Circuits and Systems
Mekelweg 4,
2628 CD Delft
The Netherlands
https://cas.tudelft.nl

Indoor Localization using Narrowband Radios and Switched Antennas in Indoor Environments

by

Ye Cui

Student number:	4735315	
Defense date:	17-10-2019	
Committee member:	Prof.dr.ir. A.J. van der Veen,	TU Delft, supervisor
	Dr.ir. J. Romme,	IMEC-Holst Center/TU Delft, daily supervisor
	Dr.ir. J.N. Driessen,	TU Delft

Indoor Localization using Narrowband Radios and Switched Antennas in Indoor Environments

Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science

in

Electrical Engineering

by

Ye Cui
born in Zibo, China

This work was performed in:

Circuits and Systems Group
Department of Microelectronics
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

Delft University of Technology

Abstract

Indoor positioning using Bluetooth addressed a great concern. The properties of narrowband usage, low energy consumption and universality on devices attract numerous customers and promote researchers to investigate potential of Bluetooth in indoor positioning field.

It has already supported Angle-of-Arrival (AoA) and Angle-of-Departure (AoD) in angle domain for indoor localization. In range domain, using Received Signal Strength (RSS) is practicable but it cannot provide enough resolution. It is needed to develop a technique that can improve resolution for range finding feature. The challenges of localization in indoor environments are mainly from the multipath propagation of signals. In this thesis, a subspace-based super-resolution algorithm for time delay dispersion estimate is developed. Range finding is realized by computing time-of-arrival (ToA) parameter of signal arriving at direct line-of-sight (DLoS). An essential procedure before the developed algorithm is applied is that to correctly separate subspace into signal space and noise space. Techniques for subspace separation are investigated in this thesis.

In this thesis, we want to explore the potential of indoor localization using Bluetooth narrowband radios. To start with, a data model according to the property of the conducted measurement data is developed. The conducted measurement data is radio channel measurements based on channel sounding technique. Then the data model is developed as channel impulse response model and multipath signals are indicated by different time delays. Since an accurate covariance matrix of measurement data is required for super-resolution algorithm, smoothing techniques is employed. The smoothing techniques considered are forward smoothing technique and forward-backward smoothing technique.

For the purpose of obtaining an accurate subspace separation, two techniques are investigated in this thesis, namely MDL criteria algorithm and the threshold method. in order to investigate the performance and reliability of those two techniques, experiments are taken out using different parameter values. Comparison is made between the results of these two techniques. Afterwards, subspace-based super-resolution algorithm is taken into consideration. In this thesis, the super-resolution algorithm implemented is MUSIC algorithm. The functionality of MUSIC algorithm on narrowband radios measurements is tested and evaluated firstly by simulation experiments, which demonstrates the practicability of applying MUSIC algorithm on narrowband radios measurements. Then experiments are extended to the measurement data that conducted from real indoor environments, for the purpose of indoor localization realization using narrowband radios.

Acknowledgments

I would like to express my sincere gratitude to all the people guided and supported me during the period of this thesis project.

Firstly, I would like to thank my supervisor Prof. Alle-Jan van der Veen. He supported me all the time with patience and encouragement from the beginning. His brilliant ideas always give me inspirations and help me to refine the works in this thesis.

Then I would like to thank my daily supervisor Dr. Jac Romme. He gave me the opportunity to do research on this project. He guided and encourage me with professional knowledge and attitude, to helpe me overcome all the difficulties during the work.

Finally, I would like to thank my family and friends for accompanying and supporting me all the time. Everything we experienced together left a great memories in my life and I had an enjoyable time in TU Delft.

Ye Cui
Delft, The Netherlands
17-10-2019

Contents

List of Figures

List of Tables

1 Introduction

In recent years, a more accurate and reliable indoor positioning technique has drawn a great attention. In outdoor positioning, a high accuracy can be achieved using Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS). GPS has the most widely applications because of its large coverage and applicability to various devices [1]. However, GPS will fail in indoor environments because of the complexity of indoor environments. In indoor environments, the geometry of the room, furniture and equipment or humans can be obstacles that influence the propagation of signals and lead to multipath loss [2, 3, 4].

Indoor positioning can provide various position-based service applications in different fields and situations, making it worth to develop. Over the past few years, many wireless communication techniques such as ultrasound, radio-frequency identification (RFID), ultra-wideband (UWB), wireless local area network (WLAN) and Bluetooth are employed and adopted for indoor positioning service (IPS).

Bluetooth is a popular choice for IPS because it only uses a narrow bandwidth and it is energy efficiently, especially for the Bluetooth low energy (BLE) system. Besides, the universality of Bluetooth also takes benefits to it, it is predicted that there are over 30% of 48 billion Internet-connected devices will be equipped with Bluetooth technology in a short future [5]. As a result of it, there will be a huge number of consumers of Bluetooth technology based IPS. All of those promote researchers to explore applications of Bluetooth in indoor positioning field.

1.1 Motivation

In the next generation of Bluetooth standard, the Bluetooth SIG wants to apply Bluetooth low energy system into indoor localization field. Currently it has already supported Angle-of-Arrival (AoA) or Angle-of-Departure (AoD) in angle domain, as well as distance measurements using Received Signal Strength (RSS). The multipath interference and noise in indoor environment take challenge to localization and tracking using range estimation. The performance is not satisfied when using RSS to localize and track an objective in indoor environments. Another approach to compute the range between the target and every receiver antenna is to estimate the time-of-arrival (ToA), and the most important thing for positioning in range domain is to estimate ToA parameter of the radio signal arriving from the direct line-of-sight (DLoS).

In order to realize localization in indoor environments, it is needed to obtain an accurate estimation of ToA parameter from DLoS path. Due to multipath interference and noise, this can be challenging. It is necessarily to develop an algorithm for ToA parameter estimate to obtain an accurate range estimate.

1.2 Research Objectives

Based on the introduction and motivation expressed above, this project wants to test the potential of BLE system in indoor positioning applications by processing measurements under the BLE specification. This project focus on using range information by computing time-of-arrival parameter to realize indoor localization and the main research objectives are to:

- develop a data model that can describe multipath propagation of signals in indoor environments.
- develop a subspace-based super-resolution algorithm used on narrowband measurements, which should estimate dispersion of time delay with high resolution.
- investigate methods and techniques for subspace separation, carry out experiments to evaluate them.
- compare the performance of the super-resolution algorithm when it cooperates with different subspace separation techniques and investigate the advantages and limitations of different techniques.

1.3 Structure

This thesis is organized as follows: Chapter 2 will firstly gives an introduction about how measurement data is conducted. In this chapter, the data model of multipath propagation of signals in indoor environments will be proposed.

In Chapter 3, the subspace-based super-resolution algorithm – MUSIC algorithm for range finding by estimating time delay parameter will be introduced. The reasons for why a accurate subspace separation is needed will be demonstrated.

In Chapter 4, a proposed method – MDL criteria algorithm for subspace separation will be proposed and evaluated. Experiments will be carried out under BLE specification.

In Chapter 5, another subspace separation method – the threshold method will be proposed and evaluated. Comparisons between MDL criteria algorithm and the threshold method will be made.

In Chapter 6, MUSIC algorithm will be firstly tested on simulation data, after which, it will be evaluated on data that conducted from real indoor environments.

The final chapter, Chapter 7 contains conclusions of this thesis. Besides, studies that can be further considered about will also be contained in this chapter.

Data Model Development 2

In indoor environments, signals propagate in multipath propagation way. In order to realize localization in range domain, the ToA parameter of the radio signal arriving from the DLoS needs to be estimated. The super-resolution technique can be employed in the frequency domain to estimate the time delay dispersion from every multipath component. In this chapter, at first, the transmitted signal setting is introduced to obtain the frequency domain measurement. Then in order to understand the influence of indoor environments on signal propagation, channel model of indoor environments is developed. At last based on the property of channel model, two smoothing techniques are introduced.

2.1 Channel Sounding

The ToA parameter is estimated by implementing super-resolution technique in the radio channel measurements. We can apply a frequency shift to transmit signal to obtain the radio channel measurements. At transmitter antenna, signals are transmitted with different carrier frequencies within a certain band range. Those carrier frequencies are uniform distributed with a certain frequency interval Δf. During transmitting, signals are firstly transmitted at the first carrier frequency ($i.e.,\ \ f_0$) for a certain period, after which the carrier frequency is shifted by Δf to $f_0+\Delta f$. Repeat this procedure until the carrier frequency is $f_0+(N-1)\Delta f$, where N is the number of carrier frequency can be used within this frequency band range. This procedure is shown in Figure 2.1. The transmitting duration of signal with different carrier frequencies are same. The transmit signal has constant envelope which means that the power of signal of all carrier frequencies are same. At receiver antenna, signals are received with different carrier frequencies. Then radio channel measurements can be conducted using a Vector Network Analyzer (VNA).

In our case, we use narrowband radios under BLE specification. The frequency band range that can be used is 2.4∼2.48GHz, and frequency interval Δf is 2MHz.

2.2 Channel Model

In order to understand how indoor environments influence the propagation of signals, the first step is to develop a model to describe indoor environments channel. We assume that a line-of-sight propagation with delay τ, and then the channel can be modeled as a impulse response, which is:

$$h(t) = c\delta(t-\tau)$$

when considering that indoor environments lead to multipath propagation of signals, the equation above can be expressed as [6, 7]:

$$h(t) = \sum_{k=0}^{d-1} c_k \delta(t-\tau_k) \tag{2.1}$$

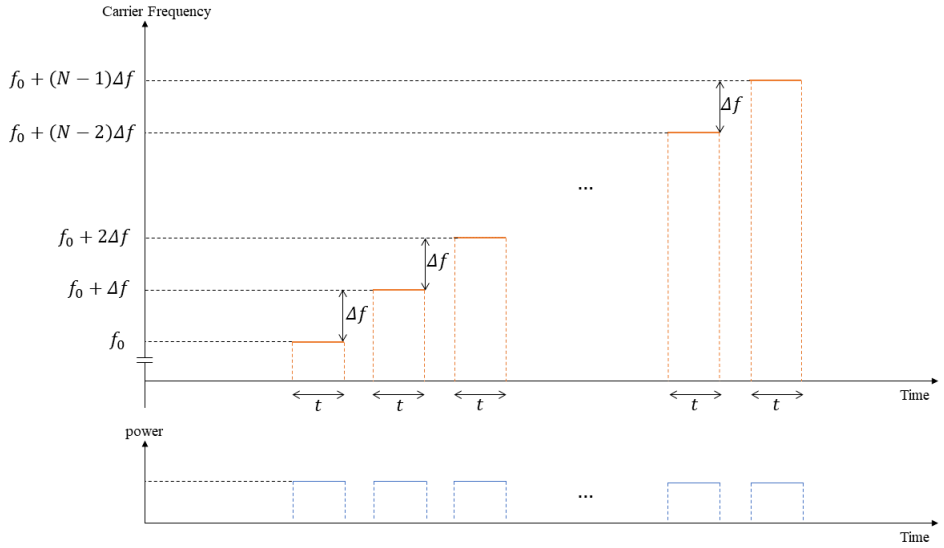


Figure 2.1: Frequency shift of transmit signal

where d is the number of multipath components, c_k is the channel gain including a real positive attenuation and a randomly phase shift caused by indoor environments [8], τ_k denotes time delay of the kth multipath and $\delta(\cdot)$ is the Dirac function.

Taking Fourier transform of Equation 2.1, then channel impulse response (CIR) in frequency domain can be expressed as:

$$H(f) = \sum_{k=0}^{d-1} c_k e^{-j2\pi f \tau_k} \tag{2.2}$$

at here j denotes imaginary unit, f is the carrier frequency, d, c_k and τ_k denote the same parameters mentioned above and they are normally assumed to be time-invariant [7]. Defining that $\tau_k < \tau_{k+1}$, then the shortest time delay is τ_0, which normally is regarded as the signal arriving form the direct line-of-sight (DLoS). In this case, multiple carrier frequencies are applied to the transmitting signals. For carrier frequency f_i, rewrite the denotation $H(f_i)$ as h_i, and then h_i denotes the frequency domain impulse response with carrier frequency f_i. The multiple carrier frequencies are assumed to be uniform distributed which denoted as $f_i = f_0 + i\Delta f$ ($i = 0, 1, 2..., N-1$), where f_0 is the first carrier frequency, Δf is the frequency interval of the frequency range and N is the number of frequencies used. In this case, the frequency range is from 2.4 to 2.48GHz.

The channel impulse response can be modeled as:

$$\begin{aligned} h_i &= \sum_{k=0}^{d-1} c_k e^{-j2\pi f_i \tau_k} \\ &= \sum_{k=0}^{d-1} c_k e^{-j2\pi (f_0 + i\Delta f)\tau_k} \\ &= \sum_{k=0}^{d-1} \underbrace{c_k e^{-j2\pi f_0 \tau_k}}_{s_k} \underbrace{e^{-j2\pi i \Delta f \tau_k}}_{a_{i,k}} \\ &= \begin{bmatrix} e^{-j2\pi i\Delta f\tau_0} & e^{-j2\pi i\Delta f\tau_1} & \dots & e^{-j2\pi i\Delta f\tau_{d-1}} \end{bmatrix} \begin{bmatrix} c_0 e^{-j2\pi f_0\tau_0} \\ c_1 e^{-j2\pi f_0\tau_1} \\ \vdots \\ c_{p-1} e^{-j2\pi f_0\tau_{d-1}} \end{bmatrix} \end{aligned} \tag{2.3}$$

i is the index of carrier frequencies and s_k is a parameter that can be regarded as the signal from k-th sources, and $a_{i,k}$ is regarded as the channel response for each source. Write the channel impulse response with different carrier frequencies into a vector:

$$\mathbf{h} = \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{N-1} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j2\pi\Delta f\tau_0} & e^{-j2\pi\Delta f\tau_1} & \dots & e^{-j2\pi\Delta f\tau_{p-1}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j2\pi(N-1)\Delta f\tau_0} & e^{-j2\pi(N-1)\Delta f\tau_1} & \dots & e^{-j2\pi(N-1)\Delta f\tau_{d-1}} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} c_0 e^{-j2\pi f_0\tau_0} \\ c_1 e^{-j2\pi f_0\tau_1} \\ \vdots \\ c_{d-1} e^{-j2\pi f_0\tau_{d-1}} \end{bmatrix}}_{\mathbf{s}} \tag{2.4}$$

which indicates that

$$\mathbf{h} = \mathbf{A}\mathbf{s} \tag{2.5}$$

with

$$\mathbf{s} = \begin{bmatrix} c_0 e^{-j2\pi f_0\tau_0} \\ c_1 e^{-j2\pi f_0\tau_1} \\ \vdots \\ c_{d-1} e^{-j2\pi f_0\tau_{d-1}} \end{bmatrix} \tag{2.6}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j2\pi\Delta f\tau_0} & e^{-j2\pi\Delta f\tau_1} & \dots & e^{-j2\pi\Delta f\tau_{d-1}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j2\pi(N-1)\Delta f\tau_0} & e^{-j2\pi(N-1)\Delta f\tau_1} & \dots & e^{-j2\pi(N-1)\Delta f\tau_{d-1}} \end{bmatrix} \tag{2.7}$$

here $\mathbf{s}$ can be regarded as a signal vector from d sources with the same carrier frequency is f_0, and $\mathbf{A}$ can be regarded as a steering matrix. To simplify this expression, we can write:

$$z_k = e^{-j2\pi\Delta f\tau_k} \tag{2.8}$$

then the expression of matrix $\mathbf{A}$ is

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} z_0^0 & z_1^0 & \cdots & z_{d-1}^0 \\ z_0^1 & z_1^1 & \cdots & z_{d-1}^1 \\ \vdots & \vdots & \ddots & \vdots \\ z_0^{N-1} & z_1^{N-1} & \cdots & z_{d-1}^{N-1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{a}(\tau_0) & \mathbf{a}(\tau_1) & ... & \mathbf{a}(\tau_{d-1}) \end{bmatrix}\end{aligned} \tag{2.9}$$

with

$$\begin{aligned}\mathbf{a}(\tau_k) &= \begin{bmatrix} 1 & e^{-j2\pi\Delta f\tau_k} & ... & e^{-j2\pi(N-1)\Delta f\tau_k} \end{bmatrix}^T \\ &= \begin{bmatrix} z_k^0 & z_k^1 & ... & z_k^{N-1} \end{bmatrix}^T\end{aligned} \tag{2.10}$$

This means that we can model the received signal as Equation 2.5, in which the transmit signals is modeled as they are from d sources with different range between them and receiver, and matrix $\mathbf{A}$ that composed by vector $\mathbf{a}(\tau_k)$ can be regarded as a frequency response.

When considering the noise, the data model can be expressed as

$$\begin{aligned}\mathbf{x} &= \mathbf{h} + \mathbf{n} \\ &= \mathbf{A}\mathbf{s} + \mathbf{n}\end{aligned} \tag{2.11}$$

where $\mathbf{n} = [n_0 \quad n_1 \quad ... \quad n_{N-1}]^T$ and $n_i \sim \mathcal{N}(0, \sigma^2)$. Then

$$\mathbf{x} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} = \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{N-1} \end{bmatrix} + \begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_{N-1} \end{bmatrix} \tag{2.12}$$

At here we define $\mathbf{x}$ is a data vector which is length N and only has one snapshot (observation).

2.3 Data Smoothing using Shift Invariance

When we implement the subspace-based super-resolution technique, an important requirement for super-resolution algorithm is that the signals are uncorrelated with each other. However, multipath signals are coherent, leading to invalidation of the super-resolution algorithm [9]. Then smoothing technique can be taken into consideration to solve the coherent signals problem.

Besides, it is normally needed to compute the subspace by taking eigenvalue decomposition of the covariance matrix $\mathbf{R}$:

$$\mathbf{R} = E\{\mathbf{x}\mathbf{x}^H\} \tag{2.13}$$

which is the theoretical covariance matrix. In reality, the covariance matrix is estimated from measurement data, which is:

$$\hat{\mathbf{R}} = \frac{1}{M}\mathbf{x}\mathbf{x}^H \tag{2.14}$$

with M is the number of observations of $\mathbf{x}$. This causes that the accuracy of subspace relies on the estimate covariance matrix $\hat{\mathbf{R}}$ significantly. Under ideal condition, the covariance matrix is a Toeplitz matrix. However, the covariance matrix is estimated based on measurement data with limited length of observations, as a result, the estimate covariance matrix loses the property of Toeplitz, which is inaccurate. In addition, in our case, the number of observation is 1 and then the rank of the estimate covariance matrix is 1. As a result, there is only 1 column contained in subspace has data information, which cannot meet our requirement of super-resolution algorithm. Smoothing techniques can improve the property of Toeplitz and increase the rank of $\hat{\mathbf{R}}$ [10].

For these two reasons above, it is necessary to apply smoothing technique to measurement date vector. There are two smoothing techniques can be taken into consideration, namely Forward smoothing technique and Forward-Backward smoothing technique. From the property of shift invariance, a Hankel structured data matrix by smoothing techniques can be constructed.

2.3.1 Shift Invariance

For one steering vector $\mathbf{a}(\tau_k)$ is contained in matrix $\mathbf{A}$ (from Equation 2.10), if we set the first L elements in $\mathbf{a}(\tau_k)$ as a subvector $\bar{\mathbf{z}}_0$, the second L elements as $\bar{\mathbf{z}}_1$ and so on, which are

$$\bar{\mathbf{z}}_0 = \begin{bmatrix} z_k^0 & z_k^1 & ... & z_k^{L-1} \end{bmatrix}^T \tag{2.15}$$

$$\bar{\mathbf{z}}_1 = \begin{bmatrix} z_k^1 & z_k^2 & ... & z_k^{L} \end{bmatrix}^T \tag{2.16}$$

$$\bar{\mathbf{z}}_2 = \begin{bmatrix} z_k^2 & z_k^3 & ... & z_k^{L+1} \end{bmatrix}^T \tag{2.17}$$

$$\vdots$$

$$\bar{\mathbf{z}}_{N-L} = \begin{bmatrix} z_k^{N-L} & z_k^{N-L+1} & ... & z_k^{N-1} \end{bmatrix}^T \tag{2.18}$$

From these expressions it can be found that $\bar{\mathbf{z}}_1 = \bar{\mathbf{z}}_0 z_k$, $\bar{\mathbf{z}}_2 = \bar{\mathbf{z}}_1 z_k$, and so on. In general $\bar{\mathbf{z}}_m = \bar{\mathbf{z}}_0 z_k^m$, This means that the steering vector has the property of shift-invariance.

