Physical Layer Techniques for Massive MIMO Sub-6 GHz LoS and Millimetre-Wave Transmission

The University Of Sheffield.

Fan Wu

Department of Electronic and Electrical Engineering

The University of Sheffield

A thesis submitted for the degree of

Doctor of Philosophy

October 2018

Dedicated to my beloved grandparents and my parents.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Professor Tim O'Farrell for the continuous support throughout my study, for his patience, motivation and subject knowledge. His guidance helped me to research and write this thesis. He has been a superior mentor.

I would also like to thank my second supervisor Professor Alan Tennant for his insightful feedback and encouragement.

I would like to dedicate this thesis to my loving grandparents, parents and family who have helped me every step of the way and supported me throughout my life, for without them, none of this could have ever been achieved.

I thank my fellow friends: Dr Qiang Bai and Dr Mohammad Reza Anbiyaei, as well as other friends for their lasting support and advice, the lengthy discussions, and time working together in the last four years.

Last but not least, I would like to thank the Sherburn and Rodgers family for their support and words of encouragement throughout writing this thesis, as well as my cat Lilly for her company on those long nights of writing.

Abstract

The explosive growth in data demand requires solutions with higher system throughput, lower energy consumption, and simultaneous support for many users. Massive multiple-input multiple-output (MIMO) and millimetre-wave (mmWave) techniques are promising candidates for next-generation wireless systems.

This thesis focuses on sub-6 GHz line-of-sight (LoS) transmissions in massive MIMO systems, which not only fulfil a variety of applications, such as small-cell back-haul but also provide a longer coherent time as the LoS channel varies more slowly and can be readily estimated compared with fading channels.This thesis also focuses on mmWave transmissions in massive MIMO systems since a large-scale antenna array can compensate for the strong pathloss of mmWave transmissions whilst the mmWave carrier frequencies enable compact BS configurations.

In this thesis, the fundamentals of the massive MIMO technique are studied comprehensively through theoretical analysis and simulations. The representative sub-6 GHz channel models of LoS and fading channels are considered. The characteristics of the LoS channel and the system performance of LoS transmissions are investigated and compared with fading channels along with the key factors that impact performance. The effective SINR expressions of the linear precoding schemes for LoS transmissions are presented. It is illustrated that the system performance of massive MIMO LoS transmissions is robust when the angles of departure are distributed within a wide range and the power of the LoS channel component is high.

The mmWave channel model and technical challenges are studied. The mmWave massive MIMO precoding problem is transformed into a beam-selecting problem. A novel channel deconstruction algorithm is proposed that enables the estimation of each received paths' parameters from the perfect or Gaussian-perturbed channel state information. Utilising the estimated path parameters, new analogue and hybrid beam-selecting (ABS and HBS) linear precoding schemes are proposed that contribute substantially to

system performance. The corresponding hardware architectures for the proposed schemes are demonstrated, which exploit low-complexity and low-cost signal processing with high energy efficiency. An enhanced hybrid beam-selecting precoding (E-HBS) scheme and hardware configuration are further proposed to achieve the optimal and near-optimal performance of digital baseband signal processing with low cost and high energy efficiency in massive MIMO systems. With E-HBS, the number of RF chains and the dimension of the baseband digital control is independent of the number of base station antennas, which is vital for massive MIMO systems.

Novel spatial user scheduling (SUS) schemes for sub-6 GHz LoS massive MIMO transmissions are proposed along with a capacity-enhancement check (CEC) scheme to further improve the system performance by mitigating the LoS channel cross-correlation.

Table of contents

List of figures

List of tables

List of Abbreviations

2-D Two-Dimensional

3-D Three-Dimensional

4G Forth-generation wireless systems

5G Fifth-generation wireless system

A-MSE Analogue estimation Mean-Square Error

ABS Analogue Beam-Selecting

AoA Angle-of-Arrival

AoD Angle-of-Departure

BS Base Station

c.d.f. cumulative distribution function

CEC Capacity-Enhancement Check

CSI Channel State Information

E-HBS Enhanced-Hybrid Beam-Selecting

FDD Frequency-Division Duplex

H-MSE Hybrid estimation Mean-Square Error

HBS Hybrid Beam-Selecting

HetNet heterogeneous network

i.i.d. independent and identically distributed

IMT-Advanced International Mobile Telecommunications-Advanced

LoS Line-of-Sight

LTE 4G Long-Term Evolution

MF Matched Filter

MIMO Multiple-Input Multiple-Output

mmWave Millimetre-Wave

NLoS Non-Line-of-Sight

OFDM Orthogonal Frequency-Division Multiplexing

p.d.f. probability density function

P2P MIMO Point-to-Point Multiple-Input Multiple-Output

RF Radio-Frequency

RUS Random User Scheduling

SINR Signal-to-Interference-plus Noise Ratio

SNR Signal-to-Noise Ratio

SUS Spatial User Scheduling

TDD Time-Division Duplex

UDN Ultra-Dense Network

ULA Uniform Linear Array

UPA Uniform Planar Array

UT User Terminal

ZF Zero Forcing

List of Notations

$\mathbf{A}$	Matrix
$\mathbf{a}$	Vector
$\mathcal{C}$	Complex number
$\mathcal{N}$	Normal distribution
$\mathbf{A} \in \mathcal{C}^{B \times C}$	A matrix $\mathbf{A}$ of size B rows and C columns with complex number elements. When B or C equals 1, it denotes a row or column vector.
$(\cdot)^H$	Hermitian transpose
$\in \{\cdot\}$	Is a member of set $\{\cdot\}$
$\mathbb{E}[\cdot]$	Mathematical expected value
$\{\cdot\}$	Set
$\mid a + b\mathbf{j} \mid$	Norm, $\mid a + b\mathbf{j} \mid = \sqrt{a^2 + b^2}$
$(\cdot)^T$	Transpose
$(\cdot)^*$	Conjugate
$\stackrel{\Delta}{=}$	Define
$\mathbf{I}_A$	Identity matrix of size A, $A \times A$ square matrix with ones on the main diagonal and zeros elsewhere

$diag(a_1,\ldots,a_B)$	Diagonal matrix of size B, $B \times B$ square matrix with $a_1,\ldots,a_B$ on the main diagonal and zeros elsewhere
$\forall$	For all
$\mathbf{A}^{-1}$	Inverse of matrix $\mathbf{A}$
∞	Infinity
$trace(\mathbf{A})$	Trace of a square matrix: sum of the elements on the main diagonal.
$\mathbf{U}[a,b]$	Uniform distribution with minimum value of a and maximum value of b
argmin	Argument of the minimum
argmax	Argument of the maximum
$\Gamma(a,\frac{1}{b})$	Gamma distribution with shape parameter a and scale parameter $\frac{1}{b}$
$\hat{a}$	Estimate of a
$\mathcal{CN}(0,1)$	Circularly symmetric complex Gaussian random variable with zero-mean and unit-variance
$\mathbb{R}$	The set of real numbers
$\mathbb{Z}$	The set of integers
$\mid \{\cdot\} \mid$	The cardinality of a set

Chapter 1

Introduction

The internet data consumption of wireless and mobile devices in 2016 was 0.59 ZB (0.59×10^{21} B). Cisco predicted that, by the year 2021, the total traffic will reach 2 ZB per year, and the mobile data traffic is expected to increase 7-fold. The amount of mobile internet traffic for video content will grow with an annual growth rate at 67% [2]. To meet this huge data traffic demand by 2020, new wireless communication techniques are required [3].

1.1 5G design aims and key techniques

Next-generation telecommunications standards, i.e. *5th-generation wireless systems* (5G), are expected to be standardised in 2020, aiming at a higher capacity than the current 4G/International Mobile Telecommunications-Advanced (IMT-Advanced) standards [4, 5]. In general, 5G supports a higher area density of broadband mobile users and allows ultra-reliable and low-latency device-to-device and massive machine-type communications [6]. Table 1.1 indicates the major design requirements for 5G developments with respect to the currently applied 4G systems [1]. The future development of 5G wireless networks must satisfy these requirements to meet the demands of different transmission scenarios. For example, Wi-Fi hotspot transmissions require a high user data rate, whereas wide-area networks require high-mobility support.

Table 1.1 5G wireless network design capabilities objectives [1]

	5G	**4G**
Peak data rate (Gbit/s)	20	1
User experienced data rate (Mbit/s)	100	10
Spectrum efficiency (bit/s/Hz)	3×	1×
Mobility (km/h)	500	350
Latency (ms)	1	10
Connection density (devices/km^2)	10×	1×
Network energy efficiency	100×	1×
Area traffic capacity (Mbit/s/m^2)	10	0.1

There are three key techniques that can be exploited to enable several orders of magnitude of performance enhancement with respect to the current 4G communications to meet the 5G requirements, such as *ultra-dense networks* (UDNs), increasing spectrum efficiency, and allocating new spectral bands [7]. UDNs have been applied in 4G communications in the form of multi-tier cellular *heterogeneous networks* (HetNets) by reusing the per-unit area of existing spectrum deployments [8]. UDNs significantly enhance the network capacity by reusing the spectrum and improve the link quality by shortening the links. In addition, enhancement of spectrum efficiency by several orders of magnitude can be achieved by *massive multiple-input multiple-output (MIMO)* techniques [9] and using higher frequency bands e.g., the *millimetre-wave* (mmWave) bands of 28 GHz-300 GHz [10].

Typically, massive MIMO techniques, as proposed in [9], utilise hundreds of antennas at the base station (BS) to serve tens of user terminals (UTs) simultaneously using the same carrier frequency in sub-6 GHz spectrum bands. The deployment of an excessive number of antennas at the BS with respect to the number of UTs allows several orders of magnitude enhancement to be achieved in not only the spectral efficiency but also the energy consumption with simple linear signal processing such as matched filter (MF) and zero forcing (ZF). Sufficient spatial multiplexing is obtained by the large-scale antenna array transmission through a unique channel of each UT in rich scattering environments. The

system capacity increases to approximately ten times higher than that of the conventional MIMO system in which the transceiver is typically equipped with fewer than ten antennas [11]. The energy efficiency is increased by the high directivity of radiation through the large-scale antenna array at the BS, with the required transmit energy per bit theoretically tending to zero as the number of antennas increases to infinity. Deploying many antennas also allows less costly components to be utilised for the hardware implementation, as the massive MIMO technique ameliorates the effects of hardware imperfections, multipath fading and noise. The effects of precision and linearity of the individual components on the system performance are no longer vital in massive MIMO systems. Instead, it is the collective effects of all components that matter, which adds robustness to the system. Furthermore, degrees of freedom in massive MIMO systems are obtained from the excessive number of antennas at the BS, which can be exploited to develop various low-complexity signal processing approaches [11–13].

Frequencies under 6 GHz are the main operating frequencies for microwave transmissions in current commercial wireless communication systems [10]. To meet the requirement for high capacity in 5G communications, a large amount of both licensed and unlicensed bandwidth is imperative to be exploited. The 28 GHz-300 GHz mmWave frequency range offers a potential solution for high-capacity future wireless communications [5, 7, 10, 14–16]. The larger the channel bandwidth, the higher the data rate that can be achieved. Although IEEE 802.11ad has standardised short-range mmWave transmissions, the potential of mmWave has not been fully exploited. The measurement of mmWave signal transmissions at frequencies of 28 GHz and 30 GHz is reported in [15], which offers new insights into the application of mmWave in mobile communications. For instance, the pathloss caused by rain attenuation and atmospheric absorption is shown to be lower than expected that is virtually negligible for mmWave at 28 GHz, 38 GHz, and 70 GHz-90 GHz in small cells with a diameter under 200 m [15, 17]. The atmospheric attenuation is approximately 0.06 dB/km at 28 GHz, 0.08 dB/km at 38 GHz, and 0.3 dB/km at 70 GHz-90 GHz. The maximum rain attenuation loss is 3 dB-6 dB in the very worst conditions.

Although HetNets, massive MIMO, and mmWave individually provide techniques to improve the system capacity, the relationship between these three techniques is symbiotic

[7, 12, 15, 18, 19]. MmWave transmissions enable interdependent techniques with HetNets and massive MIMO. In general, the utilisation of mmWave frequencies allows short carrier wavelengths. Hence, large-scale antenna arrays can be compactly packed into smaller platforms. Although mmWave signals experience severe pathloss, it can be compensated by the short propagation distance of UDNs and the substantial antenna gain of massive MIMO systems [15, 20, 21].

1.2 Research motivation and objectives

1.2.1 Aim

This research aimed to study and extend the transmission capabilities of sub-6 GHz and mmWave massive MIMO systems. Therefore, the challenges with respect to the physical layer design of massive MIMO systems can be addressed for the future deployment of 5G networks.

1.2.2 Motivation

As introduced in section 1.1, massive MIMO systems offer enormous throughput and energy efficiency, achieved by spatial multiplexing. The efficiency of the spatial multiplexing relies on the rich scattering transmission environment, such that the channel for each UT is sufficiently unique. Therefore, communication with the UTs can be executed in the same time-frequency resource using simple linear signal processing. The potential of massive MIMO systems has motivated the author to study the massive MIMO system, including investigating its spectral efficiency performance in different transmission scenarios and proposing novel low-complexity techniques for performance enhancement.

The symbiotic relationship between mmWave transmissions and massive MIMO systems has motivated the author to explore the application of mmWaves in massive MIMO systems, as massive MIMO was initially developed for sub-6 GHz transmissions. The design of a mmWave massive MIMO system is challenging. MmWave signals experience different characteristics from sub-6 GHz signals. Consequently, the signal processing

and hardware development for mmWave transmissions in 5G networks are unique [3]. The author was motivated to study and develop low-complexity linear precoding schemes to meet the low-cost and energy-efficient signal processing requirements for mmWave transmissions as well as to achieve the optimal system performance.

1.2.3 Objectives

1. **Massive MIMO systems**: This involved studying the fundamentals of massive MIMO systems through developing physical layer simulations in MATLAB for time-division duplex (TDD) mode massive MIMO downlink transmissions. To this end, the linear precoding schemes of MF and ZF are also studied. Therefore, the objective of this research was to investigate the system performance of the low-complexity MF and ZF precoding in massive MIMO systems and carry out the following investigations:

    - Study the channel models for sub-6 GHz transmissions including Rayleigh fading, line-of-sight (LoS), and Rician fading channels. Study the mmWave channel model.
    - Study fundamentals of the downlink system sum-rate capacity and the derivation of a lower bound on the sum-rate capacity as the metric of the massive MIMO system performance.
    - Study the theoretical performance of massive MIMO, i.e. the system performance of MF and ZF precoding for sub-6 GHz massive MIMO transmissions in a rich scattering environment modelled by the Rayleigh fading channel considering the number of BS antennas tends to infinity and the number of UTs is fixed.

2. **Massive MIMO LoS transmissions**: When there is a dominant propagation along the LoS between the BS and the UT, the performance of massive MIMO was unclear, especially when the number of antennas at the BS is moderately large, because of the different characteristics of LoS channels with respect to rich scattering channels. Moreover, the performance of massive MIMO relies highly on the channel state

information (CSI). The reuse of pilot sequences for massive MIMO cellular networks degrades the system performance by pilot contamination. Compared with the CSI of fading channels, the LoS CSI can be easily estimated. Due to the slow-varying nature of LoS channels, precoding using LoS CSI results in a longer coherence time for the system. Consequently, more transmission slots can be utilised for information data transmission, which contributes to the system's spectral efficiency. Therefore, LoS transmission in massive MIMO must be investigated. The following steps describe how the investigation was approached:

- Analyse the performance of MF and ZF precoding in LoS channel and Rician fading channel when the LoS CSI is exploited for precoding.
- Analyse the key factors that restrict the performance of LoS transmissions in massive MIMO systems.
- Develop simulators and compare the performance of LoS transmissions and fading channel transmissions in massive MIMO systems.

3. **MmWave massive MIMO**: MmWave signals experience unique characteristics compared with sub-6 GHz signals. Therefore, the signal processing restrictions for mmWave transmissions in massive MIMO systems must be investigated for the development of precoding schemes. The following steps illustrate the investigation of the precoding scheme for mmWave massive MIMO systems:

   - Study the channel model and signal processing constraints for mmWave massive MIMO systems.
   - Design low-complexity linear precoding schemes for TDD-mode massive MIMO systems along with low-cost and energy-efficient hardware architectures that satisfy mmWave signal processing constraints.
   - Analyse the system performance of the designed schemes. Develop simulators to evaluate the performance of the proposed methods and compare the performance with that of the conventional massive MIMO MF precoding scheme, which is achieved by the unconstrained digital baseband signal processing.

4. **Spatial user scheduling**: Based on the study of sub-6 GHz LoS massive MIMO transmissions, to improve the system performance with linear precoding, a user scheduling scheme is imperative. The existing user scheduling schemes are of high complexity for large-scale systems and fail to address the characteristics of LoS transmissions. The objective of this research was to develop a low-complexity user scheduling scheme using LoS CSI to improve the system performance for LoS massive MIMO transmissions.

1.3 Original contributions

This thesis delivers important insights into the downlink sub-6 GHz LoS and mmWave transmissions in TDD-mode massive MIMO systems. The precoding schemes are evaluated in the context of their impacts on the system performance metric of the lower bound on the sum-rate capacity in the deployments considered. As an outcome of the research, the following contributions are made:

1. **Performance evaluation for massive MIMO LoS transmissions**: The evaluation involves extensive simulations and theoretical analysis of LoS channel characteristics in two transmission scenarios: two-dimensional (2-D) transmissions from a uniform linear array equipped at the BS and three-dimensional (3-D) transmissions from a uniform planar array at the BS. The performance of MF and ZF precoding for LoS transmissions in both the LoS channel and Rician fading channel are analysed and evaluated considering the number of BS antennas is moderately large. The spatial locations of the UTs are taken into account through the distributions of the angle of departure. The impact of the system factors, namely the signal-to-noise ratio (SNR), number of antennas at the BS, Rician κ-factor in LoS transmissions, and fading channel transmissions of Rayleigh fading and Rician fading, are evaluated and compared.

2. **Proposal of a novel channel deconstruction scheme**: The mmWave channel is modelled by the limited-scattering clustered channel model. A novel channel deconstruction method is proposed for the estimation of the individual mmWave

path parameters from the CSI. The limited-scattering clustered characteristic of the mmWave channel and the asymptotic orthogonal property of the massive MIMO channel are exploited by the proposed scheme. The effectiveness of the proposed scheme is analysed for both cases of the perfect CSI adoption and the Gaussian-perturbed CSI adoption.

3. **Proposal of novel mmWave massive MIMO analogue and hybrid (analogue and digital) beam-selecting precoding schemes and hardware architecture designs**: Signal processing constraints on mmWave transmissions are studied. The mmWave massive MIMO precoding challenge is transformed into a beam-selecting problem. An analogue linear precoding scheme and a hybrid (analogue and digital) linear precoding scheme that utilise channel deconstruction results are proposed to achieve spatial multiplexing for mmWave transmissions with low complexity. The hardware architectures for the proposed precoding schemes are demonstrated that satisfy the mmWave signal processing constraints of high energy efficiency and low fabrication costs. The theoretical system performance of the proposed precoding schemes is analysed and evaluated by simulations and numerical analysis. A comparison of the system performance between the proposed schemes and the conventional massive MIMO precoding using the unconstrained digital baseband signal processing is drawn from simulation results. The hardware configurations are also compared.

4. **Proposal of the mmWave enhanced hybrid precoding scheme and hardware architecture design**: Utilising the property of the mmWave channel, an enhanced hybrid precoding scheme is proposed based on the proposed hybrid precoding scheme. The corresponding hardware architecture is illustrated wherein the adopted number of radio-frequency (RF) chains is no longer related to the number of antennas at the BS, as it is in the conventional digital baseband signal processing, but by the number of paths selected and the number of UTs served. This design is vital for the energy-efficiency improvement of massive MIMO systems. The theoretical analysis of the system performance for the proposed scheme is provided along with the evaluation by simulations and numerical analysis. The impact of the

distribution of angles-of-departure (AoDs) on the system performance is discussed and transformed into the investigation of the correlation between elements of the path vector. The uncorrelated and fully correlated cases are introduced and analysed, which respectively yield the average and lower bound on the effective signal-to-interference-plus-noise ratio (SINR). The impact of the number of selected paths on the precoder's capability of achieving the optimal and near-optimal system performance is evaluated by simulations. The effects of the SNR and distribution of AoDs are also evaluated.

5. **Proposal and evaluation of novel spatial user scheduling algorithms for sub-6 GHz massive MIMO LoS transmissions**: To improve the system performance of LoS transmissions, both transmission scenarios of 2-D transmissions from a uniform linear array (ULA) and 3-D transmissions from a uniform planar array (UPA) are analysed. The correlation-free constraints on the LoS channel that enable the grouping of LoS orthogonal UTs based on the statistical CSI of AoDs are demonstrated. Spatial user scheduling schemes are proposed for each transmission scenario based on the corresponding correlation-free constraints, which mitigate the LoS channel cross-correlation using the CSI of AoDs. A capacity-enhancement check scheme is also proposed to cooperate with the proposed spatial user scheduling schemes for further performance enhancement. The performance of the proposed schemes is evaluated through simulations and compared with the random user scheduling scheme.

The above contributions provide enhanced capacity and low complexity implementations for massive MIMO sub-6 GHz LoS and mmWave transmissions.

1.3.1 List of publications

[1] F. Wu and T. O'Farrell, "MmWave TDD-Mode Massive MIMO System Channel Deconstruction and Low-complexity, Low-cost Precoding Schemes," **(under review)** in *IEEE J. Sel. Areas Commun*.

[2] F. Wu and T. O'Farrell, "Spatial User-Scheduling for Line-of-Sight Massive MIMO Transmission," **(in preparation)** in *IEEE Trans. Wireless Commun*.

[3] F. Wu and T. O'Farrell, "An -path Diversity Hybrid Beam-Selecting Precoder for MmWave TDD Massive MIMO Systems," **(in preparation)** in *IEEE Trans. Commun*.

1.4 Thesis outline

The first chapter of this thesis summarises massive MIMO and mmWave techniques in 5G network developments and introduces the research undertaken. The motivation and objectives of the research are presented, followed by the research contributions.

Chapter 2 presents an overview of massive MIMO systems and mmWave techniques. The TDD-mode massive MIMO downlink system model and sub-6 GHz channel models of Rayleigh fading, LoS, and Rician fading channels are introduced. Conventional linear precoding schemes, MF and ZF, are presented, followed by the system performance metric of the lower bound on the sum-rate capacity. The theoretical performance of MF and ZF in Rayleigh fading channel derived using the asymptotic analysis are presented. The mmWave channel model is introduced at the end.

Chapter 3 focuses on sub-6 GHz LoS transmissions in massive MIMO systems. The advantages of LoS transmissions are reviewed. The theoretical system performance of MF and ZF precoding using the LoS CSI in both LoS and Rician fading channels is analysed. The characteristics of the LoS channel, including mean-square cross-correlation and singular value distributions, are also analysed. The impact of AoDs distribution on LoS transmissions' system performance is illustrated. The system performance of LoS and fading channel transmissions is simulated and compared.

Chapter 4 focuses on mmWave transmissions in massive MIMO systems. The mmWave massive MIMO signal processing constraints are reviewed along with the analogue and hybrid precoding solutions. The channel deconstruction approach is proposed. The effectiveness of the channel deconstruction approach is analysed for both perfect CSI adoption and Gaussian-perturbed CSI adoption. Low-complexity analogue and hybrid beam-selecting linear precoding schemes are proposed along with the low-cost, energy-

efficient design of hardware architectures. The theoretical performance of the precoders is analysed using random matrix theory and order statistic principles. This chapter also evaluates the system performance of the proposed schemes through simulations and numerical analysis of the theoretical results. The impact of key factors on system performance is evaluated.

Based on the investigation in Chapter 4, an enhanced hybrid beam-selecting precoding scheme is proposed in Chapter 5 to achieve the optimal system performance with low-cost and energy-efficient signal processing. The corresponding hardware architecture is presented. The theoretical performance is analysed, wherein the average and lower bound on the effective SINR are demonstrated. The effectiveness of the proposed precoding scheme at achieving the optimal and near-optimal system performance is evaluated through simulations and the numerical analysis of the theoretical results, followed by the impact of the system parameters.

In Chapter 6, the user scheduling schemes for massive MIMO and MIMO systems are reviewed. The LoS channel cross-correlation is analysed, which contributes to revealing the correlation-free constraints. The LoS orthogonal users' group is demonstrated based on which spatial user scheduling schemes for both transmission scenarios are proposed to reduce the LoS channel cross-correlation. The capacity-enhancement check scheme is proposed for further improvement of the system performance. The system performance of the proposed schemes for 3-D transmissions from a UPA is evaluated through simulations, followed by the impact of the system parameters.

Chapter 7 summarises the thesis by presenting the overall research outcomes and providing concluding remarks. In addition, potential further works are suggested at the end of the chapter.

Chapter 2

Overview of Massive MIMO and Millimetre-Wave Transmissions

2.1 Introduction

The global mobile data traffic is predicted to increase at a rate of 45% per year between 2016 and 2022. By the end of 2022, the data traffic of all devices is expected to be 8 times that of 2016 [22]. The enormous wireless data traffic demand has already become a burden on the current 4G LTE cellular networks [23]. For example, in a 2×2 MIMO system in LTE, a theoretical peak downlink data rate of 150.8 Mbit/s is supplied. Such performance is unlikely to support the ongoing, tremendous data traffic growth.

Next-generation wireless networks are expected to be deployed worldwide in 2020 with the aim of supporting a significant enhancement in system capacity and network throughput with seamless dense cellular coverage with respect to the current 4G networks.

Massive MIMO, originally proposed in [9, 24] for sub-6 GHz transmissions, is a compelling physical layer technique that is considered as one of the key solutions for 5G networks. Massive MIMO systems allow for orders of magnitude enhancement in both energy and spectrum efficiencies that can be achieved with low-complexity algorithms. It is demonstrated in [9] that, in a realistic transmission scenario, massive MIMO adopting MF processing permits an average net throughput of 17 Mbit/s per UT in a cellular system

with 42 UTs in each cell, and the mean net throughput is 730 Mbit/s per cell with an average spectral efficiency of 36.5 bit/s/Hz.

The mmWave spectrum has also been endorsed for 5G technologies. The current radio spectrum bands from 700 MHz to 2.6 GHz are extremely crowded for wireless communications [15]. The utilisation of mmWave carrier frequencies of 28 GHz to 300 GHz not only enables a broader bandwidth, and hence a higher data rate, but also allows advantages when cooperating with massive MIMO systems.

This chapter provides an overview of massive MIMO system developments and mmWave MIMO techniques. The TDD-mode massive MIMO downlink system model utilised in this thesis is introduced along with the representative sub-6 GHz channel models, including Rayleigh fading channel, LoS channel, and Rician fading channel. The conventional massive MIMO downlink linear precoding schemes of MF and the ZF are presented. The system performance metric of the lower bound on the sum-rate is demonstrated along with the effective signal-to-interference ratio. The theoretical system performance of sub-6 GHz massive MIMO in rich scattering transmission environments modelled by the Rayleigh fading channel is introduced. This chapter ends by introducing the conventional mmWave MIMO channel model.

2.2 Massive MIMO

Massive MIMO was developed as a large-scale multi-user MIMO technique in which a large antenna array of M antennas (tens or more elements) is deployed at the base station (BS) to communicate with K (tens) of single-antenna UTs over the same time-frequency resource [9, 11, 12, 25], as illustrated in Fig. 2.1. In a massive MIMO system, only the BS learns the channel for each UT, and UTs do not cooperate. Furthermore, the number of BS antenna elements M is much larger than the number of UTs K, and low-complexity linear processing, utilised in both uplink and downlink transmissions, is achieved by digital baseband signal processing [11].

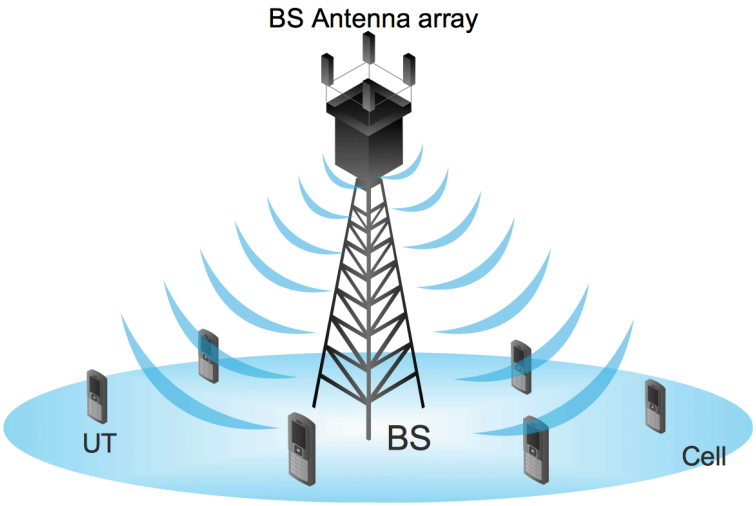


Fig. 2.1 Massive MIMO system.

2.2.1 Massive MIMO spectral and energy efficiency

As UTs are spatially located, channels to different UTs are sufficiently diverse. Massive MIMO exploits the rich scattering propagation of the radio waves to enable spatial multiplexing gain and, therefore, permits independent information signals to be transmitted to UTs in the same time-frequency resource. The spectrum efficiency advantages of massive MIMO systems are shown in [9] from random matrix theory, assuming $M \to \infty$ and $M \gg K$. Theoretically, massive MIMO is capable of achieving the MIMO system capacity upper bound with an unlimited number of BS antennas [11] in the favourable propagation scenario where channel vectors for UTs are asymptotically orthogonal. The effects of the uncorrelated noise and small-scale fading are eliminated asymptotically by linear processing, e.g. MF. The transmit power per bit from the BSs tends to zero, theoretically. Moreover, the number of supported UTs K is independent of the cell size as $M \to \infty$. M is typically designed to be significantly greater than K, whereby many degrees of freedom are available for downlink precoding. The channel from the BS to the UTs is $K \times M$. However, the actual precoding is in the K-dimensional subspace of the M dimension. This yields opportunities to achieve near-optimal performance, such

as interference elimination, or to meet the particular signal processing requirement, for example, reducing the number of adopted radio frequency chains at the BS [11].

From the energy-efficiency perspective, the large-scale antenna array deployed at the BS enables the enhancement of the array gain. Therefore, the radiation power of each element can be significantly reduced. Moreover, low-complexity signal processing schemes allow for energy savings from the power consumption on signal processing and electrical circuits [26].

