11-18-2019 9:30 AM

Exploitation of Robust AoA Estimation and Low Overhead Beamforming in mmWave MIMO System

Yuyan Zhao
The University of Western Ontario

Supervisor
Wang, Xianbin
The University of Western Ontario

Abstract

The limited spectral resource for wireless communications and dramatic proliferation of new applications and services directly necessitate the exploitation of millimeter wave (mmWave) communications. One critical enabling technology for mmWave communications is multi-input multi-output (MIMO), which enables other important physical layer techniques, specifically beamforming and antenna array based angle of arrival (AoA) estimation. Deployment of beamforming and AoA estimation has many challenges. Significant training and feedback overhead is required for beamforming, while conventional AoA estimation methods are not fast or robust. Thus, in this thesis, new algorithms are designed for low overhead beamforming, and robust AoA estimation with significantly reduced signal samples (snapshots).

The basic principle behind the proposed low overhead beamforming algorithm in time-division duplex (TDD) systems is to increase the beam serving period for the reduction of the feedback frequency. With the knowledge of location and speed of each candidate user equipment (UE), the codeword can be selected from the designed multi-pattern codebook, and the corresponding serving period can be estimated. The UEs with long serving period and low interference are selected and served simultaneously. This algorithm is proved to be effective in keeping the high data rate of conventional codebook-based beamforming, while the feedback required for codeword selection can be cut down.

A fast and robust AoA estimation algorithm is proposed as the basis of the low overhead beamforming for frequency-division duplex (FDD) systems. This algorithm utilizes uplink transmission signals to estimate the real-time AoA for angle-based beamforming in environments with different signal to noise ratios (SNR). Two-step neural network models are designed for AoA estimation. Within the angular group classified by the first model, the second model further estimates AoA with high accuracy. It is proved that these AoA estimation models work well with few signal snapshots, and are robust to applications in low SNR environments. The proposed AoA estimation algorithm based beamforming generates beams without using reference signals. Therefore, the low overhead beamforming can be achieved in FDD systems.

With the support of proposed algorithms, the mmWave resource can be leveraged to meet challenging requirements of new applications and services in wireless communication systems.

Keywords: MIMO, beamforming, codebook, overhead, AoA estimation, neural network

Lay Summary

The exploitation of communications based on new range of frequency is helpful to meet the requirement of new applications and services. Multi-input multi-output (MIMO) technique is developed in high-frequency communications, which enables other important physical layer techniques, including beamforming and antenna array based angle of arrival (AoA) estimation. However, challenges exist in the deployment of beamforming and AoA estimation. Significant training and feedback overhead is required for beamforming, while conventional AoA estimation methods are not fast or robust. Thus, in this thesis, new algorithms are designed for low overhead beamforming, and robust AoA estimation with reduced signal samples (snapshots).

The signal for transmission from each user equipment (UE) to the base station is not always available in time-division duplex (TDD) systems. Therefore, the low overhead beamforming is designed by increasing the beam serving period for the reduction of the feedback frequency. With the knowledge of location and speed of each candidate UE, the corresponding serving period can be estimated. UEs with long serving period and low interference are selected and served simultaneously. This algorithm is proved to be effective in keeping the high data rate of conventional beamforming, while the feedback required for beams generation can be reduced.

Achieving a fast and robust AoA estimation is the most important part for the low overhead beamforming design in frequency-division duplex (FDD) systems. The AoA of the signal with noise from a UE can be used as the direction of beamforming to this UE, while the AoA should be both fastly and correctly estimated for the cases with high frequent beam changing requirement. Two-step neural network models are designed for AoA estimation, which are proved to work well with few signal snapshots, and are robust to applications in low signal to noise ratio (SNR) environments. The proposed AoA estimation based beamforming can determine the directions of downlink beams with uplink transmission signals. Therefore, no extra signals for beams generation are required, and the low overhead beamforming can be achieved in FDD systems.

Acknowledgements

I would like to sincerely appreciate my supervisor, Dr. Xianbin Wang, for his guidance in the past two years. He broadened my views in the research and always encouraged me to explore the areas with novelty. With his help, I made progress not only in my research area, but also on the ways to do research work. The experience from him will also be treasure for my future life.

Then I would like to show my appreciation to Dr. Zhou, Dr. Dounavis and Dr. Rao in my examination committee. The comments and suggestions from them were meaningful for improving this thesis.

I am also grateful to my colleague, Dr. Yanan Liu, as well as other colleagues in our research group. They were selfless to share me with the research related knowledge, and also willing to help me in the past two years. Due to the excellent atmosphere for research in our group, I could focus on working and be brave enough to face any challenges.

Additionally, I would like to thank every course supervisor and administrative staff in the Western University. They were so helpful with regard to both my research work and campus life.

Last but not least, I should acknowledge my parents and friends for their support, love and encouragement throughout these years.

Contents

List of Figures

List of Tables

List of Abbreviations

4QAM   four quadrature amplitude modulation

AL   activation layer

AoA   angle of arrival

AoD   angle of departure

AWGN   additive white Gaussian noise

BS   base station

CDF   cumulative distribution functions

CI   close-in

CI-opt   close-in optimized

CNN   convolutional neural network

CSI   channel state information

DwPTS   downlink pilot time slot

ESPRIT   the estimation of signal parameters via rotational invariance technique

FCL   fully-connected layer

FDD   frequency-division duplex

GP   guard period

LAoA   LoS path AoA

LoS   line of sight

MIMO   multi-input multi-output

MLP   multi-layer perceptron

mmWave   millimeter wave

MU-MIMO   multi-user MIMO

MUSIC   multiple signal classification

NLoS	non-LoS
OFDM	orthogonal frequency division multiplexing
PLE	path loss exponent
RBFNN	radial basis function neural network
ReLU	rectified linear unit
RF	radio frequency
RMSE	root mean square error
RSS	received signal strength
Rx	receiver
SC-FDMA	single carrier frequency division multiple access
SDD	sequential downlink-downlink
SDU	sequential downlink-uplink
SINR	signal to interference and noise ratio
SISO	single input single output
SNR	signal to noise ratio
SR	sum data rate
SU-MIMO	single user MIMO
SWR	sum weighted data rate
TCC	two step classification and classification model
TCR	two step classification and regression model
TDD	time-division duplex
TDoA	time difference of arrival
ToA	time of arrival
TTI	transmission time interval
Tx	transmitter
UE	user equipment
ULA	uniform linear array
UMa	urban macrocell
UpPTS	uplink pilot time slot

Chapter 1

Introduction

1.1 Overview of Communications with Large Scale Antenna Array

Nowadays, the limited spectral resource for wireless communications can hardly satisfy the capacity requirements of emerging applications and services. The resource in high frequency band has already been explored for a long time [1], which pushes forward the development of millimeter wave (mmWave) communications. The frequency band in mmWave communications ranges from 30 to 300 GHz, the high carrier frequency leads to the small wavelength, while the wavelength is always considered to be proportional to the space between adjacent antennas in antenna array design. Therefore, on the antenna array for mmWave signals transmission, the intervals among adjacent elements are smaller, and large numbers of antennas can be packed on a small-size antenna array. It makes the large scale antenna array (antenna array with plenty of elements) transmission become possible at both transmitter and receiver in the wireless communications. The large scale antenna arrays play important roles in many applications such as radar, navigation, remote sensing, vehicular communications and biomedical imaging as shown in Fig. 1.1. They also promote the development of communications in multiple input multiple output (MIMO) systems.

In cellular communications, the large scale antenna arrays, which are always equipped at the BS, make it possible to separate UEs using the same frequency/time resource by spatial selectivity. Beamforming is a widely utilized technology to make full use of spatial domain resource, which is designed based on large scale antenna arrays. It can improve the sum data rate

1
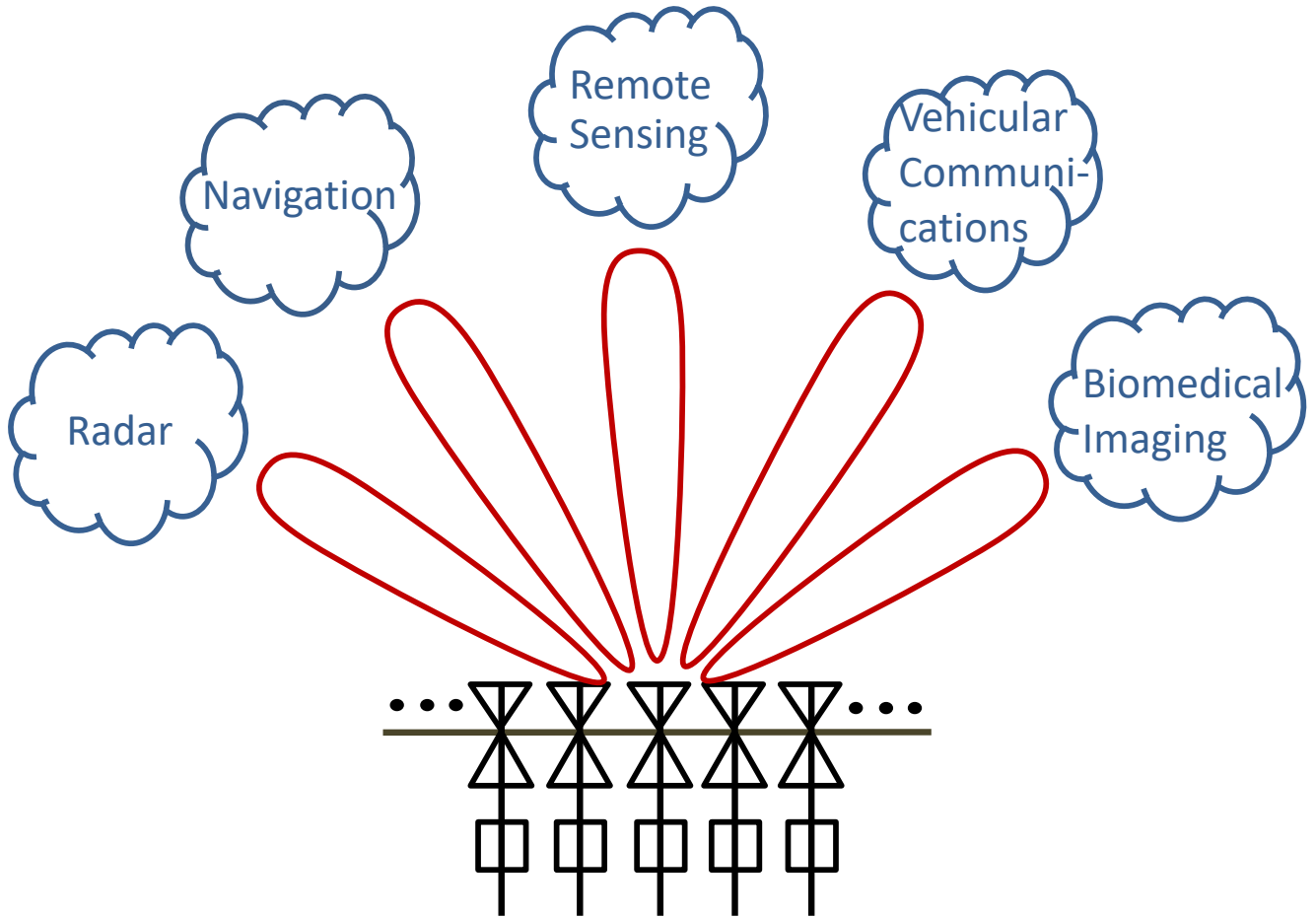
Figure 1.1: The applications using large scale antenna arrays.

and energy efficiency in multiuser systems by interference reduction in spatial domain [4][5]. However, the whole channel state information (CSI) is required in beamforming algorithms design, which is obtained by sending preambles or reference signals. The frequency/time re-source for these signals and their corresponding feedbacks are viewed as overhead for beam-forming, in the whole CSI-based beamforming, huge resources are required as overhead [6]. In order to reduce the overhead of beamforming, angle-based beamforming and codebook-based beamforming algorithms are explored. Both of them are viewed as part CSI-based beamform-ing, which generate beams with a part of channel features such as angle of departure (AoD) and multipath delay [7].

The angle-based beamforming utilizes the AoD of the dominant paths to determine the beams design. In downlink angle-based beamforming, the angle of arrival (AoA) of uplink

paths from UEs can be viewed as the AoD of the corresponding downlink paths due to the angle reciprocity in both time-division duplex (TDD) and frequency-division duplex (FDD) systems, when the time elapse between uplink and downlink signals can be neglected [8][9]. The extra uplink signals for AoA estimation are always in need, although AoA changes slower than channel matrix in mmWave channel, it still brings significant overhead. In conventional codebook-based beamforming algorithms, the estimation for AoA and CSI are not required any more. Instead, the training signals are necessary for suitable codeword selections, and both training and feedback signals should be considered as overhead, the amount of which is decided by the number of codewords for selection and the changing rate of channel. In all, for both angle-based and codebook-based beamforming, although the amount of overhead is reduced compared with the whole CSI-based beamforming, to make a further reduction is still a challenge.

The large scale antenna arrays are also helpful in applications with requirement for localization. In conventional localization methods, the fingerprinting is widely used in indoor scenarios, but this method is not adaptive to changing environments [10]. The environment-adapted localization can be achieved by location dependent parameters detection, which includes time difference of arrival (TDoA), time of arrival (ToA), AoA, received signal strength (RSS) or the combination of them. These parameters are detected by multiple reference nodes or anchor nodes, and the time-costing synchronization step is required for each time of localization [11]. In cellular system with single base station (BS), the requirement for multiple reference nodes is unrealistic. In this case, the large scale antenna arrays employed at the BS can be considered to achieve the single-anchor based localization, which means that the location information of user equipment (UE) can be obtained by uplink array signal processing at single receiver.

To specify the single-anchor based localization method with large scale antenna array, for a UE location detection based on the receiving multipath signal with line of sight (LoS) path, the ToA of the first arriving path can be utilized to calculate the straight-line distance from the UE [12]. At the same time, the AoA of LoS path can be extracted from the receiving array signal vector. The conventional antenna array based AoA estimation methods, such as the Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT) and Multiple SIgnal Classification (MUSIC) methods, cannot work well in environments with large noise [15]. Be-

sides, these methods can only provide high resolution AoA estimation with a large number of snapshots [13], for the cases such as estimating AoA for critical automotive applications, when only a limited number of snapshots are available in each time of estimation, the unignorable error may be made. To conclude, these conventional AoA estimation methods cannot achieve the requirement for fast and robust AoA estimation. Hence, based on few snapshots, achieving the high accuracy AoA estimation in low SNR environments is still a challenge.

1.2   Thesis Motivations

The mmWave large scale antenna array system is widely utilized in different scenarios, techniques such as beamforming and single-anchor based localization are designed with the help of antenna arrays. The large overhead for beamforming and the slow AoA estimation with low robustness in single-anchor based localization are pointed out as challenges in the last section. The work in thesis is motivated by the aforementioned challenges in order to achieve better performance in different applications with large scale antenna arrays.

Significant overhead for beamforming: In different beamforming algorithms, the information of channel is always required for beam generation. For both the whole CSI-based beamforming, and the part CSI-based beamforming including angle-based and codebook-based beamforming, the CSI or channel features should be estimated with less error, in order to make an appropriate beamforming design to UEs. The high quality CSI or channel features are obtained from multiple reference signals or training/feedback signals, which is viewed as overhead for beamforming and consumes plenty of transmitting resource. Therefore, the contradiction exists between the required overhead for high quality beamforming and the sacrifice of transmission signals, which brings a motivation for the research work on the high data rate and low overhead beamforming.

Non-robust AoA estimation: Due to the fact that the conventional subspace-based AoA estimation algorithms require numerous snapshots to formulate reliable signal and the noise sub-

spaces [14], these methods only work well with a large number of snapshots or in environments with high signal to noise ratio (SNR). However, the requirement for AoA estimation exists in environments with low SNR while only few snapshots are available, the aforementioned methods should be evaluated as non-robust methods for these cases (not robust to large noise and few snapshots use cases). As a result, it remains a problem to achieve accurate AoA estimation for applications in low SNR environments with only few snapshots.

1.3 Research Objectives

The objectives of this thesis are dealing with the challenges mentioned in the last section, which can be generally described as reducing the overhead in beamforming, and designing a robust AoA estimation method. To be mentioned that, in downlink beamforming, due to the difference between TDD and FDD systems, which is the availability of uplink signal in the downlink transmission. Therefore, the reduction of beamforming overhead can be discussed in TDD and FDD systems, respectively.

Low overhead beamforming in TDD systems: In TDD systems, the uplink transmission is not available all the time. Therefore, the beams should be ensured to adapt to UEs for the period between two times of uplink transmission. It can be found that the longer the period is, the less frequently the beams change, and the lower overhead required for beamforming when the overhead for each time of beam change is fixed. As a result, the objective of cutting down the beamforming overhead can be transformed into the long serving time connection beamforming design in TDD systems.

Low overhead beamforming in FDD systems: Different from TDD systems, in FDD systems, the uplink transmission is always available while transmitting the downlink signals. Therefore, for the beamforming design in FDD systems, the aim is to obtain channel features complying with reciprocity from uplink transmission signals directly without utilizing reference signals. Therefore, the downlink beams can be regenerated frequently while the low

overhead objective can be achieved.

Robust AoA estimation: In the environment with any SNR, the received array signal should be decided by the AoA of the dominant paths signal. The neural network is mentioned to have the function that classifies objects into groups according to their hidden features. Therefore, the neural network design becomes the objective, it should work to achieve the high accuracy AoA estimation with different SNR and small numbers of signal snapshots.

1.4   Technical Contributions of the Thesis

The main contributions of this thesis are summarized as follows:

- The multi-pattern codebook is designed in Chapter 3. In the cellular wireless communications system, it assumes that the candidate UEs with similar linear speed are in need of service in the serving space. Their angular speeds change with the straight-line distances to the BS, and the path loss is also distance related. In this chapter, the serving space is divided into zones according to the straight-line distances. UEs with similar distance to the BS are allocated into the same zone. A specific pattern codebook is designed to adapt to the UEs in a zone, where the beams are generated with the same beamwidth in spatial frequency domain and with fixed beam gain.

- Two UE selection algorithms are proposed in Chapter 4 based on the multi-pattern codebook design. With the AoA and straight-line distance information detected at the BS, the zone for all candidate UEs are specified and serving beams can be initialized firstly for both UE selection algorithms, then the UEs served with low interference are selected for a long serving time beamforming service in TDD systems. Two UE selection algorithms are designed as ultra low overhead oriented and large connection oriented, separately, which can be adaptive to different use cases.

- Two-step neural network models are designed to achieve the fast and robust AoA estimation in Chapter 5. In the propagation channel with single dominant path from the BS to

each UE, the proposed model can provide AoA estimation with high accuracy based on few snapshots in the low SNR environments. Besides, the proposed neural network models estimate AoA by signals with arbitrary transmitted symbols. Therefore, the uplink transmission signals can be utilized for AoA estimation, which can avoid the overhead required for sending reference signals in conventional beamforming algorithms, while the high sum data rate is kept in FDD systems.

1.5   Thesis Outline

The rest of the thesis is organized as below:

In Chapter 2, a literature survey starts from the introduction to the mmWave communications and multi-input multi-output (MIMO) communications based on the large scale antenna array system. Then the principles and procedures of conventional AoA estimation methods are detailed, in order to make performance comparisons with the novel method in Chapter 5. At the same time, some widely used neural network structures proposed are surveyed for the neural network structure selection. After that, the frame structures for TDD and FDD systems shown in 3GPP TS 36.211 are compared to prove the requirement of separate beamforming designs in two kinds of systems. Finally, different kinds of beamforming methods are introduced and compared.

In Chapter 3, the straight-line distance based zoned serving space and the propagation model in mmWave channel are firstly introduced in the multiuser MIMO system. Then the multi-pattern codebook design is proposed with the concept basis of spatial frequency domain. The beamwidth and beam gain are adaptive to candidate UEs in different zones. In the simulation work, the influence of three kinds intra-zone interference are compared.

Two UE selection algorithms are put forward in Chapter 4 on the basis of multi-pattern codebook design. In the system introduced in the last chapter, two UE selection algorithms are designed based on location information, one of them is ultra low overhead oriented algorithm, and the other one is large connection oriented algorithm. Simulation work makes comparisons among two UE selection algorithms based beamforming and a conventional beamforming

method with regard to their UEs connection ratio, sum data rate performance and the number of cumulative feedback bits required.

In Chapter 5, the application of two-step neural network based AoA estimation in FDD systems beamforming is presented. With the defination of uplink and downlink signals in simplified channel model, the structure of neural network model for uplink signal AoA estimation is firstly detailed. With the estimated AoA, the codeword selection can be employed for downlink beamforming. The simulation work shows the AoA estimation accuracy of the novel algorithm is improved compared with MUSIC algorithm with few snapshots in low SNR environments, and the higher sum weighted data rate (SWR), which is a parameter defined to reflect both the performance of sum data rate (SR) and the percentage of signal used as overhead for beamforming, can be achieved.

Finally, all the contributions presented in the previous chapters are concluded in Chapter 6, while the future plan for research is discussed in this Chapter as well.

Chapter 2

AoA Estimtion and Beamforming Design in mmWave MIMO System

The antenna array based AoA estimation and beamforming methods have been researched for a long time in mmWave MIMO systems. In this chapter, several kinds of mmWave MIMO systems are introduced firstly. Then the conventional antenna array based AoA estimation methods are detailed. Due to the difference of TDD and FDD systems, some partial CSI beamforming methods are reviewed as the basis of low overhead beamforming designs.

2.1 mmWave Band Communications

2.1.1 Characteristics of mmWave Band Communications

In the enabling technologies explored for the fifth generation (5G) mobile system, mmWave is developed as one dominant technology. It shows potentials to provide significantly rise on user throughput, spectral and energy efficiency. Also, the mmWave communication shows increase on the capacity of mobile networks with the joint capabilities of the huge available bandwidth in the mmWave frequency bands [16]. This section starts from the definition of mmWave communications. After that, the parameters for mmWave applications are compared, and a path loss model in mmWave transmission is detailed. In the end, the problem brought by the high carrier frequency is given, while the corresponding solving method is mentioned.

For the definition of mmWave communication, it spans a wide frequency range from 30 GHz to 300 GHz. Especially for outdoor applications in 5G cellular mobile systems, many

Table 2.1: Summary of Reported Outdoor mmWave Channel Measurement Campaigns

Carrier frequency	Service site	Radio frequency bandwidth	Max Tx-Rx distance	Target applications
28 GHz	Urban (street)	5 kHz	1.5 km	Point-to-point
	Orchard	5 kHz	0.9 km	Point-to-point
	Suburban	18 MHz	6 km	LMDS
	Dense urban	800 MHz	500 m	Cellular/Backhaul
38 GHz	Urban moderately dense	800 MHz	930 m	Cellular
73 GHz	Urban (campus)	800 MHz	200 m	E-band mobile /cellular backhaul
81-86 GHz	Urban (street canyon)	5 GHz	685 m	E-band radio

research interests can be found on the frequency bands in the 28-38 GHz and 70-90 GHz range [17]. This frequency band range can be further divided to adapt to different target applications. As shown in Table 2.1, part of the summary of reported measurement campaigns for characterization of outdoor millimeter wave propagation channels is given [17], which includes the information of service sites, radio frequency (RF) bandwidth, maximum Tx-Rx distance and target applications for transmission with different carrier frequencies.

Path loss is one of significant problems arisen by the high frequency carrier of mmWave communications. As mentioned in [18], the equation for close-in (CI) path loss model is

$$PL(f, d) = F(f, d_0) + 10n\log_{10}\left(\frac{d}{d_0}\right) + \chi_\sigma, (d \geqslant d_0), \qquad (2.1)$$

$$F(f, d_0) = 20\log_{10}\left(\frac{4\pi f d_0 \times 10^9}{c}\right). \tag{2.2}$$

In this model, the predefinitions are required for the reference distance $d_0(m)$ and carrier frequency $f(GHz)$. The model named as CI model or close-in optimized (CI-opt) model with different values for d_0. When $d_0 = 1m$, the model is simply CI model, while the CI-opt model can be called when the optimized d_0 is selected based on f. n and χ_σ are viewed as path loss exponent (PLE) and shadowing parameter, respectively. $d(m)$ is the Tx-Rx distance, σ is the standard deviation for zero-mean Gaussian random variable χ_σ, in mmWave transmission, σ for NLoS paths is much higher than that of LoS path, which means the signals transmitted from NLoS paths suffer from more severe path loss than the ones transmitted from LoS path. The specific value for n, χ_σ and the range for d can be determined based on the selection of d_0 and f. c represents the speed of light in Eq. (2.2). In 3GPP/ITU [19], the path loss model is rewritten in the form of

$$
\begin{aligned}
PL(f, d) &= F(f, d_0) + 10n\log_{10}\left(\frac{d}{d_0}\right) + \chi_\sigma, (d > d_0) \\
&= 20\log_{10}\left(\frac{4\pi d_0 \times 10^9}{c}\right) + 20\log_{10}(f) \\
&\quad + 10n\log_{10}\left(\frac{d}{d_0}\right) + \chi_\sigma \\
&= C + 20\log_{10}(f) + 10n\log_{10}\left(\frac{d}{d_0}\right) + \chi_\sigma \\
where \quad C &= 20\log_{10}\left(\frac{4\pi d_0 \times 10^9}{c}\right)
\end{aligned}
\tag{2.3}
$$

The drawback of mmWave transmission can be proved that when d and d_0 are fixed values, the channel path loss increases with the rising carrier frequency f. However, in large scale antennas array design, the distance between adjacent antennas is inversely proportional to f, which means that more antennas can be packed on an antenna array in high frequency band transmission, this is helpful to improve the transmitting power gain to overcome the problem of path loss. In the next section, the development of MIMO system will be explored with mmWave signal.

2.1.2   MIMO in mmWave Communications

As mentioned in the last section, the high frequency carrier in mmWave communications makes the utilization of antenna arrays with plenty of elements become common. In this section, different types of antenna arrays are shown firstly. After that, the definitions for different systems using large scale antenna arrays are introduced and compared.

The large scale antenna arrays can be classified into planar arrays and linear arrays according to their geometrical configurations [2]. As shown in Fig. 2.1, commonly the planar array structures are the rectangular or square array with equal space between adjacent elements in columns and rows, while the circular arrays are also viewed as planar arrays, they are equipped with the antenna elements arranged in concentric circles. As one of the advantages, the planar arrays can achieve beam steering in both the azimuth and elevation planes. Besides, the circular arrays is mentioned to be able to scan in the minimum change of beamwidth azimuthally. However, the utilization of planar arrays brings high cost and complexity because of the complicated array arrangement. Compared with planar arrays, linear arrays are much simpler. The antenna elements in this kind of arrays are uniformly arranged in a line. Therefore, with linear phase shifters, the main antenna lobes beam steering of linear arrays are obtained in single plane. The linear arrays conform to a Vandermonde structure, which makes the receiving signal precessing to be less challenging than that of some planar arrays. Therefore, the linear arrays are still used in many research works.