We write this expression into matrix form for the whole steering matrix $\mathbf{A}$. We set the

first L rows in $\mathbf{A}$ as a sub-matrix $\bar{\mathbf{Z}}_0$, second L rows as $\bar{\mathbf{Z}}_1$ and so on, which are

$$\bar{\mathbf{Z}}_0 = \begin{bmatrix} z_0^0 & z_1^0 & \cdots & z_{d-1}^0 \\ z_0^1 & z_1^1 & \cdots & z_{d-1}^1 \\ \vdots & \vdots & & \vdots \\ z_0^{L-1} & z_1^{L-1} & \cdots & z_{d-1}^{L-1} \end{bmatrix}, \tag{2.19}$$

$$\bar{\mathbf{Z}}_1 = \begin{bmatrix} z_0^1 & z_1^1 & \cdots & z_{d-1}^1 \\ z_0^2 & z_1^2 & \cdots & z_{d-1}^2 \\ \vdots & \vdots & & \vdots \\ z_0^{L} & z_1^{L} & \cdots & z_{d-1}^{L} \end{bmatrix}, \tag{2.20}$$

$$\bar{\mathbf{Z}}_2 = \begin{bmatrix} z_0^2 & z_1^2 & \cdots & z_{d-1}^2 \\ z_0^3 & z_1^3 & \cdots & z_{d-1}^3 \\ \vdots & \vdots & & \vdots \\ z_0^{L+1} & z_1^{L+1} & \cdots & z_{d-1}^{L+1} \end{bmatrix}, \tag{2.21}$$

$$\vdots$$

$$\bar{\mathbf{Z}}_{N-L} = \begin{bmatrix} z_0^{N-L} & z_1^{N-L} & \cdots & z_{d-1}^{N-L} \\ z_0^{N-L+1} & z_1^{N-L+1} & \cdots & z_{d-1}^{N-L+1} \\ \vdots & \vdots & & \vdots \\ z_0^{N-1} & z_1^{N-1} & \cdots & z_{d-1}^{N-1} \end{bmatrix} \tag{2.22}$$

with $N - L = M - 1$, and it can be obtained that $\bar{\mathbf{Z}}_1 = \bar{\mathbf{Z}}_0\mathbf{\Phi}$, $\bar{\mathbf{Z}}_2 = \bar{\mathbf{Z}}_1\mathbf{\Phi}$, ..., with

$$\mathbf{\Phi} = \begin{bmatrix} z_0 & 0 & \cdots & 0 \\ 0 & z_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_{d-1} \end{bmatrix} \tag{2.23}$$

and in general

$$\bar{\mathbf{Z}}_m = \bar{\mathbf{Z}}_0\mathbf{\Phi}^m \tag{2.24}$$

In conclusion, the steering matrix $\mathbf{A}$ has the property of shift-invariance. Then smoothing techniques can be taken into consideration. The smoothing techniques we consider about includes forward smoothing technique and forward-backward smoothing technique.

2.3.2 Forward Smoothing Technique

As mentioned above, indoor multipath signals are coherent. We need to take the decorrelation effect to the covariance matrix. At the same time, the covariance matrix is Hermitian and Toeplitz since the date measurement is assumed to be stationary. However, in reality the covariance matrix is estimated by data with limited length of observations, and then the property of Toeplitz lost [10]. The forward smoothing technique can be employed to solve coherent multipath signals, meanwhile, improve the property of Toeplitz of covariance matrix.

When processing data from an antenna array, a spatial smoothing technique is normally considered. The received data vector from antenna array can be divided into a certain number

of segments. Each segment is regarded as a subvector, and the length of each subvector are same. Each subvector generates a covariance matrix and the final estimated covariance matrix is taking the average of all the covariance matrices. This spatial smoothing technique is initially introduced in [11] and is improved and completed in [12, 13] to solve the coherent signals in array signal processing for direction finding. At here we call it forward smoothing technique.

The forward smoothing technique is introduced as follows. The received data vector is divided into M subvector and the length of each subvector is L Equation 2.11. Since total length of the received data vector is N, the relationship between these parameters is $M = N - L + 1$. For $m = 0, 1, 2, ..., M - 1$, the subvectors can be expressed as:

$$\mathbf{x}(m) = \begin{bmatrix} x_m \\ x_{m+1} \\ \vdots \\ x_{m+L-1} \end{bmatrix} = \begin{bmatrix} h_m \\ h_{m+1} \\ \vdots \\ h_{m+L-1} \end{bmatrix} + \begin{bmatrix} n_m \\ n_{m+1} \\ \vdots \\ n_{m+L-1} \end{bmatrix} \tag{2.25}$$

and then the final covariance matrix estimate can be expressed as:

$$\hat{\mathbf{R}} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{\mathbf{R}}_m = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{x}(m)\mathbf{x}(m)^H \tag{2.26}$$

The data matrix can be constructed as a Hankel matrix using subvectors $\mathbf{x}(m)$, which is

$$\mathbf{X} = \begin{bmatrix} x_0 & x_1 & \dots & x_{M-1} \\ x_1 & x_2 & \dots & x_M \\ \vdots & \vdots & \ddots & \vdots \\ x_{L-1} & x_L & \dots & x_{N-1} \end{bmatrix} \tag{2.27}$$

$$= \begin{bmatrix} \mathbf{x}(0) & \mathbf{x}(1) & \dots & \mathbf{x}(M-1) \end{bmatrix}$$

Using the property of shift invariance, Hankel structured data matrix can be expressed as:

$$\mathbf{X} = \begin{bmatrix} \bar{\mathbf{Z}}_0\mathbf{s} & \bar{\mathbf{Z}}_0\mathbf{\Phi}\mathbf{s} & \bar{\mathbf{Z}}_0\mathbf{\Phi}^2\mathbf{s} & \dots & \bar{\mathbf{Z}}_0\mathbf{\Phi}^{M-1}\mathbf{s} \end{bmatrix} + \mathbf{N} \tag{2.28}$$

with $\bar{\mathbf{Z}}_0$ is from Equation 2.19, $\mathbf{s}$ is from Equation 2.6, and $\mathbf{N}$ is the Hankel structure noise matrix:

$$\mathbf{N} = \begin{bmatrix} n_0 & n_1 & \dots & n_{M-1} \\ n_1 & n_2 & \dots & n_M \\ \vdots & \vdots & \ddots & \vdots \\ n_{L-1} & n_L & \dots & n_{N-1} \end{bmatrix} \tag{2.29}$$

Then the final covariance matrix estimate can be expressed as:

$$\hat{\mathbf{R}} = \frac{1}{M}\mathbf{X}\mathbf{X}^H \tag{2.30}$$

It is shown that in array signal processing when the number of subvectors is larger than or equal to the number of signals, the covariance matrix of signals is nonsingular [12]. In this case, when $M \geqslant d$ (d is the number of multipath signals), the forward smoothing technique can solve the coherent signals and then the super-resolution algorithm can be applied.

2.3.3 Forward-Backward Smoothing Technique

In the previous section, a data matrix is constructed from a data vector by forward smoothing technique. In order to further improve the property of Toeplitz of the estimate covariance matrix, forward-backward smoothing technique can be imaged.

Meanwhile, the forward-backward smoothing is employed for a more significant decorrelation effect. It has been shown that forward-backward smoothing technique can solve same number of correlated signals using less sensors. When the number of correlated signals is d, the number of sensors needed for forward-backward smoothing technique is $\frac{3}{2}d$, in comparison with $2d$ for forward smoothing technique [14, 15]. The comparison of the decorrelation effects between forward smoothing technique and forward-backward smoothing technique is shown and analyzed in [16, 17, 18, 10].

The forward-backward smoothing is firstly introduced in [17] to decorrelate the covariance matrix of coherent signals. At here, we try to apply it not only to take more decorrelation effects to the covariance matrix, but also to improve the property of Toeplitz of the covariance matrix estimate.

In forward-backward smoothing technique, an exchange matrix is introduced as

$$\mathbf{J} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{2.31}$$

Same as what we have done in forward smoothing technique, we divide the data vector into M subvectors with length is L. The estimate covariance matrix using each subvector $\mathbf{x}(m)$ is:

$$\hat{\mathbf{R}}_m = \mathbf{x}(m)\mathbf{x}(m)^H \tag{2.32}$$

with $\mathbf{x}(m)$ is a subvector from Equation 2.25. Then a new matrix $\hat{\mathbf{R}}_B$ can be generated using the exchange matrix for each $\mathbf{x}(m)$ as:

$$\hat{\mathbf{R}}_B = \mathbf{J}\hat{\mathbf{R}}_m^*\mathbf{J} \tag{2.33}$$

Then the smoothed covariance matrix estimate using forward-backward smoothing technique can be expressed as

$$\begin{aligned} \hat{\mathbf{R}}_{FB} &= \frac{1}{2M}\sum_{m=0}^{M-1}\left(\hat{\mathbf{R}}_m + \hat{\mathbf{R}}_B\right) \\ &= \frac{1}{2M}\sum_{m=0}^{M-1}\left(\hat{\mathbf{R}}_m + \mathbf{J}\hat{\mathbf{R}}_m^*\mathbf{J}\right) \end{aligned} \tag{2.34}$$

where '$*$' denotes the conjugation operation and M is the number of snapshots (observations) of the Hankel structure data matrix.

If we apply the result from forward smoothing technique (Equation 2.26), the covariance matrix estimate smoothed by forward-backward smoothing technique can be expressed as

$$\hat{\mathbf{R}}_{FB} = \frac{1}{2}\left(\hat{\mathbf{R}} + \mathbf{J}\hat{\mathbf{R}}^*\mathbf{J}\right) \tag{2.35}$$

Range Finding Algorithm 3

In localization field, using range parameter is one of the most popular approach. Localization based on range finding is realized by computing time-of-arrival (ToA) of the signal arriving at direct line-of-sight (DLoS). Due to the complexity of indoor environments, the ToA parameter of signal from DLoS cannot always be accurately computed [19, 20]. Then super-resolution algorithm can be employed on the frequency domain measurement to increase resolution for time dispersion estimate [21, 22, 23, 24].

In this chapter, subspace-based super-resolution algorithm (MUSIC algorithm) is proposed for range finding using frequency domain measurement. Then the requirements for subspace separation are proposed.

3.1 MUSIC Algorithm

MUSIC algorithm is the abbreviation of MUltiple SIgnal Classification algorithm and it is introduced for for direction finding and multiple frequency estimation [25, 26]. Using the data model developed in chapter 2, we can apply MUSIC algorithm for range finding by estimating time dispersion [21].

Using the data model proposed in chapter 2, radio measurement can be expressed as Equation 2.11. In absence of noise, it can be expressed as

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{3.1}$$

and then the covariance matrix is

$$\mathbf{R} = E\{\mathbf{x}\mathbf{x}^H\} = \mathbf{A}\mathbf{S}\mathbf{A}^H \tag{3.2}$$

where $\mathbf{S} = E\{\mathbf{s}\mathbf{s}^H\}$.

MUSIC algorithm is a subspace-based algorithm, it is needed to take eigenvalue decomposition of covariance matrix to generate subspace. We take eigenvalue decomposition to covariance matrix $\mathbf{R}$ and then

$$\begin{aligned} \mathbf{R} &= \mathbf{A}\mathbf{S}\mathbf{A}^H \\ &= \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \\ &= \begin{bmatrix}\mathbf{U_s} & \mathbf{U_n}\end{bmatrix} \begin{bmatrix}\mathbf{\Lambda_s} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix} \begin{bmatrix}\mathbf{U_s}^H \\ \mathbf{U_n}^H\end{bmatrix} \end{aligned} \tag{3.3}$$

where $\mathbf{U}$ is a unitary matrix contains eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix contains eigenvalues in non-increasing order ($\lambda_0 \geq \lambda_1 \geq ... \geq \lambda_{d-1} \geq 0$). Since here noise is absent, the matrix $\mathbf{\Lambda}$ only contains d eigenvalues with d is the number of multipath signals.

In real indoor environments, there is noise. We assume that noise is additional white Gaussian noise and then the data model can be expressed as in Equation 2.11, which is

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n}$$

We assume that the variance of white Gaussian noise is σ_n^2, and noise is independent with signals. Then the eigenvalue decomposition of covariance matrix $\mathbf{R}$ is

$$\begin{aligned}\mathbf{R} &= \mathbf{A}\mathbf{S}\mathbf{A}^H + \sigma_n^2\mathbf{I} \\ &= \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \\ &= \begin{bmatrix}\mathbf{U_s} & \mathbf{U_n}\end{bmatrix}\begin{bmatrix}\mathbf{\Lambda_s} + \sigma_n^2\mathbf{I_d} & \mathbf{0} \\ \mathbf{0} & \sigma_n^2\mathbf{I_{N-d}}\end{bmatrix}\begin{bmatrix}\mathbf{U_s}^H \\ \mathbf{U_n}^H\end{bmatrix}\end{aligned} \tag{3.4}$$

From the expression above, the subspace of covariance matrix can be separated into signal space $\mathbf{U_s}$ and noise space $\mathbf{U_n}$. Since the columns of $\mathbf{U_s}$ and the columns of $\mathbf{A}$ span the same signal space, the noise space $\mathbf{U_n}$ is orthogonal to the matrix $\mathbf{A}$, and this indicates that

$$\mathbf{U_n}^H\mathbf{A} = 0 \tag{3.5}$$

and this also implies that

$$\mathbf{U_n}^H\mathbf{a}(\tau_k) = 0 \tag{3.6}$$

To estimate time delay parameter τ is equal to find the minima of the cost function among a scanning τ. The cost function is

$$\begin{aligned}J(\tau) &= ||\mathbf{U_n}^H\mathbf{a}(\tau)||^2 \\ &= \mathbf{a^H}(\tau)\mathbf{U_n}\mathbf{U_n}^H\mathbf{a}(\tau)\end{aligned} \tag{3.7}$$

For MUSIC algorithm, rather than finding the minima of the above cost function, we normally try to find the maxima of the cost function that

$$\begin{aligned}J_{MUSIC}(\tau) &= \frac{1}{||\mathbf{U_n}^H\mathbf{a}(\tau)||^2} \\ &= \frac{1}{\mathbf{a}^H(\tau)\mathbf{U_n}\mathbf{U_n}^H\mathbf{a}(\tau)}\end{aligned} \tag{3.8}$$

In eigenvalue decomposition, if data is a $L \times M$ matrix, the signal space $\mathbf{U_s}$ is a $L \times d$ matrix while the noise space is a $L \times (L-d)$ matrix, where d is the number of multipath signals. However, when processing experiments, the number of multipath signals d is unknown. To obtain an accurate estimate of noise space, we need to make an accurate estimation of d.

MUSIC algorithm is based on the decomposition of covariance matrix, which means that the performance relies on the accuracy of the covariance matrix estimate. Before MUSIC algorithm is applied, smoothing techniques proposed in section 2.3 needed to be processed on the data vector.

3.2 Subspace Separation

To obtain a good performance when using the subspace-based super-resolution algorithm, the estimate of noise subspace $\hat{\mathbf{U}}_\mathbf{n}$ needs to be accurate. This means that before the super-resolution algorithm is applied, it is needed to separate signal subspace $\hat{\mathbf{U}}_\mathbf{s}$ and noise subspace

$\hat{\mathbf{U}}_{\mathbf{n}}$, which is equal to obtain the number of source signals d (or the number of multipath signals for our case).

Under ideal condition, the eigenvalues of the covariance matrix from i.i.d. white Gaussian noise are equal to the noise variance. Then the direct approach to obtain the number of signals is to take the number of eigenvalues that larger than noise variance as the number of signals. However, the covariance matrix used for eigenvalue decomposition is estimated by limited length of data, the eigenvalues from noise are not equal to each other. As a result, this approach fails. In order to separate the signal subspace and noise subspace by finding the number of signals (or the number of multipath signals), some other methods need to be considered.

In this article, two methods are proposed in order to realize subspace separation. The first proposed method is the MDL criteria algorithm. The second method is to find a threshold to the eigenvalues, we call it the threshold method. These two methods are proposed and tested in the following two chapters.

MDL Criteria Algorithm and Simulation Result

4

It is an important procedure to make an accurate estimation of the number of signals (or the number of multipath signals) before super-resolution algorithm is applied. In this chapter, firstly MDL criteria algorithm is introduced. Then the functionality of the MDL criteria for number of signals estimation is tested. When processing the test, experiments are firstly operated under the condition that the data contains multiple observations (snapshots). Afterwards, we apply MDL criteria algorithm to Hankel structured data matrix, in order to test the performance of the MDL criteria under Hankel structured data condition.

4.1 MDL Criteria Algorithm

Based on the eigenvalues of covariance matrix, an information theoretic criteria method introduced by Akaike in [27] for model selection and then based on his work Rissanen develop a shortest data description method in [28]. The information theoretic criteria including Akaike information theoretic criteria (AIC) and Rissanen minimum descriptive length (MDL) criteria can be employed to our case to estimate the number of multipath components [10]. Here we call this method as MDL criteria algorithm.

Using the smoothing model discussed in section 2.3 and obtain the covariance matrix estimate $\hat{\mathbf{R}}$. At here the matrix $\hat{\mathbf{R}}$ is a $L \times L$ matrix and the eigenvalue matrix Λ contain L eigenvalues in non-increasing order ($\lambda_0 \geq \lambda_1 \geq ... \geq \lambda_{L-1}$). For $k = 0, 1, 2, ..., L-1$, the expression of minimum descriptive length (MDL) algorithm is proposed as [29]

$$\mathrm{MDL}(k) = -M(L-k)\ln\left(\frac{\prod_{i=k}^{L-1}\lambda_i^{\frac{1}{L-k}}}{\frac{1}{L-k}\sum_{i=k}^{L-1}\lambda_i}\right) + \frac{1}{2}k(2L-k)\ln(M) \tag{4.1}$$

where M is the number of snapshots used in received data, then estimate of the number of multipath components d is equal to k that minimize the MDL function:

$$d = \arg\min_k \mathrm{MDL}(k) \tag{4.2}$$

when the estimate of the number of multipath d is equal to the real number of multipath p, the subspace-based MUSIC algorithm is expected to have the best performance, which implies that we want d as closer with p as possible.
It has shown that when applying forward-backward method to the covariance matrix estimate, the MDL criteria needs to be changed as the second term in equation 3.8 is rewritten as $\frac{1}{4}k(2L-k+1)\ln(M)$, and then the MDL criteria function for forward-backward method is [30] :

$$\mathrm{MDL}_{FB}(k) = -M(L-k)\ln\left(\frac{\prod_{i=k}^{L-1}\lambda_i^{\frac{1}{L-k}}}{\frac{1}{L-k}\sum_{i=k}^{L-1}\lambda_i}\right) + \frac{1}{4}k(2L-k+1)\ln(M) \tag{4.3}$$

4.2 MDL Criteria Algorithm Result

Since our research purpose is to realize localization under BLE specification, we need to evaluate MDL criteria under BLE specification. The BLE specification is showing in Table 4.1.

Parameter	**Values**
Frequency range	$2.4 \sim 2.48$ GHz
Frequency interval (Δf)	2 MHz
Number of channel tones (length of data vector N)	40

Table 4.1: BLE specification

To start with, MDL criteria algorithm is applied to measurement data that contains multiple observations (snapshots) and its performance is tested. Then MDL criteria algorithm is applied to Hankel structured data (from smoothing techniques). Before the smoothing techniques are applied, the corresponding parameters (L and M) for smoothing techniques are determined based on several limitations and reasons. Then experiments for functionality testing of MDL criteria algorithm on Hankel structured data are processed.

In following experiments, we assume that the received signal power at receiver antenna is 1. Then the parameter c_k in Equation 2.4 is set to be 1 for all k. The time delay parameter τ_k is randomly generated for each k, in order to ensure the generality of experiments. We set that $\tau_0 < \tau_1 < \cdots < \tau_{d-1}$.

4.2.1 Multiple Snapshots Contained in Measurement Data

At first, we test the effectiveness of MDL criteria algorithm under normal case: each signal has multiple observations, then the measurement data contains multiple observations (snapshots). Each signal has 100 observations, which means that the measurement data has 100 snapshots (M =100). The number of receiver antennas is set to be 40 ($L = N$ =40), this is because the smoothing techniques are not needed to be applied in this case. Then the measurement data matrix is a $L \times M$ matrix with dimensions are L=40 and M=100. The number of signals is d, and signals are set to be uncorrelated and separate enough when $d \geqslant 2$.