2.2.2 Channel state information and TDD mode in massive MIMO

The information of the radio wave propagation channel gain is referred to as the *channel state information* (CSI). With the knowledge of CSI, on uplink transmission slots, the UTs transmit signals to the BS, and the BS uses the estimated CSI to recover the individual information from each UT, whereas on downlink transmission slots, the BS precodes the desired data streams to each UT. In practice, the CSI needed for the BS and UTs can be obtained by either pilot estimation or UT feedback. The operation modes of *time-division duplexing* (TDD) and *frequency-division duplexing* (FDD) determine the method to obtain the CSI and are suited for different transmissions. TDD mode is usually considered for massive MIMO systems.

In TDD operation mode, both uplink and downlink transmissions share the same frequency spectrum and operate in turns. The channel reciprocity property of TDD operation allows the CSI required for downlink precoding to be estimated through uplink pilot training. Let τ be the channel coherence interval (τ = coherence time $\times$ coherence bandwidth) measured in symbols, over which (time and frequency intervals) the channel is considered to be invariant. The UTs simultaneously transmit ideally pairwise-orthogonal pilot sequences of a minimum symbol length K to the BS. The pilot sequences are known at the BS and are used for the CSI estimation from the received signals. The uplink detection of the transmitted information from the UTs and the downlink precoding for UTs are then performed by the BS using the estimated CSI. The slot structure of the TDD operation mode is shown in Fig. 2.2.

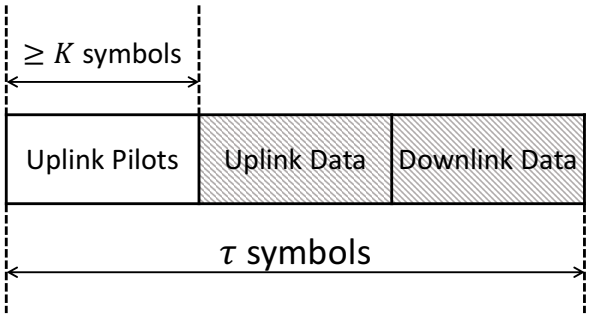


Fig. 2.2 Slot structure of the massive MIMO TDD operation mode.

In FDD operation mode, different frequency bands are exploited for uplink and downlink transmissions meaning that the CSI is different for uplink and downlink and must be estimated respectively. The uplink CSI estimation requires K UTs to transmit pilot sequences with a minimum of K symbol-length. The downlink CSI is acquired through feedback from the UTs after the BS transmits the downlink pilot of a minimum of M symbol-length. Therefore, the time that UTs spend on transmitting the feedback of downlink CSI is proportional to M. This is infeasible for massive MIMO systems since more coherence interval will be taken by the feedback transmission, which does not yield enough interval for data transmissions with the scaling up of the system, especially when the coherence interval is short in transmission scenarios for high-mobility UTs.

Compared with the FDD operation mode, the TDD mode is superior for massive MIMO systems. Only a minimum of K symbols is required during each coherence interval for pilot training, which enables no CSI feedback limitation with the scaling up of M.

2.2.3 Massive MIMO system limitations

A limited coherence interval is still the primary restriction of TDD-mode massive MIMO systems, which causes pilot contamination and disables the application of massive MIMO in high-mobility cellular environments, such as highways [27–31]. For massive MIMO cellular systems, the limited coherence interval imposes the length of symbols for pilot training. Hence, the reuse of pilot sequences among adjacent cells is inevitable to support the huge data traffic. Pilot contamination refers to the phenomenon in which, in a cellular system, the received pilot signals at one BS are contaminated by the pilot

signal transmitted by the UTs in adjacent cells using non-orthogonal pilot sequences [9, 32–34]. Moreover, in the high-mobility transmission scenarios, the channel allows a shorter coherence interval. Therefore, fewer pilot symbols can be exploited for pilot training. Both cases of the shorter pilot symbol-length and the reuse of pilot sequences cause accuracy degradation of the estimated CSI. Consequently, the massive MIMO spectral efficiency will be significantly reduced.

2.3 MmWave MIMO

A large amount of bandwidth is available in higher frequencies that enables the mmWave transmissions to achieve a dramatic throughput and spectral efficiency enhancement. MmWave MIMO utilising the less crowded frequency bands ranging from 28 GHz to 300 GHz is considered as another key technique for 5G networks. Bands including but not limited to the local multipoint distribution service band at 28 GHz-30 GHz, the unlicensed band at 60 GHz, and E-bands at 71 GHz–76 GHz, 81 GHz–86 GHz and 92 GHz–95 GHz [5, 35] could be utilised in 5G networks. A few mmWave transmission technologies have already been standardised. WirelessHD is the first standardised consumer specification in 60 GHz band for high-definition video transmissions [36]. IEEE802.15.3c is the standard for indoor wireless personal area networks (WPAN), which is the first IEEE wireless standard with over 1 Gbit/s data rate in 60 GHz [37, 38]. IEEE802.11ad Wi-Fi standard allows high data traffic in a short range [39] using 60 GHz band. MmWave has also been considered for the wireless back-haul transmissions between the core networks and the small sub-networks [40, 41].

MmWave signal processing constraints and potential solutions

Signal processing is significant for mmWave systems and is different from that for sub-6 GHz massive MIMO and MIMO systems. First, the channel model is different in mmWave, as signals of small wavelength have unique propagation characteristics, i.e. the limited scattering clustered property. The antenna receives signals that arrive in clusters with different delays, power, and central angle-of-departure (AoD) and angle-of-arrival

(AoA). The diffuse reflections are characterised by the angles and delay spreads along the paths within each cluster [42]. This property of mmWave signals must be taken into consideration when designing the mmWave signal processing methods. Second, for hardware architectures, the implementation of antenna arrays is commonly considered in mmWave systems, mainly to compensate for pathloss. IEEE802.11ad employs 32 antenna elements whilst a 256-element array has been investigated [43]. To avoid array radiation grating lobes, all of the components, including power amplifiers, RF chains, and baseband connections must be densely packed. This is extremely challenging for space-limited platforms at the BSs. Third, hardware constraints on power efficiency are also vital, especially for high pathloss mmWave transmissions. The power amplifier, analogue-to-digital, and digital-to-analogue converters are all power-demanding devices. In the conventional massive MIMO achieved by digital baseband signal processing, each antenna element is associated with a set of the above devices. Due to the high frequency of mmWave signals, the system power consumption of digital signal processing is immense [20, 42].

To overcome such mmWave signal processing constraints, the separation of the signal processing into analogue and digital domains can be executed with the aim of reducing the number of RF chains or analogue-to-digital converters. There are many potential solutions, such as hybrid beamforming, which reduces the amount of RF chains by including analogue and digital precoding stages [44]; the beamspace signal processing, which accesses the low-dimensional subspace of the mmWave channel through the Discrete Fourier Transform [45, 46], and the lens antenna, which accesses the beamspace of the channel directly with low complexity by the special antenna structure [47, 48].

MmWave massive MIMO in small cell networks

The mmWave transmission is compatible with both massive MIMO systems and HetNets. First, mmWave high-frequency signals enable the antenna separation to be very small ($\geq$ wavelength$/2$). Hence a large number of antennas can be packed on a small platform at the BS. Second, the large antenna array gain of massive MIMO compensates the strong pathloss of mmWave signals. Also, the small-cell deployment of HetNets avoids the

long-distance transmission which is unsuitable for mmWave transmissions. Moreover, the high massive MIMO antenna array gain allows the expansion of the coverage of mmWave signals. Therefore, the cooperation of mmWave Massive MIMO in small-cell deployment is one of the crucial solutions for next-generation wireless communications [10, 15, 20].

2.4 Models and preliminaries

In this section, a single-cell downlink massive MIMO system model is introduced, followed by sub-6 GHz channel models of Rayleigh fading, line-of-sight (LoS) and Rician fading. A review of the most commonly exploited massive MIMO linear precoding schemes of matched filter (MF) and zero-forcing (ZF) in the Rayleigh fading channel is presented. Throughout this thesis, a lower bound on the system sum-rate capacity is exploited as the metric of the system performance. The effective SINR for the downlink precoding is introduced, followed by the derivation of the lower bound. The system performance of MF and ZF in sub-6 GHz massive MIMO transmissions in the Rayleigh fading channel are also reviewed. The mmWave MIMO limited-scattering clustered channel model is introduced at the end of the chapter. Such preliminaries help to establish the notations used throughout the thesis.

2.4.1 Massive MIMO downlink system model

Consider a single-cell massive MIMO system, in which a BS equipped with an M-element antenna array serves K single-antenna UTs simultaneously, with $M \gg K$.

For downlink transmissions from the BS to UTs, the received signal vector $\mathbf{y} \in \mathcal{C}^{K\times 1}$ at K UTs can be denoted as

$$\mathbf{y} = \sqrt{\rho}\mathbf{G}\mathbf{x} + \mathbf{n}, \tag{2.1}$$

where ρ denotes the transmit power, and $\mathbf{G} \in \mathcal{C}^{K\times M}$ is the complex valued channel gain matrix from the BS to the UTs that embraces the joint effects of the large and small-scale fading. $\mathbf{x} \in \mathcal{C}^{M\times 1}$ is the transmit signal vector by the BS antennas. $\mathbf{n} \in \mathcal{CN}^{K\times 1}$ denotes the complex *additive white Gaussian noise* (AWGN) vector with each element

being independent and identically distributed (i.i.d.) zero mean unit variance circularly symmetric complex Gaussian random variable, i.e. $n_k \sim$i.i.d.$\mathcal{CN}(0,1)$.

The downlink $K \times M$ complex valued channel gain matrix $\mathbf{G}$ from the BS to K UTs can be further given by

$$\mathbf{G} = \sqrt{\boldsymbol{\beta}}\mathbf{H}, \tag{2.2}$$

where the diagonal matrix $\boldsymbol{\beta} = diag(\beta_1, \ldots, \beta_K)$ denotes the large-scale fading coefficient matrix with diagonal element β_k being a positive real number, which indicates the *large-scale fading*, including the pathloss and shadow fading effect, and is constant across the BS array to the k-th UT. $\mathbf{H} \in \mathcal{C}^{K \times M}$ is the channel coefficient matrix that denotes the links between the array at the BS and the UTs within a coherence interval τ. For multipath propagation, $\mathbf{H}$ is the *small-scale fading* coefficient matrix that embodies the phase shift and interference among different propagation paths in multipath fading propagation [31]. $\mathbf{H}$ is often referred to as the channel realisation. For LoS propagation, $\mathbf{H}$ is the array response matrix, which denotes the phase shift of the electromagnetic wave propagation. $\mathbf{H}$ can be further denoted as

$$\mathbf{H} = [\mathbf{h}_1^T, \ldots, \mathbf{h}_K^T]^T, \tag{2.3}$$

where the vector $\mathbf{h}_k \in \mathcal{C}^{1 \times M}$ denotes the small-scale fading channel vector or array response LoS channel vector from the BS to the k-th UT and can be further given by

$$\mathbf{h}_k = [h_{k,1}, \ldots, h_{k,M}], \tag{2.4}$$

where the element $h_{k,m} \in \mathcal{C}$ denotes the complex coefficient from the m-th antenna element at the BS to the k-th UT.

In (2.1), the transmitted signal vector $\mathbf{x} \in \mathcal{C}^{M \times 1}$ by the BS antenna array is given by

$$\mathbf{x} = \mathbf{W}\mathbf{s}, \tag{2.5}$$

where $\mathbf{W} \in \mathcal{C}^{M \times K} = [\mathbf{w}_1, \ldots, \mathbf{w}_K]$ denotes the precoding matrix with the k-th column vector $\mathbf{w}_k \in \mathcal{C}^{M \times 1}$. $\mathbf{s} \in \mathcal{C}^{K \times 1} = [s_1, \ldots, s_K]^T$ denotes the transmit symbol vector with the element s_k indicating the desired transmit symbol to the k-th UT. Also, $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = 1$ ensures the total

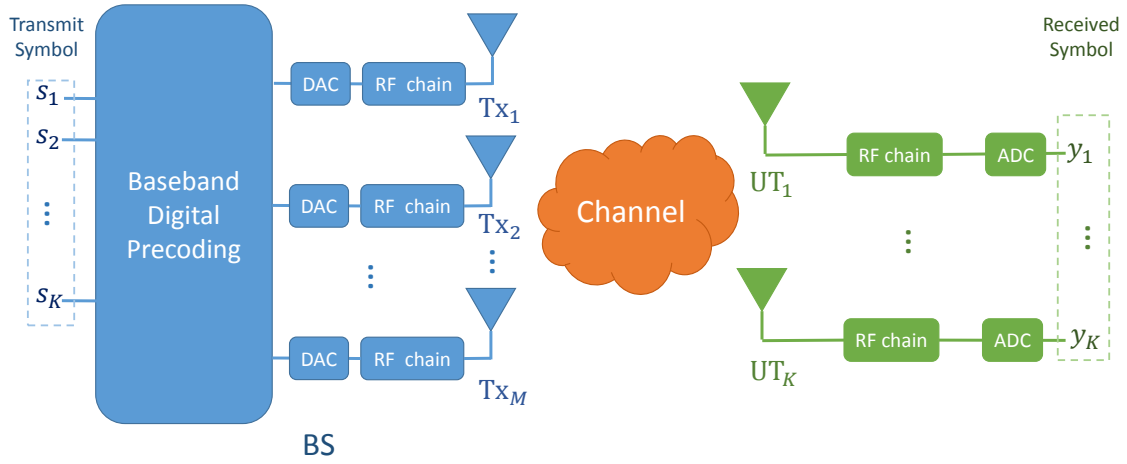


Fig. 2.3 Schematic of massive MIMO system architecture of digital baseband signal processing at frequency ≤ 6 GHz.

transmit signal power to be unity. For downlink precoding, by substituting (2.5) into (2.1), the received signal at the k-th UT is given by

$$y_k = \sqrt{\rho}\beta_k \mathbf{h}_k \mathbf{W}\mathbf{s} + n_k, \tag{2.6}$$

where $y_k \in \mathcal{C}$ is the complex signal received by the k-th UT. Fig. 2.3 illustrates the conventional sub-6 GHz massive MIMO system architecture. As shown, each antenna element at the BS is connected to a dedicated RF chain. The baseband precoding of the transmit symbols is executed through digital control. As for the UTs, no cooperation is performed.

2.4.2 Sub-6 GHz Channel Model

The narrow-band complex valued channel gains between the transmit antenna elements and the received UTs are considered to be constant within a coherence interval τ. For a narrow-band massive MIMO transmission with carrier frequency ≤ 6 GHz, the modelling of $\mathbf{H}$ is determined by the transmission scenario. This section introduces the Rayleigh fading, LoS, and Rician fading channels, which are commonly encountered. The corre-

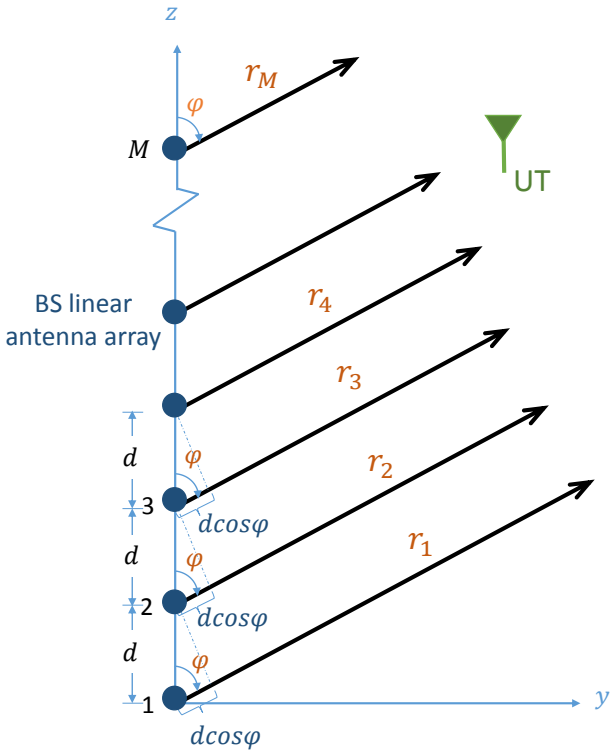
sponding notations of *Ray*, *LoS*, and *Ric* are used in the subscript of $\mathbf{H}$ for the indication of the transmission channel.

- Rayleigh fading channel model

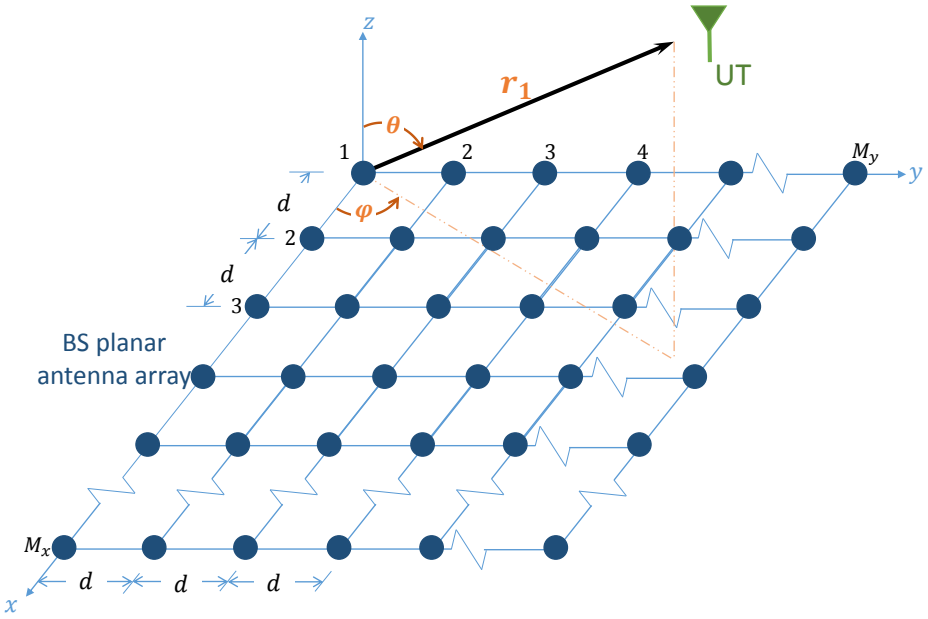
  Massive MIMO systems exploit the rich scattering propagation of the signals for the spatial multiplexing and, therefore, achieve an enormous spectrum efficiency. To model the rich scattering transmission environment, the elements of the small-scale fading coefficient channel matrix $\mathbf{H}_{Ray}$ follow the i.i.d. Rayleigh distribution [31]. Denote the Rayleigh fading channel $\mathbf{H}_{Ray}$ as

  $$\mathbf{H}_{Ray} = [\tilde{\mathbf{h}}_1^T, \ldots, \tilde{\mathbf{h}}_K^T]^T, \tag{2.7}$$

  where $\mathbf{H}_{Ray} \in \mathcal{CN}^{K \times M}$. The elements of $\mathbf{H}_{Ray}$ follow the i.i.d. circularly symmetric complex Gaussian distribution with zero mean and unit variance, i.e. $\tilde{h}_{k,m} \sim$i.i.d. $\mathcal{CN}(0,1)$. $\tilde{\mathbf{h}}_k \in \mathcal{CN}^{1 \times M}$ is the Rayleigh fading channel vector from the BS to the k-th UT. The Rayleigh fading assumption of the small-scale fading is justified by an isotropic rich-scattering transmission scenario, wherein the signal propagates to the receiver from a large number of independent scattering objects. By central limit theorem, the superposition of many independent paths facilitates the channel coefficient to be approximated by a random variable following the circularly symmetric Gaussian distribution. The Rayleigh fading channel is typically utilised for the modelling of densely built urban environments, in which only non-line-of-sight (NLoS) links are present. The CSI of the Rayleigh fading channel is treated as instantaneous CSI, due to the fast-changing nature of the small-scale fading.



(a) 2-D LoS transmissions geometry from a BS equipped with a ULA



(b) 3-D LoS transmissions geometry from a BS equipped with a UPA

Fig. 2.4 LoS transmission geometries.

- LoS channel model

  When there are LoS links from the BS to UTs, without considering the antenna mutual coupling and polarisation effects, the LoS channel matrix, i.e. the array response channel matrix is given by [31, 49]

$$\mathbf{H}_{LoS} = [\mathbf{a}_1^T, \ldots, \mathbf{a}_K^T]^T, \tag{2.8}$$

  where $\mathbf{H}_{LoS} \in \mathcal{C}^{K \times M}$, and $\mathbf{a}_k \in \mathcal{C}^{1 \times M}$ is the LoS channel vector from the BS to the k-th UT with the m-th element denoted as $\mathrm{a}_{k,m}$. The LoS channel vector is determined by the array configuration at the BS, carrier frequency, and AoDs of the UT [50, 51]. Since the UTs' spatial positions change slowly, the LoS CSI is treated as the statistical CSI, which varies more slowly with time compared with the fading channel CSI.

  For a BS equipped with a linear antenna array (1-D configured) as shown in Fig. 2.4a, the radio wave control on the horizontal dimension is disabled, meaning only the azimuth control of the transmitted signal remains. Therefore, the BS and UTs are considered to be spatially located on the same horizontal 2-D plane, thus forming 2-D transmissions. For a BS equipped with a planar array (2-D configured) as shown in Fig. 2.4b, both azimuth and horizontal dimensions' radio wave controls can be achieved. Therefore, 3-D transmissions from the BS to UTs are enabled. In this thesis, a uniform separation between the adjacent antenna elements of the array is considered. Denoting the elevation and azimuth AoDs as θ and φ, the modelling of the LoS transmission scenarios are introduced as follows [49]:

  1. 2-D transmissions from a BS equipped with a uniform linear array (ULA) When the BS and UTs are located on the same horizontal plane, the elevation dimension is ignored since $\theta_k = 90^\circ$ for all $k = 1, \ldots, K$. Hence, the far-field observation of the BS geometry positioned along the z-axis is shown in Fig. 2.4a. r_m denotes the distance from the m-th element of the array to the UT. The differences between $r_1, \ldots, r_M$ in the far-field are omitted, therefore $r_1, \ldots, r_M$ are normalised to unity. Let the first antenna element be the origin of the 2-D

geometry system, the array response vector, i.e. the LoS channel vector for the k-th UT can be denoted as

$$\mathbf{a}_k(\varphi_k) = \left[1, e^{-\mathbf{j}\frac{2\pi}{\lambda}d\cos(\varphi_k)}, \ldots, e^{-\mathbf{j}(M-1)\frac{2\pi}{\lambda}d\cos(\varphi_k)}\right], \tag{2.9}$$

where φ_k is the azimuth AoD of the k-th UT, λ is the carrier wavelength and d is the separation between the adjacent antenna elements.

2. 3-D transmissions from a BS equipped with a uniform planar array (UPA)

   Compared with the 2-D transmissions from a ULA, the additional dimension of the planar array provides both azimuth and elevation radio wave controls. The 3-D transmissions' geometric model is shown in Fig. 2.4b, in which an M-element planar array is equipped at the BS with M_x elements along the x-axis and M_y elements along the y-axis, $M_x \times M_y = M$. The separations between the adjacent elements along each dimension are uniform. The differences between $r_1, \ldots, r_M$ in far-field are omitted, therefore $r_1, \ldots, r_M$ are also normalised to unity. By setting the first element of the antenna array as the origin of the 3-D geometry system, the array response vector, i.e. the LoS channel vector for the k-th UT can be denoted as

$$\begin{aligned}\mathbf{a}_k(\theta_k, \varphi_k) = \Big[&1, .., e^{-\mathbf{j}\frac{2\pi}{\lambda}d((m_x-1)\sin(\theta_k)\cos(\varphi_k)+(m_y-1)\sin(\theta_k)\sin(\varphi_k))}, \\ &\ldots, e^{-\mathbf{j}\frac{2\pi}{\lambda}d((M_x-1)\sin(\theta_k)\cos(\varphi_k)+(M_y-1)\sin(\theta_k)\sin(\varphi_k))}\Big],\end{aligned} \tag{2.10}$$

   where θ_k and φ_k are the elevation and azimuth AoDs of the k-th UT from the BS. $1 \leq m_x \leq M_x$, $1 \leq m_y \leq M_y$ are the indices of the antenna elements on the x and y dimensions, respectively.

- Rician fading channel model

  The Rician fading channel includes both the LoS channel and the Rayleigh fading channel and represents a transmission scenario, in which both LoS and NLoS links

are present. The Rician fading channel is expressed as

$$\mathbf{H}_{Ric} = \sqrt{\frac{\boldsymbol{\kappa}}{\boldsymbol{\kappa}+\mathbf{I}_K}}\mathbf{H}_{LoS} + \sqrt{\frac{\mathbf{I}_K}{\boldsymbol{\kappa}+\mathbf{I}_K}}\mathbf{H}_{Ray}, \tag{2.11}$$

where $\mathbf{H}_{Ric} \in \mathcal{C}^{K\times M}$, $\mathbf{H}_{LoS}$ is denoted in (2.8), and $\mathbf{H}_{Ray}$ is given by (2.7). The diagonal matrix $\boldsymbol{\kappa} = diag\{\kappa_1, \ldots, \kappa_K\}$ denotes the Rician κ-factor matrix with the k-th diagonal element κ_k denoting the Rician κ-factor for the k-th UT that is the average power ratio between the deterministic and random scattering components. $\mathbf{I}_K$ denotes a K-dimensional identity matrix with ones on the main diagonal. From (2.11), the higher the κ_k is, the more power the LoS channel has, compared with the power of the scattered paths. When $\kappa_k \to \infty$, only the LoS channel presents itself, and when $\kappa_k = 0$, the only transmission channel involved is the Rayleigh fading channel. Therefore, the Rician fading CSI is instantaneous CSI due to the presence of the NLoS channel. The LoS channel is considered to be the deterministic component in a Rician fading channel since it varies (due to the motion of the UTs) slowly compared with the NLoS channel. Correspondingly, the Rayleigh fading channel is often referred to as the random component in a Rician fading channel that accounts for the scattering effects.

2.4.3 System performance metric

The definition of the system data traffic throughput that indicates the maximum data rate provided by the system is given by

$$\text{Throughput (bit/s)} \triangleq \text{Bandwidth (Hz)} \times \text{Spectral efficiency (bit/s/Hz)}. \tag{2.12}$$

It is evident that the increase of either the operational bandwidth or the system spectral efficiency enhances the system throughput. For a massive MIMO system, the spatial multiplexing is achieved on the same time-frequency resource. Therefore, the spectral efficiency is focused on in this thesis, which can be further given by

$$\text{Spectral efficiency (bit/s/Hz)} = (1 - \frac{\tau_p}{\tau_c})C^e, \tag{2.13}$$

where C^e is the system ergodic sum-rate capacity, τ_p is the samples in each coherence interval that are used for pilot transmissions, and τ_c is the total samples in each coherence interval. Hence, $(1-\frac{\tau_p}{\tau_c})$ denotes the samples for data transmission in each coherence interval. In TDD-mode systems, pilot transmissions are implemented in the uplink. Therefore, all samples in each downlink transmission coherence interval are used for data transmissions that facilitate $\tau_p = 0$. Hence, $(1-\frac{\tau_p}{\tau_c}) = 1$, and C^e represents the downlink spectral efficiency.

Lower bound on the system sum-rate capacity and effective SINR

Throughout this thesis, a lower bound on the system sum-rate capacity is exploited as the system performance metric. Such bounding method is widely utilised for massive MIMO systems [31, 52–54]. In this section, the utilised bounding method is introduced. From (2.6), consider uniform large-scale fading coefficients for all of the K UTs that are normalised to unity for simplicity. Therefore, $\beta_1 = \cdots = \beta_K = 1$. The received signal of the k-th UT in (2.1) is further given by

$$y_k = \underbrace{\sqrt{\rho}\mathbf{h}_k\mathbf{w}_k s_k}_{\text{Desired signal}} + \underbrace{\sqrt{\rho}\sum_{\substack{j=1,\\ j\neq k}}^{K}\mathbf{h}_k\mathbf{w}_j s_j}_{\text{Inter-user Interference}} + \underbrace{n_k}_{\text{Noise}} . \tag{2.14}$$

where the received desired signal is the intended information signal for the k-th UT transmitted via the channel from the BS to the k-th UT, and the inter-user interference is the information signals intended for other UTs received by the k-th UT.

From (2.14), the received signal-to-interference-plus-noise-ratio (SINR) of the k-th UT is defined as

$$\text{SINR}_k \triangleq \frac{\rho|\mathbf{h}_k\mathbf{w}_k|^2}{\rho\sum\limits_{\substack{j=1,\\ j\neq k}}^{K}|\mathbf{h}_k\mathbf{w}_j|^2 + |n_k|^2}, \tag{2.15}$$

where $|\cdot|^2$ denotes the squared norm of a complex number, i.e. the power of a signal.

The ergodic rate of the k-th UT can be given as [31]

$$C_k^e = \mathbb{E}[log_2(1+\text{SINR}_k)]. \tag{2.16}$$

Hence, the system sum-rate capacity can be denoted as

$$C^e = \sum_{k=1}^{K} C_k^e, \tag{2.17}$$

where C^e is the summation of the ergodic rates for all of the UTs.

The exact value of the ergodic system sum-rate capacity is difficult to acquire. For the system performance metric, the capacity bound (also called the achievable rate) can be exploited [31]. Next, the bounding method used in this thesis that provides a lower bound on the system sum-rate capacity in (2.17) from the effective SINR is introduced.

Note that when the CSI is unavailable at the UTs, the interference plus noise become the effective noise (with respect to the desired signal), which is non-Gaussian and has an entropy upper-bounded by the entropy of the Gaussian noise with the same variance. Hence, a lower bound on the capacity is obtained by considering the effective noise to be Gaussian [31, 52, 53]. From (2.14), the effective SINR for the k-th UT is given by

$$\begin{aligned}\text{SINR} &= \frac{variance(\sqrt{\rho}\mathbf{h}_k\mathbf{w}_k s_k)}{variance(\sqrt{\rho}\sum_{\substack{j=1,\\ j\neq k}}^{K}\mathbf{h}_k\mathbf{w}_j s_j) + variance(n_k)} \\ &\overset{(a)}{=} \frac{\rho\mathbb{E}[|\mathbf{h}_k\mathbf{w}_k|^2]}{\rho\mathbb{E}[|\sum_{\substack{j=1,\\ j\neq k}}^{K}\mathbf{h}_k\mathbf{w}_j|^2]+1},\end{aligned} \tag{2.18}$$

where (a) holds when the expected value of each term in (2.14) equals zero. The lower bound on the system sum-rate capacity of the k-th UT can be obtained from [31, 52, 53]

$$C_k = \log_2(1+\text{SINR}) \leq C_k^e, \tag{2.19}$$

where C_k^e is the ergodic rate in (2.16). Note that the UTs are independent. Therefore, the lower bound on the system sum-rate capacity is given by

$$C = KC_k = K\log_2(1+\text{SINR}) \leq C^e, \tag{2.20}$$

where C^e is the system sum-rate capacity in (2.17).

