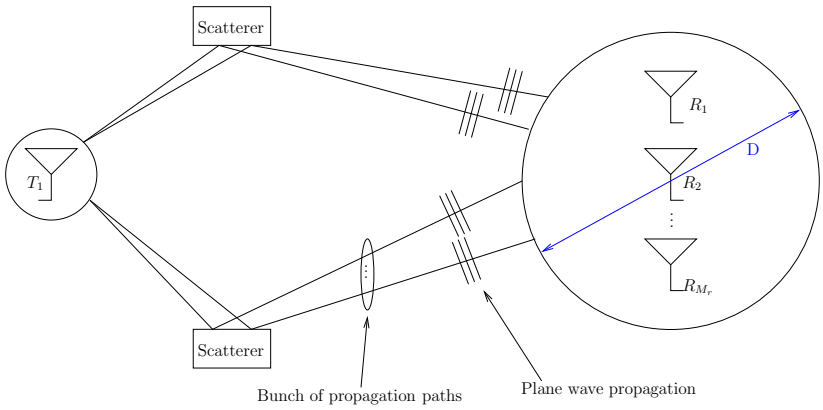


Figure 2.3: A SIMO system where a signal is transmitted by T_1 and altered in amplitude, phase, delay etc. by several scatters. The plane waves (illustrated by the parallel lines) impinge on the Rx array which has a maximum geometric extend of D.

Along the propagation the transmitted signal, $s(t)$, is reflected or diffracted by objects such as walls, trees, vehicles or crowds which cause the change of amplitude and phase of the signal. The signal waves reach the Rx at different delays and directions. The objects which cause the change of the signal are defined as scatterers. The paths the signal follows are defined as propagation path components. The radio channel impulse response is determined by the characteristics of the propagation paths.

The SIMO system is only capable to describe the incident waves at the Rx. The information includes the Direction of Arrival (DoA) and the delay (τ).

The radio propagation channel of a MISO system shares a similar structure with the SIMO system. By exchanging all the Tx information with the Rx information, i.e. by exchanging the labels Tx and Rx and inverting the direction of the signal transmission in Fig. 2.3. Discussion of the following parts will focus on the channel modeling and estimation method for the SIMO systems. All the discussions can be reciprocally used for the analysis of the MISO systems.

The following radio propagation channel model starts from one specular path with a single delay and a single DoA by this assumption, the sampled signal at sounding period n is modified into:

$$\begin{aligned} \boldsymbol{y}_n(t) &= \int_{t_n}^{t_n+T_s} \boldsymbol{g}_n(\tau_i) s(t-\tau_i)\delta(\tau-\tau_i)d\tau_i + \boldsymbol{w}_n(t) \\ &= \boldsymbol{g}_n(\tau) s(t-\tau) + \boldsymbol{w}_n(t) \end{aligned} \tag{2.3}$$

2.3.1 Channel assumption

The channel impulse response introduced in Eq. (2.3) is complex due to the radio propagation channel and the antenna radiation pattern. To set up an effective channel model, several assumptions are introduced.

1. Far field assumption and plane wave propagation

   The far field assumption considers the distance between the Tx and Rx to be larger than the Rayleigh Distance [20, page 220] given by:

   $$r \gg 2D^2/\lambda_{RF} \tag{2.4}$$

   where λ_{RF} is the wavelength and D represents the largest distance among any two antennas in the antenna array (antenna dimension).

   The far field assumption considers the signal wave as a plane wave whose wavefront are infinite parallel planes with constant amplitude normal to the velocity vector of the signal wave. By the assumption of specular path propagation, the general plane wave can be expressed as:

   $$s(\boldsymbol{p}, t) = a \exp\left(j(\boldsymbol{k}^T \boldsymbol{p} - \omega t)\right), \tag{2.5}$$

   where a is the complex amplitude, $\boldsymbol{k}$ is the velocity vector of the wave, $\boldsymbol{p}$ is the propagation vector and $\omega = 2\pi f_{RF}$. One property of plane waves is that the delay of the transmitted signal is linear to the length of the propagation path. From Eq. (2.3), the plane wave propagation form the arriving signal at the Rx side is given:

   $$\boldsymbol{x}(t) = a(\tau, p)s(t - \tau), \tag{2.6}$$

   where a is the complex amplitude of the received signal and it is now a function determined by the specific propagation path p and delay τ.

   Eq. (2.6) is a quite compact expression for the signal which is the basis of further study. Fig. 2.3 also illustrates the characteristics of a plane wave.

2. Small scale assumption

   The arriving signal is further modified by the receive antenna radiation pattern. The radiation pattern of the antenna array modifies the arrival signal in the form of:

   $$\boldsymbol{y}(t) = \boldsymbol{C}a(\tau, p)s(t - \tau), \tag{2.7}$$

   where $\boldsymbol{C}$ is the radiation pattern of the antenna array determined by the direction of the arrived signal and by the position of the antennas. The small scale characteristics of the antenna array assume that the geometry distribution of the receiver antenna array is within small range that the received signal at each receiver antenna is not influenced by the geometric positions of the antennas. As a result, the antenna radiation pattern is only determined by the direction of the arriving signal and the antenna radiation pattern $\boldsymbol{C}$ then becomes a function only determined by the DoA.

   The radiation pattern of the antenna array is represented as:

   $$\boldsymbol{C}(\boldsymbol{\Omega}) = \begin{bmatrix} c_1(\boldsymbol{\Omega}) \\ c_2(\boldsymbol{\Omega}) \\ \vdots \\ c_{M_R}(\boldsymbol{\Omega}) \end{bmatrix},$$

   where $\boldsymbol{\Omega}$ is the DoA. Such vector is called a steering vector with direction $\boldsymbol{\Omega}$.

The visual interpretation of the small scale assumption is shown in Fig. 2.3. The receive antennas at different positions can be equally viewed as in the center of the ball.

Apply the arriving signal model in Eq. (2.6) and assume that the noise is white Gaussian noise, the output of the Rx antenna array in sounding period n is modified into:

$$\boldsymbol{y}(t) = \boldsymbol{C}(\boldsymbol{\Omega})a(\tau, \boldsymbol{\Omega})s(t-\tau) + \boldsymbol{w}_n(t). \tag{2.8}$$

3. Vertical polarization

   In this project, the polarization and the cross polarization is not in the scope of the research. As a default, only vertical polarization is considered since "Vertical polarization is always used in mobile communication" [21, page 9].

4. Doppler frequency

   Doppler frequency is not considered at first in order to obtain a simpler channel model. Therefore this investigation will only focus on environments where the relative speed transmitter and receiver negligible.

When the antenna polarization and Doppler frequency are omitted, the signal model of a SIMO system under the single specular path propagation assumption is simplified to Eq. (2.8) and the complex amplitude $a(\tau, \boldsymbol{\Omega})$ is only a function of τ and $\boldsymbol{\Omega}$.

2.3.2 Dispersive path component model

Eq. (2.8) describes the signal model of one specular path component. The signal model of specular multi-path components can be written as a sum of the individual specular path components. Now suppose that L independent specular path components [21] exist in the channel. Signal $\boldsymbol{y}(t)$ is written as:

$$\boldsymbol{y} = \sum_{l=1}^{L} a_l(\tau_l, \Omega_l)\boldsymbol{C}_l(\Omega)s(t-\tau_l) + \boldsymbol{w}(t). \tag{2.9}$$

In Fig. 2.3, the specular path components is clustered by different scatters. Specular path components coming from the same scatters form a bunch of specular path components. The research is interested on the relationships of specular path components within one bunch. By assuming that the size of the scatterers are small, the specular path components within one bunch is believed to share similar propagation properties in delay and DoA. Consequently, the specular path components are no longer viewed as independent of each other but correlated. To represent this correlation, the complex amplitude is extended as a continuous distributed function in delay (τ) and direction ($\boldsymbol{\Omega}$). Thereby, one scatterer in Eq. (2.9) is replaced with a continuous model:

$$\boldsymbol{y}(t) = \int_{\mathcal{S},\mathcal{T}} a(\Omega, \tau)\boldsymbol{C}(\Omega)s(t-\tau)d\Omega d\tau + \boldsymbol{w}(t), \tag{2.10}$$

where the $\mathcal{S}$ and $\mathcal{T}$ are the integral intervals of $\boldsymbol{\Omega}$ and τ.

Naturally, signals from different scatterers are still considered to be uncorrelated.

The signal model given in Eq. (2.10) describes a continuous distributed path component model. This is a way to parametrically describe the clustering of correlated propagation paths and such will be denoted as dispersive path components. The dispersive path component model is believed more accurate corresponding the true environment.

2.4 Summary

In this chapter, the dispersive path component model was defined. We primarily research on the modeling and estimation of dispersive path components. In the following parts, by no specification the expression "path component" is considered to be dispersive path components.

Part I

Time Variant Scenario

Power Spectrum Modeling in Direction and Delay

3

For the time variant scenario estimation of the path component is only related to the power distribution of $a(\tau, \mathbf{\Omega})$. This chapter provides the modeling of the power distribution of the path components. The models are analogue to probability density functions (pdf). Maximum entropy modeling is introduced. A joint power distribution for direction and delay is proposed and an approximated power distribution is also proposed by using a multivariate Gaussian distribution.

3.1 Problem statement

The power spectrum of the path component can be viewed as the average power distribution of the continuous function, $a(\tau, \mathbf{\Omega})$, if the training sequence $s(t)$ has unit power within one sounding period. For the time variant scenario, the average power distribution is expressed as $E[a(\tau, \mathbf{\Omega})a^*(\tau, \mathbf{\Omega})]$. The power distribution can be modeled in the form of a probability density function by the following facts:

- The entire power for the received signal from one SDS is finite in one sounding period. That is,

$$\int_{t_n}^{t_n+T_s} \int_{\mathcal{S},\mathcal{T}} a_n(\mathbf{\Omega}, \tau) a_n^*(\mathbf{\Omega}, \tau) s(t-\tau) s^*(t-\tau) dt d\tau d\mathbf{\Omega} = P_n \tag{3.1}$$

- The average power distribution $E[a(\tau, \mathbf{\Omega})a^*(\tau, \mathbf{\Omega})]$ is stationary within the observation.

Introduce the power density function $f(\tau, \mathbf{\Omega})$, the expectation $E[a(\mathbf{\Omega}, \tau)a^*(\mathbf{\Omega}, \tau)]$ is replaced by $Pf(\mathbf{\Omega}, \tau)$ where $\int_{\mathcal{S},\mathcal{T}} f(\mathbf{\Omega}, \tau) d\tau d\mathbf{\Omega} = 1$. In this chapter the power density function $f(\mathbf{\Omega}, \tau)$ is derived.

A power distribution across time and space domain is usually complex. Under the condition that no prior information is given, our analysis will follow the maximum entropy modeling. It is necessary to define the support space for the direction and delay variables first.

The support space is the space a variable belongs to. The support space for a real variable is the real space $\mathbb{R}$, e.g. and the support space for a phase variable can be $[0, 2\pi)$.

The description of a direction in 3-D space requires two angles, one for the horizontal angle description and one for the vertical angle description. With polar coordinates, any 3×1 vector can be defined by its radius, the azimuth (ϕ) for the horizontal angle and the elevation (θ) for the vertical angle. We define the range of azimuth as $[-\pi, \pi)$ and the range of elevation

as $[-\pi/2, \pi/2]$. Without the interest of the radius of the vector, define the unit length vector in cartesian coordinates:

$$\mathbf{\Omega} = \begin{bmatrix} \cos(\phi)\cos(\theta) \\ \sin(\phi)\cos(\theta) \\ \sin(\theta) \end{bmatrix} \tag{3.2}$$

as the direction vector. Therefore the support space for $\mathbf{\Omega}$ is $|\mathbf{\Omega}| = 1$ which is a unit sphere.

Fig. 3.1 provides the visual expression for the direction vector.

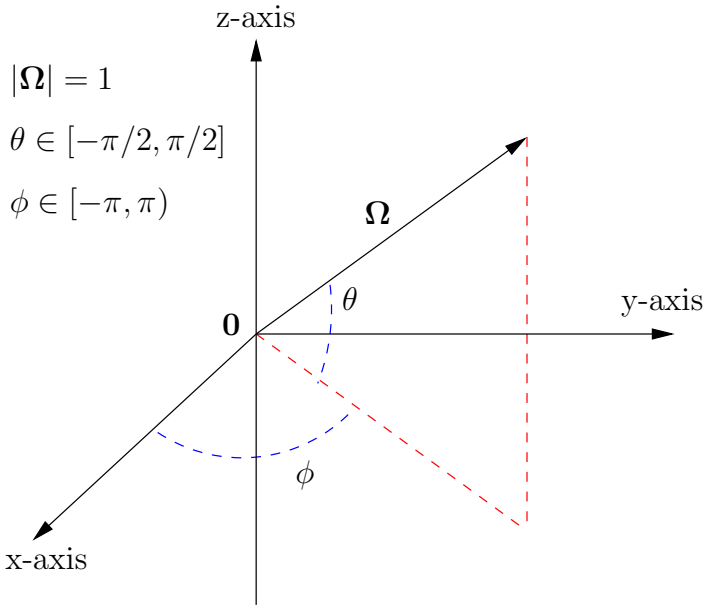


Figure 3.1: Illustration of how the direction vector is obtained.

The range of the delay is normalized by the length of one sounding period. E.g. if for one sounding period, a total 500 samples was obtained, then the the sampling rate is normalized to 500 Hertz.

3.2 Maximum entropy modeling

Maximum entropy modeling can be viewed as a common model giving the maximum likelihood for exponential models when given special observation for the variables. Principles for the construction of maximum entropy modeling are given in [22].

Theorem 1 *For any given pdf $f(\mathbf{x})$ fulfilling the following conditions:*

1. $\mathbf{S}^*$ *is the support of* $f(\mathbf{x})$ *where* $\mathbf{x} \in \mathcal{S}^*$,
2. $E[t_i(\mathbf{x})] = a_i$, $i = 1, \ldots, q$, *where* t_i *is a function of* $\mathbf{x}$ *and* a_i *is a constant value for the expectation.*

Then the entropy of f is maximized, if the pdf f is of the form:

$$f(\mathbf{x}) = \exp\left\{b_0 + \sum_{i=1}^{q} b_i t_i(\mathbf{x})\right\}, \quad \mathbf{x} \in \mathcal{S}^* \tag{3.3}$$

In Eq. (3.3), function $t_i(\mathbf{x})$ can be any kind of elementary function if only $E[t_i(\mathbf{x})]$ is observable. From a statistics view, the most common observation is $E[\boldsymbol{x}]$ and $E[\boldsymbol{x}\boldsymbol{x}^H]$.

One popular maximum entropy description is the Gaussian distribution. The following will try to interpret the Gaussian distribution as a maximum entropy model.

The general description for a real N-variate Gaussian distribution can be written as:

$$f(\boldsymbol{x}) = C \exp\left[-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu_x})^T \boldsymbol{R}(\boldsymbol{x} - \boldsymbol{\mu_x})\right] \tag{3.4}$$

where $\boldsymbol{x} \in \mathbb{R}^N$ is $N \times 1$ random vector , $\boldsymbol{R} = \boldsymbol{\Sigma}^{-1}$ is the $N \times N$ inverse covariance matrix with rank N, $\boldsymbol{\mu_x}$ is the mean of $\boldsymbol{x}$ and C is the normalization coefficient.

Expand the expression inside the exponent part in Eq. (3.4):

$$\begin{aligned} f(\boldsymbol{x}) &= C \exp\left[-\frac{1}{2}(\boldsymbol{x}^T \boldsymbol{R}\boldsymbol{x} - 2\boldsymbol{\mu_x}^T \boldsymbol{R}\boldsymbol{x} + \boldsymbol{\mu_x}^T \boldsymbol{R}\boldsymbol{\mu_x})\right] \\ &= C \exp\left[-\frac{1}{2}\boldsymbol{x}^T \boldsymbol{R}\boldsymbol{x} + \boldsymbol{\mu_x}^T \boldsymbol{R}\boldsymbol{x} - \frac{1}{2}\boldsymbol{\mu_x}^T \boldsymbol{R}\boldsymbol{\mu_x}\right] \\ \text{Let} \quad C_1 &= C \exp\left[-\frac{1}{2}\boldsymbol{\mu_x}^T \boldsymbol{R}\boldsymbol{\mu_x}\right] \\ f(\boldsymbol{x}) &= C_1 \exp\left[-\frac{1}{2}\boldsymbol{x}^T \boldsymbol{R}\boldsymbol{x} + \boldsymbol{\mu_x}^T \boldsymbol{R}\boldsymbol{x}\right] \end{aligned} \tag{3.5}$$

In Eq. (3.5), the general multivariate Gaussian distribution is seen to follow the maximum entropy modeling where $-\frac{1}{2}\boldsymbol{x}^T \boldsymbol{R}\boldsymbol{x}$ corresponds to the observation $E[\boldsymbol{x}\boldsymbol{x}^H]$ and $\boldsymbol{\mu_x}^T \boldsymbol{R}\boldsymbol{x}$ corresponds to the observation $E[\boldsymbol{x}]$.

Furthermore, the full rank matrix $\boldsymbol{\Sigma}$ indicates that it is positive definite, as a result the inverse matrix $\boldsymbol{R}$ is positive definite. The eigenvalues of $\boldsymbol{R}$ represent the power spread along the eigenvectors.

To get a clearer view for the observation of expectation $E[\boldsymbol{x}^T \boldsymbol{R}\boldsymbol{x}]$, Let the matrix $\boldsymbol{R} = \sum_{i=1}^{N} \lambda_i \boldsymbol{b}_i \boldsymbol{b}_i^H$ as in eigenvalue decomposition, thereby convert Eq. (3.5) into:

$$\begin{aligned} f(\boldsymbol{x}) &= C_1 \exp\left[-\frac{1}{2}\boldsymbol{x}^T (\sum_{i=1}^{N} \lambda_i \boldsymbol{b}_i \boldsymbol{b}_i^H)\boldsymbol{x} + \boldsymbol{\mu_x}^T \boldsymbol{R}\boldsymbol{x}\right] \\ &= C_1 \exp\left[-\frac{1}{2}\sum_{i=1}^{N} \lambda_i (\boldsymbol{x}^T \boldsymbol{b}_i)^2 + \boldsymbol{\mu_x}^T \boldsymbol{R}\boldsymbol{x}\right] \end{aligned} \tag{3.6}$$

From Eq. (3.6), it is not hard to see, as the eigenvectors of $\boldsymbol{R}$ forms a basis for the N-dimension space $\mathbb{R}^N$, the expression $\sum_{i=1}^{N} \lambda_i (\boldsymbol{x}^T \boldsymbol{b}_i)^2$ implies that a spread is present in all directions of $\mathbb{R}^N$. Denote this kind of second moment spread as full moment description. For a variable $\boldsymbol{x}$ following the Gaussian distribution with its support as $\boldsymbol{x} \in \mathbb{R}^N$, its variance spread is observable along any direction in the N-dimension space. The condition will be modified if the support of $\boldsymbol{x}$ is a subspace of $\mathbb{R}^N$ and the variance spread is not observable in all directions.

3.3 Fisher Bingham 5 distribution for direction distribution

The maximum entropy modeling for a direction vector leads to the Fisher Bingham 5(FB5) distribution which was originally proposed by Kent in [23]. The FB5 distribution provides a maximum entropy modeling for a 3×1 direction vector $\boldsymbol{\Omega} \quad (|\boldsymbol{\Omega}| = 1)$ and is written as:

$$f(\boldsymbol{\Omega}) = C \exp\left(\kappa \boldsymbol{\gamma}_1^T \boldsymbol{\Omega} + \frac{1}{2}\eta\kappa((\boldsymbol{\gamma}_2^T \boldsymbol{\Omega})^2 - (\boldsymbol{\gamma}_3^T \boldsymbol{\Omega})^2)\right), \tag{3.7}$$

where $\kappa > 0$ is the direction spread coefficient and $\eta \in [0, 1)$ is the ovalness ratio for the angular spread in FB5.

The 3 vectors $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ $\boldsymbol{\gamma}_3$ are orthonormal vectors and $\boldsymbol{\gamma}_1 = \bar{\boldsymbol{\Omega}}$ represents the mean direction of $\boldsymbol{\Omega}$. The vectors $\boldsymbol{\gamma}_2$ and $\boldsymbol{\gamma}_3$ represent the angular power spread direction of $\boldsymbol{\Omega}$. It can be seen that in FB5 distribution, the angular spread of $\boldsymbol{\Omega}$ is orthogonal to its mean direction.

Given that the mean direction $\boldsymbol{\gamma}_1$ is known, the direction of $\boldsymbol{\gamma}_2$ and $\boldsymbol{\gamma}_3$ can be determined by one relative tilt angle on the orthogonal plane of $\boldsymbol{\gamma}_1$. To arrive at such expression, define two direction vectors:

$$\alpha_\phi = \begin{bmatrix} \sin\bar{\phi} \\ -\cos\bar{\phi} \\ 0 \end{bmatrix} \quad \alpha_\theta = \begin{bmatrix} -\sin\bar{\theta}\cos\bar{\phi} \\ -\sin\bar{\theta}\sin\bar{\phi} \\ \cos\bar{\theta} \end{bmatrix}$$

which is the basis for the orthogonal plane of $\boldsymbol{\gamma}_1$. The direction vector $\boldsymbol{\alpha}_\phi$ is in the X-Y plane.

Fig. 3.2 shows the relation among the three direction vector and the tilt angle.

Define the matrix,

$$\boldsymbol{\Gamma} = [\boldsymbol{\gamma}_1 \quad \boldsymbol{\gamma}_2 \quad \boldsymbol{\gamma}_3]$$

as the direction matrix. The analytical expression for the construction of $\boldsymbol{\Gamma}$ is:

$$\boldsymbol{\Gamma} = \begin{bmatrix} \cos\bar{\theta}\cos\bar{\phi} & \sin\bar{\phi} & -\sin\bar{\theta}\cos\bar{\phi} \\ \cos\bar{\theta}\sin\bar{\phi} & -\cos\bar{\phi} & -\sin\bar{\theta}\sin\bar{\phi} \\ \sin\bar{\theta} & 0 & \cos\bar{\theta} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}, \tag{3.8}$$

where $\bar{\theta}$ is the mean elevation, $\bar{\phi}$ is the mean azimuth and α is the direction tilt angle.

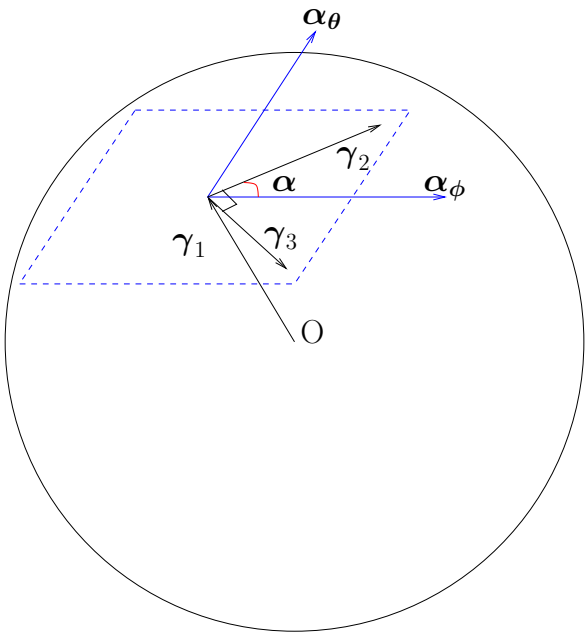


Figure 3.2: The relations between γ_1, γ_2, γ_3 and the tilt angle α

To relate the FB5 distribution to the maximum entropy modeling, the following substitution is performed.

Expand Eq. (3.7):

$$\begin{aligned} f(\mathbf{\Omega}) &= C\exp\left(\kappa\boldsymbol{\gamma}_1^T\mathbf{\Omega} + \frac{1}{2}\eta\kappa((\boldsymbol{\gamma}_2^T\mathbf{\Omega} - \boldsymbol{\gamma}_3^T\mathbf{\Omega})(\boldsymbol{\gamma}_2^T\mathbf{\Omega} + \boldsymbol{\gamma}_3^T\mathbf{\Omega}))\right) \\ &= C\exp\left(\kappa\boldsymbol{\gamma}_1^T\mathbf{\Omega} + \frac{1}{2}\eta\kappa(\boldsymbol{\gamma}_2^T - \boldsymbol{\gamma}_3^T)\mathbf{\Omega}(\boldsymbol{\gamma}_2^T + \boldsymbol{\gamma}_3^T)\mathbf{\Omega}\right) \\ \text{Let} \quad &\sqrt{2}\boldsymbol{\gamma}_2' = \boldsymbol{\gamma}_2 - \boldsymbol{\gamma}_3 \quad \text{and} \sqrt{2}\boldsymbol{\gamma}_3' = \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3 \\ f(\mathbf{\Omega}) &= C\exp\left(\kappa\boldsymbol{\gamma}_1^T\mathbf{\Omega} + \eta\kappa(\boldsymbol{\gamma}'^T_2\mathbf{\Omega}\boldsymbol{\gamma}'^T_3\mathbf{\Omega})\right) \end{aligned} \tag{3.9}$$

Fig. 3.3 shows that the γ_2', γ_3' and γ_1 are still orthonormal to each other. FB5 can then be redefined by applying γ_1, the new $\gamma_2 = \gamma_2'$ and the new $\gamma_3 = \gamma_3'$.

$$f(\mathbf{\Omega}) = C\exp(\kappa\boldsymbol{\gamma}_1^T\mathbf{\Omega} + \zeta\kappa\boldsymbol{\gamma}_2^T\mathbf{\Omega}\boldsymbol{\gamma}_3^T\mathbf{\Omega}) \tag{3.10}$$

To get a better understanding of the modified FB5 distribution as a maximum entropy model, the following transition is applied.

For the direction vector $\mathbf{\Omega}$, apply a linear transformation by using the direction matrix $\mathbf{\Gamma}$:

$$\mathbf{\Omega}' = \mathbf{\Omega}^T\mathbf{\Gamma} = \begin{pmatrix} \mathbf{\Omega}^T\boldsymbol{\gamma}_1 \\ \mathbf{\Omega}^T\boldsymbol{\gamma}_2 \\ \mathbf{\Omega}^T\boldsymbol{\gamma}_3 \end{pmatrix}$$

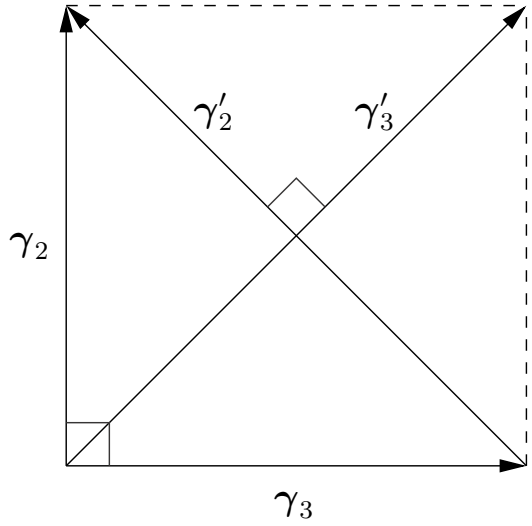


Figure 3.3: Relation between γ_2, γ_3, γ_2' and γ_3'

Introduce the $\boldsymbol{R}$ matrix for random vector $\boldsymbol{\Omega}$ as it is introduced in Gaussian distribution in Eq. (3.6).

$$\boldsymbol{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

The matrix $\boldsymbol{R}$ is a symmetric real matrix where each entry r_{ij} represents the observation of $E[\boldsymbol{\Omega}^T\boldsymbol{\gamma}_i\boldsymbol{\Omega}^T\boldsymbol{\gamma}_j]$. Now rewrite the Eq. (3.10) with the introduction of $\boldsymbol{R}$.

$$\begin{aligned} f(\boldsymbol{\Omega}') &= C\exp(\kappa\boldsymbol{\Omega}'^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta\boldsymbol{\Omega}'^T\boldsymbol{R}\boldsymbol{\Omega}') \\ &= C\exp(\kappa\boldsymbol{\Omega}'^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta\boldsymbol{\Omega}'^T \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \boldsymbol{\Omega}') \end{aligned} \tag{3.11}$$

By comparing Eq. (3.11) and Eq. (3.10), it can be seen that for FB5 distribution entries of $\boldsymbol{R}$ are zeroes except for $r_{23} = r_{32} = 1/2$. Comparing Eq. (3.11) with Eq. (3.6), it is easy to see that the FB5 distribution fulfills the maximum entropy modeling with an confined observation of $E[\boldsymbol{\Omega}\boldsymbol{\Omega}^T]$. To be more precise, the FB5 distribution is based on the observable mean direction $E[\boldsymbol{\Omega}]$ and the observable cross covariance $E[\boldsymbol{\gamma}_2^T\boldsymbol{\Omega}\boldsymbol{\gamma}_3^T\boldsymbol{\Omega}]$ which is orthogonal to its mean direction. We say the covariance observation of $E[\boldsymbol{\Omega}\boldsymbol{\Omega}^T]$ for FB5 is confined on the sphere because the support of the direction $\boldsymbol{\Omega}$ is on the unit sphere and thereby the angular spread outside or inside the sphere is not observable.

3.4 Truncated Gaussian distribution for delay distribution

The support for the delay is $\tau \in \mathbb{R}$. With TDM channel sounding technique the delay of the signal is distinguishable within one sounding period. The maximum entropy model of the delay is proposed to be a truncated Gaussian distribution within one sounding period. The spread for the delay has to be small enough for the truncation.

The truncated Gaussian distribution is given as below:

$$f(\tau) = C \exp(-B(\tau - \mu_\tau)^2) \tag{3.12}$$

$B = \frac{1}{2\sigma^2} > 0$, σ^2 is the variance of the delay. B is denoted as delay spread coefficient. and μ_τ is the mean of the delay. The range for the delay is confined in one sounding period defined as $\tau \in [0, 1)$ and $\mu_\tau \in [0, 1)$.

3.5 Joint power distribution in direction and delay

The FB5 distribution and truncated Gaussian distribution is viewed as conditional power distribution in direction and delay respectively. By considering the joint power distribution in direction ($\mathbf{\Omega}$) and delay (τ), the conditional power distribution for this joint power distribution still follows the FB5 and truncated Gaussian in direction and delay respectively. Meanwhile, the maximum entropy principles has to be satisfied again.

By analysis previously the multivariates Gaussian distribution is a general maximum entropy model which is able to describe the first moment and second moment of a variable. Because the description for the mean and the variance for a variable is enough to give a maximum entropy model to describe how the power of the variable is distributed, it is sensible to start the joint power distribution in the form of a Gaussian distribution. The FB5 distribution can also be viewed as a confined Gaussian distribution when its support is limited to $\mathbf{\Omega} \in \mathbb{R}^3, |\mathbf{\Omega}| = 1$. For the joint power distribution, there are three free variables: the azimuth, the elevation and the delay. By comparison of the 3-variate Gaussian distribution to the joint power distribution in direction and delay the following requirements need to be fulfilled.

1. The conditional distribution of the direction and delay should follow the FB5 and truncated Gaussian respectively.

2. For a 3-variate Gaussian distribution, the parameters which describe it can be categorized as: 3 parameters for the mean of the variables, 6 variables for the variance matrix of the variables. Accordingly one can expect the same number of parameters to describe the joint power distribution. For the direction description, the parameters are: 2 parameters for the mean direction ($\bar{\phi}$, $\bar{\theta}$), 3 parameters for the power spread (κ, ζ, α). For the delay description, the parameters are: 1 parameter for the mean delay($\bar{\tau}$), 1 parameter for the power spread ($B = \frac{1}{2\sigma^2}$). The cross correlation of the power spread in direction and delay needs 2 extra parameters to describe it. Finally, the number of parameters to describe the joint power distribution has to be equal to the number of parameters to describe a 3 variables Gaussian distribution.

Now construct a 4×1 random vector as a stack of a direction vector and a delay variable:

$$\begin{bmatrix} \boldsymbol{\Omega} \\ \tau \end{bmatrix}$$

The mean for this random vector is:

$$E\begin{bmatrix} \boldsymbol{\Omega} \\ \tau \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{\Omega}} \\ \bar{\tau} \end{bmatrix}$$

The derivation for the joint power distribution starts in an expression similar to Gaussian distribution:

$$f(\boldsymbol{\Omega}, \tau) = C \exp\left(\begin{bmatrix} \boldsymbol{\Omega} - \bar{\boldsymbol{\Omega}} \\ \tau - \bar{\tau} \end{bmatrix}^T \boldsymbol{R} \begin{bmatrix} \boldsymbol{\Omega} - \bar{\boldsymbol{\Omega}} \\ \tau - \bar{\tau} \end{bmatrix} \right) \tag{3.13}$$

The matrix $\boldsymbol{R}$ is 4×4 used to describe the covariance of the direction and delay. The exponential maximum entropy is determined by the components formation inside the exponent. Let $q(\boldsymbol{\Omega}, \tau)$ represent the expression inside the exponent of $f(\boldsymbol{\Omega}, \tau)$.

$$q(\boldsymbol{\Omega}, \tau) = \begin{bmatrix} \boldsymbol{\Omega} - \bar{\boldsymbol{\Omega}} \\ \tau - \bar{\tau} \end{bmatrix}^T \boldsymbol{R} \begin{bmatrix} \boldsymbol{\Omega} - \bar{\boldsymbol{\Omega}} \\ \tau - \bar{\tau} \end{bmatrix}$$

Reform $\boldsymbol{R}$ in the following:

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{c} \\ \boldsymbol{c}^T & -b \end{bmatrix},$$

where $\boldsymbol{A}$ is a 3×3 matrix describing the covariance of the direction, b is a scalar spread coefficient of the delay and $\boldsymbol{c}$ is a 3×1 vector describing the cross covariance between the direction and delay. This decomposition divides $\boldsymbol{R}$ into 3 parts, $\boldsymbol{A}$ and b for the individual covariance descriptions of $\boldsymbol{\Omega}$ and τ, $\boldsymbol{c}$ for the cross covariance description for $\boldsymbol{\Omega}$ and τ.