In MIMO systems, more than one antennas are equipped at both the transmitting side and receiving side. The conventional MIMO system mentioned in [20] is fully digital, which means all signal processing work are finished in baseband. Therefore, when the large scale antenna array is employed at any side in transmission, a large number of RF chains are required, which leads to the unaffordable energy consumption and cost in hardware. As a result, a fully analog architecture can be utilized to reduce the number of RF chains required. The phase shifter are employed to adjust signal to achieve an array gain, which process can be explored as analog beamforming. In Fig. 2.2 and Fig. 2.3, the signal process with fully digital architecture and fully analog architecture are shown respectively.
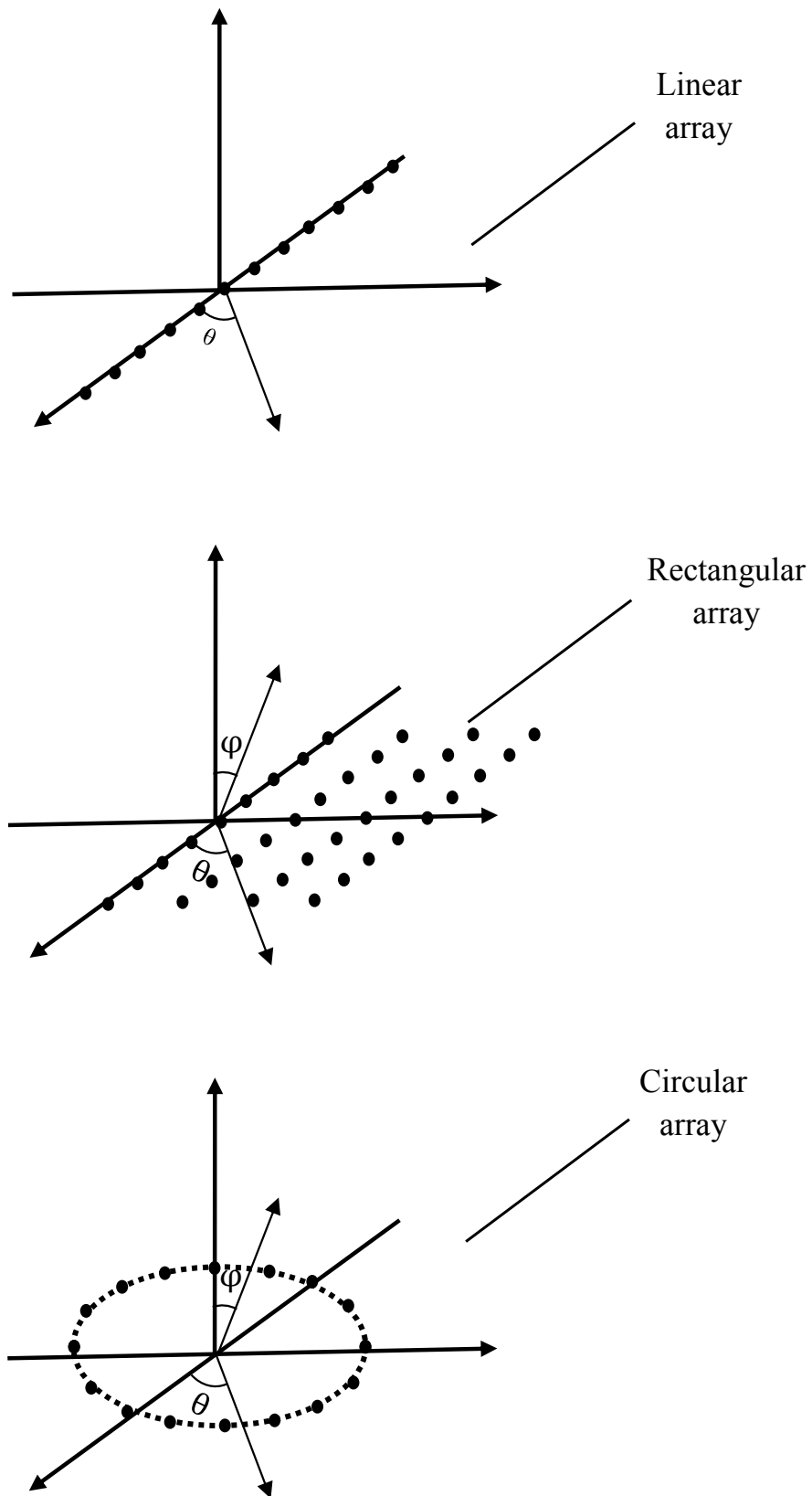
Figure 2.1: Different types of array antenna geometries.
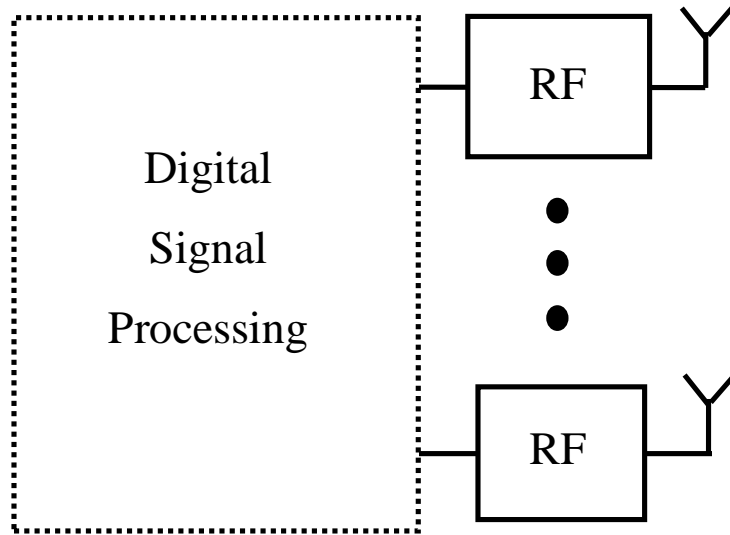
Figure 2.2: Fully digital MIMO architecture.
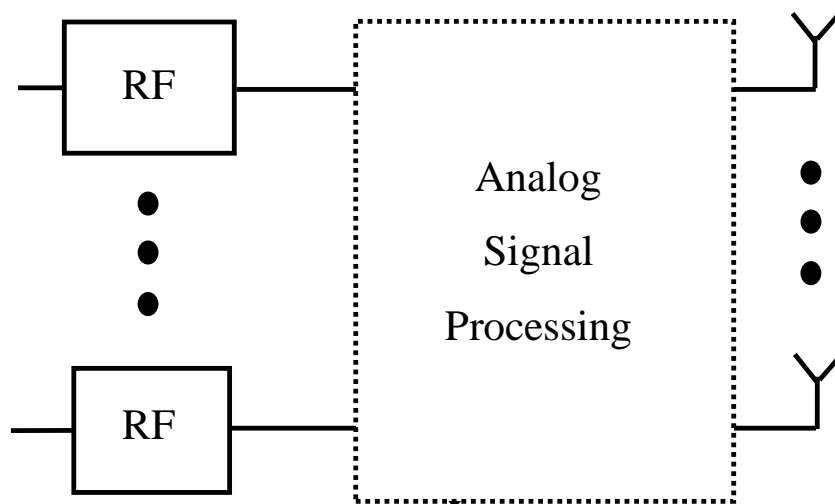
Figure 2.3: The architecture of fully analog MIMO system.
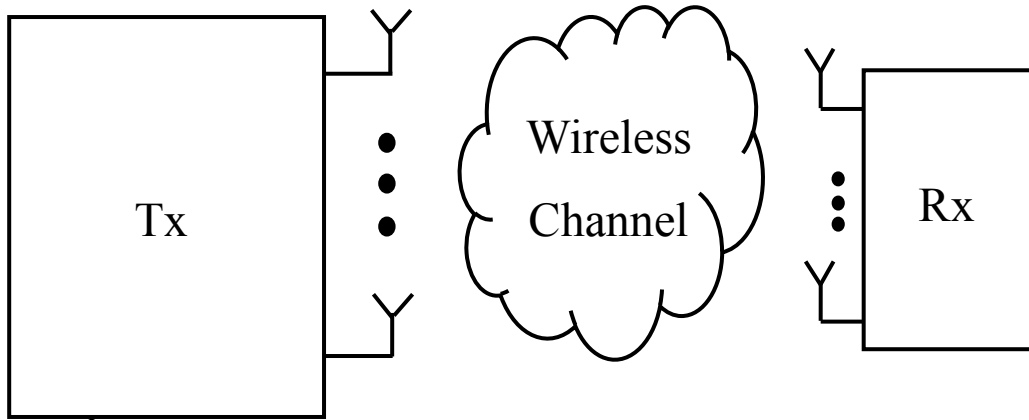
Figure 2.4: The architecture of SU-MIMO system.
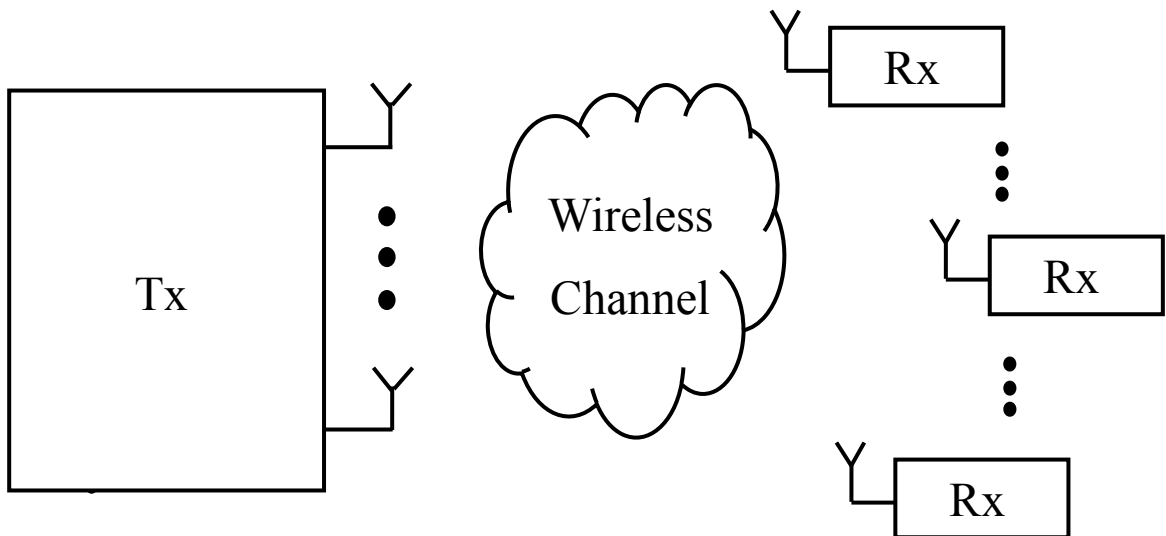
Figure 2.5: The MU-MIMO system architecture.

As defined in [16], compared with single input single output (SISO) systems, which only have single antenna equipped at both the transmitter and the receivers, MIMO systems use multiple antennas at both sides. MIMO offers higher capacity and reliability. Also, in terms of the multiplexing, diversity and array gains, the MIMO channels have considerable advantages over SISO channels.

MIMO systems have two basic configurations, which can be divided into the single user MIMO (SU-MIMO) and multi-user MIMO (MU-MIMO) systems, which are shown in Fig. 2.4 and Fig. 2.5, respectively. Linear antenna array is considered in both systems. Single active UE is served or scheduled in a transmission time interval (TTI) in SU-MIMO systems, while the large scale antenna arrays are equipped at both transmitter and receiver. In this kind of system, the single UE transmission means that all time-frequency resource are allocated to one terminal, which means the diversity in spatial domain is ignored. Compared with SU-MIMO, in MU-MIMO systems, the same time-frequency resource are reused by multiple UEs. The multi-UE diversity is considered in spatial domain, which brings large gains compared with SU-MIMO, especially in the cases that channels are spatially correlated.

Except from the advantage of the spatial diversity usage, compared with SU-MIMO systems, single antenna can be employed by each UE in MU-MIMO systems. In this case, the expensive equipment is just needed at the BS, which results in a large cost reduction.

In the comparison between SU-MIMO and MU-MIMO systems, last but not least, a rich scattering is seldom required in MU-MIMO systems because they are relatively less sensitive to the propagation environment. This feature may be beneficial to reduce the complexity of channel estimation in MU-MIMO systems.

2.2   AoA Estimation in MIMO System

2.2.1   Receiving Signal Model with Uniform Linear Array
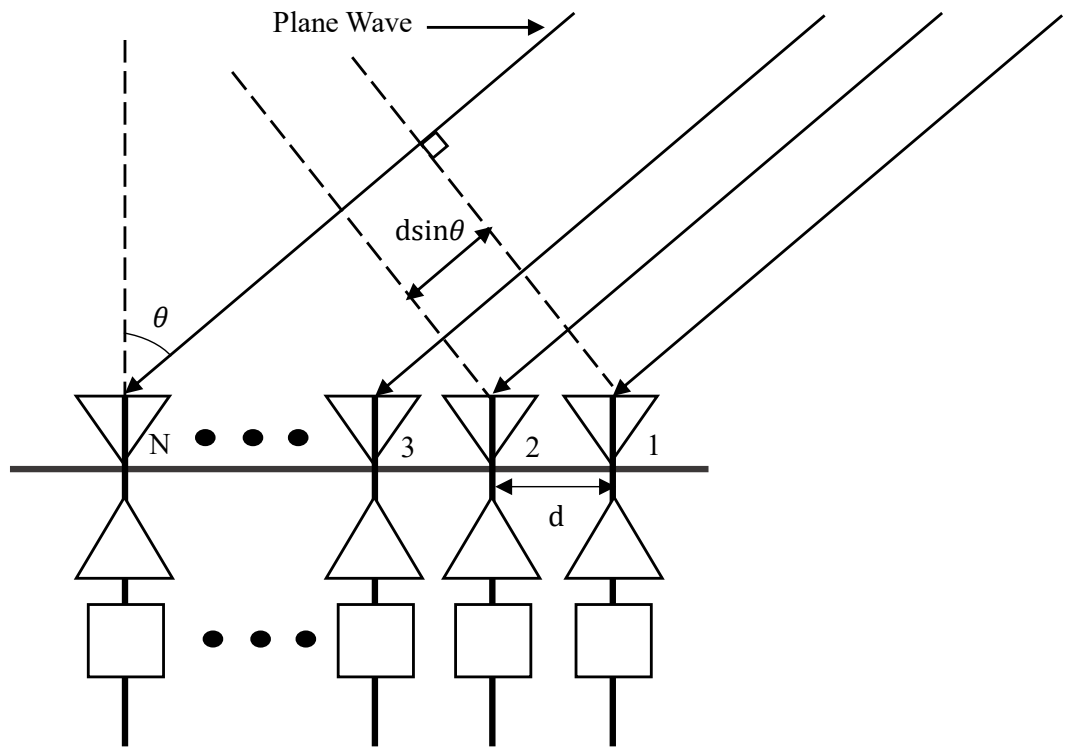
Figure 2.6: Plane wave arrives at ULA with AoA θ.

The uniform linear array (ULA) with N elements is employed to receive signals. As shown in Fig. 2.6, the distance between adjacent antennas is d, the plane wave is assumed to hit the antenna array with an angle θ. The angle θ is the AoA of the receiving signal plane wave, which is measured clockwise from the vertical line of the antenna array. In this section, the model of receiving signal vector on a ULA for AoA estimation is detailed.

In Fig. 2.6, a time delay $(i-1)\tau$ can be found for the signal received at the $i^{th}(i = 1, \cdots, N)$ element of the array antenna compared with the first element. Therefore, the signal at the ith

element in the time domain is expressed as

$$x_i(t) = x(t - (i-1)\tau) = e^{j2\pi ft} \cdot e^{-j2\pi f(i-1)\tau}, \tag{2.4}$$

where j represents the imaginary unit of a complex number, f is the carrier frequency, x is the transmitted signal. The time delay between adjacent antennas τ is

$$\tau = \frac{d\sin\theta}{c} = \frac{d\sin\theta}{f\lambda_0}. \tag{2.5}$$

λ_0 is the corresponding wavelength for the carrier frequency f. Substitute τ in Eq. (2.4) with Eq. (2.5)

$$\begin{aligned} x_i(t) &= e^{j2\pi ft} \cdot e^{-j2\pi f(i-1)\frac{d\sin\theta}{f\lambda_0}} \\ &= e^{j2\pi ft} \cdot e^{-j2\pi(i-1)\frac{d\sin\theta}{\lambda_0}} \\ &= e^{j2\pi ft} \cdot e^{-j(i-1)\psi} \end{aligned} \tag{2.6}$$

$\psi = \frac{2\pi d\sin\theta}{\lambda_0}$ is only decided by the AoA of signal on a specific antenna array with a fixed carrier frequency, it is seen as a new expression of AoA in the following equations for simplification. Considering the Additive White Gaussian Noise (AWGN) in the channel for transmission, the receiving signal vector with size $N \times 1$ at the antenna array can be written as

$$y(t) = \begin{bmatrix} 1 \\ e^{-j\psi} \\ e^{-j2\psi} \\ \vdots \\ e^{-j(N-1)\psi} \end{bmatrix} x(t) + \begin{bmatrix} n_1(t) \\ n_2(t) \\ n_3(t) \\ \vdots \\ n_N(t) \end{bmatrix}. \tag{2.7}$$

When M signals are received at the antenna array simultaneously with different AoA $[\psi_1, \psi_1, \cdots, \psi_M]$, the receiving signal vector model at ULA is

$$\begin{aligned} y(t) &= \sum_{m=1}^{M} \alpha(\psi_m) x^m + \mathbf{n}(t) \\ &= \mathbf{A}\mathbf{x}(t) + \mathbf{n}(t) \end{aligned} \tag{2.8}$$

where $\alpha\left(\psi_m\right) = \left[1, e^{-j\psi_m}, \cdots, e^{-j(N-1)\psi_m}\right]^T$ is defined as steering vector for the mth signal, $[\cdot]^T$ represents the transpose of a matrix or vector, $\mathbf{A} = \left[\alpha\left(\psi_1\right), \alpha\left(\psi_2\right), \cdots, \alpha\left(\psi_M\right)\right]$. $\mathbf{x}\left(t\right) = \left[x^1, x^2, \cdots, x^M\right]^T$ represents all the transmitting signals, and $\mathbf{n}\left(t\right) = \left[n_1\left(t\right), n_2\left(t\right), \cdots, n_N\left(t\right)\right]^T$ is the noise vector. $y\left(t\right)$ is the temporal sample of array receiving signal for the time instant t (snapshot t). A large number of temporal samples taken in the period with unchanged $\mathbf{A}$ are helpful for improving the AoA estimation accuracy.

2.2.2 Conventional AoA Estimation Methods

The antenna array based AoA estimation methods can be divided into the subspace methods and non-subspace methods [22]. it is mentioned that compared with subspace methods, some non-subspace methods, such as the maximum likelihood method, have computational intensive searching step and have poor performance in terms of estimation resolution. The MUSIC algorithm and ESPRIT algorithm belong to subspace methods, both of them are popular and widely utilized. In this section, the principles and procedures of MUSIC and ESPRIT algorithms are detailed, while their estimation resolution are compared in the cases with different numbers of signal snapshots, and in manifold SNR environments.

The main idea of MUSIC algorithm is to separate the signal space and noise space according to the orthogonality, the separation is achieved by the eigendecomposition step operated on the average array covariance matrix of the receiving signal. As the algorithm for comparison with the novel algorithm in the thesis, the MUSIC algorithm is detailed as follows.

In the MUSIC AoA estimation algorithm, the number of antennas in the receiving array should be ensured to be larger than the number of signals, i.e. $N > M$. The whole algorithm can be divided into four steps:

- Construct the average array covariance matrix based on the receiving signal vector;
  According to the receiving signal vector defined in Eq. (2.8), the matrix can be expressed as

$$\bar{\mathbf{R}}_{yy} = E\left[\mathbf{y}\mathbf{y}^H\right] = \mathbf{A}E\left[\mathbf{x}_1\mathbf{x}_1{}^H\right]\mathbf{A}^H + \sigma_n^2\mathbf{I}, \qquad (2.9)$$

  where $[\cdot]^H$ represents the conjugate transpose of a matrix or vector, σ_n^2 represents the

noise variance, $E[\cdot]$ is the expectation operator based on multiple signal snapshots.

- Obtain the eigenvalues and eigenvectors of correlation matrix by eigendecomposition, resort eigenvalues and eigenvectors;

  After the operation of eigendecomposition, N eigenvalues and their corresponding eigenvectors can be obtained. Sort the eigenvalues from the largest value to the smallest value, at the same time change the order of eigenvectors. The reordered matrix for eigenvectors is $[\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_M, \mathbf{v}_{M+1}, \cdots, \mathbf{v}_N]$.

- Form $N-M$ eigenvectors with the smallest eigenvalues into the matrix of noise subspace;

  The matrix of noise subspace is $\mathbf{V}_n = [\mathbf{v}_{M+1}, \cdots, \mathbf{v}_N]$, it satisfies

$$
\begin{aligned}
&[\mathbf{v}_{M+1}, \cdots, \mathbf{v}_N] \perp [\alpha(\psi_1), \alpha(\psi_2), \cdots, \alpha(\psi_M)] \\
&\Rightarrow \quad \alpha^H(\psi_m) \mathbf{V}_n \mathbf{V}_n^H \alpha(\psi_m) = 0, \quad (m = 1, \cdots, M)
\end{aligned}
, \quad (2.10)
$$

Eq. (2.10) means the signal space, which is represented by the steering vectors of signal components are orthogonal to the noise subspace eigenvectors.

- Scan the range of angle to create spatial spectrum, the angles located at peaks of spatial spectrum are the estimated AoA $\hat{\theta}_m$ $(m = 1, \cdots, M)$.

  With the matrix of noise subspace, the MUSIC spatial spectrum can be generated by the steering vectors scanning in the range of angle as below

$$
\begin{aligned}
P(\psi) &= \frac{1}{\alpha^H(\psi) \mathbf{V}_n \mathbf{V}_n^H \alpha(\psi)} \\
&\left(\psi = \frac{2\pi \sin\theta}{\lambda_0}, \quad \theta \in [-90°, 90°] \right)
\end{aligned}
, \quad (2.11)
$$

By locating the peaks of the spatial spectrum, the AoA of the signals can be estimated.

Fig. 2.7 gives an example of MUSIC spatial spectrums. Comparing different curves, the influence of SNR and snapshots can be found. 8 antennas are employed on the receiving array, and the AoA for two signals are estimated at the same time, for which the AoA are fixed as -30 degrees and 30 degrees. It shows that the spikes are becoming more definite with the increase of SNR, while the performance rises sharply with more available snapshots.
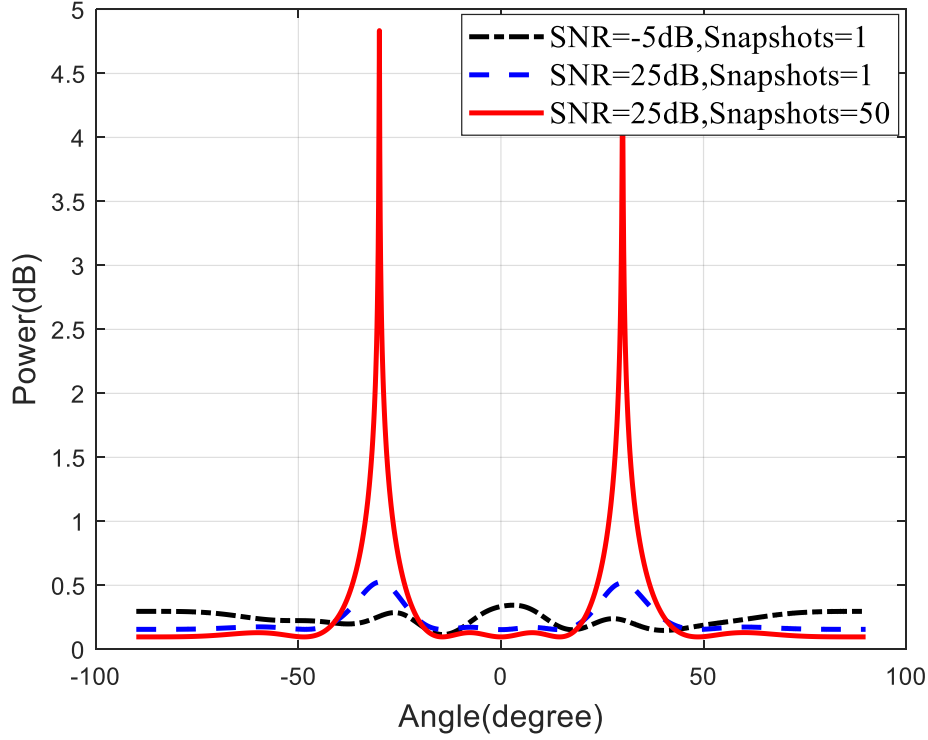
Figure 2.7: The MUSIC spatial spectrums for two signals AoA estimation with different SNR and numbers of snapshots.

The ESPRIT algorithm decomposes an antenna array with N sensors into two identical subarrays, N_e elements are equipped at each subarray, while the distance between two subarrays is known. It is the objective of this algorithm that estimating AoA by calculating the operator φ_e for the rotation from signal on one subarray to that on the other one.

Suppose that $N_e = N - 1$ for each subarray, and the distance between two subarrays is d. Matrixes of steering vectors for both subarrays are part of $\mathbf{A}$ for the whole antenna array, which is

$$
\mathbf{A}_1 = \begin{bmatrix} 1 & \cdots & 1 \\ e^{-j\psi_1} & \cdots & e^{-j\psi_M} \\ \vdots & \cdots & \vdots \\ e^{-jN_e\psi_1} & \cdots & e^{-jN_e\psi_M} \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} e^{-j\psi_1} & \cdots & e^{-j\psi_M} \\ e^{-j2\psi_1} & \cdots & e^{-j2\psi_M} \\ \vdots & \cdots & \vdots \\ e^{-j(N_e+1)\psi_1} & \cdots & e^{-j(N_e+1)\psi_M} \end{bmatrix}. \quad (2.12)
$$

It can be noticed that $\mathbf{A}_2 = \mathbf{A}_1 \mathbf{f}_T$, where $\mathbf{f}_T$ is a diagonal matrix decided by ψ of each path

of signal. The expression of $\mathbf{f}_T$ is

$$
\mathbf{f}_T =
\begin{bmatrix}
e^{-j\psi_1} & 0 & \cdots & 0 \\
0 & e^{-j\psi_2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & e^{-j\psi_M}
\end{bmatrix},
\tag{2.13}
$$

When the subarray distance and the carrier frequency of each signal are known, the AoA of each signal can be calculated from $\mathbf{f}_T$. However, the matrix of steering vectors of subarrays cannot be recovered from receiving signals directly, hence, $\mathbf{f}_T$ cannot be obtained. In the ES-PRIT algorithm, based on the fact that the steering vectors in matrix $\mathbf{A}$ span the same subspace as $\mathbf{V}_s = [\mathbf{v}_1, \cdots, \mathbf{v}_M]$, therefore, the signal subspace $\mathbf{V}_s$ is utilized to generate the matrixes of eigenvectors for $\mathbf{A}_1$ and $\mathbf{A}_2$, which are represented by $\mathbf{V}_1$ and $\mathbf{V}_2$, respectively. There exists an invertible matrix $\mathbf{T}$ to achieve the transform from $\mathbf{A}_1$, $\mathbf{A}_2$ to $\mathbf{V}_1$ and $\mathbf{V}_2$, which brings

$$
\begin{aligned}
\mathbf{V}_1 &= \mathbf{A}_1 \mathbf{T} \\
\mathbf{V}_2 &= \mathbf{A}_2 \mathbf{T} = \mathbf{A}_1 \mathbf{f}_T \mathbf{T}
\end{aligned}
,
\tag{2.14}
$$

the relationship between $\mathbf{V}_1$ and $\mathbf{V}_2$ is

$$
\mathbf{V}_1 = \mathbf{V}_2 \mathbf{T}^{-1} \mathbf{f}_T^{-1} \mathbf{T},
\tag{2.15}
$$

$[\cdot]^{-1}$ represents the inverse of a matrix. Record $\mathbf{F}_T^{-1} = \mathbf{T}^{-1} \mathbf{f}_T^{-1} \mathbf{T}$, therefore, $\mathbf{V}_1 = \mathbf{V}_2 \mathbf{F}_T^{-1}$, i.e. $\mathbf{V}_2 = \mathbf{V}_1 \mathbf{F}_T$. $\mathbf{f}_T$ is a diagonal matrix of the eigenvalues of $\mathbf{F}_T$. With $\mathbf{F}_T$, $\mathbf{f}_T$ can be obtained, and the AoA of signals can be calculated.