2.4.4 Sub-6 GHz massive MIMO linear precoding schemes for Rayleigh fading channel transmissions

Massive MIMO exploits the diversity of rich scattering user channels to achieve the spatial multiplexing for spectrum and energy efficiency improvements. Independent information signals are transmitted by precoding to the channel with the knowledge of CSI at the BS. In this section, the most commonly researched low-complexity linear precoding schemes of matched filter (MF) and zero forcing (ZF) are reviewed. It has been shown that such precoders provide nearly optimal system performance for sub-6 GHz massive MIMO transmissions in rich scattering environments modelled by the Rayleigh fading channel [11, 12, 31, 52]. Throughout this section, the perfect CSI adoption for precoding is assumed.

- The MF precoding matrix using the CSI of a Rayleigh fading channel can be denoted from (2.7), and given by

$$\mathbf{W}_{Ray}^{MF} = \frac{1}{\sqrt{(\gamma_{Ray}^{MF})}}\mathbf{H}_{Ray}^{H}, \tag{2.21}$$

  where γ_{Ray}^{MF} denotes the power normalisation factor of the precoder that is given by

$$\gamma_{Ray}^{MF} = \mathbb{E}[trace(\mathbf{H}_{Ray}\mathbf{H}_{Ray}^{H})], \tag{2.22}$$

  where *trace* denotes the trace of the matrix, i.e. the sum of the elements on the main diagonal.

  From (2.21), despite the power normalisation factor, the MF precoding matrix is the conjugate transpose of $\mathbf{H}_{Ray}$, hence, it maximises the received signal-to-noise ratio

(SNR) for each UT and ignores the effect of the inter-user interference. Therefore, MF is preferable in transmissions with limited inter-user interference. The signal processing of MF precoding is simple, as the BS only performs the multiplication of the desired signal with the conjugate-transpose of the channel matrix.

- The ZF precoding matrix exploiting the CSI of a Rayleigh fading channel is denoted as

$$\mathbf{W}_{Ray}^{ZF} = \frac{1}{\sqrt{\gamma_{Ray}^{ZF}}} \mathbf{H}_{Ray}^{H} (\mathbf{H}_{Ray} \mathbf{H}_{Ray}^{H})^{-1}, \tag{2.23}$$

where γ_{Ray}^{ZF} denotes the power normalisation factor of the precoder that is expressed as

$$\gamma_{Ray}^{ZF} = \mathbb{E}[trace(\mathbf{H}_{Ray}^{H}(\mathbf{H}_{Ray}\mathbf{H}_{Ray}^{H})^{-1}(\mathbf{H}_{Ray}^{H}(\mathbf{H}_{Ray}\mathbf{H}_{Ray}^{H})^{-1})^{H})] = \mathbb{E}[trace((\mathbf{H}_{Ray}\mathbf{H}_{Ray}^{H})^{-1})]^{*\dagger}. \tag{2.24}$$

From (2.23), despite the power normalisation factor, the ZF precoder is the pseudo-inverse of $\mathbf{H}_{Ray}$. Different with MF, the ZF precoder eliminates the inter-user interference, however, ignoring the noise effect. Therefore, ZF is preferable in transmissions with limited noise. Compared with the MF, the signal processing complexity for ZF is higher, as the pseudo-inverse is required.

2.4.5 Sub-6 GHz massive MIMO system performance in the Rayleigh fading channel

In this section, the theoretical performance of sub-6 GHz massive MIMO downlink transmissions in the Rayleigh fading channel is reviewed. The linear precoding schemes of MF and ZF are exploited. The asymptotic analysis and random matrix theory are used for the derivation of the closed-form expressions of the lower bounds assuming $M, K \to \infty$ with a fixed ratio $\chi = M/K$.

Effective SINR for MF precoding

Note that for MF precoding, expected value of each term in (2.14) is zero. Therefore, to obtain the effective SINR, only the expected value of the power of each term is required.

*if $\mathbf{A}$ is invertible, $(\mathbf{A}^{-1})^{H} = (\mathbf{A}^{H})^{-1}$.
†$trace(\mathbf{A}^{H}\mathbf{A}) = trace(\mathbf{A}\mathbf{A}^{H})$.

From the Rayleigh fading channel model denoted in (2.7), the expected value of the power of the desired signal for MF precoding in (2.18) can be given as

$$\begin{aligned}\mathbb{E}[|\tilde{\mathbf{h}}_k \mathbf{w}_k|^2] &= \frac{1}{\gamma_{Ray}^{MF}} \mathbb{E}[|\tilde{\mathbf{h}}_k \tilde{\mathbf{h}}_k^H|^2] = \frac{1}{\gamma_{Ray}^{MF}} \mathbb{E}[|\sum_{m=1}^{M} \tilde{h}_{k,m} \tilde{h}_{k,m}^*|^2] \\ &\overset{(a)}{=} \frac{1}{\gamma_{Ray}^{MF}} |M\mathbb{E}[\tilde{h}_{k,m} \tilde{h}_{k,m}^*]|^2 = \frac{M^2}{\gamma_{Ray}^{MF}},\end{aligned} \tag{2.25}$$

where (a) holds since $\tilde{h}_{k,m}$ for $\forall k = \{1,\ldots,K\}, \forall m = \{1,\ldots,M\}$ are i.i.d. $\mathcal{CN}(0,1)$.

The expected value of the power of the inter-user interference in (2.18) can be derived as

$$\begin{aligned}\mathbb{E}[|\tilde{\mathbf{h}}_k \mathbf{w}_j|^2] = \frac{1}{\gamma_{Ray}^{MF}} \mathbb{E}[|\sum_{m=1}^{M} \tilde{h}_{k,m} \tilde{h}_{j,m}^*|^2] &\overset{(a)}{=} \frac{1}{\gamma_{Ray}^{MF}} \mathbb{E}[\sum_{m=1}^{M} |\tilde{h}_{k,m} \tilde{h}_{j,m}^*|^2] \\ &\overset{(b)}{=} \frac{M}{\gamma_{Ray}^{MF}} \mathbb{E}[|\tilde{h}_{k,m} \tilde{h}_{j,m}^*|^2] = \frac{M}{\gamma_{Ray}^{MF}},\end{aligned} \tag{2.26}$$

where $(a),(b)$ hold since $\tilde{h}_{k,m}$ for $\forall k = \{1,\ldots,K\}, \forall m = \{1,\ldots,M\}$ are i.i.d. $\mathcal{CN}(0,1)$, therefore, $\mathbb{E}|[\tilde{h}_{k,m} \tilde{h}_{k,m}^*]|^2 = 1$.

The power normalisation factor in (2.22) is derived as

$$\mathbb{E}[\gamma_{Ray}^{MF}] = \mathbb{E}[trace(\mathbf{H}_{Ray} \mathbf{H}_{Ray}^H)] = \mathbb{E}[\sum diag(\mathbf{H}_{Ray} \mathbf{H}_{Ray}^H)] \underset{(a)}{=} MK, \tag{2.27}$$

where (a) holds since $\tilde{h}_{k,m} \sim$ i.i.d. $\mathcal{CN}(0,1)$ for $\forall k = \{1,\ldots,K\}, \forall m = \{1,\ldots,M\}$.

Therefore, by substituting (2.25),(2.26),(2.27) into (2.18), the effective SINR for MF precoding in the Rayleigh fading channel is obtained and given by

$$\text{SINR}_{Ray}^{MF} = \frac{\rho \frac{M}{K}}{\rho \frac{K-1}{K} + 1} \approx \frac{\rho \frac{M}{K}}{\rho + 1}. \tag{2.28}$$

By substituting (2.28) into (2.20), the lower bound on the sum-rate capacity of MF precoding in the Rayleigh fading channel for massive MIMO transmissions is obtained.

Effective SINR for ZF precoding

For ZF precoding in the Rayleigh fading channel, the inter-user interference is eliminated by the pseudo inverse of $\mathbf{H}_{Ray}$. Therefore, the effective SINR becomes the effective

SNR as the interference term is not present. Again, the expected value of each term in (2.14) equals zero. Therefore, only the expected values of the power of each term are required.

The expected value of the power of the desired signal for ZF precoding in the Rayleigh fading channel can be obtained from (2.14) and (2.23) and expressed as

$$\mathbb{E}[|\tilde{\mathbf{h}}_k \mathbf{w}_k|^2] = \frac{1}{\gamma_{Ray}^{ZF}}. \tag{2.29}$$

The power normalisation in (2.24) is derived in [31] and given by

$$\mathbb{E}[\gamma_{Ray}^{ZF}] = \mathbb{E}[trace((\mathbf{H}_{Ray}\mathbf{H}_{Ray}^H)^{-1})]^{\ddagger\S} \overset{(a)}{=} K\mathbb{E}[(\mathbf{H}_{Ray}\mathbf{H}_{Ray}^H)_{k,k}^{-1}] = \frac{K}{M-K}, \tag{2.30}$$

where (a) holds since $\tilde{h}_{k,m}$ for $\forall k = \{1,\ldots,K\}, \forall m = \{1,\ldots,M\}$ are i.i.d. $\mathcal{CN}(0,1)$.

Therefore, the effective SINR (i.e. the effective SNR for ZF precoding) in the Rayleigh fading channel can be denoted from (2.18), (2.29) and (2.30) and given by

$$\text{SINR}_{Ray}^{ZF} = \rho(\frac{M}{K} - 1). \tag{2.31}$$

By substituting (2.31) into (2.20), the lower bound on the sum-rate capacity of ZF precoding in the Rayleigh fading channel for massive MIMO transmissions is obtained.

2.4.6 MmWave channel model

MmWave signals experience high pathloss that restricts the scattering effect. Meanwhile, the large-scale antenna array packed in a small platform increases the correlation between antennas. Consequently, the conventional statistical models of fading channel, e.g. Rayleigh fading or Rician fading that are typically utilised in sub-6 GHz massive MIMO transmissions, become unrealistic for modelling mmWave massive MIMO transmissions. Instead, the narrowband limited-scattering clustered channel model based on the extended Saleh-Valenzuela model is commonly adopted in mmWave MIMO transmissions

[‡] if $\mathbf{A}$ is invertible, $(\mathbf{A}^{-1})^H = (\mathbf{A}^H)^{-1}$

[§] $trace(\mathbf{A}^H\mathbf{A}) = trace(\mathbf{A}\mathbf{A}^H)$

[10, 20, 42]. For a MIMO transmission from an M-element transmit array to a K-element receive array, the limited scattering clustered mmWave channel is given by

$$\mathbf{H} = \sum_{l=1}^{L} \sum_{n_{path}=1}^{N_{path}} \alpha_{l,n_{path}} \Lambda_r(\theta^r_{l,n_{path}}, \varphi^r_{l,n_{path}}) \Lambda_t(\theta^t_{l,n_{path}}, \varphi^t_{l,n_{path}}) \mathbf{a}_r^H(\theta^r_{l,n_{path}}, \varphi^r_{l,n_{path}}) \mathbf{a}_t(\theta^t_{l,n_{path}}, \varphi^t_{l,n_{path}}), \tag{2.32}$$

where $\mathbf{H} \in \mathcal{C}^{K \times M}$ denotes the discrete-time narrow-band mmWave channel that is contributed by L clusters of paths with N_{path} paths in each cluster. $\alpha_{l,n_{path}}$ denotes the random complex path gain of the n_{path}-th path in the l-th cluster, which takes into account the maximum Doppler shift, the AoA of the l-th path relative to the k-th UT's direction of motion, and the omnidirectional pathloss. $\varphi^r_{l,n_{path}}, \varphi^t_{l,n_{path}}$ denote the receive and transmit (AoA/AoD) azimuth angles, while $\theta^r_{l,n_{path}}, \theta^t_{l,n_{path}}$ denote the receive and transmit elevation angles. $\Lambda_r(\theta^r_{l,n_{path}}, \varphi^r_{l,n_{path}})$ and $\Lambda_t(\theta^t_{l,n_{path}}, \varphi^t_{l,n_{path}})$ represent the receive and transmit antenna gains with respect to the AoA and AoD. $\mathbf{a}_r(\theta^r_{l,n_{path}}, \varphi^r_{l,n_{path}}) \in \mathcal{C}^{1 \times K}$ and $\mathbf{a}_t(\theta^t_{l,n_{path}}, \varphi^t_{l,n_{path}}) \in \mathcal{C}^{1 \times M}$ are the array response vectors with respect to the AoA and AoD of each path. For 2-D transmissions from a ULA, the array response vector is indicated in (2.9), whilst for 3-D transmissions from a UPA, it is given by (2.10).

Clearly, the mmWave channel has distinct and common characteristics compared with the sub-6 GHz fading and LoS channels. The array response vector is the common component that appears in both the sub-6 GHz LoS channel and the mmWave channel. The high pathloss of mmWave propagation results in the mmWave channel model, which embodies the limited spatial selectivity, i.e. limited-scattering clustered effect. When applying mmWave transmissions in massive MIMO systems, such a limited-scattering clustered channel property must be taken into consideration to achieve a performance enhancement as well as meet the signal processing constraints.

2.5 Summary

Massive MIMO and mmWave techniques have attracted much research attention as they enable great performance enhancement, and are compelling for 5G networks to meet the increasing demands of the network capacity along with the spectral efficiency.

With the deployment of a large-scale antenna array at the BS, low-complexity linear processing methods enable near optimal system performance at sub-6 GHz spectrum bands in massive MIMO systems. Random matrix theory indicates that the effects of uncorrelated noise and small-scale fading are asymptotically eliminated as the number of antennas at the BS approaches infinity.

The less crowded mmWave bands from 28 GHz to 300 GHz allow the seamless integration of dense HetNets deployment and the placement of a large number of antennas on a small transceiver platform. The high pathloss of the mmWave signal and the tightly packed transmitter or receiver antenna array result in a limited-scattering clustered mmWave channel model. The signal processing for mmWave transmissions in massive MIMO systems faces unique challenges from the channel characteristic and hardware configuration. Consequently, the conventional digital baseband signal processing method is still infeasible for mmWave massive MIMO systems with todays technology.

In this chapter, massive MIMO and mmWave technologies are introduced along with the system limitations and challenges, followed by the formation of the massive MIMO system model. The sub-6 GHz massive MIMO channel models are presented and the linear precoding schemes, and the system performance in Rayleigh fading channel are reviewed followed by the introduction of the mmWave MIMO channel model.

In the upcoming chapter, the advantages of massive MIMO LoS transmissions are investigated, followed by the system performance evaluations and comparisons with the fading channel transmissions.

Chapter 3

Line-of-sight Transmission in Massive MIMO Systems

3.1 Introduction

It is learned from Chapter 2 that massive MIMO systems potentially enable orders of magnitude enhancement of both the spectrum and energy efficiencies compared with the conventional MIMO systems. As the number of antennas equipped at the BS $M \to \infty$, the effects of small-scale fading and uncorrelated noise vanish along with the transmitted energy per bit [9]. Linear algorithms in massive MIMO systems not only enable near optimal system performance, but also greatly reduce the signal processing complexity and hardware power consumption. These appealing advantages promote massive MIMO as one of the most promising techniques for 5G networks [1, 13, 20, 31, 55].

Massive MIMO achieves spatial multiplexing gain from the rich scattering propagation of the radio wave. The system performance also relies on the accuracy of the estimated CSI. As the number of BS antennas increases, the TDD operation mode is preferable in massive MIMO systems, whereby the instantaneous CSI is estimated at the BS through uplink pilot training [9, 11–13, 31]. Fewer training resources are required in TDD mode compared with FDD mode, in which the downlink CSI can only be acquired by the BS

through feedback [56]. The channel reciprocity in TDD mode enables the estimated CSI to be exploited for both downlink precoding and uplink detection.

However, the limited coherence time in TDD mode limits the length and, hence, the number of orthogonal pilot sequences [29]. Any pilot reuse among the adjacent cells causes pilot contamination in massive MIMO cellular systems. Moreover, pilot contamination will not vanish as $M \to \infty$ and significantly degrades the system performance. Consequently, it is considered one of the major challenges in massive MIMO systems [9, 32, 33].

The LoS CSI, channel covariance matrix and fading distribution, for example, are statistical characterisations of the channel. The corresponding information of the characterisation is called the statistical CSI, which mainly depends on the antenna configuration, transmission environment, and spatial locations of the UTs. The statistical CSI, therefore, varies relatively slowly compared with the instantaneous CSI of the fading channel. Consequently, the statistical CSI can be readily obtained through long-term feedback or averaging over channel samples [57], and has been utilised for the development of massive MIMO systems for various purposes [58–60]. In [61, 62], the statistical CSI is exploited by the analogue-precoding part of the proposed hybrid precoding algorithm, whereas in [63, 64], the grouping of UTs along with the reduction in inter-group interference is achieved based on the statistical CSI for microwave and mmWave FDD systems. However, for these approaches, the instantaneous CSI is still essential for the precoding.

When the instantaneous CSI is unavailable due to, for example, severe pilot contamination in cellular systems, precoding approaches relying on the instantaneous CSI fail to achieve the desired system performance [25].

There are some downlink precoding approaches that exploit the LoS CSI. In [65], a lower bound on the average signal-to-leakage-and-noise ratio of multi-user MIMO downlink systems in a Rician fading channel is derived. The effect of the LoS CSI on the sum-rate is studied, and a statistical-eigenmode space-division multiple-access scheme for the maximisation of the signal-to-leakage-and-noise ratio is proposed. It is demonstrated that as the power of the LoS component in the Rician fading channel becomes stronger, the achievable sum-rate of the proposed scheme approaches the performance of ZF precoding with a perfect CSI at the BS. A beamforming approach using the statistical LoS CSI is

proposed in [66] for 3-D transmissions from a UPA in FDD large-scale MIMO systems. The ergodic sum-rate is derived, by which the optimal beamforming vector is acquired for the maximisation of the sum-rate. It is clear that the statistical LoS CSI is promising for massive MIMO downlink precoding in transmission scenarios such as the small-cell back-haul [67, 68], indoor stadiums [69], and satellite communications [70, 71], wherein the LoS channel presents. Moreover, the directional antenna modulation technique can be deployed for LoS transmissions to achieve further improvement of the energy efficiency by replacing the digital baseband signal processing with the analogue RF-domain control [72–76].

Most of the current developments [65, 66] for LoS transmissions assume that, the LoS channel vectors of the UTs are orthogonal. Such assumptions fail to catch the LoS transmissions' characteristics when random UTs are served and are only valid with the assistance of user grouping or scheduling schemes, in which only the users that satisfy the orthogonal condition can be served.

Massive MIMO developments in the Rayleigh fading channel achieves its capacity upper bound with a high probability as $M \to \infty$, and $K \ll M$ [31, 77]. For massive MIMO transmissions through LoS channels, it is indicated using asymptotic analysis that when there are no AoDs overlapping, the asymptotically favourable propagation happens less often than it would in Rayleigh fading channels when M is finite [31, 77]. Consequently, the developments of massive MIMO based on the Rayleigh fading channel and the corresponding system performance cannot be applied directly in LoS massive MIMO transmissions due to the difference in the probability of experiencing the favourable channel condition. The study of LoS transmissions is still very limited yet vital, especially for transmissions with a moderately large-scale UPA equipped at the BS serving single-antenna UTs spatially located in 3-D space.

This chapter focuses on sub-6 GHz TDD-mode massive MIMO downlink LoS transmissions, in which the statistical LoS CSI is known at the BS and adopted for precoding [‡]. The low-complexity linear precoding schemes of MF and ZF are considered. Note that

[‡]Different with LoS transmissions, the transmissions investigated in Chapter 2, in which the CSI of Rayleigh fading channel is adopted are Rayleigh fading channel transmissions. If the CSI of Rician fading channel is adopted, the transmissions are Rician fading channel transmissions. Both transmissions are fading channel transmissions.

LoS transmissions can be executed in both the LoS channel and Rician fading channel. Thanks to the slow-varying AoDs of the UTs, not only can the statistical LoS CSI be easily obtained, but also the coherence time for the systems using the statistical LoS CSI is much longer than it is for those using the instantaneous CSI of the fading channel. In this chapter, the received signal through the NLoS channel component is treated as interference for LoS transmissions in the Rician fading channel. The effective SINRs of the precoding schemes are analysed using random matrix theory. The two LoS transmission scenarios introduced in section 2.4.2 are considered, including 2-D transmissions from a ULA and 3-D transmissions from a UPA. The arguments show that there is no closed-form expression for the effective SINR in LoS transmissions when M is finite and $K \ll M$, i.e. $\frac{M}{K} \ll \infty$. However, both effective SINRs of the MF and ZF precoding for LoS transmissions are primarily determined by the AoDs distribution of the UTs. The inter-user interference term in the effective SINR for the MF is given as a function of the AoDs distribution and is different between the two LoS transmission scenarios, i.e. the 2-D transmissions from a ULA and 3-D transmissions from a UPA. The power normalisation term for ZF in the effective SINR is obtained as a function of the singular value distribution of the LoS channel matrix.

Furthermore, the LoS channel characteristics that are related to the effective SINRs are investigated from two perspectives: channel mean-square cross-correlation and the distribution of squared ordered channel matrix singular values. The simulation results depict that both characteristics are determined by the number of BS antennas M and AoDs distribution when $\frac{M}{K} \ll \infty$. With a wider AoDs distribution range, the LoS channel mean-square cross-correlation can be reduced, and the singular values of the channel matrix will increase, which yields a higher effective SINR for LoS transmissions.

The lower bounds on the system sum-rate for MF and ZF precoding in LoS transmissions are then evaluated through theoretical analysis and simulations, by which the impact of system parameters, namely SNR, Rician κ-factor, and M is examined. Comparisons of system performance are made between the LoS transmissions and fading channel transmissions.

The results show that linear precoding schemes of MF and ZF for LoS transmissions in the Rician fading channel are robust. The system performance of LoS transmissions in the

Rician fading channel is improved with the increase of the LoS channel power or the enlargement of AoDs distribution range. For both LoS and fading channel transmissions, the scaling-up of the BS array or the increase of SNR level enhances the system performance.

3.2 System model and sub-6 GHz channel models

Consider single-cell sub-6 GHz TDD-mode massive MIMO downlink transmissions with the system model demonstrated in (2.4.1). Assume that, the perfect CSI of the channel is adopted for the precoding. The LoS channel matrix is denoted in (2.8), where the LoS channel vector for 2-D transmissions from a ULA is given by (2.9), and 3-D transmissions from a UPA is shown in (2.10). The Rayleigh fading channel matrix is given in (2.7), and Rician fading channel matrix is shown in (2.11).

3.3 MF and ZF precoding using LoS CSI and Rician fading CSI

In this section of the thesis, as the starting point of the research, massive MIMO linear precoding schemes for LoS transmissions, in which the statistical LoS CSI is adopted, are introduced followed by the precoding schemes for Rician fading transmissions, in which the instantaneous Rician fading CSI is adopted. For both cases, the construction of the MF and ZF precoding schemes is the same as in Rayleigh fading channel transmissions introduced in Chapter 2 section 2.4.4. Note that LoS transmissions can be executed when the channel is pure LoS or Rician fading. For the case of LoS transmissions in the Rician fading channel, the NLoS component of the channel results in extra interference to the UTs.

The notation of the precoding matrix $\mathbf{W}$ with subscripts of *LoS*, *Ray*, and *Ric* are used to denote the CSI utilisation for the precoding matrix, namely the statistical CSI of LoS channel, instantaneous CSI of Rayleigh fading channel, and instantaneous CSI of Rician fading channel. The superscripts of MF and ZF denote the methods of precoding, i.e. matched filter and zero forcing.

3.3.1 Precoding schemes for LoS transmissions

In LoS transmissions, the statistical CSI of the LoS channel shown in (2.8) is utilised by the precoders.

- MF precoding matrix using the LoS CSI can be denoted from (2.8) and given by

$$\begin{aligned}\mathbf{W}_{LoS}^{MF} &= \frac{1}{\sqrt{\gamma_{LoS}^{MF}}}\mathbf{H}_{LoS}^{H} \\ &= \frac{1}{\sqrt{\gamma_{LoS}^{MF}}}[\mathbf{a}_1^H,\ldots,\mathbf{a}_K^H],\end{aligned} \tag{3.1}$$

  where $\mathbf{H}_{LoS}$ is the LoS channel matrix, and γ_{LoS}^{MF} is the power normalisation factor that is denoted as

$$\gamma_{LoS}^{MF} = \mathbb{E}[trace(\mathbf{H}_{LoS}\mathbf{H}_{LoS}^{H})]. \tag{3.2}$$

- ZF precoding matrix using the LoS CSI can be denoted from (2.8) and expressed as

$$\mathbf{W}_{LoS}^{ZF} = \frac{1}{\sqrt{\gamma_{LoS}^{ZF}}}\mathbf{H}_{LoS}^{H}(\mathbf{H}_{LoS}\mathbf{H}_{LoS}^{H})^{-1}, \tag{3.3}$$

  where the power normalisation factor γ_{LoS}^{ZF} is given by

$$\gamma_{LoS}^{ZF} = \mathbb{E}[trace(\mathbf{H}_{LoS}^{H}(\mathbf{H}_{LoS}\mathbf{H}_{LoS}^{H})^{-1}(\mathbf{H}_{LoS}^{H}(\mathbf{H}_{LoS}\mathbf{H}_{LoS}^{H})^{-1})^{H})] = \mathbb{E}[trace((\mathbf{H}_{LoS}\mathbf{H}_{LoS}^{H})^{-1})] \tag{3.4}$$

3.3.2 Precoding schemes for Rician fading channel transmissions

In Rician fading channel transmissions, the instantaneous CSI of Rician fading channel shown in (2.11) is adopted by the precoders.

- MF precoding using the Rician fading CSI is expressed from (2.11) and given by

$$\mathbf{W}_{Ric}^{MF} = \frac{1}{\sqrt{(\gamma_{Ric}^{MF})}}\mathbf{H}_{Ric}^{H}, \tag{3.5}$$

where the power normalisation factor is denoted as

$$\gamma_{Ric}^{MF} = \mathbb{E}[trace(\mathbf{H}_{Ric}\mathbf{H}_{Ric}^{H})]. \tag{3.6}$$

- ZF precoding using the Rician fading CSI is given by

$$\mathbf{W}_{Ric}^{ZF} = \frac{1}{\sqrt{\gamma_{Ric}^{ZF}}}\mathbf{H}_{Ric}^{H}(\mathbf{H}_{Ric}\mathbf{H}_{Ric}^{H})^{-1}, \tag{3.7}$$

where the power normalisation factor can be expressed as

$$\gamma_{Ric}^{ZF} = \mathbb{E}[trace(\mathbf{H}_{Ric}^{H}(\mathbf{H}_{Ric}\mathbf{H}_{Ric}^{H})^{-1}(\mathbf{H}_{Ric}^{H}(\mathbf{H}_{Ric}\mathbf{H}_{Ric}^{H})^{-1})^{H})] = \mathbb{E}[trace((\mathbf{H}_{Ric}\mathbf{H}_{Ric}^{H})^{-1})]. \tag{3.8}$$

3.4 Effective SINR analysis for LoS transmissions

Following the effective SINR analysis for the Rayleigh fading channel transmissions in section 2.4.5, in this section, the theoretical effective SINRs for LoS transmissions in both the LoS channel and Rician fading channel are analysed and examined using random matrix theory.

3.4.1 Effective SINR for LoS transmissions in the LoS channel

For LoS transmissions in a LoS channel, the received signal is given by (2.14) with the precoding matrices and power normalisation factors of MF and ZF shown in (3.1), (3.2), (3.3) and (3.4), respectively. Note that the expected value of each term (the desired signal, interference and noise) in (2.14) all equal zero, the effective SINR for LoS transmissions in the LoS channel is, therefore, expressed as

$$\text{SINR}_{LoS} = \frac{\rho\mathbb{E}[|\mathbf{a}_k\mathbf{w}_k|^2]}{\rho\sum_{\substack{j=1,\\ j\neq k}}^{K}\mathbb{E}[|\mathbf{a}_k\mathbf{w}_j|^2]+1}, \tag{3.9}$$

where $\mathbf{a}_k$ denotes the LoS channel vector of the k-th UT that is denoted in (2.9) and (2.10) for both LoS transmission scenarios. The term $\mathbb{E}[|\mathbf{a}_k\mathbf{w}_k|^2]$ is the expected value of the power of the desired signal, and the term $\sum_{\substack{j=1,\\ j\neq k}}^{K}\mathbb{E}[|\mathbf{a}_k\mathbf{w}_j|^2]$ is the expected value of the power of the inter-user interference.

An asymptotic analysis for LoS transmissions in the LoS channel is made in [77] assuming $M \to \infty$, $K \ll M$. Such an assumption substantially fails to reveal the characteristics of the LoS channel in realistic system configurations. In this part of the thesis, the LoS channel property for a moderately large M is studied through the analysis of the effective SINRs without the asymptotic assumption. Hence, the analysis applies to LoS transmissions when M is finite and $K \leq M$.

- Effective SINR of MF precoding using the LoS CSI in the LoS channel

  From the MF precoding scheme expressed in (3.1) and (3.2), and the effective SINR shown in (3.9), the expected value of the power of the desired signal for the k-th UT when the transmission channel is LoS can be denoted as

$$\mathbb{E}[|\mathbf{a}_k\mathbf{w}_k|^2] = \mathbb{E}[|\frac{1}{\gamma_{LoS}^{MF}}\mathbf{a}_k\mathbf{a}_k^H|^2] = \frac{1}{\gamma_{LoS}^{MF}}M^2. \tag{3.10}$$

  The power normalisation factor is given from (3.2) and expressed as

$$\mathbb{E}[\gamma_{LoS}^{MF}] = \mathbb{E}[trace(\mathbf{H}_{LoS}\mathbf{H}_{LoS}^H)] = trace(\mathbf{H}_{LoS}\mathbf{H}_{LoS}^H) = MK. \tag{3.11}$$

  The expected value of the power of the inter-user interference from the signal intended for the j-th UT received by the k-th UT can be denoted as

$$\mathbb{E}[|\mathbf{a}_k\mathbf{w}_j|^2] = \frac{1}{\gamma_{LoS}^{MF}}\mathbb{E}[|\mathbf{a}_k\mathbf{a}_j^H|^2]. \tag{3.12}$$

  As illustrated in Chapter 2 section 2.4.2, LoS transmissions can be divided into two scenarios based on the configuration of the antenna array at the BS, i.e. 2-D transmissions from a ULA and 3-D transmissions from a UPA. Often, research studies assume that the theory of the 2-D transmissions from a ULA can be applied

directly to the 3-D transmissions from a UPA, which is not always true. Next, the inter-user interference of the two LoS transmission scenarios for each array type is illustrated considering $\frac{M}{K} \ll \infty$, which indicates the differences between the two scenarios.