We investigate the conditional direction distribution at certain delay τ_0. The expression is given below.

$$q(\boldsymbol{\Omega}|\tau_0) = \begin{bmatrix} \boldsymbol{\Omega} - \bar{\boldsymbol{\Omega}} \\ \tau_0 - \bar{\tau} \end{bmatrix}^T \begin{bmatrix} \boldsymbol{A} & \boldsymbol{c} \\ \boldsymbol{c}^T & -b \end{bmatrix} \begin{bmatrix} \boldsymbol{\Omega} - \bar{\boldsymbol{\Omega}} \\ \tau_0 - \bar{\tau} \end{bmatrix}$$

$$\begin{aligned} q(\boldsymbol{\Omega}|\tau_0) &= (\boldsymbol{\Omega} - \bar{\boldsymbol{\Omega}})^T \boldsymbol{A}(\boldsymbol{\Omega} - \bar{\boldsymbol{\Omega}}) + 2\boldsymbol{c}^T(\boldsymbol{\Omega} - \bar{\boldsymbol{\Omega}})(\tau_0 - \bar{\tau}) - b(\tau_0 - \bar{\tau})^2 \\ &= \boldsymbol{\Omega}^T \boldsymbol{A}\boldsymbol{\Omega} - 2\bar{\boldsymbol{\Omega}}^T \boldsymbol{A}\boldsymbol{\Omega} + \bar{\boldsymbol{\Omega}}^T \boldsymbol{A}\bar{\boldsymbol{\Omega}} + 2\boldsymbol{c}^T\boldsymbol{\Omega}(\tau_0 - \bar{\tau}) - 2\boldsymbol{c}^T\bar{\boldsymbol{\Omega}}(\tau_0 - \bar{\tau}) - b(\tau_0 - \bar{\tau})^2 \\ &= \boldsymbol{\Omega}^T \boldsymbol{A}\boldsymbol{\Omega} + 2(-\bar{\boldsymbol{\Omega}}^T \boldsymbol{A} + \boldsymbol{c}^T(\tau_0 - \bar{\tau}))\boldsymbol{\Omega} + C_0 \\ &\left(C_0 = \bar{\boldsymbol{\Omega}}^T \boldsymbol{A}\bar{\boldsymbol{\Omega}} - 2\boldsymbol{c}^T\bar{\boldsymbol{\Omega}}(\tau_0 - \bar{\tau}) - b(\tau_0 - \bar{\tau})^2\right) \end{aligned} \tag{3.14}$$

If the conditional direction distribution still follows the FB5 distribution, by comparing Eq. (3.14) and Eq. (3.10), the conditional mean direction $\bar{\mathbf{\Omega}}|_{\tau_0}$ and the conditional direction spread coefficient $\kappa|_{\tau_0}$ from Eq. (3.14) follow:

$$\kappa|_{\tau_0}\bar{\mathbf{\Omega}}|_{\tau_0} = \kappa\bar{\mathbf{\Omega}}^T - 2\boldsymbol{c}^T(\tau_0 - \bar{\tau}), \tag{3.15}$$

where κ and $\bar{\mathbf{\Omega}}$ are the global direction spread coefficient and the mean direction respectively.

From Eq. (3.15) it can be seen that the conditional mean direction moves along the direction of $\boldsymbol{c}$ and the conditional $\kappa|_{\tau_0}$ is determined by both κ and $\boldsymbol{c}(\tau_0 - \bar{\tau})$. Eq. (3.14) and Eq. (3.15) also indicate that the component $\mathbf{\Omega}^T\boldsymbol{A}\mathbf{\Omega}$ is not affected by the given delay τ_0. However, to maintain the FB5 distribution for conditional direction distribution, the $\mathbf{\Omega}^T\boldsymbol{A}\mathbf{\Omega}$ needs to be modified.

To reveal this relation, firstly assume that the azimuth, elevation and delay follow the joint Gaussian distribution. Fig. 3.4 (a), (b) and (c) provide the conditional Gaussian distribution of the direction given different delays on the plane. It can be seen that the mean direction of the conditional Gaussian distribution is located along a line with the global mean direction locates in the center of the line. One can also observe that the power spread ovalness ratio and angular spread direction at different delays stay the same. By extending these properties to the conditional direction FB5 distribution in Eq. (3.15), it can be seen that the ovalness radio ζ and the direction tilt angle α stay constant in different conditional direction distributions. As a result, the component $\mathbf{\Omega}^T\boldsymbol{A}\mathbf{\Omega}$ is required to be consistent with the component $\zeta\kappa|\tau_0\boldsymbol{\gamma}_2^T|\mathbf{\Omega}\boldsymbol{\gamma}_3^T\mathbf{\Omega}$ given in Eq. (3.10) for different delays, where the conditional angular spread direction is determined by α and $\kappa|_{\tau_0}$.

Fig. 3.4 (d), (e) and (f) provide the analogue relation for the conditional direction distribution which follows the FB5 distribution. The conditional mean direction moves along an arch on the sphere and the ovalness radio for the angular spread of the conditional direction distribution stay the same at different delays.

To maintain the number of parameters, two parameters are used to describe vector $\boldsymbol{c}$. Normally to determine the 3×1 vector $\boldsymbol{c}$, 3 parameters are needed (radius, azimuth and elevation). Now confine the vector $\boldsymbol{c}$ orthogonal to the mean direction $\bar{\mathbf{\Omega}}$, then the direction of $\boldsymbol{c}$ is determined by one relative tilt angle to describe the position of $\boldsymbol{c}$ on the orthogonal plane of $\bar{\mathbf{\Omega}}$. Denote the delay tilt angle for $\boldsymbol{c}$ as $\boldsymbol{\beta}$ and decompose the $\boldsymbol{c}$ into the amplitude and direction parts, the vector $\boldsymbol{c}$ is modified into:

$$\boldsymbol{c} = \eta \begin{bmatrix} \sin\bar{\phi}\cos\beta - \sin\bar{\theta}\cos\bar{\phi}\sin\beta \\ -\cos\bar{\phi} - \sin\bar{\theta}\sin\bar{\phi}\sin\beta \\ \cos\bar{\theta}\sin\beta \end{bmatrix} \tag{3.16}$$

The general joint direction delay distribution is:

$$f(\mathbf{\Omega}, \tau) = C\exp\left((\kappa\bar{\mathbf{\Omega}} - \eta\boldsymbol{c}(\tau - \bar{\tau}))^T\mathbf{\Omega} + \mathbf{\Omega}^T\boldsymbol{A}\mathbf{\Omega} - 2\boldsymbol{c}^T\bar{\mathbf{\Omega}}(\tau - \bar{\tau}) - b(\tau - \bar{\tau})^2\right) \tag{3.17}$$

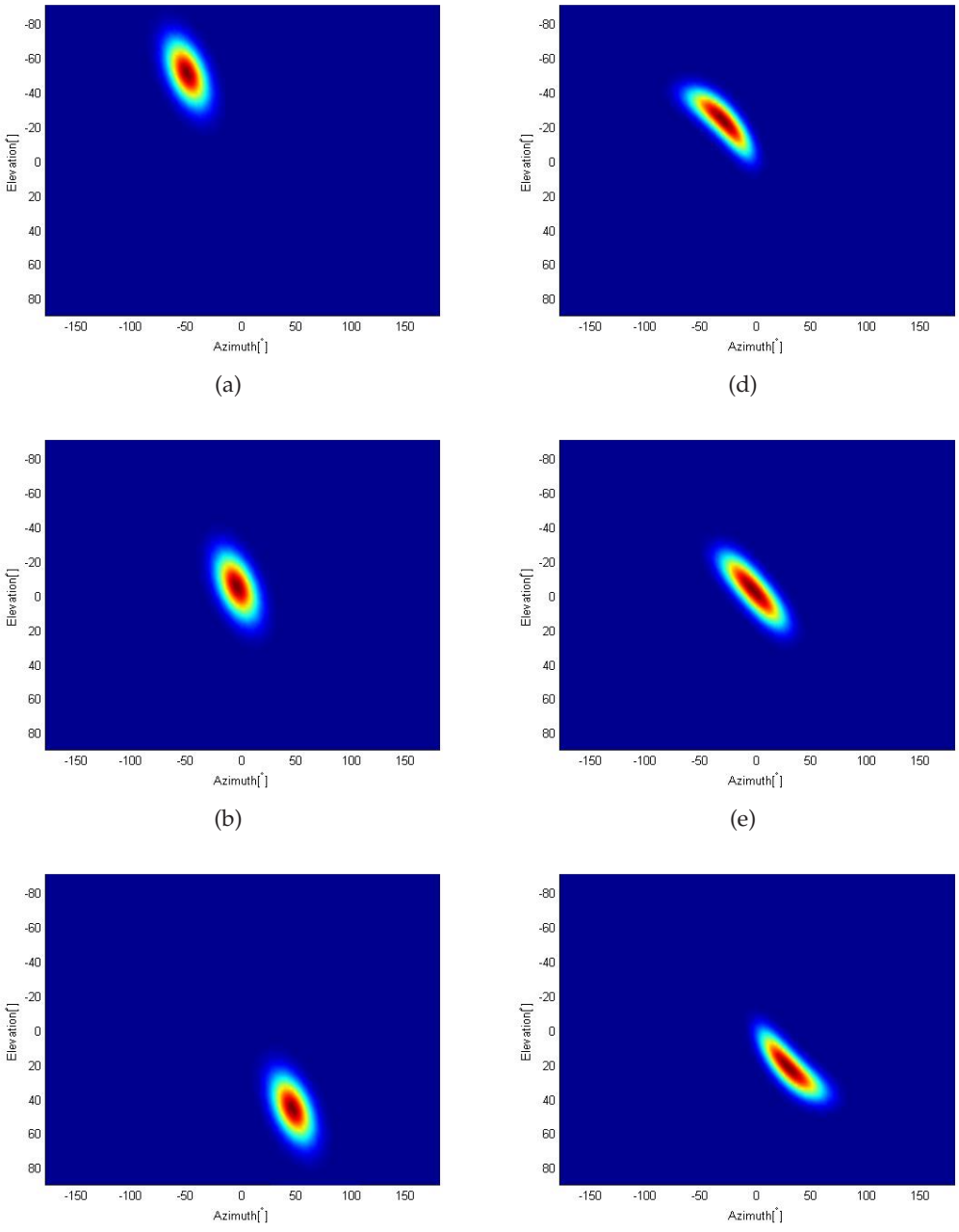


Figure 3.4: A comparison of the conditional Gaussian distribution (a, b, c) and the conditional direction distribution of the joint FB5 and Gaussian distribution (d, e, f)

The parameters inside Eq. (3.17) is explained in the following:

- C is the normalization coefficient.
- κ is the direction spread coefficient.
- $\bar{\mathbf{\Omega}}$ is the mean direction.
- η is the cross direction delay spread coefficient.
- $\boldsymbol{c}$ is the cross direction-delay spread direction by Eq. (3.16)
- $\boldsymbol{A}$ is determined by the conditional $\kappa|_{\tau_0}$, $\mathbf{\Omega}|_{\tau_0}$, the ovalness ratio of direction distribution ζ and direction tilt angle α.
- b is the delay spread coefficient.
- $\bar{\tau}$ is the mean of the delay.

The practical procedure to construct the joint pdf of direction and delay is given following.

1. For a given delay τ_0, calculate the vector $\mathbf{\Omega}_0' = \kappa\bar{\mathbf{\Omega}} - \eta\boldsymbol{c}(\tau_0 - \bar{\tau})$.
2. Let $\kappa_0 = |\mathbf{\Omega}_0'|$ be the conditional κ and $\bar{\mathbf{\Omega}}_0 = \mathbf{\Omega}_0'/\kappa_0$ be the mean direction for the conditional FB5 direction distribution.
3. Use $\kappa_0, \bar{\mathbf{\Omega}}_0$ and α to calculate the conditional angular spread direction $\boldsymbol{\gamma}_{20}$ and $\boldsymbol{\gamma}_{30}$ by Eq. (3.8).
4. Construct the matrix $\boldsymbol{A}$. $\boldsymbol{A} = \kappa_0\zeta\boldsymbol{\gamma}_{20}\boldsymbol{\gamma}_{30}^T$
5. Calculate $f(\mathbf{\Omega}, \tau_0)$ as following FB5 distribution, with condition κ_0, $\mathbf{\Omega}_0$, $\boldsymbol{\gamma_{20}}$, $\boldsymbol{\gamma_{30}}$, α and ζ.
6. Combine all the conditional direction distribution to form the joint direction delay distribution $f(\Omega, \tau)$.
7. Normalize $f(\Omega, \tau)$.

3.6 Power distribution approximations

Kent's article mentioned one property of the FB5 distribution that when the spread of the FB5 distribution is pretty small, it is close to a Gaussian distribution on the plane. From this point, it is possible to approximate the joint power distribution in direction and delay with the Gaussian model when the spread is small enough.

Recall the real 3-variate Gaussian distribution in the form of:

$$f(x, y, z) = C \exp\left(-\frac{1}{2}\begin{bmatrix} x-\mu_x \\ y-\mu_y \\ z-\mu_z \end{bmatrix}^T \mathbf{\Sigma}^{-1} \begin{bmatrix} x-\mu_x \\ y-\mu_y \\ z-\mu_z \end{bmatrix}\right) \tag{3.18}$$

where $\mathbf{\Sigma}$ is the covariance matrix.

Now define $\Delta_x = x - \mu_x$. To approximate the joint power distribution in direction and delay, the problem is to find a effective approximation to replace the

$$\begin{bmatrix} \mathbf{\Omega} - \bar{\mathbf{\Omega}} \\ \tau - \bar{\tau} \end{bmatrix}$$

with

$$\begin{bmatrix} \Delta_\phi \\ \Delta_\theta \\ \tau - \bar{\tau} \end{bmatrix}.$$

To find the approximation Δ_ϕ and Δ_θ, the X-Y plane is wrapped onto a sphere. Consider the arch connecting $\mathbf{\Omega}$ and $\bar{\mathbf{\Omega}}$ as the difference of the two direction vector on the sphere. This leads to an analogue formation of a difference between two points on the X-Y plane. Fig. 3.5(a) shows the visual formation of Δ_x and Δ_y on the plane. Applying the same method, forms of Δ_θ and Δ_ϕ on the sphere is given in Fig. 3.5(b) as two arches. Note that Δ_θ and Δ_ϕ are not related with the cartesian coordinate of the space. They are only related with the relative position of $\bar{\mathbf{\Omega}}$, α_θ and α_ϕ.

The computation of Δ_ϕ is described in Fig. 3.5(c). The following are the procedures:

1. Let $\mathbf{\Omega}_\phi = \mathbf{\Omega} - \frac{\mathbf{\Omega}^T \alpha_\theta}{|\mathbf{\Omega}||\alpha_\theta|}$ as the projection of $\mathbf{\Omega}$ onto the $\bar{\mathbf{\Omega}}$, α_ϕ plane.
2. Let $|\Delta_\phi|$ represents the absolute difference arch on the sphere between $\mathbf{\Omega}_\phi$ and $\bar{\mathbf{\Omega}}$.

$$|\Delta_\phi| = \left| \arccos \frac{\mathbf{\Omega}_\phi^T \bar{\mathbf{\Omega}}}{|\mathbf{\Omega}_\phi||\bar{\mathbf{\Omega}}|} \right|$$

3. Let $\text{sgn}(\mathbf{\Omega}_\phi^T \alpha_\phi)$ as the sign of Δ_ϕ.
4. Finally Δ_ϕ is in the form of :

$$\Delta_\phi = \text{sgn}(\mathbf{\Omega}_\phi^T \alpha_\phi) \left| \arccos \frac{\mathbf{\Omega}_\phi^T \bar{\mathbf{\Omega}}}{|\mathbf{\Omega}_\phi||\bar{\mathbf{\Omega}}|} \right|$$

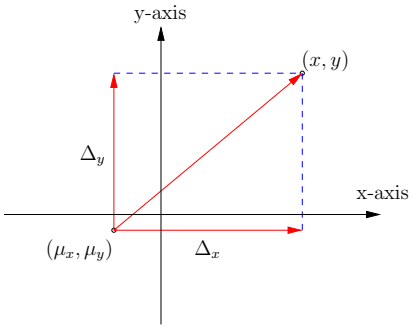
In the similar way Δ_θ can be obtained.

A simpler approximation is to directly apply the joint Gaussian distribution for:

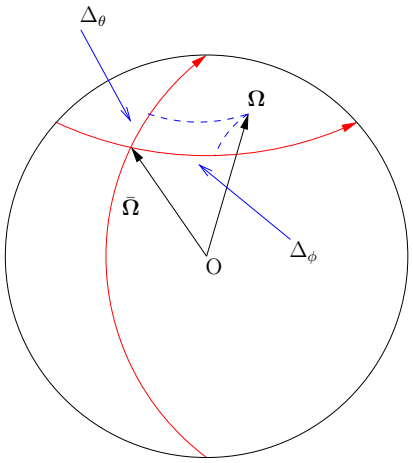
$$\begin{bmatrix} \Delta_\phi \\ \Delta_\theta \\ \Delta_\tau \end{bmatrix} = \begin{bmatrix} \phi - \bar{\phi} \\ \theta - \bar{\theta} \\ \tau - \bar{\tau} \end{bmatrix}$$

This approximation works within the following limitations:

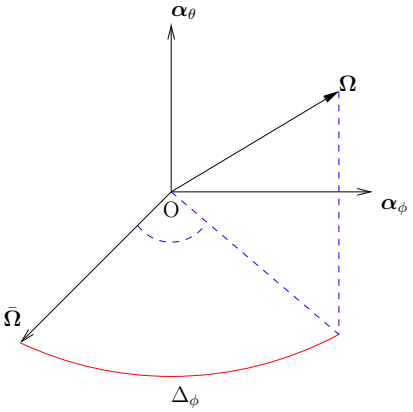
- The spread of the distribution is small
- The distribution is far away from the boundaries of the plane($\phi = \pm\pi$ and $\theta = \pm\pi/2$).
- The distribution is not close to the top or the bottom ($\theta = \pm\pi/2$) where the Jaccobi modification will cause the distortion of the distribution.



(a) The formation of Δ_x and Δ_y on the X-Y plane.



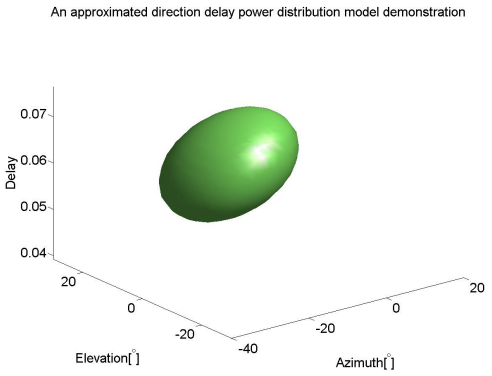
(b) The formation of Δ_θ and Δ_ϕ on the sphere.



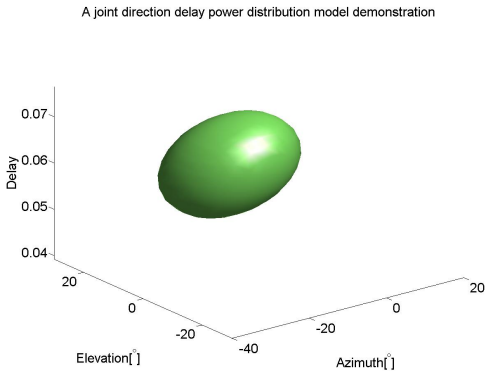
(c) The computation of Δ_ϕ

3.7 Comparison of power distribution models

Until now, there are two kinds of power distribution in direction and delay at hand. One is the joint direction delay power distribution applying the conditional FB5 and Gaussian. The other one is the approximated power distribution by using Gaussian distribution. To make a comparison, Fig. 3.7 provides a demonstration of the two different model in Fig. 3.5(d) and Fig. 3.5(e). The settings for the joint power distribution and the approximated power distribution are selected with the same mean position ($\bar{\theta}$, $\bar{\phi}$, $\bar{\tau}$). The power spread and the power spread direction are defined close between the two models. It can be seen that the both power distribution model are able to describe the spread and spread direction along direction and delay.



(d) A isosurface plot of the joint power distribution in direction and delay



(e) A isosurface plot of the approximated power distribution in direction and delay

Parameter settings	
$\bar{\theta}[^\circ]$	0
$\bar{\phi}[^\circ]$	0
$\bar{\tau}$(index)	30
κ	100
ζ	0.01
$\alpha[^\circ]$	60
$\beta[^\circ]$	270
η	1000
B	40000

(f) Parameter settings of model in (a)

Parameter settings	
$\bar{\theta}[^\circ]$	0
$\bar{\phi}[^\circ]$	0
$\bar{\tau}$(index)	30
δ_θ	0.11
δ_ϕ	0.12
δ_τ	0.004
$\rho_{\phi\theta}$	0.2
$\rho_{\phi\tau}$	0.4
$\rho_{\theta\tau}$	-0.1

(g) Parameter settings of model in (b)

Figure 3.5: Plots of the joint power distribution model and the approximated power distribution model with the tabularized parameter settings.

Radio Propagation Channel Estimation

4

A time variant radio channel means that the channel impulse response is a random process in each sounding period. Stochastic characteristics of the channel impulse response is discussed at the beginning of this chapter. The SAGE is introduced for multi-path component estimation. Estimators such as the MLE and the LSE are applied in SAGE. The analysis is only provided for the SIMO systems. All the methods can be reused for the MISO system.

4.1 Channel impulse response estimation with multi-path components

The signal model for multi-path components is the sum of all the individual path components. For L path components, the signal model is:

$$\boldsymbol{y}(t) = \sum_{l=1}^{L} \int_{\mathcal{S},T} a_l(\boldsymbol{\Omega}, \tau) \boldsymbol{C}(\boldsymbol{\Omega}) s(t-\tau) d\boldsymbol{\Omega} d\tau + \boldsymbol{w}(t). \tag{4.1}$$

In the time variant scenario, the WSSUS model [6] is assumed for the path components expressed as:

$$E[a(\boldsymbol{\Omega}_i, \tau_i) a^*(\boldsymbol{\Omega}_j, \tau_j)] = 0, \quad i \neq j \tag{4.2}$$

Furthermore the radio propagation channel model assumes that different path components are assumed to be uncorrelated:

$$E[a_i(\boldsymbol{\Omega}, \tau) a_j^*(\boldsymbol{\Omega}, \tau)] = 0, \quad i \neq j \tag{4.3}$$

The WSSUS assumption will lead to a simple expression for the stochastic characteristics of the channel impulse response.

The correlation of the output signal $\boldsymbol{y}(t)$ with $s^*(t-\tau')$ provides an estimate of the channel impulse response at the sampled delay τ'. Denote the estimated channel impulse response at delay τ' as $\boldsymbol{h}(\tau')$.

$$\begin{aligned}\boldsymbol{h}(\tau') &= \int_0^T \boldsymbol{y}(t)s^*(t-\tau')dt \\ &= \int_{\mathcal{S},\mathcal{T}} \sum_{l=1}^{L} a_l(\boldsymbol{\Omega},\tau)\boldsymbol{C}(\boldsymbol{\Omega}) \int_{t_n}^{t_n+T} s(t-\tau)s^*(t-\tau')dtd\boldsymbol{\Omega}d\tau + \int_{t_n}^{t_n+T} \boldsymbol{w}(t)s^*(t-\tau')dt \\ &= \sum_{l=1}^{L} \int_{\mathcal{S},\mathcal{T}} a_l(\boldsymbol{\Omega},\tau)\boldsymbol{C}(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau')d\boldsymbol{\Omega}d\tau + \int_{t_n}^{t_n+T} \boldsymbol{w}(t)s^*(t-\tau')dt, \end{aligned} \tag{4.4}$$

where $\mathrm{Rss}(\tau)$ is the autocorrelation function of signal $s(t)$, i.e. $\mathrm{Rss}(\tau) = \int_{t_n}^{t_n+T} s(t)s^*(t-\tau)$.

The component $\int_{t_n}^{t_n+T} \boldsymbol{w}(t)s^*(t-\tau')dt$ is additive Gaussian noise.

The sampled vector $\boldsymbol{h}(\tau')$ forms a complete estimation for the channel impulse response within one sounding period. Now stack the sampled channel impulse response estimation denoted as $\boldsymbol{H}$. The stacked channel impulse response estimation vector is given as:

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{h}(\tau_1) \\ \\ \boldsymbol{h}(\tau_2) \\ \vdots \\ \boldsymbol{h}(\tau_k) \\ \vdots \\ \boldsymbol{h}(\tau_K) \end{bmatrix} \tag{4.5}$$

where τ_k is the kth delay sample of all the K delay samples within one sounding period.

4.2 Stochastic analysis

In the time variant scenario the complex amplitude $a_l(\boldsymbol{\Omega},\tau)$ is a random process across different sounding periods. The $a_l(\boldsymbol{\Omega},\tau)$ is assumed to have uniform phase distribution in $[0, 2\pi)$. This uniform phase distribution makes the complex amplitude $a(\boldsymbol{\Omega},\tau)$ have zero mean. The average power distribution of $a_l(\boldsymbol{\Omega},\tau)$ follows the models given in Chapter 3 as $E[a_l(\boldsymbol{\Omega},\tau)a_l^*(\boldsymbol{\Omega},\tau)] = P_l f_l(\boldsymbol{\Omega},\tau)$. By Eq. (4.3) it is assumed that different path components are uncorrelated. Taking all the above into the investigation, the stochastic characteristics of $\boldsymbol{H}$ is studied.

By the central limit theorem, the random $\boldsymbol{H}$ vector given in Eq. (4.5) follows the Gaussian distribution. The mean and covariance matrix of $\boldsymbol{H}$ is derived in the following.

The mean of $\boldsymbol{H}$ is given by:

$$E[\boldsymbol{H}] = \begin{bmatrix} E[\boldsymbol{h}(\tau_1)] \\ E[\boldsymbol{h}(\tau_2)] \\ \vdots \\ E[\boldsymbol{h}(\tau_k)] \\ \vdots \\ E[\boldsymbol{h}(\tau_K)] \end{bmatrix}, \tag{4.6}$$

and derivation in A.1 shows that $E[\boldsymbol{H}] = \boldsymbol{0}$.

The covariance of $\boldsymbol{H}$ is given by:

$$E[\boldsymbol{H}\boldsymbol{H}^H] = \begin{bmatrix} E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_1)] & E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_2)] & \dots & & E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_K)] \\ \vdots & & \ddots & & \vdots \\ & & & E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)] & \\ E[\boldsymbol{h}(\tau_K)\boldsymbol{h}^H(\tau_1)] & & & & E[\boldsymbol{h}(\tau_K)\boldsymbol{h}^H(\tau_K)] \end{bmatrix} \tag{4.7}$$

where each sub-matrix $E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)]$ has the expression given in Appendix A.1.

$$E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)] = \sum_{l=1}^{L} \int_{\mathcal{S},T} P_l f_l(\boldsymbol{\Omega}, \tau) \boldsymbol{C}(\boldsymbol{\Omega}) \boldsymbol{C}^H(\boldsymbol{\Omega}) \mathrm{Rss}(\tau - \tau_i) \mathrm{Rss}(\tau - \tau_j) d\boldsymbol{\Omega} d\tau + \sigma_n^2 \boldsymbol{I} \mathrm{Rss}(\tau_i - \tau_j), \tag{4.8}$$

where P_l and $f_l(\boldsymbol{\Omega}, \tau)$ are the power and power density function for the lth path component.

The covariance sub-matrix in Eq. (4.8) can be decomposed into signal parts and a noise part. Let:

$$\begin{aligned} \Sigma_{\boldsymbol{s}_l \boldsymbol{s}_l} &= \int_{\mathcal{S},T} P_l f_l(\boldsymbol{\Omega}, \tau) \boldsymbol{C}(\boldsymbol{\Omega}) \boldsymbol{C}^H(\boldsymbol{\Omega}) \mathrm{Rss}(\tau - \tau_i) \mathrm{Rss}(\tau - \tau_j) d\boldsymbol{\Omega} d\tau \\ \Sigma_{\boldsymbol{w}\boldsymbol{w}} &= \sigma_n^2 \boldsymbol{I} \mathrm{Rss}(\tau_i - \tau_j). \end{aligned} \tag{4.9}$$

Then,

$$E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)] = \sum_{l=1}^{L} \Sigma_{\boldsymbol{s}_l \boldsymbol{s}_l} + \Sigma_{\boldsymbol{w}\boldsymbol{w}}. \tag{4.10}$$

The $\Sigma_{\boldsymbol{s}_l \boldsymbol{s}_l}$ represents the signal part in the covariance matrix and $\Sigma_{\boldsymbol{w}\boldsymbol{w}}$ is the noise part. Due to the uncorrelated multi-path components, the stochastic characteristics of $\boldsymbol{H}$ can be

viewed as the summation of the stochastic characteristics of the individual path components.

For the time variant scenario, the stochastic characteristics of the channel impulse response estimation is only determined by the power distribution of each path component. The phase distribution is not considered.

Whether the sub-matrix $E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)]$ in Eq. (A.4) is a zero-matrix or nonzero depends only on Rss(τ) and the distance between τ_i and τ_j when applying the WSSUS assumption. The Fig. 4.1 gives a example of the covariance matrix by assuming Rss(τ) $= \delta(\tau)$. It can be observed that only the sub-matrices on the diagonal of the covariance matrix are nonzero matrices.

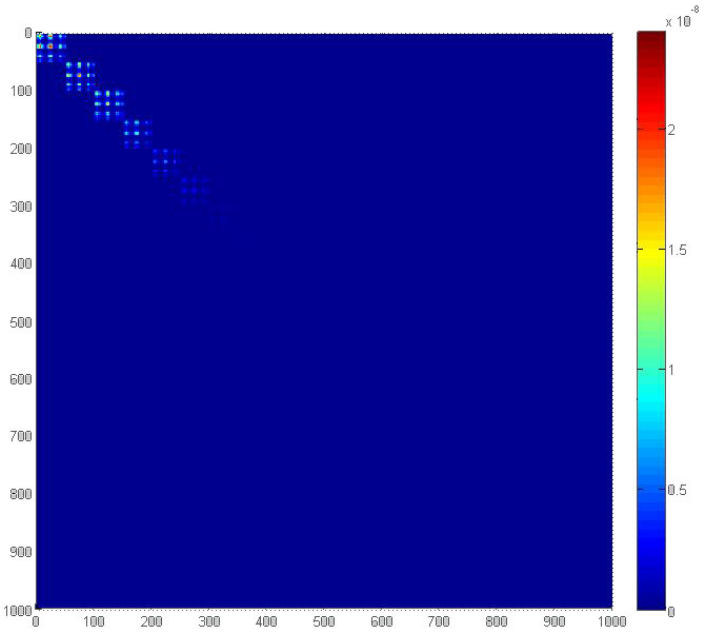


Figure 4.1: A illustration of the covariance matrix when $Rss(\tau) = \delta(\tau)$. The color in the figure represents the absolute amplitudes of the entries. From Eq. (4.8) it is known that a sub-matrix is nonzero only when $\tau_i = \tau_j$.

4.3 The SAGE

The radio propagation channel estimation in the time variant scenario is equivalent to the estimation of power distribution of the propagation path components. When the estimation expands to multiple path component, one can use the EM method to get a joint estimation for multi-path components [15] where for each estimation iteration the power distributions for all path components are estimated simultaneously. There are two drawbacks of the EM method: slow convergence and difficult maximization steps due to coupling when smoothness penalties are used. As an alternative the SAGE algorithm [24] is applied which estimate parameters in smaller groups sequentially by alternating between several small hidden data spaces defined by the algorithm designer.

4.3.1 General structure of the SAGE

In the multi-path component problem, the signal model is given in Eq. (4.1). The individual signal model for each path component can be viewed as the individual signal corrupted by additive Gaussian noise. The received signal from the lth path component can be written as:

$$\boldsymbol{x}_l(t) = \int_{S,T} a_l(\Omega, \tau; \boldsymbol{\theta}_l)\boldsymbol{C}(\Omega)s(t-\tau)d\Omega d\tau + \sqrt{\beta}\,\boldsymbol{w}(t) \qquad \beta \in [0,1], \tag{4.11}$$

where $\boldsymbol{\theta}_l$ represents the parameter set for the lth path component. The collection of all parameter sets is represented by:

$$\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \ldots, \boldsymbol{\theta}_L]$$

Joint estimation for the whole parameter set $\boldsymbol{\theta}$ is impractical due to the computation complexity and convergence rate. The key idea of SAGE is to design a well defined data subspace that is only related with one parameter sub set, $\boldsymbol{\theta}_k$. By analytical expression, for each estimation iteration, the SAGE replace:

$$\Lambda(\boldsymbol{\theta}_1^m, \boldsymbol{\theta}_2^m, \ldots, \boldsymbol{\theta}_L^m; \boldsymbol{\theta}_1^{m-1}, \boldsymbol{\theta}_2^{m-1}, \ldots, \boldsymbol{\theta}_L^{m-1})$$

with

$$\Lambda_l(\boldsymbol{\theta}_l^m; \boldsymbol{\theta}_1^{m-1}, \boldsymbol{\theta}_2^{m-1}, \ldots, \boldsymbol{\theta}_L^{m-1}),$$

where Λ represents the estimator of the joint parameters estimation and Λ_l represents the estimator of the lth parameter.

In short, SAGE works only in a data subspace extracted from the total received signal $\boldsymbol{y}(t)$. The $\boldsymbol{y}(t)$ is observable data but the internal composition given in Eq. (4.1) is unknown. The signal model for each path component is known as given in Eq. (4.11) but is observable from the receiver. The $\boldsymbol{y}(t)$ is called observable incomplete data and the $\boldsymbol{x}_i(t)$ is called complete unobservable data or hidden data if each path component is uncorrelated. The realization of SAGE relies on the hidden data for each path component. This is achieved by introducing the estimation for the hidden data based on previous estimation results.

Fig. 4.2 provides a general structure for the SAGE algorithm. The SAGE can be divided into two main steps, the Expectation step and the Maximization step. For each estimation iteration, one path component is selected and the conditional expectation of the likelihood of the parameter set of the path component is computed by previous estimation results. After the estimation is done, the algorithm checks whether the convergence of the estimation is achieved. If not, the algorithm updates the parameter subset with the estimation of current path component and continues to estimate the next path component.