On the basis of the aforementioned principles, the ESPRIT algorithm can be achieved by the following steps:

- Construct the input correlation matrix $\bar{\mathbf{R}}_{yy}$ based on the receiving signal vector, find its signal subspace $\mathbf{V}_s$ by eigendecomposition;

- Extract signal subspaces of subarrays $\mathbf{V}_1$ and $\mathbf{V}_2$ from $\mathbf{V}_s$;

- Estimate matrix $\mathbf{F}_T$ based on the least square criterion, and for the model $\mathbf{V}_2 = \mathbf{V}_1 \mathbf{F}_T$,

the estimated $\mathbf{F}_T$ is $\hat{\mathbf{F}}_T = \left(\mathbf{V}_1^H \mathbf{V}_1\right)^{-1} \mathbf{V}_1^H \mathbf{V}_2$;

- The diagonal matrix for the eigenvalues of $\hat{\mathbf{F}}_T$ is the estimated $\hat{\mathbf{f}}_T$, the estimated AoA of signals in degrees are obtained by $\hat{\theta}_m = -\frac{180}{\pi} \sin^{-1}\left(\frac{\arg\left(e^{-j\psi_m}\right)\lambda_0}{2\pi d}\right)$, $(m = 1, \cdots, M)$.
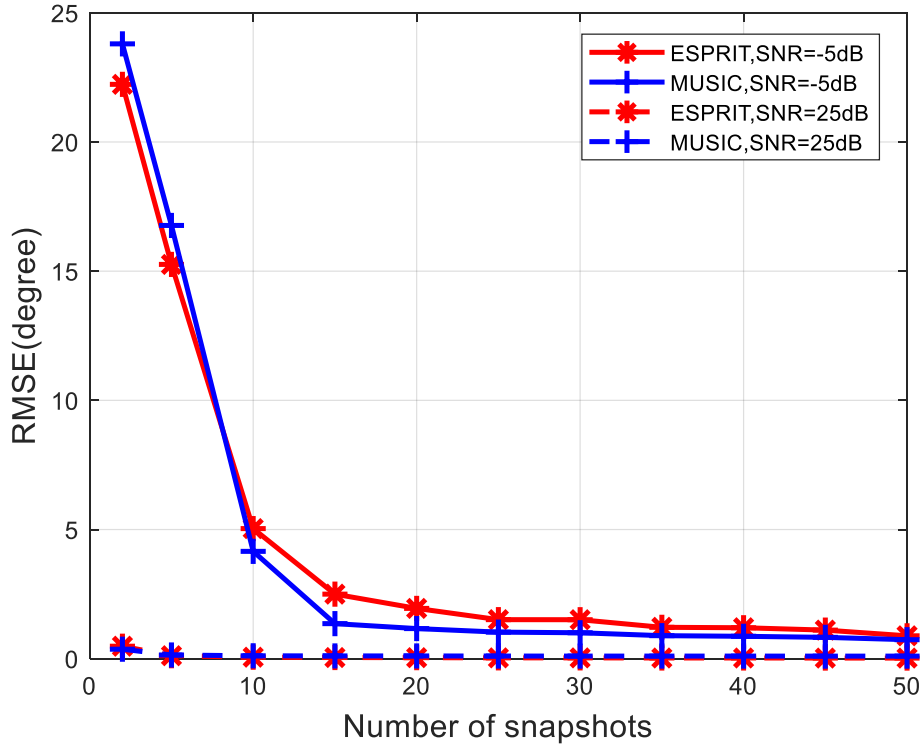
Figure 2.8: The comparison of MUSIC algorithm and ESPRIT algorithm with different SNR and increasing numbers of snapshots.

In Fig. 2.8, The root mean square error (RMSE), which is defined as Eq. (2.16), is compared between MUSIC algorithm and ESPRIT algorithm in the environment with low and high SNR ($-5dB$ and $25dB$) SNR and with different numbers of snapshots. N_{ep} times of experiments are done to calculate each point in the figure to prevent the influence of abnormal results. The number of antennas in the whole array is 8, while 7 antennas are allocated to each subarray. Similar performance can be found for both of the algorithms with any SNR or number of snapshots, while the low RMSE, which means the high estimation resolution can be provided by two algorithms in high SNR environments with only several snapshots. However, with low SNR, more snapshots are required for high estimation resolution. Therefore, the high

resolution AoA estimation is still an open issue in low SNR environment with few snapshots.

$$RMSE = \left[\frac{1}{MN_{ep}} \sum_{n=1}^{N_{ep}} \sum_{m=1}^{M} \left(\theta_m - \widehat{\theta}_m \right)^2 \right]^{1/2}, \tag{2.16}$$

2.2.3   Applications of Neural Network in AoA Estimation

Many research works are focusing on the neural network based high resolution AoA estimation methods, in order to achieve accurate AoA estimation for different cases. The convolutional neural network (CNN), radial basis function neural network (RBFNN) and the multi-layer perceptron (MLP) based neural networks are three kinds of popular neural network models which have been studied for a long time. Many papers can be found about their applications in AoA estimation.

The structure of the CNN model for beam based AoA estimation designed in [23] is shown in Fig. 2.9. It can be found that the CNN model is consisted of convolutional layers, sub-sampling layers and fully-connected layers. The convolutional layers are used to extract the high-level features such as edges, from the input nodes. In this model, each convolution plane is connected to one or more subsampling planes. In fact, the subsampling layers are always set after convolutional layers, which are designed to downsample the output of convolutional layers. The fully-connected layer is utilized as the output layer of this model. The CNN model designed in [23] is mentioned to be capable to select the suitable radiation beams without the knowledge of the source signals number, while the signals are received from different directions. The input covariance matrix defined in Eq. (2.9) is utilized to generate input matrix $\mathbf{X}_{in}$, while the output $\mathbf{Y}_{out}$ are the beams indices selected for each signal. In this AoA algorithm, although ranges of the selected beams can always cover the directions of signals, the estimation resolution is not high for it is limited by the coverage of beams in angle domain.

The RBFNN models are utilized with relatively fixed structure, so the number of neurons in hidden layer is very important for the performance of model [24]. The model is named as this because the radial basis function is used in the hidden layer of model. It is said that the RBFNN with enough hidden neurons can approach to functions with any resolution. To
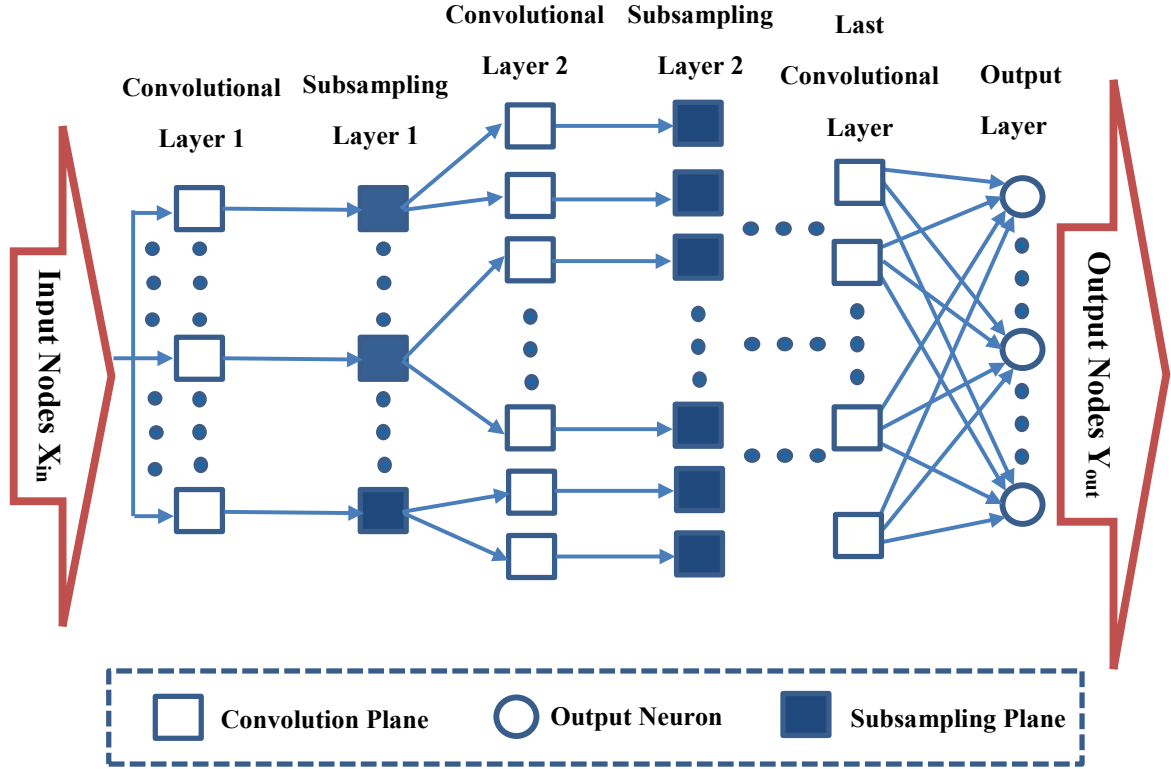
Figure 2.9: The structure of a CNN model.

be mentioned that the input nodes shown in Fig. 2.10 for this model should be the vectored covariance matrix, while the output nodes are the AoA of signals. The mapping function of RBFNN in hidden layer is given in Eq. (2.17)

$$\mathbf{Y}_{out} = F(\mathbf{X}_{in}) = \sum_{l=1}^{L} \mathbf{w}_l \phi(\|\mathbf{b}_l - \mathbf{X}_{in}\|), \tag{2.17}$$

where $\phi(x)$ is the radial basis function of x, and $\|\cdot\|$ represents the 2-norm operation. L is the number of neurons in the hidden layer, $\mathbf{w}_l$ represents the weight of hidden neurons, while the $\mathbf{b}_l$ is the center vector for the radial basis function, which is usually determined through K-means clustering. The radial basis function $\phi(x)$ is usually assumed to be un-normalized Gaussian function, which is

$$\phi(x) = e^{\frac{-x^2}{2\sigma^2}}, \tag{2.18}$$

where σ denotes the standard deviation of radial basis function. In fact, due to the simple

structure of RBFNN, the AoA estimation error of this model can hardly be low enough.
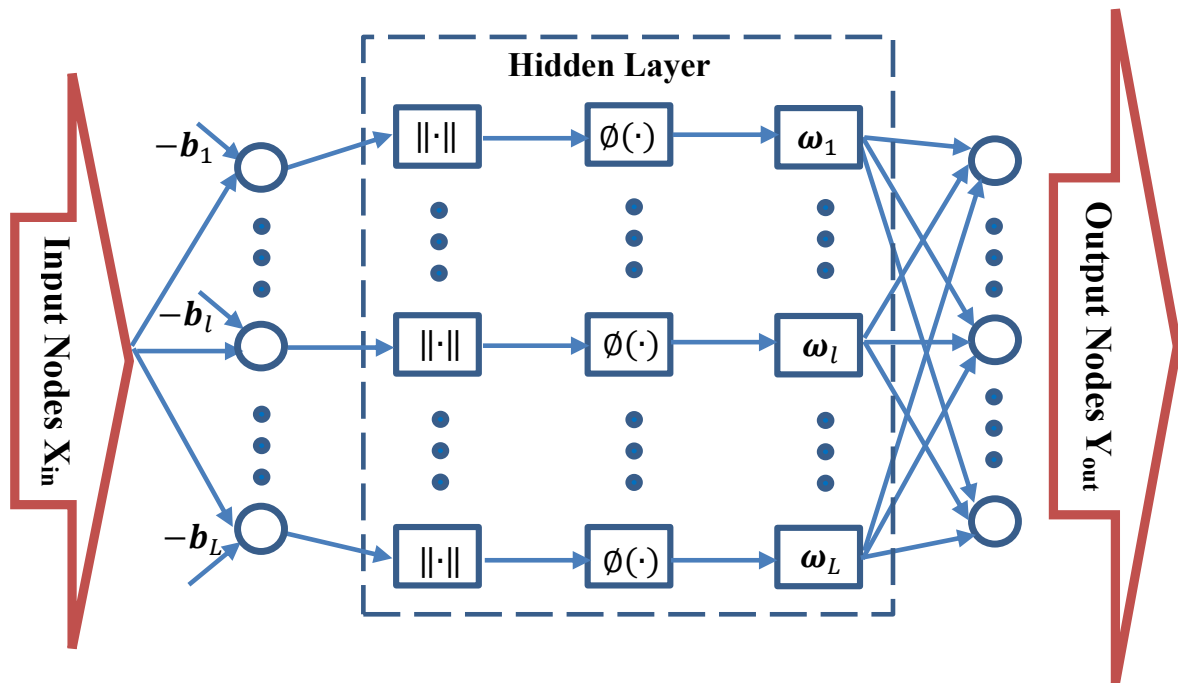
Figure 2.10: The block diagram representation of the RBFNN.

The MLP neural network is a kind of multi-layer and nonlinear neural network [25]. Commonly a MLP neural network is consisted of multiple layers, which include an input layer, hidden layers and an output layer. Nonlinear activation functions can be selected for each fully-connected layer (FCL) neuron, after the activation layer (AL) these neurons are connected with the neurons in the next layer. Multiple hyperparameters are needed for the MLP model structure design, including the number of hidden layers, the number of neurons for each layer, the activation functions selection, and so on. These predefined hyperparameters make the MLP neural network design become flexible, in order to satisfy any kinds of input-output mapping requirements. The structure of MLP neural network model is shown in Fig. 2.11. Many MLP neural network based AoA estimation algorithms have been proposed to provide high resolution results, while the large potential still can be found for both high resolution and robust estimation by MLP neural network model due to its high flexibility.
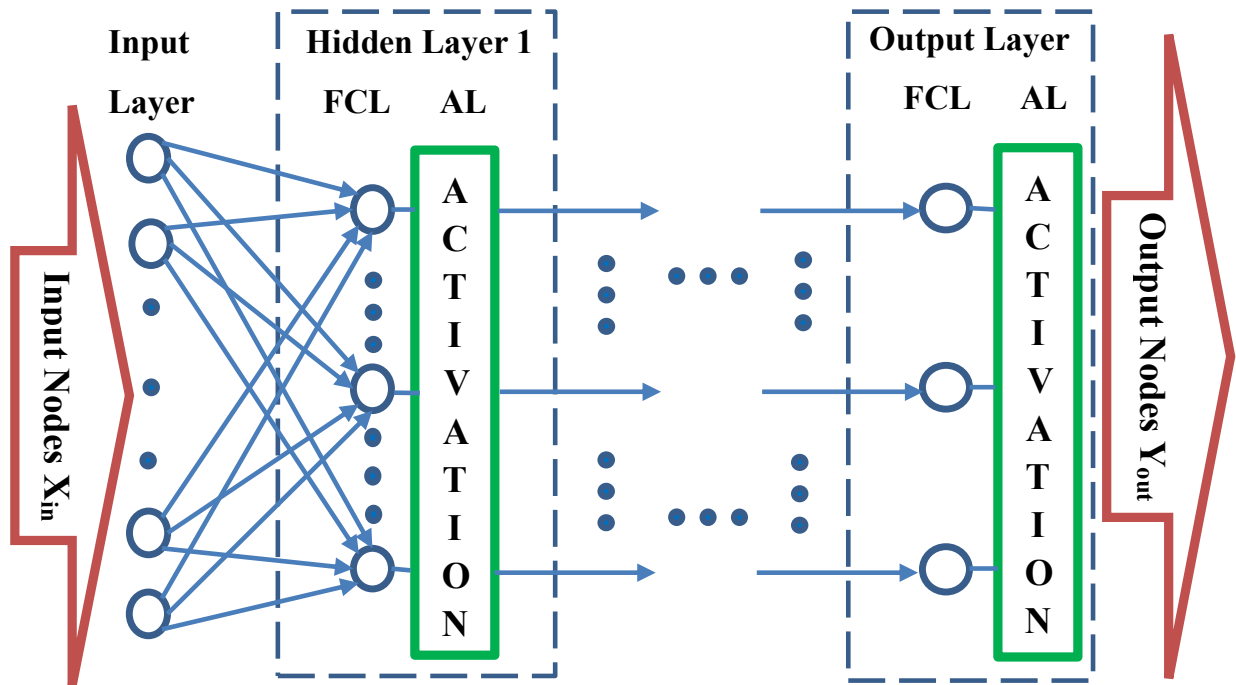
Figure 2.11: The structure of MLP neural network.

2.3 FDD/TDD Communications

2.3.1 Frame Structures for FDD/TDD
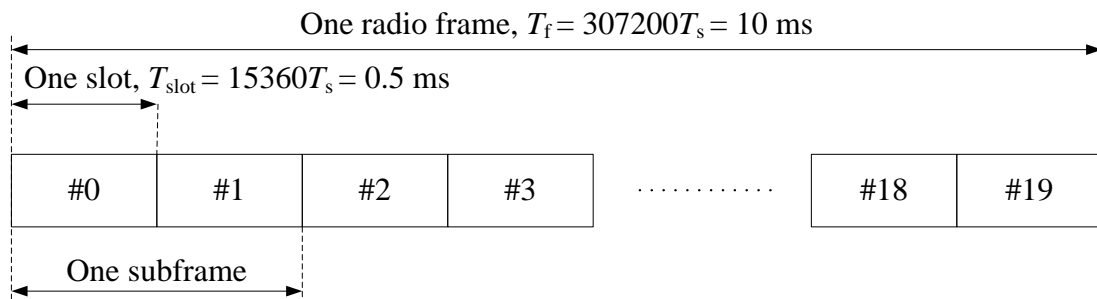
Figure 2.12: Frame structure type 1.

In this section, the frame structures for TDD systems and FDD systems are compared to prove the requirement of separate beamforming designs for different systems. The frame structure type 1 is defined for FDD scenario in 3GPP TS 36.211 [21], which is shown Fig. 2.12. For both

full duplex and half duplex FDD, 10 subframes, 20 slots, or up to 60 subslots are available for both uplink and downlink transmissions in each 10*ms* interval. Compared with the half-duplex FDD systems, the full-duplex FDD has fewer restrictions, it can transmit and receive signal at the same time.
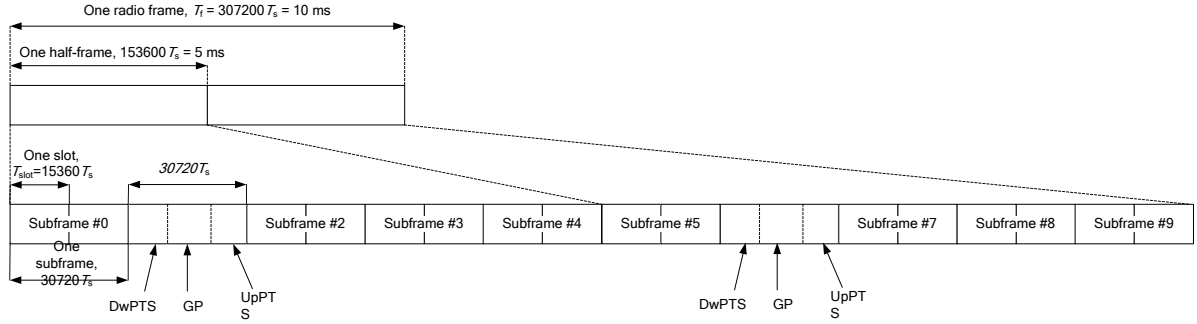
Figure 2.13: Frame structure type 2 (for 5 ms switch-point periodicity).

Frame structure type 2 is only applicable to TDD systems. As shown in Fig. 2.13, the frame structure is similar to that of FDD systems, for each 10*ms* radio frame, 10 subframes are equally divided, each of which can be further split into two slots. The signals cannot be transmitted and received simultaneously, therefore, it is mentioned that this type of frame structure can support two kinds of switch-point periodicities for the conversion of uplink and downlink signals, one subframe is required for the downlink pilot time slot (DwPTS), uplink pilot time slot (UpPTS) and guard period (GP). The smallest downlink-to-uplink switch-point periodicity is 5*ms*, which is much higher than the coherence time of mmWave channel even if the UEs are moving in a low speed. Here the coherence time is the period over which the channel impulse response (channel matrix) is considered to be unchanged. When carrier frequency $f = 38GHz$, UE moving speed $v = 5m/s$, the coherence time T_{co} is

$$T_{co} = \frac{0.423}{vf/c} \approx 0.63ms < 5ms, \tag{2.19}$$

here $c = 3 \times 10^8$ is the speed of light. Therefore, the uplink signal is not always available in each coherence time, which means the beams can hardly be selected for downlink transmission adaptively. Based on the difference of frame structures for TDD and FDD systems, it can be

concluded that the most significant difference between two systems is the availability of uplink signal during the downlink transmission. Due to this difference, the beamforming algorithms should be designed separately to adapt to different systems.

2.4 Beamforming in MIMO System

2.4.1 Types of Beamforming

With the development of massive MIMO system, the beamforming technique is explored based on the large scale antenna array for its advantage of large power gain. In MU-MIMO system, the beamforming design can not only improve the link quality, but also reduce the influence of interference [26].

In many beamforming methods, the transmitting beamformers are designed with the assumption of available full CSI. However, in practise, the transmitters with larger scale antenna arrays can only obtain partial CSI, such as the AoD of paths, based on the feedback in TDD systems communications or the channel reciprocity in FDD systems communications [27]. The codebook based beamforming and angle based beamforming are two kinds of partial CSI beamforming methods which are widely explored.

2.4.2 Codebook based Beamforming

The codebook based beamforming works to improve transmitting power and reduce interference by the perfect selection of codewords, which are generated based on the codewords from a codebook. Here the codebook is defined as a matrix consisted of vectors for precoding (beams generation). The beam training step is always required to find out the best transmitting/receiving beams for UEs, which should be suitable to their CSI. In the beam training step for regular codebook based beamforming, the beams using all codewords as precoders are transmitted, after which the indices of codewords brings large power gain and small interference are feedbacked for the beams generation in the data transmission. This kind of training

process is called exhaustive search training, this training method is conceptually straightfor-ward. However, because of the large number of candidate codewords in the codebooks for mmWave communication, the overall search time is prohibitive, and a large amount of resource is wasted as the training and feedback overhead.

To reduce the overhead for the beam training step while keeping the advantage of perfect codeword selection, the hierarchical codebook is designed in [28], which is utilizing codewords from several layered codebooks for training, the beams generated by codewords from different layered codebooks have diverse beamwidth and power gain. The principle of the hierarchical codebook is to train beams generated by codewords from the lowest layer to the highest layer. Not all of the codewords are needed to be trained in each layered codeword because the training range on the angle domain is narrowed by the selected codeword range in the last layer. Two basic criteria should be satisfied to design the hierarchical codebook, which are:

1. The beam coverage union generated by all the codewords in each codebook should be able to cover the whole angle domain;

2. The coverage of each beam generated by a codeword from a non highest-layer codebook should be covered by the coverage union of a couple of codewords in the next layer.

The hierarchical codebook is designed for the codeword selection at both the transmitting side and receiving side, for the two sides the designing principles are totally the same. For the cth $(c = 1, \ldots, C - 1)$ layer codebook, the range of beams generated by C_w adjacent codewords in the $(c + 1)^{th}$ layer should cover the beamwidth of a codeword in the cth layer. The highest-layer codebook is firstly designed with N codewords, which number is equal to the number of antennas. Each codeword can be expressed as

$$\mathbf{w}(C, in) = \alpha(N, -1 + (2in - 1)/N), \quad (in = 1, \ldots, N), \tag{2.20}$$

$$\alpha(N, \sin\theta) = \frac{1}{\sqrt{N}}\left[1, e^{j\pi\sin\theta}, \ldots, e^{j\pi(N-1)\sin\theta}\right]^T. \tag{2.21}$$

The number of layers $C = \log_{C_w}(N)$. For the codewords in the layer $c = \log_{C_w}(N) - k$, where $k = 1, 2, \ldots, \log_{C_w}(N) - 1$, the codewords can be generated by the following steps

1. The elements in $\mathbf{w}(c, 1)$ can be uniformly divided into $L = C_w^{floor\left(\frac{k+1}{2}\right)}$ groups, each group is $\mathbf{w}_l(c, 1) = [\mathbf{w}(c, 1)]_{(l-1)\frac{N}{L}+1:l\frac{N}{L}}, \quad (l = 1, \ldots, L)$;

2. If k is odd, set $N_L = L/2$; else, $N_L = L$. $\mathbf{w}_l(c, 1)$ is set as

$$\mathbf{w}_l = \begin{cases} e^{-jl\frac{N-L}{N}\pi}\alpha\left(\frac{N}{L}, -1 + \frac{L(2l-1)}{N}\right), & l = 1, \ldots, N_L \\ \mathbf{0}_{\frac{N}{L}\times 1}, & l = N_L + 1, \ldots, L \end{cases} ; \qquad (2.22)$$

3. $\mathbf{w}(c, in) = \mathbf{w}(c, 1) \circ \sqrt{N}\alpha\left(N, \frac{2(in-1)}{C_w^c}\right), \quad (in = 2, 3, \ldots, C_w^c)$.

4. Normalize $\mathbf{w}(c, in)$.

Where $\circ$ represents entry-wise product. An example of the hierarchical codebook is shown on Fig. 2.14. A multi-sectional search beamforming algorithm is proposed in [28] utilizing the hierarchical codebook. When the hierarchical codebooks utilized at the transmitting side and the receiving side are totally the same, it requires $2\log_{C_w} N$ bits for the training, and other $2\log_{C_w} N$ bits for the feedback for each receiver in order to select suitable codewords for both precoding and receiving. For the application of hierarchical codebook in the MU-MIMO system, a large scale antenna array serves multiple single-antenna UEs in the downlink, therefore, $\log_{C_w} N$ bits are needed for downlink training and uplink feedback, separately, which means $2\log_{C_w} N$ are used for the perfect precoding codeword selection of a UE. It can be found that the large number of antennas at the transmitting side brings large overhead for codewords selection.

The SWR oriented beamforming designed based on the hierarchical codebook further reduces the overhead for codebook based beamforming while the SR is partly sacrificed [29][67]. This algorithm also searches the suitable codeword for a UE from the lowest layer to the highest layer. However, the SWR is considered in the algorithm design, which is the parameter reflects both the influence of overhead and SR. In the training process, when the SWR is estimated to reduce with the further training, the training step stops and the best codeword in the current layer codebook is selected.
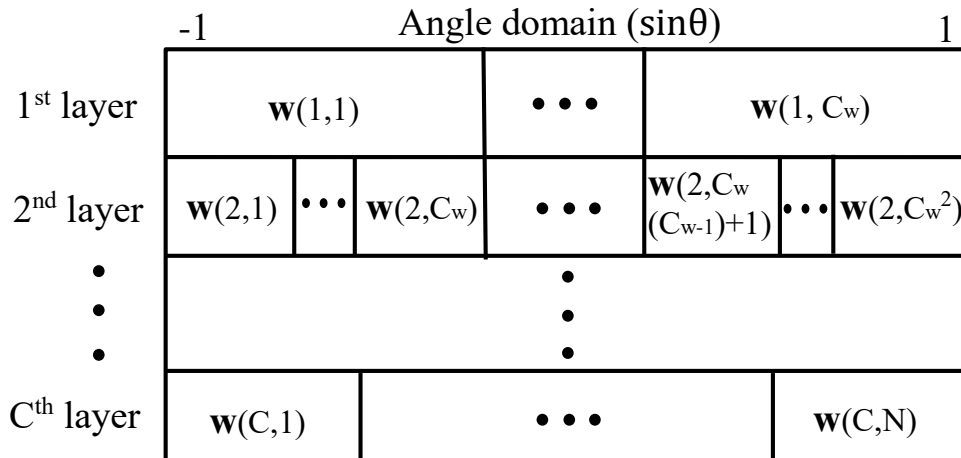
Figure 2.14: Hierarchical codebook for analog beamforming.

2.4.3   Angle based Beamforming

Angle based beamforming makes full use of the AoD of dominant paths in the transmission channel. As a partial CSI based beamforming, getting the angle information of paths is much easier than obtaining the full CSI information. In TDD systems, the uplink pilot signal can be utilized to estimate the downlink CSI directly due to the channel reciprocity; While in FDD systems, for the cases that the carrier frequencies of uplink and downlink channels are not far away from each other, the AoA estimated for the uplink paths can be viewed as the AoD for the corresponding downlink paths, when the time elapse between uplink and downlink signals can be neglected [9][30].

With the AoD information of dominant paths, the angle based beamforming can be designed to increase the power gain in the required directions, while reduce the power gain in the interference directions. The beamforming design can be applied on both analog and digital parts. The angle based analog beamforming vector for the mth UE with AoD θ_m is

$$\mathbf{w}_A^m(N, \sin\theta_m) = \frac{1}{\sqrt{N}}\left[1, e^{j\pi\sin\theta_m}, \ldots, e^{j\pi(N-1)\sin\theta_m}\right]^T. \tag{2.23}$$
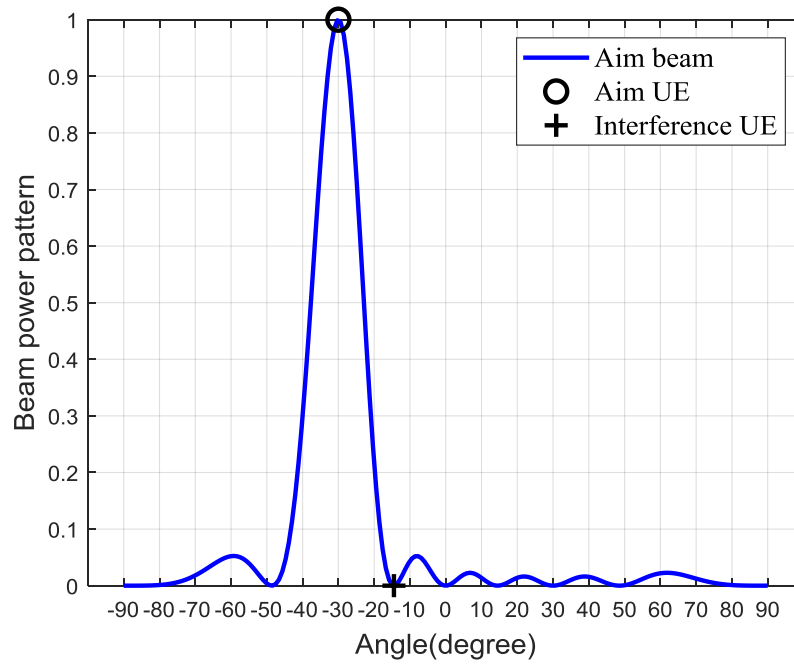
(a)
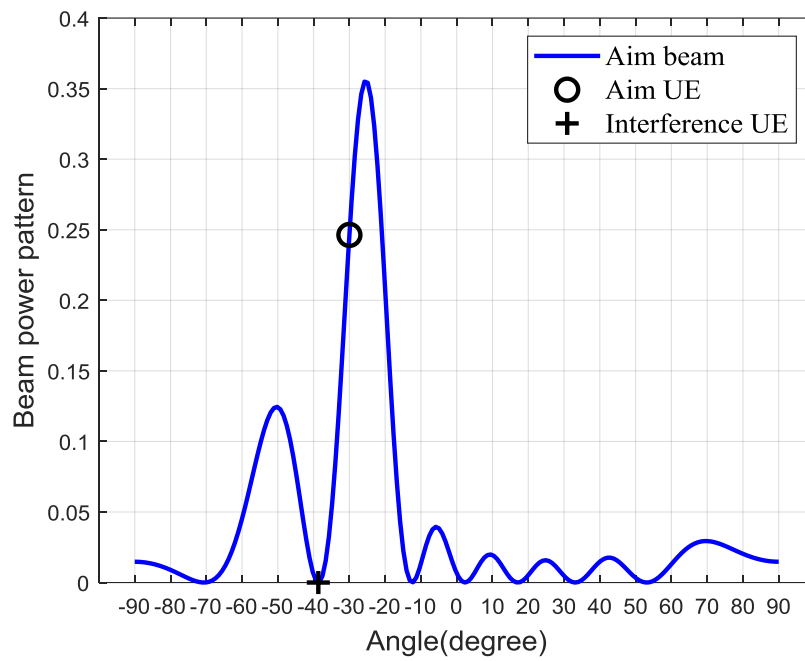
(b)

Figure 2.15: The comparisons among hybrid beamforming with $N = 8$ of (a) AoD of interference UE is $-15°$ (b) AoD of interference UE is $-38°$.