– For 2-D transmissions from a ULA, the inter-user interference shown in (3.12) is obtained from (2.9) and (3.11) and expressed as

$$\frac{1}{\gamma_{LoS}^{MF}}\mathbb{E}[|\mathbf{a}_k\mathbf{a}_j^H|^2] = \frac{1}{MK}\mathbb{E}[|\sum_{m=1}^{M} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m-1)(\cos\varphi_k-\cos\varphi_j)}|^2]. \tag{3.13}$$

It can be seen from (3.13) that in 2-D transmissions from a ULA, the inter-user interference of MF precoding in LoS transmissions is a function of the AoDs of the spatially located UTs. With the knowledge of the distribution of the AoDs, (3.13) can be further denoted as

$$\frac{1}{\gamma_{LoS}^{MF}}\mathbb{E}[|\mathbf{a}_k\mathbf{a}_j^H|^2] = \frac{1}{MK}\int\int|\sum_{m=1}^{M} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m-1)(\cos\varphi_k+\cos\varphi_j)}|^2 f(\varphi_k,\varphi_j)d\varphi_k d\varphi_j, \tag{3.14}$$

where $f(\varphi_k,\varphi_j)$ denotes the joint probability density function (p.d.f.) on the azimuth AoDs φ_k and φ_j of the UTs. When UTs have independent geometrical distributions, $f(\varphi_k,\varphi_j) = f(\varphi_k)f(\varphi_j)$, where $f(\varphi_k)$ and $f(\varphi_j)$ are the p.d.f. on the k-th and the j-th UTs' azimuth AoDs. A variety of distributions have been proposed for the modelling of the AoDs' distributions, such as uniform [54, 77] and Laplacian [42] distributions.

– For 3-D transmissions from a UPA, the inter-user interference shown in (3.12) is obtained from (2.10) and (3.11) and given by

$$\frac{1}{\gamma_{LoS}^{MF}}\mathbb{E}[|\mathbf{a}_k\mathbf{a}_j^H|^2] = \frac{1}{MK}\mathbb{E}[|\sum_{m_x=1}^{M_x}\sum_{m_y=1}^{M_y} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(\mathcal{A}+\mathcal{B})}|^2], \tag{3.15}$$

where

$$\mathcal{A} = (m_x-1)\sin\theta_k\cos\varphi_k + (m_y-1)\sin\theta_k\sin\varphi_k, \tag{3.16}$$

and

$$\mathcal{B} = (m_x - 1)\sin\theta_j \cos\varphi_j + (m_y - 1)\sin\theta_j \sin\varphi_j. \tag{3.17}$$

Hence, (3.15) can be further denoted as

$$\frac{1}{\gamma_{LoS}^{MF}}\mathbb{E}[|\mathbf{a}_k\mathbf{a}_j^H|^2] = \frac{1}{MK}\int\int\int\int |\sum_{m_x=1}^{M_x}\sum_{m_y=1}^{M_y} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(\mathcal{A}+\mathcal{B})}|^2 f(\theta_k,\varphi_k,\theta_j,\varphi_j)d\theta_k d\varphi_k d\theta_j d\varphi_j, \tag{3.18}$$

where $f(\theta_k,\varphi_k,\theta_j,\varphi_j)$ denotes the joint p.d.f. on the elevation and azimuth AoDs θ_k, φ_k, θ_j and φ_j of the UTs. When the UTs are independently distributed, $f(\theta_k,\varphi_k,\theta_j,\varphi_j) = f(\theta_k)f(\varphi_k)f(\theta_j)f(\varphi_j)$ where $f(\theta_k)$, $f(\varphi_k)$, $f(\theta_j)$ and $f(\varphi_j)$ are the p.d.f. on the elevation and azimuth AoDs.

From the inter-user interference of MF precoding in LoS transmissions shown in (3.12), (3.14), and (3.18), when $\frac{M}{K} \ll \infty$, apart from the array configuration at the BS, i.e. the separation between the antenna elements d and the number of elements M, the inter-user interference is directly limited by the spatial locations of the UTs. The expected value of the power of the inter-user interference can be obtained from the p.d.f. on the AoDs. By comparing (3.14) and (3.18), it is evident that the structure of the inter-user interference expression is significantly different between the two LoS transmission scenarios. This results in difficulties in applying the analytical results obtained from one scenario to the other.

From (3.9),(3.10),(3.11),(3.12), and (3.14), the effective SINR of MF precoding in 2-D LoS transmissions from a ULA can be obtained from the p.d.f. on AoDs and given by

$$\begin{aligned}\mathrm{SINR}_{LoS}^{MF} &= \frac{\rho\frac{M}{K}}{\rho\frac{1}{KM}\sum_{j=1,j\neq k}^{K}\int\int|\sum_{m=1}^{M}e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m-1)(\cos\varphi_k+\cos\varphi_j)}|^2 f(\varphi_k,\varphi_j)d\varphi_k d\varphi_j + 1} \\ &\overset{(a)}{=} \frac{\rho\frac{M}{K}}{\rho\frac{K-1}{KM}\int\int|\sum_{m=1}^{M}e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m-1)(\cos\varphi_k+\cos\varphi_j)}|^2 f(\varphi_k)f(\varphi_j)d\varphi_k d\varphi_j + 1},\end{aligned} \tag{3.19}$$

where (a) holds when φ_k is i.i.d. with p.d.f. $f(\varphi_k)$ for $\forall k \in \{1,\ldots,K\}$.

For 3-D transmissions from a UPA, from (3.9),(3.10),(3.11),(3.12), and (3.18), the effective SINR is given by

$$\begin{aligned}\mathrm{SINR}_{LoS}^{MF} &= \frac{\rho\frac{M}{K}}{\rho\frac{1}{KM}\sum\limits_{j=1,j\neq k}^{K}\int\int\int\int|\sum\limits_{m_x=1}^{M_x}\sum\limits_{m_y=1}^{M_y}e^{-\mathbf{j}\frac{2\pi d}{\lambda}(\mathcal{A}+\mathcal{B})}|^2 f(\theta_k,\varphi_k,\theta_j,\varphi_j)d\theta_k d\varphi_k d\theta_j d\varphi_j+1} \\ &\overset{(a)}{=} \frac{\rho\frac{M}{K}}{\rho\frac{K-1}{KM}\int\int\int\int|\sum\limits_{m_x=1}^{M_x}\sum\limits_{m_y=1}^{M_y}e^{-\mathbf{j}\frac{2\pi d}{\lambda}(\mathcal{A}+\mathcal{B})}|^2 f(\theta_k)f(\varphi_k)f(\theta_j)f(\varphi_j)d\theta_k d\varphi_k d\theta_j d\varphi_j+1},\end{aligned} \tag{3.20}$$

where (a) holds when θ_k and φ_k are i.i.d. with p.d.f. $f(\theta_k)$ and $f(\varphi_k)$ for $\forall k \in \{1,\ldots,K\}$. $\mathcal{A}$ and $\mathcal{B}$ are given by (3.16) and (3.17).

- Effective SINR of ZF precoding using the LoS CSI in the LoS channel

  From the precoder expressed in (3.3), (3.4), and the effective SINR denoted in (3.9), the expected value of the desired signal power for the k-th UT is given by

$$\mathbb{E}[|\mathbf{a}_k\mathbf{w}_k|^2] = \frac{1}{\gamma_{LoS}^{ZF}}. \tag{3.21}$$

  Due to the elimination of the inter-user interference by the ZF precoder, the effective SINR in (3.9) becomes the effective SNR as the elimination of the inter-user interference term.

  For the power normalisation of ZF precoding using the LoS CSI, the following theorem reveals the relationship between the normalisation factor and the channel singular values, which can be used for the calculation of γ_{LoS}^{ZF} when the distribution of the channel matrix singular values is known. Note that this theorem applies to any type of CSI adoption, therefore the footnote of channel type is omitted.

  Theorem 1. *γ^{ZF} is the expected value of the sum of the inverse-square of the channel matrix singular values, and can be given by*

$$\gamma^{ZF} = \mathbb{E}[\sum_{k=1}^{K}\mu_k^2] = \mathbb{E}[\sum_{k=1}^{K}\frac{1}{\varsigma_k^2}], \tag{3.22}$$

where μ_k denotes the k-th singular value of $\mathbf{Z}$ with $\mathbf{Z} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}$, and ς_k denotes the k-th singular value of $\mathbf{H}$, where $\mathbf{H}$ denotes the CSI adopted by the ZF precoder.

Proof. The normalisation factor of ZF is given by

$$\gamma^{ZF} = \mathbb{E}[trace(\mathbf{Z}\mathbf{Z}^H)] = \mathbb{E}[trace(\mathbf{Z}^{\mathbf{H}}\mathbf{Z})], \tag{3.23}$$

where $rank(\mathbf{H}) = rank(\mathbf{H}\mathbf{H}^H) = K$ when there are no identical vectors in $\mathbf{H}$, i.e. $\mathbf{h}_k \neq \mathbf{h}_j$ for $\forall k, j \in \{1,\ldots,K\}, k \neq j$, hence $\mathbf{H}\mathbf{H}^H$ is invertible.[‡]

The conjugate transpose of $\mathbf{Z}$ is denoted as

$$\mathbf{Z}^H = (\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1})^H = ((\mathbf{H}\mathbf{H}^H)^{-1})^H(\mathbf{H}^H)^H, \tag{3.24}$$

with $((\mathbf{H}\mathbf{H}^H)^{-1})^H = ((\mathbf{H}\mathbf{H}^H)^H)^{-1}$[§], (3.24) can be further denoted by

$$\mathbf{Z}^H = ((\mathbf{H}\mathbf{H}^H)^H)^{-1}\mathbf{H} = (\mathbf{H}\mathbf{H}^H)^{-1}\mathbf{H}, \tag{3.25}$$

therefore

$$\mathbf{Z}^H\mathbf{Z} = (\mathbf{H}\mathbf{H}^H)^{-1}\mathbf{H}\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1} = (\mathbf{H}\mathbf{H}^H)^{-1}. \tag{3.26}$$

Hence, for the eigenvalues of $\mathbf{Z}^H\mathbf{Z}$ and $\mathbf{H}\mathbf{H}^H$,

$$\mu_k^2 = \frac{1}{\varsigma_k^2}, \tag{3.27}$$

where μ_k^2 is the k-th eigenvalue of $\mathbf{Z}^H\mathbf{Z}$ and ς_k^2 is the k-th eigenvalue of $\mathbf{H}\mathbf{H}^H$ with ς_k the k-th singular value of $\mathbf{H}$[¶].

Therefore, for ZF precoder, the power normalisation factor $\gamma^{ZF} = \mathbb{E}[trace(\mathbf{Z}\mathbf{Z}^H)]^{\dagger} = \mathbb{E}[\sum_{k=1}^{K}\mu_k^2] = \mathbb{E}[\sum_{k=1}^{K}\frac{1}{\varsigma_k^2}]$, i.e. the expected value of the sum of the inverse of the

[‡]If LoS CSI is adopted, this requires that there's no spatially overlapped UTs (i.e. $\theta_k \neq \theta_j, \varphi_k \neq \phi_j$, for $\forall k, j \in \{1,\ldots,K\}, k \neq j$).

[§]If $\mathbf{A}$ is invertible, $(\mathbf{A}^H)^{-1} = (\mathbf{A}^{-1})^H$

[¶]The square roots of the non-zero, positive eigenvalues of $\mathbf{A}\mathbf{A}^H$ or $\mathbf{A}^H\mathbf{A}$ are the singular values of $\mathbf{A}$.

[†]$trace(\mathbf{A}) = \sum_{n=1}^{N}\psi_n$ if $\mathbf{A}$ is $n \times n$ square matrix, and ψ_n is the n-th eigenvalue of $\mathbf{A}$.

eigenvalues of $\mathbf{H}\mathbf{H}^H$ or the sum of the inverse of the squared singular values of the CSI matrix. □

From Theorem 1, with the knowledge of the p.d.f. on the singular values of $\mathbf{H}$, the power normalisation factor of the ZF precoder can be calculated by

$$\gamma^{ZF} = \int \dots \int \sum_{k=1}^{K} \frac{1}{\varsigma_k^2} f(\varsigma_1, \dots, \varsigma_K) d\varsigma_1 \dots d\varsigma_K. \tag{3.28}$$

where $f(\varsigma_1, \dots, \varsigma_K)$ denotes the joint p.d.f. on the singular values.

Hence, from (3.21), (3.28), and (3.9), the effective SINR for ZF precoding LoS transmissions in the LoS channel can be given by

$$\text{SINR}_{LoS}^{ZF} = \frac{\rho}{\int \dots \int \sum_{k=1}^{K} \frac{1}{\varsigma_k^2} f(\varsigma_1, \dots, \varsigma_K) d\varsigma_1 \dots d\varsigma_K}. \tag{3.29}$$

Note that (3.28) applies to any invertible channel matrices. Therefore, (3.29) can also be used for fading transmissions analysis. For the Rayleigh fading channel, the joint p.d.f. on the eigenvalues of $\mathbf{H}_{Ray}\mathbf{H}_{Ray}^H$ is derived in [78–80]. The research on the joint singular or eigenvalues distribution for LoS channel matrices is very limited with respect to an arbitrary AoDs distribution, which yields no mathematical expression of the p.d.f. available. However, it can be obtained through measurements from channel realisations.

Although, (3.29) is not a closed-form expression, clearly, the only term that restricts the effective SINR calculation is the p.d.f. on the singular values. Since the power normalisation factor of the ZF precoder is the expected value of the sum of the inverse-square of the singular values, when the singular values have high probabilities of being small (the channel matrix condition [‡] has a high probability of being poor [79]), the effective SINR of the ZF precoder will be significantly decreased. To better understand this property, the following quantitative analysis can be implemented.

[‡]The Matrix condition number$= \varsigma_K^2/\varsigma_1^2$. The matrix is poorly conditioned when the matrix condition number is small.

For example, if the k-th ordered singular value is $\varsigma_k = 50.1$, the contribution of this value to the normalisation factor in (3.22) is $\frac{1}{\varsigma_k^2} = 0.0004$. If the k-th ordered singular value increases by 50 and becomes $\varsigma_k = 100.1$, it yields approximately zero contribution to the normalisation factor. However, if it decreases by 50 and becomes $\varsigma_k = 0.1$, it yields the contribution to be $\frac{1}{\varsigma_k^2} = 100$, and the normalisation factor increases by 100. Therefore, when the singular values tend to have high probabilities of being small, the power normalisation factor of the precoding will be greatly increased.

From (3.9), the increase of the normalisation factor γ_{ZF} indicates that the precoder requires more power to perform interference elimination. Consequently, with the same total transmit power at the BS, the effective SINR decreases, which yields a worse system performance. In LoS transmissions, such degradation on the system performance arises when UTs are likely to be overlapped, which results in a higher probability of having a poor channel condition.

3.4.2 Effective SINR for LoS transmissions in the Rician fading channel

For a Rician fading channel, precoding with the statistical LoS CSI enables a longer coherence time for the transmissions compared with the adoption of the instantaneous Rician fading CSI. However, the downlink received signal at a UT suffers from extra intra-user interference from the NLoS channel component. Moreover, the inter-user interference received by a UT includes the transmitted signal intended for other UTs via both LoS and NLoS paths, namely LoS and NLoS components of the inter-user interference. For MF and ZF precoders using the LoS CSI only, the received signal of the k-th UT in the Rician fading channel is denoted from (2.6) and given by

$$y_k = \underbrace{\rho \frac{\kappa}{1+\kappa} \mathbf{a}_k \mathbf{w}_k x_k}_{\text{Desired Signal}} + \underbrace{\rho \frac{1}{1+\kappa} \tilde{\mathbf{h}}_k \mathbf{w}_k x_k}_{\text{Intra-user Interference}} + \underbrace{\rho \frac{\kappa}{1+\kappa} \sum_{\substack{j=1, \\ j \neq k}}^{K} \mathbf{a}_k \mathbf{w}_j x_j}_{\text{LoS Component of Inter-user Interference}} + \underbrace{\rho \frac{1}{1+\kappa} \sum_{\substack{j=1, \\ j \neq k}}^{K} \tilde{\mathbf{h}}_k \mathbf{w}_j x_j}_{\text{NLoS Component of Inter-user Interference}} + \underbrace{n_k}_{\text{Noise}}, \tag{3.30}$$

where the Rician fading channel $\mathbf{w}_k$ is given by (2.11). Consequently, the effective SINR for LoS transmissions in the Rician fading channel is given by

$$\text{SINR}_{Ric} = \frac{\rho \frac{\kappa}{1+\kappa} \mathbb{E}[|\mathbf{a}_k \mathbf{w}_k|^2]}{\rho \left(\frac{1}{1+\kappa} \mathbb{E}[|\tilde{\mathbf{h}}_k \mathbf{w}_k|^2] + \frac{\kappa}{1+\kappa} \sum_{\substack{j=1, \\ j \neq k}}^{K} \mathbb{E}[|\mathbf{a}_k \mathbf{w}_j|^2] + \frac{1}{1+\kappa} \sum_{\substack{j=1, \\ j \neq k}}^{K} \mathbb{E}[|\tilde{\mathbf{h}}_k \mathbf{w}_j|^2] \right) + 1}. \tag{3.31}$$

Next, the analysis of theoretical effective SINRs is provided. Since if the NLoS component occupies the non-dominant position in a Rician fading channel compared with the LoS component, the intra-user interference and NLoS inter-user interference received via NLoS channel is relatively small. They can be derived assuming $M \to \infty$. Such an asymptotic analysis is more accurate when the LoS component becomes stronger, i.e. κ increases, and $M \to \infty$.

- Effective SINR of MF precoding using the LoS CSI in the Rician fading channel

  The analysis of the LoS component of the inter-user interference is the same with the derivation of it in the LoS channel as illustrated in (3.12).

  From (3.30), both the intra-user and NLoS component of the inter-user interferences are the MF precoded signals received via independent Rayleigh fading channels. Therefore, the same analysis can be applied to both terms. The derivation of the intra-user interference is given in the following sequel.

From (3.31), the expected power of the intra-user interference via the NLoS channel vector of $\tilde{\mathbf{h}}_k$ is given by

$$\mathbb{E}[|\tilde{\mathbf{h}}_k \mathbf{w}_k|^2] = \frac{1}{\gamma_{LoS}^{MF}} \mathbb{E}[|\tilde{\mathbf{h}}_k \mathbf{a}_k^H|^2] = \frac{1}{\gamma_{LoS}^{MF}} \mathbb{E}[|\sum_{m=1}^{M} \tilde{h}_{k,m} \mathrm{a}_{k,m}^*|^2] \overset{(a)}{=} \frac{1}{\gamma_{LoS}^{MF}} M \tag{3.32}$$

where, (a) holds since the LoS and NLoS channels are independent, and the elements $\mathrm{a}_{k,m}$ for $\forall m = \{1,\ldots,M\}$ of the MF precoding vector $\mathbf{a}_k$ do not colour the distribution of the NLoS Rayleigh fading channel elements as $M \to \infty$, therefore, $\tilde{h}_{k,m}\mathrm{a}_{k,m}^* \sim \mathcal{CN}(0,1)$ for $\forall k \in \{1,\ldots,K\}, \forall m \in \{1,\ldots,M\}$.

Hence, for MF precoding in the Rician fading channel using the LoS CSI, the effective SINR can be derived from (3.31),(3.10),(3.11),(3.32),(3.12), and (3.14) for 2-D transmissions from a ULA and is given by

$$\begin{aligned}
&\mathrm{SINR}_{Ric}^{MF} \\
&= \frac{\rho \frac{\kappa}{\kappa+1} \frac{M}{K}}{\rho \frac{1}{\kappa+1} + \rho \frac{\kappa}{\kappa+1} \frac{1}{KM} \sum_{j=1, j\neq k}^{K} \int\int |\sum_{m=1}^{M} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m-1)(\cos\varphi_k + \cos\varphi_j)}|^2 f(\varphi_k,\varphi_j) d\varphi_k d\varphi_j + 1} \\
&\overset{(a)}{=} \frac{\rho \frac{\kappa}{\kappa+1} \frac{M}{K}}{\rho \frac{1}{\kappa+1} + \rho \frac{\kappa}{\kappa+1} \frac{K-1}{KM} \int\int |\sum_{m=1}^{M} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m-1)(\cos\varphi_k + \cos\varphi_j)}|^2 f(\varphi_k) f(\varphi_j) d\varphi_k d\varphi_j + 1},
\end{aligned} \tag{3.33}$$

where (a) holds when φ_k is i.i.d. with p.d.f. $f(\varphi_k)$ for $\forall k \in \{1,\ldots,K\}$.

For 3-D transmissions from a UPA, from (3.31),(3.10),(3.11),(3.32),(3.12), and (3.18), the effective SINR can be given by

$$\begin{aligned}
&\mathrm{SINR}_{Ric}^{MF} \\
&= \frac{\rho \frac{\kappa}{\kappa+1} \frac{M}{K}}{\rho \frac{1}{\kappa+1} + \rho \frac{\kappa}{\kappa+1} \frac{1}{KM} \sum\limits_{j=1, j\neq k}^{K} \int\int\int\int |\sum\limits_{m_x=1}^{M_x} \sum\limits_{m_y=1}^{M_y} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(\mathcal{A}+\mathcal{B})}|^2 f(\theta_k,\varphi_k,\theta_j,\varphi_j) d\theta_k d\varphi_k d\theta_j d\varphi_j + 1} \\
&\overset{(a)}{=} \frac{\rho \frac{\kappa}{\kappa+1} \frac{M}{K}}{\rho \frac{1}{\kappa+1} + \rho \frac{\kappa}{\kappa+1} \frac{K-1}{KM} \int\int\int\int |\sum\limits_{m_x=1}^{M_x} \sum\limits_{m_y=1}^{M_y} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(\mathcal{A}+\mathcal{B})}|^2 f(\theta_k) f(\varphi_k) f(\theta_j) f(\varphi_j) d\theta_k d\varphi_k d\theta_j d\varphi_j + 1},
\end{aligned} \tag{3.34}$$

where $\mathcal{A}$ and $\mathcal{B}$ are denoted in (3.16) and (3.17), and (a) holds when θ_k and φ_k are i.i.d. with p.d.f. $f(\theta_k)$ and $f(\varphi_k)$ for $\forall k \in \{1,\ldots,K\}$.

- Effective SINR of ZF using the LoS CSI in the Rician fading channel

In the effective SINR of ZF precoding denoted in (3.30), the LoS component in the inter-user interference is eliminated leaving the intra-user interference term and the NLoS component in the inter-user interference term. The remaining interference terms both arise from ZF precoded signals received via the independent Rayleigh fading channels. Therefore, the same analysis can be implemented. In the following sequel, the derivation of the intra-user interference term is given.

To obtain the expected value of the power of the intra-user interference $\mathbb{E}[|\tilde{\mathbf{h}}_k \mathbf{w}_k|^2]$, it is feasible to obtain the variances of the independent elements of $\tilde{h}_{k,m}$ and $w_{m,k}$, respectively. Since $\tilde{h}_{k,m} \sim$i.i.d.$\mathcal{CN}(0,1)$, the variance of the ZF precoder element $w_{m,k}$ is required that can be obtained from the desired signal power as demonstrated in the following content.

As $M \to \infty$, for each element $w_{m,k}$ of the ZF precoding matrix $\mathbf{W}_{LoS}^{ZF}$, $variance(w_{m,k})$ $\to 0$. Therefore, for each UT, the desired received signal components from each antenna element at the BS are approximately the same, i.e. $\mathrm{a}_{k,m} w_{m,k} \to \mathrm{a}_{k,n} w_{n,k}$ for $\forall m,n \in \{1,\ldots,M\}, m \neq n$. This results in $\mathbb{E}[|\mathbf{a}_k \mathbf{w}_k|^2] \to \mathbb{E}[|M \mathrm{a}_{k,m} w_{m,k}|^2]$. Consequently,

$$\mathbb{E}[|\mathbf{a}_k \mathbf{w}_k|^2] \overset{(a)}{=} \frac{1}{\gamma_{LoS}^{ZF}} \to \mathbb{E}[|M \mathrm{a}_{k,m} w_{m,k}|^2] = M^2 \mathbb{E}[\mathrm{a}_{k,m} w_{m,k}|^2] \overset{(b)}{=} M^2 \mathbb{E}[|w_{m,k}|^2], \quad (3.35)$$

where (a) is from (3.21), and (b) holds since $\mathrm{a}_{k,m} \sim \mathcal{C}(0,1)$ for $\forall k \in \{1,\ldots,K\}, \forall m \in \{1,\ldots,M\}$. Hence, by rearranging (3.35), the variance of the elements of ZF precoding matrix when the LoS CSI is adopted is given by

$$\mathbb{E}[|w_{m,k}|^2] = \frac{1}{\gamma_{LoS}^{ZF}} \frac{1}{M^2}. \quad (3.36)$$

Therefore, the expected value of the power of NLoS intra-user interference at the receiver in (3.31) is obtained from (3.36) and given by

$$\mathbb{E}[|\tilde{\mathbf{h}}_k \mathbf{w}_k|^2] \overset{(a)}{=} M \mathbb{E}[|w_{m,k}|^2] = M \times \frac{1}{\gamma_{LoS}^{ZF}} \frac{1}{M^2} = \frac{1}{\gamma_{LoS}^{ZF}} \frac{1}{M}, \quad (3.37)$$

where (a) holds since $\tilde{h}_{k,m}$ is independent from $w_{m,k}$, and $\tilde{h}_{k,m} \sim$i.i.d.$\mathcal{CN}(0,1)$. Hence, for ZF in the Rician fading channel exploiting the LoS CSI, the effective SINR can be obtained from (3.31),(3.21),(3.28), and (3.37) and given by

$$\text{SINR}_{Ric}^{ZF} = \frac{\rho \frac{\kappa}{\kappa+1}}{\rho \frac{1}{\kappa+1}\frac{K-1}{M} + \gamma_{LoS}^{ZF}} \tag{3.38}$$

where γ_{LoS}^{ZF} is given by (3.28).

From (3.19),(3.20),(3.29),(3.33),(3.34), and (3.38), it is evident that there is no closed-form expression readily available for the effective SINRs of MF and ZF precoders using the LoS CSI in LoS and Rician fading channels when $\frac{M}{K} \ll \infty$. System parameters of SNR, M, K and Rician κ-factor impact on the effective SINR when M is moderately large. Moreover, the LoS inter-user interference of the MF using the LoS CSI is proved to be determined by the UTs spatial distribution, and hence can be obtained from the p.d.f. on AoDs. The normalisation factor of ZF is constrained by the singular values of the channel matrix, and hence can be obtained from the p.d.f. on the singular values of the LoS channel matrix. Although there's no solution convenient for the quantification of the impact of the AoDs distribution on the singular values' p.d.f., it has been noticed that the AoDs of the UTs directly limit the performance of the ZF precoder through the power normalisation factor.

3.5 Simulation results for LoS and fading channel transmissions

In the first part of this section, the characteristics of the LoS channel are investigated. The channel mean-square cross-correlation that restricts the effective SINR of MF precoding is defined and examined. The singular values' distribution that limits the ZF precoding effective SINR is also demonstrated. Monte Carlo simulations are implemented for the experiments. The impact of the scale of the BS array and AoDs' distribution is studied. The Rayleigh fading channel is also examined for comparison.

In the second part of this section, the theoretical and simulated system performance of the lower bound on the sum-rate capacity for LoS transmissions and fading channel transmissions are examined and compared. Due to the fact that there's no closed-form expressions for the theoretical results, the numerical analysis is used to quantify the theoretical results. The impact of system parameters, such as SNR, Rician κ-factor and M are studied.

3.5.1 Channel mean-square cross-correlation

Define the channel mean-square cross-correlation for $\forall k, j \in \{1, \ldots, K\}$, and $k \neq j$ [81–84] as

$$c \triangleq \mathbb{E}[|\mathbf{h}_k \mathbf{h}_j^H|^2]. \tag{3.39}$$

By comparing the effective SINRs of MF precoder using the LoS CSI denoted in (3.19) and (3.20) with it using the Rayleigh fading CSI denoted in (2.28), the only difference is the inter-user interference term. Hence, the mean-square cross-correlation defined in (3.39) can be examined to demonstrate the property of the channel as well as the MF precoder performance when the corresponding CSI is used.

The LoS channel mean-square cross-correlation can be denoted as

$$c_{LoS} = \mathbb{E}[|\mathbf{a}_k \mathbf{a}_j^H|^2]. \tag{3.40}$$

From (3.12), the LoS channel mean-square cross-correlation is the inter-user interference for the MF precoder using the LoS CSI in the LoS channel normalised by the factor of $\frac{1}{\gamma_{LoS}^{MF}} = \frac{1}{MK}$ as derived in (3.14) and (3.18) for 2-D transmissions from a ULA and 3-D transmissions from a UPA.

For the Rayleigh fading channel, the mean-square cross-correlation is

$$c_{Ray} = \mathbb{E}[|\tilde{\mathbf{h}}_k \tilde{\mathbf{h}}_j^H|^2]. \tag{3.41}$$

From (2.26), the Rayleigh fading channel mean-square cross-correlation is the inter-user interference of MF in Rayleigh fading transmissions normalised by the factor of $\frac{1}{\gamma_{Ray}^{MF}} = \frac{1}{MK}$.

Table 3.1 2-D configured UPA.