4.3.2 Expectation step in the SAGE

The original definition of hidden data space is given in [24]:

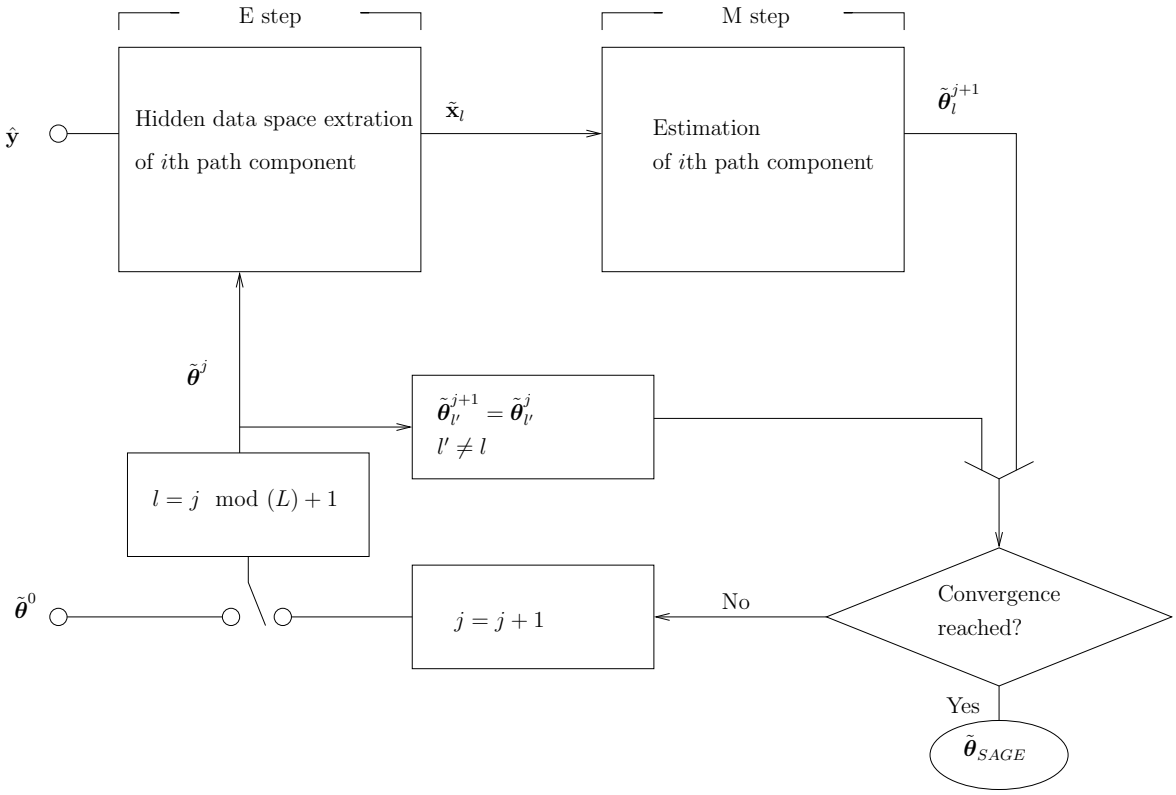


Figure 4.2: The general SAGE structure. In the structure, L path components are taken into the estimation. The switching of the path component into one estimation iteration is determined by the $l = j \mod (L)$ where j is the iteration counter. The SAGE will stop if the required convergence is reached.

Definition 1 A random vector $\boldsymbol{X}^S$ with probability density function $f(\boldsymbol{x};\boldsymbol{\theta})$ is an admissible hidden data space with respect to $\boldsymbol{\theta}_S$ for $f(\boldsymbol{y};\boldsymbol{\theta})$ when the joint density of $\boldsymbol{X}^S$ and $\boldsymbol{Y}$ satisfies:

$$f(\boldsymbol{y},\boldsymbol{x};\boldsymbol{\theta}) = f(\boldsymbol{y}|\boldsymbol{x};\boldsymbol{\theta}_{\tilde{S}})f(\boldsymbol{x};\boldsymbol{\theta}), \tag{4.12}$$

where $(\boldsymbol{\theta}_{\tilde{S}},\boldsymbol{\theta}_S) = \boldsymbol{\theta}$, $\boldsymbol{\theta}_{\tilde{S}}$ contains the rest of the parameter sets $\boldsymbol{\theta}$ when $\boldsymbol{\theta}_S$ is excluded. Eq. (4.12) indicates that the conditional distribution $f(\boldsymbol{y}|\boldsymbol{x};\boldsymbol{\theta}_{\tilde{S}})$ must be independent of $\boldsymbol{\theta}_S$. In other words, $\boldsymbol{X}^S$ must be a complete data space for $\boldsymbol{\theta}_S$ given that $\boldsymbol{\theta}_{\tilde{S}}$ is known.

One way to estimate the hidden data space $\boldsymbol{x}_l$ is to apply the conditional expectation expressed as:

$$\tilde{\boldsymbol{x}}_l(t) = E[\boldsymbol{x}_l(t)|\boldsymbol{y}(t) = \hat{\boldsymbol{y}};\tilde{\boldsymbol{\theta}}]. \tag{4.13}$$

The vector $\hat{\boldsymbol{y}}$ represents the observation of the data and $\tilde{\boldsymbol{\theta}}$ represents the estimated parameter set.

For the time variant scenario, the covariance matrix is taken into estimation. The hidden data is therefore in the form of covariance matrix for each path component as derived in

Eq. (4.8). Let:

$$\Sigma_{\boldsymbol{H}} = \sum_{l=1}^{L} \Sigma_{\boldsymbol{s}_l \boldsymbol{s}_l} + \Sigma_{\boldsymbol{ww}} \tag{4.14}$$

represents the incomplete data for the multi-path components. The hidden data for the lth path component becomes:

$$\Sigma_{\boldsymbol{H}_l} = \Sigma_{\boldsymbol{s}_l \boldsymbol{s}_l} + \beta \Sigma_{\boldsymbol{ww}} \quad \beta \in [0, 1], \tag{4.15}$$

where $\Sigma_{\boldsymbol{s}_l \boldsymbol{s}_l}$ and $\Sigma_{\boldsymbol{ww}}$ is given in Eq. (4.9).

Denote the $\boldsymbol{H}_l^j$ as the estimation of $\boldsymbol{H}_l$ from the jth estimation iteration, $\boldsymbol{H}^j$ as the estimation of $\boldsymbol{H}$ from the jth estimation iteration. By central limit theorem, both $\boldsymbol{H}_i$ and $\boldsymbol{H}$ follows Gaussian distribution. The conditional expectation then turns to the conditional covariance matrix of two correlated Gaussian random vector. From Appendix A.3, the conditional expectation of the hidden data space for $\boldsymbol{H}_l^j$ at the iteration j reads:

$$\Sigma_{\boldsymbol{H}_l^j} = \Sigma_{\boldsymbol{H}_l^{j-1}} - \Sigma_{\boldsymbol{H}_l^{j-1}} \Sigma^{-1}_{\boldsymbol{H}^{j-1}} \Sigma_{\boldsymbol{H}_l^{j-1}} + \Sigma_{\boldsymbol{H}_l^{j-1}} \Sigma^{-1}_{\boldsymbol{H}^{j-1}} \overline{\hat{\boldsymbol{H}} \hat{\boldsymbol{H}}^H} \Sigma^{-1}_{\boldsymbol{H}^{j-1}} \Sigma_{\boldsymbol{H}_l^{j-1}} \tag{4.16}$$

where $\overline{\hat{\boldsymbol{H}} \hat{\boldsymbol{H}}^H} = \frac{1}{N} \sum_{n=1}^{N} \hat{\boldsymbol{H}}_n \hat{\boldsymbol{H}}_n^H$ is the observed covariance matrix of the data.

4.3.3 Maximization step in the SAGE

The Maximization step in SAGE is the estimation of the individual path component parameter set based on the hidden data expectation. There are two kinds of estimators available: the Maximum Likelihood Estimator (MLE) and the Least Square Estimator (LSE).

MLE

The general principles for MLE is given in Appendix A.4. For a random variable x following the probability density function $f(x; \boldsymbol{\theta})$, the MLE for the parameter set $\boldsymbol{\theta}$ is expressed as:

$$\tilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \prod_{i=1}^{n} f(x_i; \boldsymbol{\theta}) \tag{4.17}$$

Further, the logarithm of the MLE turns it into a Maximum Loglikelihood Estimator.

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \ln \left(\prod_{i=1}^{n} f(x_i; \boldsymbol{\theta}) \right) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \sum_{i=1}^{n} \ln f(x_i; \boldsymbol{\theta}) \tag{4.18}$$

By the central limit theorem the channel impulse response is zero Gaussian random vector. A.5 provides the MLE in the time variant scenario using Eq. (4.18).

$$\tilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmax}} \quad \left(- \ln |\tilde{\boldsymbol{\Sigma}}(\boldsymbol{\theta})| - trace(\tilde{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}) \right) \tag{4.19}$$

The $\tilde{\boldsymbol{\Sigma}}(\boldsymbol{\theta})$ is the estimated hidden data for the path component. The $\hat{\boldsymbol{\Sigma}}$ is the expectation of the hidden data from previous estimation results.

LSE

Least Square Error (LSE) estimation can be viewed as a fitting operation between the estimation and the measurement. The estimator is given in a simple expression.

$$\tilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \left(||\hat{\boldsymbol{\Sigma}} - \tilde{\boldsymbol{\Sigma}}|| \right) \tag{4.20}$$

where the $|| \cdot ||$ denote the norm operation.

The LSE estimator is an approximation to the MLE and because of its low computational complexity it may be preferably when the complexity of other estimators are too high.

4.3.4 Performance analysis

The reason to apply the SAGE is its lower complexity and fast convergence. Suppose the computation complexity for one parameter set is $O(f(n))$, then the computation complexity for joint estimation of K parameter sets is $O(f(n)^K)$. The SAGE, on the other hand, lower the computation complexity into $O(Kf(n))$.

The SAGE is an extension of the EM algorithm and an approximation towards the global maximum likelihood. The structure of SAGE does not guarantee that it reaches the global maximum likelihood. In fact for most cases the SAGE is limited around a local optimal estimation result. The performance of SAGE relies heavily on the initialization step. A good initialization can guarantee the estimation results of SAGE closer to global maximum likelihood.

The hidden data space introduced in Eq. (4.15) consists of two parts, the signal $\Sigma_{\boldsymbol{s}_l\boldsymbol{s}_l}$ and the noise $\beta\Sigma_{\boldsymbol{ww}}$. The β in Eq. (4.15) is a controllable parameter when extracting the hidden data space. the The coefficient β determines how much noise and estimation errors that are introduced into the hidden data. The expected hidden data is then taken into the Maximization step to reduce the introduced estimation errors. In theory, the introduced estimation errors will be decreased along with iteration. The SAGE can also be viewed as to reduce the estimation errors with the estimation iteration. This explains the effect of β. As β get closer to 0, less estimation errors are introduced and the convergence of SAGE get slower but the performance of the estimation will not fluctuate too much. As β is closer to 1, more estimation errors are introduced and the convergence of SAGE get faster but the performance of the estimation may change quite a lot due to e.g. model mismatch.

Fig. 4.3 provides a visual interpretation of the constitution of the hidden data space, it can be seen that the selection of β is also a important factor for the performance of SAGE.

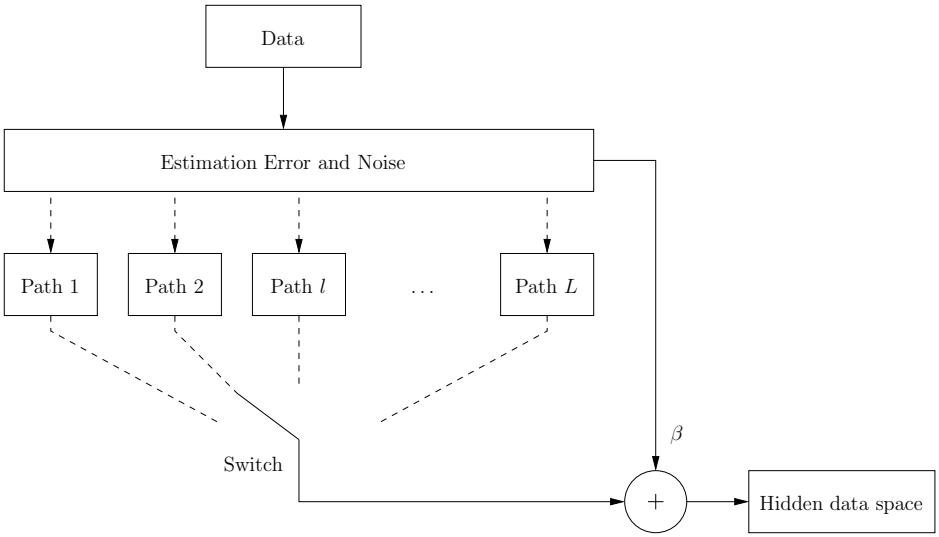


Figure 4.3: The constitution of hidden data space.

5 Algorithm Design

This chapter describes the specific algorithm design for the radio propagation channel estimation in the time variant scenario. The overall structure of the algorithm is described at the beginning. The initialization design for the SAGE is a critical part in the algorithm design and several initialization strategies are offered and compared. The data structure and the flow chart of the algorithm are also presented.

5.1 General structure

Fig. 5.1 shows the basic steps of the algorithm, where the SAGE estimates path components. Generally, the algorithm can be divided into 4 block: Data preparation, initialization for the estimation, the SAGE estimator and saving and displaying the estimation resuls.

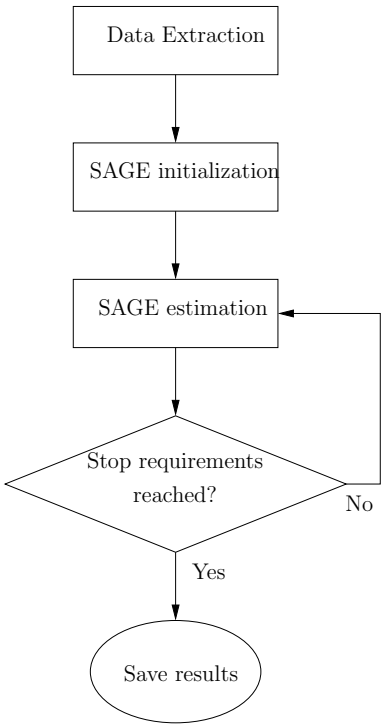


Figure 5.1: The general framework of the algorithm

5.2 Data structures

Basically there are 4 kinds of data that are used in the algorithm. They are:

- The information register for the algorithm and measurement data settings which includes:
  - The directory of where the measurement data and the corresponding antenna response should be extracted.
  - Which scenario is going to be estimated.
  - Whether a SIMO or a MISO is taken into estimation.
  - How many sounding period samples are taken into estimation
  - How the extracted measurement data and antenna response data should be further truncated for the purpose of estimation.
  - Whether time variant scenario or time invariant scenario is selected.
  - Which power distribution model that is selected.
  - The upper boundary of the number of path components in the estimation
  - Which initialization strategy the SAGE applies.
  - The estimation control of the SAGE including the beta value, the maximum iteration number, etc.
- The measurement data structure which includes the following information:
  - Statistical information of the measurement data such as the mean and the covariance matrix.
  - The index for the selected delay samples which is taken into estimation
  - The autocorrelation function $Rss(\tau)$ for signal $s(t)$.
  - The power delay profile for each sub channel.
- The antenna response structure which include the information:
  - The calibrated antenna response.
  - The minimal antenna response.
  - The antenna resolution.
  - The delay range in index.
  - The delay resolution.
- The parameter container which is used to store the parameter subset for each path component.

5.3 Data preparation

The measurement data extraction is done by channel sounding techniques.

The data needs to be truncated to alleviate the work load of the estimator. The algorithm selects the delay samples with the most power. The rest of the delay samples are considered as noise. The truncation of delay samples will not deteriorate the performance of the estimation because it only excludes part of the noise.

The antenna can be partially selected to alleviate the estimation work load. The antenna truncation will somehow reduce the resolution of the estimation thereby more error are likely to be introduced.

From Eq. (4.8) the noise can be estimated from the average power from the power delay profile where signals are believed not to be present. Therefore, the noise is estimated during the data preparation step.

5.4 Initialization

The initialization design is crucial for the performance of SAGE. A good initialization should resemble the true environment. In the power distribution estimation problem beamforming techniques are proposed as a first estimation for the power spectrum in the initialization step.

5.4.1 Power spectrum analysis

The beamformer is a filter bank approach to convert the power spectrum from eigenmode into angular view.

The methods is similar to the standard Fourier transform. Denotes the filter representation of beamformer as: $\boldsymbol{a}(\boldsymbol{\Omega})$. For a certain delay τ_0, the power spectrum in angular view is:

$$p(\boldsymbol{\Omega}, \tau_0) = \boldsymbol{a}(\boldsymbol{\Omega})^H \hat{\boldsymbol{\Sigma}}(\tau_0) \boldsymbol{a}(\boldsymbol{\Omega}),$$

where $\hat{\boldsymbol{\Sigma}}(\tau_0) = \frac{1}{N} \boldsymbol{y}_n \boldsymbol{y}_n^H$ is the data covariance matrix.

The ideal power spectrum in angular view is able to remove the antenna response effect as:

$$p(\boldsymbol{\Omega}, \tau) = Pf(\boldsymbol{\Omega}, \tau)$$

However, because the antenna array has a finite number of antennas the power spectrum will be blurred by sidelobes similar to a finite length Fourier transform. Such limitation will reduce the resolution of the power spectrum. Currently, there are several kinds of beamformers that can be used for the estimation of the power spectrum.

Bartlett beamforming

Bartlett beamforming tries to maximize the power in a specific direction [25].

For the time variant scenario, the Bartlett spectrum at delay τ_k is given in the form of :

$$p(\mathbf{\Omega}, \tau_k) = \frac{\boldsymbol{C}(\mathbf{\Omega})^H \hat{\mathbf{\Sigma}}(\tau_k) \boldsymbol{C}(\mathbf{\Omega})}{\boldsymbol{C}(\mathbf{\Omega})^H \boldsymbol{C}(\mathbf{\Omega})} \tag{5.1}$$

The $\hat{\mathbf{\Sigma}}(\tau_k) = \frac{1}{N} \sum_{n=1}^{N} \hat{\boldsymbol{h}}(\tau_k)_n \hat{\boldsymbol{h}}(\tau_k)_n^H$ is the measurement covariance matrix at delay τ_k.

Eq. (5.1) is derived from the condition that:

$$\max \boldsymbol{a}(\mathbf{\Omega})^H \hat{\mathbf{\Sigma}}(\tau_k) \boldsymbol{a}(\mathbf{\Omega})$$

In eigenmode view, decompose the $\hat{\mathbf{\Sigma}}(\tau_k)$ as:

$$\hat{\mathbf{\Sigma}}(\tau_k) = \sum_{i=1}^{N} \lambda_i \boldsymbol{n} \boldsymbol{b}_i \boldsymbol{b}_i^H.$$

Bartlett beamforming can be viewed as the summation of the eigenvalue projections of the covariance matrix on the specific direction $\mathbf{\Omega}$.

The number of the antennas determine the resolution of Bartlett spectrum. As more antennas are included both along the elevation and azimuth, the resolution of Bartlett beamforming is improved and the sidelobes from Bartlett beamforming decreases. Fig. 5.2 illustrates this effect.

Capon beamforming

Capon beamforming tries to maintain the power gain level for a specific direction while it minimizes the power from other directions [25]. Capon beamforming contains several different expression for different scenario. The study so far investigates three kinds of Capon beamforming, the Capon I, the Capon II, and the robust Capon beamforming.

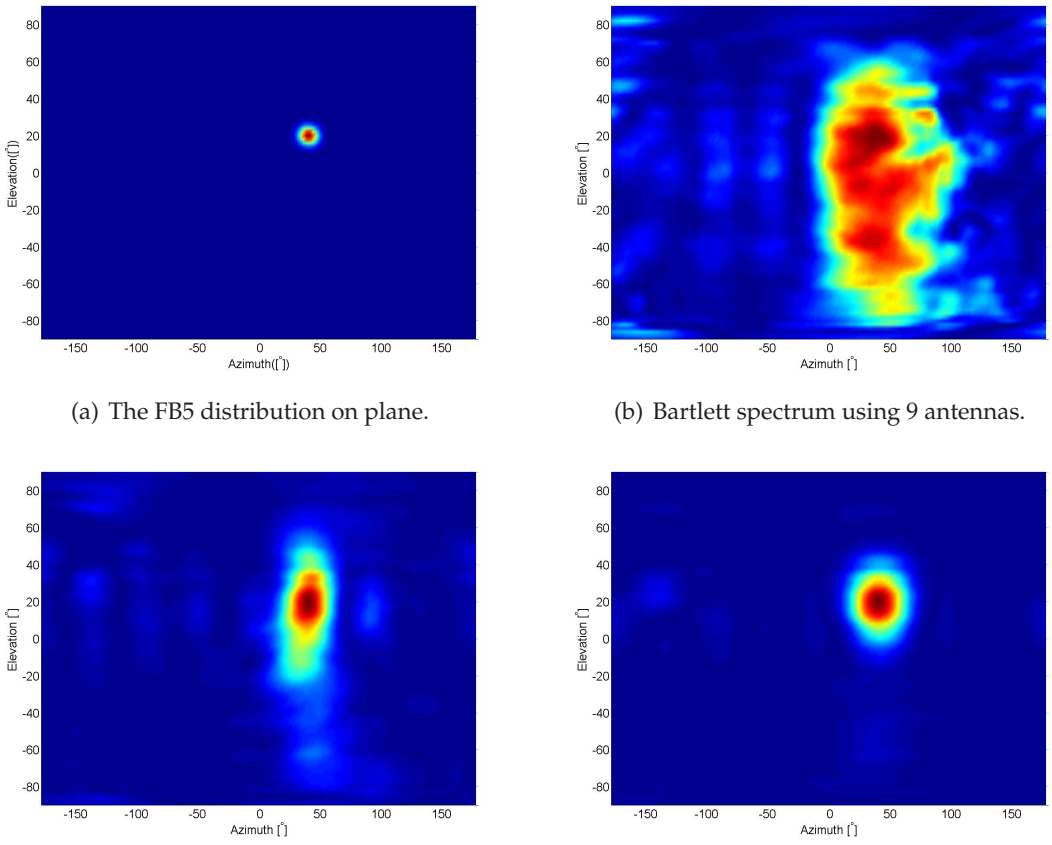
For the time variant scenario, Capon I spectrum for at delay τ_k is then in the form of:

$$p(\mathbf{\Omega}, \tau_k) = \frac{1}{\boldsymbol{C}(\mathbf{\Omega})^H \hat{\mathbf{\Sigma}}(\tau_k)^{-1} \boldsymbol{C}(\mathbf{\Omega})} \tag{5.2}$$

It is derived from the condition that

$$\min \boldsymbol{a}^H \hat{\mathbf{\Sigma}}(\tau_k) \boldsymbol{a} \quad \text{while} \quad \boldsymbol{a}^H \boldsymbol{C}(\mathbf{\Omega}) = 1$$

The Capon I spectrum is sensitive to the antenna response because it is not normalized to the antenna response $\boldsymbol{C}(\mathbf{\Omega})$. The effect is dominant where the $|\boldsymbol{C}(\mathbf{\Omega})|$ is quite small.



(a) The FB5 distribution on plane.

(b) Bartlett spectrum using 9 antennas.

(c) Bartlett spectrum using 32 antennas.

(d) Bartlett spectrum using 50 antennas.

Figure 5.2: The Bartlett power spectrums for the FB5 distribution given in (a) with different number of antennas. Color of the figures represent the linear power level. It is seen that the resolution of the Bartlett spectrum increases as more antennas are included.

The Capon II spectrum is able to normalize the power spectrum. The Capon II has a general expression as given :

$$p(\boldsymbol{\Omega}) = \frac{\boldsymbol{C}(\boldsymbol{\Omega})^H \hat{\boldsymbol{\Sigma}}(\tau_k)^{-(i-1)} \boldsymbol{C}(\boldsymbol{\Omega})}{\boldsymbol{C}(\boldsymbol{\Omega})^H \hat{\boldsymbol{\Sigma}}(\tau_k)^{-i} \boldsymbol{C}(\boldsymbol{\Omega})} \tag{5.3}$$

where i is a positive integer $i \geq 1$.

as $i = 1$, Capon II spectrum can be viewed similar to a Bartlett spectrum where it is based on the projection of the eigenvalues to the specific direction.

The robust Capon spectrum is introduced in [26]. The robust Capon spectrum is applied when the the calibration of the antenna response $\boldsymbol{C}(\boldsymbol{\Omega})$ is imprecise. The robust Capon spectrum is only able to improve the spectrum when the antenna response is not precisely calibrated.

All kinds of the above Capon spectrums call for the inverse operations for matrix $\hat{\boldsymbol{\Sigma}}(\tau_k)$. This requires that the samples for the measurement data should be no less than the dimension

of the sub-covariance matrix $\frac{1}{N}\sum_{n=1}^{N}\boldsymbol{y}_n\boldsymbol{y}_n^H$ otherwise it will cause the problem of singular matrix inversion.

In our algorithm the Bartlett spectrum is selected a trustable power spectrum estimator given the condition that the resolution of the antenna is high enough and the effect of side-lobes are small.

The construction of the power spectrum along the delay and direction can be viewed as stacking of the individual power spectrum for each delay sample. This provides an easy way to construct the direction and delay power spectrum.

5.4.2 Initialization strategy

The general principle for the initialization of SAGE is to search the peaks in the Bartlett power spectrum as the central of the path components. After the mean values ($\bar{\tau}$ and $\bar{\boldsymbol{\Omega}}$) and the power of one path component are initialized, parameters to describe the power distribution of each path component are initialized according to different strategies described below.

Power initialization

The power P initialization of path components can be estimated from the absolute values of the peaks in Bartlett spectrum with some modifications. Consider a single specular path propagation, the covariance matrix at the peak delay is:

$$\boldsymbol{\Sigma}(\tau_0) = P\boldsymbol{C}(\boldsymbol{\Omega}_0)\boldsymbol{C}(\boldsymbol{\Omega}_0)^H$$

The Bartlett spectrum for the direction $\boldsymbol{\Omega}_0$ is :

$$p(\boldsymbol{\Omega},\tau_0) = \frac{\boldsymbol{C}(\boldsymbol{\Omega}_0)^H P\boldsymbol{C}(\boldsymbol{\Omega}_0)\boldsymbol{C}(\boldsymbol{\Omega}_0)^H\boldsymbol{C}(\boldsymbol{\Omega}_0)}{\boldsymbol{C}(\boldsymbol{\Omega}_0)\boldsymbol{C}^H(\boldsymbol{\Omega}_0)}$$

$$p(\boldsymbol{\Omega},\tau_0) = P\boldsymbol{C}(\boldsymbol{\Omega}_0)^H\boldsymbol{C}(\boldsymbol{\Omega}_0)$$

Therefore, the estimated power using the Bartlett spectrum is:

$$\tilde{P} = \left|\frac{p(\boldsymbol{\Omega},\tau_0)}{\boldsymbol{C}(\boldsymbol{\Omega}_0)^H\boldsymbol{C}(\boldsymbol{\Omega}_0)}\right| \tag{5.4}$$

Note that the power initialization given in Eq. (5.4) is based on the specular path. As the estimation is carried, the optimal power estimation will be different for a distributed path component. Experience showed that the performance of SAGE gets stable when the power estimation has reached a local maximum likelihood.

Power spread initialization

After the initialization of the power and the mean position of the path components, other parameters for each path component are initialized. These parameters describe the power distribution in delay and direction. One can design different procedures using different criterias. These approaches are categorized into two groups: blind initialization and estimated initialization.

Blind initialization Blind initialization stands for providing the parameters with default values. This method avoids to give each path component specific parameter initialization. However, such default initialization may cause a lot of estimation errors for the initialization step and will affect the performance of the later SAGE step. One safe design of the blind initialization is to initialize each path component with quite small power spread. Currently there are two approaches as the candidates of blind initialization.

1. The default initialization which sets the path component as specular path component at the beginning. This initialization is believed to be safe because any kind of path component distribution can be viewed as the summation of large numbers of specular path components.
2. The initialization which sets the path component as having quite small power spread. This strategy can be viewed as extension of the first one while the later one has a easier transition for specular path initialization to the distributed path estimation.

Estimated initialization Since blind initialization method may introduce unnecessary errors, an improved strategy can be realized by give a specific initialization of the power spread for each path component. This calls for the estimation at the initialization step.

The following describe several candidates for the estimation strategy.

1. Parallel path estimation using SAGE: One can apply the normal SAGE to parallel estimate the multi-path components. To differentiate the SAGE in initialization step and in estimation step, SAGE in initialization part can set $\beta = 1$ in Eq. (4.15) for the hidden data expectation to include the most estimation errors. In the estimation step, the β is set for a smaller value to maintain the stability of the estimation performance.
2. Sequential path estimation: In this method, the initialization starts with the strongest power as the first path component. The programme estimates the first path within the entire data. After the estimation is done, the reconstructed data for it will be reduced from the measurement data. For the next iteration, the programme searches for the strongest power in the residual data as the second path component, same estimation is conducted. The initialization continuous until the threshold is reached. This method can reduce the influence of previous path components initialization to the next path component initialization. Meanwhile, the estimation error can still make some false initialization due to the erroneous residual measurement data.
3. Sequential path estimation with path management: The path management is an extension for the sequential path estimation. After reducing one reconstructed data from the previous path component initialization, the next path component is not found in the

strongest power peak of the residual power spectrum. As an alternative, the next path component is redirected to find the strongest peak within a predefined division of the whole power spectrum. The division can be made from the beginning by dividing the whole power spectrum with the peak distribution. One benefit of the introduction of path management is that the estimation error from the previous path component initialization is reduced in the residual data. Furthermore the algorithm can find multi-path components in a larger power range.

5.5 SAGE

The SAGE follows the general SAGE structure in Chapter 4. The β value is selected to optimize the estimation performance.

For the Maximization step in SAGE, the joint estimation of all parameters of one path component will increase the computation complexity. A solution to reduce the complexity is to adopt a sequential parameter estimation in the Maximization step. The order of the parameter estimation should be carefully sorted. One sensible strategy is to group the parameters with regards to the moment description. The parameter describing the mean position of the path component and the parameters describing the power spread should be grouped respectively. The power estimation determine the stability of the estimation performance and should always be estimated at the beginning of the Maximization step.

Complete flow chart

The Fig. 5.3 depicts a complete flow chart of the algorithm using the sequential estimation steps.

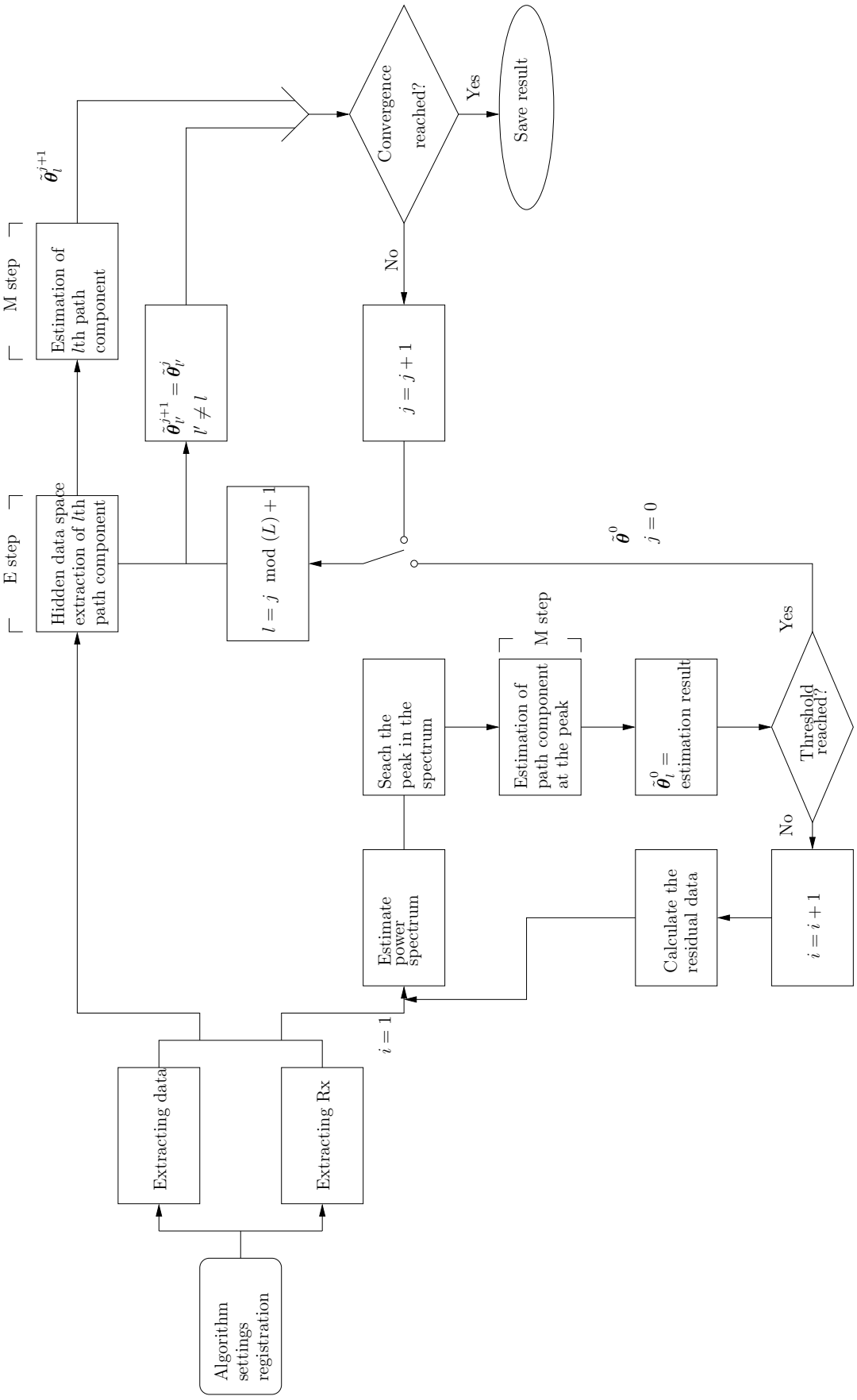


Figure 5.3: A flow chart of the algorithm

Numerical and Experimental Results

6

The discussion in this chapter is divided into two parts, the test of the programme in simulation environment and the numerical and experimental results of the programme with measurement data. The test in the simulation environment is done to validate the performance of the programme. The measurement conditions and environment is described in details such as the channel sounder equipment and the sounding signal. The model mismatch for time variant scenario is analyzed for measurement data and optimization is made.

6.1 Synthetic data estimation

6.1.1 Synthetic data creation principles and methods

Estimation of synthetic data is carried out in order to evaluate the performance of the algorithm. The principles behind the synthetic data creation include:

1. The synthetic data is created with known and controllable power distributions of multi-path components. Different sets of synthetic data created by the same parameter settings are supposed to have the same power distribution in direction and delay.
2. The model mismatch is not considered in the synthetic data. Instead of applying the Monte Carlo simulation to generate a random process for the path components, the simulation directly generate the covariance matrix for the synthetic data. The procedure of the covariance matrix generation follows the model in Chapter 2 where the distribution of the path component follows the model in Chapter 3.
3. The synthetic data shares the antenna response information with the measurement data in the time variant scenario. It also shares the same training sequence $s(t)$ and its autocorrelation function Rss(τ). The similar settings can anticipate the programmes performance in the measurement data.