The power of each signal is set to be 1, and the SNR is 20dB per signal. The number of signals is set to be d =1, 2, 3, 5, 10, 15, 20, 30, 40 respectively. The MDL function plot (for one experiment run) is shown in Figure 4.1

From Equation 4.2 and Figure 4.1 it shows that the estimate number of signals are:

True number of signals d	1	2	3	5	10	15	20	30	40
Estimate number of signals $\hat{d}$	1	2	3	5	10	15	20	30	0

Table 4.2: MDL estimate result for normal data matrix

This means that MDL criteria can give an accurate estimate result under this condition. When d =40, MDL criteria cannot generate a valid result because it can only produce a valid result when $d < L$. In fact, if d is close to L (*i.e.*, d =38, 39), it is hard for MDL criteria to generate a correct estimate result. This is because when d is relatively large, the curve will increase before it decreases at around $k = d$. The decrease can be less significant than the

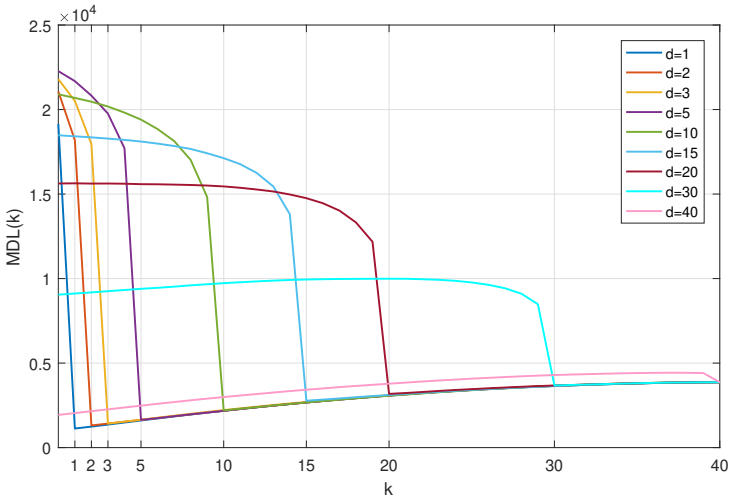


Figure 4.1: The MDL criteria function

increase (*i.e.*, d =37, 38, 39, result is showing in Figure A.1), then MDL criteria will fail under this condition. This problem can be solved by finding the local minimum of instead of the global minimum of MDL function.

Then the performance of MDL criteria algorithm with varying SNR is investigated. The parameter that the probability of detection (or probability of correct estimation) P_D is introduced to indicate the performance. The parameter P_D is computed as:

$$P_D = \frac{\text{Number of experiments that correctly estimate the signal number}}{\text{Total number of experiments}} \tag{4.4}$$

The SNR varies from -20 to 20dB, and for each SNR choice, 2000 Monte-Carlo experiments are conducted. The result is shown in Figure 4.2.

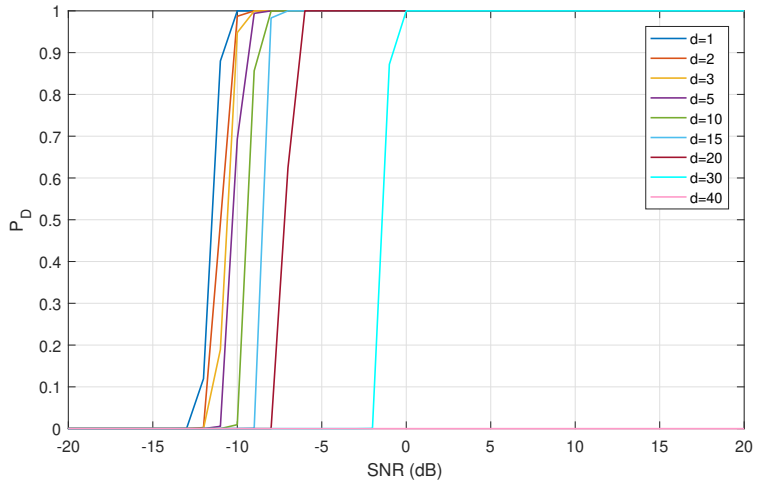


Figure 4.2: P_D-SNR plot for different number of signals

From Figure 4.2 it shows that with the increase of SNR, P_D increases, which indicates

that the performance of MDL criteria will be better with a relatively higher SNR. When the number of signals is relatively small (*i.e.*, $d < 5$), the MDL criteria can generate an accurate estimate under low SNR condition.

Besides, with MDL criteria, there is no false alarm control for the estimate result. For example, if there is only noise contained in measurement data, for different noise power (σ_n^2=0.01, 0.1, 0.5, 1, 2, 4 respectively), the MDL function plot (for one experiment run) is showing in Figure A.2 $\sim$ A.7 (It has shown that for a large number of Monte-Carlo experiments, the result plot curves have similar shape). Though for different values for σ_n^2, the result curves have similar shapes, and the estimate number of signals from MDL criteria is 0 for all experiments. Then it can be concluded that for noise only data, MDL criteria can produce an accurate result, which is that estimate number of signals is 0 since no signal contained in data. It also can be concluded that the performance of MDL criteria depends on SNR significantly.

In conclusion, when measurement data contains multiple snapshots, MDL criteria algorithm can estimate the number of signals with high accuracy.

4.2.2 Determination of L and M for Smoothing Techniques

In order to implement subspace-based super-resolution algorithm, the first step is to compute subspace by taking eigenvalue decomposition of the covariance matrix. To solve coherent signals problem and estimate the covariance matrix with high accuracy, use the smoothing method proposed in chapter 2 to reconstruct the data vector (length N) into an Hankel structured data matrix which is a $L \times M$ matrix with M is the number of snapshots (subvectors) and L is the length of each subvector. The relationship between these parameters is $N = M + L - 1$. Before processing the measurement data using super-resolution algorithm, it is necessary to determine the parameters of L and M.

Firstly, the smoothing techniques are applied for the purpose of solving coherent multipath signals problem. Assuming that the number of multipath signals is d, for forward smoothing technique, it is needed that $M \geqslant d$ and for the forward-backward smoothing technique, it only needs that $M \geqslant (1/2)d$.

Then limitations from respective of super-resolution algorithm also need to be taken into consideration. The dimension of the subspace $\mathbf{U}$ of covariance matrix $\mathbf{R}$ is $L \times L$, the dimension of subspace from signal $\mathbf{U}_s$ is $L \times d$, and the dimension of subspace from noise $\mathbf{U}_n$ is $L \times L - d$. According to super-resolution algorithm, subspace from noise $\mathbf{U}_n$ is needed. which means that $L > d$. It is normally to expect that there is large dimension of $\mathbf{U}_n$ to generate an accurate result. At the same time, the number of signals d normally is unknown. As a result, the parameter L needs to be set as large as possible.

Besides, MDL criteria and the threshold method are based on the eigenvalues of covariance matrix. The rank deficient of covariance matrix will influence the eigenvalues, and further influence the results of MDL criteria and the threshold method. In order to avoid this influence, it is normally to ensure that the covariance matrix is full rank, which means that $L \leqslant M$ for the Hankel structured data matrix.

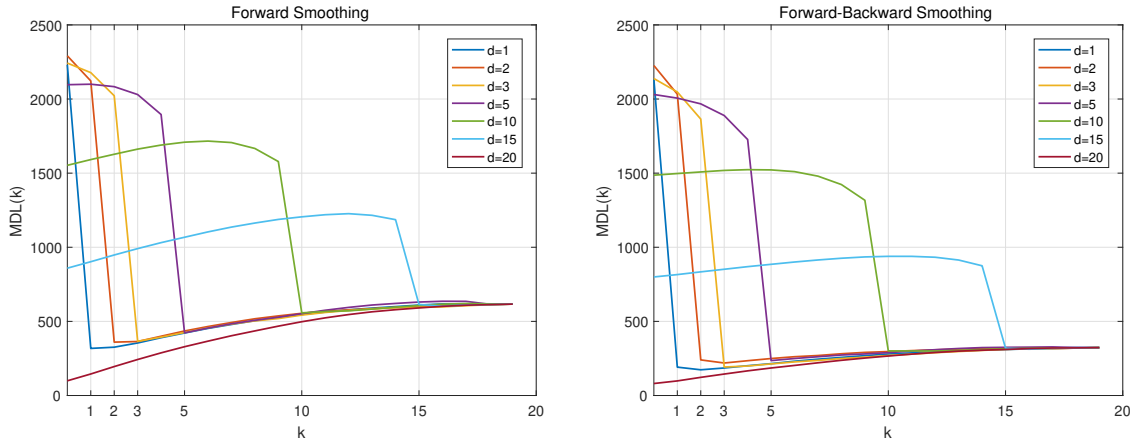
Another purpose for employment of smoothing technique is that to generate multiple snapshots for a more Toeplitz covariance matrix, and further to generate a more accurate subspace. From this respective, we require that M is as large as possible.

Taking the limitations and requirements listed above into consideration, we chose $L = (1/2)N$ when processing data under the BLE specification.

4.2.3 Hankel Structured Measurement Data

In this section, we evaluate the functionality of MDL criteria on Hankel structured data matrix. The data used is generated by the data model and smoothing techniques under the BLE specification which is shown in Table 4.1. Choose the parameter L as $L = (1/2)N$ as proposed in subsection 4.2.2 for the smoothing techniques. The number of signals here is the number of multipath signals, denoted as d. Setting the channel gain for each multipath signal c_k as 1. The power of each multipath signal is 1, and set SNR is 20dB per signal. The number of multipath signal is set to be d=1, 2, 3, 5, 10, 15, 20.

The results of MDL criteria function for forward smoothing case and forward-backward smoothing case are shown in Figure 4.3 and the related estimate result $\hat{d}$ are shown in Table 4.3



(a) Forward smoothing (b) Forward-Backward smoothing

Figure 4.3: the MDL criteria simulation result

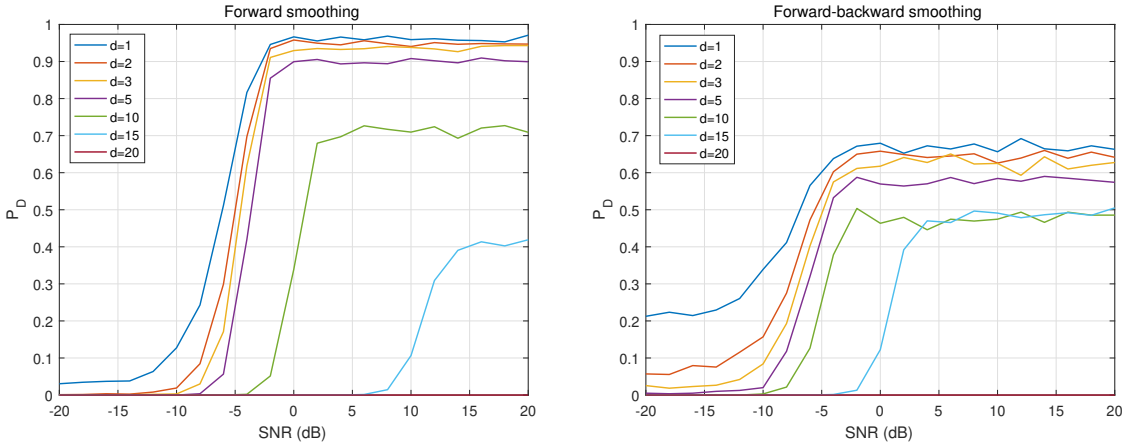
d	1	2	3	5	10	15	20
$\hat{d}$ (FS)	1	2	3	5	10	16	0
$\hat{d}$ (FBS)	2	3	3	5	12	16	0

Table 4.3: MDL estimate result for Hankel data matrix

From Figure 4.3 and Table 4.3 it can be observed that for the forward smoothing (FS) case the estimate results are relatively accurate when $d < 20$, especially when d is small. In comparison with the forward-backward smoothing (FBS) case, the deviation happens even though the d is small. It is notable that there is a slight increase before the function drops when $k \approx d$, this increase can be greater than the decrease at $k \approx d$ when d is relatively large under the low SNR condition (i.e., taking d=14, 16, 18 respectively, results are showing in Figure A.8). As a result, the minimum point always happen at $k = 0$. The extent of this increase for FS case is more significant than it in FBS case, which makes the performance of MDL for FS case is worse than FBS case. There is a possible approach to solve this problem, rather than finding the global minimum point of MDL function, we can try to find the local minima. However, this approach only works when the SNR is still at a considerable level.

If the SNR continues to decrease, more fluctuations will be introduced and MDL criteria algorithm will fail.

In order to show the performance of MDL criteria for Hankel structured data with varying SNR, the P_D proposed by Equation 4.4 is computed. The SNR varies from -20 to 20dB. For each SNR choice, 2000 Monte-Carlo simulation experiments are conducted. The result is showing in Figure 4.4.



(a) Forward smoothing

(b) Forward-Backward smoothing

Figure 4.4: P_D-SNR result for MDL criteria

From Figure 4.4 it can be concluded that for FS case, with the increase of SNR, the P_D increases. When the number of multipath signals d is relatively small (*i.e.*, $d \leqslant 5$), with the increase of SNR, P_D can reach to a high level, which indicates that the MDL criteria can produce an accurate result. If d increases, the performance degrades. For example, when d =10 and 15, P_D can only reach to 0.7 and 0.4 respectively even if SNR is high. For FBS case, with the increase of SNR, P_D increases. However, P_D only fluctuates at a relatively low level even though d is small and SNR is high. Besides, under low SNR condition, if d is large, the FBS can generate a better performance than the FS. For example, if we use FBS, when d =10 and 15, P_D starts to increase at a lower SNR and P_D for d =15 reaches to a higher level.

Another thing noticeable is that even though there is a small number of signals (*i.e.*, d =1, 2, and 3), P_D cannot reach to 1 with the increase of SNR. For example, for FS case, this number fluctuate around 0.95 and cannot not reach to 1. This indicates that even if SNR is considerable high, MDL criteria still make mistakes. The reason for this is that, from 4.3(a) the curves have an extremely slight increase at the area that k just passes d, and sometimes decrease happens during this period. Then it leads to mistakes for estimating d. This decrease happens easily and more significantly for FBS case (from Figure 4.3), which explains that P_D cannot reach to a satisfied level even SNR is high.

In conclusion, for FS case, when the number of multipath signals d is small, the MDL criteria estimate result is reliable under the condition that the SNR is at a relatively high level (*i.e.*, SNR$\geqslant$ 0). By contrast, for the FBS case, the performance of MDL criteria is not satisfied since it cannot estimate the number of multipath signals with a high probability

even though the number of multipath signals is small and the SNR is high. As a result, it is preferential that using MDL criteria algorithm cooperating with forward smoothing technique.

4.3 Conclusion

In this chapter, MDL criteria algorithm for subspace separation is proposed, and the performance of it is evaluated by simulation experiments.

When data contains multiple snapshots, MDL criteria algorithm can produce a reliable result. With the increase of SNR, the performance of MDL criteria can be improved.

If data matrix is Hankel structured, two cases are investigated, the first one is only forward smoothing (FS) is applied and the second is that forward-backward smoothing (FBS) is applied. For FS case, with the increase of SNR, when the number of multipath signals d is small, the probability of detection P_D can reach to a high value (larger than 0.9). For FBS case, P_D can only reach to a relatively low level even if SNR is high. Overall, MDL criteria algorithm is more reliable when cooperating with forward smoothing technique.

The Threshold Method and Simulation Result

5

In this chapter, another approach, namely the threshold method, will be proposed for subspace separation. In the first section, the threshold method is introduced mathematically. In the second section, the functionality of the threshold method is tested based on simulation data. It will be shown that the threshold method is not ideally proper for Hankel structured data model, and an approach that adjusts the threshold to a reasonable level will be introduced in this section.

5.1 The Threshold Method

The basic ideal of the threshold method is setting a threshold to the eigenvalues of the covariance matrix. The number of eigenvalues that are larger than this threshold is regarded as the number of signals. Under ideal condition that measurement data contains infinite observations, the eigenvalues from noise are same and they are equal to the noise variance. Then the threshold can be equal to the noise variance. However, in practical the data has limited length, then the eigenvalues from noise are no longer equal to each other. As a result, this method fails in practical.

Another approach to compute the threshold needs to be taken into consideration. Normally, assuming that the eigenvalues from signal are much larger than the eigenvalues from noise. Under this assumption, it is reasonable to take the largest eigenvalue from noise as the threshold, as showing in Figure 5.1.

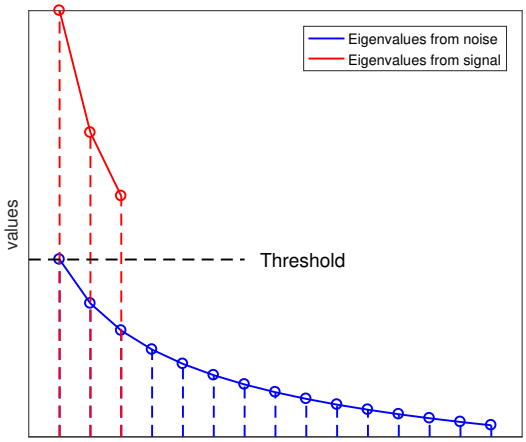


Figure 5.1: The threshold indication

We assume that noise is random complex white Gaussian noise with the covariance matrix

is

$$\mathbf{R}_n = \sigma_n^2 \mathbf{I} \tag{5.1}$$

with σ_n^2 is the variance of noise. If the noise is a noise matrix (denoted by $\mathbf{N}$) with dimension $L \times M$, the estimate covariance is

$$\hat{\mathbf{R}}_n = \frac{1}{M} \mathbf{N}\mathbf{N}^H \tag{5.2}$$

The largest eigenvalue of $\hat{\mathbf{R}}$ is denoted by $\hat{\lambda}_1$, when $L \to \infty$, it has shown that [31, 32]

$$\hat{\lambda}_1 \to \sigma_n^2 \rho^2 \tag{5.3}$$

with

$$\rho = \left(1 + \frac{\sqrt{L}}{\sqrt{M}}\right) \tag{5.4}$$

At here we introduce the unrefined threshold denoted as ϵ_0 and computed as:

$$\epsilon_0 = \sigma_n^2 \left(1 + \frac{\sqrt{L}}{\sqrt{M}}\right)^2 \tag{5.5}$$

Normally, we can take ϵ_0 as the threshold. However, the unrefined threshold ϵ_0 is not an upper bound of the largest eigenvalue.

Since the threshold needs to be an accurate upper bound of the largest eigenvalue from noise, ϵ_0 is not a proper threshold. In [32], a refinement to compute a more reasonable threshold is proposed, in which the unrefined threshold ϵ_0 is scaled by a factor β^2.

In order to compute the refined threshold, the distribution of the largest eigenvalue needs to be known. It has proposed that the distribution can be approximated by writing the centering and scaling constant:

$$\mu = \left(\sqrt{M + a_1} + \sqrt{L + a_2}\right)^2 \tag{5.6}$$

$$\nu = \sqrt{\mu} \left(\frac{1}{\sqrt{M + a_1}} + \frac{1}{\sqrt{L + a_2}}\right)^{\frac{1}{3}} \tag{5.7}$$

If we write expression

$$w := \frac{M\hat{\lambda}_1/\sigma_n^2 - \mu}{\nu} \tag{5.8}$$

it approximately converges to Tracy-Widom distribution ($\mathcal{TW}_\eta$) [33, 34]. The Tracy-Widom distribution and its importance for computing distribution of the largest eigenvalue of random matrices are proposed in [35, 36, 37]. The order of Tracy-Widom can be η =1, 2 and 4 [38, 39, 34]. When noise is real Gaussian, $\eta = 1$ and $a_1 = a_2 = -1/2$ in Equation 5.6 and 5.7 [34]. When noise is complex Gaussian, $\eta = 2$ and $a_1 = a_2 = 0$ in Equation 5.6 and 5.7 [40, 41]. At here, noise is assumed to be random complex white Gaussian noise and the expression of centering μ and scaling constant ν can be rewritten as:

$$\mu = \left(\sqrt{M} + \sqrt{L}\right)^2 = M\rho^2 \tag{5.9}$$

$$\nu = \sqrt{\mu}\left(\frac{1}{\sqrt{M}} + \frac{1}{\sqrt{L}}\right)^{\frac{1}{3}} = M^{\frac{1}{2}}L^{-\frac{1}{6}}\rho^{\frac{4}{3}} \tag{5.10}$$

and w approximately converges to Tracy-Widom distribution of order 2, which is:

$$w := \frac{M\hat{\lambda}_1/\sigma_n^2 - \mu}{\nu} \sim \mathcal{TW}_2 \tag{5.11}$$

It has shown that this approximation is more accurate when the value of the largest eigenvalue is relatively large, which is the interested part for threshold computing [42]. The threshold can be computed by setting the probability that the expected largest eigenvalue is larger than the threshold is 0, which is $P(\hat{\lambda}_1 > \epsilon) = 0$. Using CDF of the Tracy-Widom distribution, it can be expressed as

$$\begin{aligned} &P\left(\frac{M\hat{\lambda}_1/\sigma_n^2 - \mu}{\nu} > \frac{M\epsilon/\sigma_n^2 - \mu}{\nu}\right) = 0 \\ &\Leftrightarrow \frac{M\epsilon/\sigma_n^2 - \mu}{\nu} = F_2^{-1}(1) \end{aligned} \tag{5.12}$$

where $F_2^{-1}(\cdot)$ is the inverse CDF, and then the threshold ϵ can be expressed as:

$$\epsilon = \sigma_n^2\rho^2\beta^2 \tag{5.13}$$

with

$$\beta^2 = 1 + M^{-\frac{1}{2}}L^{-\frac{1}{6}}\rho^{-\frac{2}{3}}f_1 \tag{5.14}$$

where $f_1 = F_2^{-1}(1) \approx 2.24$. Then the upper bound of the largest eigenvalue is computed and we can this upper bound as the refined threshold.