M	$=$	M_x	$\times$	M_y
16		4		4
32		8		4
64		8		8
128		16		8
256		16		16
512		32		16
1024		32		32

In the following investigations, 3-D transmissions from an M-element UPA are considered. The configurations of the UPA used in simulations are shown in Table 3.1. The LoS channel is modelled considering the AoDs of the UTs θ_k, φ_k are uniformly distributed with the elevation angle $\theta \sim$i.i.d.$\mathbf{U}[0^\circ, \theta_{max}^\circ]$ and the azimuth angle $\varphi \sim$i.i.d.$\mathbf{U}[-180^\circ, 180^\circ]$ for all k. Hence, θ_{max} indicates the distribution range of the AoDs. 10^4 incidences of the channel realisations are simulated. The mean-square cross-correlations of the LoS channel and Rayleigh fading channel are calculated by (3.40) and (3.41), and normalised by a factor of M. Note that the assumptions adopted here differ from the assumption made in [31, 77] that assumed 2-D transmissions from a ULA at the BS with $\sin\varphi \sim$i.i.d.$\mathbf{U}[-1, 1]$ for all k. Such an assumption is inappropriate to reveal the characteristics of LoS channel in 3-D transmissions from the randomness of AoDs perspective since i.i.d. random AoDs distribution does not yield uniform distribution on $\sin\varphi$.

Fig. 3.1 illustrates the relationship between normalised mean-square cross-correlation and M. From Fig. 3.1, the normalised Rayleigh fading channel mean-square cross-correlation remains constant (=1) for any value of M, which is confirmed in (2.26). The LoS channel is much more correlated compared with the Rayleigh fading channel. The LoS channel mean-square cross-correlation increases with M, especially when θ_{max} is small, i.e. UTs are more likely to be spatially close to each other.

Fig. 3.2 shows the impact of θ_{max} on the normalised mean-square cross-correlation. The mean-square cross-correlation of the Rayleigh fading channel is unity for any M as shown in Fig. 3.1. Therefore, the average value of unity is plotted in Fig. 3.2. It

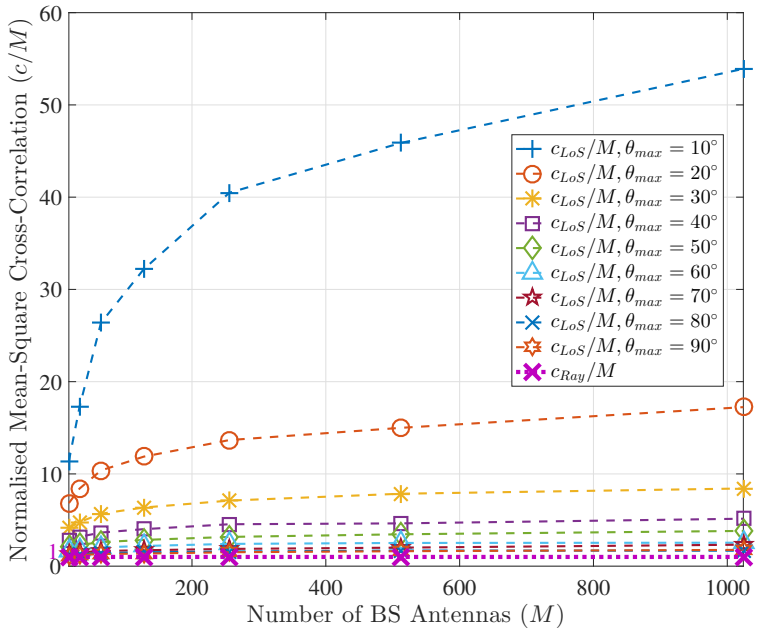


Fig. 3.1 Normalised channel mean-square cross-correlation vs number of BS antennas (M). $\theta \sim \mathbf{U}[0°, \theta^{\circ}_{max}]$ and $\varphi \sim \mathbf{U}[-180°, 180°]$ are considered with $f = 2.6$ GHz, $d = \lambda/2$. 10^4 channel incidences.

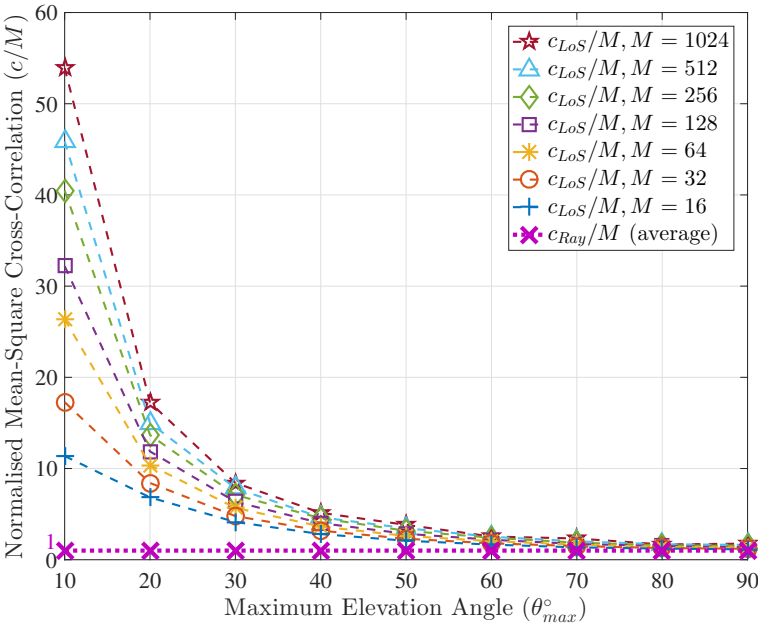


Fig. 3.2 Normalised channel mean-square cross-correlation vs θ_{max}. θ_{max} denotes the range of the AoDs distribution. $\theta \sim \mathbf{U}[0°, \theta^{\circ}_{max}]$ and $\varphi \sim \mathbf{U}[-180°, 180°]$ are considered with $f = 2.6$ GHz, $d = \lambda/2$. 10^4 channel incidences.

is shown that the normalised LoS channel mean-square cross-correlation decreases and approaches unity when θ_{max} increases. For a common BS coverage of $\theta_{max} = 60°$ [31],

the mean-square cross-correlation of LoS channel is severe, e.g. when $M = 128$, the LoS channel mean-square cross-correlation is doubled compared with that of the Rayleigh fading channel.

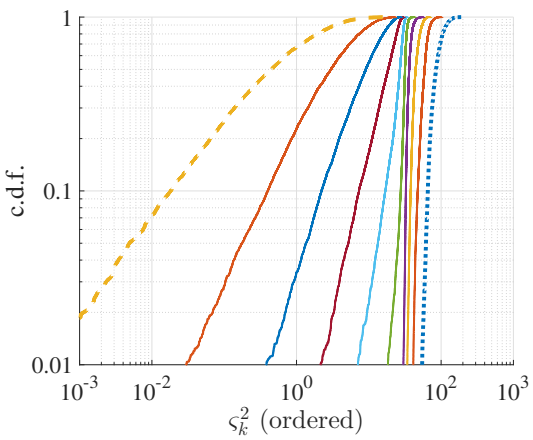
Fig. 3.1 and Fig. 3.2 show that MF precoding performs worse in LoS transmissions than it would in the Rayleigh fading channel transmissions due to the higher mean-square cross-correlation of the LoS channel. The mean-square cross-correlation of the LoS channel is reduced when the AoDs are distributed within a wider range.

3.5.2 The distribution of channel matrix singular values

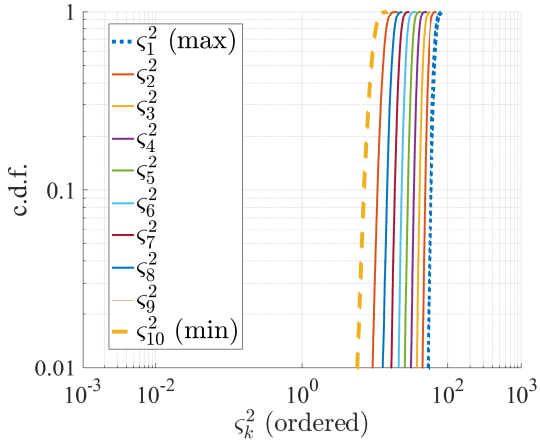
It is indicated in section 3.4.1 that singular values of the channel matrix determine the effective SINR of the ZF precoder via the normalisation factor. To further ascertain the impact of M and distribution of AoDs on the singular values distribution, the cumulative distribution function (c.d.f.) on each ordered squared singular value ς_k for $k = 1, \ldots, K$ with $\varsigma_1^2 \leq \cdots \leq \varsigma_K^2$ is investigated in this section. 3-D transmissions from a UPA to $K = 10$ single antenna UTs are considered. For LoS channel matrices, the AoDs follow uniform distributions with $\theta \sim \mathbf{U}[0^\circ, \theta_{max}]$ and $\varphi \sim \mathbf{U}[-180^\circ, 180^\circ]$ for all k. 10^4 channel realisations are simulated. Therefore 10^4 groups of ordered squared singular values are obtained. The c.d.f. on each ordered squared singular value is hence examined. Comparisons are made between the LoS channel and Rayleigh fading channel.

Fig. 3.3a & 3.3b, Fig. 3.4a & Fig. 3.4b, and Fig. 3.5a & 3.5b present the c.d.f. on the ordered squared singular values of the channel matrix for the LoS channel with $\theta_{max} = 50^\circ$ and the Rayleigh fading channel for $M = 32, 128$, and 512. It is illustrated that all of the singular values increase with M. As a consequence, the power normalisation factor of the ZF precoder decreases when M increases. Therefore, when the total transmit power at the BS is limited, ZF precoding performs better with the scaling up of the antenna array at the BS for both cases of transmissions in the LoS channel and Rayleigh fading channel.

Fig. 3.6a, Fig. 3.4a, and Fig. 3.6b show the c.d.f. on $K = 10$ ordered squared singular values of the LoS channel matrices when $M = 128$ for different AoDs distributions, i.e. $\theta_{max} = 10^\circ, 50^\circ$ and 90°. It is shown that the LoS channel is likely to have small singular values along with a large value spread $(\varsigma_K^2 - \varsigma_1^2)$. As θ_{max} decreases, i.e. the range of

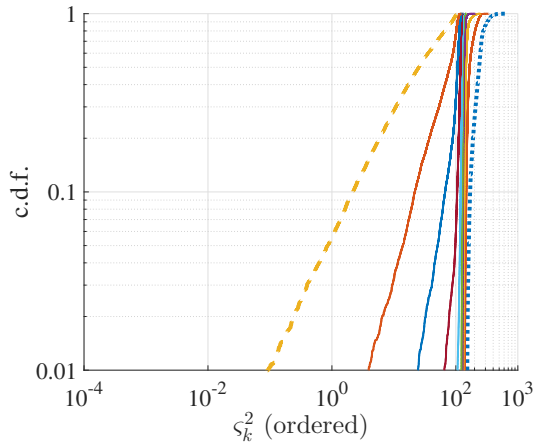


(a) LoS channel. $\theta_{max} = 50°$.

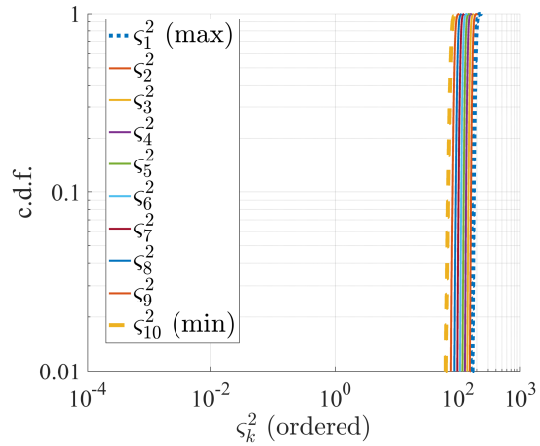


(b) Rayleigh fading channel.

Fig. 3.3 Cumulative distribution of ordered squared singular values ς_k^2, for $k = 1, \ldots, K$ of (a) LoS channel matrices with $\theta_{max} = 50°$ and (b) Rayleigh fading channel matrices. $M = 8 \times 4 = 32, K = 10, f = 2.6$ GHz, $d = \lambda/2$. 10^4 channel incidences.



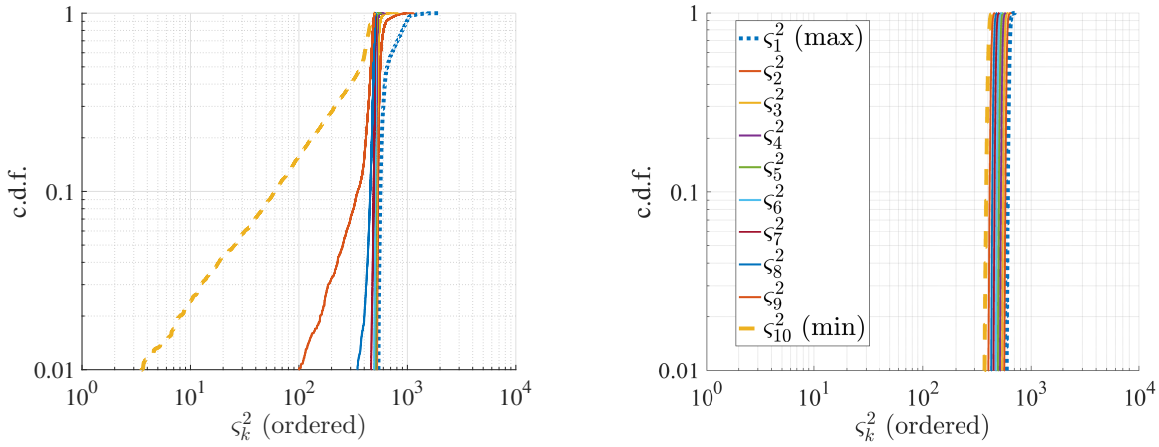
(a) LoS channel. $\theta_{max} = 50°$.



(b) Rayleigh fading channel

Fig. 3.4 Cumulative distribution of ordered squared singular values ς_k^2, for $k = 1, \ldots, K$ of (a) LoS channel matrices with $\theta_{max} = 50°$ and (b) Rayleigh fading channel matrices. $M = 16 \times 8 = 128, K = 10, f = 2.6$ GHz, $d = \lambda/2$. 10^4 channel incidences.

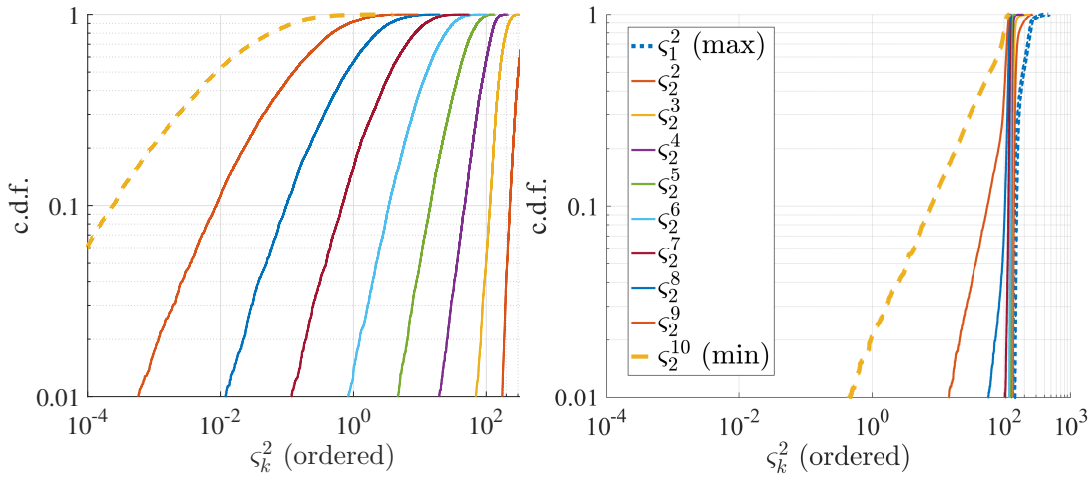
AoDs distribution narrows, the values of the smaller ordered singular values decrease significantly. For example, in Fig. 3.6a for $\theta_{max} = 10°$, the leftmost curve (the c.d.f. on the smallest squared singular value ς_1^2 distribution) indicates that the smallest squared singular value is less than 1 with a 100% probability. However, the leftmost curve of Fig. 3.4a shows that there is a 5% probability for the smallest squared singular value to be less than 1 when $\theta_{max} = 50°$. When $\theta_{max} = 90°$, there is only a 2% probability as shown by the leftmost curve in Fig. 3.6b. For all of the figures, there's no significant variation of



(a) LoS channel. $\theta_{max} = 50°$.

(b) Rayleigh fading channel.

Fig. 3.5 Cumulative distribution on ordered squared singular values ς_k^2, for $k = 1,\ldots,K$, of (a) LoS channel matrices with $\theta_{max} = 50°$ and (b) Rayleigh fading channel matrices. $M = M_x \times M_y = 32 \times 16 = 512, K = 10, f = 2.6$ GHz, $d = \lambda/2$. 10^4 channel incidences.



(a) LoS channel. $\theta_{max} = 10°$.

(b) LoS channel. $\theta_{max} = 90°$.

Fig. 3.6 Cumulative distribution on ordered squared singular values ς_k^2, for $k = 1,\ldots,K$, of LoS channel matrices with (a) $\theta_{max} = 10°$ and (b) $\theta_{max} = 90°$. $M = M_x \times M_y = 16 \times 8 = 128, K = 10, f = 2.6$ GHz, $d = \lambda/2$. 10^4 channel incidences.

the c.d.f. on the maximum squared singular value ς_K^2 (the rightmost curve of each figure), which is around 10^2.

Fig. 3.4b illustrates the c.d.f. on ς_k^2 for Rayleigh fading channel. It is obvious that the curves have uniform spread around 10^2 with little value spread.

Therefore, compared with the Rayleigh fading channel matrix, the LoS channel matrix is more likely to have higher singular values, especially when AoDs have a wider distribution range, i.e. the UTs are more likely to be spatially away from each other. This illustrates

that ZF precoding performs worse in the LoS channel than it would in the Rayleigh fading channel as analysed in section 3.4.1.

In general, from investigations of the mean-square cross-correlation and c.d.f. on the ordered squared singular values for LoS and Rayleigh fading channels, it can be concluded that the performance of MF and ZF in LoS transmissions is highly limited by M and the distribution of AoDs. A wider range of AoDs distribution improves the percoders' performance by decreasing the inter-user interference for MF and normalisation factor for ZF.

3.5.3 System performance of LoS transmissions and fading channel transmissions

In this section, the performance of the lower bound on the system sum-rate capacity for LoS transmissions and fading channel transmissions are examined through Monte Carlo simulations to further study the effectiveness of MF and ZF precoding in the Rician fading channel when the statistical LoS CSI is utilised. The impact of the system parameters, i.e. SNR, Rician κ-factor and M on system performance is also analysed.

A single-cell massive MIMO system is considered assuming the BS is equipped with an $M = M_x \times M_y$-element UPA, and K single antenna UTs are served. For LoS transmissions and Rician fading channel transmissions, 3-D transmissions with $\varphi \sim$i.i.d.$\mathbf{U}[-180^\circ, 180^\circ]$ and $\theta \sim$i.i.d.$\mathbf{U}[0, 90^\circ]$ are considered. The carrier frequency f is 2.6 GHz and the separation between adjacent BS antenna elements d is $\frac{1}{2}\lambda$. For simplicity but without loss of generality, assume that the UTs experience the same large-scale fading that is normalised to unity, and the same Rician κ-factor.

Numerical analysis is used to approximate the theoretical results which are not closed-form expressions. The system sum-rate capacity lower bounds are obtained by (2.20), in which the numerical effective SINRs for LoS transmissions in LoS channel and Rician fading channel are calculated by (3.20),(3.34) for MF precoding, and (3.29),(3.38) for ZF precoding. Since the expected value of the power of the inter-user interference for MF precoding utilising the LoS CSI does not yield a closed-form expression, the simulation results from Fig. 3.1 are adopted in the numerical effective SINR results. As for the ZF

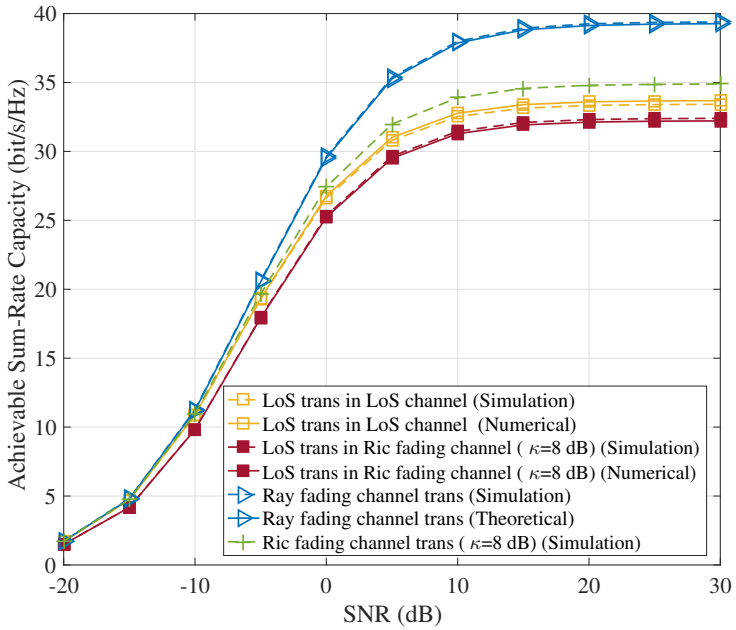


Fig. 3.7 Lower bounds on the sum-rate capacities of MF precoding vs SNR (dB) for LoS and fading channel transmissions. 3-D transmissions from a UPA at the BS are considered with $M = M_x \times M_y = 16 \times 8 = 128, K = 10, f = 2.6$ GHz, $d = \lambda/2$.

precoding utilising the LoS CSI, the measurements from simulation results are adopted for the power normalisation factor calculation in numerical effective SINRs. For Rayleigh fading channel transmissions, the theoretical effective SINR for MF is calculated by (2.28) and for ZF is by (2.31). The simulated results of the effective SINRs for LoS transmissions in the LoS channel are calculated by (3.9), for Rayleigh fading channel transmissions is calculated by (2.18), and for Rician fading channel transmissions is calculated by the similar method as shown in (3.9).

Fig. 3.7 illustrates the sum-rate capacity lower bounds of MF precoding versus the SNR (dB) for LoS and fading channel transmissions. For the Rician fading channel, $\kappa = 8$ dB is considered. It can be seen that for LoS and Rayleigh fading channel transmissions, the numerical results (LoS transmissions) and theoretical result (Rayleigh fading channel transmissions) agree closely with the simulated results. All of the curves show a similar shape with no crossover. Lower bounds increase logarithmically with SNR (dB) and converge to different non-zero limits due to the existence of inter-user interference. Rayleigh fading transmissions perform the best, which confirms that the Rayleigh fading channel has the lowest mean-square cross-correlation. The precoding in the Rician fading channel

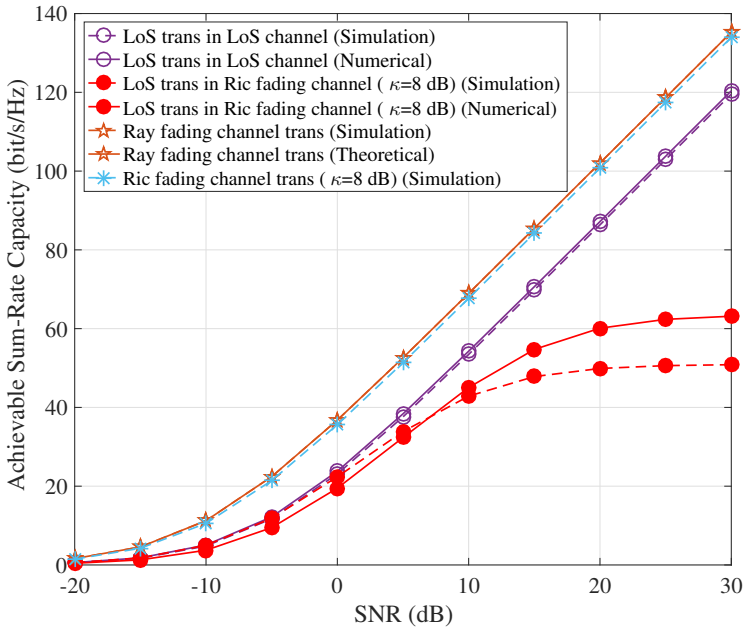


Fig. 3.8 Lower bounds on the sum-rate capacities of ZF precoding vs SNR (dB) in LoS and fading channel transmissions. 3-D transmissions from a UPA are considered with $M = M_x \times M_y = 16 \times 8 = 128, K = 10, f = 2.6$ GHz, $d = \lambda/2$.

using the LoS CSI yields a close performance to it using the Rician fading CSI with a minor penalty (maximum 2 bit/s/Hz). Therefore, the interference arising from the NLoS component in the Rician fading channel has little impact on the system performance for LoS transmissions as SNR increases.

Fig. 3.8 illustrates the sum-rate capacity lower bound performance of ZF precoding for the same transmissions in Fig. 3.7. The numerical results (LoS transmissions) and theoretical result (Rayleigh fading channel transmissions) agree closely with the simulated results, except for the case of using the LoS CSI in the Rician fading channel, especially in the high SNR region. This is due to the fact that, the asymptotic analysis adopted for the theoretical intra-user interference and NLoS component of inter-user interference is less accurate when M is moderately large. All of the transmissions show that the lower bounds grow without limit and increase linearly at the high SNR region, except the LoS transmissions in the Rician fading channel, due to the presence of interference. The performance in the LoS channel is within a gap of 4 dB from the fading channel transmissions. The lower bound of the LoS transmissions in the Rician fading channel only approximates the LoS channel performance in the low SNR region. As SNR increases,

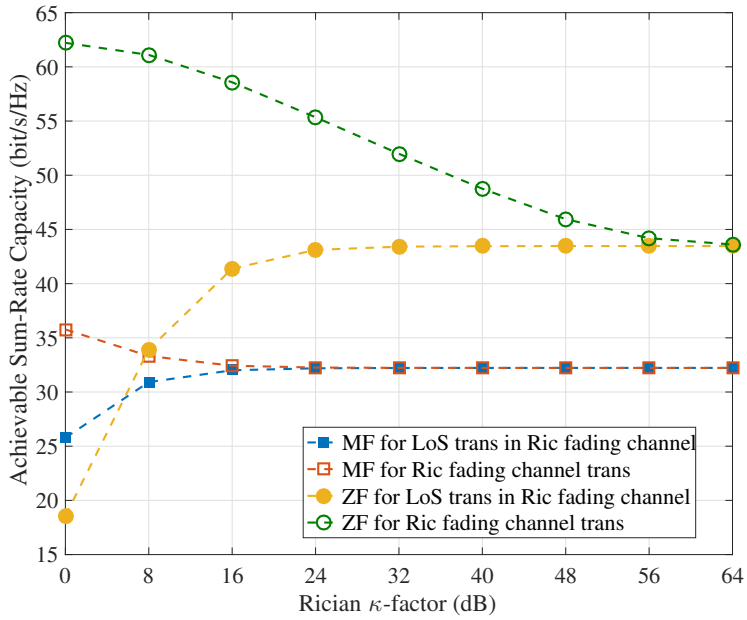


Fig. 3.9 Lower bounds on the sum-rate capacities of the MF and ZF vs Rician κ-factor in LoS and Rician fading channel transmissions when the channel is Rician fading. 3-D transmissions from a UPA at the BS are considered with $M = M_x \times M_y = 16 \times 8 = 128, K = 10$, SNR=8 dB, $\theta_k \sim \mathbf{U}[0^\circ, 90^\circ], \varphi_k \sim \mathbf{U}[-180^\circ, 180^\circ], f = 2.6$ GHz, $d = \lambda/2$.

the interference from the NLoS component seriously harms the lower bound that yields a maximum rate of 50 bit/s/Hz. This indicates that the NLoS interference is detrimental to ZF precoding for LoS transmissions in the Rician fading channel, especially in the high SNR region.

To further examine the impact of Rician κ-factor on the system performance of LoS transmissions in the Rician fading channel, Fig. 3.9 shows the simulation results of the sum-rate lower bound performance versus the Rician κ-factor for LoS transmissions and Rician fading channel transmissions in the Rician fading channel. SNR=8 dB is considered. It can be seen that the lower bounds of MF and ZF decrease to non-zero limits when κ (dB) increases when the Rician fading CSI is adopted. This is due to the increase of channel mean-square cross-correlation with the reinforcing of the LoS channel component. However, when the LoS CSI is adopted, the lower bounds of MF and ZF increase with κ to non-zero limits. The maximum lower bound of MF using the LoS CSI equals the minimum lower bound for the Rician fading CSI. The ZF precoder shows a similar trend.

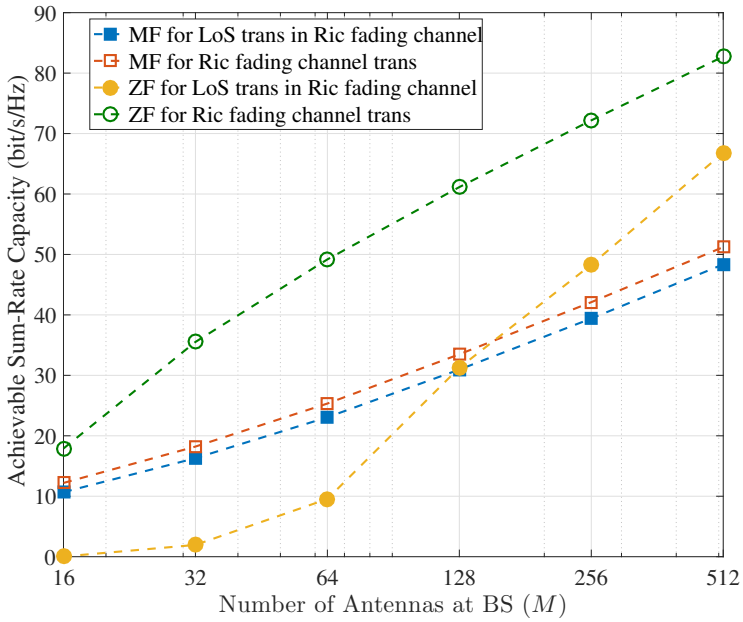


Fig. 3.10 Lower bounds on the sum-rate capacities of MF and ZF vs number of antennas at the BS in LoS and Rician fading channel transmissions when the channel is Rician fading. 3-D transmissions from a UPA at the BS are considered with $K = 10$, SNR=8 dB, κ=8 dB, $\theta_k \sim \mathbf{U}[0^\circ, 90^\circ], \varphi_k \sim \mathbf{U}[-180^\circ, 180^\circ], f = 2.6$ GHz, $d = \lambda/2$..

Therefore, the system performance of both MF and ZF using the LoS CSI is robust in the Rician fading channel when the power of LoS component is high.