The truncated sampled $Rss(\tau)$ for the synthetic data is given:

$$Rss(k) = \begin{cases} 1 & k = 0 \\ 0.5 & |k| = 1 \\ 0 & |k| > 1 \end{cases}$$

An omnidirectional antenna array is considered in the synthetic environment. This is the same antenna array of the Tx antenna array for the measurement data shown in Fig. 6.7. The antenna array has 50 antennas with calibrated antenna response. The sum of antenna response is given in Fig. 6.1. The effective radiation of the antenna array is:

Azimuth $[-180°, 180°)$, elevation: $[-70°, 90°]$.

The resolution of the Bartlett spectrum for this antenna array is illustrated in Fig. 6.2. It can be seen that the effect of the sidelobe in the spectrum is small.

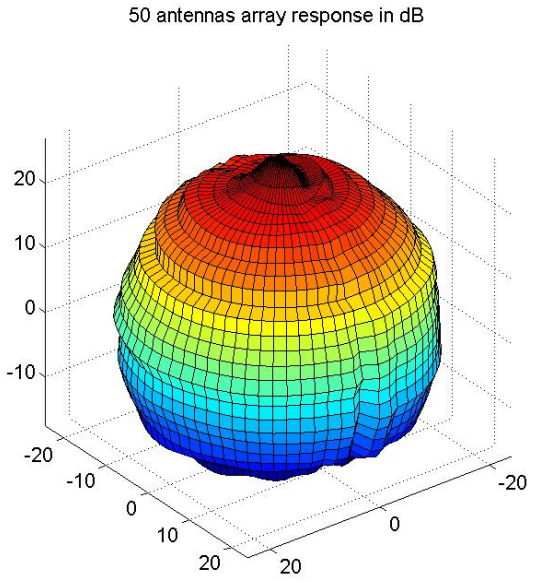


Figure 6.1: Antenna array response of the selected 50 antennas for the synthetic data creation

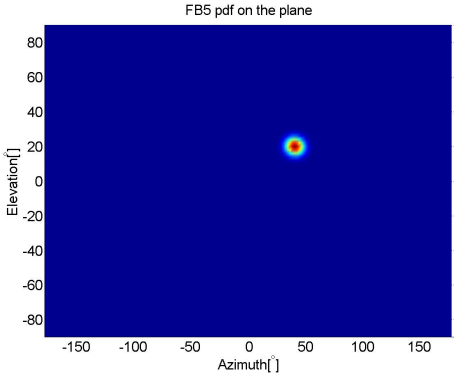
6.1.2 Estimation results

The synthetic data is created with 2 path components following the joint power distribution in direction and delay given in Eq. (3.17). The estimation is divided into two groups. One without path order mismatch and one with path order mismatch where the estimation tries to estimate 3 path component.

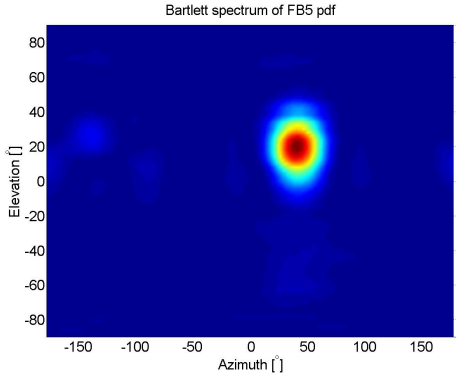
The estimation result without model order mismatch is given in Fig. 6.3. It can be seen that the position of the path components are well captured, the delay spread and direction spread estimations for each path component are also quite consistent with the settings of the synthetic data.

The estimation result with path order mismatch is given in Eq. (6.4). It can be seen that the extra path estimation converges towards a quite low power that can be ignored. This supports the performance of the programme under path order mismatch.

The algorithm shows reliable and accurate estimation results in the synthetic environment. Based on that, we believe the algorithm is applicable for the measurement data.



(a) The FB5 distribution shown on the plane

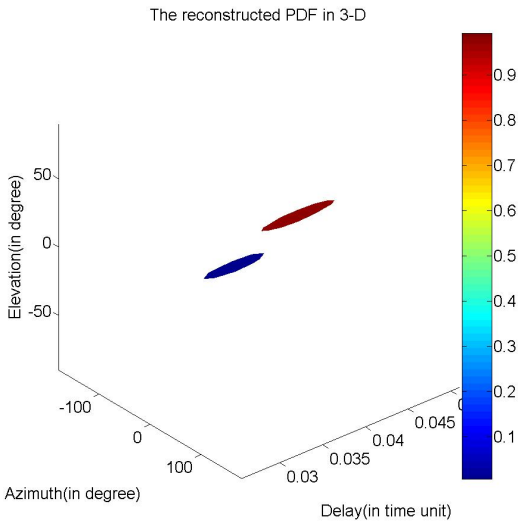


(b) Bartlett spectrum for the FB5 distribution shown on the plane

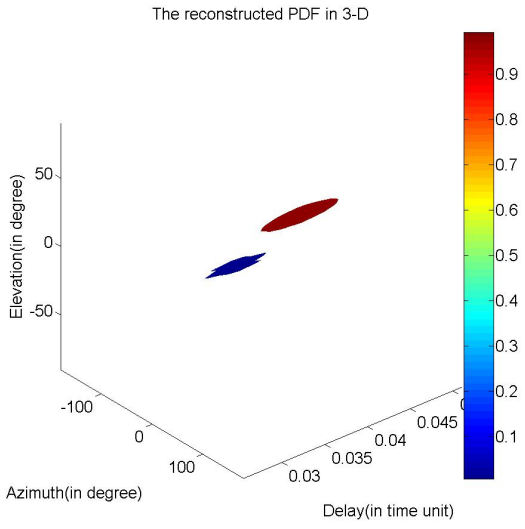
Parameter of FB5	Value
κ	200
ζ	0.1
α	$40°$
$\bar{\theta}$	$40°$
$\bar{\phi}$	$20°$
P	$2.961 \cdot 10^{-11}$
$\delta^2_{\boldsymbol{w}}$	$3.776 \cdot 10^{-12}$

Notice that the sidelobes can still be observed from the Bartlett spectrum. However comparing Fig. 6.2(a) and Fig. 6.2(b), the effect of the sidelobes is considered to be small.

Figure 6.2: An illustration of the 50 antennas array resolution.

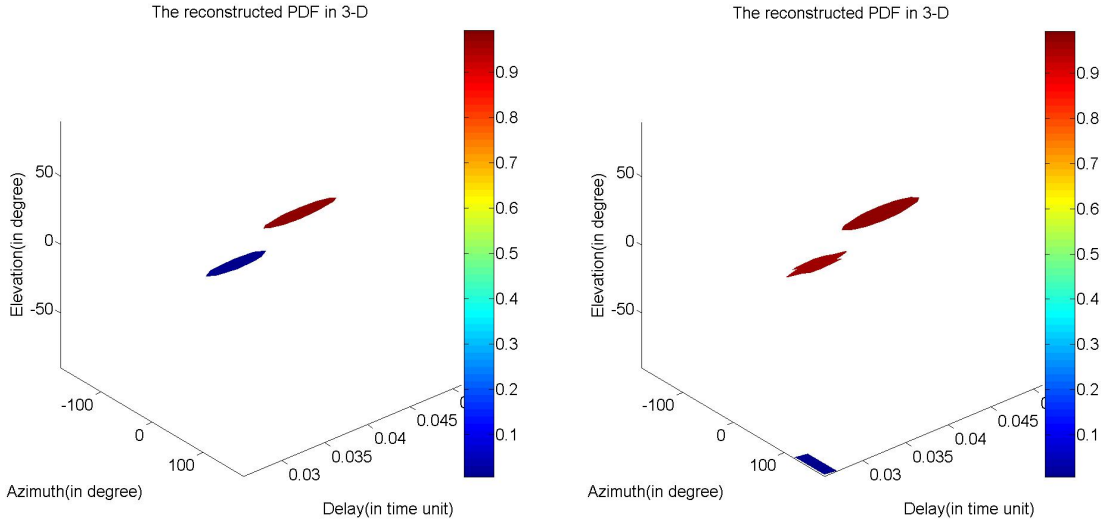


(a) Plot of the true path components



(b) Plot of the estimated path components

Figure 6.3: Estimation of synthetic data without path order mismatch.



(a) Plot of the true path components

(b) Plot of the estimated path components

Figure 6.4: Estimation of synthetic data with path order mismatch.

6.2 Measurement data estimation

6.2.1 Introduction to Electrobit PropSound CS

The measurement data is collected by a channel sounder called EB PropSound CS, designed and produced by Elektrobit Group Ltd[1]. The sounder is a wide band MIMO radio channel sounder and it is able to capture and record fast fading impulse responses, Doppler frequency, polarization as well as DoD and DoA in the radio channel. Electrobit also provides tools in Matlab environment to proceed the measurement data.

For the current project, the channel sounder operates as a TDM channel sounder. Basic information for the settings of the channel sounder is offered in Table 6.1.

The sounding signal (or so-called training sequence) is selected as a known Pseudo-Noise (PN) training sequence with unit power for each sounding period:

$$\int_{t_n}^{t_n+T} s(t)s^*(t)dt = 1 \tag{6.1}$$

Notice that the sampling frequency is twice the chip rate given in Table 6.1. This leads to the sampled autocorrelation function $\mathrm{Rss}(k)$ as given in Fig. 6.5. As seen the $\mathrm{Rss}(k)$ has small values when $|k| > 1$. The algorithm truncates $\mathrm{Rss}(k)$ which greatly reduces the computation complexity:

$$\mathrm{Rss}(k) = \begin{cases} 1 & k = 0 \\ 0.5 & |k| = 1 \\ 0 & |k| > 1 \end{cases},$$

[1] http://www.elektrobit.com

Basic parameters	
Carrier frequency	5.25GHz
Transmit power	+26dBm
Bandwidth	200MHz
Chip rate	100MChip/s
Sampling frequency	200MHz
Measurement distance	
Code length	255 chips
Measurement excess delay	2.55 μs
Measurement excess distance	765 m
Spatial resolution parameters	
Number of Tx antennas	50
Number of Rx antennas	32
Array scan time	8.42 ms
Tx antenna height	1.53 m
Rx antenna height	1.05 m
Sounder setting	
Samples per chip	2
Total channels	1650
Channels in use	1600

Table 6.1: Basic channel sounder parameters

where k stands for the index of the delay sample τ_k.

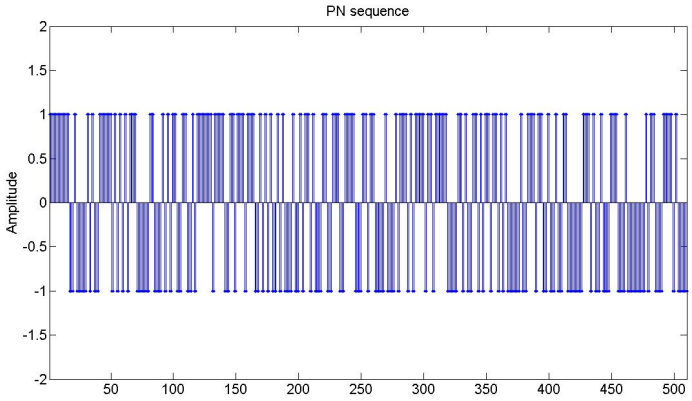
The sounder switching mode is given in Fig. 6.6. Notice that there are 50 sub-channels that are used as guard intervals at the beginning of each Rx switching cycle. The sub-channel extraction should disregards those guard interval sub-channels.

6.2.2 Description of the measurement environment and equipment

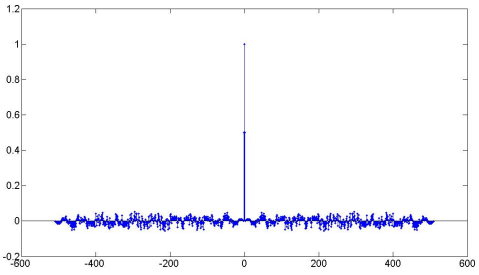
Measurement data for time variant scenario is collected in Oulu university in Finland. The data is collected indoors at different locations including large and small rooms. The LOS and NLOS paths may exist in different scenarios. The equipments of the Tx and Rx are installed in trolleys and by assumption there is no vertical movement for the equipments. Both Tx and Rx antenna array are omni-direction antenna arrays. Photographs for the antenna array is given in Fig. 6.7 and basic information of the antenna arrays are given in Table 6.2.

The information in Table 6.2 tells that antennas in the second ring from the bottom(19 to 36) of the Rx antenna array do not work. In [27] analysis shows that this defect results in a distance larger than half of the wavelength between the antennas at the bottom and at the top for Rx antenna array. Consequently, it leads to ambiguity for the elevation of arrival. To this point, the estimation for the elevation of arrival becomes unreliable. Fortunately, the Tx antenna array utilizes all the antennas in the antenna array and provide little ambiguity on the elevation of departure. The estimation of elevation of departure is then much more reliable than the estimation of elevation of arrival. The following estimation is done with the data extracted for a MISO system.

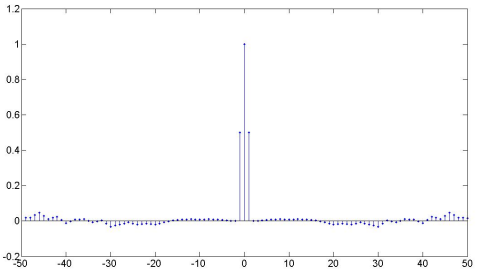
The selected environment is TxR11, the snapshots of TxR11 environment is given in Fig. 6.8.



(a) The PN sequence



(b) The biased autocorrelation of the signal code



(c) The partial autocorrelation around 0

Figure 6.5: The PN signal and it autocorrelation

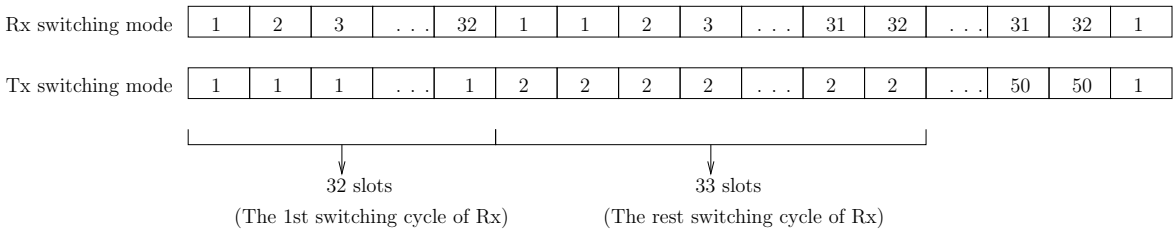
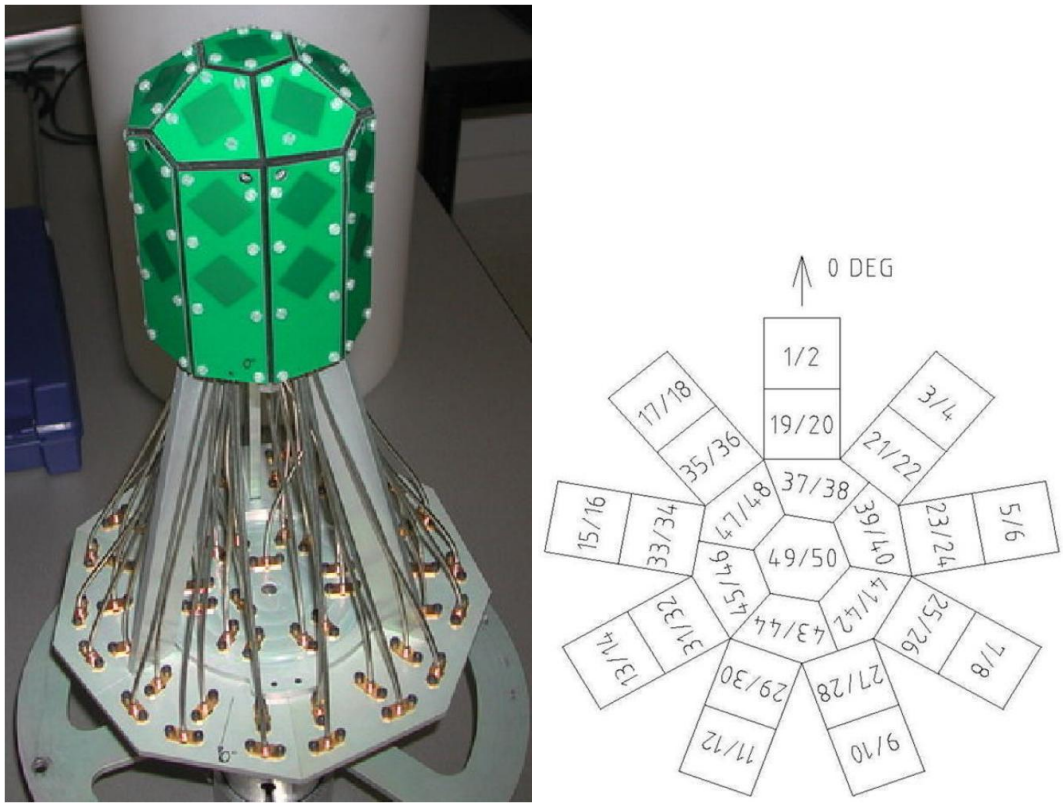


Figure 6.6: The sounder switching mode for the Tx and Rx antennas



(a) Photograph (b) Schematic plot

Figure 6.7: Omni-direction antenna array used at 5.25GHz at both, Tx and Rx

Transmitter	
Antenna designation	2x9ODA_5G25_T1
Frequency/Bandwidth	5.25GHz/420MHz
Radiation	$[-180°\quad 180°]$ in azimuth, $[-70°\quad 90°]$ in elevation
Antenna type	Dual polarized ($\pm 45°$) path array, 50 elements (2×25)
Antenna in use (Index Fig. 6.7(b))	All elements
Receiver	
Antenna designation	2x9ODA_5G25_T2
Frequency/Bandwidth	5.25GHz/420MHz
Radiation	$[-180°\quad 180°]$ in azimuth, $[-70°\quad 90°]$ in elevation
Antenna type	Dual polarized ($\pm 45°$) path array, 50 elements (2×25)
Antenna in use (Index Fig. 6.7(b))	1, 2, . . . 18, 37,. . . , 50

Table 6.2: Information about the antennas arrays

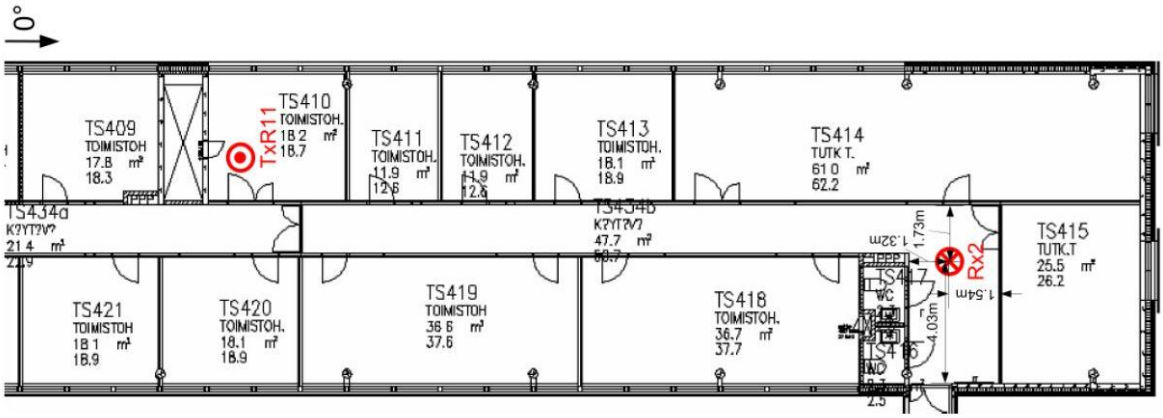
General information	
Tx location	TxR11
Rx location	Rx2
Tx orientation	0°
Rx orientation	0°
Scenario	Time variant
Notes	People moving around
LOS	not available

Table 6.3: General information about the TxR11 environment

Basic information for TxR11 environment is offered in Table 6.3.

(a) Rx environment

(b) Tx environment



(c) Route map of TxR11

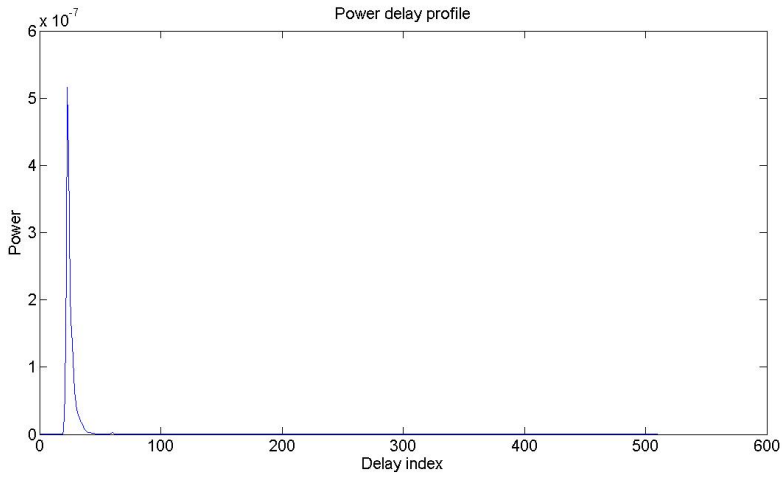
Figure 6.8: The TxR11 environment

Measurement data description

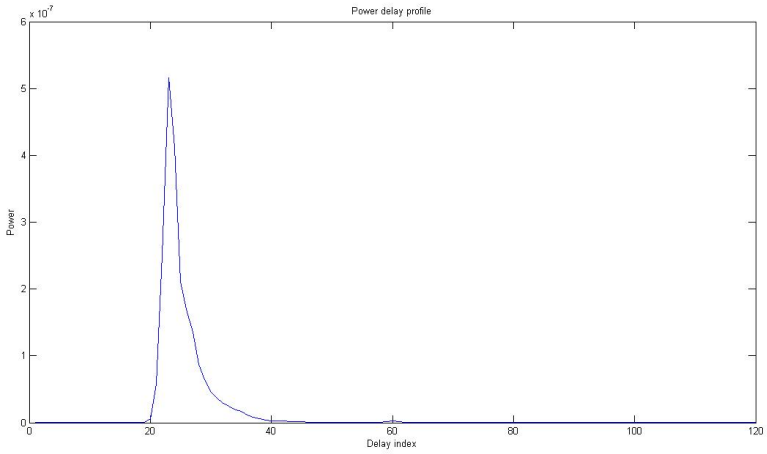
The MISO system is extracted from the MIMO system with the entire Tx antenna array and receive antenna R_9.

The power delay profile for the MISO system is depicted in Fig. 6.9. The power delay profile

is obtained by calculate the average power at different delay as $P(\tau_k) = \sum_{n=1}^{N} \boldsymbol{h}_n(\tau_k)^H \boldsymbol{h}_n(\tau_k)$. Delay samples from τ_{21} to τ_{40} are selected from the observation. The delay samples truncation is sensible that the selected delay samplings contain most of the signal power. The noise estimation is obtained by calculating the average of the power delay profile from delay τ_{100} to the end.



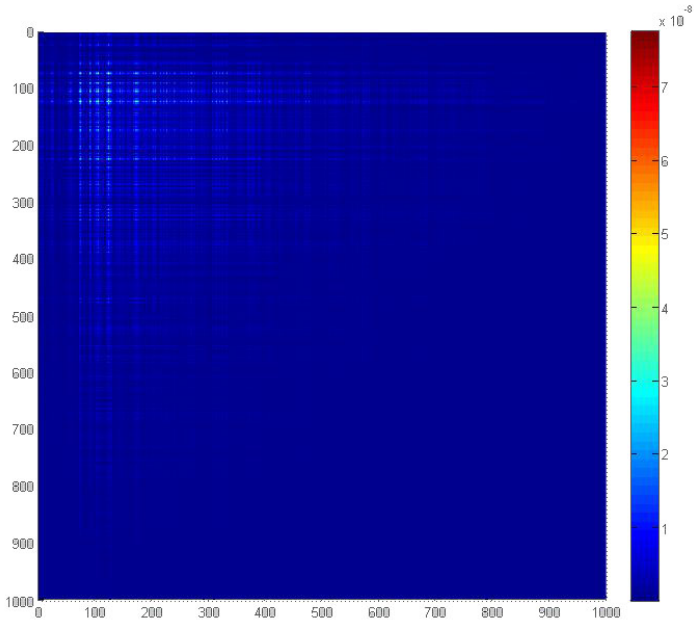
(a) Power delay profile, global view



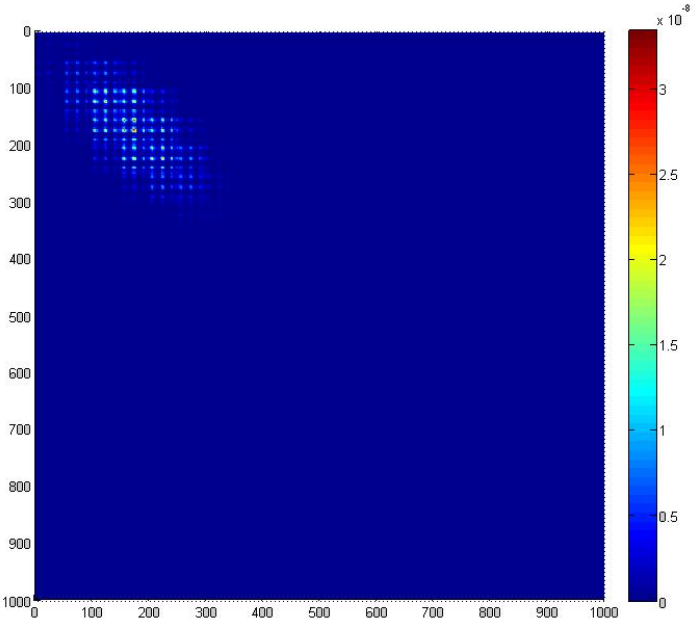
(b) Power delay profile, local view

Figure 6.9: The power delay profile of environment TxR11, MISO system with all Tx antennas and R_9 receive antenna selected

The covariance matrix of the measurement is shown in Fig. 6.10(a). A reconstructed covariance matrix following the WSSUS assumption in Chapter 2 is shown in Fig. 6.10(b). Comparing Fig. 6.10(a) and Fig. 6.10(b) it is found that for the covariance matrix of measurement data, the nonzero entries distributed along the columns and rows that are relatively large which implies that the WSSUS assumption does not hold. That reveals the existence of model mismatch.



(a) Covariance matrix of the channel response for TxR11



(b) The reconstructed channel response covariance matrix using WSSUS model

Figure 6.10: The measurement covariance matrix has nonzero entries that is distributed along the column and row around position $(100, 100)$. With a WSSUS assumption, those positions are suppose to have zero entries or close to zero entries. The dominant nonzero entries given the Rss are supposed to be around the diagonal of the matrix in WSSUS model. It shows the model mismatch for the measurement data.

Discussion on model mismatch

The observation of the model mismatch indicates that the WSSUS assumption of the path components is not true for the observations. In the time variant scenario the WSSUS assumption requires that the complex amplitudes $a_l(\boldsymbol{\Omega}, \tau)$ are uncorrelated for different path component.

Recall the expression:

$$E[a_i(\boldsymbol{\Omega}, \tau)a_j(\boldsymbol{\Omega}, \tau)] = 0. \tag{6.2}$$

As mentioned earlier the covariance matrix derived for a WSSUS model have its nonzero sub-matrices distributed along the diagonal of the covariance matrix. The following discussion provides an investigation of the constitution of the covariance matrix under non-WSSUS condition where different path components are correlated.

Suppose there are 2 specular scatters existing in the environment, transmitted signal for these two scatters can be denoted as:

$$\begin{aligned}\boldsymbol{s}_1 &= a_1(\boldsymbol{\Omega}_1, \tau_1)\boldsymbol{c}(\boldsymbol{\Omega}_1)s(t-\tau_1)\\ \boldsymbol{s}_2 &= a_2(\boldsymbol{\Omega}_2, \tau_2)\boldsymbol{c}(\boldsymbol{\Omega}_2)s(t-\tau_2)\end{aligned}$$

The received signal $\boldsymbol{y}(t)$ is the sum of $\boldsymbol{s}_1(t)$ and $\boldsymbol{s}_2(t)$

$$\boldsymbol{y}(t) = \boldsymbol{s}_1(t) + \boldsymbol{s}_2(t) \tag{6.3}$$

Recall the representation of the estimated channel impulse response given by:

$$\boldsymbol{h}_n(\upsilon) = \int_{t_n}^{t_n+T} \boldsymbol{y}(t)s(t-\upsilon)dt \tag{6.4}$$

Let,

$$\begin{aligned}a_1(\boldsymbol{\Omega}_1, \tau_1) = a_1, \quad \boldsymbol{c}(\boldsymbol{\Omega}_1) = \boldsymbol{c}_1, \quad \boldsymbol{s}_1 = a_1\boldsymbol{c}_1 s(t-\tau_1)\\ a_2(\boldsymbol{\Omega}_2, \tau_2) = a_2, \quad \boldsymbol{c}(\boldsymbol{\Omega}_2) = \boldsymbol{c}_2, \quad \boldsymbol{s}_2 = a_2\boldsymbol{c}_2 s(t-\tau_2)\end{aligned}$$

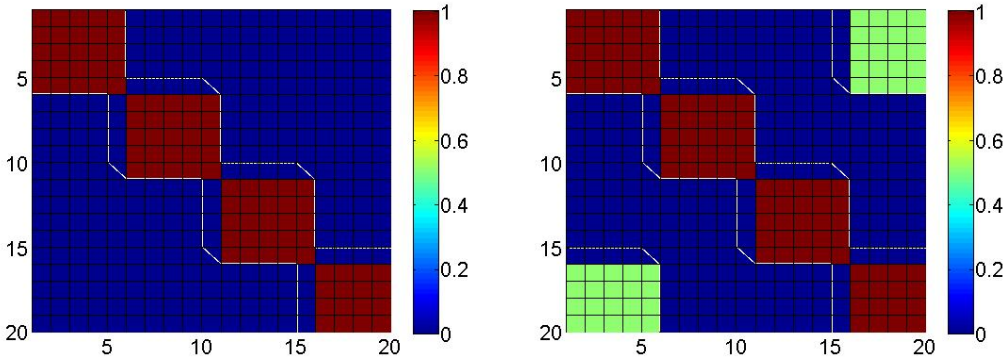
for short. The cross covariance matrix for the channel response $\boldsymbol{h}(\upsilon)$ at two different delays υ_1 and υ_2 is derived in A.7:

$$\begin{aligned}&E[\boldsymbol{h}(\upsilon_1)\boldsymbol{h}^H(\upsilon_2)] =\\ &\int_{t_n}^{t_n+T}\int_{t_n}^{t_n+T}[E[a_1a_1^*]\boldsymbol{c}_1\boldsymbol{c}_1^H s(t_1-\tau_1)s^*(t_2-\tau_1)s(t_1-\upsilon_1)s^*(t_2-\upsilon_2)+\\ &\qquad E[a_1a_2^*]\boldsymbol{c}_1\boldsymbol{c}_2^H s(t_1-\tau_1)s^*(t_2-\tau_2)s(t_1-\upsilon_1)s^*(t_2-\upsilon_2)+\\ &\qquad E[a_2a_1^*]\boldsymbol{c}_2\boldsymbol{c}_1^H s(t_1-\tau_2)s^*(t_2-\tau_1)s(t_1-\upsilon_1)s^*(t_2-\upsilon_2)+\\ &\qquad E[a_2a_2^*]\boldsymbol{c}_2\boldsymbol{c}_2^H s(t_1-\tau_2)s^*(t_2-\tau_2)s(t_1-\upsilon_1)s^*(t_2-\upsilon_2)]dt_1dt_2\end{aligned}$$

Suppose the autocorrelation of $s(t)$ is a delta function $\text{Rss}(\tau) = \delta(\tau)$.

$$\begin{aligned}\boldsymbol{h}(\upsilon_1)\boldsymbol{h}^H(\upsilon_2) =&[E[a_1a_1^*]\boldsymbol{c}_1\boldsymbol{c}_1^H\delta(\tau_1-\upsilon_1)\delta(\tau_1-\upsilon_2) + E[a_1a_2^*]\boldsymbol{c}_1\boldsymbol{c}_2^H\delta(\tau_1-\upsilon_1)\delta(\tau_2-\upsilon_2)+\\ &E[a_2a_1^*]\boldsymbol{c}_2\boldsymbol{c}_1^H\delta(\tau_2-\upsilon_1)\delta(\tau_1-\upsilon_2) + E[a_2a_2^*]\boldsymbol{c}_2\boldsymbol{c}_2^H\delta(\tau_2-\upsilon_1)\delta(\tau_2-\upsilon_2)]\end{aligned} \quad (6.5)$$

Under non-WSSUS assumption, as long as $E[a_1a_2^*] \neq 0$ or equivalently a_1 and a_2 are correlated, sub-matrices $E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_2)]$ and $E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_2)]$ will become nonzero sub-matrices. These nonzero sub-matrices contain the information of the correlated path component. Fig. 6.11 gives a brief instance of the correlation in covariance matrix.



(a) WSSUS model (b) non WSSUS model

Figure 6.11: A comparison of covariance matrices with WSSUS assumption and with non-WSSUS assumption with 4 antennas and 5 delay samples. $\text{Rss}(\tau) = \delta(\tau)$. The color of the figure represents the absolute value of the entries. In non-WSSUS model, path components in the first delay and in the last delay are considered correlated. This leads to two nonzero sub-matrices in the corner.

To decrease the effect of model mismatch it is suggested to apply LSE method. LSE can be viewed as a fitting technique to approximate the maximum likelihood. d

To further reduce the effect of model mismatch using LSE, one can recheck the constitution of the covariance matrix. By the given model, it can be seen that the sub-matrices along the diagonal of the covariance matrix contain all the power distribution information in direction and delay. The sub-matrices in other positions of the covariance matrix are just repeated combination of the the power spectrum for different delays. For this point, the model mismatch can be reduced further by only fitting the diagonal sub-matrices using LSE. This is called selective LSE.