The lower bound of the smallest eigenvalue can also be estimated from reflected Tracy-Widom distribution of order 2 [43, 32]. Since here we try to find the threshold on eigenvalues that separate the signal space and noise space, we are only interested in the upper bound.

When applying this refined threshold ϵ, the number of eigenvalues that larger than this threshold is regarded as the number of signals.

5.2 The Threshold Method Result

In this section, the functionality of the threshold method proposed in section 5.1 is evaluated. Similar process will be operated as in the MDL criteria experiments: At first, we consider the case that the signals contain multiple observations and there are multiple snapshots (observations) contained in data matrix, the smoothing techniques are not applied. Under this situation, we evaluate the functionality of the threshold method.

Then we consider about the second case that there is only a data vector, the data matrix is Hankel structured from smoothing techniques. According to section 5.1, the threshold is based on the distribution of the largest eigenvalue of the covariance matrix from white Gaussian noise. It will be shown later that the distribution of the largest eigenvalue will be

changed by smoothing techniques, and as a result of this, the threshold needs to be adjusted. In this section, a practicable approach based on experimental results is introduced to adjust the threshold to a reasonable level.

In this section, since the threshold is based on the largest eigenvalue of the covariance matrix of white Gaussian noise, for both two cases mentioned above, the experiments are firstly processed under the condition that the data only contains pure random white Gaussian noise, in order to test whether the threshold is reasonable and proper, after which we can consider about the situation that the data contains signal and noise.

5.2.1 Multiple Snapshots Contained in Measurement Data

In this section, data is generated with multiple observations (length M), the number of receiver antennas (responses) is L, then the data matrix is a $L \times M$ matrix. Firstly, the experiments are processed on the data that only contains random white Gaussian noise. Then the threshold method is applied to the data that contains signal and noise.

5.2.1.1 Pure White Gaussian Noise Case

At first, the generated data only contains a pure random complex white Gaussian noise, which is

$$\begin{aligned}\mathbf{X} &= \mathbf{N} \\ &= \begin{bmatrix} n_0(1) & n_0(2) & \dots & n_0(M) \\ n_1(1) & n_1(2) & \dots & n_1(M) \\ \vdots & \vdots & \ddots & \vdots \\ n_{L-1}(1) & n_{L-1}(2) & \dots & n_{L-1}(M) \end{bmatrix}\end{aligned} \tag{5.15}$$

with $n_i \sim \mathcal{N}(0, \sigma_n^2)$. And the covariance matrix estimate is

$$\hat{\mathbf{R}} = \frac{1}{M}\mathbf{N}\mathbf{N}^H \tag{5.16}$$

We investigate the relationship between the threshold ϵ proposed by Equation 5.13 and the largest eigenvalue λ_1 of the covariance matrix $\hat{\mathbf{R}}$. Since the threshold is the upper bound of the largest eigenvalue, it is normally expected that $\epsilon \geqslant \lambda_1$.

In this experiment, we assume that we have 40 tones can be used, then we choose $L = 40$, M varies from 20 to 80 (interval is 5 for a clear plot). The variance of white Gaussian nois is $\sigma_n^2 = 1$. For each choice of M, 5000 Monte-Carlo experiments are conducted. For each experiment, the white Gaussian noise data is randomly generated. We use Equation 5.13 to calculate the threshold ϵ, and plot the related curves with varying M. The result is shown in Figure 5.2

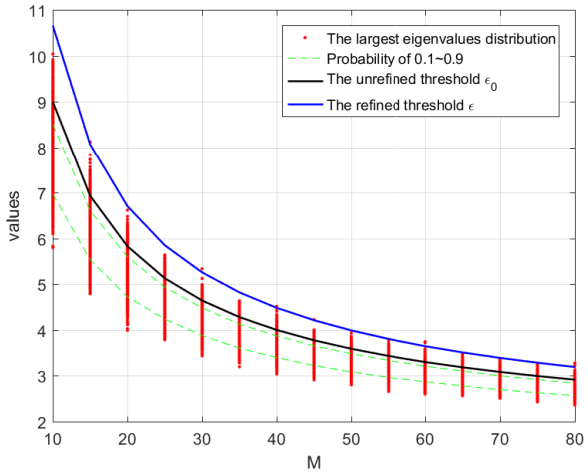


Figure 5.2: Pure random white Gaussian noise case with varying M

In Figure 5.2, the unrefined threshold ϵ_0 is computed by Equation 5.5. It shows that because of the variance, there are a considerable number of eigenvalues are larger than ϵ_0, then it is not proper to using ϵ_0 as the threshold. In comparison with the refined threshold ϵ, ϵ is larger than almost all the largest eigenvalue. There is only an extremely small amount of points are larger than ϵ, indicating that the refined threshold ϵ can allow more variance and it is a proper threshold. Besides, this also indicates that the probability in Equation 5.12 is approaching to 0.

It is shown that this threshold ϵ is still proper when noise variance σ_n^2 changed. We choose σ_n^2=0.01, 0.1, 0.5, 2 separately and repeat this experiments. The result plots are shown in Appendix B, Figure B.1 $\sim$ B.4.

It is also shown that this threshold is proper when choosing different L. we choose L=5, 10, 20 and 30 separately and keep σ_n^2 =1, the results are shown in Appendix B, Figure B.5 $\sim$ B.7.

5.2.1.2 Signal and Noise Case

In this section, we consider about the case that measurement data contains both signal and additional white Gaussian noise. Use Equation 5.13 to calculate the threshold ϵ. The number of eigenvalues that larger than the threshold ϵ is regarded as the number of signals $\hat{d}$.

In order to illustrate the performance of refined threshold for finding the number of signals, compute the probability of detection (P_D) introduced by Equation 4.4 under different SNR conditions. In these experiments, set the number of snapshots (observations) of data is M=100, the number of antennas (array responses) is L=40. The SNR ranges from -20 to 20dB. The number of signals is set to be d=1, 2, 3, 5, 10, 15, 20, 30 and 40 respectively. The P_D-SNR plot result is shown in Figure 5.3

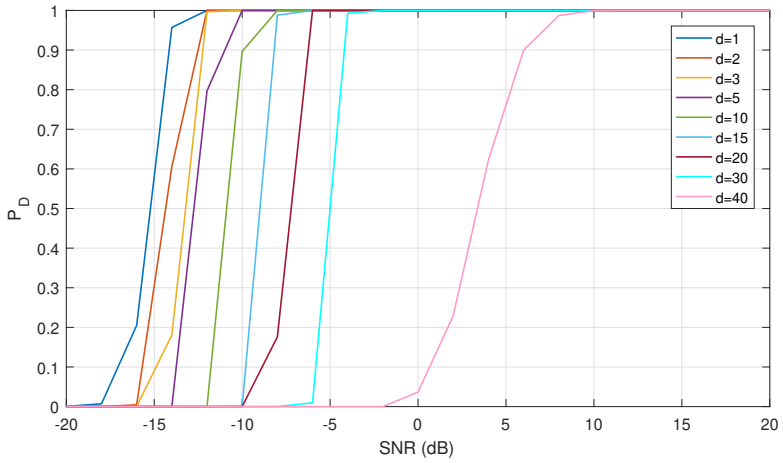


Figure 5.3: P_D-SNR result for different choices of d

From Figure 5.3 it shows that the refined threshold ϵ can generate an accurate result of number of signals, since the P_D will reach to 1 even under the condition of low SNR and large number of signals. When comparing this result to the result from MDL criteria under same conditions (Figure 4.2), it can be shown that P_D computed by the threshold method starts to increase at a lower SNR and reaches to 1 at a lower SNR, which indicates that the threshold method produces an accurate result under a lower SNR condition than MDL criteria. Therefore, the threshold method is preferred.

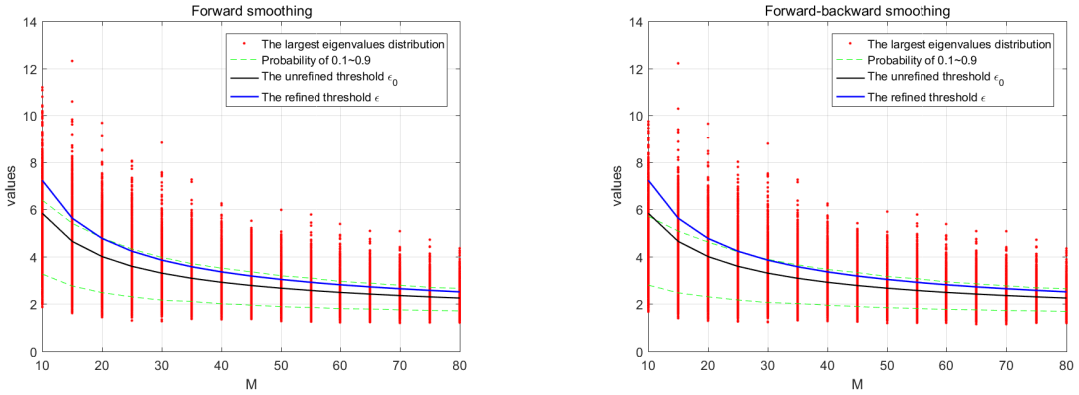
In conclusion, under the condition that the measurement data contains multiple snapshots (observations), the refined threshold ϵ is a tight upper bound of the largest eigenvalue of covariance matrix. Then the refined threshold ϵ can generate a reliable estimate for the number of signals.

5.2.2 Hankel Structured Measurement Data

In this section, Hankel structured measurement data case is taken into consideration. To start, we assume that the data only contains random white Gaussian noise. Afterwards, the threshold is applied to the case that the data contains signal and noise in order to realize subspace separation.

5.2.2.1 Pure White Gaussian Noise with Smoothing

At first, experiments are processed on the data that only contains random white Gaussian noise, in order to test whether the threshold computed by Equation 5.13 is a proper upper bound of the largest eigenvalue. Hankel structure is obtained by smoothing techniques. In this experiment, we choose $L = 20$ unchanged and M varying from 10 to 80, the total length of the original data sequence N is calculated as $N = M+L-1$. For each M 5000 Monte-Carlo experiments are conducted and results are shown in Figure 5.4.



(a) Forward smoothing (b) Forward-backward smoothing

Figure 5.4: Pure random white Gaussian noise data with smoothing techniques

From Figure 5.4 it shows that there is a considerable number of the largest eigenvalues that are larger than the refined threshold ϵ, which makes ϵ is no longer a proper threshold.

In conclusion, Hankel structure will change the distribution of the largest eigenvalue. Then the refined threshold ϵ computed by Equation 5.13 is not proper for the Hankel structured data. In order to solve this problem, the threshold needs to be further adjusted to a reasonable level.

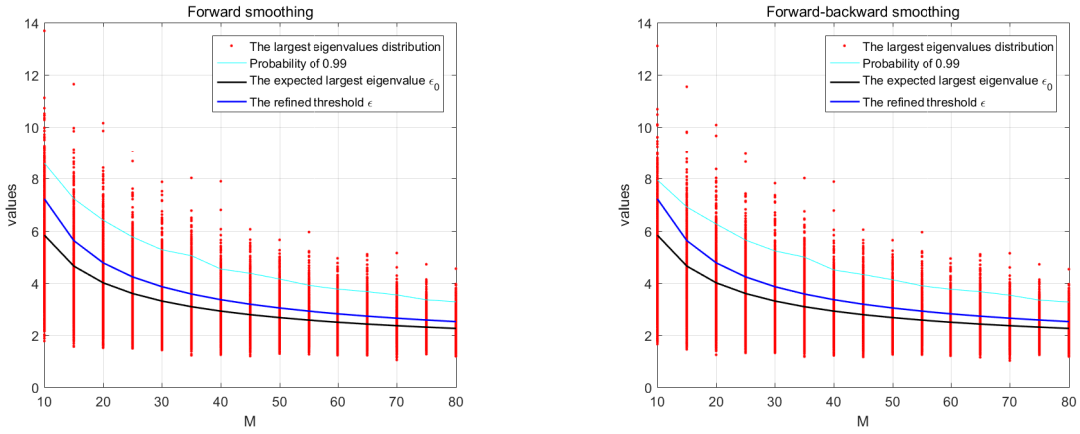
5.2.2.2 Threshold Adjustment Approach

In this experiment, we want to adjust the threshold to a proper level. The requirement is that we need to ensure that the threshold after adjustment is a tight upper bound of the largest eigenvalue.

At here, we set the adjusted threshold is larger than the largest eigenvalues from noise with the the probability of 0.99, which indicate that the threshold only allow a constant false alarm rate (CFAR) of 1% (probability of 0.01). In this experiment, we define a proper threshold is that a threshold with a CFAR of 1%. This can be expressed as

$$P(T_a > \lambda_1) = 0.99 \tag{5.17}$$

with T_a is the threshold after adjustment and λ_1 is the largest eigenvalue of noise covariance matrix. The adjusted threshold T_a is required to overlap with the bound line of cumulative probability of 0.99 (cyan line showing in Figure 5.5). The related parameters are L=20, σ_n^2=1.



(a) Forward smoothing (b) Forward-backward smoothing

Figure 5.5: Pure random white Gaussian noise data with smoothing techniques

From the above figures, for L=20, it is needed to raise the refined threshold ϵ (blue line) to the position where the 0.99 probability bound (cyan line) is.

An approach can be taken into consideration, which will be proposed in followings.

From Figure 5.5, the result distributions of the largest eigenvalues for forward smoothing case and forward-backward smoothing case are similar. In followings, take the forward smoothing case as an example to explain the adjustment approach.

From the expression of the refined threshold in Equation 5.13, which is:

$$\epsilon = \sigma_n^2 \rho^2 \beta^2$$

with σ_n^2 is the noise variance and

$$\beta^2 = 1 + M^{-\frac{1}{2}} L^{-\frac{1}{6}} \rho^{-\frac{2}{3}} f_1$$

$$\rho = 1 + \sqrt{\frac{L}{M}}$$

In the expression of β^2, there is a constant f_1 ≈2.24 from the Tracy-Widom distribution in the second term. Inspired by this, the approach to raise the threshold is that scaling the second term of β^2 by another factor α, then the threshold after adjustment (T_a) can be expressed as:

$$T_a = \sigma_n^2 \rho^2 \beta_a^2 \tag{5.18}$$

with

$$\beta_a^2 = 1 + M^{-\frac{1}{2}} L^{-\frac{1}{6}} \rho^{-\frac{2}{3}} f_1 \alpha \tag{5.19}$$

Then the scaling factor α can be computed using the known parameters. An important assumption: for a certain L, assuming that the factor α will not change with varying M. The observation of cumulative probability of 0.99 is denoted as O, and it is expected to satisfy that $O \approx T_a$.

Write above expression into vector form with different value of M:

$$\mathbf{O} = \mathbf{T}_{un} \cdot \boldsymbol{\beta}_a^2 \tag{5.20}$$

$$= \mathbf{T}_{un} \cdot (\mathbf{1} + \mathbf{M}^{-\frac{1}{2}} L^{-\frac{1}{6}} \rho^{-\frac{2}{3}} f_1 \alpha) \tag{5.21}$$

with

$$\mathbf{O} = [O_{M10}, O_{M15}, O_{M20}, \cdots, O_{M80}]^T \tag{5.22}$$

$\mathbf{M}$ is a vector containing different values for M, $\mathbf{T}_{un}$ is a vector containing T_{un} that computed by different values for M, and T_{un} is the unrefined threshold computed by Equation 5.5 ($T_{un} = \epsilon_0$), the operation '$\cdot$' denotes pointwise product.

If we write

$$\beta_0 = M^{-\frac{1}{2}} L^{-\frac{1}{6}} \rho^{-\frac{2}{3}} f_1 \tag{5.23}$$

then

$$\mathbf{O} = \mathbf{T}_{un} \cdot (\mathbf{1} + \boldsymbol{\beta}_0 \alpha) \tag{5.24}$$

and the factor α can be computed as

$$\alpha = (\mathbf{T}_{un} \cdot \boldsymbol{\beta}_0)^{\dagger} (\mathbf{O} - \mathbf{T}_{un}) \tag{5.25}$$

To evaluate this result, a similar experiment is taken out. For L=20, the noise variance is σ_n^2=1, the factor α is computed as $\alpha \approx$3.79 and the result plot is shown in Figure 5.6.

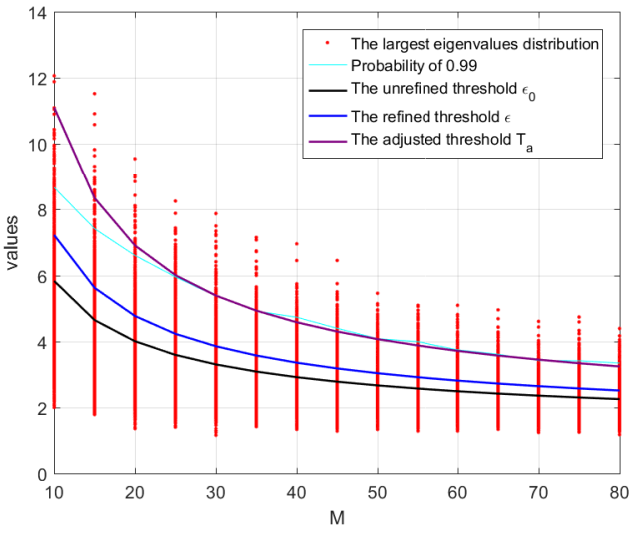


Figure 5.6: The result threshold by adjustment

From Figure 5.6 it shows that the adjusted threshold T_a overlaps with the bound line of probability of 0.99 to a high extent when $M \geqslant$20. Since L=20, the T_a is much larger than the 0.99 probability observation line when $M < 20$, this is from that the rank deficient influences the distribution of the largest eigenvalue. The figure also shows that the scaling factor α is proper for different value of M, which means that the assumption we made before the iterations is reasonable.

If noise variance σ_n^2 is changed, it can be shown that the scaling factor is still proper. Choosing different values for σ_n^2=0.01, 0.1, 0.5 and 2. Set L=20 is unchanged. The results are showing in Figure B.9 $\sim$ B.12.

5.2.2.3 Fitting $\alpha - L$

From the discussion above, the scaling factor α will not change with different value of M. However, it shows that the scaling factor α will change with the change of L. For example, in Figure 5.7 and Figure 5.8, take L=10 and 30 separately and repeat the experiment, the scaling factor α computed using L=20 is not proper anymore.