Compared with the analog beamforming design, the digital beamforming considers not only the AoD of the aim UE, but also that of other UEs. Therefore, the estimated AoD matrix $\mathbf{H} = [\alpha(N, \sin\theta_1), \alpha(N, \sin\theta_2), \cdots, \alpha(N, \sin\theta_m)]^H$ should be utilized for digital beamforming design. The zero forcing (ZF) digital algorithm is widely used in digital precoder design, which is chosen as an example [8]

$$\mathbf{W}_{ZF} = \mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H\right)^{-1}. \tag{2.24}$$

However, in the applications with large scale antenna arrays, the size of matrix $\mathbf{H}$ is too large to apply the digital beamforming in a limited time. The hybrid beamforming is designed to face this challenge, it utilizes both analog and digital beamforming at the same time. This means the ZF algorithm can work on the dimension-reduced AoD matrix in this example, which is

$$\mathbf{W}_{AZF} = \mathbf{H}_A^H\left(\mathbf{H}_A\mathbf{H}_A^H\right)^{-1}, \tag{2.25}$$

$$\mathbf{H}_A = \mathbf{H}\mathbf{W}_A, \tag{2.26}$$

where $\mathbf{W}_A = \left[\mathbf{w}_A^1, \mathbf{w}_A^2, \cdots, \mathbf{w}_A^M\right]$ is the matrix of analog beamforming vectors for M UEs. It is the fact that compared with analog beamforming, the digital beamforming is designed based on all UEs AoD vector, i.e. for a specific aim UE with fixed AoD, the difference on the AoD of interference UEs brings different digital beamforming design, while the analog beamforming for this UE is always the same. Fig. 2.15 compares the hybrid beamforming design for an aim UE with fixed AoD, while the interference UE has different AoD. The totally different beam patterns (power gain and beamwidth) prove our conclusion. We can also find that the hybrid beamforming can work to reduce interference to other UEs with any dominant path AoD with flexibility. However, this may reduce the power of signal to the aimed UE.

2.5   Chapter Summary

This chapter starts from the introduction to the mmWave channel and the MIMO system in mmWave communicatons. Except from the configurations of antenna arrays and two kinds of signal preprocessing methods with MIMO, the SU-MIMO and MU-MIMO systems are

detailed as well. After that, the antenna array based AoA estimation is review in the second section. The receiving signal model is introduced at first, based on that the procedures of AoA estimation by MUSIC and ESPRIT algorithms are shown step by step, the numerical results for the performance of MUSIC and ESPRIT algorithms with diverse SNR and snapshots are given, which show the challenge in conventional AoA estimation explicitly. Then the neural networks with different structures and used to be mentioned in AoA estimation methods design are discussed to compare their characteristics. In Section 2.3 and 2.4, the frame structures of TDD and FDD systems are compared to prove the necessity of separated beamforming designs in different systems. Then in the last section, some low overhead partial CSI beamforming methods including the codebook and angle based beamforming, are introduced and compared.

Chapter 3

Multi-Pattern Codebook based Low Overhead Beamforming in TDD mmWave MIMO System

In TDD systems, the downlink angle based beamforming requires extra uplink feedback to estimate AoA information, which increases the overhead for beamforming and leads to a sacrifice of symbols for data transmission. In this chapter, the multi-pattern codebooks are utilized to provide beams with pattern-controlled beamwidth and beamforming gain in spatial frequency domain, which works as the fundamental of long serving period beamforming with low feedback cost.

3.1 Introduction

In fulfilling the expectations of the 5G wireless systems, mmWave communications have attracted extensive research interests due to the available bandwidth and extremely high data rates promised [20]. Due to the short wavelength of mmWave band, a large number of antennas can be packed into a given small space hence massive MIMO technique with increased array gain becomes feasible [31]. However, mmWave communications face numerous challenges. One of the biggest challenges is the high free-space path loss resulting from high carrier frequency [32]. Thanks to massive MIMO, large scale antenna arrays can be employed to overcome the high path loss through beamforming [33]. Besides, reduced channel coherence time caused by high carrier frequency becomes the other challenge in mmWave massive MIMO system [34].

This is because the mmWave beamfoming need to be designed as fast as possible within the coherence time to keep up with channel changes [32]. However, in TDD systems, the uplink signal can hardly be available within each coherence time, which introduces high signalling, feedback and pilot overhead. Such huge overhead leads to a decrease in performance, particularly data throughput, which motivates us to design new mmWave beamforming schemes with low overhead.

To address the large overhead problem, different beamforming schemes have been proposed recently. The codebook based RF beamforming scheme has been employed in IEEE 802.11ad [35], which utilizes the sequential downlink-uplink (SDU) sector sweep combination method to reduce the beam training overhead. Furthermore, authors in [32] put forward a sequential downlink-downlink (SDD) sector sweep combination scheme to extend the applying scerario of [35] to MU-MIMO system. In addition, paper [36] proposes a fast beam training method which can be adapted to different modulation patterns, to reduce the training overhead.

However, the previous studies select codewords based on accurate channel information, which incurs high complexity due to the high dimensional of channel matrix in massive MIMO. Moreover, the channel effects on the transmitted signal changes substantially beyond the channel coherence time [37]. As a result, channel information has to be updated frequently in updating the serving beams. Compared with accurate channel acquisition, it is easier to obtain angular information of UEs [38]. Besides, angle domain based massive MIMO transmitter can reduce training overhead and improve channel estimation accuracy compared to the beamspace approach [39]. On top of that, compared with accurate channel information, angular information shows low changing speed. In TDD systems, the CSI for the uplink channel can be viewed as that for the downlink channel within a limited time delay, due to the channel reciprocity [40]. It means the AoA for the LoS path detected by the uplink signal, which is more stable compared with AoA of other paths, can be considered as the AoD of the corresponding path in the downlink. Therefore, it is unnecessary to frequently update serving beams for LoS path AoA based beamforming, this beamforming method is referred as angle based beamforming for short in the following content.

The overhead of angle based beamforming is decided by the beams changing frequency and the overhead required for each time of beam changing. The beam changing frequency

can be estimated with a special codebooks design. For UEs with similar moving speed in a large serving space, different angular speeds, which are defined as the changing rates of AoA for LoS paths, and the path loss of these UEs, varies with the distances from UEs to the BS. To serve UEs with different distances, serving beams with diverse patterns should be utilized based on the multi-pattern codebooks design [44], in order to provide beams with adaptable beamforming gain and beamwidth. The serving space is zoned according to the distance to BS, each zone should be served by a specific pattern of codebook. Therefore, with the knowledge of speeds and location information of UEs, the selected beam serving time can be estimated. To achieve the low overhead beamforming objective with multi-pattern codebooks, the algorithms for location estimation based UE selection with low overhead should be designed, which will be introduced in the next chapter.

In this chapter, the multi-pattern codebooks design is utilized in angle based beamforming to reduce the beamforming overhead in zoned serving space. Three kinds of intra-zone interferences are introduced, and simulation work shows that UEs in the same zone and served with beams generated by the same or adjacent codewords from a predifined codebook suffer from more serious interference, compared with the UEs only affected by non-adjacent beam interference. This fact is considered in the UE selection algorithms design in the next chapter.

The rest of this chapter is organized as follows. In Section 3.2, the multi-zone serving space, the signal propagation model and the downlink receiving signal model are introduced. After that, the multi-pattern codebooks designed for angle based beamforming in different zones are detailed in Section 3.3. In the Section 3.4, the simulation result is presented and discussed. Finally, the conclusion is given in Section 3.5.

3.2   System Model

3.2.1   Multi-zone Serving Space

In this section, we consider a single-cell multi-user massive MIMO system where a BS is located at the cell center. It is assumed that all UEs retain their moving speed without changes on their directions during one beam serving period. The angular speed of UEs is negatively

correlated to the distance between UEs and the BS, i.e. for a fixed linear speed, the angular speeds of far UEs are smaller than that of near UEs. In order to control the serving time of beams, the serving area is assumed to be a circle and is divided into Y zones with respective ranges, was shown in Fig. 3.1. From this figure, we can see that UEs in different zones have different maximum angle changes θ under the same moved distance d_0. Therefore, the beams of different zones have different patterns, which represented by beamwidth and power, and requires different numbers of antennas to generate.

The system model is shown in Fig. 3.2. The antennas of BS is divided into Y subarrays. In each zone, a ULA of $N_l \in [N_1, N_2 \cdots, N_Y]$ antennas at BS is used to serve $K_l \in [K_1, K_2 \cdots, K_Y]$ single-antenna UEs. As shown in Fig. 3.2, each RF chain is able to transmit from all subarrays [41]. We assume that there are K_c UEs in the whole serving area, where $K_c \geq K_1 + \cdots + K_Y$.
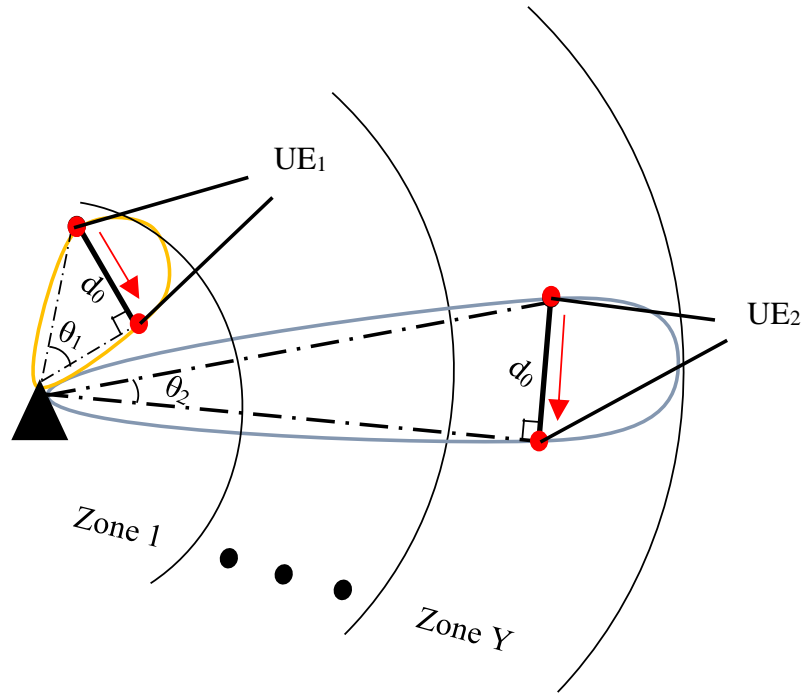
Figure 3.1: Beams with different patterns for UEs in two zones.
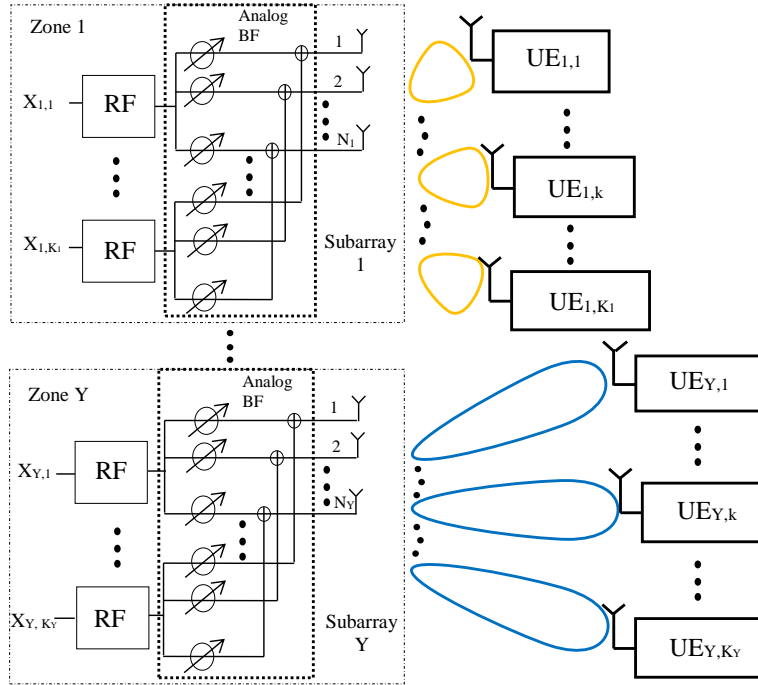
Figure 3.2: BS model with multi-antennas subarrays.

3.2.2 Propagation Model

In this section, with the serving space divided into multi-zones, we have to consider the distance influence of UEs to the signal model. As detailed in chapter 2, after the equation simplification, the CI free-space reference distance model written in 3GPP/ITU for signal attenuation can be expressed as

$$PL(f, d) = 10n\log_{10}\left(\frac{d}{d_0}\right) + 20\log_{10}(f) + \chi_\sigma + C, \tag{3.1}$$

where $d_0 = 1$ is the reference distance chosen in the CI close space environment, this choice of d_0 can provide high parameter stability and model accuracy. d $(d > d_0)$ is the distance between transmitter and receiver, its range can be larger than 500 meters in UMa environments. When reference distance d_0 is set, C is a constant in the model and $C = 20\log_{10}\left(\frac{4\pi d_0 \times 10^9}{c}\right) = 32.4dB$. n denotes the PLE, χ_σ is the zero-mean Gaussian random variable in decibels with a standard deviation σ, which represents the large-scale signal fluctuations about the mean path loss over distance and frequency. As the result, the channel gain of a receiving signal can be simply calculated by the Tx-Rx distance d when the carrier frequency f is selected.

3.2.3 Downlink Signal Model

We assume a multipath channel system in mmWave band. The space between two adjacent antenna elements of a subarray Δd is assumed to be half of the wavelength λ_0, thus there exists a high correlation between antennas. Also, the space between adjacent subarrays is assumed to be much larger than the wavelength, where independent scatters are presented for each subarray. As a result, the channels among different subarrays are independent. To make use of all subarray antennas, the fully-connected structure is chosen [42], where the transmitted signals are combined before transmission. A specific zone is chosen for each UE based on its distance before transmission, UEs in the same zone are served by the same subarray and frequency carrier. Since those for UEs in different zones are different, only intra-zone interference exists in the system. The downlink channel vector for the UE k ($k \in [1, \cdots, K_l]$) in l^{th} ($l \in [1, \cdots, Y]$) zone is shown as follows [43]

$$\mathbf{h}_{l,k} = \sqrt{\frac{N_l}{N_{path}}}\gamma_{l,k,1}\alpha(\theta_{l,k,1}) + \sqrt{\frac{N_l}{N_{path}}}\sum_{n=2}^{N_{path}}\gamma_{l,k,n}\alpha(\theta_{l,k,n}), \tag{3.2}$$

where N_l is the number of antennas of the lth subarray. N_{path} represents the number of paths, $\gamma_{l,k,n}$ denotes the channel gain for each path, $\gamma_{l,k,n} = \sqrt{PL_{l,k,n}(f,d)}$. $\alpha(\theta_{l,k,n})$ is the normalized array steering vector, $\theta_{l,k,n}$ denotes the AoD of nth ray, among which the $\theta_{l,k,1}$ denotes the AoD of the LoS path. $\alpha(\theta_{l,k,n})$ can be expressed as

$$\alpha(\theta_{l,k,n}) = \frac{1}{\sqrt{N_l}}[1, e^{j\pi \sin\theta_{l,k,n}}, \cdots, e^{j\pi(N_l-1)\sin\theta_{l,k,n}}]^T. \tag{3.3}$$

In each zone, more than one UEs are served using the same time/frequency resource. Therefore, the received signal of the kth UE in l^{th} zone includes the useful signal and the intra-zone interference, which can be expressed as

$$y_{l,k} = \sqrt{p_{l,k}}\mathbf{h}_{l,k}\mathbf{f}_{l,k}x_{l,k} + \sum_{m=1,m\neq k}^{K_l}\sqrt{p_{l,m}}\mathbf{h}_{l,k}\mathbf{f}_{l,m}x_{l,m} + n_{l,k}, \tag{3.4}$$

where the $p_{l,k}$ and $x_{l,k}$ represent the transmitting power and the transmitted signal to the kth UE, respectively. The transmit power is uniformly distributed to the K UEs, which means

$P = \sum\limits_{l=1}^{Y} \sum\limits_{k=1}^{K_l} p_{l,k}$. $\mathbf{f}_{l,k}$ is the kth vector of analog beamforming matrix $\mathbf{F}_l = [\mathbf{f}_{l,1}, \mathbf{f}_{l,2}, \ldots, \mathbf{f}_{l,K_l}]$. $n_{l,k}$ is the noise of the kth UE with mean zero and variance σ_n^2. The signal to interference and noise ratio (SINR) of the UE can be expressed as

$$SINR_{l,k} = \frac{\left| \sqrt{p_{l,k}} \mathbf{h}_{l,k} \mathbf{f}_{l,k} \right|^2}{\sum\limits_{m=1,m\neq k}^{K_l} \left| \sqrt{p_{l,m}} \mathbf{h}_{l,k} \mathbf{f}_{l,m} \right|^2 + n_{l,k}}, \tag{3.5}$$

and the SR of selected UEs can be expressed as follows

$$SR = \sum_{l=1}^{Y} \sum_{k=1}^{K_l} \log_2 \left(1 + SINR_{l,k} \right). \tag{3.6}$$

3.3 Multi-Pattern Codebook Design

3.3.1 Spatial Frequency based Multi-Pattern Codebook Design

As mentioned before, because of the same linear speed, the distance between a UE and BS is negatively correlated to angular speed and positively correlated to the path loss [37]. It means that the codebooks for different zones should be able to provide beams with different beamwidth and maximum power to promote UE fairness among UEs in all zones. While the beams in the same codebook, which serve UEs in a specific zone, should be able to provide small intra-zone interference service. Based on these requirements, in this section, a spatial frequency based orthogonal codebook design is introduced [44].

3.3.1.1 Codebook Design for the Same Zone UEs

The definition for spatial frequency scaling factor [44] in the zone l is $\Delta v_l = 1/(N_l \Delta d)$. The spatial frequency index with respect to the angle $\theta_{l,k,1}$, is defined as

$$a_{l,k} = \frac{\sin \theta_{l,k,1}}{\lambda_0 \Delta v_l} = \frac{N_l \sin \theta_{l,k,1}}{2}. \tag{3.7}$$

From Eq. (3.7), we can find that the spatial frequency indexes in $[-N_l/2, N_l/2]$ can be totally mapped to the angle domain of $[-\pi/2, \pi/2]$. Therefore, the array steering vector in Eq. (3.3) can be rewritten as

$$\alpha(a_{l,k}) = \frac{1}{\sqrt{N_l}}[1, e^{j2\pi a_{l,k}/N_l}, \cdots, e^{j2\pi(N_l-1)a_{l,k}/N_l}]^T. \tag{3.8}$$

The codebooks for the analog beamforming are designed in the same form with array steering vectors. The beams generated based on one codebook should satisfy two requirements

1. Beams generated based on codebook for zone l are uniformly distributed in the spatial frequency domain.

2. Low intra-zone interference.

The first requirement means that there should always exists one suitable beam to serve UEs with any index in the spatial frequency domain, hence the beams can serve all directions UEs in angle domain. Besides, there should exist solutions to avoid the high intra-zone interference caused by the simultaneous usage of beams generated by one codebook.

To meet the two requirements, the spatial frequency based codebook for zone l is designed as follows

$$\mathbf{f}_{l,m} = \frac{1}{\sqrt{N_l}}[1, e^{j2\pi c_{l,m}/N_l}, \cdots, e^{j2\pi(N_l-1)c_{l,m}/N_l}]^T, \tag{3.9}$$

$$c_{l,m} = \begin{cases} \pm0.5, \pm1.5, \cdots, \pm\frac{N_l-1}{2}, & if \quad N_l \bmod 2 = 0 \\ 0, \pm1, \cdots, \pm\frac{N_l-1}{2}, & if \quad N_l \bmod 2 = 1, \end{cases} \tag{3.10}$$

where $m \in [1, N_l]$. Therefore, $c_{l,m} = m - (N_l + 1)/2$, which means the antenna number of subarray l decides the number of beams generated by the codebook for this zone. In the spatial frequency domain, these beams are uniformly distributed and orthogonal to each other [44].

The power of beams on all values of spatial frequency axis is analyzed to identify the shape of beams. For array steering vector $\alpha(a)$, where $a \in [-N_l/2, N_l/2]$, we assume that the mth beam is chosen to do the beamforming. Therefore, the power pattern of mth beam in l^{th} zone

can be shown as follows [44]

$$
\begin{aligned}
\beta(c_{l,m} - a) &= \left| \alpha(a)^H \mathbf{f}_{l,m} \right|^2 \\
&= \left| \frac{1}{N_l} \sum_{n_l=0}^{N_l-1} e^{\frac{j2\pi n_l}{N_l}(c_{l,m}-a)} \right|^2 \\
&= \left| \frac{e^{j\pi \frac{(N_l-1)}{N_l}(c_{l,m}-a)}}{N_l} \times \frac{\sin(\pi(c_{l,m}-a))}{\sin(\pi \frac{(c_{l,m}-a)}{N_l})} \right|^2 \\
&= \frac{\sin^2(\pi(c_{l,m}-a))}{N_l^2 \sin^2(\pi \frac{(c_{l,m}-a)}{N_l})},
\end{aligned}
\tag{3.11}
$$

from Eq. (3.11), it can be seen that $\beta(c_{l,m} - a)$ is a periodical function with period N_l. Also, the range of variable $(c_{l,m} - a)$ for the mth beam in one codebook is $[-N_l + m - 1/2, m - 1/2]$. Therefore, the beams in one codebook are cyclic shift to each other in spatial frequency domain. The Fig. 3.3 gives an example for beam pattern in l zone with $N_l = 8$. In Fig. 3.3, the red line, obtained by connecting all crossed point of main lobes, represents the minimum value of power can be provided by beams. It shows that there always exists a suitable beam for a UE with power larger than the red line value. When the serving beam is selected, the beams generated by the remaining codewords in the codebook can be divided as the adjacent beams and non-adjacent beams. The distance on spatial frequency axis between two adjacent crossed points is regarded as the beamwidth of a beam, so all the beams in one codebook have the same beamwidth with regard to this axis.

The distance between the main lobes center of two adjacent beams is

$$
c_{l,m+1} - c_{l,m} = 1, m \in [1, N_l - 1],
\tag{3.12}
$$

the half range of main lobe is decided by the nearest zero point a_0 to the $c_{l,m}$, which is

$$
\begin{aligned}
a_0 &= \arg \min_{a} \left| c_{l,m+1} - a \right| \\
s.t. \quad &\beta\left(c_{l,m+1} - a\right) = 0 \\
&a \in [-N_l/2, N_l/2] \\
&a \neq c_{l,m+1}.
\end{aligned}
\tag{3.13}
$$

It is clear that $a_0 = c_{l,m+1} \pm 1$, hence the half range of a beam main lobe on spatial frequency
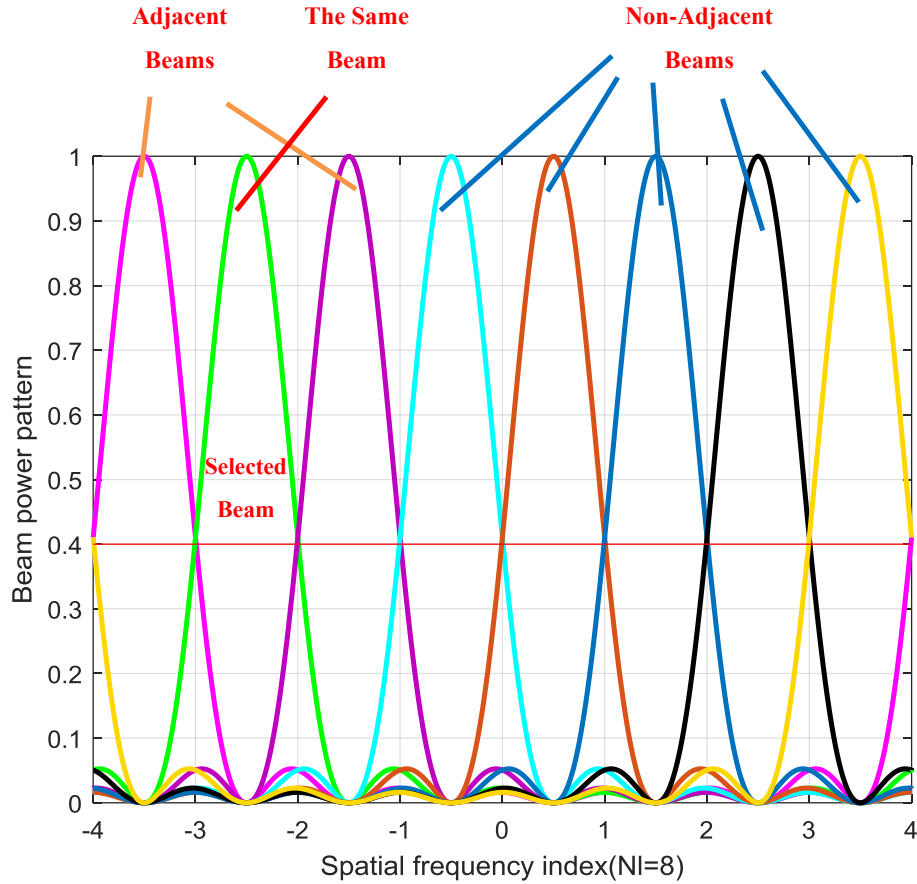
Figure 3.3: An example for beams pattern in one codebook.

axis is equal to the distance between two main lobes centres. It means that the adjacent beams always have overlap on their main lobes, which brings high interference to both of them. Therefore, adjacent beams should not be chosen together in order to prevent performance reduction caused by high intra-zone interference. To be mentioned that, the first and the last beams generated by a codebook should be viewed as adjacent beams because of their cyclic shift feature.

3.3.1.2 Codebook Design for UEs in Different Zones

Because of the angular speed and path loss of UEs vary among zones, codebooks with different patterns should be designed to adapt UEs in different zones. The beam pattern is represented by the beamwidth and beam power. As for the beamwidth, due to the uniform distribution of

beams in spatial frequency domain, the average beamwidth is negatively correlated to N_l, so the codebook with large N_l is suitable for the UEs in outer zones with low angular speed. As for the beam power, it is assumed that $\mathbf{h}$, γ and d represent the channel vector, channel gain at the spatial frequency index a and the LoS distance from the BS to the detecting point, respectively. The power of beam can be shown as

$$P_{beam} = \left\| \mathbf{h}^H \mathbf{f}_{l,m} \right\|^2 = N_l \gamma^2 \left\| \alpha(a)^H \mathbf{f}_{l,m} \right\|^2 \propto \frac{N_l}{d}. \tag{3.14}$$

We can find that the beam power increases with N_l. As a result, the subarrays with large numbers of N_l should be allocated to generate codebooks for outer zones.