Estimation results

Estimation of the given MISO system in the environment TxR11 contains 50 sub-channels and 20 delay samples. This made the covariance matrix of $\boldsymbol{H}$ have a size of 1000×1000. Matlab has its limitation on calculating such huge matrix. The MLE method introduced in Eq. (4.19) calls for the determinant computation and inversion operation on the covari-

ance matrix. The results will become unreliable in Matlab. As an alternative, the MLE was replaced by the LSE method.

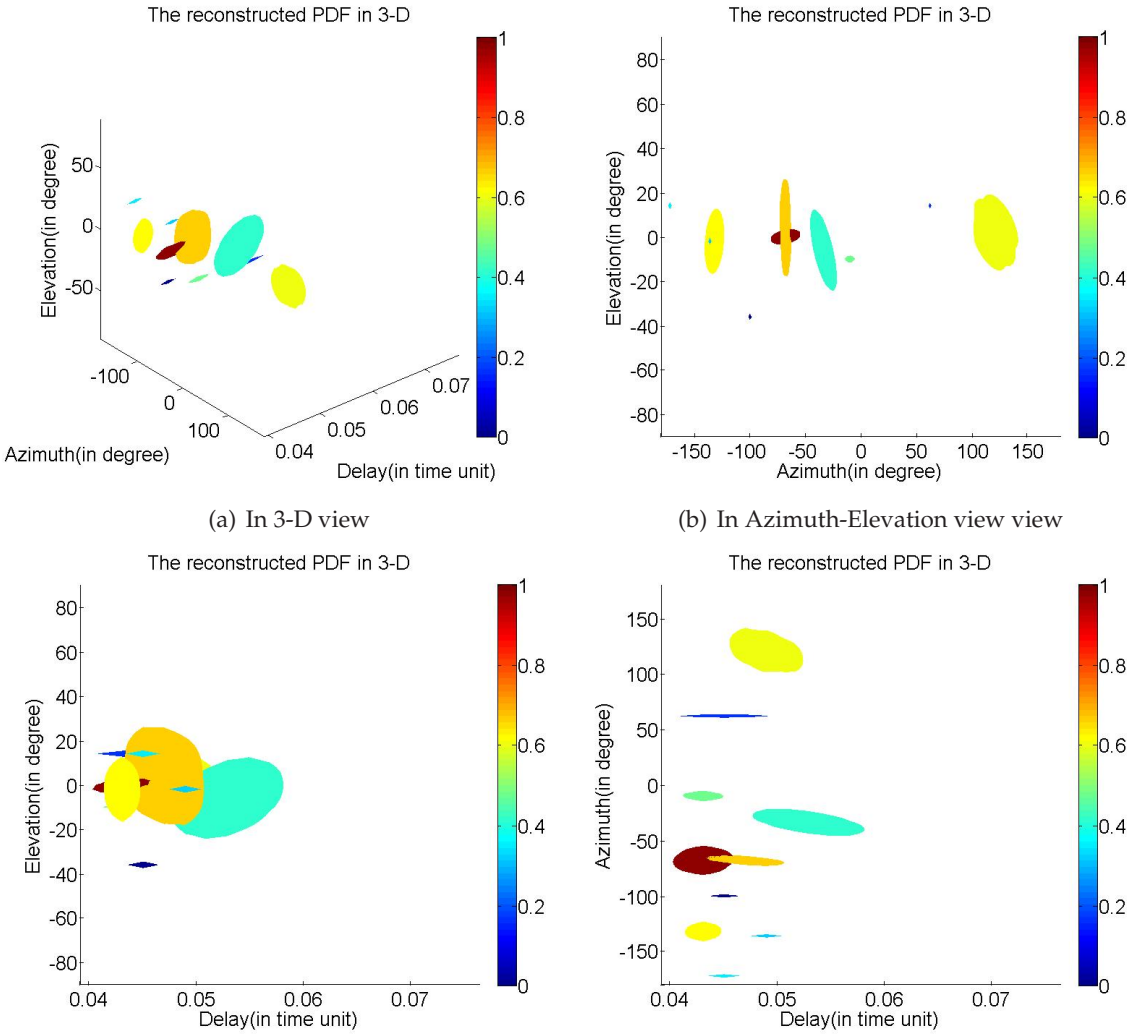
To reduce the effect of model mismatch in the measurement, the selective LSE method mentioned previously is applied which only considers the sub-matrices on the diagonal.

In short, estimation under environment TxR11 in time variant scenario adopts selective LSE method to fit the sub-matrix along the diagonal of measurement covariance matrix. The signal model is still considered to follow the WSSUS assumption.

Fig. 6.12 (a) displays the estimation results using joint direction delay power distribution given Eq. (3.17). The estimation limits the maximum number of path components to 10 and the power level threshold to 20dB below the maximum power found in the Bartlett spectrum at the initialization step.

Fig. 6.12 (b), (c), and (d) provide different views of the power distribution for the multi-path components. It can be seen that there are two path components overlapped around the position with azimuth -70° and elevation 0°. This observation indicates that the path component distribution at those positions does not fully follow the proposed model. The programme then tries using 2 path components to patch the real path component.

The list of the estimated parameter sets are given in Table B.3.



(a) In 3-D view

(b) In Azimuth-Elevation view view

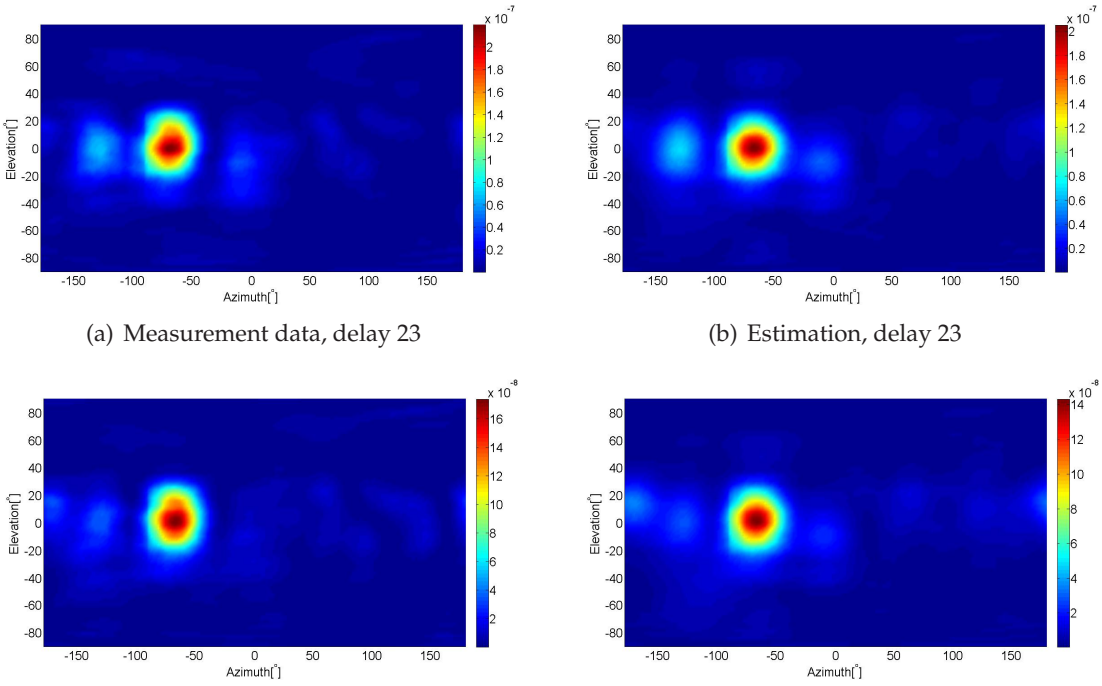
(c) In Elevation-Delay view

(d) In Azimuth-Delay view

Figure 6.12: The estimated power distribution of the path component using the joint FB5 Gaussian distribution. A MISO system with all Tx antennas and R_9 receive antenna is considered in a time variant scenario and in the environment TxR11. The color of the figure represent the signal power in dB of each path component.

To evaluate the estimation results, the reconstructed Bartlett spectrum is compared with the Bartlett spectrum of the measurement data at each delay. This is given from Fig. 6.13 to Fig. 6.15.

Comparing the strongest power delay τ_{23} and the second strongest power delay τ_{24} in Fig. 6.13, The reconstructed Bartlett spectrum is quite close to the measurement Bartlett spectrum. It is believed that the most dominant paths are captured by the estimation. Furthermore, it is worth noticing that in the reconstructed Bartlett spectrum at delays τ_{23} and τ_{24} the estimation managed to capture the path components with smaller power. This observation further confirms the performance of the estimation.



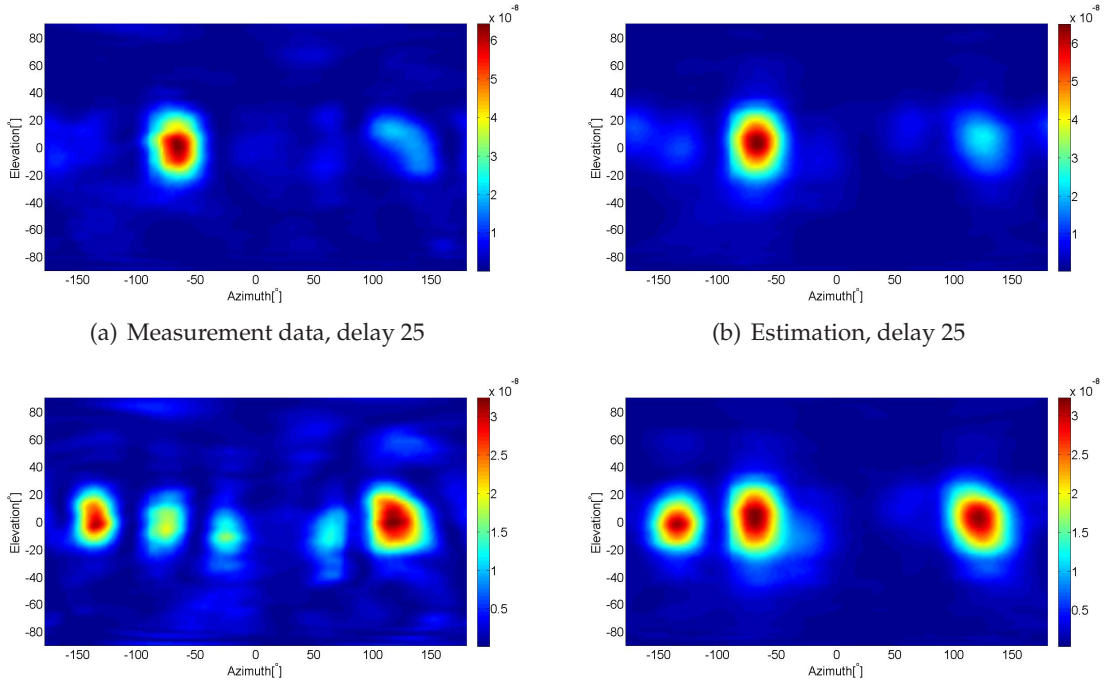
(a) Measurement data, delay 23

(b) Estimation, delay 23

(c) Measurement data, delay 24

(d) Estimation, delay 24

Figure 6.13: Bartlett spectrum comparison, delay 23 and delay 24



(a) Measurement data, delay 25

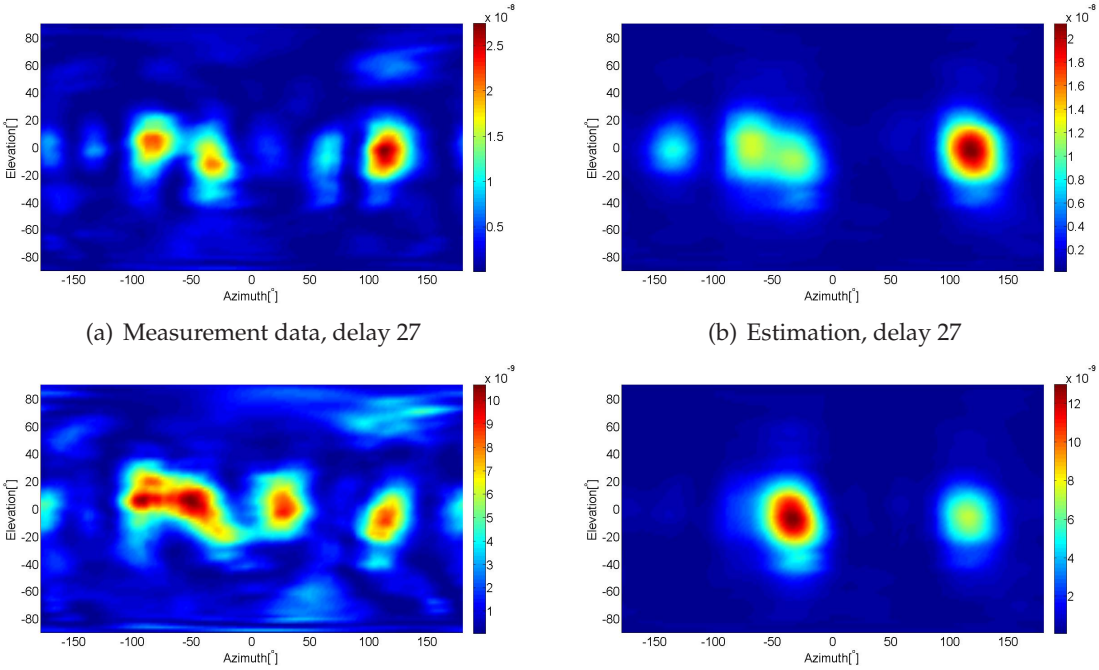
(b) Estimation, delay 25

(c) Measurement data, delay 26

(d) Estimation, delay 26

Figure 6.14: Bartlett spectrum comparison, delay 25 and 26

In Fig. 6.14 and Fig. 6.15, the estimation continues capturing the path component at later delay samples. In delay τ_{25} 2 path components are captured with a close angular spread estimation, shown in Fig. 6.14. At delay τ_{26} the estimated power for the path components are different from the measurement power spectrum. It could be due to the skewness of the measurement power spectrum that does not fully follow the proposed model. The estimation error can be alleviated by introducing supplementary path components to patch this skewed power spectrum.



(a) Measurement data, delay 27

(b) Estimation, delay 27

(c) Measurement data, delay 28

(d) Estimation, delay 28

Figure 6.15: Bartlett spectrum comparison, delay 27 and 28

In delay τ_{27}, 4 paths are captured with promising power level and angular spread. The estimation does not capture all the peaks in the Bartlett spectrum because of the power threshold. The reader may observe the difference between the estimation and the measurement at delay τ_{28} where some small power path components are not estimated. The difference between the estimation and the measurement is caused by the path order mismatch. One can anticipate that, as the threshold of the power allows larger power range to get more path components, the programme will capture more path components.

The estimation shows promising performance by using the selective LSE even there is model mismatch in the measurement data. The path order selection has to be consistent with observation to get good estimation results. It is reasonable to concentrate on the estimation for the dominant path components. The estimation results in environment TxR11 shows that the estimation is trustable by focusing on the dominant path components.

Part II

Time Invariant Scenario

Path Component Distribution Model

7

7.1 Channel model modification in the time invariant scenario

A time invariant scenario means that the path component is deterministic and that the complex amplitude $a(\mathbf{\Omega}, \tau)$ is a deterministic function of direction and delay. Thereby both the amplitude distribution and the phase distribution of $a(\mathbf{\Omega}, \tau)$ needs to be modeled in the time invariant scenario.

7.2 Path component model

The power distribution of one path component becomes just $a(\mathbf{\Omega}, \tau)a^*(\mathbf{\Omega}, \tau)$. the power distribution still fulfills the modeling condition given in Chapter 3, thereby the proposed power density function discussed in Chapter 3 can be reused to describe the power distribution in time invariant scenario. Eq. (7.1) provides the amplitude distribution model.

$$\begin{aligned} |a(\mathbf{\Omega}, \tau)|^2 &= Pf(\mathbf{\Omega}, \tau) \\ |a(\mathbf{\Omega}, \tau)| &= \sqrt{Pf(\mathbf{\Omega}, \tau)} \end{aligned} \tag{7.1}$$

Modeling of phase distribution for $a(\mathbf{\Omega}, \tau)$ can be regarded as a problem independent of the power distribution for $a(\mathbf{\Omega}, \tau)$. The phase distribution can be caused by many factors. For a reflection of the propagation wave, the phase change depends on the reflection coefficient of the reflector and the reflection angle [21, p.18]. The description of the phase distribution for dispersive path components is quite complex depending on the properties of the scatterers. In this report, the phase distribution is proposed as a identical value of one path component. This suggestion assumes that the path component is quite concentrated and the phase distribution remains close to different delays and directions. The identical modeling of the phase distribution can be viewed as a solution to simplify the path component distribution in time invariant scenario.

The proposed path component distribution reads:

$$a(\mathbf{\Omega}, \tau) = \exp(-j\vartheta)\sqrt{Pf(\mathbf{\Omega}, \tau)} \tag{7.2}$$

where ϑ is the phase of the path component and $f(\mathbf{\Omega}, \tau)$ is the power density function given in Chapter 3.

Radio Channel Propagation Estimation

8

This chapter follows the same structure as the in Chapter 4 where characterization and estimation in time variant scenario is carried. The modifications for the estimation methods are described for the time invariant scenario.

8.1 Channel impulse response estimation with multi-path components

The signal model for multi-path components is the summation of all the individual path components. For L path components, the signal model is in the form:

$$\boldsymbol{y}(t) = \sum_{l=1}^{L} \int_{\mathcal{S},T} a_l(\boldsymbol{\Omega}, \tau) \boldsymbol{C}(\boldsymbol{\Omega}) s(t-\tau) d\boldsymbol{\Omega} d\tau + \boldsymbol{w}(t), \tag{8.1}$$

where each $a_l(\boldsymbol{\Omega}, \tau)$ is a deterministic function.

Recall the channel impulse response estimation method for the time variant scenario. By applying it to the time invariant scenario, we get the same channel impulse response estimation structure:

$$\begin{aligned}
\boldsymbol{h}(\tau') &= \int_{t_n}^{t_n+T} \boldsymbol{y}(t) s^*(t-\tau') dt \\
&= \int_{\mathcal{S},\mathcal{T}} \sum_{l=1}^{L} a_l(\boldsymbol{\Omega}, \tau) \boldsymbol{C}(\boldsymbol{\Omega}) \int_{t_n}^{t_n+T} s(t-\tau) s^*(t-\tau') dt d\boldsymbol{\Omega} d\tau + \int_{t_n}^{t_n+T} \boldsymbol{w}(t) s^*(t-\tau') dt \\
&= \sum_{l=1}^{L} \int_{\mathcal{S},\mathcal{T}} a_l(\boldsymbol{\Omega}, \tau) \boldsymbol{C}(\boldsymbol{\Omega}) \mathrm{Rss}(\tau - \tau') d\boldsymbol{\Omega} d\tau + \int_{t_n}^{t_n+T} \boldsymbol{w}(t) s^*(t-\tau') dt.
\end{aligned} \tag{8.2}$$

The estimated channel impulse estimation $\boldsymbol{h}(\tau')$ is made up of two parts, the deterministic part $\sum_{l=1}^{L} \int_{\mathcal{S},\mathcal{T}} a_l(\boldsymbol{\Omega}, \tau) \boldsymbol{C}(\boldsymbol{\Omega}) \mathrm{Rss}(\tau - \tau') d\boldsymbol{\Omega} d\tau$ and the random part $\int_{t_n}^{t_n+T} \boldsymbol{w}(t) s^*(t-\tau') dt$ which is an additive Gaussian noise. It is then known that $\boldsymbol{h}(\tau')$ follows the Gaussian distribution.

Again the stacked channel response estimation $\boldsymbol{H}$ is provided:

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{h}(\tau_1) \\ \boldsymbol{h}(\tau_2) \\ \vdots \\ \boldsymbol{h}(\tau_k) \\ \vdots \\ \boldsymbol{h}(\tau_K) \end{bmatrix} \tag{8.3}$$

and $\boldsymbol{H}$ follows the Gaussian distribution.

8.2 Stochastic analysis

For the Gaussian distributed channel impulse response vector $\boldsymbol{H}$, the mean and the covariance matrix for $\boldsymbol{H}$ are studied.

The mean of $\boldsymbol{H}$ is given by:

$$E[\boldsymbol{H}] = \begin{bmatrix} E[\boldsymbol{h}(\tau_1)] \\ E[\boldsymbol{h}(\tau_2)] \\ \vdots \\ E[\boldsymbol{h}(\tau_k)] \\ \vdots \\ E[\boldsymbol{h}(\tau_K)] \end{bmatrix}, \tag{8.4}$$

In time invariant scenario, expression for the sub-vector $E[\boldsymbol{h}(\tau_1)]$ is derived in A.2.

$$E[\boldsymbol{h}(\tau_k)] = \sum_{l=1}^{L} \int_{\mathcal{S},T} a_l(\boldsymbol{\Omega}, \tau) \boldsymbol{C}(\boldsymbol{\Omega}) R_{ss}(\tau - \tau_k) d\boldsymbol{\Omega} d\tau \tag{8.5}$$

The covariance of $\boldsymbol{H}$ is given by:

$$E[\boldsymbol{H}\boldsymbol{H}^H] = \begin{bmatrix} E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_1)] & E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_2)] & \dots & E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_K)] \\ \vdots & & \ddots & \vdots \\ & & E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)] & \\ E[\boldsymbol{h}(\tau_K)\boldsymbol{h}^H(\tau_1)] & & & E[\boldsymbol{h}(\tau_K)\boldsymbol{h}^H(\tau_K)] \end{bmatrix} \tag{8.6}$$

And the derivation of a sub-matrix in the covariance matrix is derived in Appendix A.2. The expression can be equivalently viewed as:

$$E[\boldsymbol{H}\boldsymbol{H}^H] = \sigma_w^2 \boldsymbol{D}, \tag{8.7}$$

where $\boldsymbol{D}$ is a matrix determined by the $Rss(\tau)$ function. The covariance matrix of the channel impulse response in the time invariant scenario is not related to the path component but only determined by the noise variance and Rss(τ).

8.3 The SAGE

The SAGE for the time invariant scenario follows the same structure as in the time variant scenario. Only the Expectation step and the Maximization step are modified.

8.3.1 Expectation step in the SAGE

The hidden data is composed by the deterministic data of the path component and the additive noise part. Let:

$$\boldsymbol{S}_l = \int_{S,T} a_l(\boldsymbol{\Omega}, \tau)\boldsymbol{C}(\boldsymbol{\Omega})R_{ss}(\tau - \tau_k)d\boldsymbol{\Omega}d\tau$$

denote the signal part for the lth path component. Let:

$$E[\boldsymbol{H}] = \sum_{l=1}^{L} \boldsymbol{S}_l$$

denote the signal part for the multi-path components.

The hidden data for the lth path component is in the form:

$$\boldsymbol{H}_i = \boldsymbol{S}_l + \beta \boldsymbol{w}, \tag{8.8}$$

where $\boldsymbol{w}$ is the introduced noise.

The conditional expectation of the hidden data is given as:

$$\tilde{\boldsymbol{H}}_l^j = \tilde{\boldsymbol{S}}_l^{j-1} + \beta(\overline{\hat{\boldsymbol{H}}} - E[\tilde{\boldsymbol{H}}]^{j-1}), \tag{8.9}$$

where $\overline{\hat{\boldsymbol{H}}} = \frac{1}{N}\sum_{n=1}^{N} \hat{\boldsymbol{H}}_n$ is the observed mean of $\boldsymbol{H}$ from the measurement data.

The β is again the controlling parameter to determine how much estimation errors that are introduced into the hidden data for the Maximization step.

8.3.2 Maximization step in the SAGE

The MLE and LSE are candidates of the estimator for the Maximization step in the SAGE.

Recall the general principles of MLE introduced in the variant scenario, the channel impulse response for time invariant scenario follows the Gaussian distribution given the condition that the noise is additive Gaussian noise. By taking the $E[\boldsymbol{H}]$ and $E[\boldsymbol{H}\boldsymbol{H}^H]$ given in Eq. (8.5) and Eq. (8.7) into Eq. (4.18), MLE for one path component in the time invariant scenario is derived in Appendix A.6.

$$\tilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmax}} \ \left(-\sigma_w^2 \ln|\boldsymbol{\Sigma}(\boldsymbol{\theta})| - (E[\hat{\boldsymbol{H}}] - E[\tilde{\boldsymbol{H}}(\boldsymbol{\theta})])^H \boldsymbol{R} (E[\hat{\boldsymbol{H}}] - E[\tilde{\boldsymbol{H}}(\boldsymbol{\theta})])\right) \tag{8.10}$$

The $E[\tilde{\boldsymbol{H}}(\boldsymbol{\theta})]$ is the estimated hidden data for the path component. The $E[\hat{\boldsymbol{H}}]$ is the hidden data expectation from previous estimation results.

For a fixed noise level, Eq. (8.10) is equivalent to:

$$\tilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \ |E[\hat{\boldsymbol{H}}] - E[\tilde{\boldsymbol{H}}(\boldsymbol{\theta})]|, \tag{8.11}$$

which is equivalent to LSE. Therefore, for the time invariant scenario, MLE method equals to the LSE method by the condition that the noise is additive Gaussian noise.

8.3.3 Performance analysis

It is important to keep in mind that for the time invariant scenario the MLE is equivalent to the LSE by the condition the noise is a white additive Gaussian noise. Estimation in time invariant scenario is regarded as fitting the estimated channel impulse response vector $\tilde{\boldsymbol{H}}$ with the measurement. Fitting a complex vector depends on a good phase estimation. The algorithm should be carefully designed to get a good phase estimation. So far the phase distribution model for the path component is set as an identical value and the estimated path components are expected to have a small power spread to maintain the performance of the estimation.

9 Algorithm Design

9.1 General description

The algorithm described in Chapter 5 is reused in the time invariant scenario. The programme structure and the data structure are unchanged. This chapter gives a brief description of what is changed in the power distribution initialization, the phase initialization and the estimation strategies in time invariant scenario.

9.2 The initialization modification

9.2.1 Power spectrum analysis

Both the Bartlett spectrum and the Capon spectrum requires a matrix and for the time invariant scenario this matrix constructed at delay τ_k in the following:

$$\boldsymbol{R}_k = \overline{\hat{\boldsymbol{h}}}(\tau_k)\overline{\hat{\boldsymbol{h}}}^H(\tau_k), \tag{9.1}$$

where $\overline{\hat{\boldsymbol{h}}}(\tau_k) = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{h}_n(\tau_k)$ is the mean vector of the data. The expression provides the matrix which can be used to estimate the power spectrum.

The Bartlett spectrum at delay τ_k is:

$$p(\boldsymbol{\Omega}, \tau_k) = \frac{\boldsymbol{C}(\boldsymbol{\Omega})^H \boldsymbol{R}_k \boldsymbol{C}(\boldsymbol{\Omega})}{\boldsymbol{C}(\boldsymbol{\Omega})^H \boldsymbol{C}(\boldsymbol{\Omega})} \tag{9.2}$$

The Bartlett spectrum in time invariant can be just viewed as the power of the projection of the mean vector $\overline{\hat{\boldsymbol{h}}}(\tau_k)$ to the antenna response $\boldsymbol{C}(\boldsymbol{\Omega})$ in direction $\boldsymbol{\Omega}$.

The Capon spectrum is not available for the time invariant scenario because $\boldsymbol{R}_k$ is a singular matrix and an inversion operation on $\boldsymbol{R}_k$ will cause errors.

The initialization of the power and the power distribution follow the same principles as for the time variant scenario.

9.2.2 Initialization strategy modification

Initialization of the identical phase for one path component can be set to 0 at the beginning and then estimated to an optimal value. Normally the identical phase value is regarded as

a additional information for the path component distribution. The estimation is supposed to have a better performance by adjusting the identical phase value frequently to fit the estimated power distribution.

Numerical and Experimental Results

10

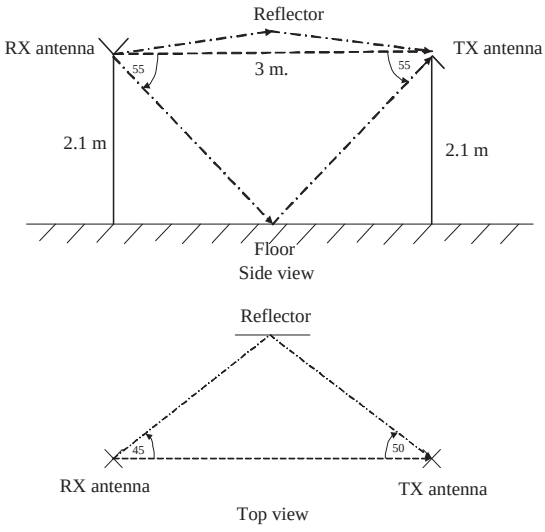
10.1 Measurement data estimation

Measurement data for the time invariant scenario is provided by the same EB PropSound CS as used in the time variant scenario and the same switching pattern and PN training sequence are used as given in Chapter 6. Descriptions of the environment is given in this chapter. The estimation is conducted in MISO and SIMO system. Model mismatch is discussed at the end.

10.1.1 Description of the measurement environment and equipment

Measurement data for the time invariant scenario is collected by Attaphongse Taparugssanagorn from Oulu Univeristy in an anechoic chamber where the walls of the room is covered by absorbing material. The measurement data is collected under control and therefore the propagation path is considered as stationary.

The environment is depicted in Fig. 10.1(a). It shows that only 3 propagation paths are possible: The LOS path, the horizonal reflection path from the side of the equipment and the ground reflection path.



(a) The site map of the environment

25 / 26	18 / 17	9 / 10	2 / 1
27 / 28	20 / 19	11 / 12	4 / 3
29 / 30	22 / 21	13 / 14	6 / 5
31 / 32	24 / 23	15 / 16	8 / 7

(b) The placement of the Rx antennas

The Tx antenna array uses the same equipment given in Fig. 6.7. The Rx antenna array

consists of 32 antennas positioned on a plane. The Rx is not viewed as a omni-directional antenna array anymore. Its effective radiation pattern in azimuth is $(-90°90°)$.

The information for the antennas is given in Table 10.1. Because both Rx and Tx utilize all of their antenna elements no ambiguity is introduced to the estimation of either Direction of Departure (DoD) or Direction of Arrival (DoA).

Transmitter	
Antenna designation	2x9ODA_5G25_T1
Frequency/Bandwidth	5.25GHz/420MHz
Radiation	$[-180° \quad 180°]$ in azimuth, $[-70° \quad 90°]$ in elevation
Antenna type	Dual polarized ($\pm 45°$) path array, 50 elements (2×25)
Antenna in use (Index Fig. 6.7(b))	All elements
Receiver	
Antenna designation	4x4_5G25_R1
Frequency/Bandwidth	5.25GHz/420MHz
Radiation	$[-90° \quad 90°]$ in azimuth, $[-70° \quad 90°]$ in elevation
Antenna type	Dual polarized ($\pm 45°$) path array, 32 elements (2×16)
Antenna in use (Index Fig. 10.1(b))	All the elements

Table 10.1: Information about the antennas arrays

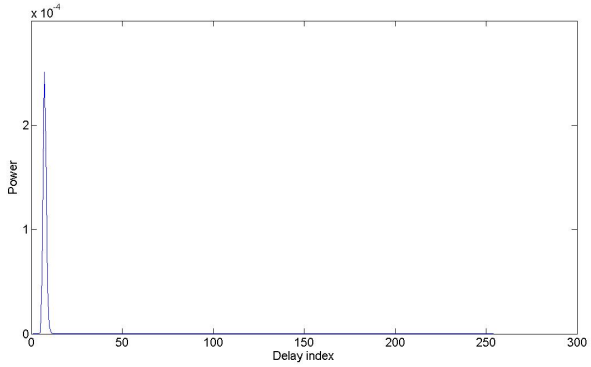
10.1.2 Estimation for measurement data

Estimation of SIMO and MISO systems are conducted respectively.

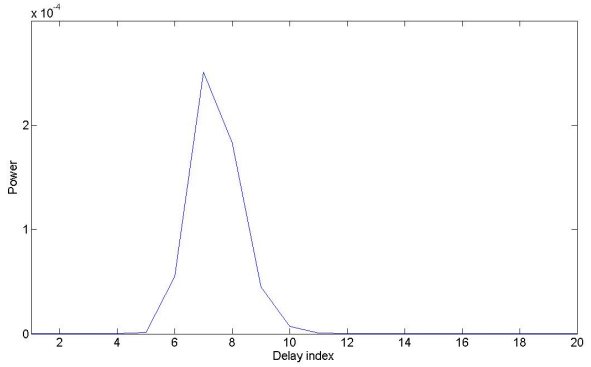
Estimation results for SIMO system

The SIMO system is extracted with T_1 transmit antenna and the even index receive antennas. In Fig. 10.1(b) one antenna is selected in each grid. The minimum distance of the selected antennas are still smaller than half wavelength. There is no ambiguity problem in this SIMO system.

The power delay profile is given in Fig. 10.1. From Fig. 10.1(d), delay τ_4 to τ_{12} are selected for the estimation.