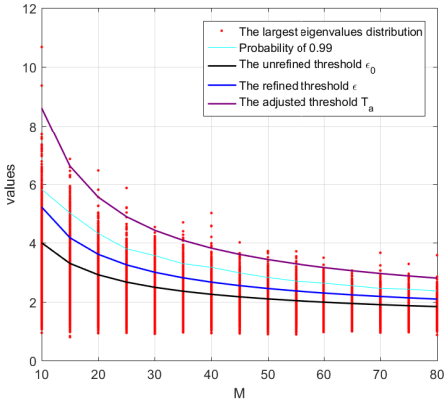


Figure 5.7: The largest eigenvalue distribution when L =10

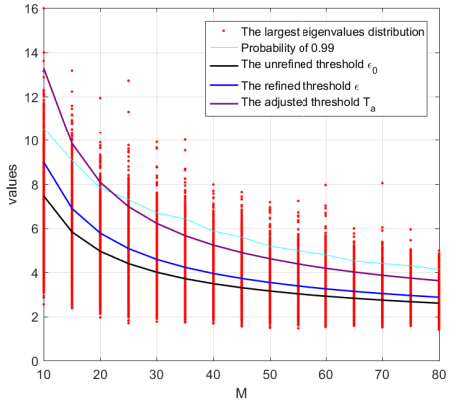


Figure 5.8: The largest eigenvalue distribution when L =30

In order to describe the relationship between the scaling factor α and L, repeat the experiment with different choice of L. In following experiments, L varies from 5 to 100 (with interval 5), and for each L, using M ranges from 5 to 200 (when applying fitting process, only use the data that generated from $M \geqslant L$). The result of $\alpha - L$ relationship is shown in Figure 5.9.

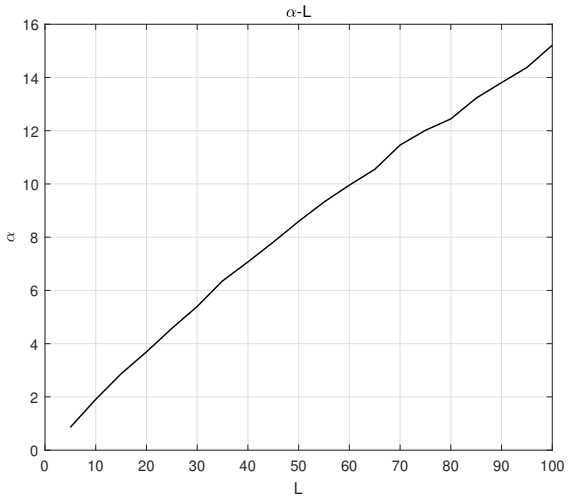


Figure 5.9: The relationship between L and α

From Figure 5.9, it shows that with the increase of L, the scaling factor α increases with an approximately linear relationship. Then we try to fit this observed relationship using a function that

$$\begin{aligned} \alpha &= f(L) \\ &= c_1 L^p + c_2 \end{aligned} \tag{5.26}$$

with $p \leqslant 1$. Write this expression into vector form for different value of L:

$$\boldsymbol{\alpha} = \mathbf{L}_{\text{fit}}\mathbf{c} \tag{5.27}$$

where

$$\boldsymbol{\alpha} = [\alpha_{L5}, \alpha_{L10}, ..., \alpha_{L100}]^T \tag{5.28}$$

$$\mathbf{L}_{\text{fit}} = \begin{bmatrix} L_5^p & 1 \\ L_{10}^p & 1 \\ \vdots & \vdots \\ L_{100}^p & 1 \end{bmatrix} \tag{5.29}$$

$$\mathbf{c} = [c_1, c_2]^T \tag{5.30}$$

Fit the relationship between α and L shown in Figure 5.9 using the function expressed as Equation 5.27, the fitting result is shown in Figure 5.10

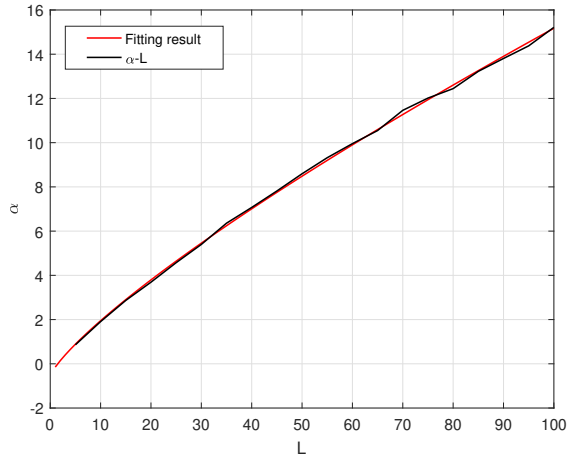


Figure 5.10: Fitting result

with the parameter in Equation 5.26 are

$$p = \frac{4}{5}$$
$$\mathbf{c} = [c_1, c_2]^T \approx [0.3948, -0.543]^T$$

From Figure 5.10 it shows that the fitting result (red line) overlaps with the experimental $\alpha - L$ to a high extent. It is expected that using the function Equation 5.26 and related value of parameter, the scaling factor α can be computed for a certain given L.

5.2.2.4 Discussion

In order to show the performance of this fitting result, we try to compute the probability of false alarm P_{FA} for different dimension data matrix by Monte-Carlo experiments, the P_{FA} factor is computed as:

$$P_{\text{FA}} = \frac{\text{Number of experiments of eigenvalues that are larger than } T_a}{\text{Total number of experiments}} \tag{5.31}$$

Here the threshold T_a is computed by the scaling factor α which is from the fitting result by Equation 5.26.

For different choices of dimension of data matrix, the parameter $M = 5, 10, ..., 100$ and $L = 5, 10, ..., 100$. For each pair of M and L, 5000 Monte-Carlo experiments are conducted and the result is shown in Figure 5.11. The fitting result is computed by CFAR of 1% ($P_{\text{FA}} = 0.01$).

From Figure 5.11 it shows that at the area of $M \geqslant L$, the result probabilities of false alarm P_{FA} fluctuate around 0.01, which is equal to the CFAR we set. At the area of $M < L$, the P_{FA} are equal (or approximately equal) to 0, which means that the threshold is not tight when $M < L$. This is because the covariance matrix is rank deficient when $M < L$, which leads to that the largest eigenvalue is smaller in comparison with the case that the covariance matrix is full rank. This conclusion can also be made by 2D plot showing in Figure 5.12 (only $M = 20, 30, 40, 50, 60$ selected for a clear plot), from which it shows that the result P_{FA} fluctuates around 0.01 when $M \geqslant L$.

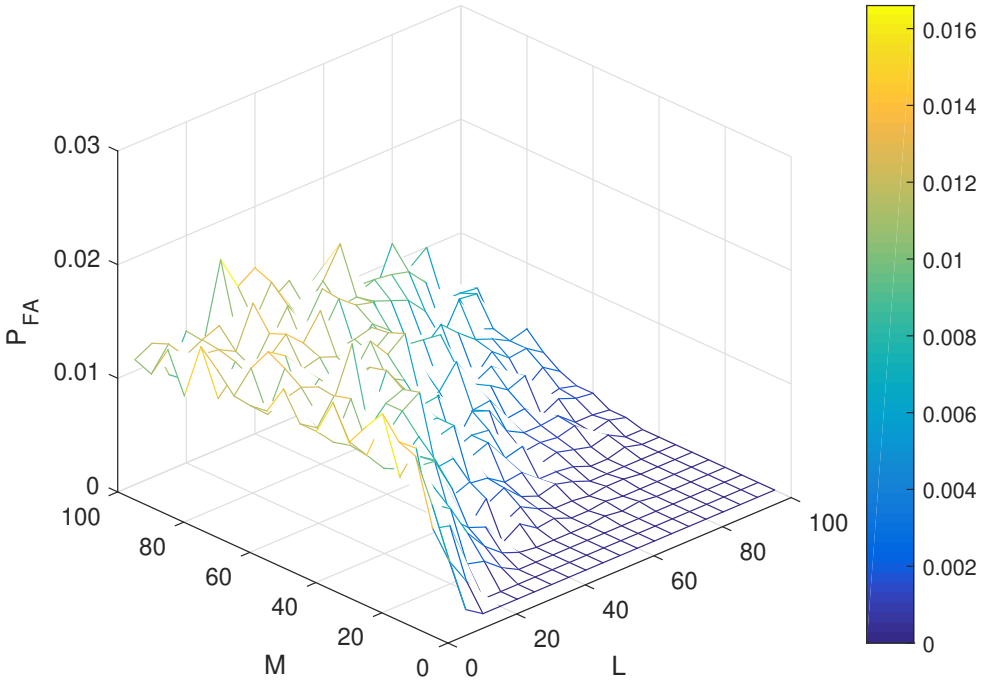


Figure 5.11: P_{FA} mesh plot

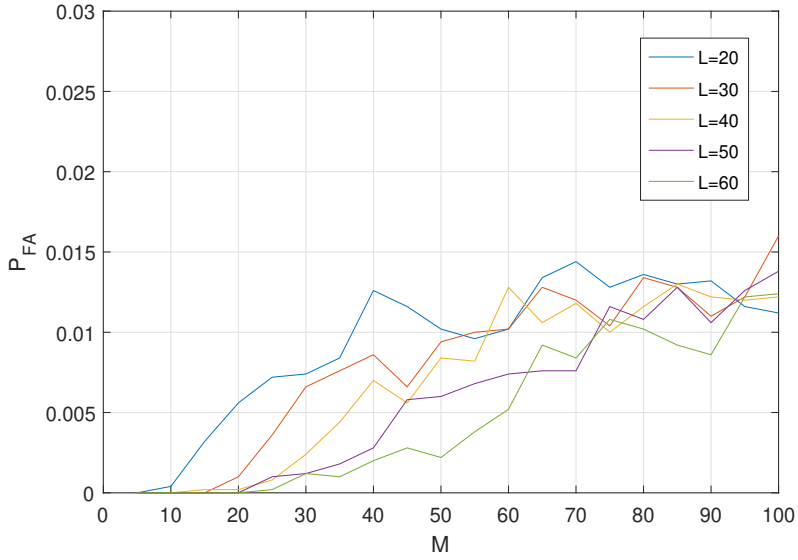


Figure 5.12: P_{FA} 2D plot

In conclusion, the scaling factor α computed by fitting method can adjust threshold to a reasonable level with a high reliability. As a result, the threshold T_a can be computed for a Hankel data matrix when the dimension (L and M) of this data matrix is given. If the requirement of probability of false alarm P_{FA} changed, the scaling factor α can still be computed by this fitting method.

Based on these results, we can consider about the case that the Hankel data contains both signal and noise.

5.2.2.5 Signal and Noise Case

When measurement data only contains random white Gaussian noise, we adjusted the refined threshold ϵ and generated a new threshold (T_a), which is proper and reasonable for the Hankel structured data. The new threshold T_a is set to be larger than the largest eigenvalues of the covariance matrix with the probability of 0.99. In this section, in order to investigate the performance of T_a for number of signals estimation and subspace separation, the threshold T_a is applied to the Hankel structured data that contains both signals and noise.

The simulation data is generated according to the data model expressed in chapter 2 and smoothed by smoothing techniques. At first, we only consider about forward smoothing technique. In order to show the performance of the threshold method using this new threshold T_a under different choices of M and L pairs. The probability of detection P_D over 2000 Monte-Carlo experiments with varying M and L ($M = 5, 10, ..., 50$ and $L = 5, 10, ..., 50$). P_D is computed by Equation 4.4. Choosing number of signals d=1, 5, 10, 15 separately. It needs to be ensured that signals are separately enough when $d \geqslant 2$. The results are shown in Figure 5.13 $\sim$ 5.16.

From Figure 5.13 $\sim$ 5.16, it can be seen that for a certain L, with the increase of M the P_D increases. This is because for a certain L, a larger M means that there is a larger dataset can be used in experiments. For the same reason, it can be observed that P_D will with be higher with larger L and M when the number of signals is same.

Furthermore, when the number of signals increased, P_D decreases under the same condition of L and M. This indicates that with a large number of signals contained in data, the performance of T_a degrades. The reason for this can be: with the increase of the number of signals, the smallest eigenvalue from signal can be smaller than the largest eigenvalue from noise (i.e., showing in Figure 5.17 and Figure 5.18), at the same time, the property of the data model (Hankel structure from smoothing operation) makes this happen more easier.

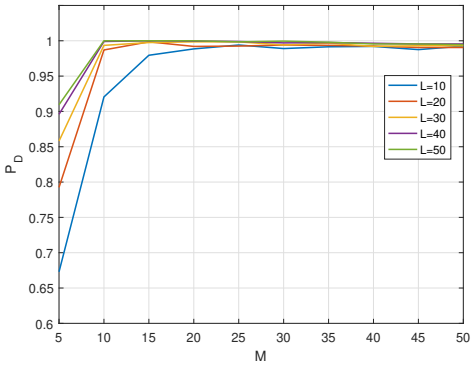


Figure 5.13: P_D plot for d=1

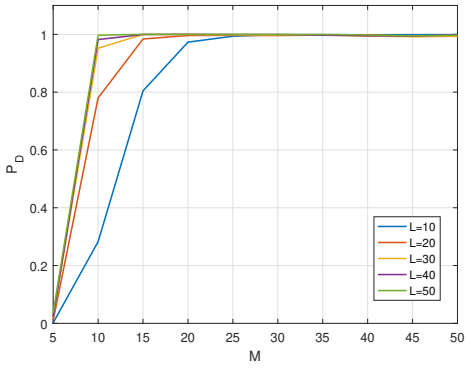


Figure 5.14: P_D plot for d=5

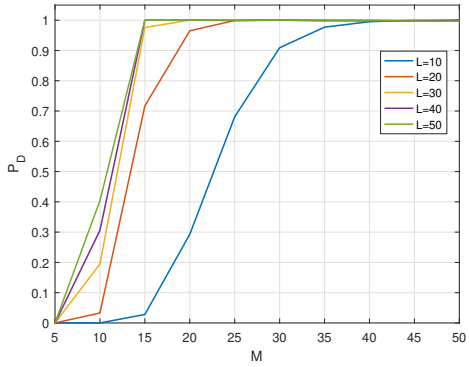


Figure 5.15: P_D plot for d=10

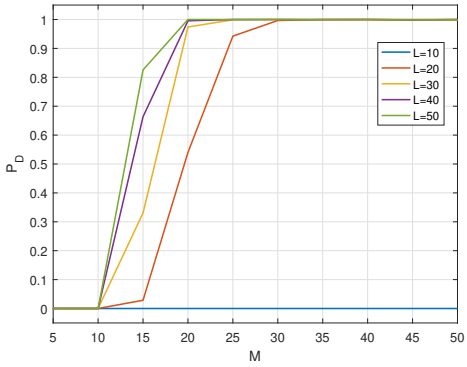


Figure 5.16: P_D plot for d=15

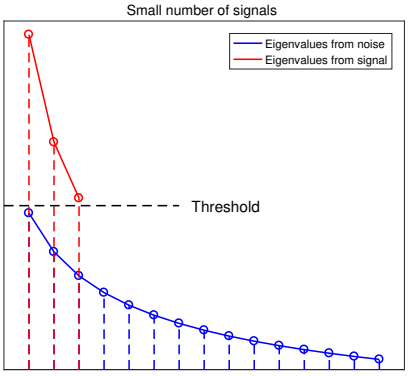


Figure 5.17: Small number of signals

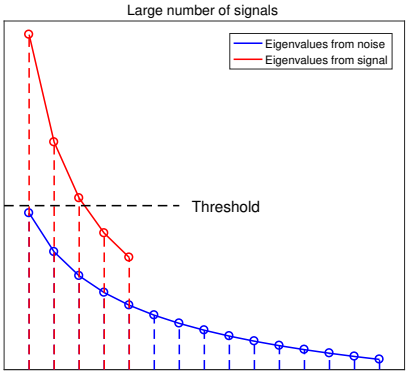


Figure 5.18: Large number of signals

In conclusion, the threshold T_a can effectively compute the number of signals under the condition of Hankel structure data. When the number of signals is small, a dataset with a small dimension can generate an accurate result. With the increase of the number of signals, a data matrix with a larger dimension is needed to achieve a reliable result.

5.2.2.6 Performance Under BLE Specification

For our case, data needs to be processed under BLE specification (narrowband radios) which means that the interested frequency range is $2.4 \sim 2.48$GHz, and the related parameters are proposed in Table 4.1. Choose $L = \lfloor (1/2)N \rfloor$ for the smoothing techniques. In order to show the performance of this threshold under different condition of SNR, compute the probability of detection (P_D) introduced by Equation 4.4 with varying SNR (dB). The signal power is set to be 1, the SNR varying from -20 to 20 (dB). Take the number of signals is d =1, 2, 3, 5, 10, 15 and 20 respectively, and it is ensured that signals are separately enough when $d \geqslant 2$. P_D is computed over 2000 Monte-Carlo experiments and the result for forward smoothing case is showing in Figure 5.19.

For the forward-backward smoothing, using the same fitting method to compute the scaling factor α and then compute the threshold T_a by the method proposed in subsubsection 5.2.2.2, the result P_D is shown in Figure 5.20.

From Figure 5.19 and Figure 5.20 it can be observed that with the increase of SNR, P_D increases from 0 to 1, which indicates that the threshold method has a better performance with a higher SNR. When the number of signals (multipath signals here) is relatively smalle ($ied \leqslant 5$), P_D can reach to 1 even at a low SNR. This illustrates that the threshold T_a can achieves an accurate result for a small number of signals even though SNR is low. The increasing of number of signals may take difficulties to the threshold method, the possible reasons are proposed in last section and Figure 5.17, 5.18.

Then make a comparison between the result of forward smoothing case and the result of forward-backward smoothing case. When the number of signals is relatively small (*i.e.*, $d \leqslant 5$), the performance of the results using two smoothing techniques are similar. If the number of signals further increases, the forward-backward smoothing outperforms forward smoothing since the parameter P_D reach to 1 at a lower SNR when using forward-backward smoothing technique.

It is also notable in the result figures that it is possible to correctly compute the number of signals d even though $d = L$. Under this situation, all the eigenvalues from the covariance matrix are larger than the threshold T_a. Even if the threshold method can generate an accurate result, there will not be data contained in noise subspace, and as a result, the subspace-based super-resolution algorithm fails. If the number of signals is larger than the dimension of covariance matrix, the only solution for this is to enlarge dimension of the measurement data.

It is also possible to set the frequency interval Δf=1MHz for the BLE specification. Then the length of measurement data is N=80. We process experiments under this condition, and the results are shown in Figure B.13 and Figure B.14. When comparing the result of Δf =2MHz (Figure 5.19 and 5.20) with the result of Δ =1MHz (Figure B.13 and B.14), it can be concluded that with the same SNR condition, P_D computed from Δf=1MHz is larger than P_D computed from Δf =2MHz for each d. This indicates that choosing the frequency interval Δf=1MHz is preferred according to the P_D computational result.