3.4 Performance Evaluation

3.4.1 System Parameters

The parameters for simulation work are shown in Table 3.1 [45]. The minimum number of subarray antennas is 4, which is the minimum number of antennas to serve more than one UEs without using adjacent beams. We assume that the UEs are evenly distributed in the serving area. Therefore, the UE density $\rho = K_c / S_{sum}$ can be viewed as a fixed value in the serving area S_{sum}. The range of one zone should be decided based on the UE capacity of the zone, and the way used to decide the edge of zones, is shown as follows

$$
\begin{aligned}
\frac{K_l}{K_c} &= \frac{C_l}{C_{sum}} \Rightarrow \frac{\rho S_l}{\rho S_{sum}} = \frac{N_l}{N_{sum}} \\
&\Rightarrow \frac{\left(r_l^2 - r_{l-1}^2 \right)}{R^2} = \frac{N_l}{N_{sum}} \\
&\Rightarrow r_l = \sqrt{\frac{N_l R^2}{N_{sum}} + r_{l-1}^2},
\end{aligned}
\tag{3.15}
$$

where C_{sum} and C_l represent the sum UE capacity and UE capacity for zone l, respectively. $l \in [1, 2, \cdots, L]$, L is the number of zones. For each zone, the UE capacity is represented by the number of subarray antennas N_l. N_{sum} is the sum antennas number of all subarrays, i.e.

$N_{sum} = N_1 + N_2 + \cdots + N_L$. K_l and K_c represent the number of UEs in zone l and the sum number of candidate UEs, which are positively correlated to the coverage area of zones. Therefore, the edge of zone l, represented by r_l, can be decided by N_l and r_{l-1}. When the distance from a UE to the BS is too small, sharp change on AoA may happen frequently. Besides, r_0 should be set to be larger than d_0. Therefore, we set $r_0 = 1$.

Table 3.1: Simulation Parameters for Beamforming Design in TDD Systems

Cell radius R (m)	200
Range of AoD/AoA	$(-\pi/2, \pi/2)$
Carrier frequency f (GHz)	38
System bandwidth B (MHz)	800
Noise density n_0 (dBm/Hz)	-174
Duplex mode	TDD
Number of zones Y	6
BS antennas N_l	4, 8, 16, 32, 64, 128 ULA
Edge of zones $r(m)$	25, 44, 67, 98, 140, 200
UE antennas	1
Path loss exponent	1.9 for LoS, 2.7 for NLoS
Large scale signal standard deviation σ (dB)	3.5 for LoS, 10.5 for NLoS
Number of served UEs K_s	50

3.4.2 Numerical Results

In this section, the simulation work compares the SR of UEs in three situations with all three kinds of interference (same beam, adjacent beam and non-adjacent beam interference), two kinds of interference (adjacent beam and non-adjacent beam interference) and one kind of interference (only non-adjacent beam interference). In the simulation work in both this chapter and next chapter, the UEs served at the same time are only affected by intra-zone interference, while the UEs in different zones utilize different frequency resource, and no inter-zone interference exists. From Fig. 3.4, it can be found that while the same numbers of UEs are served, the same beam interference have the most significant influence to bring the UEs data rate reduction. The interference power from adjacent beams to a served UE is much lower than that caused by the same beam reuse. However, when both the same beam interference and adjacent beam interference, which are high power interference compared with non-adjacent

beam interference, do not exist in the downlink data transmission, the UEs SR can be highly improved. Therefore, the UEs served at the same time should be selected to avoid the high power interference.
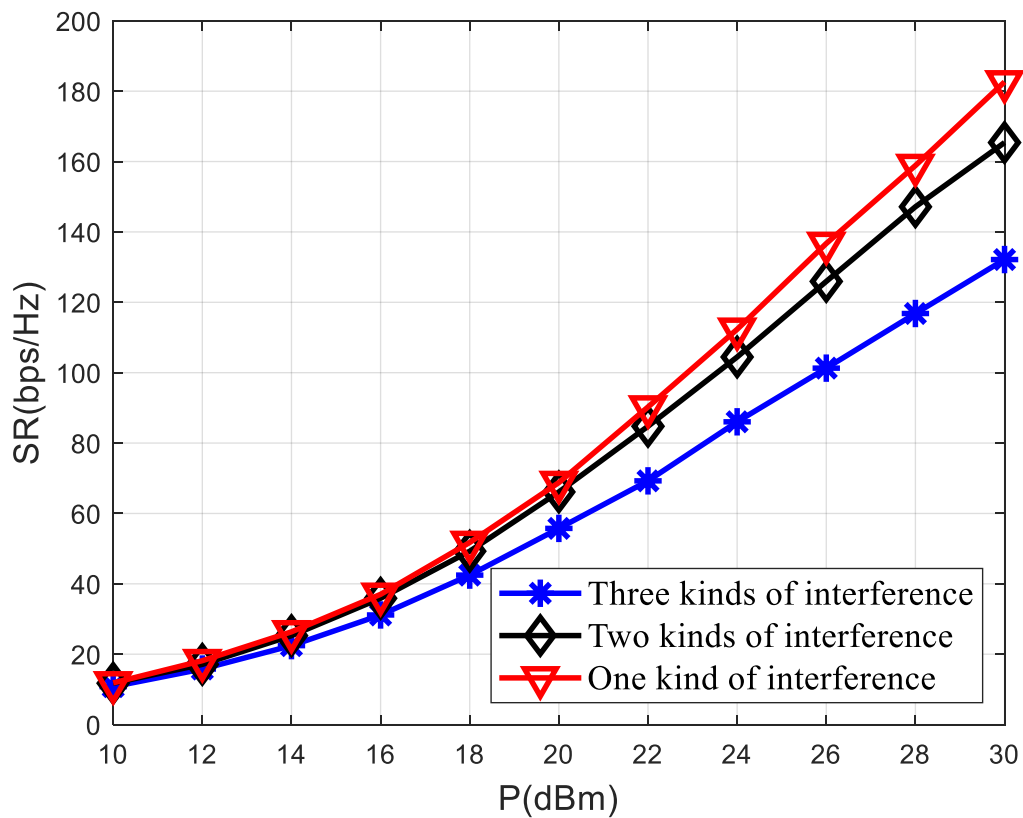
Figure 3.4: The sum data rate comparison for cases with different kinds interference.

3.5 Chapter Summary

This chapter introduces the multi-pattern codebooks design in order to adapt to the beamforming requirements in the multi-zone environment. In one specific zone, the codewords providing orthogonal beam patterns in spatial frequency domain are designed to prevent intra-zone interference, while in different zones, codebooks with changing patterns can be selected to generate beams with suitable beamwidth and power. The downlink transmission signals to UEs in different zones are divided in frequency domain, while the spatial domain based interference reduction should be achieved by beamforming for UEs in the same zone. It is proved in the

simulation work that for the served UE of the same zone, when the interference from the reuse of the same codeword and the interference caused by forming adjacent beams can be avoided, the average SR can witness a significant increase. Such a conclusion puts forward requirements for the served UEs selection at the same time. In the next chapter, different UE selection methods will be proposed to satisfy systems with different requirements.

Chapter 4

UE Selection Designs for Low Overhead Beamforming in TDD mmWave MIMO System

With the multi-pattern codebook based beamforming in zoned serving space designed in the last chapter, in this chapter, two kinds of UE selection algorithms are proposed, which are ultra low overhead oriented algorithm and large connection oriented algorithm, respectively. With advantages in different aspects, two UE selection algorithms make the proposed low overhead beamforming adaptive to different use cases.

4.1   Introduction

In order to meet the high data rate expectation of 5G wireless systems, mmWave communications become critical because of their large available bandwidth [20]. The short wavelength of mmWave communication allows plenty of antennas to be packed into a limited space, which enables the feasibility of the massive MIMO [31]. By exploiting the large scale antenna array, narrow beams can be generated to fully utilize the spatial domain resources and enhance the overall capacity. However, to fully exploit the use of beamforming, accurate CSI is required at the transmitter side [46] resulting in a large overhead for training and feedback information in FDD mmWave massive MIMO systems [47] [48]. As mentioned in the last chapter, due to the reduced coherence time associated with high carrier frequencies [49], the CSI should be estimated more frequently to keep pace with the channel variation, which also leads a larger

overhead for beamforming.

The overhead caused by channel estimation for beamforming can be divided into the downlink training overhead and the uplink feedback overhead. A training preamble is utilized to obtain the CSI, which is measured by the downlink channel vector at the user side [50]. Many effective methods are proposed to reduce the training overhead in [47], [35], [51] and [52], and the weighted sum rate is defined to evaluate the performance of the system with overhead due to the downlink training of the beamforming [52]. In order to reduce the uplink feedback overhead proportion in the data transmission, two approaches are considered in recent studies. One method is to reduce the number of bits in one feedback period directly [53], [54], [55]. Generally, the number of bits in the feedback overhead is proportional to the number of antenna elements [54]. To reduce the influence of the large scale antenna array, an antenna grouping based feedback reduction technique is designed for FDD-based massive MIMO systems in paper [53]. In paper [55], a low feedback beamforming scheme is proposed to mitigate the severe interference in diverse mobility cases. In particular, for current emerging research on the angle-based beamforming [39], the feedback overhead can be reduced because UEs only need to report their angle information instead of CSI. The angle quantization-based schemes are able to make a further reduction on the feedback overhead [54]. Therefore, angle-based beamforming is considered in our work.

The other approach is to reduce the frequency of feedback in a specific period. In [56], the feedback is divided into a long-term part and a short-term part to transmit the location related information and the fast-fading component, respectively. The feedback frequency for the long-term part can be relatively low. In angle-based beamforming, the beam generation is decided by the AoD from the BS. Compared with the change of CSI in the mmWave band, the AoD changes much slower. A concept called beam coherence time is defined to evaluate the stability of AoD in [57], which is much longer than the channel coherence time. Therefore, angle-based beamforming only needs to feedback AoD in every beam coherence time. In the previous work, both beam coherence time, channel coherence and beamwidth design are decided by the linear speed of user when the carrier frequency is fixed. However, UEs moving with the same linear speed may have different angular speeds due to their changing distance from the transmitter to the receiver. Hence, compared with the linear speed, angular speed, which represents the

speed of angular variation, is much more suitable to be considered in beamwidth design for environments with a large radius of serving space. Nevertheless, few of previous works takes angular speed into consideration.

Based on the beamforming in zoned serving space with multi-pattern codebooks design introduced in the last chapter, and the influence of different types of intra-zone interference shown in the simulation work before, in this chapter, two UE selection algorithms are proposed to satisfy different use cases requirements, which are ultra low feedback and large UE connection, respectively. For both of the algorithms, each UE is required to feedback signal with one bit information for LoS path AoA and ToA estimation initially, based on the AoA and ToA information, UEs without the same or adjacent beam interference are selected. After that, the ultra low feedback oriented algorithm selects UEs that minimum beam serving time estimated can meet the requirement of a threshold to be served for a period, a new turn of UE selection will start after that period. While for the large connection oriented algorithm, these UEs are served directly without further selection, after each time interval predefined, UEs moving out of the selected beam serving range are disconnected according to their one bit feedback for codeword reselection, and a new UE selection will start when the required serving period is satisfied or the connection ratio is lower than a predefined threshold.

As shown in the simulation work, for UEs with the same moving speed, the large connection oriented UE selection algorithm always outperforms the ultra low feedback oriented UE selection algoithm with regard to the percentage of connected UEs simultaneously and the SR of UEs, while for the cumulative feedback calculated in a period of time, the reverse is true. Therefore, two algorithms can be utilized to adapt to different use cases.

The rest of this chapter is organized as follows. In Section 4.2, the way to obtain the location information (AoA and ToA) of UEs is firstly introduced, then the procedures of two UE selection algorithms with the aid of location information are detailed. In Section 4.3, the simulation results are presented and discussed. Finally, the conclusions are given in Section 4.4.

4.2 Location-Aided UEs Selection Methods Design

4.2.1 Location Estimation based on AoA and ToA Measurement

The location information of UEs can be determined by their AoA and distance to BS, both of them can be detected by one training bit of uplink signal. In [58], the AoA of the LoS path from a UE can be detected by MUSIC method. As mentioned in [15], the high-resolution MUSIC algorithm can only be achieved with the premises of large numbers of snapshots and high SNR. When the transmitting power for UEs is not high enough, plenty of snapshots are required to reduce the influence of noise. The MUSIC based AoA estimation resolution with different numbers snapshots in changing transmitting power will be discussed in the simulation part.

In addition, the distance from a UE to the BS can be detected by estimating ToA [59]. Therefore, according to the estimation of AoA and ToA, UEs can be divided into different zones and served by suitable beams.

4.2.2 UE Selection Methods for Different Use Cases

4.2.2.1 Ultra Low Feedback Oriented UEs Selection

To serve UEs with low feedbacks while avoiding the adjacent interference, in this section, we propose a ultra low feedback oriented UE selection algorithm. The algorithm can be divided into four steps:

1. Allocate UEs to different zones according to $\mathbf{d}$ and record their minimum time $\mathbf{t}_{edge}$ moving from current locations to the edges of zones;

2. Select codewords for UEs by their AoA θ at the BS, calculate their real minimum serving period $\mathbf{t}_{min}$;

3. Sort UEs who satisfy the minimum serving period threshold T requirement according to $\mathbf{t}_{min}$;

4. Choose UEs without the same beam or adjacent beam interference.

Algorithm 1 Low Feedback Oriented UEs Selection Algorithm

Require: K_c: number of candidate UEs; **d**: distance of UEs to the BS; θ: AoA of UEs at BS in spatial frequency domain; Y: number of zones; **L**: range of zones; **N**: number of antennas for each zone; vt: UE speed; T: threshold for serving period;

Ensure: $\mathbf{t}_{min}$: real minimum serving period for selected UEs; **c**: spatial frequency index of codewords for selected UEs ; $K(l)$: selected UEs in zone l;

1: **Initialize** $K(1:Y), \Delta K(1:Y), \mathbf{t}_{edge}$ as null;
2: **Step 1:** Allocate UEs to different zones, record number of UEs in each zone with their edge time and distances.
3: **for** $k = 1$ to K_c; $l = 1$ to Y **do**
4:      **if** $\mathbf{d}(k) \in \mathbf{L}(l)$ **then**
5:          $K(l) = K(l) + 1$;
6:          Update $\mathbf{t}_{edge}, \mathbf{d}$;
7:      **end if**
8: **end for**
9: **Step 2:** Calculate the minimum time of UEs serving by one beam, record beam indexes.
10: **for** $l = 1$ to Y; $k = 1$ to $K(l)$ **do**
11:      $\mathbf{c}(k,l) = \underset{\mathbf{c}(k,l)\in\mathbf{F}(l)}{argmin} (\mathbf{c}(k,l) - \theta(k,l))$
12:      $\textbf{left}(k,l) = |angle(\mathbf{c}(k,l) - 0.5) - angle(\theta(k,l))|$;
13:      $\textbf{right}(k,l) = |angle(\mathbf{c}(k,l) + 0.5) - angle(\theta(k,l))|$;
14:      $\mathbf{t}_{min}(k,l) = min(\textbf{left}(k,l)/(vt*\mathbf{d}(l,k)),$
15:             $\textbf{right}(k,l)/(vt*\mathbf{d}(l,k)), \mathbf{t}_{edge}(k,l))$;
16: **end for**
17: **Step 3:** Select UEs satisfying serving period requirement, sort by the real minimum serving period from long to short.
18: $\Delta K(1:Y) = 0$;
19: **for** $l = 1$ to Y; $k = 1$ to $K(l)$ **do**
20:      **if** $\mathbf{t}_{min}(k,l) < T$ **then**
21:          $\Delta K(l) = \Delta K(l) + 1$;
22:          delete $\mathbf{t}_{min}(k,l), \mathbf{c}(k,l)$;
23:      **end if**
24:      **sort**$(\mathbf{t}_{min}(:,,l), downorder)$
25:      **Update c**;
26: **end for**
27: $K(1:Y) = K(1:Y) - \Delta K(1:Y)$;
28: **Step 4:** Disconnect with the UEs bring the same or adjacent beam interference.
29: $\Delta K(1:Y) = 0$;
30: **for** $l = 1$ to Y; $k = 2$ to $K(l)$ **do**
31:      $mmin = min(|\mathbf{c}(k,l) - \mathbf{c}(1:k,l)|)$;
32:      $mmax = max(|\mathbf{c}(k,l) - \mathbf{c}(1:k,l)|)$;
33:      **if** $|mmin| < 2 || |mmax| > \mathbf{N}(l) - 1$ **then**
34:          $\Delta K(l) = \Delta K(l) + 1$;
35:          **Delete** $\mathbf{t}_{min}(k,l)$ and $\mathbf{c}(k,l)$
36:      **end if**
37: **end for**
38: $K(1:Y) = K(1:Y) - \Delta K(1:Y)$;
39:

Here **d** is a vector represents distances of LoS paths from UEs to the BS, which is calculated by $\mathbf{d} = \mathbf{T}_{LoS} \times c$, where $\mathbf{T}_{LoS}$ is the ToA vector of LoS paths for candidate UEs, c is the speed of light. The real minimum serving period array, $\mathbf{t}_{min}$, means the time of UEs spending on getting out of the beam serving space in the shortest distance. Due to the $\mathbf{t}_{min}$ of the selected UEs are all longer than or equal to the threshold T, therefore, the threshold T can be utilized as the period for uplink feedback. The algorithm is detailed in Algorithm 1. To be mentioned that, the function *angle* in step 2 is shown as follows

$$angle(\mathbf{c}(k, l)) = arcsin(2\mathbf{c}(k, l)/N(l)). \tag{4.1}$$

It can be found that the UEs selected by this algorithm are the ones with longest $\mathbf{t}_{min}$ and no the same beam or adjacent beam interference.

4.2.2.2 Large Connection Oriented UEs Selection

To adapt to the high connection use case [60] in the real environment while keeping the long serving time advantage, a large UE connection oriented algorithm is proposed to select UEs. This algorithm can also be applied by four steps:

1. Allocate UEs to their corresponding zones according to **d**, and for the kth UE in zone l, find the serving beam index $c_{l,k}$ from the codebook l based on $\theta_{l,k}$;

2. Choose UEs without the same beam or adjacent beam interference as the initialized UEs;

3. At the end of the predefined interval Δt, UEs who move out of their corresponding beam serving range are found and disconnected based on the uplink feedback;

4. Use a threshold of connection ratio *Per* or serving period T to decide the time for a new serving circle.

In order to know whether the UEs which have not moved out of their original zones still in the serving range of original beams or not, a matrix consisted of three codeword vectors is designed in Eq. (4.2),

$$\mathbf{C}_{l,k} = [\mathbf{f}(\tilde{v}_1, N_l), \mathbf{f}(\tilde{v}_2, N_l), \mathbf{f}(\tilde{v}_3, N_l)], \tag{4.2}$$

$$\mathbf{f}(v, N_l) = \frac{1}{N_l}\left[1, e^{j2\pi v/N_l}, \cdots, e^{j2\pi(N_l-1)v/N_l}\right]^T,$$ (4.3)

where $\tilde{v}_1 = c_{l,k} - 1$, $\tilde{v}_2 = c_{l,k}$ and $\tilde{v}_3 = c_{l,k} + 1$, which means the matrix includes the codewords of the serving beam and its adjacent beams. Due to the reciprocal channel in TDD systems, such characteristic can be utilized to specify the location of UEs. When a short serving interval Δt is finished, only slight changes take place at the location of served UEs. Therefore, the kth served UE in zone l feedbacks one bit signal to the BS, which is used for LoS path ToA detection, and beamformed by the already selected codeword vector and corresponding adjacent codeword vectors contained in $\mathbf{C}_{l,k}$, in order to replace the time-comsuming AoA detection step with large number of signal snapshots requirement. The new suitable codeword index $\tilde{v}_{n_c}$ for the UE with channel vector $\mathbf{h}_{l,k}$ can be calculated by Eq. (4.4) [61],

$$n_c = \arg\max_{1 \leqslant j \leqslant 3}\left|\mathbf{f}(\tilde{v}_j, N_l)^H \mathbf{h}_{l,k}\right|.$$ (4.4)

At the end of each interval Δt, if $\tilde{v}_{n_c} = c_{l,k}$, i.e. the UE is still in the beam serving range, the BS will continue to serve it in the next Δt period. Otherwise, the BS should disconnect the UE. Therefore, within the beam serving time, one bit is used for feedback with a period of Δt. A new UE selection starts when the percentage of serving UEs is lower than the predefined threshold *Per* or the serving time satisfies threshold T.

4.3   Performance Evaluation

The simulation work in this section starts from the comparison of MUSIC based AoA estimation resolution with different numbers of snapshots or transmitting power. After that, Δt, which is a parameter for the large connection oriented UEs selection method is specified for different speeds scenarios. Then the UE selection percentage, the SR and the cumulative feedback are compared between two UE selection algorithms. At the same time, the SR and the cumulative feedback are compared among two proposed algorithms and the multi-sectional codebooks searching algorithm proposed in [28]. This algorithm employs beam training method to select suitable codewords for UEs, it is designed based on hierarchical codebooks introduced in

chapter 2 and can always select codewords from the highest layer of hierarchical codebooks. For each zone, a hierarchical codebook with multiple layers is designed. It can be proved that when the number of antennas is fixed, the highest layer of hierarchical codebook has the same design as the proposed multi-pattern codebook.

Most parameters for simulation work are shown in Table 3.1. What to be mentioned is that the number of candidate UEs K_c for UE selection is 100.

4.3.1 Numerical Results
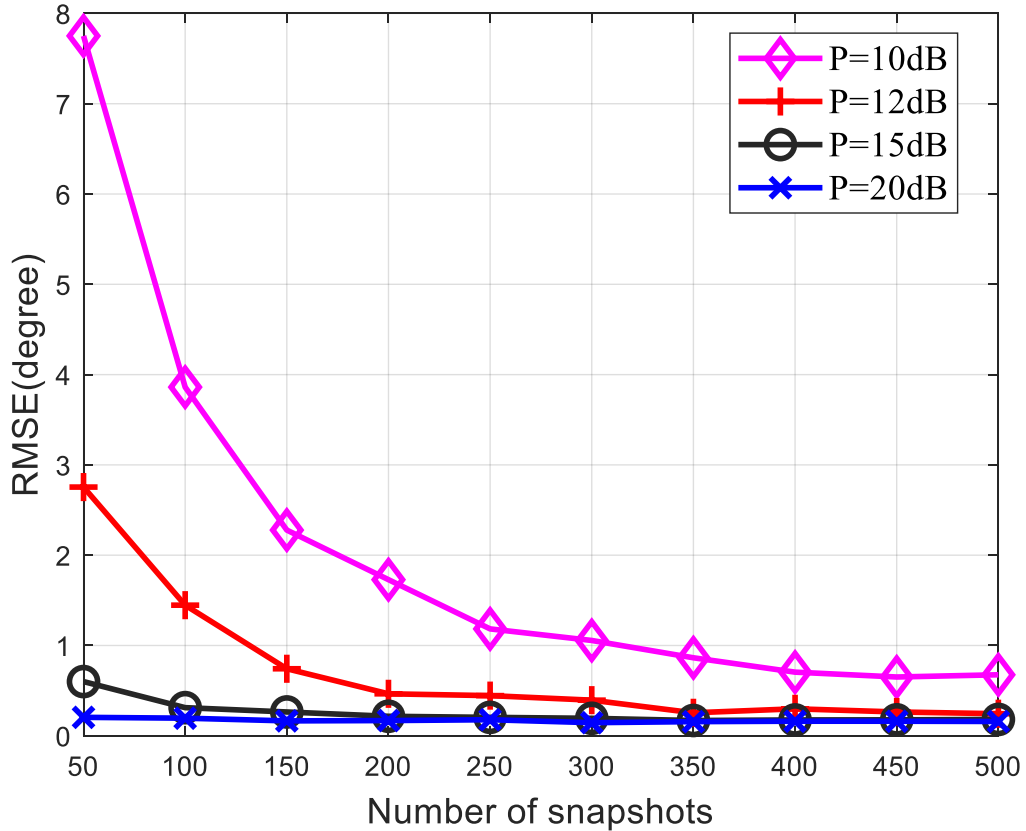
4.3.1.1 MUSIC based AoA Estimation

Figure 4.1: AoA estimation resolution based on different numbers of snapshots with diverse transmitting power.

We use MUSIC method to estimate the AoA of LoS paths based on the uplink signal from UEs to the BS, the signal is reciprocal to the downlink signal defined in Eq. (3.4). The high

resolution AoA should be available for accurate codeword selection. Due to the minimum beamwidth is 0.5 degrees, the average error of AoA estimation should not be lower than this value. In Fig. 4.1, the changing tendency of AoA estimation RMSE with increasing numbers of snapshots are compared among cases with different transmitting power. When the path loss of signal is considered as a fixed value, the receiving power of signal at BS is only decided by the power of transmitting signal. It can be easily found that for all transmitting power cases, the reducing RMSE with increasing snapshots means that the AoA estimation resolution increases while more signal snapshots are available, it is because that the large number of snapshots are helpful to reduce the influence of noise in MUSIC based AoA estimation, although these snapshots lead to the consequence of long AoA estimation time. Compared curves for signals with different transmitting power, the higher the transmitting power is, the fewer snapshots are needed to achieve the required resolution. When the number of snapshots utilized is 500, the signal with relatively low transmitting power can provide high AoA estimation resolution.

4.3.1.2 UEs Connection Ratio

Fig. 4.2 compares the changes on UE connection ratio with different length of Δt while using the large connection oriented UE selection algorithm, in order to find out a Δt which can both updates the suitability of beams for each UE on time, and avoid too much waste of feedback at the same time. We define that when UE connection change ratio is higher than one percent, a feedback should be given. It can be found from Fig. 4.2 that for UEs in speed of $5m/s$, when $\Delta t = 0.04s$, the UE connection change is around one percent, hence $\Delta t = 0.04s$ when UEs speed is $5m/s$. When the UE speed increases, the value of Δt should adaptively decrease to keep the UE connection change ratio. Therefore, for UEs in speeds of $10m/s$ and $20m/s$, Δt drops to $0.02s$ and $0.01s$, respectively. What to be mentioned is that the Δt is always too short to a sharp AoA change on each UE, so the beam training based on three codeword vectors in $\mathbf{C}_{l,k}$ for a UE is enough to detect the suitability of original beam.