Next, the impact of the number of antennas deployed at the BS on the system performance for LoS transmissions and Rician fading channel transmissions is examined. The simulation results of the sum-rate lower bounds against the number of antennas at the BS are indicated in Fig. 3.10. The configurations of the UPA are taken from Table 3.1. SNR=8 dB and κ=8 dB are considered. It can be seen that all of the lower bounds grow with M without limits. For MF precoding, using the LoS CSI yields a very close performance to the case of using the Rician fading CSI, and both lower bounds increase linearly with M. As for ZF precoding, the performance of LoS transmissions only approaches the performance of Rician fading channel transmissions in the large M region. This indicates that ZF precoding for LoS transmissions in the Rician fading channel performs well when M is large. By comparing the performance of MF and ZF in LoS transmissions, MF outperforms ZF when $M \leq 128$, however, the advantage of MF diminishes quickly as M increases. After M=128, the lower bound of ZF exceeds that of MF and has a higher rate.

Therefore, when only the statistical LoS CSI is adopted, MF is more appropriate when the scale of the array at the BS is not so large. ZF performs better when there are more antennas at the BS.

3.6 Summary

The massive MIMO technique provides orders of magnitude performance enhancement of both the spectral efficiency and energy efficiency with low-complexity linear processing and is considered to be one of the key techniques for 5G networks. Most of the current research for sub-6 GHz massive MIMO focuses on the transmissions in rich scattering environments modelled as a Rayleigh fading channel.

This chapter focuses on massive MIMO LoS transmissions, in which the statistical LoS CSI is adopted for precoding instead of the instantaneous fading CSI. The adoption of LoS CSI benefits the system, as it enables a longer coherence time and is straightforward to estimate.

The lower bounds on the sum-rate capacities for MF and ZF precoding schemes in LoS transmissions have been derived and evaluated through effective SINRs. It is illustrated that when $\frac{M}{K} \ll \infty$, there is no closed-form expression available for the effective SINR and the lower bound. However, the inter-user interference of MF precoding is given as a function of p.d.f. on AoDs, and the power normalisation factor for ZF precoding is provided as a function of p.d.f. on the channel matrices' singular values.

The LoS channel characteristics are evaluated through mean-square cross-correlation and singular values distribution. The number of antennas at the BS and the AoDs distribution determine the system performance for LoS transmissions. A wider AoDs distribution range yields a better system performance.

The lower bounds on the sum-rate capacities for LoS transmissions using MF and ZF are evaluated through theoretical (approximated by numerical results) and simulation results and compared with the lower bounds for fading channel transmissions. It is shown that LoS transmissions in the Rician fading channel are robust. For MF precoding, the adoption of the LoS CSI operates closely to the adoption of the Rician fading CSI when

the LoS component is strong. For ZF precoding, the performance of the LoS CSI adoption tends to approach the performance of the Rician CSI adoption in the low SNR region with high Rician κ-factor when the array-scale is large. For LoS and fading channel transmissions, both the scaling up of the array and the increase of SNR improve the system performance. For LoS transmissions, a wider range of AoDs distribution or a higher Rician κ-factor enhances the system performance.

At this point in the thesis, the investigation of the sub-6 GHz massive MIMO LoS transmissions has been demonstrated and compared with the fading channel transmissions. In the upcoming chapter, the focus is on the mmWave massive MIMO transmissions at 28 GHz-300 GHz. Note that the array response vector is the common component of the sub-6 GHz LoS channel in (2.8) and the limited-scattering clustered mmWave channel in (2.32). Besides, due to the high pathloss, mmWave transmissions show the limited spatial selectivity. The similarities and differences between sub-6 GHz LoS and mmWave transmissions prompt the motivation for research into the mmWave massive MIMO transmissions.

Chapter 4

MmWave Massive MIMO Precoding Schemes and Hardware Configurations

4.1 Introduction

Massive MIMO is a key technique for 5G wireless communications and refers to the principle of deploying a large-scale antenna array at the cellular BS to serve tens of UTs simultaneously [13]. Massive MIMO is both spectral and energy efficient with several orders of magnitude improvement over conventional MIMO techniques as illustrated in Chapter 3. MmWave communications exploit signals with higher frequencies from 28 GHz to 300 GHz than the microwave signals with frequencies under 6 GHz used in most of the current commercial wireless networks. MmWave transmissions are emerging as the solution to meet the current exploding bandwidth and data traffic demands. The application of mmWave transmissions in massive MIMO systems is appealing since the mmWave spectrum bands are less crowded, and the reduction in wavelength allows a large-scale antenna configuration to be packed into a small transceiver platform [20, 21]. In conventional massive MIMO transmissions at sub-6 GHz frequencies, the signals often exhibit statistically i.i.d. Rayleigh fading channels. However, severe pathloss is experienced at mmWave frequencies, which leads to sparse channels regarding multipath propagation [20] and the tightly packed array gives rise to high channel correlation. Although simple

linear precoding algorithms, e.g. MF and ZF, are proven to yield near optimal system performance with low complexity [12] in sub-6 GHz massive MIMO transmissions, full digital baseband processing requires an exclusive RF chain to be connected to each antenna to achieve both amplitude and phase control. Such a hardware architecture leads to high cost and power consumption that is undesirable yet circumventable when applied to mmWave massive MIMO systems [20, 44].

Schemes that exploit the RF-domain control have drawn attention to the signal processing design for mmWave point-to-point MIMO (P2P MIMO) transmissions, wherein a BS serves only one UT at a time, and both the BS and UT are equipped with multiple antennas. The analogue beam-forming technique developed for P2P MIMO systems only involves phase control of a constant envelop signal on each element of the transmit antenna via a network of phase shifters in the RF domain instead of using digital baseband control [45, 46, 66, 85–91]. The beamspace MIMO method is utilised for analogue beam-forming [45, 46, 66, 85, 86] where discrete Fourier transform (DFT) vectors serve as the orthonormal basis for the spatial signal space and used to beam-form the signal to the subspace of the actual channel, which captures most of the channel power. In [87–91], a codebook beamtraining method is exploited, by which the transmitter and receiver jointly select their analogue beam-forming vectors aiming to acquire the strongest received signal. The codebook with multiple resolutions is designed to measure iteratively the spatial signal power. The codes in the codebook are used for steering beams with different beamwidths to specific directions. Step by step, codes are selected to make beamwidths more narrow in order to search for the direction of the strongest signal. Extra processing for the channel estimation is no longer needed since beamforming vectors are found during the beamtraining phase. In [91], for a mmWave MIMO link, it is shown that as the number of transmit and receive antennas grows large, the optimal solution from singular value decomposition (SVD) precoding and combining using the unconstrained baseband signal processing, converges to the solution of using beam steering vectors at the transmitter and receiver. Hence, the exhaustive or iterative searching for the optimal steering vectors is not necessary under the limitation of phase-only control and predefined directions. The analogue codebook beamtraining P2P MIMO beam-forming method has already been

implemented in commercial standards such as IEEE 802.11ad [39, 92], IEEE 802.15.3c [38] and wireless HD [36] for single data-stream transmissions. Analogue beam-forming allows low-complexity signal processing along with low consumption on fabrication and energy. However, the performance is limited by the absence of signal amplitude control and the additional constraint from the available phase quantisation for the beam steering.

The hybrid precoding technique combines analogue control in the RF domain with low-dimensional digital control in the baseband [10, 42, 44, 64, 93]. The number of RF chains required is much less than the number of antennas. In [42], a mmWave large-scale P2P MIMO hybrid precoding scheme focusing on the single or multiple streams linear array is proposed assuming the CSI is known at both the transmitter and receiver. By SVD of the channel matrix, the precoder and combiner ensure the maximum capacity using the orthogonal matching pursuit (OMP) algorithm. The OMP approach iteratively searches for steering vectors along which the optimal precoder has the maximum projection and then adds this vector to the analogue precoder. The baseband precoder is calculated through the least squares approach. After removing the discovered vectors from the optimal precoder, the residual precoder becomes the updated optimal precoder that guides the next round of searching until the digital and analogue precoding matrices are formed. This method is not realisable in massive MIMO due to the lack of cooperation between the UTs. In [93], a limited feedback precoding scheme is studied for a multi-user MIMO system, in which each UT is equipped with multiple antennas. In the first stage, the codebook RF precoding is used to maximise the precoder and combiner SNR between the BS and UTs, neglecting other users' interference to noise ratio. In the second stage, the effective channel per user is estimated, quantised and fed back. Subsequently, the BS uses ZF to eliminate the inter-user interference. The requirement for channel estimation and feedback at the UTs in this two-stage process adds complexity to the system. More importantly, the quantisation of the effective channel limits the system performance. A joint spatial division and multiplexing scheme for mmWave channels is proposed in [64], which groups users with similar covariance matrices within each group. The RF precoding is applied to users within the same group to form beam patterns, and the digital precoding is applied to users between groups to cancel the interference. However, the number of

UTs per group restricts the system performance plus the number of data streams supported is limited. The superiority of the hybrid precoding approach to the analogue precoding approach is embodied in the extra degrees of freedom added by the digital control. Such degrees of freedom can be exploited, for example, to cancel the multi-stream or multi-user interference, without raising the fabrication cost and signal processing complexity to prohibitive levels. The trade-off between energy efficiency and spectrum efficiency of hybrid beam-forming for large-scale antenna systems is studied in [44]. There it is indicated that system complexity and hardware architecture are of great importance when designing the mmWave massive MIMO hybrid precoding algorithms.

Many mmWave MIMO channel estimation techniques have been developed for P2P MIMO systems focusing on the recovery of the overall channel from the estimation of each path. For example, compressive sensing techniques enable the estimation of the complex path gain, the AoD and AoA of each path when reconstructing a clustered mmWave channel [10, 94, 95]. In [94], the transmit beam pattern and receive combining vectors need to be designed using multiple channel measurements, which depend on the number of paths presented in the limited-scattering clustered mmWave channel, and each training state relies on the result of the previous training stage. Such channel estimation algorithms sequentially estimate each users' channel one after the other at both the UT and BS for multi-user MIMO systems. However, for TDD-mode massive MIMO systems with single-antenna UTs, the sequential estimation of each UT's channel is not feasible when the system scales up. The available coherence time may be completely used up when estimating the channels for all received paths for each UT sequentially. Many channel estimation algorithms for TDD-mode massive MIMO have been developed to obtain the complex channel coefficients $\mathbf{H}$ [32–34]. For a mmWave channel, each complex channel coefficient consists of a small number of individual complex path gains that can be estimated separately from $\mathbf{H}$.

This chapter focuses on mmWave massive MIMO downlink transmissions and aims to provide a low-complexity precoding algorithm that can be executed by low-cost and energy-efficient signal processing. The limited-scattering clustered channel model is adopted, which characterises the sparse nature of the mmWave channel. TDD-mode massive

MIMO systems are considered, in which the CSI contributed by all of the mmWave paths is acquired via the uplink pilot training process, and the CSI is assumed to be known at the BS. The large-scale multi-user precoding complexity and the mmWave signal processing constraints are jointly addressed with the exploitation of the limited-scattering clustered characteristic of the mmWave channel. The main contributions of this chapter is summarised below.

- Develop a novel channel deconstruction algorithm for TDD-mode mmWave massive MIMO systems. The limited-scattering clustered property of the mmWave channel and the asymptotic orthogonality of the massive MIMO channel are jointly exploited in the design. The proposed algorithm enables the BS to estimate the array response and the complex path gain from the CSI for each contributive mmWave path for each UT.

- Analysis of the effectiveness of the proposed channel deconstruction approach on the precise estimation of the path parameters with the scaling up of the array at the BS, as well as the immunity to the Gaussian noise presented in the imperfect CSI.

- Developing analogue and hybrid beam-selecting precoding algorithms for multi-user mmWave massive MIMO transmissions. The proposed algorithms aim to achieve a robust system performance while at the same time they can be implemented with low signal processing complexity, less energy consumption, and low fabrication cost with respect to the conventional massive MIMO signal processing method. To this end, the mmWave precoder design is transformed into a mmWave beam-selecting problem, and hence the mmWave massive MIMO analogue and hybrid beam-selecting (ABS and HBS) precoding approaches are proposed that employ the strongest path parameters estimated by the proposed channel deconstruction method.

- Designs of the system hardware architecture for the proposed analogue beam-selecting and hybrid beam-selecting precoding approaches. Both designs exhibit low signal processing complexity and far fewer RF chains than the conventional massive MIMO signal processing. The number of required RF chains for both designs is

independent of the number of BS antennas, which is vital with the scaling up of the system.

- Performance evaluation of the proposed ABS and HBS. The analysis of theoretical effective SINRs and lower bounds on the sum-rate capacities are completed using random matrix theory and ordered statistics principles.

The simulation results show that the proposed channel deconstruction scheme achieves the accurate estimation of the parameters for the strongest path with the scaling up of the BS from both the perfect CSI and Gaussian-perturbed CSI. The system performance of sum-rate lower bounds for the proposed ABS and HBS precoders are evaluated by simulations and compared with the optimal performance of the MF precoder exploiting the unconstrained digital baseband signal processing. Simulation results indicate that the selection of the strongest path by the proposed precoders substantially contributes to the system performance, due to the limited-scattering clustered nature of the mmWave channel and the large number of antennas deployed at the BS. The performance of both of the proposed precoders is mainly limited by the number of received paths. However, both proposed precoders maintain the performance over a large range of the CSI imperfection factor.

4.2 System Model

The massive MIMO single-cell downlink system model introduced in Chapter 2 section 2.4.1 is considered, in which an M-element antenna array is equipped at the BS and communicating with K single-antenna UTs exploiting mmWave signals.

4.3 MmWave limited-scattering clustered channel model for massive MIMO systems

MmWave signals experience strong free-space pathloss that limits the scattering effect. Hence, the narrowband limited-scattering clustered channel model, introduced in (2.32), is

considered based on the extended Saleh-Valenzuela model and is commonly exploited for the mmWave massive MIMO channel modelling [20, 42, 93].

The channel vector from the M-element array at the BS to the k-th single-antenna UT $\mathbf{h}_k \in \mathcal{C}^{1\times M}$ is considered to be composed from L clusters of paths and there are N_{path} paths within each cluster as illustrated in (2.32). For simplicity, assume that each cluster only contributes a single path to the channel, i.e. $N_{path} = 1$. Also, consider that, the antennas at the BS and UTs are isotropic antennas, which set the antenna gain functions to unity, i.e. $\Lambda_r(\theta^r_{l,n_{path}}, \varphi^r_{l,n_{path}}) = 1$ and $\Lambda_t(\theta^t_{l,n_{path}}, \varphi^t_{l,n_{path}}) = 1$ for any AoAs and AoDs [20, 42, 93]. For single-antenna UTs, the receive antenna response at the k-th UT becomes unity, i.e. $\mathbf{a}_r(\theta^r_{l,n_{path}}, \varphi^r_{l,n_{path}}) = 1$. Therefore, the mmWave massive MIMO narrowband channel vector from the BS to the k-th single-antenna UT can be obtained from (2.32) and given by

$$\mathbf{h}_k = \sum_{l=1}^{L} \alpha^l_k \mathbf{a}^l_k(\theta^l_k, \varphi^l_k) \tag{4.1}$$

where L denotes the number of received paths or clusters, and $\mathbf{h}_k \in \mathcal{C}^{1\times M}$, α^l_k denotes the complex gain of the path in the l-th cluster, i.e. the l-th path of the k-th UT. Consider that $\alpha^l_k \sim$i.i.d.$\mathcal{CN}(0, \frac{1}{L})$, i.e. each path has equal average path power. α^l_k takes into account the maximum Doppler shift, the AoA of the l-th path relative to the k-th UT's direction of motion and the omnidirectional pathloss [96]. The elevation and azimuth AoDs of the l-th path of the k-th UT are θ^l_k, φ^l_k and $\mathbf{a}^l_k(\theta^l_k, \varphi^l_k) \in \mathcal{C}^{1\times M}$ is the corresponding transmit antenna array response vector for the k-th UT. Therefore, for a BS equipped with a ULA, the transmit antenna array response of the l-th path of the k-th UT in 2-D transmissions can be denoted from (2.9) and given by

$$\mathbf{a}^l_k(\varphi^l_k) = \left[1, e^{-\mathbf{j}\frac{2\pi}{\lambda}d\cos(\varphi^l_k)}, \ldots, e^{-\mathbf{j}(M-1)\frac{2\pi}{\lambda}d\cos(\varphi^l_k)}\right], \tag{4.2}$$

where λ is the carrier wavelength and d is the distance between the adjacent antenna elements. For 3-D transmissions from a UPA, the transmit antenna array response vector

for the l-th path is give from (2.10) by

$$\begin{aligned}\mathbf{a}_k^l\left(\theta_k^l,\varphi_k^l\right) = \Big[1,..,e^{-\mathbf{j}\frac{2\pi}{\lambda}d\left((m_x-1)\sin(\theta_k^l)\cos(\varphi_k^l)+(m_y-1)\sin(\theta_k^l)\sin(\varphi_k^l)\right)},\\ \ldots,e^{-\mathbf{j}\frac{2\pi}{\lambda}d\left((M_x-1)\sin(\theta_k^l)\cos(\varphi_k^l)+(M_y-1)\sin(\theta_k^l)\sin(\varphi_k^l)\right)}\Big],\end{aligned} \tag{4.3}$$

where $1 \leq m_x \leq M_x$, $1 \leq m_y \leq M_y$ are the indices of the antenna elements on the x and y dimensions, respectively, M_x and M_y are the numbers of the antennas along the x and y dimensions, and $M = M_x \times M_y$.

Hence, the mmWave massive MIMO channel coefficients matrix $\mathbf{H} \in \mathcal{C}^{K\times M}$ from the BS to K UTs is denoted as

$$\mathbf{H} = [\mathbf{h}_1^T,\ldots,\mathbf{h}_K^T]^T, \tag{4.4}$$

where $\mathbb{E}[trace(\mathbf{H}\mathbf{H}^H)] = KM$ from (4.1).

4.4 Channel deconstruction algorithm

The previous section illustrated the channel model which incorporates the limited-scattering clustered property of the mmWave massive MIMO channel. The objective is to design a precoding algorithm for TDD-mode mmWave massive MIMO systems which can be implemented by low-complexity, low-cost and energy-efficient signal processing. Before proposing the precoding scheme, a channel deconstruction algorithm is firstly proposed in this section, which utilises the obtained CSI to obtain parameters for the proposed precoding schemes.

Consider a TDD-mode massive MIMO system, in which the instantaneous CSI of the mmWave channel for the multi-user downlink precoding process is acquired by the BS through pilot training. From (4.1) and (4.4), the CSI estimated by the BS is the complex channel coefficients contributed by all of the paths.

The channel deconstruction algorithm proposed in this section enables the estimation of the individual mmWave path parameters for each UT from the CSI, i.e. the complex path gains α_k^l and the corresponding AoDs θ_k^l, φ_k^l for $\forall l = \{1,\ldots,L\}$ and $\forall k = \{1,\ldots,K\}$. The proposed channel deconstruction algorithm exploits the limited-scattering clustered property of the mmWave channel as well as the asymptotic orthogonality of the massive

MIMO channel. Random matrix theory is used to analyse the effectiveness of the proposed algorithm assuming the number of BS antennas $M \to \infty$. The estimation from both the perfect CSI and Gaussian-perturbed CSI is considered. Finally, the estimation mean-square error is explored for the quantification of the estimation accuracy.

4.4.1 Channel deconstruction from perfect CSI

Consider that the CSI of the mmWave channel coefficients matrix $\mathbf{H}$ is perfectly known at the BS via ideal uplink training. 2-D transmissions from an M-element ULA are considered with the channel matrix $\mathbf{H}$ given by (4.2),(4.1) and (4.4) to simplify the formulation and derivation. Nonetheless, the following sequel may be applied to the 3-D transmissions from a planar array case too. Let $\mathrm{a}_{k,m}^{l}(\varphi_k^l) = e^{-\mathbf{j}\frac{2\pi}{\lambda}d(m-1)\cos(\varphi_k^l)}$ denote the m-th element of the transmit array response vector $\mathbf{a}_k^l$ of the l-th path for the k-th UT corresponding to the AoD of φ_k^l as shown in (4.2) that gives the linear BS array response vector of the l-th path of the k-th UT

$$\mathbf{a}_k^l(\varphi_k^l) = [1, \mathrm{a}_{k,2}^{l}(\varphi_k^l), \ldots, \mathrm{a}_{k,M}^{l}(\varphi_k^l)]. \tag{4.5}$$

From (4.1), the channel vector can be further denoted using (4.5) as

$$\mathbf{h}_k = [\sum_{l=1}^{L} \alpha_k^l, \sum_{l=1}^{L} \alpha_k^l \mathrm{a}_{k,2}^{l}(\varphi_k^l), \cdots, \sum_{l=1}^{L} \alpha_k^l \mathrm{a}_{k,M}^{l}(\varphi_k^l)]. \tag{4.6}$$

Consider a training matrix $\mathbf{T} \in \mathcal{C}^{N\times M}$ that is used by the BS to perform the path estimation from the measured CSI using a channel deconstruction algorithm [†]. The construction of the training matrix follows the array response vectors of the BS array to N potential spatial directions ω_n for $n = 1, \ldots, N$. Let the number of BS training beams be N with $N \gg L$. Then N non-overlapping training beam spatial directions $\omega_1, ..., \omega_N$ can be specified, which cover the whole distribution range of the AoDs φ_k^l for $\forall k, \forall l$. In particular, for uniform training, if the AoDs distribution range spans the entire 360° azimuth range, then $\omega_N - \omega_1 = 2\pi$. The relationship between N and the isolation of training directions

[†]The BS does not need to implement any transmission to UTs for the estimation. The estimation is through calculations using the training matrix $\mathbf{T}$ at the BS.

$\Delta\omega$ is $N = (2\pi/\Delta\omega) + 1$. Other training beam directions are possible, and for the case of 3-D transmissions from a planar array, non-uniform training isolation in azimuth and elevation domains are possible. Accordingly, the training matrix $\mathbf{T} \in \mathcal{C}^{N\times M}$ can be formed following array response vectors each specifying a training beam direction and is given by

$$\mathbf{T} = \left[\mathbf{t}_1^T, \ldots, \mathbf{t}_N^T\right]^T, \tag{4.7}$$

where the n-th training vector is $\mathbf{t}_n(\omega_n) = [1, \mathrm{a}_{n,2}(\omega_n), \ldots, \mathrm{a}_{n,M}(\omega_n)]$ and $\mathrm{a}_{n,m}(\omega_n) = e^{-\mathbf{j}\frac{2\pi}{\lambda}d(m-1)\cos(\omega_n)}$.

Assume that the perfect CSI is available at the BS. A response vector $r_k \in \mathcal{C}^{N\times 1}$ for the k-th UT can be obtained from the CSI of $\mathbf{h}_k$ and the training matrix $\mathbf{T}$ by

$$\mathbf{r}_k = \mathbf{T}\mathbf{h}_k^H, \tag{4.8}$$

where the n-th element $r_{k,n} = \mathbf{t}_n(\omega_n)\mathbf{h}_k^H$ of the response vector $\mathbf{r}_k$ corresponds to the spatial direction of ω_n, for $\forall n \in \{1, ..., N\}$. Then, the norm of $r_{k,n}$ with the highest value $|r_{k,n}|_{max}$, among all of the N norms of the elements of the response vector $\mathbf{r}_k$, indicates the direction of the strongest path between the BS and the k-th UT when $M \to \infty$. This outcome is then justified, which underpins the main contribution of this chapter, using random matrix theory and the proof is formulated in the following sequel.

Before the justification, the notations used are introduced. For a L-cluster mmWave channel (with a single path in each cluster) from the BS to the k-th UT, the unordered path parameters are introduced in (4.1). Sort the norms of the complex path gains α_k^l in descending order and denote the complex path gain of the l-th ordered path as $\alpha_k^{(l)}$, therefore, $|\alpha_k^{(L)}| \geq |\alpha_k^{(L-1)}| \geq \cdots \geq |\alpha_k^{(1)}|$. Let $\mathbf{a}_k^{(l)}$ and $\varphi_k^{(l)}$ denote the array response vector and AoD corresponding to the l-th ordered complex path gain $\alpha_k^{(l)}$, namely the l-th ordered array response vector and azimuth AoD. Note that for a mmWave path $\alpha_k^{(l)}\mathbf{a}_k^{(l)}$, the power of the path is given by $\alpha_k^{(l)}\mathbf{a}_k^{(l)}(\alpha_k^{(l)}\mathbf{a}_k^{(l)})^H = M|\alpha_k^{(l)}|^2$. Hence, the norm of the complex path gain directly corresponds to the power of the path.

Theorem 2. *The maximum entry of the norms of the response vector elements corresponds to the strongest mmWave path.*

Proof. An ULA or UPA transmitter with independent, continuously distributed AoDs will exhibit a row-wise orthogonal array response matrix as $M \to \infty$, and K is fixed [91]. Hence, the training matrix $\mathbf{T} \in \mathcal{C}^{N \times M}$ is also row-wise orthogonal. Consequently, $\alpha_k^{(l)*}\mathbf{t}_n(\omega_n)\mathbf{a}_k^{(l)*}(\varphi_k^{(l)}) \to 0$, for $\forall l = \{1,...,L\}$ when $M \to \infty$ and $\omega_n \neq \varphi_k^{(l)}$. When multiplying the training matrix $\mathbf{T}$ with the mmWave channel vector $\mathbf{h}_k$ as illustrated in (4.8), there are three possible conditions for the response vector elements $r_{k,n}(\omega_n)$ for $\forall n = \{1,\ldots,N\}$:

1. For the p-th entry of the response vector $r_{k,p}(\omega_p)$ such that the corresponding training direction overlaps the AoD of the strongest path, i.e. $\omega_p = \varphi_k^{(L)}$, for $p \in \{1,...,N\}$, denote the corresponding training vector as $\mathbf{t}_p(\omega_p)$. The norm of the p-th element of the response vector is given from (4.6) by

$$\begin{aligned} |r_{k,p}(\omega_p)| &= |\mathbf{t}_p(\omega_p)\mathbf{h}_k^H| \\ &= |\alpha_k^{(L)*}\mathbf{t}_p(\omega_p)\mathbf{a}_k^{(L)H}(\varphi_k^{(L)}) + \sum_{l=1}^{L-1}\alpha_k^{(l)*}\mathbf{t}_p(\omega_p)\mathbf{a}_k^{(l)H}(\varphi_k^{(l)})| \\ &\underset{(a)}{=} |M\alpha_k^{(L)*} + \sum_{l=1}^{L-1}\alpha_k^{(l)*}\mathbf{t}_p(\omega_p)\mathbf{a}_k^{(l)H}(\varphi_k^{(l)})| \\ &\underset{(b)}{\overset{M\to\infty}{\longrightarrow}} M|\alpha_k^{(L)*}| = M|\alpha_k^{(L)}|. \end{aligned} \tag{4.9}$$

   where (a) holds since $\omega_p = \varphi_k^{(L)}$ hence $\mathbf{t}_p(\omega_p) = \mathbf{a}_k^{(L)}(\varphi_k^{(L)})$, and (b) holds since when $\omega_p \neq \varphi_k^{(l)}$, $\alpha_k^{(l)*}\mathbf{t}_n(\omega_n)\mathbf{a}_k^{(l)*}(\varphi_k^{(l)}) \to 0$.

2. For the q-th entry of the response vector $r_{k,q}(\omega_q)$ such that the corresponding training direction overlaps the AoD of one of the 1st, …, $(L-1)$-th paths, i.e. $\omega_q = \varphi_k^{(l)}$, $\forall l \in \{1,...,L-1\}$, and $q \in \{1,...,N\}$ and $q \neq p$. The norm of the corresponding q-th element of the response vector is given by

$$|r_{k,q}(\omega_q)| = |\mathbf{t}_q(\omega_q)\mathbf{h}_k^H| \overset{M\to\infty}{\longrightarrow} M|\alpha_k^{(l)*}| = M|\alpha_k^{(l)}| \leq M|\alpha_k^{(L)*}| = |r_{k,p(\omega_p)}|. \tag{4.10}$$

3. For the rest of the entries of the response vector $r_{k,n}(\omega_n)$ such that $\omega_n \neq \varphi_k^{(l)}$, $\forall n \in \{1,...,N\}$, $n \neq p,q$ and $\forall l \in \{1,...,L\}$, the corresponding norm of the n-th

element of the response vector satisfies

$$|r_{k,n}(\omega_n)| \overset{M\to\infty}{\longrightarrow} 0. \tag{4.11}$$

Therefore, when perfect CSI is adopted, the element of the response vector with the greatest norm corresponds to the strongest mmWave path, i.e.

$$|r_{k,p}(\omega_p)| \overset{M\to\infty}{\geq} |r_{k,q}(\omega_q)| \overset{M\to\infty}{\geq} |r_{k,n}(\omega_n)| \tag{4.12}$$

□

From Theorem 2, as $M \to \infty$, the path with the maximum norm of the complex path gain and the elements of the response vector have the following relationship :

$$\max_{\forall n}(\left|r_{k,n}(\omega_n)\right|) = \max_{\forall l}\left|M\alpha_k^l\right| = M\left|\alpha_k^{(L)}\right|. \tag{4.13}$$

Theorem 2 also shows that the estimation of the BS array response vector, AoD, and complex path gain of the strongest path for each UT at the BS can be achieved from the perfect CSI by solving

$$\begin{aligned}\left\{\hat{\mathbf{a}}_k^{(L)} = \mathbf{t}_p, \hat{\varphi}_k^{(L)} = \omega_p, \hat{\alpha}_k^{(L)} = \frac{1}{M} r_{k,p}^*(\omega_p)\right\} &= \arg\max_{\forall n}\{\left|r_{k,n}(\omega_n)\right|\}, \\ s.t. \quad n = 1,...,N, N = \frac{2\pi}{\Delta\omega} + 1, \omega_n &= (n-1)\Delta\omega,\end{aligned} \tag{4.14}$$

where $r_{k,n}(\omega_n)$ is obtained from (4.8) by the training matrix and the obtained CSI. $\hat{\mathbf{a}}_k^{(L)}$, $\hat{\varphi}_k^{(L)}$, and $\hat{\alpha}_k^{(L)}$ denote the estimated array response vector, azimuth AoD, and complex path gain of the strongest path. Note that by sequentially subtracting an acquired path from the mmWave CSI, and repeating the above estimation process, all of the received paths can be sequentially estimated following the descending order of the path power.

4.4.2 Channel deconstruction from imperfect CSI

The channel deconstruction approach is proposed in (4.14) for the estimation of the individual mmWave path parameters with the perfect CSI that exploits both the limited-scattering clustered properties of the mmWave channel and the asymptotic orthogonality of the massive MIMO channel. This section shows the effectiveness of the proposed channel deconstruction approach on the estimation when imperfect CSI with Gaussian-perturbation is adopted.