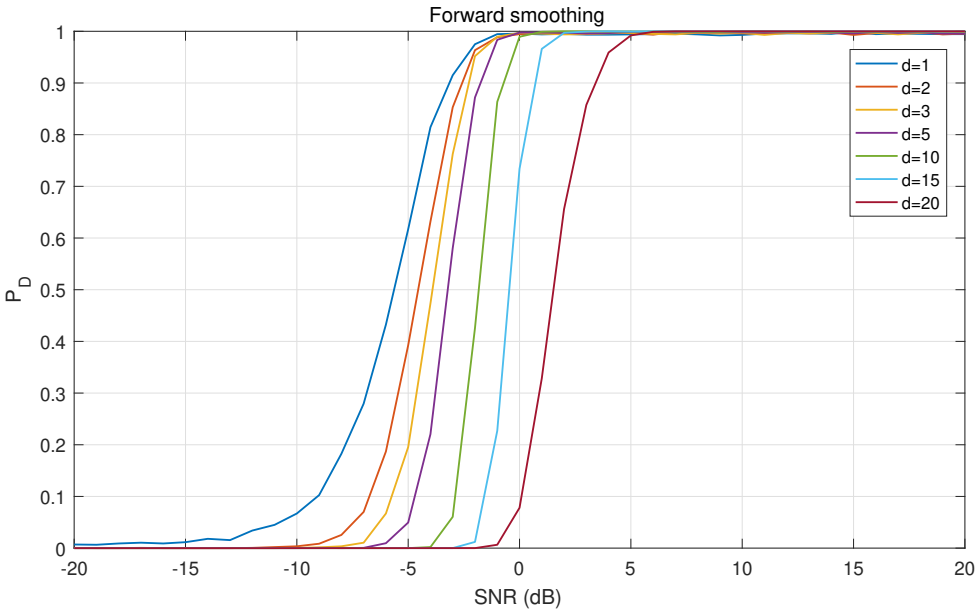


Figure 5.19: P_D-SNR plot for the threshold method (FS)

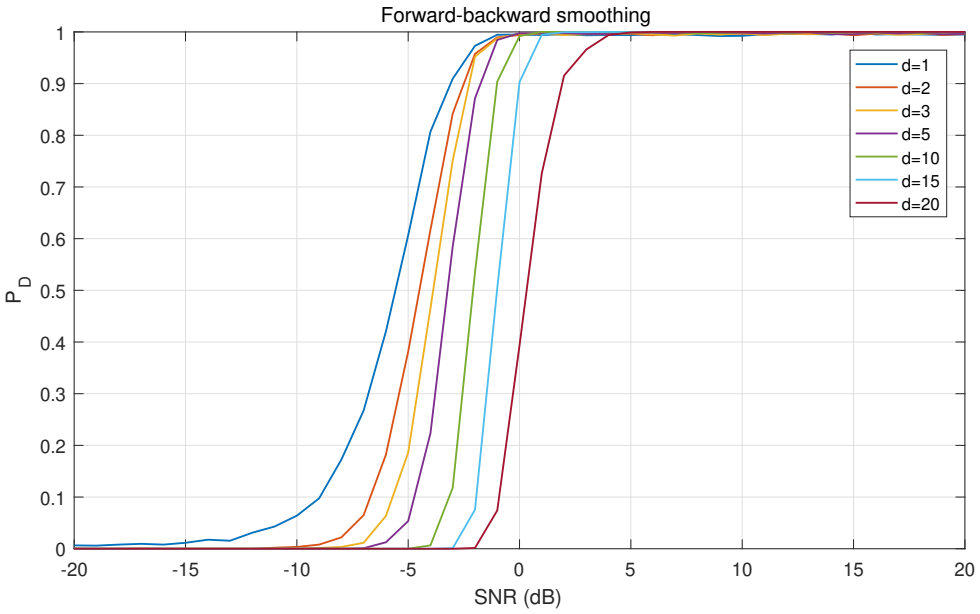


Figure 5.20: P_D-SNR plot for the threshold method (FBS)

5.3 Comparison

In this section, comparison is made between P_D results of MDL criteria (Figure 4.4) and P_D results of the threshold method (Figure 5.19 and 5.20) when the data is Hankel structure. For forward smoothing (FS) case, when d is relatively small (*i.e.*, $d \leqslant 5$), parameters of P_D for MDL criteria result and the threshold result start to increase at a similar SNR. However, with the increase of SNR, P_D for the threshold method gradually increases to 1, in contrast, P_D for MDL criteria can only reach to 0.95, at the same time, fluctuations happen in MDL result even though SNR is high. When d is relatively large (*i.e.*, $d > 5$), the threshold method can still produce an accurate result. P_D can reach to 1 with the increase of SNR. In comparison with MDL criteria, though P_D increases with the increase of SNR, it ends up at a low level (*i.e.*, 0.7 for $d = 10$ and 0.4 for $d = 15$). When forward-backward smoothing (FBS) is applied, P_D result from MDL criteria is larger than it from the threshold method when SNR is low (SNR$\leqslant$-5). This indicates that when SNR$\leqslant$-5, MDL criteria has a higher probability than the threshold method to correctly estimate d. However, with the increase of SNR, P_D result from the threshold method increases to 1, while P_D result from MDL criteria can only reach to a relatively low level.

Overall, the threshold method can produce a more stable result and can achieve a higher P_D for each d with under same SNR condition, then the threshold method is preferred.

5.4 Discussion

Based on experiments result in this chapter, a few areas of improvement on the threshold method can be imaged. As mentioned above, when the number of signals is relatively large, the situation proposed in Figure 5.18 may happen. Under this condition, the estimate result will be smaller than true value. This implies that the threshold should be dependent on the number of signals. To solve this problem, a sequential detection algorithm is required.

Another thing noticeable is that the threshold does not reach to the prescribed P_{FA} for all (M, L) pair. For example, (1)When $M < L$, the threshold is not tight. (2)For a constant L, with the increase of M, P_{FA} has a increasing trend which can exceed the prescribed P_{FA}. The technique of fitting the scaling factor α can be improved. In addition, the threshold can also be adjusted by mathematical analysis, which is needed to find the true distribution of the largest eigenvalue for Hankel structured data.

5.5 Conclusion

In this chapter, the threshold method is proposed and the functionality of it is evaluated. It shows that the threshold method proposed in section 5.1 is not proper for Hankel structure data. Then a fitting approach is taken into consideration in order to adjust the threshold to a reasonable level. From the results, the adjusted threshold can separate the signal subspace and noise subspace with a high accuracy for both FS case and FBS case. The performance of FBS case is better than FS case, since the parameter P_D will reach to 1 at a lower SNR condition when FBS is applied.

When comparing the subspace separation result of the threshold method and MDL criteria, the threshold method produce a more stable and accurate result. Then the threshold method is a preferred method.

In addition, when using the frequency interval Δf =1MHz, the performance of the threshold method will be improved. Then using narrowband radios with frequency interval Δf =1MHz is preferred if it is practicable.

Range Finding using Super-Resolution Algorithm

6

In this chapter, super-resolution algorithm (MUSIC algorithm) for range finding is verified. At first, experiments are processed on a simple data in order to test the practicability of applying super-resolution algorithm on narrowband radios (under BLE specification). Then a more complex simulation data is generated, trying to simulate the indoor multipath propagation signals to a high extent. Afterwards, super-resolution algorithm is applied to the data conducted from real indoor environment, for the purpose of exploring the potential of indoor localization using narrowband radios.

6.1 Super-Resolution Algorithm Result

In this section, the functionality of super-resolution algorithm is tested using generated simulation data. In this experiment, we generate a simple simulation data, which should be easy for super-resolution algorithm to realize range finding.

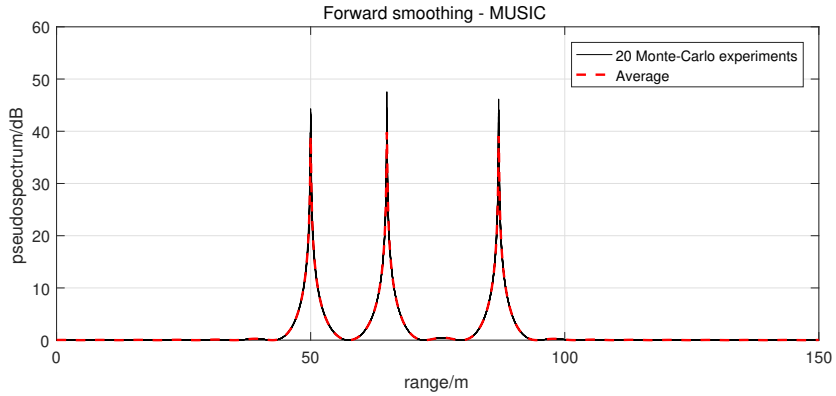
In this experiment, there is only 3 multipath signals with the distances between transmitter antenna and receiver antenna are set to be r=50, 65, 87m respectively, which ensure that the signals are separately enough. The time delay (time-of-arrival) parameters are τ=166.67, 216.67, 290.00ns respectively. The variance of complex white Gaussian noise is σ_n^2=0.01 and received signal power is 1, which indicates that SNR=20dB per signal. The data is generated according to the BLE specification from Table 4.1, in which frequency range is 2.4$\sim$2.48GHz and frequency interval is Δf =2MHz. Then subspace-based super-resolution algorithm is applied to simulation data.

Firstly, we assume that the number of (multipath) signals (d) is known, which means that the number of signals parameter is set to be d=3 when separating subspace. Afterwards, the experiments will be processed using super-resolution cooperated with the threshold method.

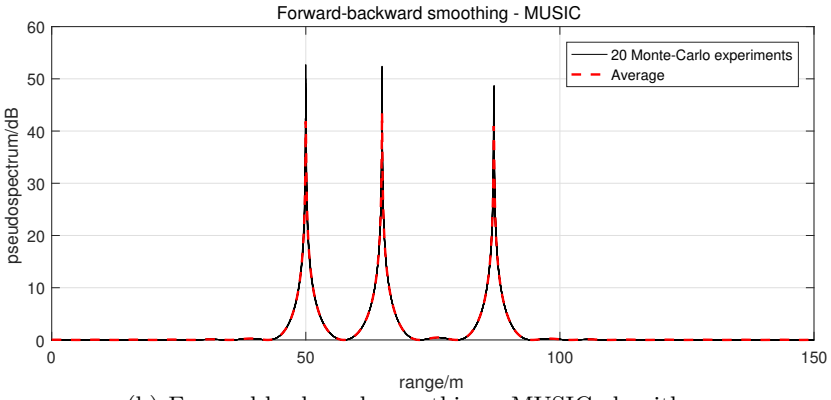
6.1.1 Known Number of Multipath Signals

At first, we assume that the number of multipath signals $d = 3$ is known. The methods applied are forward smoothing (FS) cooperated with MUSIC algorithm and forward-backward smoothing (FBS) cooperated with MUSIC algorithm. For each method, 20 Monte-Carlo experiments are conducted with randomly generated white Gaussian noise. The results are shown in Figure 6.1. The black solid lines are the 20 Monte-Carlo experiments, and the red dash line is the average value of those 20 experiments.

From Figure 6.1, with a known $d = 3$, there are three significant peaks appear at the range where signals arriving. The results produced by 20 Monte-Carlo experiments are almost same when the SNR is high (20dB). It shows that super-resolution algorithm can obtain an



(a) Forward smoothing - MUSIC algorithm



(b) Forward-backward smoothing - MUSIC algorithm

Figure 6.1: Range finding results with known d

accurate estimate of range r (or time delay τ). Since the frequency interval is Δf =2MHz, the maximum range that can be computed is 150m.

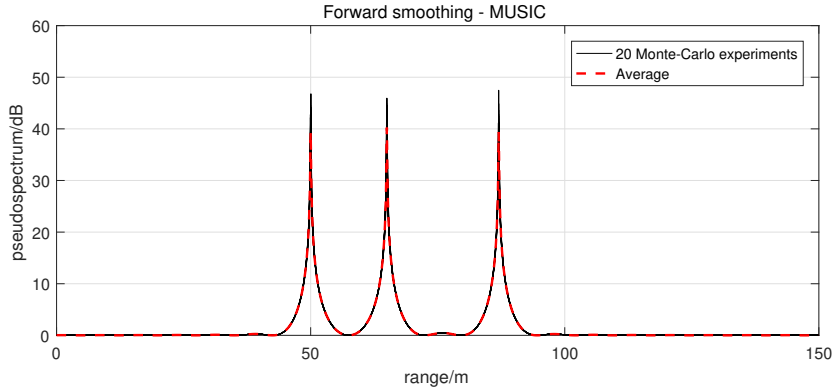
6.1.2 Estimate Number of Multipath Signals

In this part, we assume that the number of multipath signals is unknown and it needs to be estimated by the threshold method, and after which, super-resolution is applied to the simulation data for the purpose of range finding.

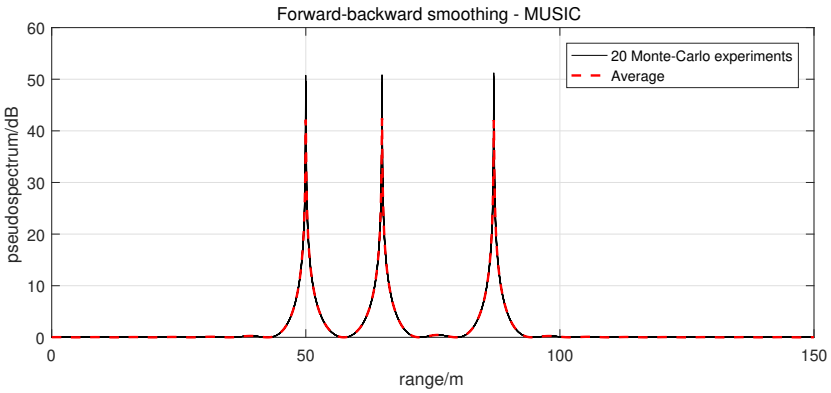
Using the parameter setting proposed at the start of section 6.1. There are 20 Monte-Carlo experiments are conducted with randomly generated white Gaussian noise. The results are shown in Figure 6.2

Figure 6.2 shows a similar result as Figure 6.1 since the threshold method can correctly estimate the number of multipath signals.

In conclusion, MUSIC algorithm can generate an accurate range dispersion of signals and MUSIC algorithm is practicable for range finding using narrowband radios.



(a) Forward smoothing - MUSIC algorithm



(b) Forward-backward smoothing - MUSIC algorithm

Figure 6.2: Range finding results with estimated d

6.2 Range Finding Simulation Result

In this section, a more complex multipath propagation condition is taken into consideration. In this experiment, the transmitted signal power is 1W and the received power is computed considering the free space path loss. We simply assume that the magnitude of multipath signals arrived after DLoS signal follow Rayleigh distribution. The signal arriving from DLoS arrived at r =7m (τ=23.33ns). The generated impulse response is shown in Figure 6.3 and 6.4.

In this experiment, received signal power is estimated as:

$$P_s = \frac{1}{N}\mathbf{x}^H\mathbf{x} \tag{6.1}$$

when $\mathbf{x}$ only contains signals which is vector of length N.

At here, we define SNR is:

$$\mathrm{SNR}_{\mathrm{dB}} = 10\log\left(\frac{P_s}{P_n}\right) \tag{6.2}$$

with P_n is the noise power and $P_n = \sigma_n^2$ for white Gaussian noise.

From Figure 6.3, there are a number of 45 multipath signal (including signal from DLoS and signals are after it). The threshold method is applied to estimate the number of multipath

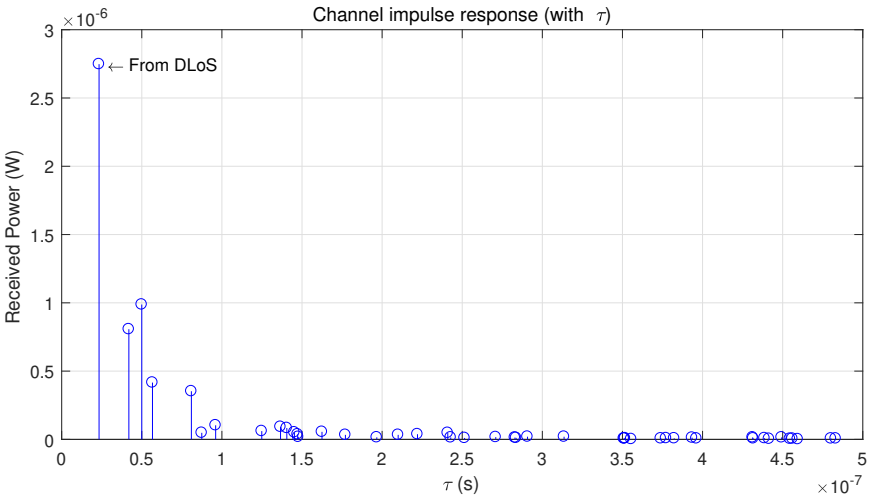


Figure 6.3: Channel impulse response with time delay (τ)

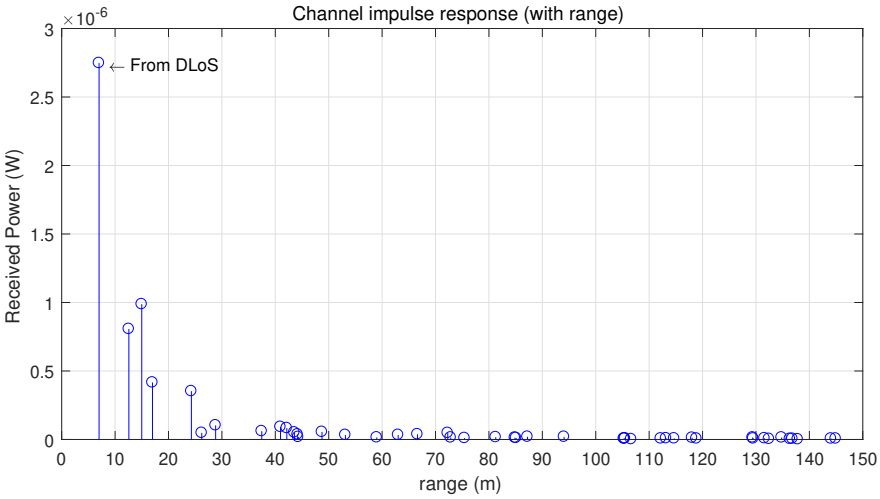
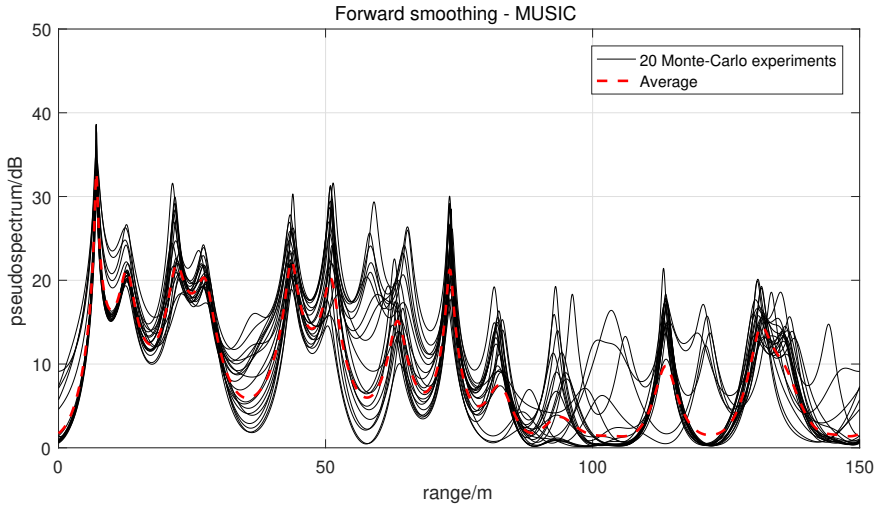


Figure 6.4: Channel impulse response with range (r)

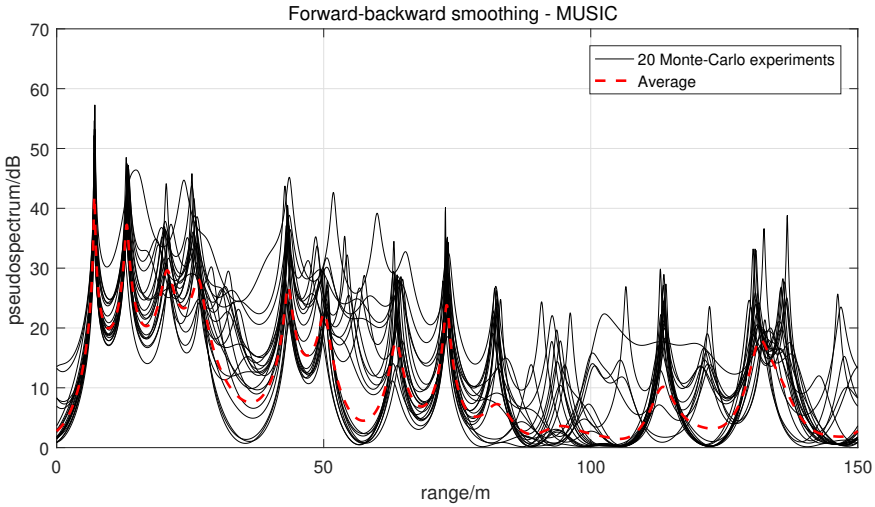
signals. Since the number of signals is larger and the signals arriving at large range (or long time delay) are received with low power, the threshold method can only find the signals with significant power. Normally, we assume that the signal arriving at DLoS is one of the most significant signals.

In this experiments, the SNR is set to be 20dB. We take 20 Monte-Carlo simulations and for each simulation the added white Gaussian noise is randomly generated. The results are shown in Figure 6.5.

From Figure 6.5 it is shown that there are multiple peaks appear in the result plots for both forward smoothing (FS) case and forward-backward smoothing (FBS) case, which indicates that MUSIC algorithm can detect multiple signals contained in measurement data. The peaks appear at the received signals that have relatively significant power and appear at the area where there are several signals that are close to each other arrived.



(a) Forward smoothing - MUSIC algorithm



(b) Forward-backward smoothing - MUSIC algorithm

Figure 6.5: Range finding results for simulation channel impulse response

In addition, when we compute the distance between transmitter antenna and receiver antenna, the signal arriving at DLoS is the only one interested. From Figure 6.5, MUSIC algorithm can estimate the range of DLoS signal with high accuracy.