In the following simulation performance figures, algorithm 1 represents the ultra low feedback oriented algorithm, while algorithm 2 is the large connection oriented algorithm. T is the threshold of the minimum serving period, which represents a predefined minimum period be-
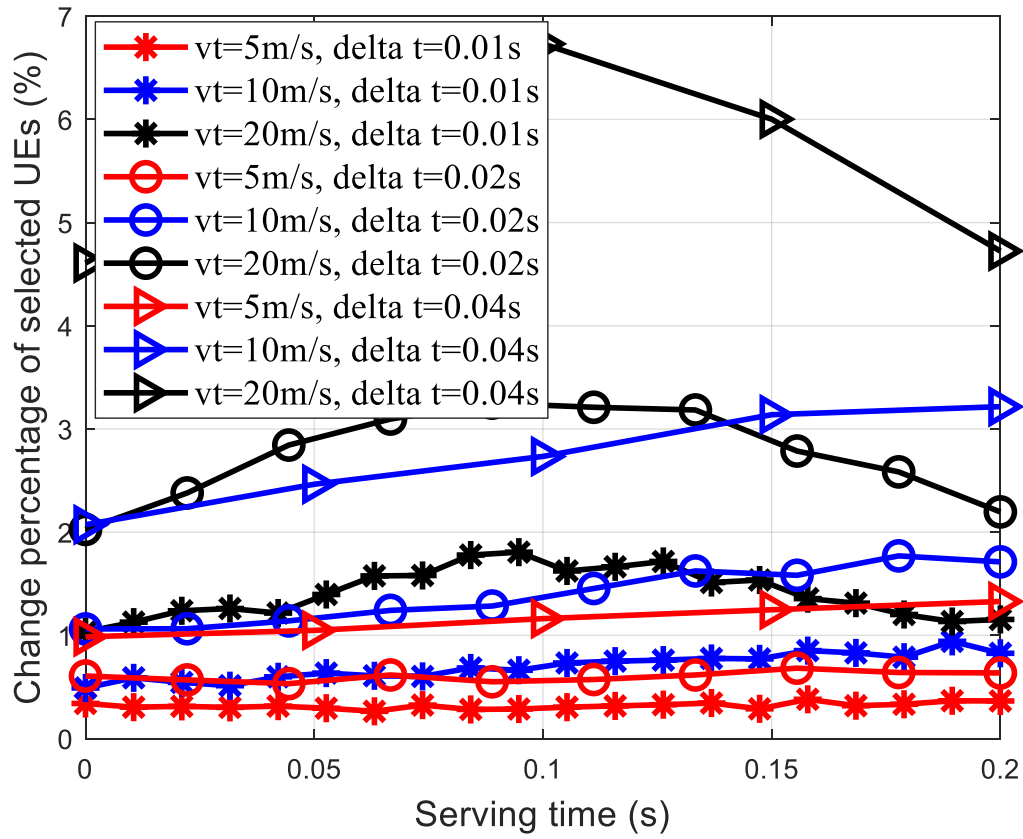
Figure 4.2: Selected UEs ratio change with different speeds and Δt.

tween two times of UE connection for both algorithm 1 and algorithm 2, the UEs are assumed to have fixed moving direction and speed within the beam serving time. The UE connection percentages of the two algorithms with changing threshold T are shown in Fig. 4.3, here the curves for algorithm 2 are consisted of UE connection percentage at the end of period T. For both algorithms, the percentages of connections decrease with the increasing threshold T and UE speed. However, the connection percentage of algorithm 2 is much higher than that provided by algorithm 1 under the same UE speeds and threshold T, which means that algorithm 2 provides a significant increase in the UE connection ratio while maintaining the advantage of a long serving time compared with algorithm 1. Therefore, the high connection ratio is the advantage of algorithm 2, while the low connection ratio is the disadvantage of algorithm 1.
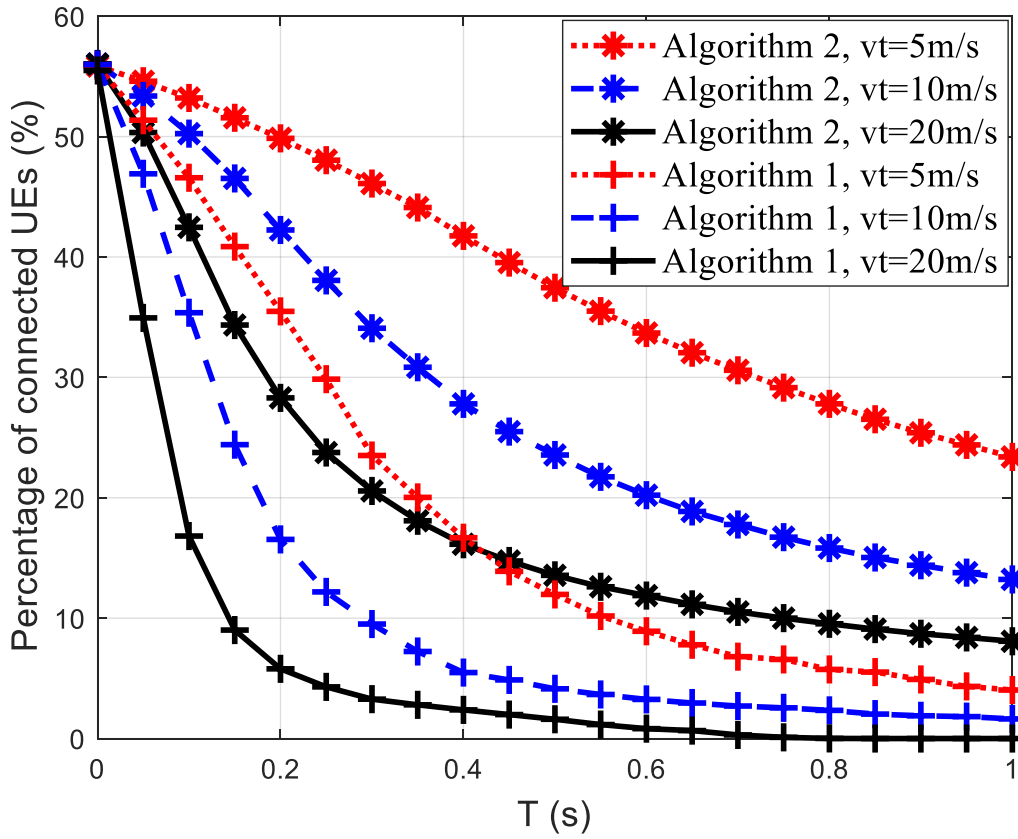
Figure 4.3: The connection ratio comparison between two UE selection algorithms.

4.3.1.3 Sum Data Rate

The sum data rates of algorithm 1, algorithm 2 and the multi-sectional codebooks searching algorithm proposed in [28] are compared in Fig. 4.4 while the speed of UEs is fixed as $5m/s$. As shown in the simulation result, the sum data rates of algorithm 2 always outperform those of algorithm 1 with different T. Such advantage of algorithm 2 is brought by the larger number of UEs served at the same time compared with algorithm 1. On top of that, the number of UEs simultaneously served decreases with the increasing threshold T for both algorithm 1 and algorithm 2, therefore, the sum data rates for both algorithms witness an increase while threshold T is declining. Finally, it should be mentioned that the UEs are reselected every coherence time in the multi-sectional codebooks searching algorithm, and only UEs who would not be affected by the same beam or the adjacent beam interference are selected. Therefore, the high UEs serving ratio can be kept and these UEs can always be served by suitable beams in
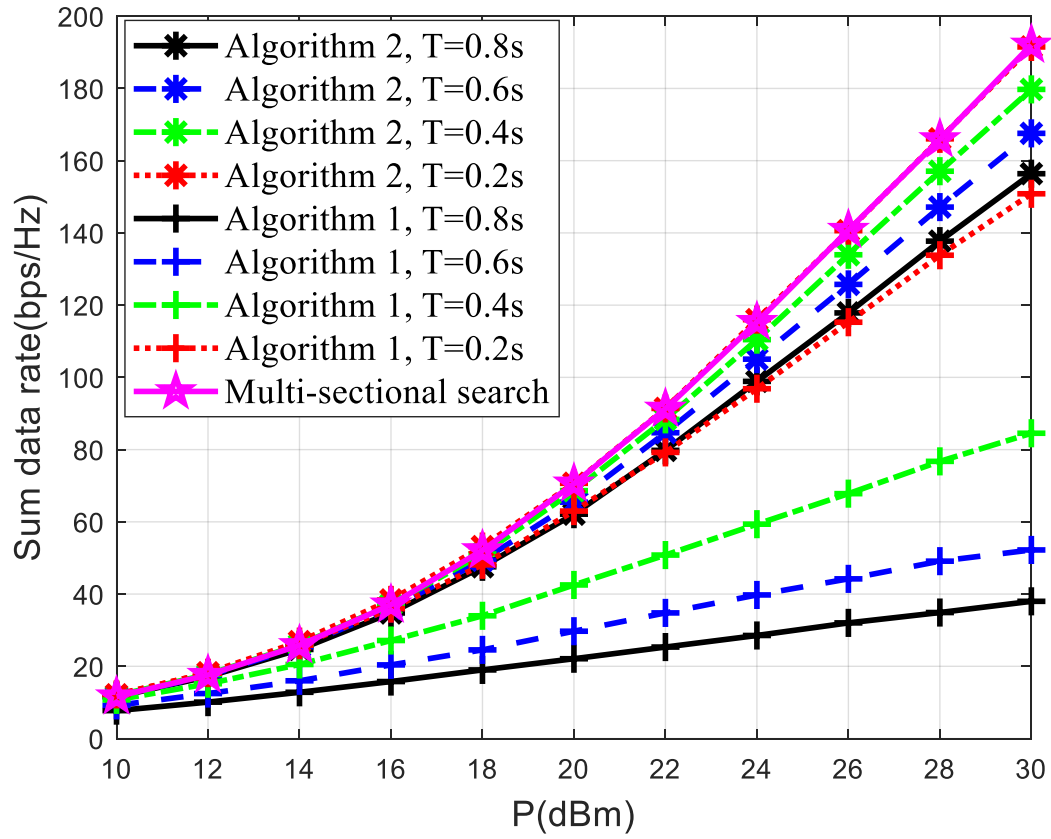
Figure 4.4: The throughput comparison between two UE selection algorithms.

this algorithm, which leads to the high SR. The overlap of the curve of this algorithm and the curve of algorithm 2 when $T = 0.2$ means that algorithm 2 can not only reduce the feedback required, but also provide high SR.

4.3.1.4 Cumulative Feedback

Table 4.1 compares the minimum serving period T with at least 50% UE connection for two algorithms. We can find the 50% connection serving time for algorithm 1 is much shorter than that of algorithm 2, however, the feedback of a UE using algorithm 2 is decided not only by the threshold T, but also the interval of feedback Δt in each period of T. Therefore, the cumulative feedback of both algorithms will be discussed. On top of that, the coherence time mentioned in the table is viewed as the feedback period of multi-sectional codebooks searching algorithm, which will be compared with two algorithms with regard to the cumulative feedback later.

Table 4.1: Codeword Selection Period Comparison

Speed / Algorithm	Algorithm 1 50% connection minimum serving time T(s)	Algorithm 2 50% connection minimum serving time T(s)	Δt for Algorithm 2 (s)	Coherence time (s)
$vt = 5m/s$	0.067	0.2	0.04	0.00063
$vt = 10m/s$	0.032	0.1	0.02	0.00031
$vt = 20m/s$	0.015	0.04	0.01	0.00016

Figure 4.5: The cumulative feedback comparison between two UE selection algorithms.

The cumulative feedback for single UE using algorithm 1, algorithm 2 and the multi-sectional codebooks searching algorit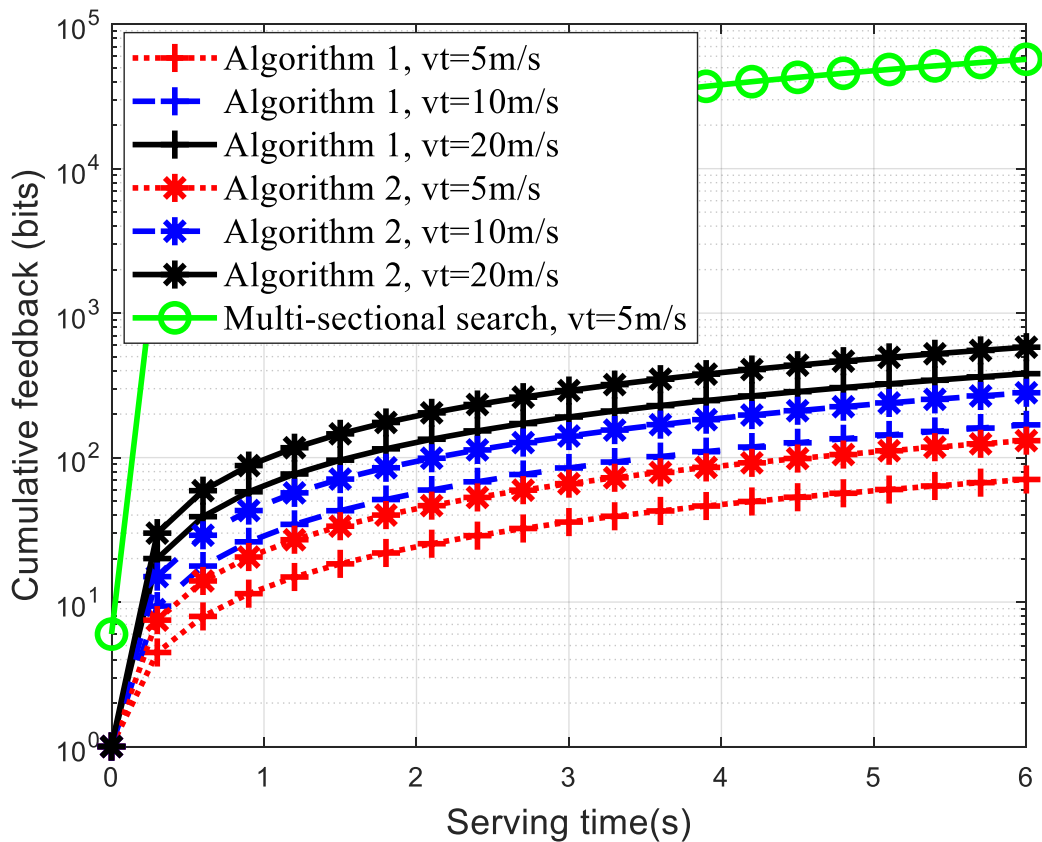hm are compared in Fig. 4.5. It can be found from the figure that for both proposed algorithms, UEs moving faster require more feedback bits in a fixed period of serving time. Compared two algorithms for UEs in same speed, the algorithm 2 sacrifices more bits for the performance improvement on connection percentage and SR. As mentioned before, the feedback period of multi-sectional codebooks searching algorithm is UE coherence time. Besides, for the application in multi-zone scenario, the number of bits required by this algorithm changes with the antennas number. 2 to 7 bits are needed respectively for the cases using 4 to 128 antennas. In each time of codeword selection, based on the distance detected by the first bit feedback, the zone is determined and the number of bits required can be known. The average feedback for each time of codeword selection is 6 bits, while only one bit required for feedback in two proposed algorithms. Therefore, the feedback of both proposed algorithms are much lower than that of the multi-sectional codebooks searching algorithm, which means the aim of the algorithms design is achieved.

4.4 Chapter Summary

In this chapter, based on the multi-pattern based beamforming introduced in the last chapter, two kinds of UE selection algorithms are detailed to make sure that UEs are served with only non-adjacent beam interference, and a relatively long serving period can be ensured. Compared with the large UE connection oriented algorithm, ultra low feedback oriented algorithm sacrifices fewer bits of feedback, while the former algorithm has better performance on SR and the percentage of connected UE at the same time. The simulation work shows that the SR of proposed low overhead beamforming utilizing large connection UE selection is very close to that of the multi-sectional search beamforming algorithm, while fewer numbers of feedback bits are required by the beamforming algorithm with any proposed UE selection methods, compared with the multi-sectional search beamforming algorithm.

Chapter 5

AoA Estimation based Low Overhead Beamforming in FDD mmWave MIMO System

In FDD channel, with the full use of the angle reciprocity, the fast estimated AoA information based on uplink signal can be helpful for downlink codeword selections. When the uplink transmission signal is always available for AoA estimation, the codeword can be adaptively selected to the changing channel, and the overhead for beamforming is avoided while the high downlink data rate is kept. In this chapter, we propose a few uplink signal snapshots based fast AoA estimation algorithm.

5.1 Introduction

With the increasing exploitations of mmWave band, large scale antenna array can be utilized to achieve very narrow beamforming and high power gain for improving the transmitting data rate of future wireless systems [62]. The angle-based beamforming and beam-training based beamforming are two widely used beamforming algorithms for different application scenarios. Thanks to the angle reciprocity between uplink and downlink, the AoD required in angle-based beamforming can be obtained by the AoA estimation in the opposite direction [63]. This algorithm can be adopted by use cases when no beam-training overhead is allowed. Unlike angle-based beamforming, codebook-based beamforming requires beam-training process. The beam-training process, assisted by transmitting reference signal, is utilized to select codewords

with the highest data rates based on the feedback from terminals [54], so AoA estimation is not needed any more.

The applications of angle-based beamforming algorithm are introduced in [64] and [30]. However, in these applications, the beamforming vector has to be adapted to keep up with the variation of AoA, which could be time-consuming. Besides, the resolution of AoA estimation plays an important role in the performance of angle-based beamforming. In conventional AoA estimation methods, the MUSIC method was proposed to provide AoA estimation with high resolution. However, the high-resolution MUSIC algorithm can only be achieved with the premises of large numbers of snapshots and high SNR [15], which fact has already been proved by the simulation work in the last chapter. As a result, the high resolution AoA estimation in low SNR environments with limited snapshots still remains an open issue. Compared with angle-based beamforming, the inaccurate AoA estimation and the time-consuming beam generation process are evitable in beam-training based beamforming. In this beamforming algorithm, the weight vectors of phase shifters are specified by selected codewords without AoA estimation requirement [65]. A new problem arises from the heavy overhead expense caused by exhaustive beam-training for high resolution codebooks. The hierarchical codebook has been widely considered to cut down the number of symbols required for beam-training [66], on the basis of which, a low beam-training overhead codeword selection algorithm is further proposed in paper [67]. However, in the case with high frequent beam changing requirement [68], the large beam-training overhead is still the bottleneck.

With the development of machine learning, its application in beamforming has been widely studied in many recent papers [69][70]. Compared with traditional machine learning methods, the multi-layer neural network is more powerful and flexible to capture complex relationships due to the combined capability of multiple hidden layers [71]. The CNN is adopted in AoA estimation in [23], but it is a coarse estimation method and the resolution is not adequate for angle-based beamforming. In this section, with the further exploration of neural network in robust AoA estimation, a high resolution angle based codeword selection algorithm, which combines the advantages in terms of the quick beam generation of codebooks and low beam-training overhead of angle-based beamforming, is proposed. As mentioned in [72], the path loss of LoS path is much lower than those of non-LoS (NLoS) paths in mmWave systems.

Therefore, we generate few snapshots-based uplink LoS path AoA (LAoA) estimation models to select codewords from hierarchical codebooks for downlink beamforming. To make a more precise distinction among received uplink signals with similar LAoA, a two-step neural network structure is proposed to learn the relationship between receiving signal and the corresponding LAoA from a UE. In the first step, coarse LAoA estimation is achieved by classifying UEs into several angle groups. After that, LAoA of the UE can be further estimated within group range by the corresponding step 2 model.

The simulation work is applied in FDD systems, compared with the TDD systems we discussed in the last two chapters, the uplink signal is always available for frequent LAoA estimation during the downlink transmission. Due to the fact that the downlink SR and SWR are decided by the codeword for beamforming, which is selected based on estimated LAoA. As a result, the accuracy of uplink LAoA estimation becomes the only factor to affect the performance of downlink transmission. Simulation results demonstrate that our proposed two step neural network based LAoA estimation method provides higher estimation resolution than the classical MUSIC method with a couple of signal snapshots, while the proposed LAoA estimation application in codeword selection outperforms the SWR optimization oriented codeword selection algorithm for both SR and SWR within a large range of SNR.

The rest of this chapter is organized as follows. In Section 5.2, the uplink and downlink system models are introduced. In Section 5.3, the uplink signal based LAoA estimation model and its application in codeword selection is detailed, while in Section 5.4, the LAoA estimation resolution and codeword selection accuracy of proposed algorithms are compared with conventional algorithms. Finally, our conclusions are given in Section 5.5.

5.2 Signal Model

5.2.1 Downlink Signal Model

Consider a MU-MIMO-OFDM broadcast system. Due to the fact that the distance between BS and UEs are not important in the beamforming design, therefore, the influence of distance to the channel model is not considered to simplify the problem. In the downlink signal transmission, the large ULA equipped with N_t antennas at the BS is considered in mmWave MIMO system,

while UEs are equipped with single antennas. Half wavelength is designed as antenna interval, and the antenna array is used to serve $J(J > 1)$ UEs every time, the transmitter employs J analog RF beamformer. We focus on the design of analog beamformer, and the receiving signal at the kth UE on the lth subcarrier can be modeled as

$$R_D^{(k)}(l) = \sqrt{P}\mathbf{h}_D^{(k)H}(l)\mathbf{w}^{(k)}(l)s^{(k)}(l) + \underbrace{\sum_{j \neq k}^{J} \sqrt{P}\mathbf{h}_D^{(k)H}(l)\mathbf{w}^{(j)}(l)s^{(j)}(l)}_{interference} + n_0(l). \qquad (5.1)$$

For convenience, the subcarrier index l will be omitted when there will be no confusion. In Eq. (5.1), $\mathbf{w} \in \mathbb{C}_{N_t \times 1}$ is the normalized beamforming vector with $\|\mathbf{w}\| = 1$, $n_0 \sim \mathcal{CN}(0, \sigma_n^2)$ represents additive white Gaussian noise and P denotes the power of transmitter. s is the transmitted symbol block. The conjugate transpose of the downlink channel vector $\mathbf{h}_D$ can be written as

$$\mathbf{h}_D^{(k)} = \sqrt{N_t}\sum_{m=2}^{M}\gamma_m^{(k)}\alpha^{(k)}(N_t, \sin\theta_m^{(k)}) + \sqrt{N_t}\gamma_1^{(k)}\alpha^{(k)}(N_t, \sin\theta_1^{(k)}), \qquad (5.2)$$

where M and $\gamma_m \sim \mathcal{CN}(0, \sigma_M^2)$ denote the number of multipath and the channel gain of the $m^{th}(m = 2, \ldots, M)$ path, while the channel gain for the LoS path is $\gamma_1 \sim \mathcal{CN}(0, \sigma_1^2)$. The AoD $\theta(\theta \in [0, \pi])$ can be divided into the AoD of LoS path θ_1 and AoDs for NLoS paths θ_m. The function $\alpha(N, \sin\theta)$ denotes steering vectors at transmitter, where

$$\alpha(N, \sin\theta) = \frac{1}{\sqrt{N}}\left[1, e^{j\pi\sin\theta}, \ldots, e^{j\pi(N-1)\sin\theta}\right]^T. \qquad (5.3)$$

Therefore, the SR of UEs can be expressed as

$$SR = \sum_{k=1}^{J}\log_2\left(1 + \frac{\left|\sqrt{P}\mathbf{h}_D^{(k)H}\mathbf{w}^{(k)}\right|^2}{\sum_{j \neq k}^{J}\left|\sqrt{P}\mathbf{h}_D^{(k)H}\mathbf{w}^{(j)}\right|^2 + |n_0|^2}\right). \qquad (5.4)$$

Assume that the unit of signal transmission is symbol blocks and the length of blocks can be viewed as a mmWave channel coherence time period, in each coherence time period, γ_m and γ_1 can be considered to be constant. The symbol blocks in a coherence period should be divided into K_t symbols blocks for beam-training and K_d symbols blocks for data transmission, i.e.

$K = K_t + K_d$. Compared with the UEs SR, the SWR is a better parameter to show the influence of beam-training symbols to the data transmission, which can be formulated as [67]

$$S\,WR = \sum_{k=1}^{J} \frac{K - K_t^{(k)}}{K} \log_2 \left(1 + \frac{\left| \sqrt{P} \mathbf{h}_D^{(k)H} \mathbf{w}^{(k)} \right|^2}{\sum_{j \neq k}^{J} \left| \sqrt{P} \mathbf{h}_D^{(k)H} \mathbf{w}^{(j)} \right|^2 + |n_0|^2} \right). \tag{5.5}$$

5.2.2 Uplink Signal Model

The single carrier frequency division multiple access (SC-FDMA) signal is utilized in uplink LTE system. For each conventional SC-FDMA symbol block [73], a snapshot in a transmitted symbol is $x(t) = \frac{1}{\sqrt{M_s}} e^{j\frac{2\pi f_0 t}{N}} \sum_{n=0}^{M_s-1} s(n) g\left(t - n\frac{N}{M_s}\right)$, where $g(t) = \frac{1}{\sqrt{N}} e^{j\frac{\pi(M_s-1)t}{N}} \sin \frac{\pi M_s t}{N} / \sin \frac{\pi t}{N}$, N is the number of points for IFFT, t is viewed as a snapshot in the symbol block, $(t = 0, 1, \cdots, N-1)$, f_0 is the carrier frequency, $s(n)$ $(n = 0, 1, \cdots, M_s - 1)$ represents the complex modulated data in a symbol block, which is modulated by four Quadrature Amplitude Modulation (4QAM), and M_s denotes the size of a symbol block. Therefore, in the time domain, a snapshot of the uplink receiving signal at an antenna is $\mathbf{R}_U^{nt}(t)$, which can be written as

$$\mathbf{R}_U^{nt}(t) = \sqrt{P} \mathbf{h}_U^{nt}(t) * x(t) + \mathbf{n}_0^{nt}(t), \tag{5.6}$$

$$\begin{aligned} \mathbf{h}_U(t) = \sqrt{N_t} \sum_{m=2}^{M} \gamma_m \alpha(N_t, \sin \theta_m) \delta(t - \tau_m) \\ + \sqrt{N_t} \gamma_1 \alpha(N_t, \sin \theta_1) \delta(t - \tau_1), \end{aligned} \tag{5.7}$$

here $*$ denotes operation of convolution, $\mathbf{h}_U(t)$ is the uplink channel matrix and $\mathbf{h}_U^{nt}(t)$ represents the channel vector for the nt^{th} $(nt = 1, 2, \ldots, N_t)$ antenna, τ_m $(m = 1, 2, \ldots, M)$ is the time delay of each path, LAoA equals to the AoD θ_1 of the corresponding downlink channel due to angle reciprocity [74]. As shown in Fig. 5.1, the uplink and downlink signals are assumed to be transmitted with the same size of symbol block and coherence time period in FDD systems.
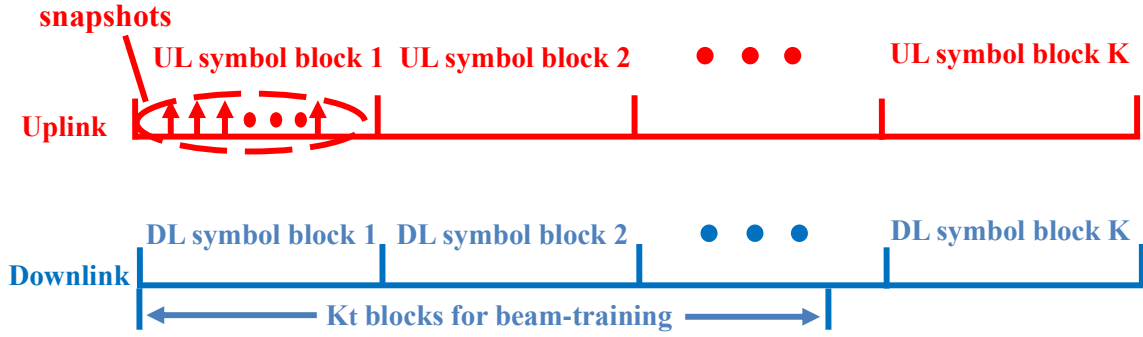
Figure 5.1: Uplink and downlink signals in a mmWave coherence time period.

5.3 AoA Estimation based Beamforming

In this section, a two-step neural network is introduced to estimate LAoA by the uplink receiving signal for codeword selection. Since the model is designed to discover the relationship between uplink receiving signal at BS and the corresponding LAoA from a UE, the uplink channel and signal models are introduced firstly. After that, the generation of training and testing signal samples are presented, and the construction of the two-step neural network models are shown. Finally, we introduce the application of the proposed neural network based LAoA estimation in codeword selection algorithm.

5.3.1 Uplink Signal based AoA Estimation

5.3.1.1 Samples Generation for Neural Network

The snapshots of the uplink receiving signal with all multipath signals received at N_s ($N_s \leqslant N_t$) adjacent antennas are utilized to generate the input nodes of training/testing samples. The first N_p snapshots are taken from an uplink symbol block of transmitted data. As mentioned in paper [76], the variability of uncertain signal waveform can be reduced by the utilization of average array covariance matrix, which are generated based on N_p uplink signal snapshots in Eq. (5.8)

$$\mathbf{R} = \frac{1}{N_p} \sum_{n=1}^{N_p} \mathbf{R_U}[n] \mathbf{R_U}^H[n]. \tag{5.8}$$

Reshape the off-diagonal upper right matrix elements of $\mathbf{R}$ as an input vector to the neural network model, which is

$$\mathbf{r} = [R_{1,2}, R_{1,3}, \ldots, R_{1,N_s}, R_{2,3}, \ldots, R_{2,N_s}, \ldots, R_{N_s-1,N_s}]$$
$$\in \mathbb{C}_{(N_s-1)N_s/2\times 1}$$
(5.9)

$$\mathbf{r}_{input} = [Real\,(\mathbf{r})\,, Imag\,(\mathbf{r})]/\|\mathbf{r}\|,$$
(5.10)

$\mathbf{r}_{input}$ is the input nodes vector for both training and testing samples, which has $N_s\,(N_s - 1)$ nodes, while the corresponding output node is decided by the selected model type, which will be explained in the next section.

5.3.1.2 Two-step Neural Network for AoA Estimation
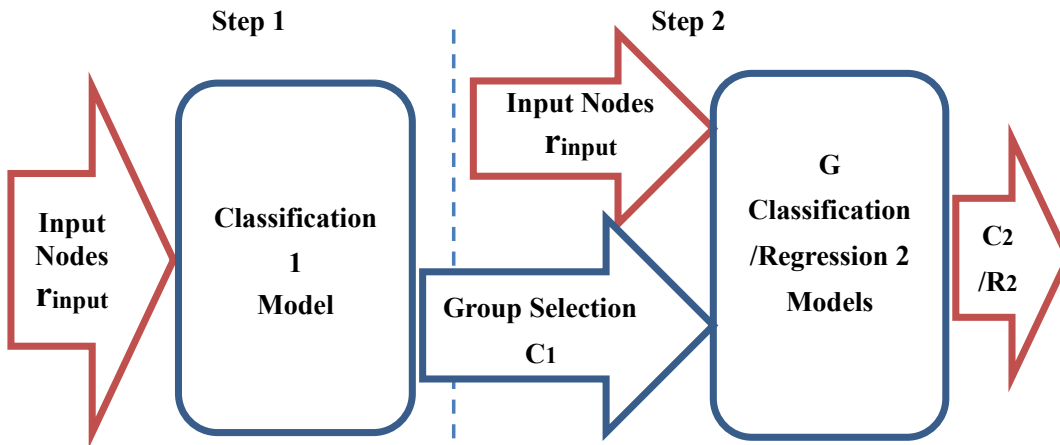
Figure 5.2: The flow chart of two-step neural network.