In TDD-mode mmWave massive MIMO systems, the CSI is obtained from pilot training. An imperfect estimation of the actual channel can result from other interference within the massive MIMO system as well as the implementation imperfections (e.g. time and frequency off-sets). The imperfection in the CSI is modelled by an independent, Gaussian-perturbation term [79, 97], and given by

$$\tilde{\mathbf{h}}_k = \sqrt{1-\delta_k^2}\mathbf{h}_k + \delta_k \boldsymbol{\Omega}_k, \tag{4.15}$$

where $\tilde{\mathbf{h}}_k$ represents the imperfect, perturbed channel vector known as the CSI at the BS, $\mathbf{h}_k \in \mathcal{C}^{1\times M}$ is the actual channel vector denoted in (4.1), and $\boldsymbol{\Omega}_k \in \mathcal{CN}^{1\times M}$ is the interference or perturbation vector, i.e. estimation error vector. $\boldsymbol{\Omega}_k \in \mathcal{CN}^{1\times M}$ is i.i.d. vector with zero mean and unit variance elements, i.e. $\Omega_{k,m} \sim$i.i.d.$\mathcal{CN}(0,1)$. The term δ_k denotes the CSI imperfection factor that quantifies the CSI accuracy with $0 \leq \delta_k \leq 1$, such that $\tilde{\mathbf{h}}_k \overset{\delta_k \to 0}{\longrightarrow} \mathbf{h}_k$ and $\tilde{\mathbf{h}}_k \overset{\delta_k \to 1}{\longrightarrow} \boldsymbol{\Omega}_k$.

Next, the effectiveness of the proposed channel deconstruction algorithm with imperfect CSI adoption is indicated. Applying $\tilde{\mathbf{h}}_k$ denoted in (4.15) to (4.8) illustrates the effect of the imperfect CSI for mmWave channel deconstruction and results in an imperfect response vector $\tilde{\mathbf{r}}_k$, with the n-th entry for the k-th UT given by

$$\tilde{r}_{k,n} = \mathbf{t}_n(\omega_n)\tilde{\mathbf{h}}_k^H = \mathbf{t}_n(\omega_n)(\sqrt{1-\delta_k^2}\mathbf{h}_k^H + \delta_k \boldsymbol{\Omega}_k^H). \tag{4.16}$$

Accordingly, Theorem 3 can be obtained in the following sequel.

Theorem 3. *The maximum entry of the norms of the response vector elements from the Gaussian-perturbed imperfect CSI corresponds to the scaled value of the corresponding quantity with the perfect CSI as $M \to \infty$.*

Proof. Following the principle of Theorem 2, and noticing that the channel training vectors $\mathbf{t}_n$ for $\forall n \in \{1,\ldots,N\}$ are orthogonal to the independent perturbed Gaussian vector $\mathbf{\Omega}_k$ for $\forall k \in \{1,\ldots,K\}$ as $M \to \infty$, i.e. $\mathbf{t}_n(\tilde{\omega}_n)\mathbf{\Omega}_k^H \to 0$ where $\tilde{\omega}_n$ denotes the training beam spatial direction used by the proposed channel deconstruction algorithm for the estimation of the path parameters from the Gaussian-perturbed imperfect CSI. Hence, when the training beam spatial direction overlaps the AoD of the strongest mmWave path, i.e. the estimated AoD from the imperfect CSI $\tilde{\omega}_p = \varphi_k^{(L)}$, the corresponding norm of the p-th element of the response vector is given by

$$\begin{aligned}
|\tilde{r}_{k,p}(\tilde{\omega}_p)| &= |\mathbf{t}_p(\tilde{\omega}_p)(\sqrt{1-\delta_k^2}\mathbf{h}_k^H + \delta_k\mathbf{\Omega}_k^H)| \\
&= |\sqrt{1-\delta_k^2}\alpha_k^{(L)*}\mathbf{t}_p(\tilde{\omega}_p)\mathbf{a}_k^{(L)H}(\varphi_k^{(L)}) \\
&\quad + \sqrt{1-\delta_k^2}\sum_{l=1}^{L-1}\alpha_k^{(l)*}\mathbf{t}_p(\tilde{\omega}_p)\mathbf{a}_k^{(l)H}(\varphi_k^{(l)}) \\
&\quad + \delta_k\mathbf{t}_p(\tilde{\omega}_p)\mathbf{\Omega}_k^H| \\
&\overset{M\to\infty}{\longrightarrow} |\sqrt{1-\delta_k^2}M\alpha_k^{(L)*}| \\
&= \sqrt{1-\delta_k^2}|r_{k,p}(\omega_p)|,
\end{aligned} \tag{4.17}$$

where $\tilde{r}_{k,p}(\tilde{\omega}_p)$ denotes the response element with the maximum norm among all of the N response elements calculated from the imperfect CSI and $\tilde{\omega}_p$ is the corresponding training beam direction with imperfect CSI adoption.

Hence, the element of the response vector with the highest norm obtained from imperfect CSI $\tilde{r}_{k,p}(\tilde{\omega}_p)$ has the following relationship with the complex path gain of the strongest path $\alpha_k^{(L)}$ and the element of the response vector with the highest norm obtained from the perfect CSI $r_{k,p}(\omega_p)$

$$\tilde{r}_{k,p}(\tilde{\omega}_p) = \sqrt{1-\delta_k^2}M\alpha_k^{(L)*} = \sqrt{1-\delta_k^2}r_{k,p}(\omega_p). \tag{4.18}$$

□

Theorem 3 implies that, for the estimation of mmWave paths for the k-th UT, although imperfect CSI is adopted, the proposed channel deconstruction algorithm denoted in (4.14) can successfully identify the appropriate array response vector and the corresponding AoD, i.e. $\tilde{\tilde{\varphi}}_k^{(L)} = \tilde{\omega}_p = \varphi_k^{(L)} = \omega_p = \hat{\varphi}_k^{(L)}$, where $\tilde{\tilde{\varphi}}_k^{(L)}$ is the estimated AoD of the strongest path from the imperfect CSI. Meanwhile, for the estimation of the complex path gain, from (4.18), the elements of the response vectors calculated for perfect CSI is a scaled value of it calculated for imperfect CSI with a factor of $\sqrt{1-\delta_k^2}$. If the BS knows δ_k, the imperfection $\sqrt{(1-\delta_k^2)}$ in the estimation of the complex path gain can be further removed. Therefore, the Gaussian-perturbed component in the adopted CSI does not affect the estimation of the mmWave path parameters by the proposed channel deconstruction approach as $M \to \infty$.

4.4.3 Estimation Mean-Square Error (MSE)

The proposed channel deconstruction algorithm enables the estimation of the mmWave path parameters, i.e. the array response vector and the corresponding complex path gain for each path and each UT using the CSI in TDD-mode massive MIMO systems. This section introduces the analogue and hybrid estimation mean-square errors (MSEs) for evaluating the accuracy of the estimation. The MSEs are significant due to the relationship with the precoding schemes proposed in the upcoming section.

Consider that the BS has no knowledge of δ_k, by solving (4.14), the mmWave individual path parameters, i.e. the complex path gain and the corresponding array response vector, spatial AoDs of the paths can be obtained. For the l-th ordered path of the k-th UT $\alpha_k^{(l)}\mathbf{a}_k^{(l)}$, denote the m-th element of the estimated array response vector $\hat{\mathbf{a}}_k^{(l)}$ as $\hat{\mathrm{a}}_{k,m}^{(l)}$, and denote the m-th element of the estimated path vector obtained by $\hat{\mathbf{h}}_k^{(l)} = \hat{\alpha}_k^{(l)}\hat{\mathbf{a}}_k^{(l)}$ as $\hat{h}_{k,m}^{(l)}$.

To quantify the estimation of the array response vector and the path vector, the analogue estimation mean-square error (A-MSE) and hybrid estimation mean-square error (H-MSE) can be defined corresponding to the ABS and HBS precoding schemes, respectively. To be specific, when estimating the array response vector of the l-th ordered path for the k-th UT,

the estimation error of the m-th array response element is given by

$$\varepsilon_{k,m}^{A,(l)} = \hat{\mathrm{a}}_{k,m}^{(l)} - \mathrm{a}_{k,m}^{(l)}, \tag{4.19}$$

where $\hat{\mathrm{a}}_{k,m}^{(l)}$ and $\mathrm{a}_{k,m}^{(l)}$ denote the m-th element of the estimated and actual array response vectors $\hat{\mathbf{a}}_k^{(l)}$ and $\mathbf{a}_k^{(l)}$. The estimation A-MSE is hence defined as

$$\overline{|\varepsilon_k^{A,(l)}|^2} \triangleq \mathbb{E}[|\varepsilon_{k,m}^{A,(l)}|^2] = \mathbb{E}[\frac{1}{M}\sum_{m=1}^{M}|\hat{\mathrm{a}}_{k,m}^{(l)} - \mathrm{a}_{k,m}^{(l)}|^2]. \tag{4.20}$$

When estimating the mmWave path contributed by the complex path gain and the array response vector, the hybrid estimation error can be denoted as

$$\varepsilon_{k,m}^{H,(l)} = \hat{h}_{k,m}^{(l)} - h_{k,m}^{(l)}, \tag{4.21}$$

hence, the estimation H-MSE can be defined as

$$\overline{|\varepsilon_k^{H,(l)}|^2} \triangleq \mathbb{E}[|\varepsilon_{k,m}^{H,(l)}|^2] = \mathbb{E}[\frac{1}{M}\sum_{m=1}^{M}|\hat{h}_{k,m}^{(l)} - h_{k,m}^{(l)}|^2], \tag{4.22}$$

where $\hat{h}_{k,m}^{(l)} = \hat{\alpha}_k^{(l)}\hat{\mathrm{a}}_{k,m}^{(l)}$ and $h_{k,m}^{(l)} = \alpha_k^{(l)}\mathrm{a}_{k,m}^{(l)}$.

4.5 MmWave analogue and hybrid beam-selecting precoding

To jointly tackle the signal processing constraints in large-scale mmWave systems and the multi-user precoding complexity in massive MIMO systems, this section proposes the analogue and hybrid beam-selecting precoding approaches exploiting the estimated parameters of the strongest path for the TDD-mode massive MIMO downlink precoding followed by the corresponding system architecture designs.

4.5.1 Analogue beam-selecting precoding

The objective is to design a low-complexity, massive MIMO analogue precoder which satisfies the RF-domain constraint of phase-only control to improve the energy efficiency and reduce the fabrication cost by reducing the number of RF chains. This section proposes an analogue beam-selecting (ABS) precoder that exploits the strongest path parameters estimated by the channel deconstruction algorithm. With the proposed ABS precoder, the RF-domain phase-only constraint is satisfied, which require the elements of the precoding matrix to be unit-magnitude complex numbers. Therefore, the precoding can be achieved by low-complexity, low-cost and energy-efficient signal processing with a significant reduction in the number of RF chains. The constraint of the quantised RF control feasibility is to be considered in future work by searching for the feasible precoder with minimum Euclidean distance to the ABS precoder.

On the downlink, the BS applies an $M \times K$ analogue precoder that is denoted as

$$\mathbf{W}_A = \frac{1}{\sqrt{\gamma_A}}\mathbf{A}, \tag{4.23}$$

where $\mathbf{W}_A \in \mathcal{C}^{M\times K}$, $\mathbf{A} \in \mathcal{C}^{M\times K}$ is the analogue precoding matrix with constant-magnitude entries, and γ_A denotes the normalisation factor, such that $\gamma_A = \mathbb{E}[trace(\mathbf{A}\mathbf{A}^H)]$. Hence, the signal vector transmitted by the BS is given by

$$\mathbf{x} = \mathbf{W}_A\mathbf{s} = \frac{1}{\sqrt{\gamma_A}}\mathbf{A}\mathbf{s}, \tag{4.24}$$

where $\mathbf{x} \in \mathcal{C}^{M\times 1}$ denotes the transmit signal vector by the M-element BS array and $\mathbf{s}$ is the $K \times 1$ desired information symbol vector to K UTs, such that $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_K$.

With the proposed channel deconstruction approach in (4.14), the estimated array response vectors of the strongest path $\hat{\mathbf{a}}_k^{(L)}$ for all UTs ensure unit-magnitude complex elements that meet the analogue precoding constraint. Therefore, the analogue precoding matrix is formed by the estimated strongest path array response vector of each UT, and given by

$$\mathbf{A} = [\hat{\mathbf{a}}_1^{(L)H}, \ldots, \hat{\mathbf{a}}_K^{(L)H}], \tag{4.25}$$

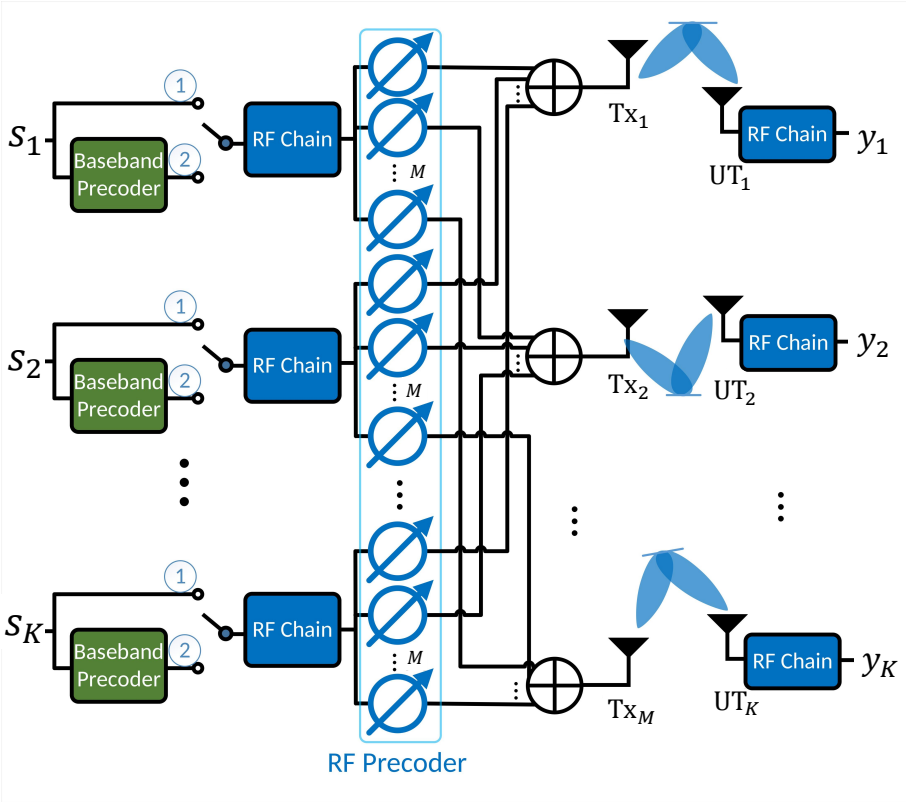


Fig. 4.1 Hardware architectures of analogue and hybrid beam-selecting precoding for TDD-mode mmWave massive MIMO systems implemented by K RF chains, followed by $M \times K$ RF-domain phase shifters.

and the normalisation factor is given by

$$\gamma_A = \mathbb{E}[trace(\mathbf{A}\mathbf{A}^H)] = M \times K. \tag{4.26}$$

For the hardware architecture of the ABS precoding, a fully connected structure can be adopted, in which each BS antenna is connected to all RF chains via a network of analogue phase shifters for the RF-domain analogue control. K RF chains are employed for multi-stream data transmissions to K independent UTs. $M \times K$ analogue phase shifters set the phases over M antenna elements to realise the $M \times K$ analogue precoder $\mathbf{A}$. The hardware architecture illustrated in Fig. 4.1 is common for both the analogue and hybrid precoders wherein switches at the input can select between the two operating modes. For the ABS precoding, the switches are in position 1.

4.5.2 Hybrid beam-selecting precoding

The hybrid precoder divides the precoding into analogue and digital domains, which enables a large-scale analogue control to be implemented in the RF domain via analogue phase shifters as well as a low-dimensional digital control executed at the baseband. The channel deconstruction approach denoted in (4.14) enables the estimation of the complex path gains as well as the array response vectors for each UT and each path. Therefore, the K-dimensional digital precoder can be formed by the estimated complex path gains $\hat{\alpha}_k^{(L)}$ of the strongest paths for all UTs, and given by

$$\mathbf{D} = diag\{\hat{\alpha}_1^{(L)}, \ldots, \hat{\alpha}_K^{(L)}\}^H, \tag{4.27}$$

where $d_{k,k} = \hat{\alpha}_k^{(L)*}$ is the k-th diagonal element of $K \times K$ matrix $\mathbf{D}$ and indicates the digital baseband control for the k-th UT. Meanwhile, the corresponding estimated array response vectors compose the analogue precoder and form the same design with the ABS precoder as shown in (4.25). Hence, the HBS precoder $\mathbf{W}_H \in \mathcal{C}^{M \times K}$ is obtained by a low-dimensional digital baseband precoder $\mathbf{D}$, followed by an RF-domain large-dimensional analogue precoder $\mathbf{A}$

$$\mathbf{W}_H = \frac{1}{\sqrt{\gamma_H}}\mathbf{AD}, \tag{4.28}$$

and the power normalisation factor is given by

$$\gamma_H = \mathbb{E}[trace(\mathbf{AD}(\mathbf{AD}^H))]. \tag{4.29}$$

The signal transmitted by the BS antenna array is therefore given by

$$\mathbf{x} = \mathbf{W}_H\mathbf{s} = \frac{1}{\sqrt{\gamma_H}}\mathbf{ADs}. \tag{4.30}$$

For the hardware architecture design of the HBS precoder, the baseband digital precoder $\mathbf{D}$ is added to the fully connected structure of the proposed ABS precoder as showed in Fig. 4.1 with the switches in position 2.

So far, the ABS and HBS precoders are proposed in this section for mmWave massive MIMO precoding. The estimated strongest mmWave path for each UT is selected by the precoders that enable the highest SNR compared with random path selecting precoding. Meanwhile, since the strongest path is estimated directly from the CSI through only one implementation of the channel deconstruction, extra estimation processes are avoided. For the architectures of both ABS and HBS precoders, only K RF chains, $M \times K$ phase shifters, and M combiners are required as illustrated in Fig. 4.1 that substantially reduce the signal processing complexity as well as increase the energy efficiency at the BS with respect to the conventional massive MIMO precoding using fully digital baseband signal precoding, which requires a dedicated baseband digital control and an RF chain for each antenna as shown in Fig. 2.3. Moreover, for the proposed precoders, the number of RF chains and the dimension of digital control are independent of the number of BS antennas. Such a design is vital especially when the BS array scales up. The obtaining of the CSI through RF limited hardware architecture is still an open research question though the BS can always use M RF chains for CSI estimation and the proposed architecture for data transmission to achieve high energy efficiency. Note that the proposed ABS and HBS precoders exploit the strongest mmWave path estimated from training beams as shown in (4.7). These training beams are formed following the same structure of the mmWave path array response vector. Therefore, selecting the path from all of the received paths in the mmWave channel for precoding corresponds to selecting the corresponding beam from all of the training beams.

4.6 Effective SINRs and lower bounds on the sum-rate capacities for ABS and HBS precoding

This section analyses the lower bound on the system sum-rate capacity for ABS and HBS precoding in mmWave massive MIMO transmissions via the effective SINR with random matrix theory and order statistics principles. The analysis is intractable with the presence of estimation errors during the channel deconstruction under realistic system configurations, wherein the numbers of BS antennas are moderately large. Therefore, consider that the perfect path parameters are obtained from the channel deconstruction

estimation denoted in (4.14). Therefore, $\hat{\alpha}_k^{(L)} = \alpha_k^{(L)}$, and $\hat{\mathbf{a}}_k^{(L)} = \mathbf{a}_k^{(L)}$ that yield zero analogue and hybrid MSEs in (4.19) and (4.21), i.e. $\varepsilon_{k,m}^{A,(l)} = \varepsilon_{k,m}^{H,(l)} = 0$.

Following the same principle of the effective SINR introduced in Chapter 2 section 2.4.3, denote the effective SINRs of ABS and HBS precoding as SINR^{ABS} and SINR^{HBS}. The following sections present the theoretical analysis of the effective SINRs for the ABS and HBS precoding, wherein the received signal through the unselected 1st,...,$(L-1)$-th paths are treated as the inter-cluster (or intra-user) interference.

4.6.1 Effective SINR of ABS precoding

For ABS precoding, the received signal at the k-th UT can be obtained from (4.1),(4.24), (4.25) and (4.26), and expressed as

$$\begin{aligned} y_k = & \underbrace{\sqrt{\frac{\rho}{\gamma_A}}\alpha_k^{(L)}\mathbf{a}_k^{(L)}\hat{\mathbf{a}}_k^{(L)H}s_k}_{\substack{Desired\\Signal}} + \underbrace{\sqrt{\frac{\rho}{\gamma_A}}\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)}\hat{\mathbf{a}}_k^{(L)H}s_k}_{\substack{Inter-cluster\\Interference}} + \\ & \underbrace{\sqrt{\frac{\rho}{\gamma_A}}\sum_{\substack{q=1\\q\neq k}}^{K}\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)}\hat{\mathbf{a}}_q^{(L)H}s_q}_{\substack{Inter-user\\Interference}} + \underbrace{n_k}_{Noise} \end{aligned} \tag{4.31}$$

$$\begin{aligned} \overset{(a)}{=} & \underbrace{\sqrt{\frac{\rho}{\gamma_A}}\alpha_k^{(L)}\mathbf{a}_k^{(L)}\mathbf{a}_k^{(L)H}s_k}_{\substack{Desired\\Signal}} + \underbrace{\sqrt{\frac{\rho}{\gamma_A}}\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_k^{(L)H}s_k}_{\substack{Inter-cluster\\Interference}} + \\ & \underbrace{\sqrt{\frac{\rho}{\gamma_A}}\sum_{\substack{q=1\\q\neq k}}^{K}\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_q^{(L)H}s_q}_{\substack{Inter-user\\Interference}} + \underbrace{n_k}_{Noise}, \end{aligned} \tag{4.32}$$

where (a) holds assuming $\hat{\mathbf{a}}_k^{(L)} = \mathbf{a}_k^{(L)}$ for all k. The desired signal term on the right side of (4.31) and (4.32) denotes the desired downlink signal received through the desired path of the k-th UT, the inter-cluster interference (or intra-user interference) term denotes the interference from the transmitted signal received by the $(L-1)$ undesired paths, and the inter-user interference denotes the received signal by the k-th UT intended for other $(K-1)$ UTs. Hence, the effective SINR of the ABS precoding for the k-th UTs is obtained from

(4.32) and given by

$$\begin{aligned}\text{SINR}^{ABS} &= \frac{variance\left\{\sqrt{\frac{\rho}{\gamma_A}}\alpha_k^{(L)}\mathbf{a}_k^{(L)}\mathbf{a}_k^{(L)H}s_k\right\}}{variance\left\{\sqrt{\frac{\rho}{\gamma_A}}\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_k^{(L)H}s_k\right\}+variance\left\{\sqrt{\frac{\rho}{\gamma_A}}\sum_{\substack{q=1\\q\neq k}}^{K}\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_q^{(L)H}s_q\right\}+1} \\ &\overset{(a)}{=} \frac{\frac{\rho}{\gamma_A}\mathbb{E}[|\alpha_k^{(L)}\mathbf{a}_k^{(L)}\mathbf{a}_k^{(L)H}|^2]}{\frac{\rho}{\gamma_A}\mathbb{E}[|\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_k^{(L)H}|^2]+\frac{\rho}{\gamma_A}\mathbb{E}[|\sum_{\substack{q=1\\q\neq k}}^{K}\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_q^{(L)H}|^2]+1},\end{aligned} \tag{4.33}$$

where (a) holds because the expected value of each term in (4.32) is zero. Therefore, the expected values of the power of each term must be calculated.

Since the complex path gain of the unsorted paths follows $\alpha_k^l \sim$i.i.d.$\mathcal{CN}(0,\frac{1}{L})$, the square norm of the random complex path gain $X=|\alpha_k^l|^2$ follows the gamma distribution, i.e. $X\sim\Gamma(1,\frac{1}{L})$. For a gamma distributed random variable $X\sim\Gamma(\alpha_g,\frac{1}{\lambda_g})$ with α_g denoting the shape parameter and $\frac{1}{\lambda_g}$ representing the scale parameter, the corresponding p.d.f. can be denoted as [98]

$$\begin{aligned}f_\Gamma(x;\alpha_g,\lambda_g) &= \frac{x^{(\alpha_g-1)}\lambda_g{}^{\alpha_g}e^{(-\lambda_g x)}}{\Gamma(\alpha_g)} \\ &= Le^{-Lx},\end{aligned} \tag{4.34}$$

where $\Gamma(\alpha_g)$ is a complete gamma function [98] with $\Gamma(\alpha_g)=(\alpha_g-1)!$ and $\alpha_g=1$, $\lambda_g=L$. Since α_g is a positive integer, the c.d.f. on X becomes the Erlang distribution that can be denoted as [98]

$$\begin{aligned}F_\Gamma(x;\alpha_g,\lambda_g) &= 1-\sum_{\mathtt{i}=0}^{\alpha_g-1}\frac{(\lambda_g x)^{\mathtt{i}}}{\mathtt{i}!}e^{-\lambda_g x} = 1-e^{-\lambda_g x} \\ &= 1-e^{-Lx}.\end{aligned} \tag{4.35}$$

The proposed ABS precoder selects the strongest estimated mmWave path for each UT among L received paths that corresponds to the selection of the L-th order statistic of L samples of the gamma distributed X (the square norm of the random complex path gain) with the p.d.f. and c.d.f. shown in (4.34) and (4.35). Denote the L-th ordered X as $X_{(L)}$, the corresponding selected complex path gain as $\alpha_k^{(L)}$ and the array factor as $\mathbf{a}_k^{(L)}$. Therefore,

the p.d.f. on the square norm of the strongest path's complex path gain, i.e. the L-th order statistic $X_{(L)} = |\alpha_k^{(L)}|^2$ is given from (4.34) and (4.35) with $\alpha_g = 1$, $\lambda_g = L$ by [99]

$$\begin{aligned} f_{(L)}(x) &= L(F_\Gamma(x;\alpha_g,\lambda_g))^{L-1} f_\Gamma(x;\alpha_g,\lambda_g) \\ &= L^2(1-e^{-Lx})^{L-1}e^{-Lx}. \end{aligned} \tag{4.36}$$

The expected power of the complex path gain for the selected strongest path can be denoted from (4.36) or from the ordered exponential distribution as it is a special case of the gamma distribution [100]. Therefore,

$$\begin{aligned} \mathbb{E}[|\alpha_k^{(L)}|^2] &= \int x f_{(L)}(x)dx \\ &= \frac{1}{L}\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}. \end{aligned} \tag{4.37}$$

(4.37) illustrates the variance of the complex path gain for the L-th ordered path, which is given by the i.i.d. Gaussian distribution of the unordered complex path gain from the order statistic principle as presented.

Next, for the effective SINR evaluation of (4.33), the expected value of the power of each term is derived.