6.3 Real Data Experiments

The simulation data cannot fully describe the multipath propagation of signal in indoor environments. To understand the potential of BLE system for indoor localization, the radio channel measurements have been conducted. In this chapter, experiments are proposed using real channel measurement conducted by Yao. The details about how the measurement is conducted is fully illustrated by Yao in [44], at here we only briefly introduce it.

6.3.1 Measurement Description

The measurement data is conducted in a rectangular room with two sides are 7 and 11.76m respectively. The transmitter antenna is an antenna pair with two monopole antennas that are set as orthogonal with each other (one is horizontal to ground and another one is vertical to ground). The transmitter antenna is used at mobile site, and there are 110 positions in the room for transmitter antenna. The receiver antenna array is also used at mobile site and it is placed at four corners of the room. When the receiver antenna array is placed at one corner receiving the signals from transmitter antenna located at the 110 different positions, it is called one round measurement, and then there are four rounds measurement in total. The receiver antenna is an antenna array contains four antennas and this antenna array employs a switch-then-sample mechanism when receiving measurement data.

When transmitting signal, channel sounding technique is applied to transmit signal at multiple carrier frequencies. Afterwards, the radio channel measurements in the frequency domain is conducted by a Vector Network Analyzer (VNA). When channel sounding technique is applied, the frequency range used is $2.2 \sim 2.6095$GHz, and this frequency range is sampled by 4096 points with frequency interval is 100kHz. Since the purpose of this experiment is to explore the potential of the BLE system in indoor localization field, the data needs to be processed under BLE specification.

6.3.2 Range Finding Result

The data is measured under extremely strict condition, which means that the channel measurement data can only describe the multipath propagation of signals, but the influence from noise is absent. However, in most of situations, the noise cannot be ignored. As a result, in this experiment, we add white Gaussian noise to the received channel measurement data.

In this experiment, we use the definition of received signal power and SNR proposed in Equation 6.1 and Equation 6.2. Then super-resolution algorithm is applied to the noise added channel measurement data, in order to estimate the dispersion of the time delay. The time delay of the signal arriving at direct line-of-sight (DLoS) is the only one interested to compute the distance between transmitter antenna and receiver antenna. The signal arrived with the shortest time delay is regarded as the signal arriving at DLoS and then we select the signal corresponding to the first peak in pseudo-spectrum plot of super-resolution result.

As mentioned before, when receiver antenna array is at one position receiving data from transmitter antenna located at 110 different positions, it is called one round measurement and we have four rounds measurement in total. As a result, there are a number of 440 sets of data vector received by each receiver antenna (in antenna array), which means that there are a number of 440 range parameters needs to be estimated for each antenna. The antenna array is composed by four antenna, then in total there are a number of 1760 range parameters need to be estimated.

In order to investigate the performance of super-resolution algorithm for range finding, the standard error is computed for those 1760 range estimated result. The standard error (e) is computed as the difference between the estimate range $\hat{r}$ and the true range r, which is

$$e = \hat{r} - r \tag{6.3}$$

Then the cumulative density function (CDF) is calculated based on the percentage of the standard error that is smaller than a certain value.

In this experiment, SNR is set to be 20dB. Data is processed under BLE specification, which is that frequency range used is 2.4∼2.48GHz and frequency interval is Δf=2MHz, then the length of each data vector is N =40. Using the threshold method to estimated the number of multipath signals contained in measurement data. The result is shown in Figure 6.6.

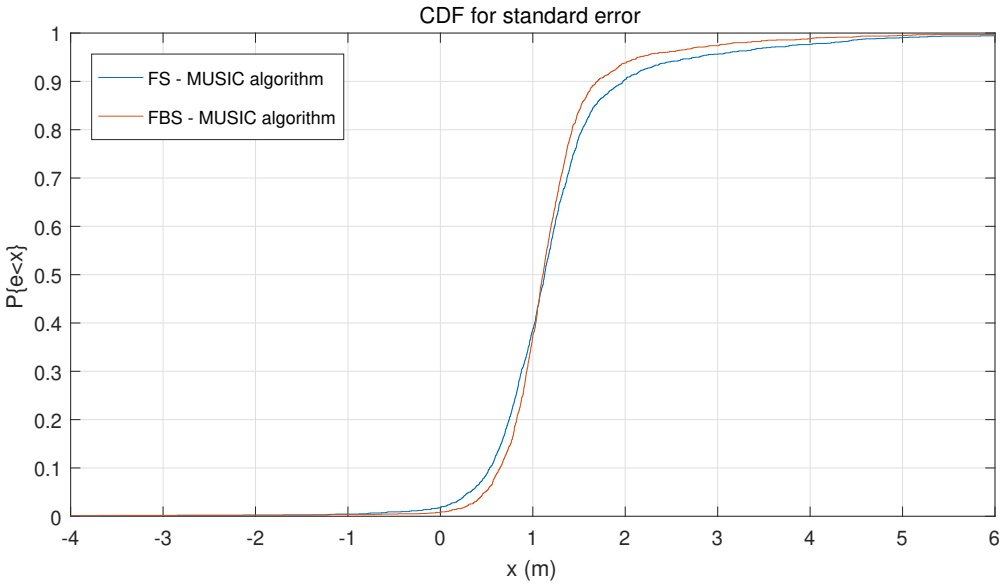


Figure 6.6: CDF of standard error

From Figure 6.6 it can be shown that when the standard error $e \leqslant$1m, the forward smoothing (FS) case performs better, to a small extent, than the forward-backward smoothing (FBS) case since the blue curve is above the red curve. Nevertheless, when $e \leqslant$1m, the percentages are still at a relatively low level (smaller than 40%) for both FS case and FBS case. When $1 < e \leqslant$ 2m, the two curves increase fast, and the number for FBS case increases faster than FS case. Besides, the percentages at x = 2m have reached to a relatively high level (94% for FBS case in comparison with 90% for FS case), which indicates that, for most of experiments, MUSIC algorithm can achieve range finding with standard error $e <$ 2m. In addition, when $x >$ 1m, the percentage curve for FBS is above the curve for FS all the time, then it can be concluded FBS technique outperforms FS technique.

In conclusion, when SNR=20dB, super-resolution algorithm can realize range finding with standard error $e <$ 1m for 40% experiments, and $e <$ 2m for 94% experiments. FBS technique cooperating with MUSIC algorithm is prior.

Besides, in BLE specification, if frequency interval is Δf =1MHz is practicable, the result of CDF of standard error is shown in Figure C.1.

6.4 Conclusion

In this chapter, the functionality of MUSIC algorithm is tested and evaluated.

In the first section, MUSIC algorithm is applied to a simple case. It has shown that MUSIC algorithm is practicable for estimating time delay dispersion using narrowband ratios.

Afterwards, we apply MUSIC algorithm to a more complex case. Although MUSIC algorithm cannot detect all the multipath signals, it is still efficient for finding the signal arriving at DLoS under a high SNR condition.

At last, MUSIC algorithm is applied to measurement data conducted from real indoor environment, in order to realize range finding. It has shown that when SNR=20dB, MUSIC algorithm can achieve that standard error is smaller than 1m for 40% of experiments, and standard error is smaller than 2m for 94% of experiments, using forward-backward smoothing technique.

Conclusions and Future Works

7

7.1 Conclusions

This thesis is to explore the potential of indoor localization using narrowband radios. The narrowband radios are under BLE specification. The purpose is to realize indoor localization in range domain, therefore the research objective of this thesis is to develop an algorithm to realize range finding by estimating time-of-arrival (ToA) of signal arriving at direct line-of-sight (DLoS). The algorithm considered is a subspace-based super-resolution algorithm, which is MUSIC algorithm. At the same time, it is an essential procedure to realize subspace separation, equally to compute the number of (multipath) signals, before subspace-based super-resolution algorithm is applied. Then another essential research objective is that to develop an algorithm to accurately separate the subspace. Different techniques have been compared.

To start with, according to the conducted measurement data, which is channel impulse response, a data model was developed. In this data model, multipath signals were indicated using different time delays. In order to obtain an accurate covariance matrix estimate for followed procedures, based on the shift invariance property of the data vector, smoothing techniques, including forward smoothing technique and forward-backward smoothing technique, were taken into consideration. Then measurement data was converted from a data vector into a Hankel structured data matrix by smoothing operations. The purposes of applying smoothing techniques are that, firstly, to improve the Teoplitz property of covariance matrix for a more accurate subspace result; secondly, to solve coherent signals problem in indoor environments.

Then the subspace separation methods, namely MDL criteria algorithm and the threshold method, were proposed and evaluated. For the MDL criteria algorithm, to evaluate the functionality of it, the probability of that the MDL criteria correctly computes the number of multipath signals was calculated for different SNR choices by Monte-Carlo experiments. For our Hankel structured data, we concluded that MDL criteria can only produce a reliable result for a small number of multipath signals under high SNR condition when cooperated with forward smoothing technique. We also concluded that the result of MDL criteria is not stable. When we evaluated the performance of the threshold method, we found that the proposed threshold method is not reasonable for our Hankel structured data. Based on experimental results, the threshold was adjusted to a proper and reasonable level, which is a tight upper bound of the largest eigenvalue from noise with a CFAR (1%). The performance is evaluated using the parameter of probability of that the threshold correctly computes the number of multipath signals. We concluded that the result of the threshold method is accurate when applying it on Hankel structured data. The threshold method outperform the MDL criteria.

Finally, super-resolution algorithm (MUSIC algorithm) was applied in order to realize range finding. To evaluate its functionality and effectiveness, simulation data was generated under BLE specification. It has shown that MUSIC algorithm is practicable under BLE

specification. Afterwards, MUSIC algorithm was applied to data that conducted from real indoor environments. Then experiments that were similar to Monte-Carlo simulations were processed since the conducted data contains numerous cases for different range parameters. Using MUSIC algorithm cooperating with forward-backward smoothing technique under BLE specification (with Δf =1MHz), the result standard error is smaller than 1m for 40% of experiments, 2m for 94% of experiments.

7.2 Future works

In future, studies can be done on further improving super-resolution algorithm. Investigation can be done on optimization algorithms, in order to improve resolution of time delay dispersion and the accuracy of range finding result.

For subspace separation section, the threshold method can achieve a reliable result based on white Gaussian noise. Studies can be done on cases that different distributions of noise are introduced. Besides, for the threshold method, the threshold for Hankel structured data is computed based on experimental results. More research can be done to figure out the mathematical result of threshold for Hankel structured data.

Once the localization part using narrowband radios is realized, the studies can be extended to objective tracking. In this thesis, subspace for MUSIC algorithm is computed by eigenvalue decomposition, which is a high computational complexity method, and it is not proper for objective tracking. A practicable method that has low computational complexity and high robustness needs to be developed. For example, subspace tracking technique can be imaged.

The MDL Criteria Result

A

MDL function plot for a relatively large number of signals.

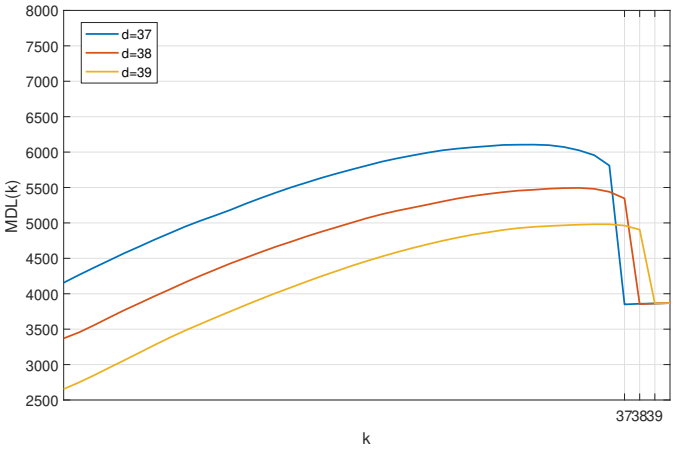


Figure A.1: MDL criteria function for d =37, 38 and 39

MDL function plot for pure white Gaussian noise measurement data.

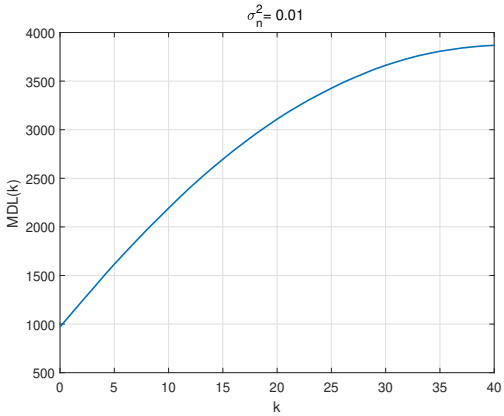


Figure A.2: MDL plot for noise only data (σ_n^2 =0.01)

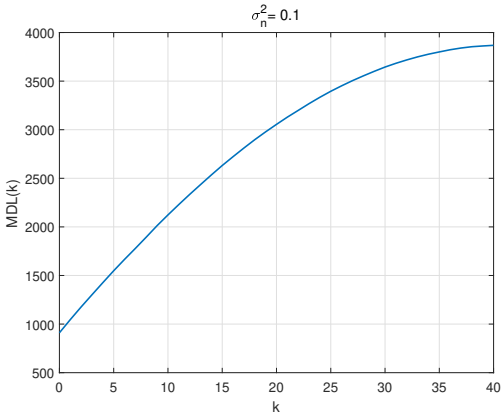


Figure A.3: MDL plot for noise only data (σ_n^2 =0.1)

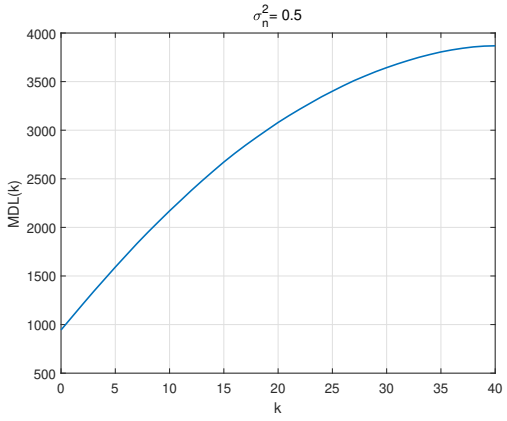


Figure A.4: MDL plot for noise only data (σ_n^2 =0.5)

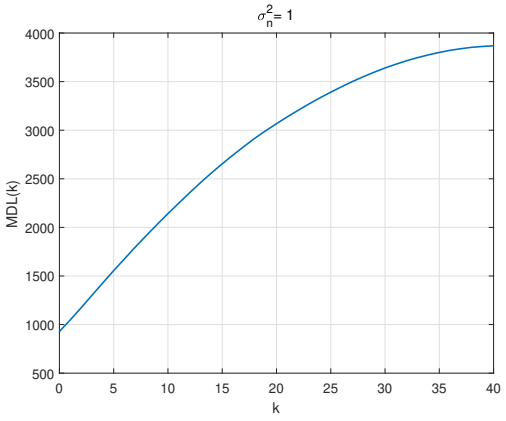


Figure A.5: MDL plot for noise only data (σ_n^2 =1)

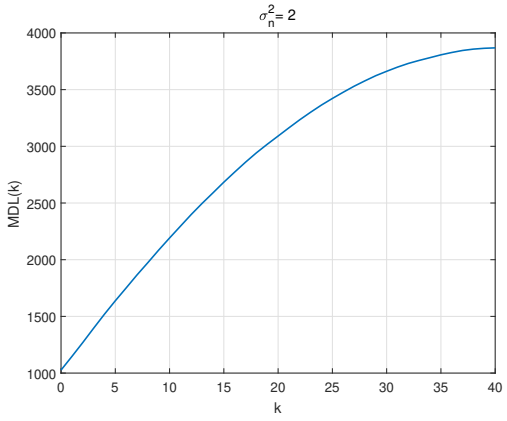


Figure A.6: MDL plot for noise only data (σ_n^2 =2)

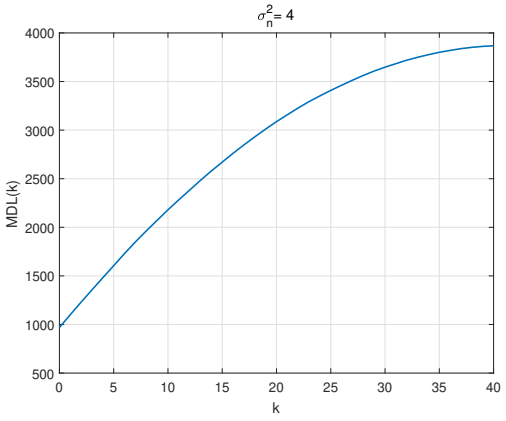
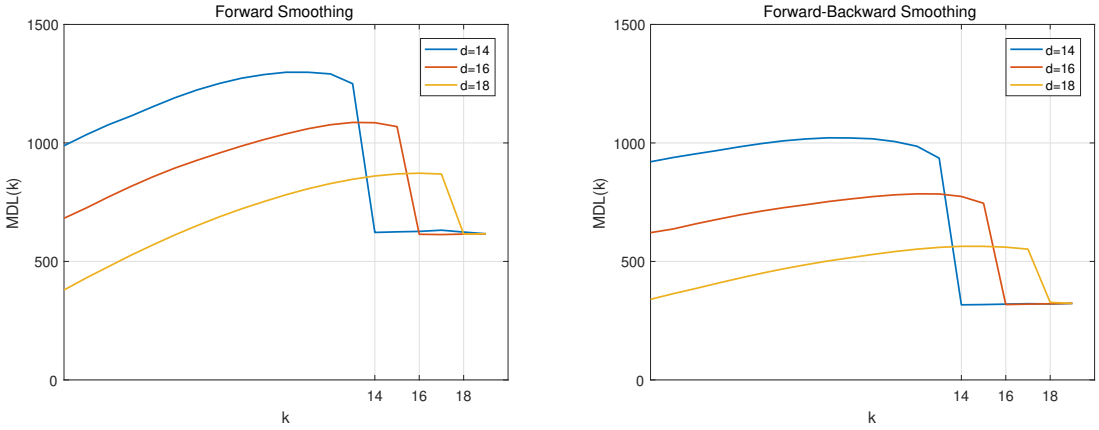


Figure A.7: MDL plot for noise only data (σ_n^2 =4)

MDL function plot for a relatively large number of signals for Hankel sturetured data.



(a) Forward smoothing

(b) Forward-backward smoothing

Figure A.8: The MDL criteria function for d =14, 16 and 18

The Threshold Method Result

B

Multiple snapshots contained in data, for pure random white Gaussian noise case. When choosing different noise variance, σ_n^2=0.01, 0.1, 0.5 and 2 respectively, L=20.

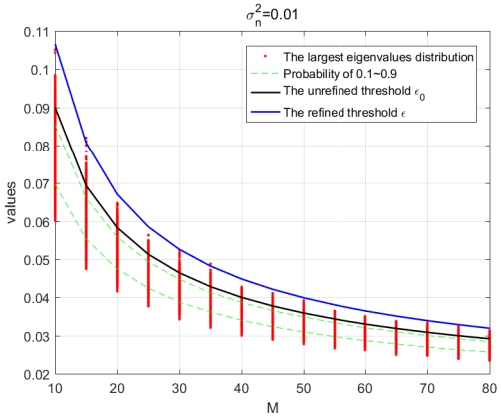


Figure B.1: Pure random white Gaussian noise (σ_n^2=0.01)

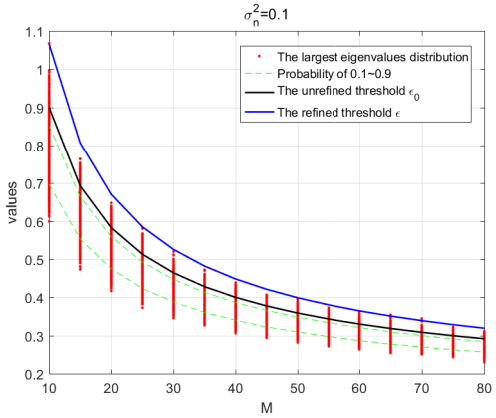


Figure B.2: Pure random white Gaussian noise (σ_n^2=0.1)

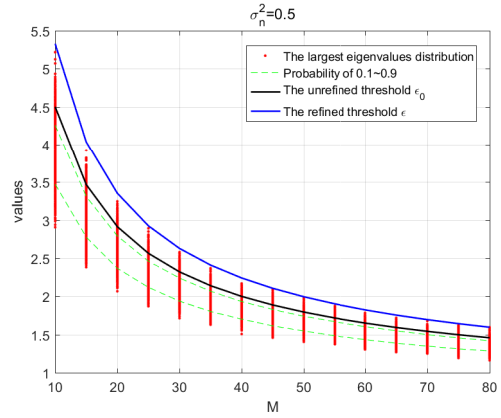


Figure B.3: Pure random white Gaussian noise (σ_n^2=0.5)

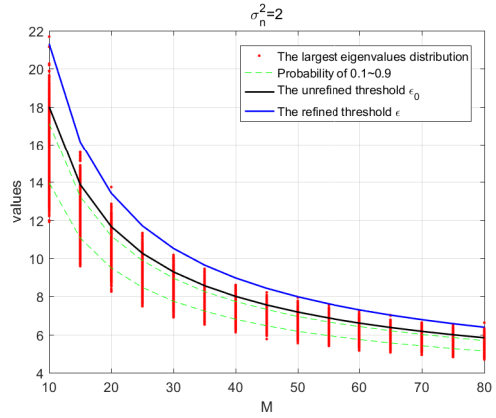


Figure B.4: Pure random white Gaussian noise (σ_n^2=2)

When choosing different L, L=5, 10, 20 and 30 respectively, σ_n^2=1.