(c) Global view



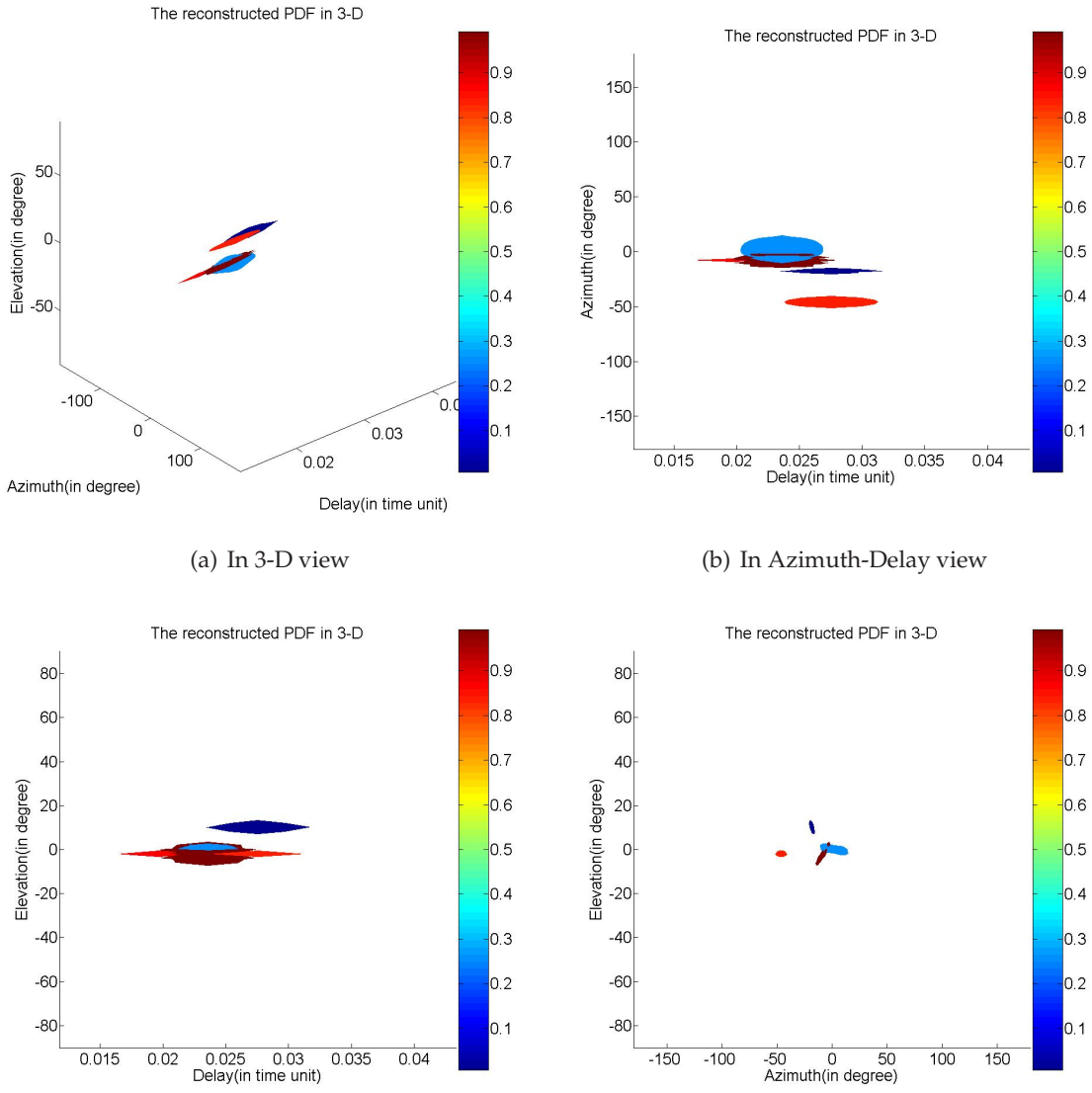
(d) Local view

Figure 10.1: The delay power profile of time invariant SIMO system.

The estimation limits the maximum number of path component estimation to 5. The sequential path initialization method is adopted.

The estimation results for the SIMO system is given in Fig. 10.2. Values of parameter estimation are given in Table B.1. The estimation provides path components at two positions, the LOS and the side reflection as shown in Fig. 10.1(a). It shows that the LOS path has the strongest power as expected.

Comparison between the reconstructed data and the measurement data is presented in Fig. 10.3.



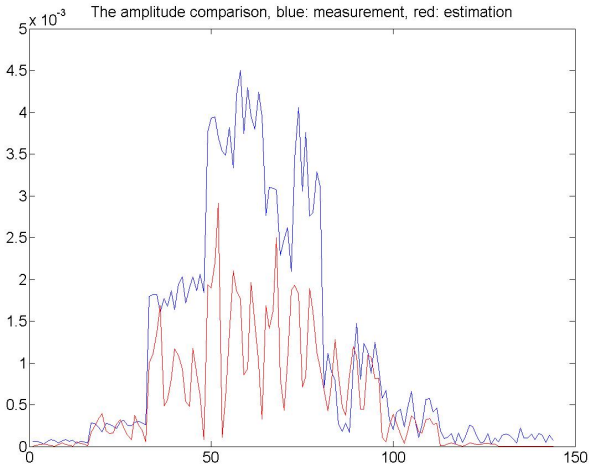
(a) In 3-D view

(b) In Azimuth-Delay view

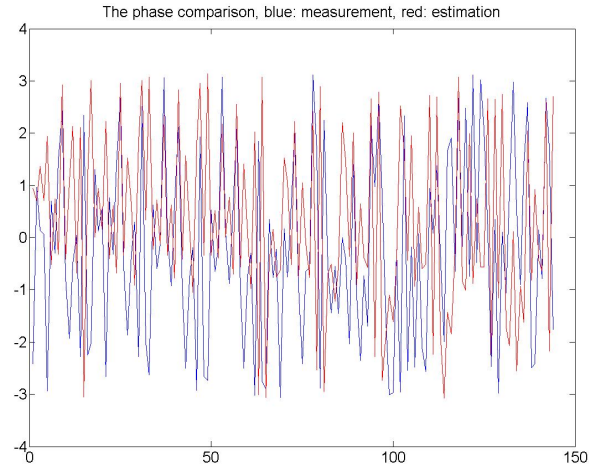
(c) In Elevation-Delay view

(d) In Azimuth-Elevation view

Figure 10.2: The path components estimation using the joint FB5 Gaussian distribution. A SIMO system with all even index Rx antennas and T_1 receive antenna is used for the time invariant scenario. The color of the figure represent the signal power in dB of each path component.



(a) Amplitude comparison between the measurement(blue) and estimation(red)



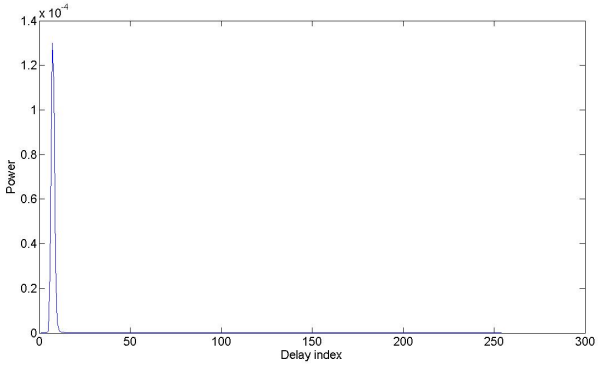
(b) Phase comparison between the measurement(blue) and estimation(red)

Figure 10.3: Estimation results comparison for the SIMO system in the time invariant scenario.

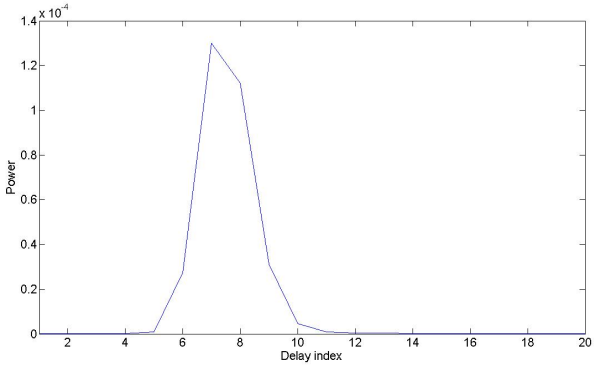
Estimation results for MISO system

The MISO system is extracted with the all Tx antennas and R_2 receive antenna. Notice the Tx antenna array is a omnidirectional antenna array, the effective radiation pattern for azimuth is $[-180°, 180°)$.

The power delay profile is given in Fig. 10.4. From Fig. 10.4(b), delay τ_4 to τ_{12} are selected for the estimation.



(a) Global view



(b) Local view

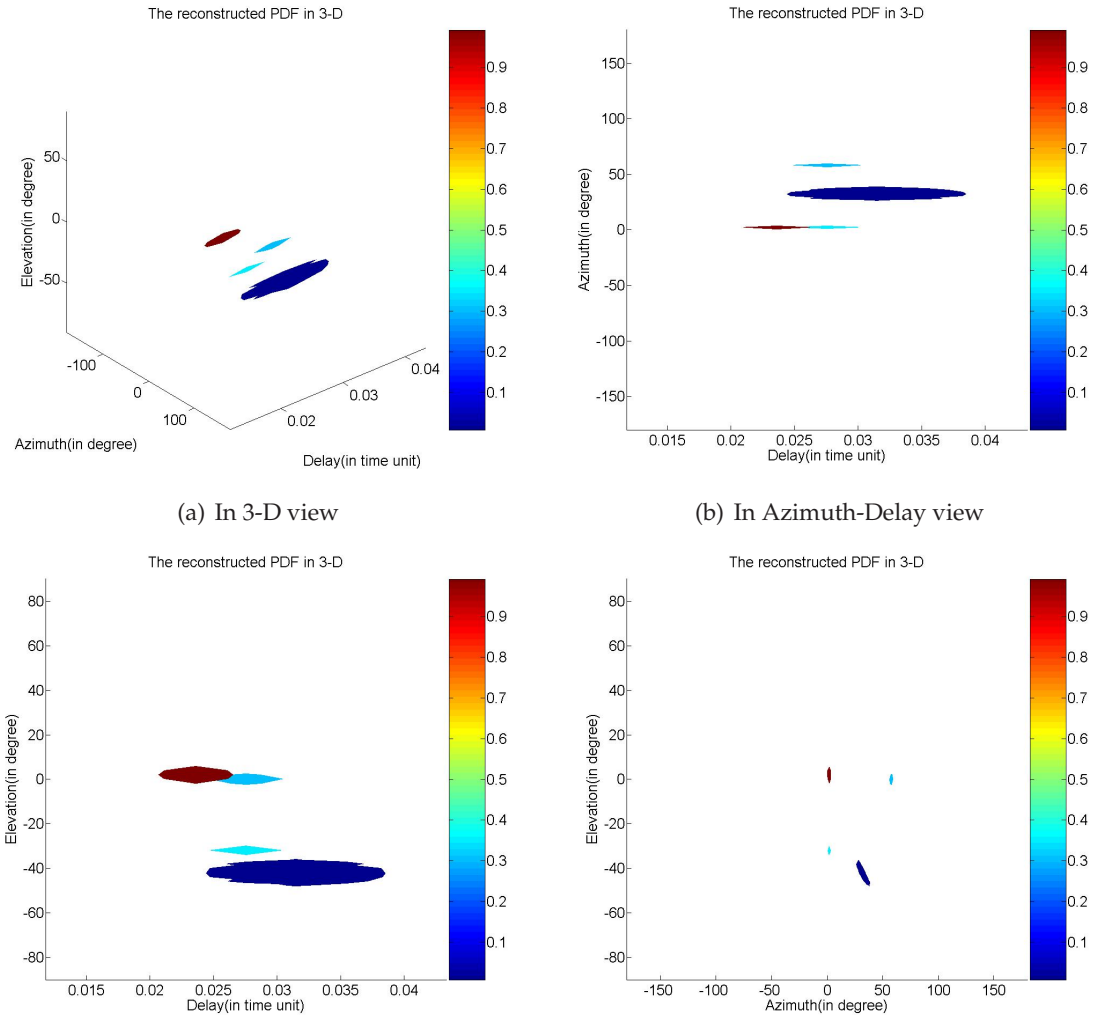
Figure 10.4: The delay power profile of time invariant MISO system.

The estimation limits the maximum number of path component estimation to 5. The sequential path initialization method is adopted.

The estimation results for the SIMO system is given in Fig. 10.5. Values of parameter estimation are given in Table B.2.

Fig. 10.5 provides the first 4 path components estimation. The last and the weakest path component is overlapped with the strongest one. It is due to the estimation error and noise made by the residual data from the initialization.

The estimation result shows 3 dominant path components with the strongest corresponding to the LOS. The other two path component can be viewed consistent with the true environment where one path component is the reflection from the floor and and the other one comes from the horizontal reflection. The delay estimation is also consistent with the environment where the LOS is estimated to be the first to arrive at the receive side.



(a) In 3-D view

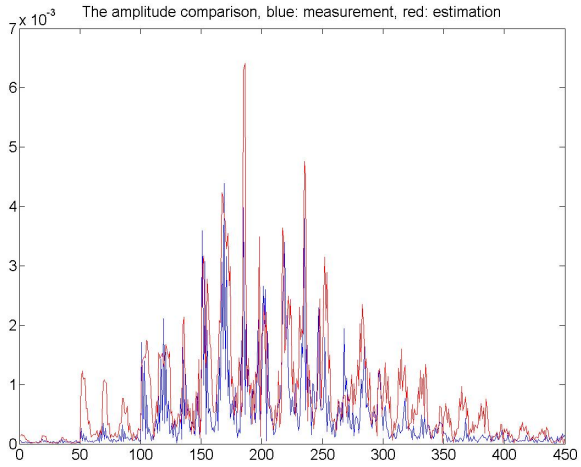
(b) In Azimuth-Delay view

(c) In Elevation-Delay view

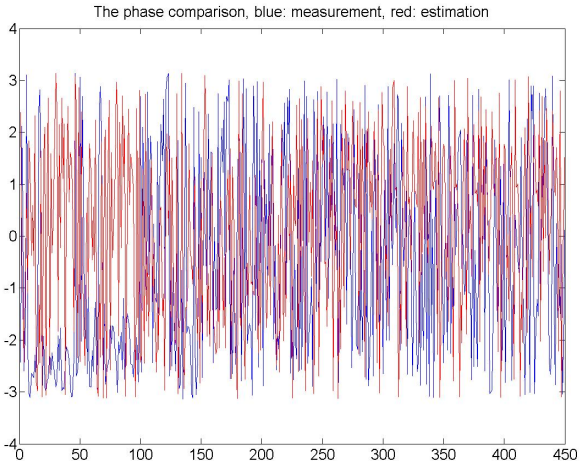
(d) In Azimuth-Elevation view

Figure 10.5: The path components estimation using the joint FB5 Gaussian distribution. A MISO system with all Tx antennas and R_2 receive antenna is used for the time invariant scenario. The color of the figure represent the signal power in dB of each path component.

Comparison between the reconstructed data and the measurement data is presented in Fig. 10.6.



(a) Amplitude comparison between the measurement(blue) and estimation(red)



(b) Phase comparison between the measurement(blue) and estimation(red)

Figure 10.6: Estimation results comparison for the MISO system in the time invariant scenario

10.2 Discussion on model mismatch

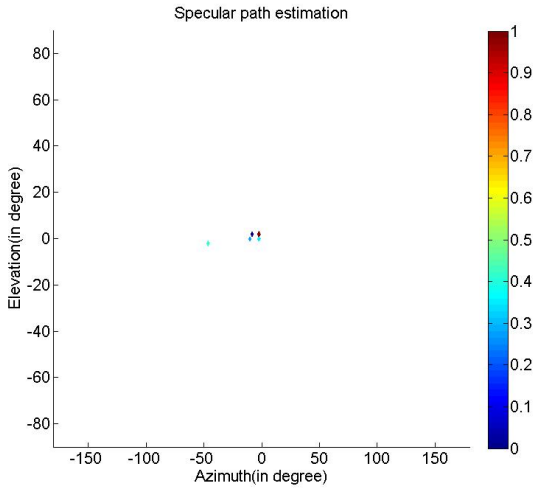
To improve the estimation performance, the model mismatch is analyzed. It is not certain that the proposed identical phase distribution for one path component is fully consistent with the true environment. This may cause estimation errors especially when the estimated paths have large power spread.

A complete estimation of the phase distribution for each path component may improve the performance of the algorithm. If no models are proposed for phase distribution, the

first consideration is to estimate the specific value of the phase at every point in one path component.

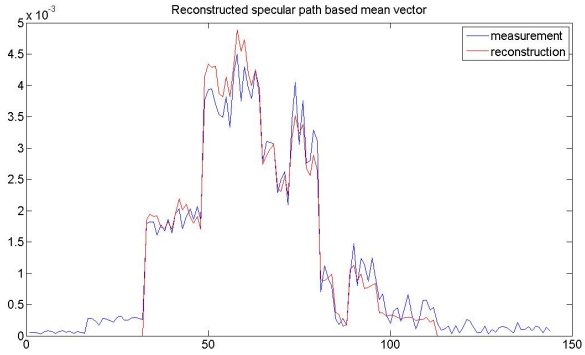
Another solution to estimate the complete phase distribution it to use the specular path model instead. Without the requirement to estimate the power spread, the estimation can be much faster. If multiple specular paths are estimated as within one cluster, this can still be viewed as the effect of dispersive path component model.

One attempt to estimate the SIMO system mentioned in Section 10.1.2 using specular path component model is given in Fig. 10.8 where 5 specular paths are estimated. The result showed the improvement of the estimation. Notice the clustering of a couple of specular paths around the position with azimuth $0°$ and elevation $0°$, it may indicate the existence of a dispersive path component with specific phase distribution.



(a) The estimated channel response mean vector

Figure 10.7: The specular path estimation in azimuth-elevation view



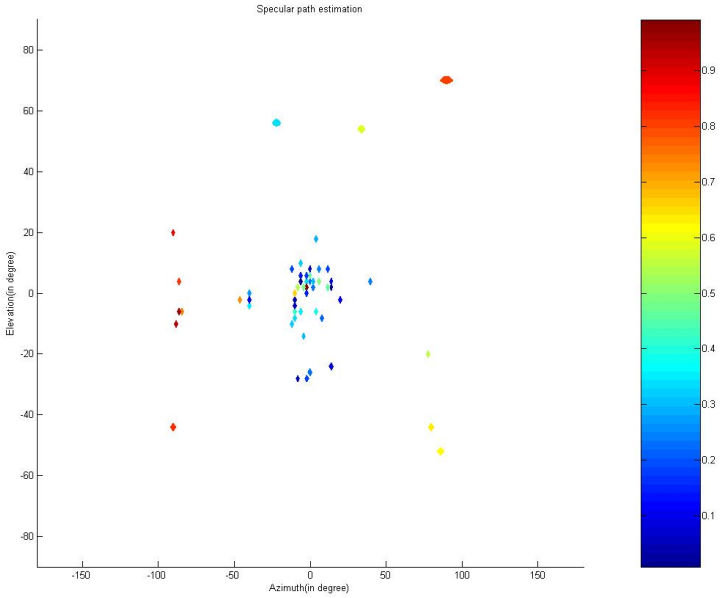
(a) Specular path estimation

Figure 10.8: The 5 specular paths estimation in the SIMO system mentioned in Section 10.1.2. The color of the figure represent the signal power in dB of each path component.

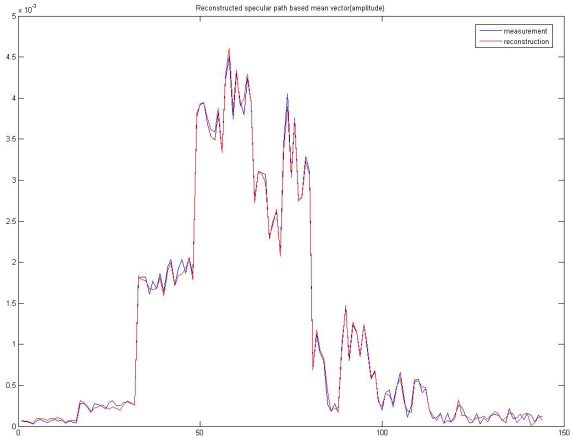
The performance of the estimation using specular path component model is supposed to get improved as more specular path components estimation are included. In Fig. 10.9, 60

specular paths are taken into estimation for the SIMO system mentioned in Section 10.1.2 and the reconstructed data of the estimation shows quite close to the measurement in both amplitude and phase.

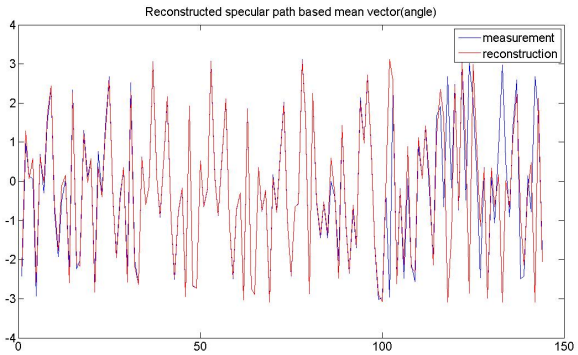
However, even when the massive number of specular path components may produce a small residual error, it is not guaranteed that the estimation results can be fully trusted. The specular path models can also be affected by the estimation errors. In Fig. 10.9, it can be seen that the estimation of specular path components contains path component that are far away from the environment setting, such as one specular path component around position of azimuth -20° and elevation 60°.



(a) The specular paths estimation



(b) The reconstructed channel response mean vector amplitude



(c) The reconstructed channel response mean vector phase

Figure 10.9: The 60 specular paths estimation in the SIMO system mentioned in Section 10.1.2. The color of the figure represent the signal power in dB of each path component.

11 Conclusion and Summary

11.1 Summary of the project

In this project a continuous distributed path component model was used to describe the radio propagation channel. An analysis and discussion were undertaken in both time variant and time invariant scenarios and the WSSUS assumption for the path components was used in order to analyze the characteristics of each path component where multi-path components are present.

To model the continuous distribution of path components, we introduced the maximum entropy modeling to describe the power distribution of each path component. In the time invariant scenario, a phase distribution of the path component was also needed and it was modeled as a constant value for each path component.

As estimation method the SAGE algorithm was applied to cope with the estimation of multi-path components in the two scenarios. The algorithm is an approximation to the global maximum likelihood, but with faster convergence. The SAGE works on the estimated data space of single path components (denoted as a hidden data space) and the performance is inherently sensitive of its design. In the computation of the hidden data space, β was introduced as a parameter to control how much noise and estimation error that was added into the hidden data space.

The performance of the estimation algorithm was evaluated in a synthetic environment and with measurement data. The estimation results in the synthetic environment were reliable and accurate. Observations and a discussion on the model mismatch in the measurement data was presented and an effort to minimize it was carried out. Also, the computational complexity of the algorithm was reduced. A LSE estimator for the time variant scenario (with the cross-correlation in delay omitted) showed promising performance for the measurements data in the environment TxR11. Estimation results of the specular path model in the time invariant scenario showed better performance than the dispersive path model. We believe a more rigorous phase model to describe the phase of the dispersive path components will improve that performance.

11.2 Discussion and proposed improvements

In both the time variant and time invariant scenario with measurement data a model mismatch was apparent. The problem is how to find an effective way to minimize the model mismatch as it is inevitable. Some possible improvements in different scenarios will be mentioned:

11.2.1 Time variant scenario

For the time variant scenario the non-WSSUS scatterer model is believed to better describe the true radio propagation channel. Usually, the WSSUS scatterer model is regarded as a simplified approximation of the real scatterers. The WSSUS assumes that different dispersive path components are uncorrelated. Correlation among the path components can be caused by the finite samples of data or simply by the physical environment of the scatterers.

Another observation from the estimation results showed that some specular paths existed. It is naturally to think of the radio channel propagation as a mixture of specular path components and dispersive path components. The specular path components are just a extreme form of continuous distributed path component (the spread is 0), and the existence of specular path components should not, to some extend, deteriorate the performance of the algorithm, but only decrease the convergence rate.

The model mismatch can also be caused by the skewness of the power distribution. A seriously skewed power distribution will cause a dominant model mismatch as well. This kind of model mismatch can be removed by using several distributed path components to model a skewed path component.

A solution to alleviate the model mismatch for the time variant scenario is to replace the MLE method with LSE method as mentioned before. The LSE is a fitting estimator. The LSE is able to capture useful information from measurement data with dominant model mismatch. On the other hand, the MLE always take the entire measurement data to estimate. With the current proposed model and available measurement data, we believe that the LSE works better than the MLE in the measurement data with model mismatch.

The discussion on the LSE also showed its limitation to give a reliable estimation that can be "connected" to the true physical environment. The LSE is sensitive to the residual data (estimation errors and noise) which may introduce false existence of path components. The reliability of LSE is primary determined by the initialization of the SAGE. For a high resolution antenna array, the Bartlett spectrum is viewed as a trustable estimation of the power spectrum. The combination of the Bartlett spectrum in the initialization and the LSE in the later estimation can then be considered as a trustable estimation of the radio channel response in time variant scenario.

11.2.2 Time invariant scenario

For the time invariant scenarios, the identical value of the phase distribution in one path component may be quite different from the true condition. So far there is no further discussion on the phase distribution of the path component in time invariant scenario. One suggested improvement is to provide a specific description of the phase at different direction and different delay. The complexity of this specific phase description can increase the computational cost.

Another suggestion to solve the phase estimation is to replace the distributed path component modeling with specular path modeling. The attempts of using specular path component for the estimation showed some improvement for the estimation though one should still be aware of the possible estimation of false "images" from the residual data.

11.3 Outlooks

So far we have focused on a solution to estimate the radio propagation channel in MISO and SIMO systems. The radio propagation channel is a complex process that contains more information than just direction, delay, doppler frequency and polarization. The experience in this project can be utilized into the radio channel propagation into MIMO system. The joint power distribution of DoA and DoD will become quite complex if the FB5 distribution for each power direction distribution is still applied. To this point, an approximation for this joint direction distribution using a joint Gaussian distribution can still become an effective candidate to solve the distribution problem. The phase distribution for deterministic path component in time invariant scenario should be investigated to obtain an accurate description of it.

Improvements to further reduce the model mismatch are required. The non-WSSUS condition, skewness of the power distribution, mixture of deterministic path components and random path components are probable reasons to cause a model mismatch.

It is worth exploring a mixture model for the non-WSSUS and WSSUS conditions of the path component.

The current measurement data processing for the estimation algorithm is heavy and quite time consuming. Furthermore, the computational limits of MATLAB^{TM}restricts some operations to be performed in the algorithms (determinant computation). For every parameter set of one path component, the sequential parameter estimation is adopted. The ideal joint estimation is impractical due to the workload. It is another challenge of future research to optimize the algorithm design and programming. In a real-time application, the speed of the programme is critical. For data analyzing purpose, speed of the programme is still an important benchmark. It is necessary to find a balance between accuracy and speed of the algorithm.

Formula Derivation

A

A.1 Derivation of $E[\boldsymbol{H}]$ and $E[\boldsymbol{H}\boldsymbol{H}^H]$ in the time variant scenario

A.1.1 Derivation of $E[\boldsymbol{H}]$

$$E[\boldsymbol{H}] = \begin{bmatrix} E[\boldsymbol{h}(\tau_1)] \\ E[\boldsymbol{h}(\tau_2)] \\ \vdots \\ E[\boldsymbol{h}(\tau_k)] \\ \vdots \\ E[\boldsymbol{h}(\tau_K)] \end{bmatrix} \tag{A.1}$$

The $\boldsymbol{h}(\tau_k)$ follows the model given in Eq. (4.4).

For each sub-vector $E[\boldsymbol{h}(\tau_k)]$ in $E[\boldsymbol{H}]$:

$$E[\boldsymbol{h}(\tau_k)] = \sum_{l=1}^{L} \int_{\mathcal{S},\tau} E[a_l(\boldsymbol{\Omega},\tau)]\boldsymbol{C}(\boldsymbol{\Omega})\text{Rss}(\tau-\tau_k)d\boldsymbol{\Omega}d\tau + \int_{t_n}^{t_n+T} E[\boldsymbol{w}(t)]s^*(t-\tau_k)dt \tag{A.2}$$

$a_l(\boldsymbol{\Omega},\tau)$ is assumed to have phase uniformly distributed and $\boldsymbol{w}(t)$ is a zero mean Gaussian noise. Therefore,

$E[a_l(\boldsymbol{\Omega},\tau)] = 0$ and $E[\boldsymbol{w}(t)] = 0$.

The mean of $\boldsymbol{H}$ is zero:

$$E[\boldsymbol{H}] = 0. \tag{A.3}$$

A.1.2 Derivation of $E[\boldsymbol{H}\boldsymbol{H}^H]$

$$E[\boldsymbol{H}\boldsymbol{H}^H] = \begin{bmatrix} E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_1)] & E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_2)] & \dots & & E[\boldsymbol{h}(\tau_1)\boldsymbol{h}^H(\tau_K)] \\ \vdots & & \ddots & & \vdots \\ & & & E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)] & \\ E[\boldsymbol{h}(\tau_K)\boldsymbol{h}^H(\tau_1)] & & & & E[\boldsymbol{h}(\tau_K)\boldsymbol{h}^H(\tau_K)] \end{bmatrix} \tag{A.4}$$

with each sub-matrix $E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)]$:

$$\begin{aligned} & E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)] \\ =& E[\sum_{l=1}^{L}(\int_{\mathcal{S},\tau} a_l(\boldsymbol{\Omega},\tau_i)\boldsymbol{C}(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau_i)d\boldsymbol{\Omega}d\tau + \int_{t_n}^{t_n+T}\boldsymbol{w}(t)s^*(t-\tau_i)dt)\cdot \\ & (\sum_{l=1}^{L}\int_{\mathcal{S},\tau} a_l^*(\boldsymbol{\Omega},\tau)\boldsymbol{C}^H(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau_j)d\boldsymbol{\Omega}d\tau + \int_{t_n}^{t_n+T}\boldsymbol{w}^H(t)s(t-\tau_j)dt)] \\ =& E[(\sum_{l=1}^{L}\int_{\mathcal{S},\tau} a_l(\boldsymbol{\Omega},\tau)\boldsymbol{C}(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau_i)d\boldsymbol{\Omega}d\tau)(\sum_{l=1}^{L}\int_{\mathcal{S},\tau} a_l^*(\boldsymbol{\Omega},\tau)\boldsymbol{C}^H(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau_j)d\boldsymbol{\Omega}d\tau)] \\ & + E[(\sum_{l=1}^{L}\int_{\mathcal{S},\tau} a_l(\boldsymbol{\Omega},\tau)\boldsymbol{C}(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau_i)d\boldsymbol{\Omega}d\tau)(\int_{t_n}^{t_n+T}\boldsymbol{w}^H(t)s(t-\tau_j)dt)] \\ & + E[(\sum_{l=1}^{L}\int_{\mathcal{S},\tau} a_l(\boldsymbol{\Omega},\tau)\boldsymbol{C}(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau_j)d\boldsymbol{\Omega}d\tau)(\int_{t_n}^{t_n+T}\boldsymbol{w}^H(t)s(t-\tau_i)dt)] \\ & + E[(\int_{t_n}^{t_n+T}\boldsymbol{w}^H(t)s(t-\tau_j)dt)(\int_{t_n}^{t_n+T}\boldsymbol{w}^H(t)s(t-\tau_j)dt)]. \end{aligned} \tag{A.5}$$

Component

$E[(\int_{\mathcal{S},\tau} a(\boldsymbol{\Omega},\tau_i)\boldsymbol{C}(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau_i)d\boldsymbol{\Omega}d\tau)(\int_{t_n}^{t_n+T}\boldsymbol{w}^H(t)s(t-\tau_j)dt)] = 0$

$E[(\int_{\mathcal{S},\tau} a(\boldsymbol{\Omega},\tau_j)\boldsymbol{C}(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau_j)d\boldsymbol{\Omega}d\tau)(\int_{t_n}^{t_n+T}\boldsymbol{w}^H(t)s(t-\tau_i)dt)] = 0$

are given by the assumption that $a(\boldsymbol{\Omega},\tau)$ and $\boldsymbol{w}(t)$ are uncorrelated.

By the assumption in Eq. (4.3) that different path components are uncorrelated.

$$E[a_i(\boldsymbol{\Omega}_i,\tau_i)a_j(\boldsymbol{\Omega}_j,\tau_j)] = 0.$$

$$E[(\sum_{l=1}^{L}\int_{\mathcal{S},T} a_l(\boldsymbol{\Omega},\tau)\boldsymbol{C}(\boldsymbol{\Omega})\text{Rss}(\tau-\tau_i)d\boldsymbol{\Omega}d\tau)(\sum_{l=1}^{L}\int_{\mathcal{S},T} a^*(\boldsymbol{\Omega},\tau)\boldsymbol{C}^H(\boldsymbol{\Omega})\text{Rss}(\tau-\tau_j)d\boldsymbol{\Omega}d\tau)]$$
$$=\sum_{l=1}^{L}\int_{\mathcal{S},T}\int_{\mathcal{S},T} E[a_l(\boldsymbol{\Omega}_1,\tau_1)a_l^*(\boldsymbol{\Omega}_2,\tau_2)]\boldsymbol{C}(\boldsymbol{\Omega}_1)\boldsymbol{C}^H(\boldsymbol{\Omega}_2)\text{Rss}(\tau_1-\tau_i)\text{Rss}(\tau_2-\tau_j)d\boldsymbol{\Omega}_1\boldsymbol{\Omega}_2 d\tau_1 d\tau_2 \tag{A.6}$$

The complex amplitude $a_l(\boldsymbol{\Omega},\tau)$ is considered as uncorrelated in time and direction.

$$E[a_l(\boldsymbol{\Omega}_1,\tau_1)a_l^*(\boldsymbol{\Omega}_2,\tau_2)] = E[a_l(\boldsymbol{\Omega}_1,\tau_2)a_l^*(\boldsymbol{\Omega}_2,\tau_2)]\delta(\Omega_1-\Omega_2)\delta(\tau_1-\tau_2)$$

Therefore,

$$E[(\sum_{l=1}^{L}\int_{\mathcal{S},T} a_l(\boldsymbol{\Omega},\tau)\boldsymbol{C}(\boldsymbol{\Omega})\text{Rss}(\tau-\tau_i)d\boldsymbol{\Omega}d\tau)(\sum_{l=1}^{L}\int_{\mathcal{S},T} a_l^*(\boldsymbol{\Omega},\tau)\boldsymbol{C}^H(\boldsymbol{\Omega})\text{Rss}(\tau-\tau_j)d\boldsymbol{\Omega}d\tau)]$$
$$=\sum_{l=1}^{L}\int_{\mathcal{S},T} E[a_l(\boldsymbol{\Omega},\tau)a_l^*(\boldsymbol{\Omega},\tau)]\boldsymbol{C}(\boldsymbol{\Omega})\boldsymbol{C}^H(\boldsymbol{\Omega})\text{Rss}(\tau-\tau_i)\text{Rss}(\tau-\tau_j)d\boldsymbol{\Omega}d\tau \tag{A.7}$$

Component $E[(\int_{t_n}^{t_n+T}\boldsymbol{w}^H(t)s(t-\tau_j)dt)(\int_{t_n}^{t_n+T}\boldsymbol{w}^H(t)s(t-\tau_j)dt)]$ is expanded as:

$$E[(\int_{t_n}^{t_n+T}\boldsymbol{w}(t)s(t-\tau_j)dt)(\int_{t_n}^{t_n+T}\boldsymbol{w}^H(t)s^*(t-\tau_j)dt)]$$
$$=\int_{t_n}^{t_n+T}\int_{t_n}^{t_n+T} E[\boldsymbol{w}(t_1)\boldsymbol{w}^H(t_2)]s(t_1-\tau_j)s^*(t_2-\tau_j)dt_1dt_2$$

As given $\boldsymbol{w}(t)$ is a white complex Gaussian noise that $E[\boldsymbol{w}(t_1)\boldsymbol{w}^H(t_2)] = \sigma_n^2\boldsymbol{I}\delta(t_1-t_2)$

Therefore:

$$E[\boldsymbol{w}(t_1)\boldsymbol{w}^H(t_2)] = \sigma_n^2\boldsymbol{I}\int_{t_n}^{t_n+T} s(t-\tau_j)s^*(t-\tau_j)dt = \sigma_n^2\boldsymbol{I}\text{Rss}(\tau_i-\tau_j) \tag{A.8}$$

Replace $E[a(\boldsymbol{\Omega},\tau)a^*(\boldsymbol{\Omega},\tau)] = Pf(\boldsymbol{\Omega},\tau)$ as a continuous power distribution of the signal. Finally, $E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)]$ is expressed as:

$$E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)] = \sum_{l=1}^{L}\int_{\mathcal{S},T} P_l f_l(\boldsymbol{\Omega},\tau)\boldsymbol{C}(\boldsymbol{\Omega})\boldsymbol{C}^H(\boldsymbol{\Omega})\text{Rss}(\tau-\tau_i)\text{Rss}(\tau-\tau_j)d\boldsymbol{\Omega}d\tau + \sigma_n^2\boldsymbol{I}\text{Rss}(\tau_i-\tau_j) \tag{A.9}$$

A.2 Derivation of $E[\boldsymbol{H}]$ and $E[\boldsymbol{H}\boldsymbol{H}^H]$ in the time invariant scenario

The formation of $\boldsymbol{H}$ in the time invariant scenario has the same structure as in the time variant scenario.