In order to achieve high accuracy LAoA estimation, we apply a two-step neural network based estimation algorithm. The flow chart of the proposed algorithm is shown in Fig. 5.2, in the first step, the serving space of LAoA is uniformly divided into G angle groups with group interval $g = 180°/G$, while the coarse LAoA estimation is done by classifying the UE into an angle group $C_1\,(C_1 = 0, 2, \ldots, G - 1)$. After that, the LAoA of the UE can be further estimated within the group range by the corresponding step-two regression/classification model. The classification model represents the predictive model with the task to approximate a mapping

Figure 5.3: The neural network framework for both steps.

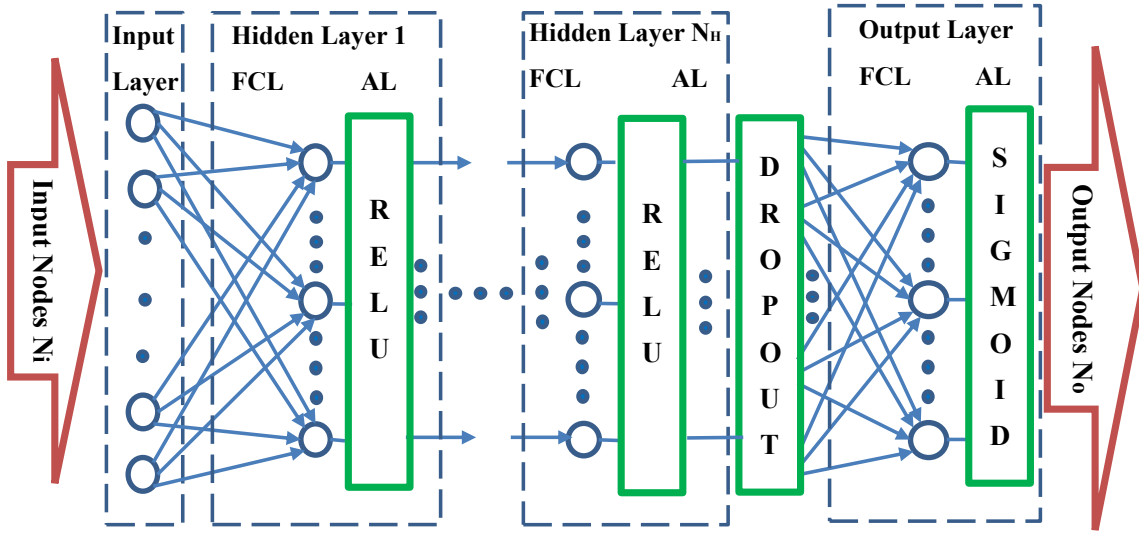function from input variables to discrete output variables, while the regression model outputs continuous variable. To simplify the writing, we write the two-step model with regression model in the second step as Two-step Classification-Regression (TCR) model, while the model with classification model in the second step is called Two-step Classification-Classification (TCC) model.

Fig. 5.3 is the framework of neural network utilized in both step 1 and step 2. The neural network has one input layer to get the input nodes into the neural network model, while N_H hidden layers and one output layer are all consisted of a FCL and an AL. The ReLU function is chosen for the AL in N_H hidden layers, while that for the output layer is sigmoid function because the output of this layer is normalized LAoA or probability, which are values between 0 and 1. The equations of two activation functions are

$$F_{sigmoid}(x) = \frac{1}{1 + e^{-x}}, \tag{5.11}$$

$$F_{ReLU}(x) = \max(0, x). \tag{5.12}$$

A dropout layer is deployed between hidden layers and the output layer for the sake of preventing overfitting [75].

It is defined that the neurons number for a FCL in each hidden layer is N_F, and that for FCL

in a output layer is equal to the number of output nodes N_o. For both step 1 and step 2 models, the number of input nodes is $N_i = N_s(N_s - 1)$, while the number of output nodes N_o for the step 1 classification model is G, and each value of output nodes is the probability of LAoA in this angle range. The index C_1 of the output nodes with the highest probability is the index of selected angle group.

For the TCR model, only one output node is required for the regression model in step 2. When the output R_2 is obtained from step 2 model, which is the normalized LAoA in the corresponding angle group, the real LAoA can be recovered by Eq. (5.13)

$$LAoA = g \times (C_1 + R_2), \quad C_1 \in [0, \cdots, G-1] \ . \tag{5.13}$$

For the TCC model, the divided UEs in G groups are further classified into smaller groups with specific LAoA values estimation. Compared with the uniform way for group division adopted in step 1, here the angle range is nonuniformly divided. As shown in Fig. 5.4, for the uniform group division, the group interval represents the angle covered by an angle group, which is the distance between adjacent edges. The largest error of this way is a group interval. However, for the nonuniform group division, the group interval is the distance between adjacent group indexes. Each point in the whole range is divided into the angle group with the shortest absolute distance from corresponding group index to the point, therefore, the largest error of this way is only half of a group interval, which can be smaller than that of the uniform way. The group interval for step 2 model is set as g_2, and the output of this step C_2 is calculated in Eq. (5.14)

$$C_2 = round\left(\frac{LAoA - g \times C_1}{g_2}\right), \quad C_1 \in [0, \cdots, G-1]. \tag{5.14}$$

The number of output nodes in the step 2 classification model can be calculated by $N_o = \frac{180°/G}{g_2} + 1$, and the LAoA can be easily recovered by $LAoA = g \times C_1 + C_2 \times g_2$. When g/g_2 is an integer, the resolution of the LAoA estimation with correct group selections in both steps is $g_2/2$.
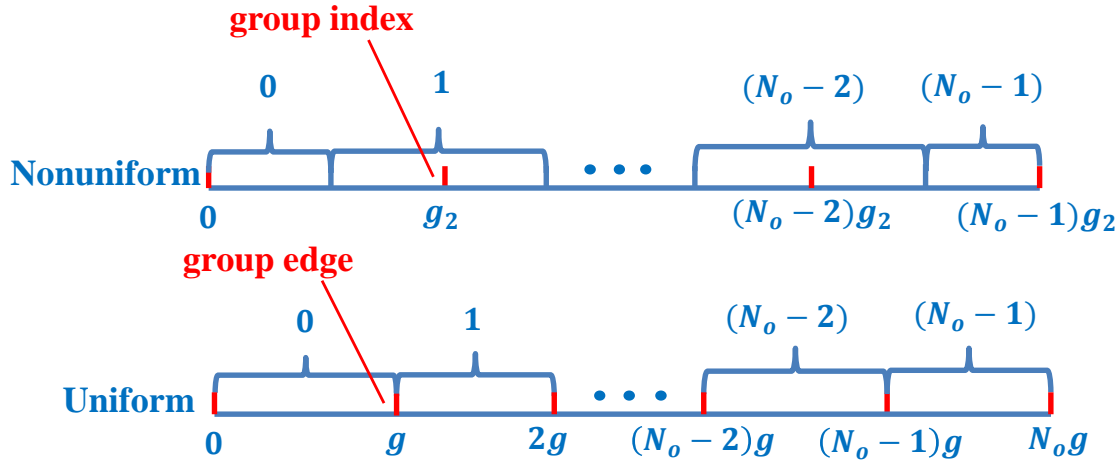
Figure 5.4: Two angle group division methods.

5.3.2 AoA based Codeword Selection for Multiple UEs

The hierarchical codebooks are designed to generate beams with varying beamwidth and beam gain based on different numbers of antennas used [66]. The codewords in the highest layer codebook can generate beams with the narrowest beamwidth and highest beam gain, which employ N_t antennas and the beam weights can be shown as Eq. (5.15)

$$\mathbf{w}(i) = \alpha\left(N_t, -1 + (2i-1)/N_t\right), (i = 1, \ldots, N_t). \tag{5.15}$$

The number of antennas for codebooks generation decreases with the layer index of codebooks. Therefore, the lower level the codeword belongs to, the wider beamwidth and lower beam gain it provides. We utilize number of antennas to evaluate the beam resolution. The more antennas used for codebook generation, the narrower the beamwidth is, the higher beam gain can be obtained, which means the higher resolution the beam has. However, a large beam-training overhead is required to select a suitable codeword with high resolution for transmission. In paper [67], a SWR optimization oriented codeword selection algorithm based on the hierarchical codebooks is proposed to make the best tradeoff between the beam resolution and the beam-training overhead, which is said to provide better performance than the multi-sectional codebooks searching algorithm [28] with regard to SWR.

In our proposed LAoA based codeword selection algorithm, with the LAoA estimated by the two-step neural network, the codebook providing the highest resolution beams can be se-

lected to serve the UEs. Due to the changing range of beams with codeword index in the highest level codebook, whether the most suitable codeword can be selected to the UE depends on the LAoA estimation accuracy and the beamwidth of beams generated by codewords in the codebook. As a classical method for AoA estimation, MUSIC can also provide estimated LAoA with high accuracy based on the uplink signal snapshots for downlink codeword selection. However, it results in less accurate AoA estimation method in low SNR environments with few snapshots [15]. In the next section, the performance of proposed algorithm will be compared with the MUSIC based codeword selection algorithm and the benchmark algorithm respectively with different SNR.

5.4 Performance Evaluation

5.4.1 System Parameters

In this section, the numerical results are given to show the performance of two-step neural network models for LAoA estimation and their applications in codeword selection algorithm. Some parameters for the simulation work are defined in Table 5.1. Noted that the training samples are generated by uplink signal snapshots with random LAoA in the angle range, which are utilized for both step 1 and step 2 models. These training samples are used to train both TCC and TCR models with different numbers of groups division, while for the step 2 models accounting for different divided ranges, the training samples with LAoA in corresponding ranges are allocated.

5.4.2 Numerical Results

5.4.2.1 LAoA Estimation Accuracy

The number of nodes utilized in hidden layers and output layers for different two-step models and single-step models are listed in Table 5.2 [67]. In Fig. 5.5 and Fig. 5.6, the testing error cumulative distribution functions (CDF) are compared among different neural network models and MUSIC algorithm. In these figures, *re* is the resolution of TCC models, which is

Table 5.1: Simulation Parameters for Beamforming Design in FDD Systems

UE distance to BS R (m)	$[0, 500]$
Complex channel gain	$\alpha_m \sim \mathcal{CN}(0, 0.01)$, $\alpha_1 \sim \mathcal{CN}(0, 1)$
number of paths M	10
number of UEs J	2
LAoA range θ	$(-\pi/2, \pi/2)$
Duplex mode	FDD
BS antennas N_t (ULA)	64/128
Samples generation antennas number N_s	32
Number of IFFT points	512
SNR (dB)	$[-5, 25]$
Symbol blocks in coherence period K	20
Neural network input nodes N_i	992
Training samples number	36000
Testing samples number	500 for each SNR
Snapshots for sample generation	3
Dropout ratio	0.25

calculated by $g_2/2$ and represents the largest error for correct classifications of both steps. Each point in figures represents the percentage of testing samples with error smaller than or equal to the value in the x axis. Due to the fact that only small error LAoA estimations can provide high accuracy codeword selections. Therefore, the error range is selected from 0.5 degree to 5 degrees in Fig. 5.5 and Fig. 5.6.

In Fig. 5.5(a)(b), we compare the performance of TCC/TCR models with different G, Single-step Classification model and Single-step Regression model. It can be found that the TCC models with the same re but different G show similar accuracy in both low and high SNR environments, and all of them can provide about 35% percents of UEs with estimation error smaller than 0.5 degree. Compared with TCC models, TCR models witness a decrease on the error CDF with the decreasing number of divided groups. Although the TCR model with 30 groups shows similar performance with TCC models in low SNR environments, its percentage with small estimation error is much lower than those of TCC models in the high SNR environments. On top of that, the Single-step Classification model outperforms the Single-step

Table 5.2: Hyperparameters for Neural Network Models

Number of models	Model type	N_H	N_F	N_o	G
1	Regression	2	500	1	0
	Classification $(re = 0.5°)$	2	500	181	
10	Classification 1	2	120	9	9
	Classification 2 $(re = 0.5°)$	2	120	21	
	Regression 2	2	120	1	
19	Classification 1	2	120	18	18
	Classification 2 $(re = 0.5°)$	2	120	11	
	Classification 2 $(re = 0.25°)$	2	200	21	
	Classification 2 $(re = 0.125°)$	2	200	41	
	Regression 2	2	120	1	
31	Classification 1	2	150	30	30
	Classification 2 $(re = 0.5°)$	2	120	7	
	Regression 2	2	120	1	

Regression model in both low and high SNR environments. The percentage of UEs with small error estimated by Single-step Classification model is even higher than that of TCR models in high SNR environments. To conclude, the TCC models are the best LAoA estimation models with different SNR.

Based on the above analysis, the TCC models with $G = 18$ are selected for further analysis. In Fig. 5.6(a)(b), the 18 groups TCC models with various resolutions are compared with the MUSIC algorithm. The TCC model with $re = 0.25$ can be found with highest percentage of UEs with small error compared with other ones. For the comparison between TCC models and MUSIC algorithm, all three TCC models outperform the MUSIC algorithm in the low SNR environments, while in the high SNR environments, MUSIC algorithm shows a little bit better performance. What's more, the TCC model with $re = 0.25$ is the one can provide the closest performance to the MUSIC algorithm in the high SNR environments.

The neural network based LAoA estimation can be compared with MUSIC based LAoA estimation in terms of the computational complexity. Both estimation methods are applied on

the average array covariance matrix in Eq. (5.1). As mentioned in [77], the MUSIC algorithm complexity is

$$O_{MUSIC} = N_s^2 (M + 2) + I (N_s + 1) (N_s - M) \ flops \ , \tag{5.16}$$

where I is the number of spectral points of the total angle range. The complexity calculation of the simple neural network framework in Fig. 5.3 is similar to that of matrix multiplications. The fully connected layers complexity calculation can be divided into the input layers part, the hidden layers part and the output layer part, the equation is

$$\begin{aligned} O_{NN} = (N_i + 1) N_F + (N_H - 1) (N_F + 1) N_F \\ + (N_F + 1) N_o \quad flops \end{aligned} \tag{5.17}$$

For the calculation of complexity for a specific testing sample input, the complexity of both step models should be added to get the sum result. Therefore, compared the complexity of 18 groups TCC models with MUSIC algorithm, for LAoA estimation with different resolutions, the MUSIC algorithm utilizes more flops than that of 18 groups TCC models. On top of that, compared with TCC models, the complexity of the MUSIC algorithm increases with resolution more dramatically.
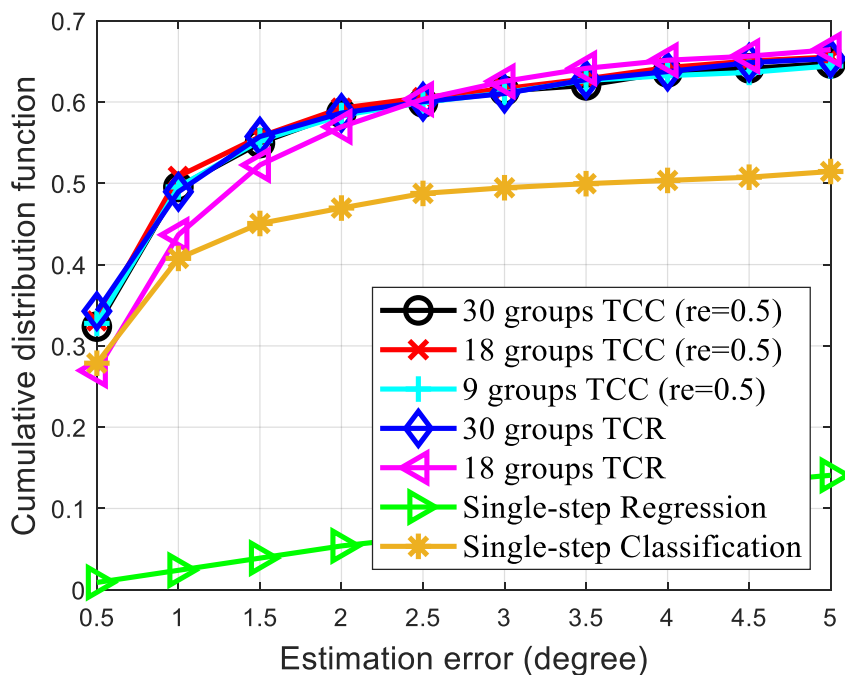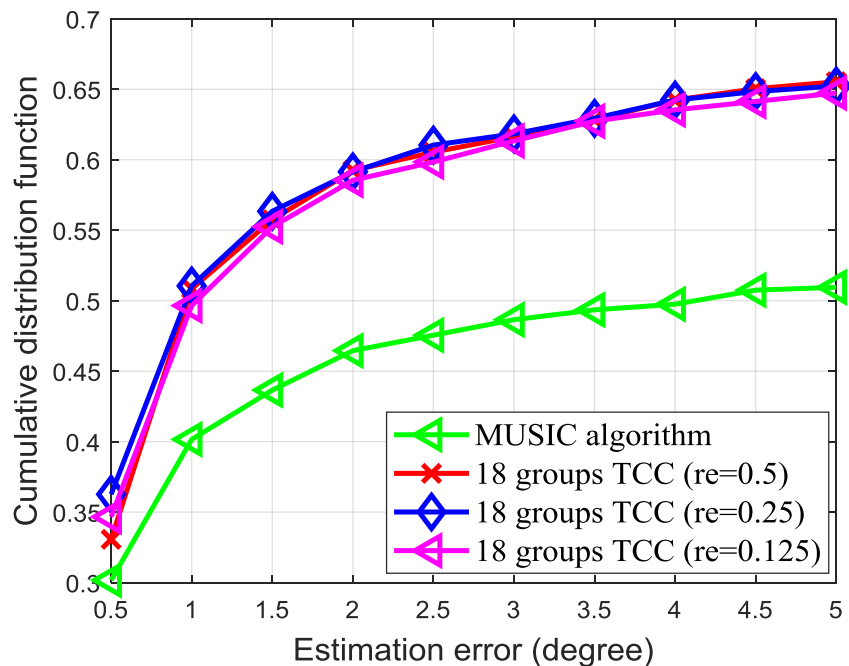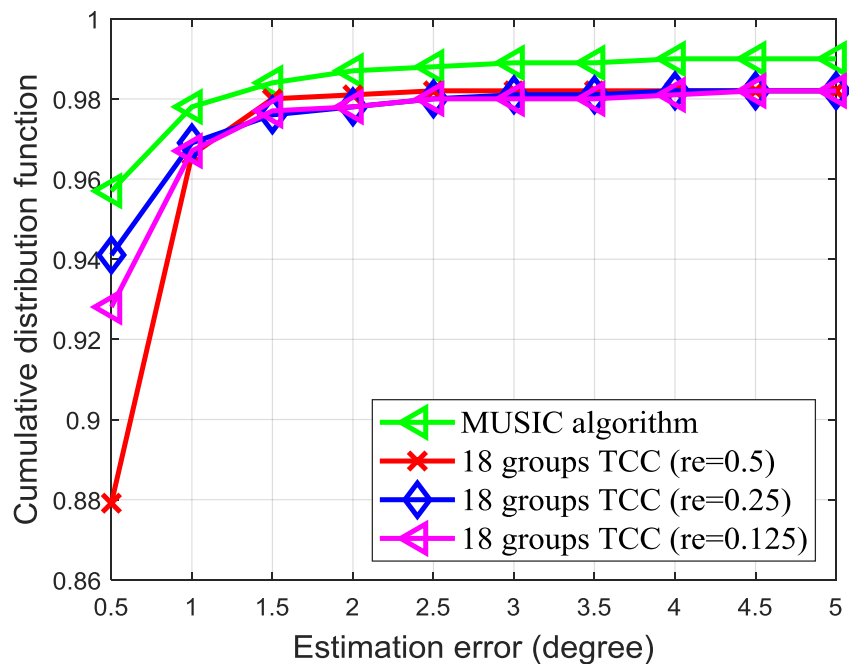
(a)
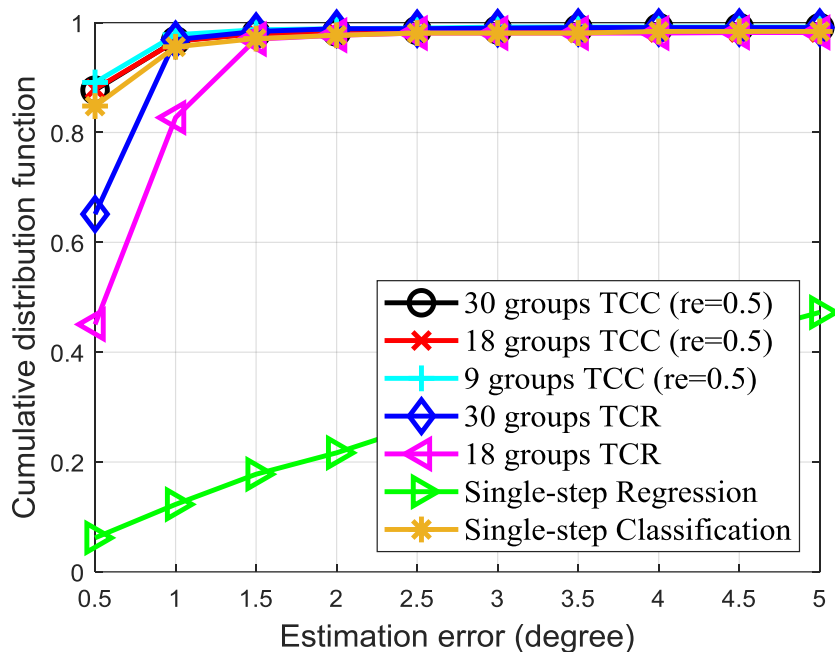
(b)

Figure 5.5: The comparisons of error CDF for different LAoA estimation models (a) SNR=-5dB (b) SNR=25dB.

(a)

(b)

Figure 5.6: The comparisons of error CDF for different resolution LAoA estimation models (a)SNR=-5dB (b) SNR=25dB.

**5.4.2.2   LAoA Estimation based Codeword Selection**

With the hierarchical codebook designed by 64 and 128 antennas, the TCC models based codeword selection algorithms with $G = 18$ and different resolutions are compared with the MUSIC based algorithm, the multi-sectional search codeword selection algorithm [28] and the SWR optimization oriented algorithm [67] in terms of SR and SWR, as shown in Fig. 5.7 and Fig. 5.8, separately. In the simulation, codewords for two UEs are considered to be selected at the same time.

When $N_t = 64$, as shown in Fig. 5.7(a), both the TCC models based codeword selection algorithms and the MUSIC based algorithm provide larger SR than that of the SWR optimization oriented algorithm when SNR>-2dB. In the low SNR environments, the TCC models outperform the MUSIC based algorithm, while in the high SNR environments, both TCC model and MUSIC based algorithms can provide SR which is equal to the multi-sectional search based algorithm, which algorithm is mentioned to provide the best codeword selections. The TCC models and the MUSIC algorithm do not require beam-training overhead for codeword selection, therefore, in Fig. 5.7(b), the TCC models based algorithms and the MUSIC based algorithm show larger SWR than the multi-sectional search based algorithm and SWR optimization oriented algorithm in all SNR environments. Noted that both SR and SWR performance of two TCC models with different resolutions are similar, which means the 0.5 degree resolution TCC model is enough for the 64 antennas based codebook.

With the hierarchical codebook designed by 128 antennas, the performance comparison of SR and SWR among two TCC models based, MUSIC based algorithm, multi-sectional search based algorithm and SWR optimization oriented based codeword selection algorithm are shown in Fig. 5.8(a)(b). The lines for four algorithms show similar changing tendencies with those shown in Fig. 5.7(a)(b), while for TCC model with $re = 0.5$ based algorithm, in the environments with SNR>13dB, it can hardly provide perfect codeword selection. To conclude that, the TCC model with $re = 0.25$ is required to adapt to beams with higher resolution when $N_t = 128$.
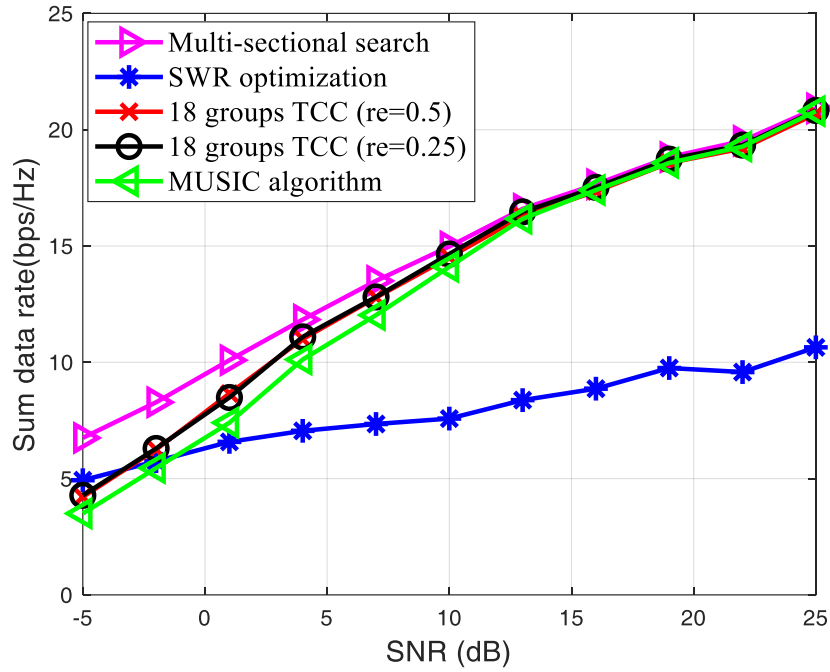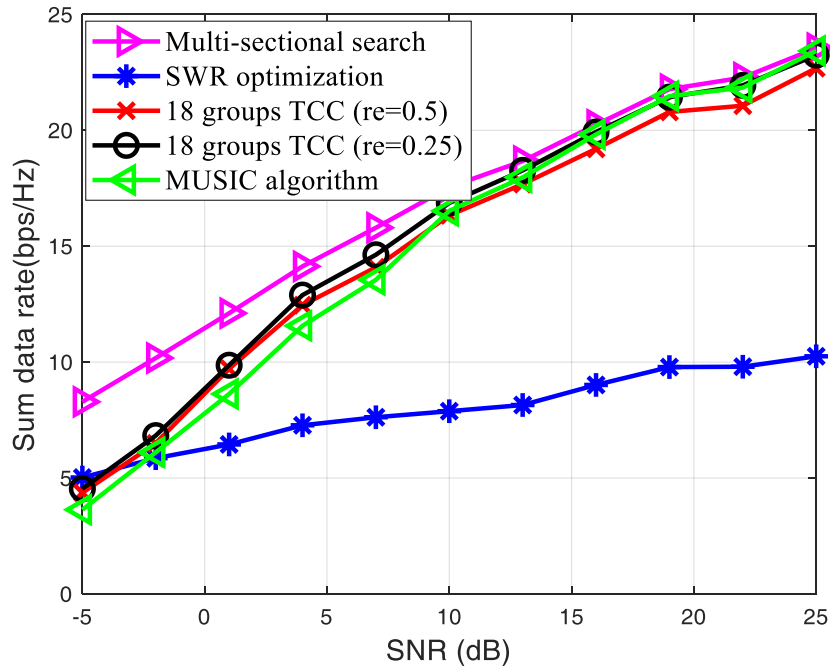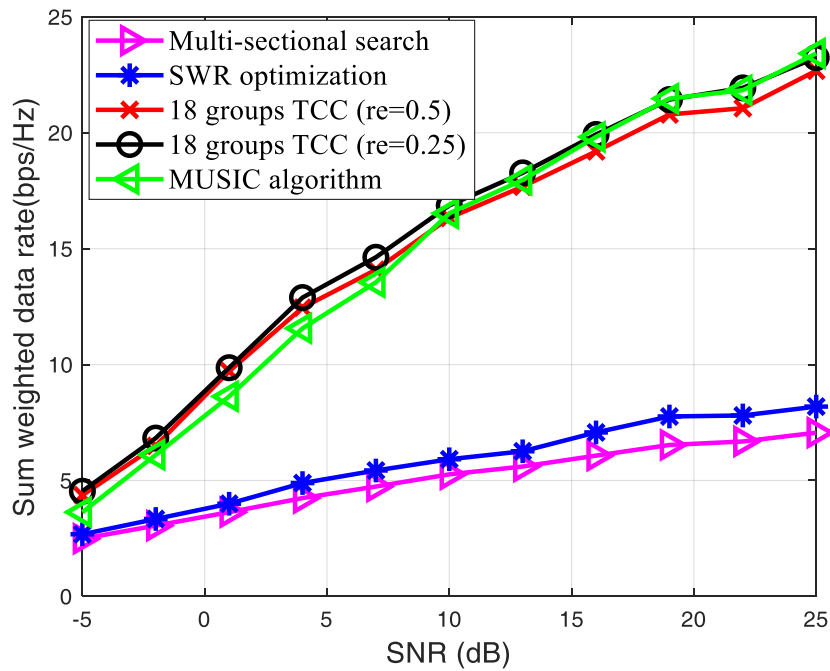
(a)
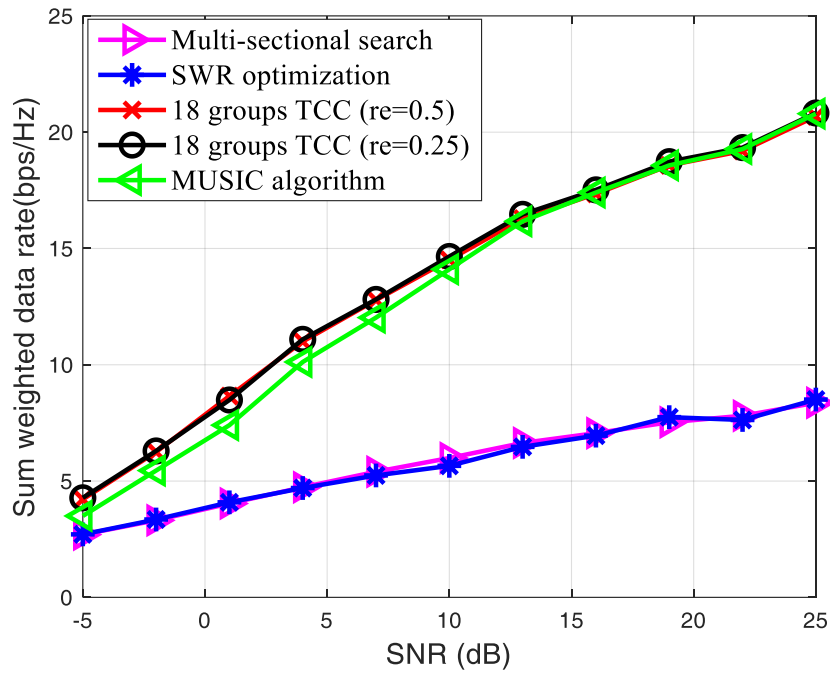
(b)

Figure 5.7: The comparisons with $N_t = 64$ of (a) SR among different algorithms (b) SWR among different algorithms

(a)

(b)

Figure 5.8: The comparisons with $N_t = 128$ of (a) SR among different algorithms (b) SWR among different algorithms.