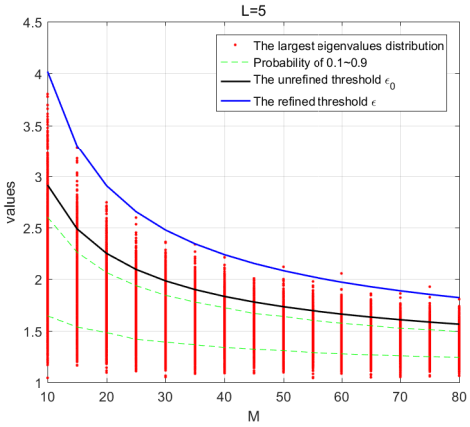


Figure B.5: Pure random white Gaussian noise (L=5)

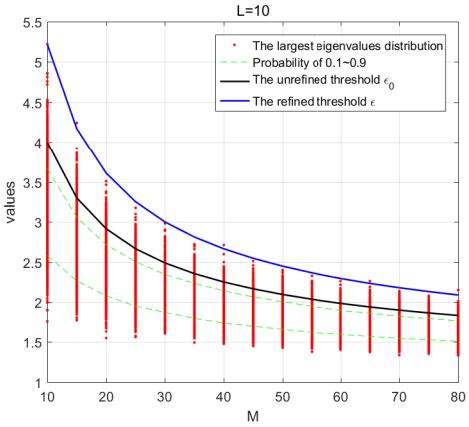


Figure B.6: Pure random white Gaussian noise (L=10)

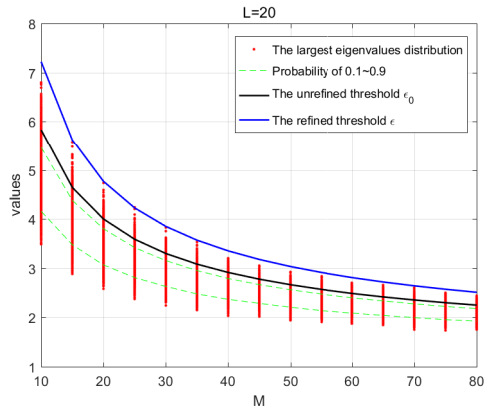


Figure B.7: Pure random white Gaussian noise (L=20)

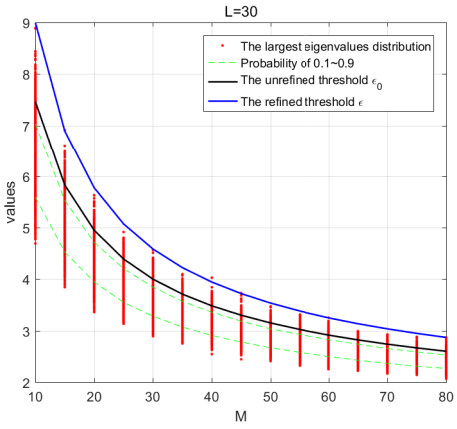


Figure B.8: Pure random white Gaussian noise (L=30)

Hankel structure data matrix, for the pure random white Gaussian noise case. Choosing different noise variance σ_n^2=0.01, 0.1, 0.5, and 2, L=20.

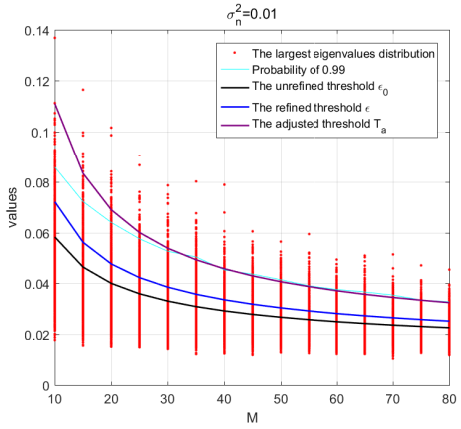


Figure B.9: Pure random white Gaussian noise (σ_n^2=0.01)

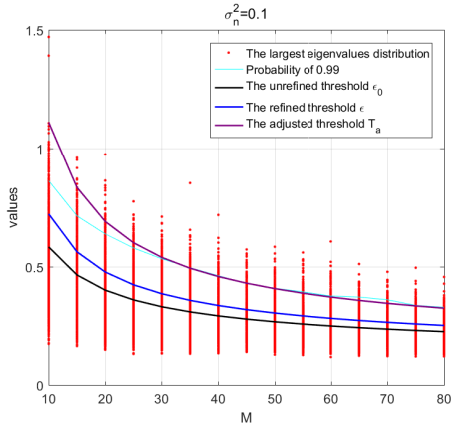


Figure B.10: Pure random white Gaussian noise (σ_n^2=0.1)

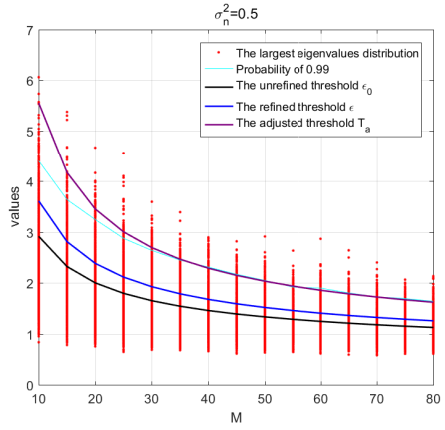


Figure B.11: Pure random white Gaussian noise (σ_n^2=0.5)

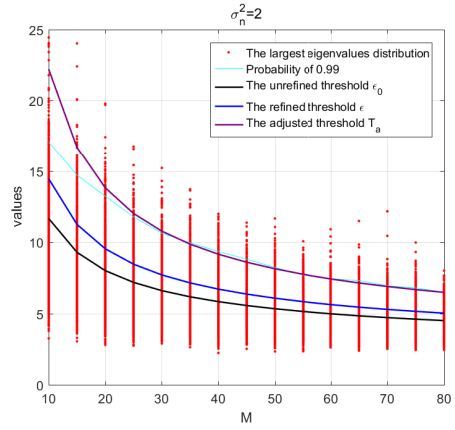


Figure B.12: Pure random white Gaussian noise (σ_n^2=2)

The P_D-SNR results when the frequency interval is Δf=1MHz in BLE specification.

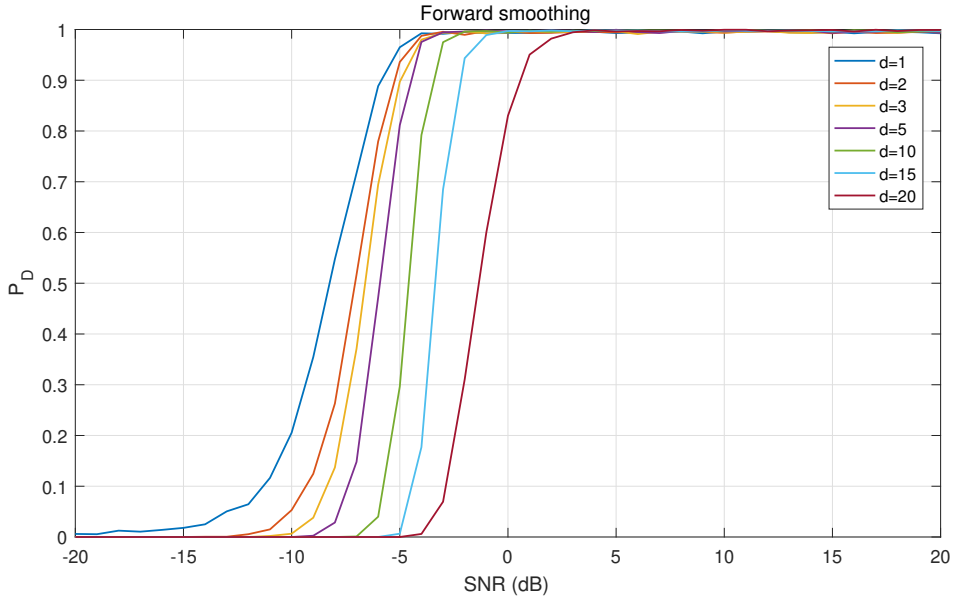


Figure B.13: P_D-SNR for different number of signals d (FS)

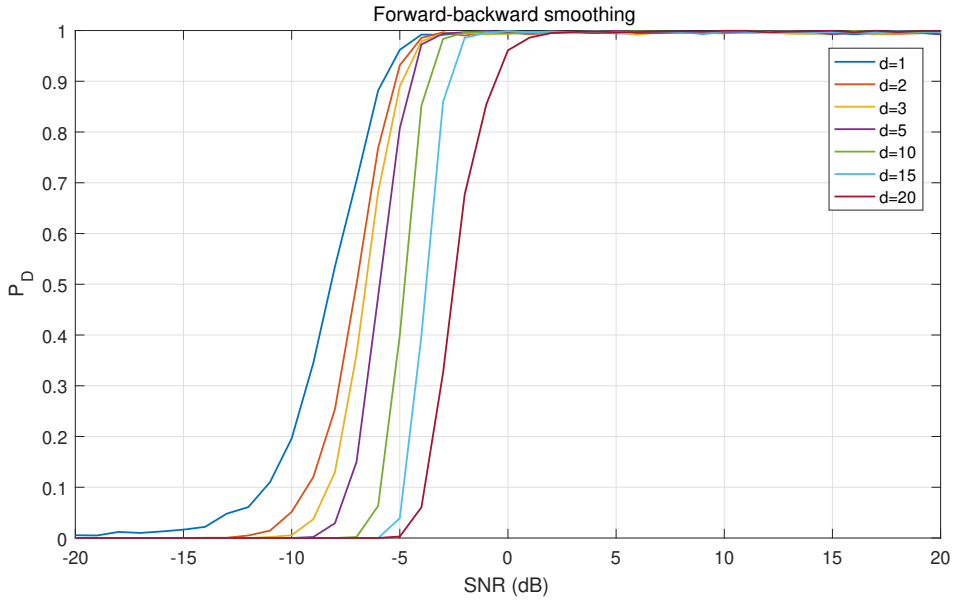


Figure B.14: P_D-SNR for different number of signals d (FBS)

Range Finding Result

C

CDF plot of standard error when frequency interval is Δf =1MHz.

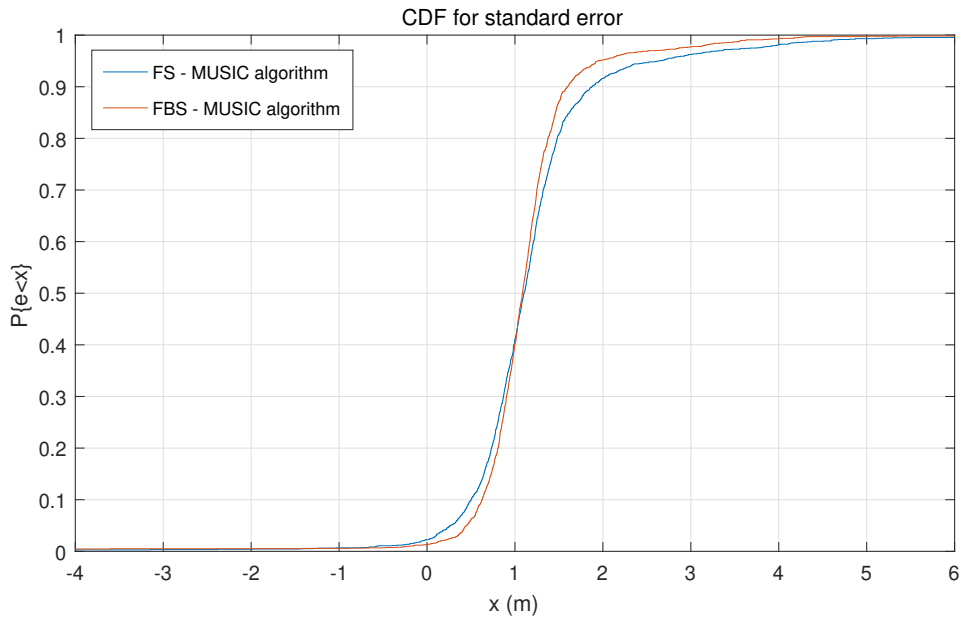


Figure C.1: CDF of standard error (Δf =1MHz)

Bibliography

[1] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, *Global positioning system: theory and practice*. Springer Science & Business Media, 2012.

[2] A. Yassin, Y. Nasser, M. Awad, A. Al-Dubai, R. Liu, C. Yuen, R. Raulefs, and E. Aboutanios, "Recent advances in indoor localization: A survey on theoretical approaches and applications," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1327–1346, 2016.

[3] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Communications Surveys & Tutorials, 11 (1), 2009*, 2009.

[4] Z. Farid, R. Nordin, and M. Ismail, "Recent advances in wireless indoor localization techniques and system," *Journal of Computer Networks and Communications*, vol. 2013, 2013.

[5] M. Woolley and S. Schmidt, "Bluetooth 5 go faster. go further," *Bluetooth SIG*, vol. 1, no. 1, pp. 1–25, 2017.

[6] T. S. Rappaport, S. Y. Seidel, and K. Takamizawa, "Statistical channel impulse response models for factory and open plan building radio communicate system design," *IEEE Transactions on Communications*, vol. 39, no. 5, pp. 794–807, 1991.

[7] A. A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on selected areas in communications*, vol. 5, no. 2, pp. 128–137, 1987.

[8] K. Pahlavan and A. H. Levesque, *Wireless information networks*. John Wiley & Sons, 2005, vol. 93.

[9] J. S. Thompson, P. M. Grant, and B. Mulgrew, "Performance of spatial smoothing algorithms for correlated sources," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 1040–1046, 1996.

[10] X. Li and K. Pahlavan, "Super-resolution toa estimation with diversity for indoor geolocation," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 224–234, 2004.

[11] J. E. Evans, D. Sun, and J. Johnson, "Application of advanced signal processing techniques to angle of arrival estimation in atc navigation and surveillance systems," Massachusetts Inst of Tech Lexington Lincoln Lab, Tech. Rep., 1982.

[12] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 806–811, 1985.

[13] T.-J. Shan and T. Kailath, "Adaptive beamforming for coherent signals and interference," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 3, pp. 527–536, 1985.

[14] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 1, pp. 8–15, 1989.

[15] W. Du and R. L. Kirlin, "Improved spatial smoothing techniques for doa estimation of coherent signals," *IEEE Transactions on signal processing*, vol. 39, no. 5, pp. 1208–1210, 1991.

[16] H. Yamada, M. Ohmiya, Y. Ogawa, and K. Itoh, "Superresolution techniques for time-domain measurements with a network analyzer," *IEEE Transactions on antennas and propagation*, vol. 39, no. 2, pp. 177–183, 1991.

[17] R. T. Williams, S. Prasad, A. K. Mahalanabis, and L. H. Sibul, "An improved spatial smoothing technique for bearing estimation in a multipath environment," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 425–432, 1988.

[18] V. Reddy, A. Paulraj, and T. Kailath, "Performance analysis of the optimum beamformer in the presence of correlated sources and its behavior under spatial smoothing," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 35, no. 7, pp. 927–936, 1987.

[19] K. Pahlavan, P. Krishnamurthy, and A. Beneat, "Wideband radio propagation modeling for indoor geolocation applications," *IEEE Communications Magazine*, vol. 36, no. 4, pp. 60–65, 1998.

[20] K. Pahlavan, X. Li, and J.-P. Makela, "Indoor geolocation science and technology," *IEEE Communications Magazine*, vol. 40, no. 2, pp. 112–118, 2002.

[21] T. Lo, J. Litva, and H. Leung, "A new approach for estimating indoor radio propagation characteristics," *IEEE Transactions on Antennas and Propagation*, vol. 42, no. 10, pp. 1369–1376, 1994.

[22] G. Morrison and M. Fattouche, "Super-resolution modeling of the indoor radio propagation channel," *IEEE transactions on vehicular technology*, vol. 47, no. 2, pp. 649–657, 1998.

[23] L. Dumont, M. Fattouche, and G. Morrison, "Super-resolution of multipath channels in a spread spectrum location system," *Electronics Letters*, vol. 30, no. 19, pp. 1583–1584, 1994.

[24] M.-A. Pallas and G. Jourdain, "Active high resolution time delay estimation for large bt signals," *IEEE Transactions on Signal Processing*, vol. 39, no. 4, pp. 781–788, 1991.

[25] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 37, no. 7, pp. 984–995, 1989.

[26] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE transactions on antennas and propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[27] H. Akaike, "A new look at the statistical model identification," in *Selected Papers of Hirotugu Akaike*. Springer, 1974, pp. 215–222.

[28] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.

[29] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 33, no. 2, pp. 387–392, 1985.

[30] G. Xu, R. H. Roy, and T. Kailath, "Detection of number of sources via exploitation of centro-symmetry property," *IEEE Transactions on Signal processing*, vol. 42, no. 1, pp. 102–112, 1994.

[31] Z.-D. Bai and Y.-Q. Yin, "Limit of the smallest eigenvalue of a large dimensional sample covariance matrix," in *Advances In Statistics*. World Scientific, 2008, pp. 108–127.

[32] M. Zhou and A.-J. van der Veen, "Blind separation of partially overlapping data packets," *Digital Signal Processing*, vol. 68, pp. 154–166, 2017.

[33] K. Johansson, "Shape fluctuations and random matrices," *Communications in mathematical physics*, vol. 209, no. 2, pp. 437–476, 2000.

[34] I. M. Johnstone *et al.*, "On the distribution of the largest eigenvalue in principal components analysis," *The Annals of statistics*, vol. 29, no. 2, pp. 295–327, 2001.

[35] C. A. Tracy and H. Widom, "Correlation functions, cluster functions, and spacing distributions for random matrices," *Journal of statistical physics*, vol. 92, no. 5-6, pp. 809–835, 1998.

[36] ——, "Level-spacing distributions and the airy kernel," *Communications in Mathematical Physics*, vol. 159, no. 1, pp. 151–174, 1994.

[37] ——, "On orthogonal and symplectic matrix ensembles," *Communications in Mathematical Physics*, vol. 177, no. 3, pp. 727–754, 1996.

[38] ——, "The distributions of random matrix theory and their applications," in *New trends in mathematical physics*. Springer, 2009, pp. 753–765.

[39] N. E. Karoui, "On the largest eigenvalue of wishart matrices with identity covariance when n, p and p/n tend to infinity," *arXiv preprint math/0309355*, 2003.

[40] Z. Ma *et al.*, "Accuracy of the tracy–widom limits for the extreme eigenvalues in white wishart matrices," *Bernoulli*, vol. 18, no. 1, pp. 322–359, 2012.

[41] I. M. Johnstone, "High dimensional statistical inference and random matrices," *arXiv preprint math/0611589*, 2006.

[42] M. Chiani, "Distribution of the largest eigenvalue for real wishart and gaussian random matrices and a simple approximation for the tracy–widom distribution," *Journal of Multivariate Analysis*, vol. 129, pp. 69–81, 2014.

[43] D. Paul, "Asymptotic distribution of the smallest eigenvalue of wishart (n, n) when n, n such that n/n 0," in *Nonparametric Statistical Methods and Related Topics: A Festschrift in Honor of Professor PK Bhattacharya on the Occasion of His 80th Birthday*. World Scientific, 2012, pp. 423–458.

[44] L. Yao, "Bluetooth direction finding," Available at http://resolver.tudelft.nl/uuid:c07eb3a2-a303-4690-ac3e-e96f0064afcd.