A.2.1 Derivation of $E[\boldsymbol{H}]$

$$E[\boldsymbol{H}] = \begin{bmatrix} E[\boldsymbol{h}(\tau_1)] \\ E[\boldsymbol{h}(\tau_2)] \\ \vdots \\ E[\boldsymbol{h}(\tau_k)] \\ \vdots \\ E[\boldsymbol{h}(\tau_K)] \end{bmatrix}, \tag{A.10}$$

where $\boldsymbol{h}(\tau_k)$ still follows the model in Eq. (4.4).

The sub-vector $E[\boldsymbol{h}(\tau_k)]$:

$$E[\boldsymbol{h}(\tau_k)] = \sum_{l=1}^{L} \int_{\mathcal{S},T} E[a_l(\boldsymbol{\Omega},\tau)]\boldsymbol{C}(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau_k)d\boldsymbol{\Omega}d\tau + \int_{t_n}^{t_n+T} E[\boldsymbol{w}(t)]s^*(t-\tau_k).dt \tag{A.11}$$

$E[a_l(\boldsymbol{\Omega},\tau)] = a_l(\boldsymbol{\Omega},\tau)$ and $E[w(t)] = 0$ in time invariant scenario.

Thereby the sub-vector $E[\boldsymbol{h}(\tau_k)]$ is given as:

$$E[\boldsymbol{h}(\tau_k)] = \sum_{l=1}^{L} \int_{\mathcal{S},T} a_l(\boldsymbol{\Omega},\tau)\boldsymbol{C}(\boldsymbol{\Omega})\mathrm{Rss}(\tau-\tau_k)d\boldsymbol{\Omega}d\tau \tag{A.12}$$

A.2.2 Derivation of $E[\boldsymbol{H}\boldsymbol{H}^H]$

$$E[\boldsymbol{H}\boldsymbol{H}^H] = \begin{bmatrix} E[(\boldsymbol{h}(\tau_1)-E[\boldsymbol{h}(\tau_1)])(\boldsymbol{h}^H(\tau_1)-E[\boldsymbol{h}(\tau_1)])] & \dots & E[(\boldsymbol{h}(\tau_1)-E[\boldsymbol{h}(\tau_1)])(\boldsymbol{h}^H(\tau_K)-E[\boldsymbol{h}(\tau_K)])] \\ \vdots & \ddots & \vdots \\ E[(\boldsymbol{h}(\tau_K)-E[\boldsymbol{h}(\tau_K)])(\boldsymbol{h}^H(\tau_1)-E[\boldsymbol{h}(\tau_1)])] & & E[(\boldsymbol{h}(\tau_K)-E[\boldsymbol{h}(\tau_K)])(\boldsymbol{h}^H(\tau_K)-E[\boldsymbol{h}(\tau_K)])] \end{bmatrix} \tag{A.13}$$

with each sub-matrix $E[\boldsymbol{h}(\tau_i)\boldsymbol{h}^H(\tau_j)]$ given as:

$$
\begin{aligned}
&E[(\boldsymbol{h}(\tau_i) - E[\boldsymbol{h}(\tau_i)])(\boldsymbol{h}^H(\tau_j) - E[\boldsymbol{h}(\tau_j)])] \\
&= E[(\int_{t_n}^{t_n+T} \boldsymbol{w}(t)s^*(t-\tau_j)dt)(\int_{t_n}^{t_n+T} \boldsymbol{w}^*(t)s^*(t-\tau_j)dt)] \\
&= \sigma_w^2 \boldsymbol{I}\text{Rss}(\tau_i - \tau_j)
\end{aligned} \tag{A.14}
$$

A.3 Hidden data expectation in the time variant scenario

Take the representation of $\Sigma_{\boldsymbol{s}_l\boldsymbol{s}_l}$ and $\Sigma_{\boldsymbol{ww}}$ in Eq. (4.9)

Let $\Sigma_{\boldsymbol{H}_l} = \Sigma_{\boldsymbol{s}_l\boldsymbol{s}_l} + \beta\Sigma_{\boldsymbol{ww}}$ denote the covariance matrix of the channel impulse response for the lth path component. Let $E[\boldsymbol{H}_l]$ denote the mean for the channel impulse response for the lth path component.

Let $\Sigma_{\boldsymbol{H}} = \sum_{l=1}^{L} \Sigma_{\boldsymbol{s}_l\boldsymbol{s}_l} + \Sigma_{\boldsymbol{ww}}$ denote the covariance matrix for all the multi-path component. Let $E[\boldsymbol{H}]$ denote the mean of the all multi-path component.

Let $\overline{\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^H} = \frac{1}{N}\sum_{n=1}^{N} \hat{\boldsymbol{H}}_n\hat{\boldsymbol{H}}_n^H$ represent the observed covariance matrix for all the multi-path components.

From A.1 it is known $E[\boldsymbol{H}_l] = 0$, $E[\boldsymbol{H}] = 0$.

Let $\boldsymbol{R}^j$ represent an estimation of matrix $\boldsymbol{R}$ at iteration j

The $\boldsymbol{H}_l$ and $\boldsymbol{H}$ are assumed to be joint Gaussian. Then the hidden data expectation in time variant scenario becomes the condition expectation of the joint Gaussian covariance matrix.

From [28][p.132],conditional expectation of the covariance matrix $\Sigma_{\boldsymbol{H}_i|\hat{\boldsymbol{H}},\boldsymbol{\theta}}$ is given as following.

$$
\Sigma_{\boldsymbol{H}_l|\hat{\boldsymbol{H}},\boldsymbol{\theta}} = \Sigma_{\boldsymbol{H}_l|\boldsymbol{\theta}} - \Sigma_{\boldsymbol{H}_l\boldsymbol{H}|\boldsymbol{\theta}}\Sigma_{\boldsymbol{H}|\boldsymbol{\theta}}^{-1}\Sigma_{\boldsymbol{H}\boldsymbol{H}_l|\boldsymbol{\theta}} + E[\boldsymbol{H}_l|\boldsymbol{H},\boldsymbol{\theta}]E[\boldsymbol{H}_l|\boldsymbol{H},\boldsymbol{\theta}] \tag{A.15}
$$

$$
E[\boldsymbol{H}_l|\boldsymbol{H},\boldsymbol{\theta}] = E[\boldsymbol{H}_l] + \Sigma_{\boldsymbol{H}_l\boldsymbol{H}|\boldsymbol{\theta}}\Sigma_{\boldsymbol{H}|\boldsymbol{\theta}}^{-1}(\hat{\boldsymbol{H}}_n - E[\boldsymbol{H}]) = \Sigma_{\boldsymbol{H}_l\boldsymbol{H}|\boldsymbol{\theta}}\Sigma_{\boldsymbol{H}|\boldsymbol{\theta}}^{-1}\hat{\boldsymbol{H}}_n \tag{A.16}
$$

$$
\Sigma_{\boldsymbol{H}_l|\boldsymbol{\theta}} = \Sigma_{\boldsymbol{H}_l\boldsymbol{H}|\boldsymbol{\theta}} = \Sigma_{\boldsymbol{H}\boldsymbol{H}_l|\boldsymbol{\theta}} \tag{A.17}
$$

Therefore,

$$
\Sigma_{\boldsymbol{H}_l|\hat{\boldsymbol{H}},\boldsymbol{\theta}} = \Sigma_{\boldsymbol{H}_l|\boldsymbol{\theta}} - \Sigma_{\boldsymbol{H}_l|\boldsymbol{\theta}}\Sigma_{\boldsymbol{H}|\boldsymbol{\theta}}^{-1}\Sigma_{\boldsymbol{H}_l|\boldsymbol{\theta}} + \Sigma_{\boldsymbol{H}_l|\boldsymbol{\theta}}\Sigma_{\boldsymbol{H}|\boldsymbol{\theta}}^{-1}\overline{\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^H}\Sigma_{\boldsymbol{H}|\boldsymbol{\theta}}^{-1}\Sigma_{\boldsymbol{H}_i|\boldsymbol{\theta}}. \tag{A.18}
$$

Replace the estimated covariance matrix with the estimation results of estimation iteration $j-1$, the conditional covariance matrix for the estimation iteration j for the lth scatter is given:

$$
\Sigma_{\boldsymbol{H}_l^j} = \Sigma_{\boldsymbol{H}_l^{j-1}} - \Sigma_{\boldsymbol{H}_l^{j-1}}\Sigma_{\boldsymbol{H}^{j-1}}^{-1}\Sigma_{\boldsymbol{H}_l^{j-1}} + \Sigma_{\boldsymbol{H}_l^{j-1}}\Sigma_{\boldsymbol{H}^{j-1}}^{-1}\overline{\hat{\boldsymbol{H}}\hat{\boldsymbol{H}}^H}\Sigma_{\boldsymbol{H}^{j-1}}^{-1}\Sigma_{\boldsymbol{H}_l^{j-1}} \tag{A.19}
$$

A.4 MLE principles

The general principles of Maximum Likelihood Estimator(MLE) can be described as follow.

Let $F(x;\boldsymbol{\theta})$ indicate the distribution of the random variable X given the parameter $\boldsymbol{\theta}$. Thus, if $X_1, X_2, \ldots, X_n$ are i.i.d.(independent identical distributed) measurement of X, then $f(x_1, x_2, \ldots, x_n; \boldsymbol{\theta})$ is the joint pdf of $X_1, \ldots, X_n$ given $\boldsymbol{\theta}$, and because of independency it can be written:

$$f(x_1, x_2, \ldots, x_n; \boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i; \boldsymbol{\theta})$$

Here $f(x_1, x_2, \ldots, x_n; \boldsymbol{\theta})$ are called a likelihood function and is denoted by:

$$L(\boldsymbol{\theta}) = f(x_1, x_2, \ldots, x_n; \boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i; \boldsymbol{\theta}) \tag{A.20}$$

and it can be considered as a function of the unknown parameter set of $\boldsymbol{\theta}$.

If there is a value $\tilde{\boldsymbol{\theta}}$ that maximizes the likelihood function, it is then called a Maximum Likelihood Estimator(MLE) of $\boldsymbol{\theta}$ expressed as:

$$\tilde{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \; L(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \; \prod_{i=1}^{n} f(x_i; \boldsymbol{\theta})$$

Further, the logarithm of the MLE then it is converted into a Maximum Loglikelihood Estimator.

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \; \ln L(\boldsymbol{\theta}) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \; \ln \left(\prod_{i=1}^{n} f(x_i; \boldsymbol{\theta}) \right) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \; \sum_{i=1}^{n} \ln f(x_i; \boldsymbol{\theta}) \tag{A.21}$$

A.5 MLE in time variant scenario

The $\boldsymbol{H}$ is believed to follow zero mean Gaussian distribution.

$$f(\hat{\boldsymbol{H}}) = \frac{1}{(2\pi)^M |\boldsymbol{\Sigma}(\boldsymbol{\theta})|} \exp\left(-\hat{\boldsymbol{H}}^H \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) \hat{\boldsymbol{H}} \right) \tag{A.22}$$

Let $\tilde{\boldsymbol{\Sigma}}$ denote the estimated covariance matrix.Insert Eq. (A.22) into Eq. (A.21), the MLE is given by:

$$\begin{aligned}
\tilde{\boldsymbol{\theta}} &= \underset{\boldsymbol{\theta}}{\text{argmax}} \ \sum_{i=1}^{n} \ln \left(\frac{1}{(2\pi)^M |\tilde{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\theta}})|} \exp \left(-\hat{\boldsymbol{H}}_i^H \tilde{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\theta}}) \hat{\boldsymbol{H}}_i \right) \right) \\
&= \underset{\boldsymbol{\theta}}{\text{argmax}} \ \sum_{i=1}^{n} \left(\ln \frac{1}{(2\pi)^M |\tilde{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\theta}})|} - \hat{\boldsymbol{H}}_i^H \tilde{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\theta}}) \hat{\boldsymbol{H}}_i \right) \\
&= \underset{\boldsymbol{\theta}}{\text{argmax}} \ \left(-nM \ln 2\pi - n \ln |\tilde{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\theta}})| - \sum_{i=1}^{n} (\hat{\boldsymbol{H}}_i^H \tilde{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\theta}}) \hat{\boldsymbol{H}}_i) \right) \\
&= \underset{\boldsymbol{\theta}}{\text{argmax}} \ \left(- \ln |\tilde{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\theta}})| - \frac{1}{n} \sum_{i=1}^{n} (\hat{\boldsymbol{H}}_i^H \tilde{\boldsymbol{\Sigma}}^{-1}(\hat{\boldsymbol{\theta}}) \hat{\boldsymbol{H}}_i) \right) \\
&= \underset{\boldsymbol{\theta}}{\text{argmax}} \ \left(- \ln |\tilde{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\theta}})| - trace(\tilde{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\Sigma}}) \right)
\end{aligned} \tag{A.23}$$

where $\hat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{H}}_i \hat{\boldsymbol{H}}_i^H$.

A.6 MLE in time invariant scenario

The $\boldsymbol{H}$ is believed to follow Gaussian distribution.

$$f_{\hat{\boldsymbol{H}}} = \frac{1}{(2\pi)^M |\boldsymbol{\Sigma}|} \exp \left(-(\boldsymbol{H} - E[\boldsymbol{H}])^H \boldsymbol{\Sigma}^{-1} (\boldsymbol{H} - E[\boldsymbol{H}]) \right) \tag{A.24}$$

Let $\tilde{\boldsymbol{\Sigma}}$ denote the estimated covariance matrix. Insert Eq. (A.24) into Eq. (A.21), the MLE is given by:

$$\begin{aligned}
\tilde{\boldsymbol{\theta}} &= \underset{\boldsymbol{\theta}}{\text{argmax}} \ \sum_{i=1}^{n} \ln \frac{1}{(2\pi)^M |\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})|} \exp \left(-(\hat{\boldsymbol{H}}_i - E[\boldsymbol{H}])^H \boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\theta}}) (\hat{\boldsymbol{H}}_i - E[\boldsymbol{H}]) \right) \\
&= \underset{\boldsymbol{\theta}}{\text{argmax}} \ \sum_{i=1}^{n} \left(\ln \frac{1}{(2\pi)^M |\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})|} - (\hat{\boldsymbol{H}}_i - E[\boldsymbol{H}])^H \boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\theta}}) (\hat{\boldsymbol{H}}_i - E[\boldsymbol{H}]) \right) \\
&= \underset{\boldsymbol{\theta}}{\text{argmax}} \ - \ln |\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})| - \frac{1}{n} \sum_{i=1}^{n} (\hat{\boldsymbol{H}}_i - E[\boldsymbol{H}])^H \boldsymbol{\Sigma}^{-1}(\hat{\boldsymbol{\theta}}) (\hat{\boldsymbol{H}}_i - E[\boldsymbol{H}]) \\
&= \underset{\boldsymbol{\theta}}{\text{argmax}} \ - \ln |\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})| - \frac{1}{n} \sum_{i=1}^{n} (\hat{\boldsymbol{H}}_i - E[\boldsymbol{H}])^H \sigma_w^{-2} \boldsymbol{R} (\hat{\boldsymbol{H}}_i - E[\boldsymbol{H}]) \\
&= \underset{\boldsymbol{\theta}}{\text{argmax}} \ - \sigma_w^2 \ln |\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})| - \frac{1}{n} \sum_{i=1}^{n} (\hat{\boldsymbol{H}}_i - E[\boldsymbol{H}])^H \boldsymbol{R} (\hat{\boldsymbol{H}}_i - E[\boldsymbol{H}])
\end{aligned} \tag{A.25}$$

A.7 Model mismatch in time variant scenario

The sub covariance matrix for two correlated specular path a_1 and a_2 is given:

$$
\begin{aligned}
&E[\boldsymbol{h}(\upsilon_1)\boldsymbol{h}^H(\upsilon_2)] \\
&= E[\int_{t_n}^{t_n+T} \boldsymbol{y}(t)s(t-\upsilon_1)dt \int_{t_n}^{t_n+T} \boldsymbol{y}^H(t)s^*(t-\upsilon_2)dt] \\
&= E[\int_{t_n}^{t_n+T}\int_{t_n}^{t_n+T} \boldsymbol{y}(t_1)\boldsymbol{y}^H(t_2)s(t_1-\upsilon_1)s^*(t_2-\upsilon_2)dt_1dt_2] \\
&= E[\int_{t_n}^{t_n+T}\int_{t_n}^{t_n+T} [a_1\boldsymbol{c}_1 s(t_1-\tau_1)+a_2\boldsymbol{c}_2 s(t_1-\tau_2)][a_1^*\boldsymbol{c}_1^H s^*(t_2-\tau_1)+a_2^*\boldsymbol{c}_2^H s^*(t_2-\tau_2)]\cdot \\
&\qquad\qquad s(t_1-\upsilon_1)s^*(t_2-\upsilon_2)dt_1dt_2] \\
&= \int_{t_n}^{t_n+T}\int_{t_n}^{t_n+T} [E[a_1a_1^*]\boldsymbol{c}_1\boldsymbol{c}_1^H s(t_1-\tau_1)s^*(t_2-\tau_1)s(t_1-\upsilon_1)s^*(t_2-\upsilon_2)+ \\
&\qquad\qquad E[a_1a_2^*]\boldsymbol{c}_1\boldsymbol{c}_2^H s(t_1-\tau_1)s^*(t_2-\tau_2)s(t_1-\upsilon_1)s^*(t_2-\upsilon_2)+ \\
&\qquad\qquad E[a_2a_1^*]\boldsymbol{c}_2\boldsymbol{c}_1^H s(t_1-\tau_2)s^*(t_2-\tau_1)s(t_1-\upsilon_1)s^*(t_2-\upsilon_2)+ \\
&\qquad\qquad E[a_2a_2^*]\boldsymbol{c}_2\boldsymbol{c}_2^H s(t_1-\tau_2)s^*(t_2-\tau_2)s(t_1-\upsilon_1)s^*(t_2-\upsilon_2)]dt_1dt_2
\end{aligned}
$$

B Estimation Results

Paths	1	2	3	4	5
P	$7.0967 \cdot 10^{-6}$	$3.9773 \cdot 10^{-7}$	$1.3225 \cdot 10^{-6}$	$2.5927 \cdot 10^{-7}$	$8.9704 \cdot 10^{-8}$
$\bar{\theta}[^\circ]$	2	2	0	-2	2
$\bar{\phi}[^\circ]$	-6	-2	-8	-46	-4
$\bar{\tau}$(index)	6	7	8	8	7
κ	5600	3460	1458	5024.8	366.74
ζ	0.14	0.78	0.94	0.94	0.98
$\alpha[^\circ]$	154	140	102	140	160
$\beta[^\circ]$	20	132	-132	-76	-60
η	700	700	700	700	700
B	285120	271524	158400	14434	65885
$\vartheta[^\circ]$	214.54	83.634	72.726	225.45	261.82

Table B.1: The estimation results of the time invariant environment. A SIMO system with all even index Rx antennas and T_1 transmit antenna is used in the time invariant scenario. The selected power distribution model is the joint direction and delay power distribution in Eq. (3.17). The LSE was applied.

Paths	1	2	3	4	5
P	$1.5622 \cdot 10^{-5}$	$2.5044 \cdot 10^{-6}$	$2.2277 \cdot 10^{-6}$	$9.7908 \cdot 10^{-7}$	$4.9591 \cdot 10^{-7}$
$\bar{\theta}[^\circ]$	2	-32	0	-42	2
$\bar{\phi}[^\circ]$	2	2	58	32	-12
$\bar{\tau}$(index)	6	7	7	8	6
κ	2033	103740	2230	1742	162
ζ	0.72	0.98	0.46	0.92	0.98
$\alpha[^\circ]$	40	44	74	8	130
$\beta[^\circ]$	84	-180	96	72	-156
η	700	0	700	700	357
B	550800	4752000	318240	30240	83635
$\vartheta[^\circ]$	113	105.45	196	0	291

Table B.2: The estimation results of the time invariant environment. A MISO system with all Tx antennas and R_2 receive antenna is used for the time invariant scenario. The selected power distribution model is the joint direction and delay power distribution in Eq. (3.17). The LSE was applied.

Paths	1	2	3	4	5	6	7	8	9	10
P	$8.785 \cdot 10^{-8}$	$2.256 \cdot 10^{-8}$	$1.325 \cdot 10^{-8}$	$8.489 \cdot 10^{-9}$	$2.135 \cdot 10^{-8}$	$7.440 \cdot 10^{-9}$	$2.677 \cdot 10^{-8}$	$1.080 \cdot 10^{-8}$	$2.414 \cdot 10^{-9}$	$4.127 \cdot 10^{-9}$
$\bar{\theta}[^\circ]$	0	-2	-10	14	2	-2	4	-6	-36	14
$\bar{\phi}[^\circ]$	-68	-132	-10	-172	120	-136	-68	-34	-100	62
$\bar{\tau}$(index)	22	22	22	23	25	25	24	27	23	23
κ	505.83	84.7	2787.8	10890	54	14641	336.96	97.978	66647	152460
ζ	0.92	0.52	0.86	0.14	0.68	0.44	0.98	0.74	0.98	0.92
$\alpha[^\circ]$	130	56	128	94	170	90	42	22	120	56
$\beta[^\circ]$	-76	16	-64	12	140	12	124	-136	68	-60
η	6855.7	1209.8	24484	89455	3460.5	182310	8229.6	3197.7	229550	76601
B	261620	522720	355450	781470	221760	2316600	200990	104360	790610	95040
$\delta^2_{\boldsymbol{w}}$	$2.7380 \cdot 10^{-12}$									

Table B.3: The estimation results for environment TxR11, MISO system with all Tx antennas and R_9 receive antenna, time variant scenario.The selected power distribution model is the joint direction and delay power distribution in Eq. (3.17). The selective LSE was adopted.

Bibliography

[1] Gerard. 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, 1998.

[2] I. Emre Telatar. "Capacity of Multi-Antenna Gaussian Channels". Technical report, AT & T Bell Lab., 1996.

[3] Vahid Tarokh, Hamid Jafarkhani, and A. R. Calderbank. Space Time Block Codes from Orthogonal Designs. *IEEE Transaction on Information Theory*, 45(5):1456, July 1999.

[4] Gerard J. Foschini. Layered Space-Time Architecture for Wireless Communication in a Fading Environment when using Multi-Element Antennas. *Bell Labs Technical Journal*, 1996.

[5] Gordon L. Stuber, John R. Barry, Steve W. Mclaughlin, Ye Li, Mary Ann Ingram, and Thomas G. Pratt. Broadband MIMO-OFDM Wireless Communications. *Porceedings of the IEEE*, 92(2):271, Feb. 2004.

[6] Arogyaswami Paulraj, Rohit Nabar, and Dhananjay Gore. *Introduction to Space-Time Wireless Communication*. Cambridge University Press, 2003.

[7] Cássio B. Ribeiro, Andreas Richter, and Visa Koivunen. Joint Maximum Likelihood Estimation of Angular and Time-Delay MIMO Propagation Parameters. *IEEE*, 2005.

[8] Xuefeng Yin. *High-Resolution Parameter Estimation for MIMO Channel Sounding*. PhD thesis, Aalborg University, July 2006.

[9] Bernard H. Fleury, Patrik Jourdan, and Andreas Stucki. "High-Resolution Channel Parameter Estimation for MIMO Applications Using the SAGE Algorithm". International Zürich Seminar on Broadband Communication, February 2002.

[10] Xuefeng Yin, Bernard H. Fleury, Troels Pedersen, and Nicolai Czink. "Parametric Characterization and Estimation of Bi-Azimuth Dispersion of Multipath Components". IEEE Workshop on Signal Processing Advances in Wireless Communication, July 2006.

[11] Terence Betlehem, Tharaka A. Lamahewa, and Thushara D. Abhayapala. "Dependence of MIMO System Performance on the Joint Properties of Angular Power". IEEE, July 2006.

[12] Hamid Krim and Mats Viberg. Two Decades of Array Signal Processing Research. *IEEE Signal Processing Magazine*, page 67, July 1996.

[13] Ralph O. Schmidt. "Multiple Emitter Location and Signal Parameter Estimation". *IEEE Transactions on Antennas and Propagation*, AP-34(3):276–280, March 1986.

[14] Richard Roy and Thomas Kailath. ESPRIT-Estimation of Signal Parameters via Rotational Invariance Techniques. *IEEE Transaction on Acoustics, Speech and Signal Processing*, 37(7):984, July 1989.

[15] Todd K. Moon. The Expectation-Maximization Algorithm. 1996.

[16] Xuefeng Yin, Lingfeng Liu, Daniel K. Nielsen, Nicolai Czink, and Bernard H. Fleury. "Characterization of the Azimuth-Elevation Power Spectrum of Individual Path Components". Vienna, Oestrig, Feb. 2007. International ITG-IEEE Workshop on Smart Antennas.

[17] Xuefeng Yin, Lingfeng Liu, Daniel K. Nielsen, Troels Pedersen, and Bernard H. Fleury. "a SAGE Algorithm for Estimation of Direction Power Spectrum of Individual Path Components". December 2007.

[18] Dave Laurenson and Peter Grant. "A Review of Radio Channel Sounding Techniques". Technical report, Institute for Digital Communications, The University of Edinburgh, 2006.

[19] Beata J. Wysocki, Tadeusz A. Wysocki, and Hans-Jurgen Zepernick. "Walsh-Chirp Sequence for Wireless Applications". *Journal of Telecommunications and Information Technology*, 3 2001.

[20] Sathish Chandran. *Adaptive Antennas Arrays*. Springer-Verlag Berlin Heidelberg New York, 2004.

[21] John D. Parsons. *The Mobile Radio Propagation Channel*. John Wiley & Sons Ltd., 2nd edition, 2001.

[22] Kanti V. Mardia. "Statistics of Directional Data". *Journal of the Royal Statistical Society*, 37(3):349–393, 1975.

[23] John T. Kent. "The Fisher-Bingham Distribution on the Sphere". *Journal of the Royal Statistical Society*, 44(1), 1982.

[24] Jeffrey A. Fessler and Alfred O. Hero. "Space-Alternating Generalized Expextation-Maximization Algorithm". *IEEE Transactions on Signal Processing*, 42(10), 1994.

[25] Petre Stoica and Randolph Moses. *Introduction To Spectral Analysis*. Prentice-Hall Inc., 1997.

[26] Petre Stoica, Zhisong Wang, and Jian Li. "Robust Capon Beamforming". *IEEE Signal Processing Letters*, 10(6):172, June 2003.

[27] Nicolai Czink. "Second Milestone Report - MIMO Radio Channel Measurements". Technical report, Institut für Nachrichtentechnik und Hochfrequenztechnik Technische Universität, Wien, 2005.

[28] Sergei M. Prigarin. *Spectral Models of Random Fields in Monte-Carlo Methods*. VSP, 2001.

A SAGE Algorithm for the Estimation of Direction Power Spectrum of Individual Path Components

Xuefeng Yin[1], Lingfeng Liu[1], Daniel K. Nielsen[1], Troels Pedersen[1] and Bernard H. Fleury[1,2]

[1]Department of Electronics Systems, Aalborg University, DK-9220 Aalborg, Denmark
[2]Forschungszentrum Telekommunikation Wien (ftw.), Vienna, Austria

***Abstract*— In this contribution, the Fisher-Bingham-5 (FB_5) probability density function (pdf) is used to model the shape of the direction power spectrum of individual path components in the radio channel response. The FB_5 pdf is selected because among all distributions of direction, the FB_5 distribution maximizes the entropy under the constraints that the first moment of the distribution and its second moment are specified. A SAGE (Space-Alternating Generalized Expectation-maximization) algorithm is derived based on this model for estimation of the parameters characterizing the power spectrum of individual path components. The performance of the SAGE algorithm is evaluated using measurement data. Preliminary results show that the estimated power spectra of individual path components exhibit different ovalness and tilt angle. They are noticeably more concentrated than the corresponding footprints in the power spectrum estimated using the Bartlett beamformer.**

***Index Terms*— Path component, Fisher-Bingham-5 distribution, SAGE algorithm**

I. Introduction

Due to the heterogeneity of the propagation environment, the response of the radio channel is the superposition of a certain number of components. Each component, which we refer to as "path component", is contributed by an electromagnetic wave propagating along a path from the transmitter (Tx) to the receiver (Rx). Along this path, the wave interacts with a certain number of objects called scatterers. Due to the geometrical extent and the nonhomogeneous electromagnetic properties of the scatterers, a path may be dispersive in delay, direction of departure, direction of arrival, polarizations, as well as in Doppler frequency when the environment is time-variant. As a consequence, an individual path component may be spread in these dispersion dimensions. Modeling of these dispersion phenomena is required for the design and optimization of mobile communication systems and thus, the knowledge of the dispersive characteristics of path components is necessary.

In recent years, estimation of the dispersive characteristics of individual path components in multiple dimensions has attracted much attention. Some of the techniques are derived using the assumption that the shape of the power spectrum of individual path components can be described using a probability density function (pdf). In [1], the product of the von-Mises pdf and the exponential pdf is used to describe the shape of the delay–AoA (Azimuth of Arrival) power spectrum. In [2] and [3], the von–Mises–Fisher and Fisher–Bingham–5 (FB_5) pdfs are used to characterize the shape of the AoA–AoD (azimuth of departure) power spectrum and the direction (azimuth and elevation) power spectrum respectively. The shape of the delay–AoA–AoD power spectrum can be described using a 3-variate pdf derived in [4].

In this contribution, we derive a SAGE algorithm which is used to estimate the direction power spectrum of the individual path components. The power spectrum is modeled using the FB_5 pdf. The performance of the SAGE algorithm is evaluated using measurement data.

This contribution is organized as follows. In Section II, a signal model for channel sounding is presented and the characterization of the power spectrum based on the FB_5 pdf is introduced. In Section III, the estimators of the model parameters are derived within the SAGE framework. Section IV shows the experimental results. Finally concluding remarks are stated in Section V.

II. Signal Model

In this contribution, we are interested in the dispersive characteristics of individual path components in direction of arrival (DoA). The channel sounding system considered has a SIMO (single-input multiple-ouput) configuration with a single Tx antenna and a M-element Rx antenna array. The signal model, the characterization method, and the estimation method derived here can be easily modified to handle a MISO (multiple-input single-ouput) channel sounding configuration where dispersion in direction of departure (DoD) is of interest.

We consider narrow-band transmission, which implies that the product of the signal bandwidth times the channel delay spread is much smaller than one. Following the nomenclature in [5], the continuous-time output signal of

This work was jointly supported by the Network of Excellence in Wireless COMmunications (NEWCOM) and Elektrobit Group.

the Rx array in the SIMO system reads

$$\begin{aligned}\boldsymbol{Y}(t) &= \boldsymbol{H}(t)u(t) + \boldsymbol{W}(t) \in \mathbb{C}^M \\ &= \left[\int_{\mathbb{S}_2} \boldsymbol{c}(\boldsymbol{\Omega})h(t;\boldsymbol{\Omega})\mathrm{d}\boldsymbol{\Omega}\right]u(t) + \boldsymbol{W}(t). \quad (1)\end{aligned}$$

The complex vector $\boldsymbol{Y}(t)$ contains the output signals of the Rx array observed at time instance t. The scalar function $u(t)$ denotes the complex envelope of the transmitted sounding signal at time t. The vector $\boldsymbol{H}(t)$ represents the time-variant impulse response of the SIMO system. We assume that $u(t)$ is known to the Rx and that $\int_0^T u(t)u(t)^*\mathrm{d}t = 1$, where $[\cdot]^*$ denotes complex conjugate and T represents the duration of one observation interval. The function $h(t;\boldsymbol{\Omega})$ is the (time-variant) DoA spread function of the propagation channel [5]. Here, $\boldsymbol{\Omega}$ denotes the DoA, which is defined to be a unit vector with initial point anchored at the origin $\mathcal{O}$ of a coordinate system where the Rx array is confined. The end point of $\boldsymbol{\Omega}$ locates at a unit sphere $\mathbb{S}_2$ centered at $\mathcal{O}$. The DoA $\boldsymbol{\Omega}$ is uniquely determined by the spherical coordinates $(\phi,\theta) \in [-\pi,\pi) \times [0,\pi]$ of its end point according to the relation

$$\boldsymbol{\Omega} = [\cos(\phi)\sin(\theta), \sin(\phi)\sin(\theta), \cos(\theta)]^{\mathrm{T}} \quad (2)$$

with $[\cdot]^{\mathrm{T}}$ denoting transposition. The angles ϕ and θ are referred to as the azimuth and elevation of the DoA respectively. The noise $\boldsymbol{W}(t)$ in (1) is a vector-valued circularly symmetric, spatially and temporally white Gaussian process with component spectral height σ_w^2. We assume that σ_w^2 can be measured and therefore is known in advance. The complex vector

$$\boldsymbol{c}(\boldsymbol{\Omega}) \doteq [c_1(\boldsymbol{\Omega}), c_2(\boldsymbol{\Omega}), \ldots, c_M(\boldsymbol{\Omega})]^{\mathrm{T}}$$

arising in (1) is the responses of the Rx array. In a scenario where the electromagnetic energy propagates from the Tx to the Rx via D paths, the DoA spread function $h(t;\boldsymbol{\Omega})$ can be decomposed as

$$h(t;\boldsymbol{\Omega}) = \sum_{d=1}^{D} h_d(t;\boldsymbol{\Omega}). \quad (3)$$

The summand $h_d(t;\boldsymbol{\Omega})$ denotes the dth path component in $h(t;\boldsymbol{\Omega})$.