5.5   Chapter Summary

In this chapter, angle-based beamforming algorithm without beam-training process is proposed to reduce the overhead for codeword selection while keeping high transmission data rate. According to the estimated LAoA, the beam to serve a UE is generated by selecting codeword from the highest layer of hierarchical codebook directly. Furthermore, two-step neural network models are designed to provide high resolution and fast LAoA estimation with only a couple of snapshots in the varing SNR environments. To find out the best model structure, the CDF of error for TCC models and that for the TCR models are compared in numerical results, where TCC models show better robustness for the changing SNR. After that, compared with the MUSIC based codeword selection algorithm and the SWR optimization oriented algorithm, the TCC model based algorithm with different resolutions shows better SR and SWR. The numerical results show that for codebooks with different beam resolutions, there always exists one kind of TCC model which can provide more accurate and robust codeword selection than other algorithms.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this thesis, two challenges in the mmWave systems with large scale antenna arrays were investigated in detail, which are the significant overhead for beamforming and the low robustness of antenna array based AoA estimation. To face challenges, low overhead beamforming algorithms and a fast and robust AoA estimation method are proposed in Chapter 3, 4 and 5.

The overhead for beamforming can be reduced by increasing the beams serving time in TDD systems, i.e. cutting down the times for beams change in a fixed period of time. In Chapter 3, the serving space is divided into zones according to the straight distances from UEs to the BS. Due to the relationship between the straight distances and the angular speeds of UEs with similar linear speed, the UEs in the same zone are considered as ones with close angular speed. Then the multi-pattern codebook is designed to adapt to UEs in different zones. Beams generated by codewords from a specific pattern codebook have the same beamwidth and beam gain in the spatial frequency domain, while for beams generated by different patterns codewords, antenna arrays with different numbers of elements are used to generate beams with suitable beamwidth and beam gain to serve UEs in different zones. The simulation work shows that the large intra-zone interference is caused by the reuse of a specific beam at the same time, or the use of adjacent beams simultaneously, which brings the requirement for UE selection step before beamforming.

Based on the multi-pattern codebook designed in Chapter 3, in Chapter 4, two UE selection algorithms are proposed to prevent the serious interference. With the knowledge of UEs loca-

tion information, two UE selection algorithms are designed as ultra low overhead oriented and large connection oriented, respectively, in order to adapt to different use cases. In the ultra low overhead oriented UE selection algorithm, the minimum estimated serving period for selected UEs should be longer than or equal to the predefined threshold; While in the large connection oriented algorithm, all UEs without the same beam or adjacent beam interference are served together at the beginning, while the uplink feedback every time interval enables the BS to disconnect with the UEs moving out of the selected beam serving space. Simulation results prove that the percentage of connection for the second algorithm is much higher than that of the first algorithm under the same beam serving period. The SR of the second algorithm is higher than that of the first algorithm, which can even be close to the coherence time based beamforming algorithm. While the cumulative feedback of the first algorithm is fewer than that required by the second algorithm, both proposed algorithms need fewer feedbacks than the coherence time based algorithm, which indecates that our proposed algorithms achieve the objective of overhead reduction in beamforming.

In Chapter 5, a robust and fast AoA estimation method is proposed for the low overhead beamforming in FDD systems. Due to the availability of uplink transmission signal, the few uplink transmission signal snapshots based two-step neural network models are proposed to estimate AoA of the LoS path. Then the fast estimated LAoA can be used for real-time beam update in angle based codeword selection for beamforming. The results of simulation work show that the AoA estimated by the proposed two-step classification models is more precise than that obtained by the MUSIC method in low SNR environment with only a couple of snapshots. On top of that, compared with the performance of multi-sectional codebook search algorithm, which can always make the perfect codeword selection, the SR of the proposed algorithms are lower in the cases with low SNR; while with regard to SWR, which is the parameter taking the overhead for beamforming into consideration, the proposed algorithms always outperform the multi-sectional search algorithm. For the performance comparison between the proposed algorithms and the SWR optimization algorithm, the proposed ones work better than the latter one in terms of both SR and SWR.

To summarize, this thesis propose several solutions to face the challenges in mmWave large scale antenna array system, which improve the communications performance in this system.

6.2   Future Work

In the future, the work in this thesis can be further extended from the following aspects:

- The low overhead beamforming design in TDD systems in this thesis utilizes codebooks for beam generation, which is a kind of analog signal processing method. The digital signal processing performs better on the interference reduction compared with analog signal processing, which means the exploration on hybrid beamforming may bring benefits to reduce the interference and keep the beamforming gain at the same time.

- Our proposed algorithms focus on the beamforming of signals for UEs with single dominant path to the BS. However, in some environments, more than one dominant paths may exist. The beamforming methods should be further developed to adapt to these circustances.

- For the application of physcial layer air interface techniques, the accurate CSI is always important for signal transmission and receiving. The utilization of large scale antenna arrays brings high dimensional wireless channel, which makes the accurate channel estimation become an open issue.

Bibliography

[1] Ibrahim A. Hemadeh, Katla Satyanarayana, Mohammed El-Hajjar, and Lajos Hanzo. Millimeter-Wave Communications: Physical Channel Models, Design Considerations, Antenna Constructions, and Link-Budget. *IEEE Communications Surveys & Tutorials*, 20(2):870-913,2018.

[2] Saswati Ghosh, and Debarati Sen. An Inclusive Survey on Array Antenna Design for Millimeter-Wave Communications. *IEEE Access*, 7:83137-83161,2019.

[3] Qiang Li, Tao Su, and Kai Wu. Accurate DOA Estimation for Large-Scale Uniform Circular Array Using a Single Snapshot. *IEEE Communications Letters*, 23(2):302-305,2019.

[4] Ali Cagatay Cirik, Sudip Biswas, Satyanarayana Vuppala, and Tharmalingam Ratnarajah. Beamforming Design for Full-Duplex MIMO Interference Channels-QoS and Energy-Efficiency Considerations, *IEEE Transactions on Communications*, 64(11):4635-4651,2016.

[5] Sean Huberman, and Tho Le-Ngoc. MIMO Full-Duplex Precoding: A Joint Beamforming and Self-Interference Cancellation Structure, *IEEE Transactions on Wireless Communications*, 14(4):2205-2217,2015.

[6] Naveen K. D. Venkategowda, Nitin Tandon, and Aditya K. Jagannatham. MVDR-Based Multicell Cooperative Beamforming Techniques for Unicast/Multicast MIMO Networks With Perfect/Imperfect CSI, *IEEE Transactions on Vehicular Technology*, 64(11):5160-5176,2015.

[7] Shiva Navabi, Chenwei Wang, Ozgun Y. Bursalioglu, and Haralabos Papadopoulos. Predicting Wireless Channel Features Using Neural Networks, *2018 IEEE International Conference on Communications (ICC)*, pages 1-5. IEEE, 2018.

[8] Xing Zhang, John Tadrous, Evan Everett, Feng Xue, and Ashutosh Sabharwal. Angle-of-arrival Based Beamforming for FDD Massive MIMO, *2015 49th Asilomar Conference on Signals, Systems and Computers*, pages 1-5. IEEE, 2015.

[9] Umut Ugurlu, Risto Wichman, Cássio B. Ribeiro, and Carl Wijting. A Multipath Extraction-Based CSI Acquisition Method for FDD Cellular Networks With Massive Antenna Arrays, *IEEE Transactions on Wireless Communications*, 15(4): 2940-2953,2016.

[10] Zahid Farid, Rosdiadee Nordin, and Mahamod Ismail, Recent Advances in Wireless Indoor Localization Techniques and System, *Journal of Computer Networks and Communications*, Article ID 185138, 2013.

[11] Yuan Xue, Wei Su, Hongchao Wang, Dong Yang, and Jian Ma. A Model on Indoor Localization System Based on the Time Difference Without Synchronization. *IEEE Access*, 6:34179-34189, 2018.

[12] Fuxi Wen, and Chen Liang. Fine-Grained Indoor Localization Using Single Access Point With Multiple Antennas. *IEEE Sensors Journal*, 15(3):1538-1544, 2015.

[13] Jonas Fuchs, Robert Weigel, and Markus Gardill. Single-Snapshot Direction-of-Arrival Estimation of Multiple Targets using a Multi-Layer Perceptron. *2019 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, pages 1-5. IEEE, 2019.

[14] Renzheng Cao, Binyue Liu, Feifei Gao and Xiaofei Zhang. A Low-Complex One-Snapshot DOA Estimation Algorithm with Massive ULA. *IEEE Communications Letters*, 21(5): 1071-1074, 2017.

[15] Giorgos A. Ioannopoulos, Dimitris E. Anagnostou, and Michael T. Chryssomallis. Evaluating the effect of small number of snapshots and signal-to-noise-ratio on the efficiency of MUSIC estimations, *Antennas Propagation IET Microwaves*, 11(5):755-762, 2017.

[16] Sherif Adeshina Busari, Kazi Mohammed Saidul Huq, Shahid Mumtaz, Linglong Dai, and Jonathan Rodriguez. Millimeter-Wave Massive MIMO Communication for Future Wireless Systems: A Survey. *IEEE Communications Surveys & Tutorials*, 20(2):836-869, 2018.

[17] Shajahan Kutty, and Debarati Sen. Beamforming for Millimeter Wave Communications: An Inclusive Survey. *IEEE Communications Surveys & Tutorials*, 18(2):949-973, 2016.

[18] Shu Sun, Theodore S. Rappaport, Timothy A. Thomas, Amitava Ghosh, Huan C. Nguyen, István Z. Kovács, Ignacio Rodriguez, Ozge Koymen and Andrzej Partyka, Investigation of Prediction Accuracy, Sensitivity, and Parameter Stability of Large-Scale Propagation Path Loss Models for 5G Wireless Communications. *IEEE Transactions on Antennas and Propagation*, 65(5): 2843-2860, 2016.

[19] Study on 3D Channel Model for LTE (Rel. 12), 3GPP TR 36.873, V12.1.0, 2015.

[20] Ming Xiao, Shahid Mumtaz, Yongming Huang, Linglong Dai, Yonghui Li, Michail Matthaiou, George K. Karagiannidis, Emil Björnson, Kai Yang, Chih-Lin I, and Amitabha Ghosh. Millimeter Wave Communications for Future Mobile Networks. *IEEE Journal on Selected Areas in Communications*, 35(9):1909-1935, 2017.

[21] 3GPP, TS 36.211, V15.4.0, Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation, 2018.

[22] Ramin Zaeim, "Direction of Arrival Estimation Technique for Narrow-Band Signals Based on Spatial Discrete Fourier Transform," MASc Thesis, University of Victoria, 2018.

[23] Youssef Harkouss, Hassan Shraim, and Hussein Bazzi. Direction of arrival estimation for smart antenna in multipath environment using convolutional neural network. *International Journal of RF and Microwave Computer-Aided Engineering*, 28(6):1-11, 2018.

[24] Marija Agatonović, Zoran Stanković,Nebojša Dončov, Leen Sit, Bratislav Milovanović, and Thomas Zwick. Application of Artificial Neural Networks for Efficient High-Resolution 2D DOA Estimation. *Radioengineering*, 21(4):1178-1186, 2012.

[25] Subhash Kumar Mishra, Ram Narayan Yadav, and Ripandeep Singh. A Survey on Applications of Multi Layer Perceptron Neural Networks in DOA Estimation for Smart Antennas. *International Journal of Computer Applications*, 83(17):22-28, 2013.

[26] Mamta Agiwal, Abhishek Roy, and Navrati Saxena. Next Generation 5G Wireless Networks: A Comprehensive Survey. *IEEE Communications Surveys & Tutorials*, 18(3):1617-1655, 2016.

[27] Xiaoming Chen, Derrick Wing Kwan Ng, Wolfgang H. Gerstacker, and Hsiao-Hwa Chen. A Survey on Multiple-Antenna Techniques for Physical Layer Security. *IEEE Communications Surveys & Tutorials*, 19(2):1027-1053, 2017.

[28] Zhenyu Xiao, Tong He, Pengfei Xia, and Xiang-Gen Xia. Hierarchical Codebook Design for Beamforming Training in Millimeter-Wave Communication. *IEEE Transactions on Wireless Communications*, 15(5):3380-3392, 2016.

[29] Yuhan Sun, and Chenhao Qi. Weighted Sum-Rate Maximization for Analog Beamforming and Combining in Millimeter Wave Massive MIMO Communications. *IEEE Communications Letters*, 21(8):1883-1886, 2017.

[30] Seungnyun Kim, and Byonghyo Shim. AoD-Based Statistical Beamforming for Cell-Free Massive MIMO Systems. *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, pages 1-5. IEEE, 2018.

[31] Dantong Liu, Lifeng Wang, Yue Chen, Maged Elkashlan, Kai-Kit Wong, Robert Schober, and Lajos Hanzo. User Association in 5G Networks: A Survey and an Outlook. *IEEE Communications Surveys and Tutorials*, 18(2):1018-1044, 2016.

[32] Taeseok Oh, Changick Song, Jaehoon Jung, and Inkyu Lee. A new RF beam training method for multi-user millimeter wave systems. *IEEE International Conference on Communications (ICC)*, pages 1-6, May. 2017.

[33] Xin Liu, Yanan Liu, Xianbin Wang, and Hai Lin. Highly Efficient 3-D Resource Allocation Techniques in 5G for NOMA-Enabled Massive MIMO and Relaying Systems. *IEEE Journal on Selected Areas in Communications*, 35(12):2785-2797, 2017.

[34] Hadi Ghauch, Taejoon Kim, Mats Bengtsson, and Mikael Skoglund. Sum-Rate Maximization in Sub-28-GHz Millimeter-Wave MIMO Interfering Networks. *IEEE Journal on Selected Areas in Communications*, 35(7):1649-1662, 2017.

[35] IEEE 802.11 WG. IEEE 802.11ad, Amendment 3: Enhancements for Very High Throughput in the 60GHz Band. 2012.

[36] Joongheon Kim, and Andreas F. Molisch. Fast millimeter-wave beam training with receive beamforming. *Journal of Communications and Networks*, 16(5):512-522, 2014.

[37] Theodore. S. Rappaport, Wireless Communications: Principles and Practice. Upper Saddle River, NJ: Prentice Hall, 2nd ed. edition, 2002.

[38] Rui Peng, and Yafei Tian. Robust Wide-Beam Analog Beamforming With Inaccurate Channel Angular Information. *IEEE Communications Letters*, 22(3):638-641, 2018.

[39] Hai Lin, Feifei Gao, Shi Jin, and Geoffrey Ye Li. A New View of Multi-User Hybrid Massive MIMO: Non-Orthogonal Angle Division Multiple Access. *IEEE Journal on Selected Areas in Communications*, 35(10):2268-2280, 2017.

[40] Jakob Hoydis, Stephan ten Brink, and Merouane Debbah. Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need? *IEEE Journal on Selected Areas in Communications*, 31(2):160-171, 2013.

[41] C. Lin and G. Y. Li, Adaptive Beamforming with Resource Allocation for Distance-Aware Multi-User Indoor Terahertz Communications. *IEEE Transactions on Communications*, 63(8):2985-2995, 2015.

[42] Jingbo Du, Wei Xu, Hong Shen, Xiaodai Dong and Chunming Zhao. Hybrid Precoding Architecture for Massive Multiuser MIMO with Dissipation: Sub-Connected or Fully-Connected Structures? *IEEE Transactions on Wireless Communications*, 17(8):5465-5479, 2018.

[43] Pan Cao, and John Thompson. Low complexity energy efficiency analysis in millimeter wave communication systems. *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pages 1-5, 2017.

[44] Hsiao-Lan Chiang, Tobias Kadur, and Gerhard Fettweis. Analyses of orthogonal and non-orthogonal steering vectors at millimeter wave systems. *2016 IEEE 17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1-6, 2016.

[45] Theodore S. Rappaport, Felix Gutierrez, Eshar Ben-Dor, James N. Murdock, Yijun Qiao, and Jonathan I. Tamir. Broadband Millimeter-Wave Propagation Measurements and Models Using Adaptive-Beam Antennas for Outdoor Urban Cellular Communications. *IEEE Transactions on Antennas and Propagation*, 61(4):1850-1859, 2013.

[46] Yonghee Han, Jungwoo Lee, and David J. Love. Compressed Sensing-Aided Downlink Channel Training for FDD Massive MIMO Systems. *IEEE Transactions on Communications*, 65(7):2852 - 2862, 2017.

[47] Panos N. Alevizos, Xiao Fu, Nicholas D. Sidiropoulos, Ye Yang, and Aggelos Bletsas. Limited Feedback Channel Estimation in Massive MIMO With Non-Uniform Directional Dictionaries. *IEEE Transactions on Signal Processing*, 66(19):5127 - 5141, 2018.

[48] Wiroonsak Santipach, and Michael L. Honig. Optimization of Training and Feedback Overhead for Beamforming Over Block Fading Channels. *IEEE Transactions on Information Theory*, 61(4):6103 - 6115, 2010.

[49] Hadi Ghauch, Taejoon Kim, Mats Bengtsson, and Mikael Skoglund. Sum-Rate Maximization in Sub-28-GHz Millimeter-Wave MIMO Interfering Networks. *IEEE Journal on Selected Areas in Communications*, 35(7):1649-1662, 2017.

[50] Wenqian Shen, Linglong Dai, Guan Gui, Zhaocheng Wang, Robert W. Heath, and Fumiyuki. AoD-Adaptive Subspace Codebook for Channel Feedback in FDD Massive MIMO Systems. *2017 IEEE International Conference on Communications (ICC)*, pages 1-5, 2017.

[51] Ahmed Alkhateeb, Omar El Ayach, Geert Leus, and Robert W. Heath. Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):831-846, 2014.

[52] Cunhua Pan, Hong Ren, Maged Elkashlan, Arumugam Nallanathan, and Lajos Hanzo. Weighted Sum-Rate Maximization for the Ultra-Dense User-Centric TDD C-RAN Downlink Relying on Imperfect CSI. *IEEE Transactions on Wireless Communications*, 18(2):1182-1198, 2019.

[53] Cunhua Pan, Hong Ren, Maged Elkashlan, Arumugam Nallanathan, and Lajos Hanzo. Antenna Grouping Based Feedback Compression for FDD-Based Massive MIMO Systems. *IEEE Transactions on Communications*, 63(9):3261-3274, 2015.

[54] Fangchao Zhang, Shaohui Sun, Qiubin Gao, and Hui Li. Hybrid CSI-RS transmission mechanism-based 3D beamforming scheme for FDD massive MIMO system. *China Communications* , 13(2):109-119, 2016.

[55] Yinglei Teng, Weiqi Sun, An Liu, Ruizhe Yang, and Vincent K. N. Lau. Mobility-Aware Transmit Beamforming for Ultra-Dense Networks With Sparse Feedback. *IEEE Transactions on Vehicular Technology*, 68(2):1968-1972, 2019.

[56] Yo-Seb Jeon, and Moonsik Min. Large System Analysis of Two Stage Beamforming With Limited Feedback in FDD Massive MIMO Systems. *IEEE Transactions on Vehicular Technology*, 67(6):4984 - 4997, 2018.

[57] Vutha Va, Junil Choi, and Robert W. Heath. The Impact of Beamwidth on Temporal Channel Variation in Vehicular Channels and Its Implications. *IEEE Transactions on Vehicular Technology*, 66(6):5014-5029, 2017.

[58] D. Inserra and A. M. Tonello, "A Frequency-Domain LoS Angle-of-Arrival Estimation Approach in Multipath Channels", *IEEE Transactions on Vehicular Technology*, vol. 62, no. 6, pp. 2812-2818, Jul. 2013.

[59] M. Jamalabdollahi and S. A. R. Zekavat, "Joint Neighbor Discovery and Time of Arrival Estimation in Wireless Sensor Networks via OFDMA", *IEEE Sensors Journal*, vol. 15, no. 10, pp. 5821-5833, Oct. 2015.

[60] NGMN Alliance. NGMN 5G White Paper, March 2015.

[61] Taesang Yoo, Nihar Jindal, and Andrea Goldsmith. Multi-Antenna Downlink Channels with Limited Feedback and User Selection. *IEEE Journal on Selected Areas in Communications*, 25(7):1478-1491, 2007.

[62] Irfan Ahmed, Hedi Khammari, Adnan Shahid, Ahmed Musa, Kwang Soon Kim, Eli De Poorter, and Ingrid Moerman Survey on Hybrid Beamforming Techniques in 5G: Architecture and System Model Perspectives. *IEEE Communications Surveys Tutorials*, 20(4):3060-3097, 2018.

[63] Hongxiang Xie, Feifei Gao, Shi Jin, Jun Fang, and Ying-Chang Liang. Channel Estimation for TDD/FDD Massive MIMO Systems With Channel Covariance Computing. *IEEE Transactions on Wireless Communications*, 17(6):4206-4218, 2018.

[64] Ming-Fu Tang, Yi-Ying Huang, and Borching Su. Beam-Time Block Coding With Joint User Grouping and Beamforming for FDD Massive MIMO Systems. *IEEE Access*, 6:52519-52530, 2018.

[65] Shajahan Kutty, and Debarati Sen. Beamforming for Millimeter Wave Communications: An Inclusive Survey. *IEEE Communications Surveys Tutorials*, 18(2):949-973, 2015.

[66] Zhenyu Xiao, Tong He, Pengfei Xia, and Xiang-Gen Xia. Hierarchical Codebook Design for Beamforming Training in Millimeter-Wave Communication. *IEEE Transactions on Wireless Communications*, 15(5):3380-3392, 2016.

[67] Yuhan Sun, and Chenhao Qi. Analog Beamforming and Combining Based on Codebook in Millimeter Wave Massive MIMO Communications. *in GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pages 1-6, 2017.

[68] Xuhong Chen, Jiaxun Lu, Pingyi Fan, and Khaled Ben Letaief. Massive MIMO Beamforming With Transmit Diversity for High Mobility Wireless Communications. *IEEE Access*, 5:23032-23045, 2017.

[69] Xiangrong Wang, Pengcheng Wang, and Xianghua Wang. Adaptive Sparse Array Reconfiguration based on Machine Learning Algorithms. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1159-1163, 2018.

[70] Carles Antón-Haro, and Xavier Mestre. Learning and Data-Driven Beam Selection for mmWave Communications: An Angle of Arrival-Based Approach. *IEEE Access*, 7:20404-20415, 2019.

[71] Cheng Yang, Maosong Sun, Wayne Xin Zhao, Zhiyuan Liu, and Edward Y.Chang. A neural network approach to jointly modeling social networks and mobile trajectories. *ACM Transactions on Information Systems*,35(5):124-161, 2017.

[72] Tianyang Bai, Vipul Desai, and Robert W. Heath. Millimeter wave cellular channel models for system evaluation. *in 2014 International Conference on Computing, Networking and Communications (ICNC)*, pages 178-182, 2014.

[73] Jinwei Ji, Guangliang Ren, and Huining Zhang. PAPR Reduction of SC-FDMA Signals Via Probabilistic Pulse Shaping. *IEEE Transactions on Vehicular Technology*, 64(9):3999-4008, 2015.

[74] Ming-Fu Tang, Szu-Yu Wang, and Borching Su. Beamforming designs for multiuser transmissions in FDD massive MIMO systems using partial CSIT. *in 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 1-5, 2016.

[75] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *Journal of Machine Learning Research (JMLR)*, 15(1):1929-1958, 2014.

[76] Zhang-Meng Liu, Chenwei Zhang, and Philip S. Yu. Direction-of-Arrival Estimation Based on Deep Neural Networks With Robustness to Array Imperfections. *IEEE Transactions on Antennas and Propagation*, 66(12):7315-7327, 2018.

[77] Fenggang Yan, Ming Jin, and Xiaolin Qiao. Low-Complexity DOA Estimation Based on Compressed MUSIC and Its Performance Analysis, *IEEE Transactions on Signal Processing*, 61(8):1915-1930, 2013.

Curriculum Vitae

Name:	Yuyan Zhao
Post-Secondary Education and Degrees:	2017 - present, M.E.Sc
	Electrical and Computer Engineering
	Western University
	London, Ontario, Canada

2013 - 2017, B.Sc
Communications Engineering
Tongda College of Nanjing University of Posts and Telecommunications
Yangzhou, Jiangsu, China

Honours/Awards:	Best Paper Award, IEEE WCSP 2017
Related Work Experience:	Teaching Assistant
	Western University
	2018 - 2019

Research Assistant
Western University
2017 - 2019

Publications:

[1] Y. Zhao, Y. Liu, G. Boudreau, A. B. Sediq, H. A. Zeid and X.Wang, "A Two-Step Neural Network Based Beamforming in MIMO Without Reference Signal", in *IEEE GLOBECOM 2019*, Dec. 2019.

[2] Y. Zhao, Y. Liu, G. Boudreau, A. B. Sediq, and X.Wang, "Angle-based Beamforming in mmWave Massive MIMO Systems with Low Feedback Overhead Using Multi-Pattern Codebooks", *China Communications*, 16(9):18-30, 2019.

[3] Y. Zhao, Y. Liu, G. Boudreau, A. B. Sediq, and X.Wang, "A Low Overhead Angle-based Beamforming Using Multi-Pattern Codebooks for mmWave Massive MIMO Systems", in *IEEE ICCC*, Aug. 2018.

[4] Y. Zhao, W. Xu, and S. Jin, "An minorization-maximization based hybrid precoding in NOMA-mMIMO", in *IEEE WCSP*, Dec. 2017, pp. 1-6.

[5] D. Qiao, W. Tan, Y. Zhao, C. Wen, and S. Jin, "Spectral efficiency for massive MIMO zero-forcing receiver with low-resolution ADC", in *IEEE WCSP*, Dec. 2016, pp. 1-6.