We assume that the transfer vector $\boldsymbol{H}(t)$ fluctuates over the overall sounding period, but remains constant within individual observation intervals:

$$\boldsymbol{H}(t) \doteq \boldsymbol{H}_n, \quad t \in [t_n, t_n + T).$$

Similarly, the spread functions $h_d(t;\boldsymbol{\Omega}),\ d = 1,\ldots,D$ arising in (3) are constant within individual observation intervals:

$$h_d(t;\boldsymbol{\Omega}) = h_d(t_n;\boldsymbol{\Omega}) \doteq h_{d,n}(\boldsymbol{\Omega}), \quad t \in [t_n, t_n + T). \quad (4)$$

The processes $h_{d,n}(\boldsymbol{\Omega})$, $n \in [1,\ldots,N],\ \ d \in [1,\ldots,D]$ with N denoting the number of observation intervals, are assumed to be uncorrelated complex (zero-mean) orthogonal stochastic measures, i.e.

$$\mathrm{E}[h^*_{d,n}(\boldsymbol{\Omega})h_{d',n'}(\boldsymbol{\Omega}')] = P_d(\boldsymbol{\Omega})\delta_{nn'}\delta_{dd'}\delta(\boldsymbol{\Omega} - \boldsymbol{\Omega}'). \quad (5)$$

Here, $\delta_{(\cdot)}$ and $\delta(\cdot)$ represent the Kronecker delta and the Dirac delta function respectively, and $P_d(\boldsymbol{\Omega}) \doteq \mathrm{E}[|h_{d,n}(\boldsymbol{\Omega})|^2]$ is the direction power spectrum of the dth path component. Identity (5) implies that the spread functions of different individual path components or at different observation intervals are uncorrelated. This scenario is referred to as the *uncorrelated scattering* case in the literature (see e.g. [5]).

The spectrum $P_d(\boldsymbol{\Omega})$ describes the manner the average power of the dth path component is distributed on the unit sphere $\mathbb{S}_2$. We assume

$$P_d(\boldsymbol{\Omega}) = P_d \cdot f_d(\boldsymbol{\Omega}) \quad (6)$$

with P_d representing the average power of the dth path component and $f_d(\boldsymbol{\Omega})$ the DoA density function of the dth path component.

In this contribution, we assume that $f_d(\boldsymbol{\Omega})$ is of the form of the FB_5 pdf [6]. Among all distributions on the unit sphere $\mathbb{S}_2$, the FB_5 distribution [6] maximizes the entropy under the constraints that the first moment of the distribution and its second moment are specified. The first moment of the distribution are parameterized by the center of gravity of the power spectrum, while the second moments are characterized by the parameters describing the concentration and the ovalness of the spreads on the surface of the unit sphere. The direction density function $f_{\mathrm{FB}_5}(\boldsymbol{\Omega})$ reads

$$\begin{aligned}f_{\mathrm{FB}_5}(\boldsymbol{\Omega}) &= C(\kappa,\eta)^{-1}\exp\{\kappa\boldsymbol{\gamma}_1^{\mathrm{T}}\boldsymbol{\Omega} \\ &\quad +\kappa\cdot\eta[(\boldsymbol{\gamma}_2^{\mathrm{T}}\boldsymbol{\Omega})^2 - (\boldsymbol{\gamma}_3^{\mathrm{T}}\boldsymbol{\Omega})^2]\}, \quad (7)\end{aligned}$$

where $\kappa \geq 0$ represents the concentration parameter and $\eta \in [0, 1/2)$ describes the ovalness of the distribution. In (7), $C(\kappa,\eta)$ denotes a normalization constant number depending on κ and η, $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$, and $\boldsymbol{\gamma}_3 \in \mathbb{R}^3$ are unit vectors. The matrix $\boldsymbol{\Gamma} \doteq [\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3]$ is uniquely determined by three angular parameters $\bar{\theta}$, $\bar{\phi}$ and α according to

$$\begin{aligned}\boldsymbol{\Gamma} = &\begin{bmatrix}\sin(\bar{\theta})\cos(\bar{\phi}) & -\sin(\bar{\phi}) & \cos(\bar{\theta})\cos(\bar{\phi}) \\ \sin(\bar{\theta})\sin(\bar{\phi}) & \cos(\bar{\phi}) & \cos(\bar{\theta})\sin(\bar{\phi}) \\ \cos(\bar{\theta}) & 0 & -\sin(\bar{\theta})\end{bmatrix} \\ &\cdot\begin{bmatrix}1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha)\end{bmatrix}. \quad (8)\end{aligned}$$

In (8), $\bar{\phi}$ and $\bar{\theta}$ coincide with respectively the azimuth and the elevation of the mean direction, i.e. the first

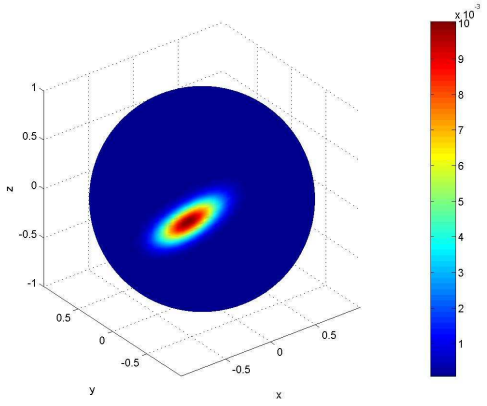


Fig. 1. The FB_5 density function with $\bar{\phi} = 135°, \bar{\theta} = 18°, \alpha = 144°, \kappa = 80$ and $\eta = 0.375$. The color bar to the right of the plot shows density expressed in linear scale.

moment of the distribution. The angle α describes how the density function is tilted on $\mathbb{S}_2$. A detailed description of the meanings of $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ and $\boldsymbol{\gamma}_3$ can be found in [6]. Note that when η equals 0, the FB_5 pdf does not depend on the values of α and the equal-density contours of $f_{\mathrm{FB}_5}(\boldsymbol{\Omega})$ are circles. For $\eta \in (0, 1/2)$, the equal-density contours of the pdf exhibit the ovalness, which becomes significant as η increases. Fig. 1 depicts the FB_5 density function for the parameter setting reported in the figure caption.

The parameters of $f_d(\boldsymbol{\Omega})$ are denoted by $\tilde{\boldsymbol{\theta}}_d \doteq [\bar{\phi}_d, \bar{\theta}_d, \kappa_d, \eta_d, \alpha_d]$. We use a vector $\boldsymbol{\theta}$ to represent all unknown model parameters in (1), i.e.

$$\boldsymbol{\theta} \doteq [P_1, P_2, \ldots, P_D, \tilde{\boldsymbol{\theta}}_1, \tilde{\boldsymbol{\theta}}_2, \ldots, \tilde{\boldsymbol{\theta}}_D].$$

III. Estimation of the Model Parameters

In a scenario with multiple path components, as depicted by (1), the problem at hand is to estimate the parameter vector $\boldsymbol{\theta}$. We now derive a SAGE algorithm [7] as an approximation of the maximum likelihood estimator of $\boldsymbol{\theta}$.

A. Admissible hidden data

We choose the subsets of parameters updated in the iterations of the SAGE algorithm to be the sets including the parameters characterizing individual path components. At Iteration $i = 1, 2, \ldots$, the parameter subset $\boldsymbol{\theta}_d \doteq [P_d, \tilde{\boldsymbol{\theta}}_d]$ with $d = [(i-1) \mod D] + 1$ is updated.

We define the admissible hidden data associated with $\boldsymbol{\theta}_d$ as

$$\begin{aligned}\boldsymbol{X}_d(t) &\doteq \boldsymbol{H}_d(t)u(t) + \boldsymbol{W}(t) \\ &= \left[\int_{\mathbb{S}_2} \boldsymbol{c}(\boldsymbol{\Omega}) h_d(t; \boldsymbol{\Omega}) \mathrm{d}\boldsymbol{\Omega}\right] u(t) + \boldsymbol{W}(t). \quad (9)\end{aligned}$$

It follows from the properties of $h_d(t; \boldsymbol{\Omega})$ that $\boldsymbol{H}_d(t)$ is constant within individual observation intervals, i.e.

$$\boldsymbol{H}_d(t) \doteq \boldsymbol{H}_{d,n} = \int_{\mathbb{S}_2} \boldsymbol{c}(\boldsymbol{\Omega}) h_{d,n}(\boldsymbol{\Omega}) \mathrm{d}\boldsymbol{\Omega}. \quad (10)$$

The output of a correlator

$$\tilde{\boldsymbol{H}}_{d,n} \doteq \int_{t_n}^{t_n+T} \boldsymbol{x}_d(t) u(t)^* \mathrm{d}t, \quad n = 1, \ldots, N. \quad (11)$$

with the input as the observation $\boldsymbol{X}_d(t) = \boldsymbol{x}_d(t)$ can be written as

$$\tilde{\boldsymbol{H}}_{d,n} = \boldsymbol{H}_{d,n} + \boldsymbol{N}_n, \quad (12)$$

where $\boldsymbol{N}_n \in \mathbb{C}^M$, $n = 1, \ldots, N$ is a sequence of N independent random vectors, the entries of which are independent circularly symmetric Gaussian random variables with variance σ_w^2. Invoking the central limit theorem, the elements of $\tilde{\boldsymbol{H}}_{d,n}$ in (9) are assumed to be Gaussian random variables. The vectors $\tilde{\boldsymbol{H}}_{d,1}, \ldots, \tilde{\boldsymbol{H}}_{d,N}$ form a sufficient statistic for the estimation of $\boldsymbol{\theta}_d$.

B. Expectation Step

In the Expectation (E-) step of Iteration i, we compute the expectation of the likelihood of $\boldsymbol{\theta}_d$ conditioned on the observation $\boldsymbol{Y}(t) = \boldsymbol{y}(t)$ and assuming that $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}^{[i-1]}$:

$$Q(\boldsymbol{\theta}_d | \widehat{\boldsymbol{\theta}}^{[i-1]}) \doteq \mathrm{E}\big[\Lambda(\boldsymbol{\Omega}_d; \boldsymbol{X}_d) | \boldsymbol{Y}(t) = \boldsymbol{y}(t), \widehat{\boldsymbol{\theta}}^{[i-1]}\big]. \quad (13)$$

Here, $\widehat{\boldsymbol{\theta}}^{[i-1]}$ denotes the parameter estimates obtained in the $(i-1)$th iteration. It can be shown that (13) is of the form

$$\begin{aligned}Q(\boldsymbol{\theta}_d | \widehat{\boldsymbol{\theta}}^{[i-1]}) = &- \ln |\boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}_d}(\boldsymbol{\theta}_d)| - \mathrm{tr}\big[(\boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}_d}(\boldsymbol{\theta}_d))^{-1} \\ &\cdot \widehat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{H}}_d | \boldsymbol{y}(t)}(\widehat{\boldsymbol{\theta}}^{[i-1]})\big], \quad (14)\end{aligned}$$

where $\mathrm{tr}[\cdot]$ is the trace of the matrix given as an argument and $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}_d}(\boldsymbol{\theta}_d)$ is the covariance matrix of $\tilde{\boldsymbol{H}}_{d,n}$ as a function of the parameter subset $\boldsymbol{\theta}_d$:

$$\boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}_d}(\boldsymbol{\theta}_d) = P_d \int_{\mathbb{S}_2} \boldsymbol{c}(\boldsymbol{\Omega}) \boldsymbol{c}(\boldsymbol{\Omega})^{\mathrm{H}} f_d(\boldsymbol{\Omega}) \mathrm{d}\boldsymbol{\Omega} + \sigma_w^2 \boldsymbol{I}_M$$

with $[\cdot]^{\mathrm{H}}$ denoting the Hermitian operator. In (14), $\widehat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{H}}_d | \boldsymbol{y}(t)}(\boldsymbol{\theta})$ is the conditional covariance matrix of $\tilde{\boldsymbol{H}}_{d,n}$ given the observation $\boldsymbol{y}(t)$. It can be shown that

$$\begin{aligned}\widehat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{H}}_d | \boldsymbol{y}(t)}(\widehat{\boldsymbol{\theta}}^{[i]}) &= \boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}_d}(\widehat{\boldsymbol{\theta}}_d^{[i]}) \\ &- \boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}_d}(\widehat{\boldsymbol{\theta}}_d^{[i]}) \big[\boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}}(\widehat{\boldsymbol{\theta}}^{[i]})\big]^{-1} \boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}_d}(\widehat{\boldsymbol{\theta}}_d^{[i]}) \\ &+ \boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}_d}(\widehat{\boldsymbol{\theta}}_d^{[i]}) \big[\boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}}(\widehat{\boldsymbol{\theta}}^{[i]})\big]^{-1} \widehat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{H}}} \big[\boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}}(\widehat{\boldsymbol{\theta}}^{[i]})\big]^{-1} \\ &\quad \boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}_d}(\widehat{\boldsymbol{\theta}}_d^{[i]}), \quad (15)\end{aligned}$$

where

$$\boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}}(\widehat{\boldsymbol{\theta}}) = \sum_{d=1}^{D} \boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}_d}(\widehat{\boldsymbol{\theta}}_d) + \sigma_w^2 \boldsymbol{I}_M,$$

$$\widehat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{H}}} = \frac{1}{N} \sum_{n=1}^{N} \tilde{\boldsymbol{H}}_n \tilde{\boldsymbol{H}}_n^{\mathrm{H}}$$

with $\tilde{\boldsymbol{H}}_n \doteq \int_{t_n}^{t_n+T} \boldsymbol{y}(t) u(t)^* \mathrm{d}t$, $n = 1, \ldots, N$.

C. Maximization Step

In the M-step, the estimate $\widehat{\boldsymbol{\theta}}_d^{[i]}$ is calculated as

$$\widehat{\boldsymbol{\theta}}_d^{[i]} = \arg\max_{\boldsymbol{\theta}_d} Q(\boldsymbol{\theta}_d|\widehat{\boldsymbol{\theta}}^{[i-1]}).$$

By applying a coordinate-wise updating procedure similar to that used in [8], the required multiple-dimensional maximization can be reduced to multiple one-dimensional maximization problems. This coordinate-wise updating still remains within the SAGE framework with the admissible data given in (9).

D. Initialization Step

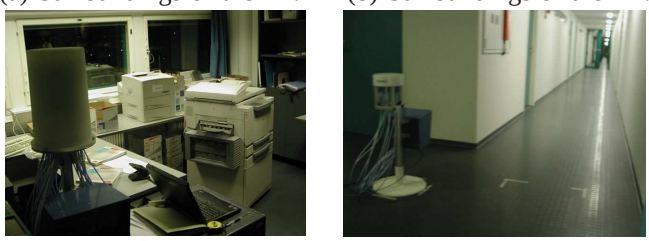
In the initialization step, the nominal azimuths of arrival and elevations of arrival of the path components are estimated using a SAGE algorithm derived based on the specular-path model [8]. The parameters which the Bartlett beamformer is incapable to estimate are set to certain predefined values. So, the estimates of the concentration parameters κ_d, $d = 1, \ldots, D$ are set to 100 and the ovalness parameters are sent to zero. With this setting it is assumed a priori that the path components are close to specular path components and that the spectrum is not tilted on $\mathbb{S}_2$. This initialization method procedure has worked well for measurement data in the scenarios where it was tested.

IV. Experimental Investigations

The measurement data was collected using an Elektrobit Propsound CS switched channel sounder [2] in an office building. The sounder was configured with a MISO (multiple-input single-output) structure where the Rx has a single antenna and the Tx is equipped with a 50-element omnidirectional antenna array. A detailed description of the sounder, the array and the measurement settings can be found in [2]. In the measurement, the Rx was located in a corridor and the Tx was located in an office room. Two photographs and the map shown in Fig. 2 depict the surroundings of the Rx and Tx. In the map, the locations of the Tx and the Rx are marked with the symbols $\odot$ and $\otimes$ respectively. In the measurement, both Tx and Rx were fixed. People were moving in the office where the Tx was located. These movements created the randomness of the radio channel. Due to this reason, the uncorrelated scattering condition as depicted in (5) is considered to be valid.

The data obtained from 50 consecutive measurement cycles within a period of 3.3 seconds are considered. A measurement cycle is referred to as the interval within which all 50 subchannels are sounded once. In this preliminary study, we investigate dispersion of individual path components in direction of departure and neglect dispersion in other dimensions. As delay dispersion is not considered, we consider the output of the Rx antenna

(a) Surroundings of the Tx. (b) Surroundings of the Rx.



(c) Map of the premises.

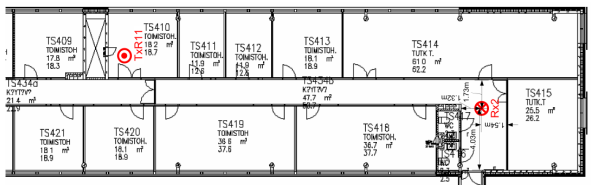


Fig. 2. Photographs and map of the premises where the measurement experiment was conducted.

at 160 ns. The narrow-band signal model (1) is applicable in the considered scenario. The parameter estimators derived based on the SAGE algorithm can be easily modified to estimate the parameters of the DoD power spectrum of individual path components.

For the implementation of the SAGE algorithm, we assume the number of the path components is known and equals 4 in the considered scenario. Totally 10 SAGE iteration cycles are performed. Here, an iteration cycle is referred to as the procedure in which the estimates of all elements in $\boldsymbol{\theta}$ are updated once. In the M-step we select the quantization step to be 2° in both azimuth and elevation. This is due to the fact that the Rx array response is measured at points with minimum spacing of 2° in both azimuth and elevation.

Fig. 3 depicts the estimation results obtained using the SAGE algorithm. The parameter estimates are reported in Table I. The notation $\mathrm{Bartlett}(\cdot)$ in Fig. 3 denotes the Bartlett spectrum computed with the covariance matrix given as an argument and $\hat{P}(\boldsymbol{\Omega})$ represents the estimated power spectrum of the radio channel response. It can be observed that the power spectrum of individual path components estimated using the SAGE algorithm are noticeably more concentrated as shown in $\hat{P}(\boldsymbol{\Omega})$ than the corresponding footprints depicted in $\mathrm{Bartlett}(\boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}}(\widehat{\boldsymbol{\theta}}))$. These path components differ in concentration, ovalness and tilt angle. The "blurring" effect arising in the Bartlett spectrum is due to the response of the Rx array.

The footprints of the path components shown in $\mathrm{Bartlett}(\boldsymbol{\Sigma}_{\tilde{\boldsymbol{H}}}(\widehat{\boldsymbol{\theta}}))$ and $\mathrm{Bartlett}(\widehat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{H}}})$ are observed to be similar. This implies that the reconstructed covariance matrix computed using the parameter estimates is close to the sample covariance matrix. We also observe some

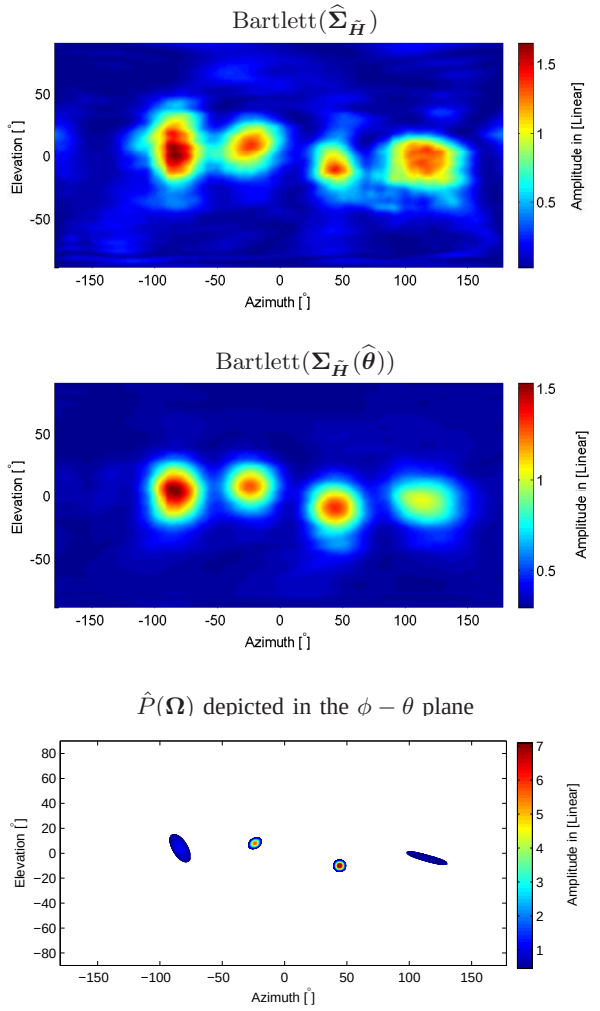


Fig. 3. Estimated direction power spectrum at delay 160 ns. The estimates of the parameters are shown in Table I.

differences on the shapes and the maximum spectral heights of the corresponding footprints. These differences can be caused by the approximation of the "true" power spectrum of individual path components with a FB_5 pdf. Another possible reason for the difference is that dispersion in other dimensions, e.g. in delay, is not considered.

V. Conclusions

In this contribution, we derived a SAGE algorithm for estimation of the parameters characterizing the direction power spectrum of individual path components in a radio propagation channel. The Fisher-Bingham-5 probability density function (pdf) was used to model the shape of the direction power spectrum of individual path components. The performance of the SAGE algorithm was evaluated using measurement data. From the results we observed that the Bartlett spectra obtained from the signal covariance matrix computed using the SAGE parameter estimates and from the sample covariance matrix are similar. The estimated power spectra of individual path components exhibit different ovalness and tilt angle. They are more concentrated than the corresponding footprints in the Bartlett spectrum. These results indicated that dispersive path components exist in real propagation channels. In such a case, the conventional algorithms derived based on the specular-path model are inappropriate for estimation of the parameters of the radio channel. As shown in [9], the mismatch between the specular-path model and the "true" dispersive feature of path components results in significant estimation errors with high probabilities of occurrence.

TABLE I

The estimates of the parameters obtained using the SAGE algorithm.

d	$\bar{\phi}_d$ [°]	$\hat{\bar{\theta}}_d$ [°]	$\hat{\kappa}_d$	$\hat{\eta}_d$	$\hat{\alpha}_d$ [°]	$\hat{P}_d[10^{-10}]$	$\hat{P}_d$ [dB]
1	-84	4	140	0.33	59.3	7.10	0
2	114	-4	160	0.49	15.8	5.72	-1
3	44	-10	923	0.00	26.5	5.19	-1
4	-24	8	923	0.17	144.0	4.10	-2

References

[1] C. B. Ribeiro, A. Richter, and V. Koivunen, "Stochastic maximum likelihood estimation of angle- and delay-domain propagation parameters," in *Proceedings of the 16th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Berlin, Germany, 2005.

[2] X. Yin, T. Pedersen, N. Czink, and B. H. Fleury, "Parametric characterization and estimation of bi-azimuth dispersion of path components," in *Proceedings of the 7th IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, Nice, France, July 2006.

[3] X. Yin, L. Liu, D. Nielsen, N. Czink, and B. H. Fleury, "Characterization of the azimuth-elevation power spectrum of individual path components," in *Proceedings of the International ITG/IEEE Workshop on Smart Antennas (WSA)*, Vienna, Austria, Feb. 2007.

[4] X. Yin, T. Pedersen, N. Czink, and B. H. Fleury, "Parametric characterization and estimation of bi-azimuth and delay dispersion of path components," in *Proceedings of The First European Conference on Antennas and Progation (EuCAP)*, Acropolis, Nice, France, November 2006.

[5] B. H. Fleury, "First- and second-order characterization of direction dispersion and space selectivity in the radio channel," *IEEE Trans. Information Theory*, no. 6, pp. 2027–2044, Sept. 2000.

[6] J. T. Kent, "The Fisher-Bingham distribution on the sphere," *Journal of the Royal Statistical Society, Serieal B (Methodological)*, vol. 44, pp. 71–80, 1982.

[7] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. on Signal Processing*, vol. 42, no. 10, pp. 2664–2677, Oct. 1994.

[8] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. L. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 434–450, Mar. 1999.

[9] M. Bengtsson and B. Völcker, "On the estimation of azimuth distributions and azimuth spectra," in *Proceedings of the 54th IEEE Vehicular Technology Conference (VTC-Fall)*, vol. 3, no. 12, 2001, pp. 1612–1615.

Characterization of the Azimuth-Elevation Power Spectrum of Individual Path Components

Xuefeng Yin[1], Lingfeng Liu[1], Daniel K. Nielsen[1], Nicolai Czink[3,2] and Bernard H. Fleury[1,2]

[1]Department of Electronics Systems, Aalborg University, DK-9220 Aalborg, Denmark
[2]Forschungszentrum Telekommunikation Wien (ftw.), Vienna, Austria
[3]Institut für Nachrichtentechnik und Hochfrequenztechnik, Technische Universität Wien, Vienna, Austria

***Abstract*— In this contribution, we propose to use the density function of the Fisher-Bingham-5 distribution to characterize the shape of the azimuth-elevation power spectrum of individual path components in the response of the radio channel. The maximum likelihood estimator of the parameters of the power spectrum is derived and applied to estimate the dispersive characteristics of individual path components from measurement data. Preliminary results are presented that illustrate the applicability of the method.**

I. Introduction

Due to the heterogeneity of the propagation environment, the response of the radio channel is the superposition of a certain number of components. Each component, which we call a "path component", is contributed by an electromagnetic (EM) wave propagating along a path from the transmitter (Tx) to the receiver (Rx). Along this path, the EM wave interacts with a certain number of objects that we call scatterers. Due to the geometrical extent and nonhomogeneous electromagnetic properties of the scatterers, a path may be dispersive in delay, direction of departure, direction of arrival, polarization, as well as in Doppler frequency when the environment is time-variant. Thus, an individual path component may be spread in these dispersion dimensions.

Recently, different methods have been proposed for estimation of dispersive characteristics of individual path components. Some of these methods make use of the assumption that the shape of the power spectrum of individual path components can be described using a density function of a probability distribution. In the case where dispersion in one dimension, e.g. azimuth of arrival (AoA), is considered, the shape of the AoA power spectrum can be described using the density function of the (truncated) Gaussian distribution [1], the uniform distribution confined within a certain azimuth range [2], and the von-Mises distribution [3]. Recently, dispersion in multiple dimensions of individual path components have also been investigated. In [4], a von-Mises-Fisher density function is used to characterize the shape of the biazimuth (azimuth of departure and AoA) power spectrum. In [5], a 3-variate density function is derived which can be used to model the shape of the biazimuth-delay power spectrum of individual path components.

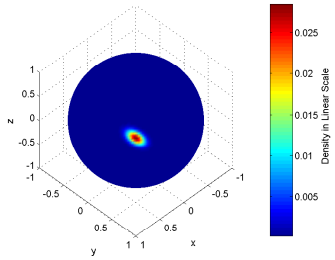
In this contribution, we propose to use the Fisher-Bingham-5 (FB_5) density function to describe the shape of the azimuth-elevation power spectrum of individual path components. Among all distributions on the unit sphere, the FB_5 distribution maximizes the entropy under the constraints that the distribution's first moment and second moment are specified [6]. The maximum likelihood (ML) estimator of the parameters of the power spectrum is derived and applied to estimate the dispersive characteristics of individual path components from measurement data.

II. FB_5 Density Function, Signal Model and Maximum Likelihood Estimator

Following the nomenclature in [7], a direction can be characterized using a unit vector $\mathbf{\Omega}$. This vector has its initial point anchored at the origin O of a coordinate system, and terminal point located on a unit sphere $\mathbb{S}_2$ centered at O. This vector $\mathbf{\Omega}$ is uniquely determined by its elevation θ and azimuth ϕ. The FB_5 distribution [6] is a probability distribution on $\mathbb{S}_2$ which maximizes the entropy under the constraint that the distribution's first moment and second moment are specified. The density function of the FB_5 distribution is of the form

$$f_{\mathrm{FB}_5}(\mathbf{\Omega}) = c \cdot \exp\{\kappa \boldsymbol{\gamma}_1^{\mathrm{T}} \mathbf{\Omega} + \beta[(\boldsymbol{\gamma}_2^{\mathrm{T}} \mathbf{\Omega})^2 - (\boldsymbol{\gamma}_3^{\mathrm{T}} \mathbf{\Omega})^2]\},$$

where c represents a normalization constant, κ and β are respectively the concentration parameter and the ovalness parameter, the vectors $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ and $\boldsymbol{\gamma}_3 \in \mathbb{R}^{3\times 1}$



$\bar{\phi}$ [°]	$\bar{\theta}$ [°]	α [°]	κ	β
45	70	35	200	100

Fig. 1. The FB_5 density function with the parameter setting given above.

This work was jointly supported by the Network of Excellence in Wireless COMmunications (NEWCOM) and Elektrobit Group.

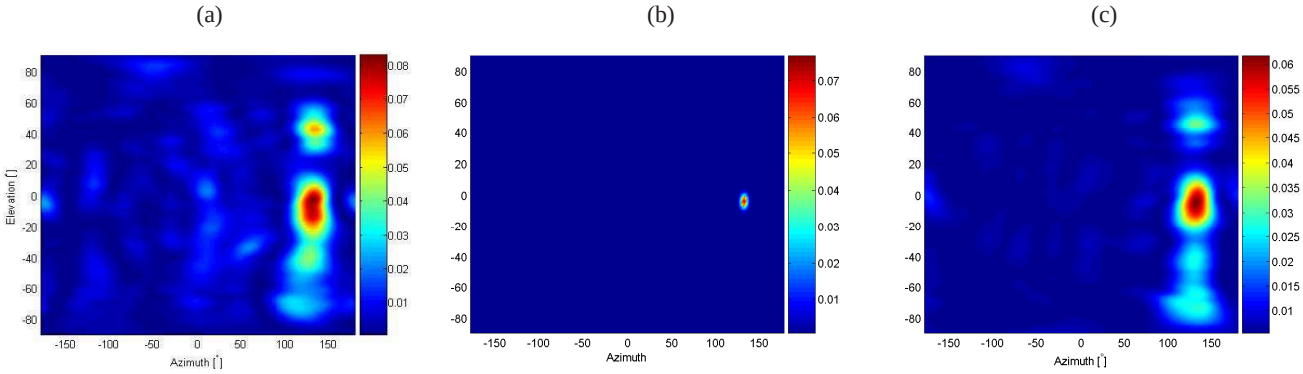


Fig. 2. Estimated azimuth-elevation power spectra: (a), Bartlett spectrum computed from original received data; (b), Proposed ML estimate of the power spectrum; (c), Bartlett spectrum reconstructed using the ML estimate of the power spectrum. Color bars to the right of these plots show the spectral height in linear scale.

are determined by three angular parameters $\bar{\theta}$, $\bar{\phi}$ and α. Here, $\bar{\theta}$ and $\bar{\phi}$ specify respectively the elevation and the azimuth of the mean direction, and α describes how the distribution is tilted on $\mathbb{S}_2$. Fig. 1 depicts a surface calculated using the FB_5 density function on the unit sphere for the parameter setting also reported in the figure. A detailed description of the features of the FB_5 density function will be presented in the full version of the paper.

In the full version of the paper, we will also present a signal model for channel sounding, in which the FB_5 density function is used to characterize the shape of the azimuth-elevation spectrum of individual path components. The ML estimator of the parameters of the power spectrum will be derived and its properties described in the full paper.

III. Preliminary Experimental Investigation

We apply the derived MLE to measurement data collected using the MIMO wideband radio channel sounder Elektrobit Propsound CS [4]. The description of the sounder setting as well as of the environment where the measurements were conducted is given in [4].

In a preliminary investigation, we consider a SIMO system which consists of one Tx antenna and a 32-element Rx array. As delay dispersion is not investigated in this study, the output of the Rx array at a specific delay is considered. Fig. 2 (a) shows the estimate of the AoA-EoA power spectrum computed from the selected data by using the Bartlett beamformer [8]. In the following discussion, we refer to this spectrum as Bartlett spectrum. The proposed ML estimate of the power spectrum of a single path component is depicted in Fig. 2 (b). Fig. 2 (c) shows the Bartlett spectrum calculated from the reconstructed signal covariance matrix computed based on the ML estimate of the power spectrum illustrated in Fig. 2 (b). It can be observed that the ML estimate of the power spectrum is more concentrated than the corresponding footprint observed in the Bartlett spectrum in Fig. 2 (c). The blurring effect observed in the Bartlett spectrum is due to the ambiguity function of the Rx array response. It can also be observed that the "reconstructed" Bartlett spectrum shown in Fig. 2 (c) is similar to the "original" Bartlett spectrum depicted in Fig. 2 (a). A more comprehensive discussion of the results will be presented in the full paper.

IV. Conclusions

In this contribution, we proposed to use the Fisher-Bingham-5 density function to model the shape of the azimuth-elevation power spectrum of individual path components. We derived the maximum likelihood estimator of the parameters of the power spectrum and used it to estimate the dispersive characteristics of individual path components from measurement data. From a preliminary result, we found that the estimated power spectra of the path components are noticeably more concentrated than the corresponding footprints in the Bartlett spectrum.

References

[1] T. Trump and B. Ottersten, "Estimation of nominal direction of arrival and angular spread using an array of sensors," *Signal Processing*, vol. 50, pp. 57–69, Apr. 1996.

[2] O. Besson and P. Stoica, "Decoupled estimation of DoA and angular spread for spatially distributed sources," *IEEE Trans. Signal Processing*, vol. 49, pp. 1872–1882, 1999.

[3] C. B. Ribeiro, E. Ollila, and V. Koivunen, "Stochastic maximum likelihood method for propagation parameter estimation," in *Proceedings of the 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, vol. 3, Sept. 5-8 2004.

[4] X. Yin, T. Pedersen, N. Czink, and B. H. Fleury, "Parametric characterization and estimation of bi-azimuth dispersion of path components," in *Proceedings of the 7th IEEE International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, Nice, France, July 2006.

[5] ——, "Parametric characterization and estimation of bi-azimuth and delay dispersion of path components," in *Proceedings of The First European Conference on Antennas and Progation (EuCAP)*, Acropolis, Nice, France, November 2006.

[6] J. T. Kent, "The fisher-bingham distribution on the sphere," *Journal of the Royal Statistical Society, Serieal B (Methodological)*, vol. 44, pp. 71–80, 1982.

[7] B. H. Fleury, "First- and second-order characterization of direction dispersion and space selectivity in the radio channel," *IEEE Trans. Information Theory*, no. 6, pp. 2027–2044, Sept. 2000.

[8] M. Bartlett, "Smoothing periodograms from time series with continuous spectra," *Nature*, vol. 161, 1948.