- The expected value of the power of the desired signal for ABS precoding denoted in (4.33) can be derived from (4.37), and given by

$$\begin{aligned} &\mathbb{E}[|\alpha_k^{(L)}\mathbf{a}_k^{(L)}\mathbf{a}_k^{(L)H}|^2] = \mathbb{E}[|M\alpha_k^{(L)}|^2] = M^2\mathbb{E}[|\alpha_k^{(L)}|^2] \\ &= \frac{M^2}{L}\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}. \end{aligned} \tag{4.38}$$

- The power normalisation of the ABS precoder is derived from (4.25), and given by

$$\gamma_A = \mathbb{E}[(\mathbf{A}\mathbf{A}^H)] = MK. \tag{4.39}$$

- The inter-cluster interference of ABS precoding arises from the unselected 1st $,\ldots,(L-1)$-th ordered paths as demonstrated in (4.33). It is obvious that the analogue array response elements, which have mean zero and variance 1, do not

change the mean and the variance of the complex path gains. Hence, consider that the elements of the inter-cluster interference follow $\alpha_k^{(i)} \mathrm{a}_{k,m}^{(i)} \mathrm{a}_{k,m}^{(L)*} \sim \mathcal{CN}(0, \sigma_{(i)}^2)$ for $\forall m \in \{1,\ldots,M\}$, where $\sigma_{(i)}^2$ denotes the variance of the complex path gain of the i-th ordered path $\alpha_k^{(i)}$, i.e. $\sigma_{(i)}^2 = \mathbb{E}[|\alpha_k^{(i)}|^2]$. This assumption is the *uncorrelated* case for the elements of the inter-cluster interference that is the integrated result of the AoDs distribution, carrier frequency, d, and M, etc. [96]. Further discussion about this assumption of the correlation is illustrated in the next chapter. Note that from the order statistic, the variances $\sigma_{(1)}^2,\ldots,\sigma_{(L-1)}^2$ of the ordered complex path gains for the $(L-1)$ unselected paths are jointly distributed for each UT. Hence, the expected value of the power of the inter-cluster interference shown in (4.33) is given by

$$\begin{aligned}\mathbb{E}[|(\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)})\mathbf{a}_k^{(L)H}|^2] &= \mathbb{E}[\sum_{i=1}^{L-1}|\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_k^{(L)H}|^2] \\ &= M\int\ldots\int(\sum_{i=1}^{L-1}x_{(i)})f\left(x_{(1)},\ldots,x_{(L-1)}\right)dx_{(1)}\ldots dx_{(L-1)}.\end{aligned} \tag{4.40}$$

where $X_{(i)} = |\alpha_k^{(i)}|^2$, and $f\left(x_{(1)},\ldots,x_{(L-1)}\right)$ is the joint p.d.f. on $(X_{(1)},\ldots,X_{(L-1)})$ order statistics. From (4.34) and (4.35), $f\left(x_{(1)},\ldots,x_{(L-1)}\right)$ is given by the joint p.d.f. on the order statistics principle introduced in [99], and denoted as

$$f\left(x_{(1)},\ldots,x_{(L-1)}\right) = L^{L-1}L!e^{-L(x_{(1)}+\ldots+2x_{(L-1)})}. \tag{4.41}$$

To further simplify the analysis, an approximation of the variance of the elements of the inter-cluster interference can be made by considering the complex path gains of the unselected paths $\alpha_k^{(i)}$ to be unordered i.i.d. with zero mean and $\frac{1}{L}$ variance. This approximation corresponds to the case that each complex path gain of the unselected paths has its maximum possible variance at the same time. Correspondingly, the elements of the inter-cluster interference can be approximated by $\alpha_k^{(i)}\mathrm{a}_{k,m}^{(i)}\mathrm{a}_{k,m}^{(L)*} \sim \mathcal{CN}(0,\frac{1}{L})$ for $\forall m \in \{1,\ldots,M\}$, $\forall i \in \{1,\ldots,L-1\}$, and $\forall k \in \{1,\ldots,K\}$. Hence, the approximation of the expected value of the power of the inter-cluster interference

received via the i-th unselected path is given by

$$\mathbb{E}[|\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_k^{(L)H}|^2] = \frac{M}{L}. \tag{4.42}$$

Thus, the approximated expected value of the power of the total inter-cluster interference is expressed as

$$\begin{aligned}\mathbb{E}[|\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_k^{(L)H}|^2] &= \mathbb{E}[\sum_{i=1}^{L-1}|\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_k^{(L)H}|^2] \\ &= (L-1)\mathbb{E}[|\alpha_k^{(i)}\mathbf{a}_k^{(i)}\mathbf{a}_k^{(L)H}|^2] = \frac{(L-1)M}{L}.\end{aligned} \tag{4.43}$$

- The inter-user interference of the ABS precoding arises from the received signal intended for the other UTs via the channel of the k-th UT as shown in (4.33). Since $\alpha_k^l \sim$i.i.d.$\mathcal{CN}(0,\frac{1}{L})$, the summation of the complex path gains equals the summation of all ordered complex path gains, i.e. $\sum_{l=1}^{L}\alpha_k^l = \sum_{l=1}^{L}\alpha_k^{(l)} \sim$i.i.d.$\mathcal{CN}(0,1)$. Similar with the assumption utilised in the analysis of the inter-cluster interference, the *uncorrelated* case for elements of the path vectors can be considered for the analysis of the inter-user interference, i.e. $\sum_{i=1}^{L}\alpha_k^{(i)}\mathrm{a}_{k,m}^{(i)}\mathrm{a}_{q,m}^{(L)*} \sim$i.i.d.$\mathcal{CN}(0,1)$ for $\forall k,q \in \{1,\ldots,K\}, k\neq q, \forall m \in \{1,\ldots,M\}$. Therefore, the expected value of the power of the inter-user interference of ABS precoding is given by

$$\mathbb{E}[|(\sum_{\substack{q=1\\q\neq k}}^{K}\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)})(\mathbf{a}_q^{(L)H})|^2] = \mathbb{E}[\sum_{\substack{q=1\\q\neq k}}^{K}|(\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)})(\mathbf{a}_q^{(L)H})|^2] = (K-1)M. \tag{4.44}$$

Hence, with the inter-cluster interference calculated by (4.40) and (4.41), the effective SINR for ABS precoding shown in (4.33) can be obtained by (4.38),(4.39), and (4.44), and denoted as

$$\mathrm{SINR}^{ABS} = \frac{\frac{\rho M}{KL}\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}}{\frac{\rho L^{L-1}L!}{K}\int\ldots\int(\sum_{i=1}^{L-1}x_{(i)})e^{-L(x_{(1)}+\ldots+2x_{(L-1)})}dx_{(1)}\ldots dx_{(L-1)} + \frac{\rho(K-1)}{K} + 1}. \tag{4.45}$$

Using the approximated inter-cluster interference denoted in (4.43), the effective SINR can also be given as

$$\mathrm{SINR}^{ABS} = \frac{\frac{\rho M}{KL}\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}}{\frac{\rho(L-1)}{KL}+\frac{\rho(K-1)}{K}+1}. \tag{4.46}$$

4.6.2 Effective SINR for HBS precoding

For HBS precoding, the received signal by the k-th UT can be obtained from (4.1), (4.28),(4.29), and (4.30), and given by

$$\begin{aligned} y_k = & \underbrace{\sqrt{\frac{\rho}{\gamma_H}}\alpha_k^{(L)}\mathbf{a}_k^{(L)}(\hat{\alpha}_k^{(L)}\hat{\mathbf{a}}_k^{(L)})^H s_k}_{Desired\quad Signal} + \underbrace{\sqrt{\frac{\rho}{\gamma_H}}\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)}(\hat{\alpha}_k^{(L)}\hat{\mathbf{a}}_k^{(L)})^H s_k}_{Inter-cluster\quad Interference} + \\ & \underbrace{\sqrt{\frac{\rho}{\gamma_H}}\sum_{\substack{q=1\\q\neq k}}^{K}\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)}(\hat{\alpha}_q^{(L)}\hat{\mathbf{a}}_q^{(L)})^H s_q}_{Inter-user\quad Interference} + \underbrace{n_k}_{Noise} \end{aligned} \tag{4.47}$$

$$\begin{aligned} \overset{(a)}{=} & \underbrace{\sqrt{\frac{\rho}{\gamma_H}}\alpha_k^{(L)}\mathbf{a}_k^{(L)}(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H s_k}_{Desired\quad Signal} + \underbrace{\sqrt{\frac{\rho}{\gamma_H}}\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)}(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H s_k}_{Inter-cluster\quad Interference} + \\ & \underbrace{\sqrt{\frac{\rho}{\gamma_H}}\sum_{\substack{q=1\\q\neq k}}^{K}\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)}(\alpha_q^{(L)}\mathbf{a}_q^{(L)})^H s_q}_{Inter-user\quad Interference} + \underbrace{n_k}_{Noise}\ , \end{aligned} \tag{4.48}$$

where (a) holds assuming perfect parameters are obtained from channel deconstruction, i.e. $\hat{\alpha}_k^{(L)} = \alpha_k^{(L)}$ and $\hat{\mathbf{a}}_k^{(L)} = \mathbf{a}_k^{(L)}$. The effective SINR of the HBS precoding is hence denoted

from (4.48) and given by

$$\begin{aligned}
&\mathrm{SINR}^{HBS} \\
&= \frac{variance\left\{\sqrt{\frac{\rho}{\gamma_H}}\alpha_k^{(L)}\mathbf{a}_k^{(L)}(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H s_k\right\}}{variance\left\{\sqrt{\frac{\rho}{\gamma_H}}\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)}(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H s_k\right\} + variance\left\{\sqrt{\frac{\rho}{\gamma_H}}\sum_{\substack{q=1\\q\neq k}}^{K}\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)}(\alpha_q^{(L)}\mathbf{a}_q^{(L)})^H s_q\right\} + 1} \\
&\overset{(a)}{=} \frac{\frac{\rho}{\gamma_H}\mathbb{E}[|\alpha_k^{(L)}\mathbf{a}_k^{(L)}(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H|^2]}{\frac{\rho}{\gamma_H}\mathbb{E}[|\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)}(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H|^2] + \frac{\rho}{\gamma_H}\mathbb{E}[|\sum_{\substack{q=1\\q\neq k}}^{K}\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)}(\alpha_q^{(L)}\mathbf{a}_q^{(L)})^H|^2] + 1}.
\end{aligned} \tag{4.49}$$

where (a) holds because the expected values of the terms of the desired signal, inter-user interference and inter-cluster (or intra-user) interference in (4.48) are all zero. Next, the expected value of the power of each term in (4.49) is derived.

- The expected value of the power of the desired signal for HBS precoding denoted in (4.49) can be obtained from the p.d.f. on the power of the complex path gain of the selected L-th ordered path shown in (4.36) or the order statistic of the exponential distribution [100], and is given by

$$\begin{aligned}
&\mathbb{E}[|\alpha_k^{(L)}\mathbf{a}_k^{(L)}(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H|^2] = M^2\mathbb{E}[||\alpha_k^{(L)}|^2|^2] \\
&= M^2\int x^2 f_{(L)}(x)dx \\
&= \frac{M^2}{L^2}(\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}^2} + (\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}})^2).
\end{aligned} \tag{4.50}$$

- For the power normalisation, since UTs are independent and $\mathbb{E}[\alpha_k^{(L)}\mathbf{a}_k^{(L)}(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H] = \mathbb{E}[M|\alpha_k^{(L)}|^2]$, the normalisation factor γ_H can be calculated from (4.25) and (4.27),

and given by

$$\begin{aligned}\gamma_H &= \mathbb{E}[trace(\mathbf{AD}(\mathbf{AD})^H)] \\ &= \mathbb{E}[M\sum_{k=1}^{K}|\alpha_k^{(L)}|^2] \\ &= MK\mathbb{E}[|\alpha_k^{(L)}|^2] \\ &= \frac{MK}{L}\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}.\end{aligned} \tag{4.51}$$

where the expected square-norm of the L-th ordered path is given by (4.37).

- For HBS precoding, the inter-cluster interference arises from the unselected 1st to the $(L-1)$-th ordered paths as shown in (4.49). Note that the element $\mathrm{a}_{k,m}^{(i)}$ of the array response vector follows $\mathrm{a}_{k,m}^{(i)} \sim \mathcal{C}(0,1)$ that does not colour the mean and variance of the corresponding complex path gain. Consider that the elements of the path vectors follow $\alpha_{k,m}^{(i)}\mathrm{a}_k^{(i)} \sim$i.i.d.$\mathcal{CN}(0,\sigma_{(i)}^2)$. Meanwhile, the joint distribution of the variances of the complex path gains for the ordered paths remains. This assumption also corresponds to the uncorrelated case for the path elements as mentioned in the inter-cluster interference analysis for ABS precoding. Hence, the expected value of the power of the inter-cluster interference for HBS precoding is given by

$$\begin{aligned}&\mathbb{E}[|(\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)})(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H|^2] \\ &= \mathbb{E}[\sum_{i=1}^{L-1}|(\alpha_k^{(i)}\mathbf{a}_k^{(i)})(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H|^2] \\ &= M\int\ldots\int(\sum_{i=1}^{L-1}x_{(i)})(x_{(L)})f(x_{(1)},\ldots,x_{(L)})dx_{(1)}\ldots dx_{(L)},\end{aligned} \tag{4.52}$$

where the joint p.d.f. on the L order statistics is obtained from [99], and given by

$$f(x_{(1)},\ldots,x_{(L)}) = L^L L! e^{-L(x_{(1)}+\ldots+x_{(L)})}, \tag{4.53}$$

where $X_{(i)} = |\alpha_k^{(i)}|^2$ for $\forall i \in \{1,\ldots,L\}$.

To further simplify the analysis, similar approximation with the ABS precoding inter-cluster interference can be exploited that assumes all the unselected path vectors' elements have the same mean and variance with the unordered complex path gain, i.e. $\alpha_k^{(i)}\mathrm{a}_{k,m}^{(i)} \sim$i.i.d.$\mathcal{CN}(0,\frac{1}{L})$ for $\forall m \in \{1,\ldots,M\}$, $\forall i \in \{1,\ldots,L-1\}$, $\forall k \in \{1,\ldots,K\}$. Hence, the simplified approximation of the expected value of the power of the inter-cluster interference received by the k-UT is given by

$$\begin{aligned}
&\mathbb{E}[|(\sum_{i=1}^{L-1}\alpha_k^{(i)}\mathbf{a}_k^{(i)})(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H|^2] \\
&= M(L-1)\mathbb{E}[|\alpha_k^{(i)}\mathrm{a}_{k,m}^{(i)}|^2]\mathbb{E}[|\alpha_k^{(L)*}\mathrm{a}_{k,m}^{(L)*}|^2] \\
&= \frac{M(L-1)}{L}\mathbb{E}[|\alpha_k^{(L)}|^2] \\
&= \frac{M(L-1)}{L^2}\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}.
\end{aligned} \tag{4.54}$$

where the expected value of the square norm of the L-th ordered complex path gain is given by (4.37).

- Following a similar analysis implemented for ABS inter-user interference, for the inter-user interference of HBS precoding, the complex channel gain of the k-th UT $\sum_{i=1}^{L}\alpha_k^{(i)}\mathrm{a}_{k,m}^{(i)}$ is considered to be $i.i.d. \sim \mathcal{CN}(0,1)$ for $\forall k \in \{1,\ldots,K\}, \forall m \in \{1,\ldots,M\}$. Therefore, the expected value of the power of the inter-user interference is obtained from (4.37), and given by

$$\begin{aligned}
&\mathbb{E}[|(\sum_{\substack{q=1\\q\neq k}}^{K}\sum_{i=1}^{L}\alpha_q^{(i)}\mathbf{a}_q^{(i)})(\alpha_k^{(L)}\mathbf{a}_k^{(L)})^H|^2] \\
&= M(K-1)\mathbb{E}[|(\alpha_k^{(L)}\mathrm{a}_{k,m}^{(L)})^*|^2] \\
&= M(K-1)\mathbb{E}[|\alpha_k^{(L)}|^2] \\
&= \frac{M(K-1)}{L}\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}.
\end{aligned} \tag{4.55}$$

Hence, with the inter-cluster interference calculated by (4.52) and (4.53), the effective SINR of HBS precoding can be obtained from (4.37),(4.50),(4.51), and (4.55), and denoted

as

$$\mathrm{SINR}^{HBS} = \frac{\frac{\rho M}{LK}\left(\frac{\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}^2}}{\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}} + \sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}\right)}{\frac{\rho L^{L+1}L!}{K}\frac{1}{\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}}\int\ldots\int\left(\sum_{i=1}^{L-1}x_{(i)}\right)(x_{(L)})e^{-L(x_{(1)}+\ldots+x_{(L)})}dx_{(1)}\ldots dx_{(L)} + \frac{\rho(K-1)}{K} + 1}. \tag{4.56}$$

Using the approximation for the inter-cluster interference denoted in (4.54), the effective SINR can also be expressed as

$$\mathrm{SINR}^{HBS} = \frac{\frac{\rho M}{LK}\left(\frac{\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}^2}}{\sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}} + \sum_{\mathtt{i}=1}^{L}\frac{1}{\mathtt{i}}\right)}{\frac{\rho(L-1)}{LK} + \frac{\rho(K-1)}{K} + 1}. \tag{4.57}$$

So far, the effective SINRs for the proposed ABS and HBS precoding are analysed, and the theoretical results are derived using order statistic principles and random matrix theory. The *uncorrelated* case for the path vectors' elements is considered in the interference analysis.

To obtain the lower bound on the system sum-rate capacity for the proposed ABS and HBS precoders, the effective SINRs as well as their simplified approximations denoted in (4.45), (4.56), (4.46), and (4.57) can be substituted into (2.19).

4.7 Simulation results for channel deconstruction approach, ABS, and HBS schemes

This section first evaluates the A-MSE and H-MSE performance of the proposed channel deconstruction approach using Monte Carlo simulations to validate the effectiveness of the proposed channel deconstruction approach on the estimation of the strongest path parameters. Lower bounds on the system sum-rate capacities for the proposed ABS and HBS precoders are also simulated, in which the path parameters exploited by the ABS and HBS precoders are estimated by the proposed channel deconstruction algorithm. The proposed theoretical lower bounds are evaluated using approximations from the numerical

analysis. The performance of the conventional massive MIMO MF precoding using the perfect CSI of the mmWave channel is simulated for comparison and is considered as the optimal system performance that is implemented by the unconstrained digital baseband precoding. The impact of the CSI accuracy constrained by δ^2 on the system performance of the proposed precoders is also evaluated.

The single-cell mmWave massive MIMO downlink transmissions are considered. An M-element UPA is equipped at the BS and serving K single-antenna UTs simultaneously. The array configurations adopted are shown in Table 3.1. The limited-scattering clustered mmWave channel model presented in (4.1) is adopted, wherein each UT receives L clusters of paths, and there's a single path within each cluster. Also, consider that the complex path gain of each path follows $\alpha_k^l \sim$i.i.d.$\mathcal{CN}(0, \frac{1}{L})$ for all paths and UTs. The carrier frequency f=28 GHz and BS antenna spacing $d = 0.5\lambda$ are considered. The AoDs of each path are modelled by $\theta \sim$i.i.d.$\mathbf{U}[0, 90^\circ]$, $\varphi \sim$i.i.d.$\mathbf{U}[-180^\circ, 180^\circ]$ for the generation of the channel realisation. For path estimation by the proposed channel deconstruction algorithm, consider that the number of the training beam spatial directions N is large enough that the training beam spatial directions cover the actual AoDs of each path.

4.7.1 MSEs performance of channel deconstruction approach on estimating parameters for the strongest path

This section conducts experiments to validate the effectiveness of the proposed mmWave channel deconstruction approach denoted in (4.14) on the estimation of the strongest path parameters from perfect CSI or Gaussian-perturbed CSI. Such an evaluation is vital since the estimation results will be used for the proposed ABS and HBS precoders. Monte Carlo simulations are implemented assuming the BS has no knowledge of the channel imperfection factor δ^2, and the A-MSE and H-MSE are calculated from (4.20) and (4.22) that demonstrate the analogue and hybrid estimation accuracies.

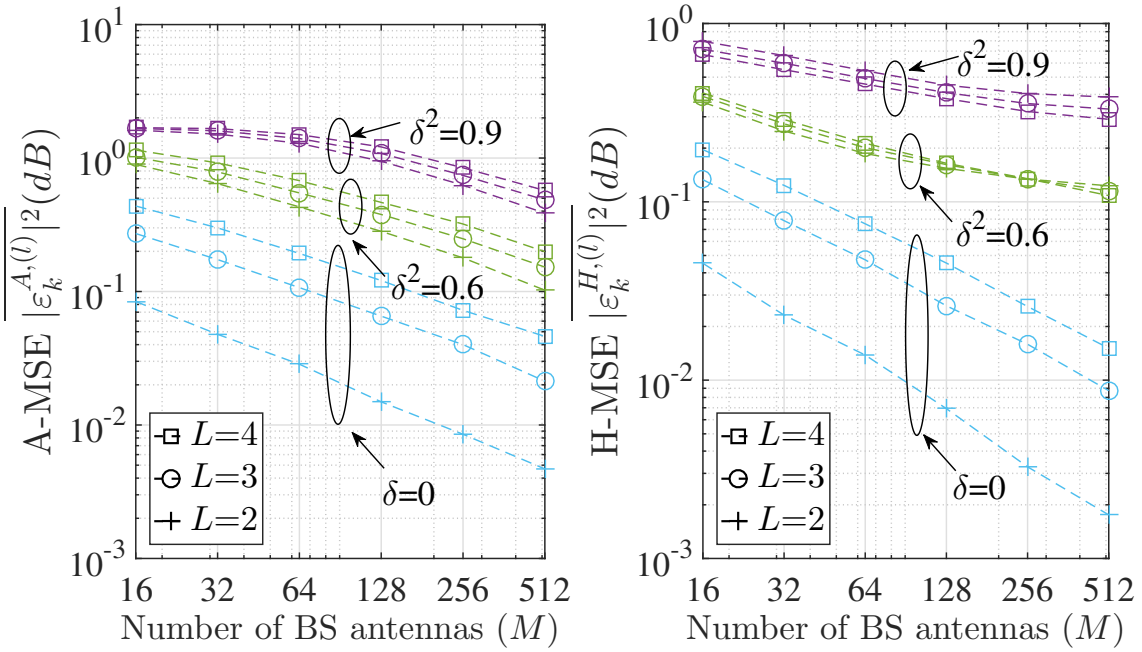
First, the effectiveness of the channel deconstruction approach on the estimation of the strongest path with the scaling up of the BS array is evaluated. Fig. 4.2a and Fig. 4.2b indicate the simulated A-MSE and H-MSE versus the number of BS antennas M for the total number of paths $L = 2, 3$ and 4 [20]. The perfect CSI and Gaussian-perturbed

imperfect CSI adoptions are considered with $\delta^2 = 0, 0.6$ and 0.9. As expected both A-MSE and H-MSE decrease with the increase of M, e.g. the A-MSE decreases from 0.27 to 0.02 and the H-MSE decreases from 0.14 to 0.009 when M increases from 16 to 512 and $L = 3, \delta^2 = 0$. Therefore, both analogue and hybrid estimations of the strongest path achieve a higher accuracy when M increases thanks to the asymptotic orthogonal property of the channel with the scaling up of the BS. Note also that the A-MSE decreases if the channel becomes less scattered, e.g. when $M = 128$ and $\delta^2 = 0$, the A-MSE decreases from 0.11 to 0.015 when L decreases from 4 to 2 as shown in Fig. 4.2a. However, the same trend only holds for H-MSE as illustrated in Fig. 4.2b when CSI is not so accurate, i.e. δ^2 is high, e.g. the H-MSE decreases from 0.45 to 0.38 when $\delta^2 = 0.9$. An opposite trend is shown for H-MSE when δ^2 is small.

To further evaluate the effectiveness of the proposed channel deconstruction algorithm when the CSI is Gaussian-perturbed as denoted in (4.15), Fig. 4.3a and Fig. 4.3b evaluate the estimation MSEs of the strongest path versus the CSI imperfection factor δ^2. The figures indicate that the A-MSE and H-MSE increase when the CSI adopted becomes less accurate, i.e. δ^2 increases. Therefore, the estimation is more accurate when the CSI is more accurate. As M increases, the estimation accuracy is improved and remained over a larger range of δ^2, e.g. when $L = 2$, the range of $0 \leq \delta^2 \leq 0.4$ corresponds to A-MSE ≤ 0.4 for $M = 32$, the range expands to $0 \leq \delta^2 \leq 0.9$ for $M = 512$. This indicates that with the scaling up of the BS array, the accuracy of the estimation by the proposed channel deconstruction approach on the strongest path can be maintained over a wider tolerance range of CSI quality.

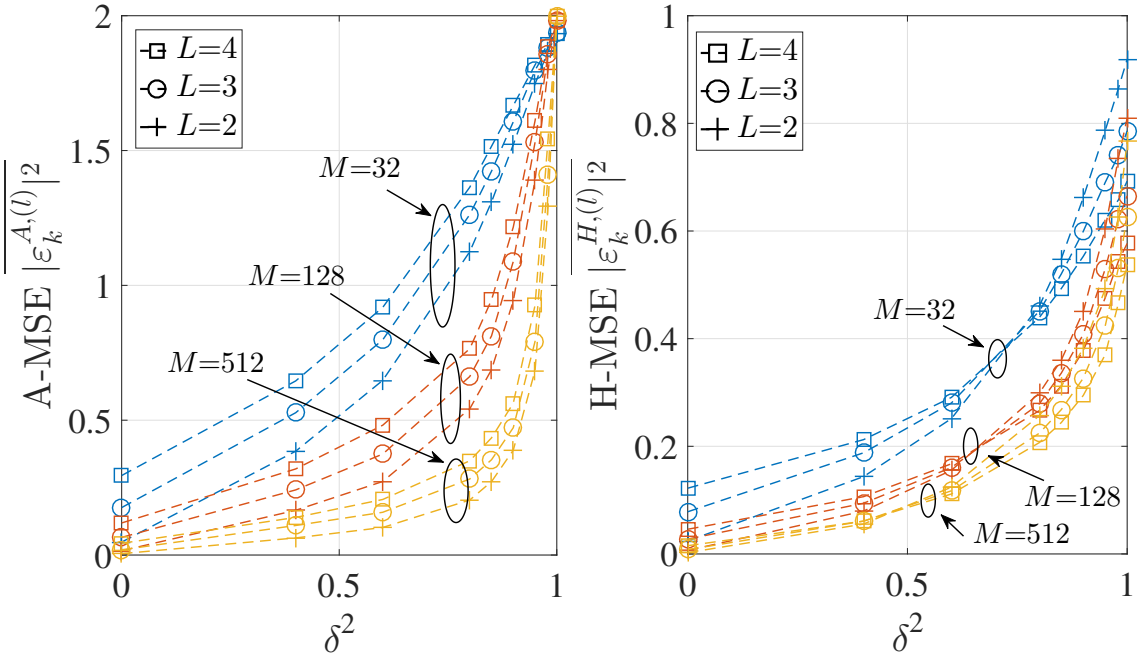
4.7.2 Lower bounds on the sum-rate capacities for ABS and HBS precoding

This section shows the simulation results of lower bounds on the system sum-rate capacities denoted in (2.19) that are obtained from Monte Carlo simulations of the effective SINRs for ABS and HBS expressed in (4.33) and (4.49). The path parameters adopted by the proposed precoders are estimated by the proposed channel deconstruction algorithm as shown in (4.14). The theoretical results are approximated by numerical analysis of



(a) A-MSE (b) H-MSE

Fig. 4.2 Estimation MSEs (dB) of the strongest path vs M, for total number of paths $L = 4,3,2$ from perfect CSI $\delta^2 = 0$ and imperfect CSI $\delta^2 = 0.6, 0.9$. $K = 10$ UTs are served simultaneously with f=28 GHz, $d = 0.5\lambda$.



(a) A-MSE (b) H-MSE

Fig. 4.3 Estimation MSEs of the strongest path vs δ^2, from perfect CSI $\delta^2 = 0$ to no CSI $\delta^2 = 1$ for number of BS antennas $M = 32, 128$, and 512. $K = 10$ UTs are served simultaneously with f=28 GHz, $d = 0.5\lambda$.

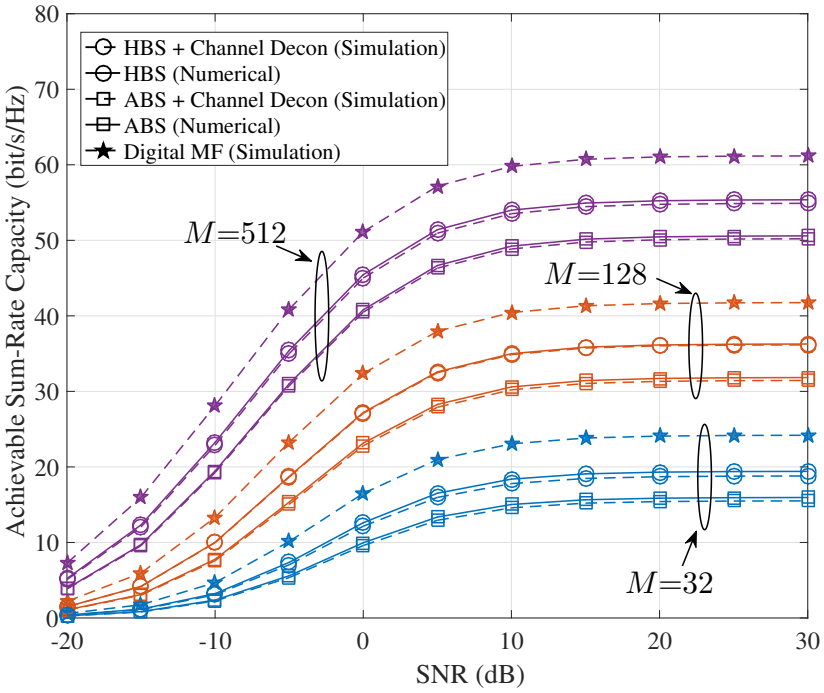


Fig. 4.4 Lower bounds on the sum-rate capacities vs SNR (dB) for ABS, HBS implemented by the proposed hardware architectures, and MF achieved by the unconstrained digital baseband signal processing when $M = 32, 128, 512$ BS antennas, with total number of paths $L = 3$. Perfect CSI is adopted and $K = 10$ UTs are served simultaneously.

the SINRs in (4.45) and (4.56), and presented for verification. The performance of the conventional massive MIMO MF precoding that is executed by the unconstrained digital baseband signal processing is simulated and considered as the optimal system performance assuming perfect CSI is adopted.

Fig. 4.4 shows the achieved lower bounds for the precoding schemes and the impact of SNR on the sum-rate performance assuming perfect CSI is adopted by the proposed channel deconstruction algorithm. $M = 32, 128$ and 512 are evaluated for $K = 10$ and $L = 3$ with the UPA configurations shown in Table. 3.1. It can be observed that the numerical results of the theoretical analysis agree with the simulation results that ascertain the theoretical analysis made and the reliability of the proposed channel deconstruction algorithm. When SNR increases, the lower bounds of all precoding approaches increase and converge to different limits. The HBS outperforms ABS, due to the extra digital precoding. However, the MF outperforms both of the proposed precoders, due to its fully digital baseband control. Although the MF performs the best, the selection of the strongest

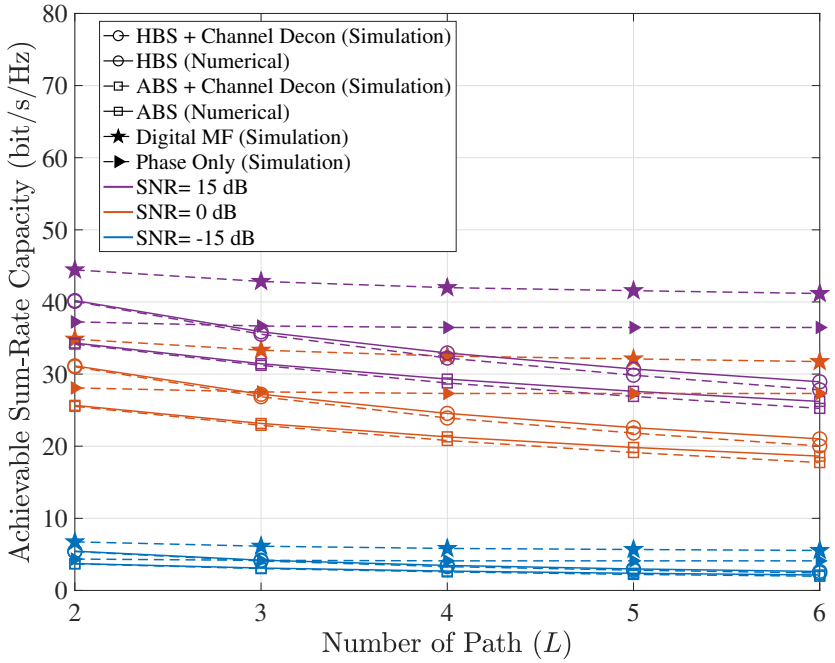


Fig. 4.5 Lower bounds on the sum-rate capacities vs L for ABS, HBS implemented by the proposed hardware architectures, MF achieved by the unconstrained digital baseband signal processing, and phase only control achieved by analogue RF control when SNR= -15 dB, 0 dB, and 15 dB. Perfect CSI is adopted and $K = 10$ UTs are served simultaneously by $M = 128$ BS antennas.

path from the limited-scattering clustered mmWave channel allows a large percentage of the optimal system performance to be achieved by the proposed precoders. For example, the maximum system performance achieved by MF via 128 RF chains and 128×10-dimensional digital baseband precoding is 43 bit/s/Hz, when $M = 128$, $K = 10$. The maximum system performance achieved by HBS precoding via 10 RF chains and 10-dimensional digital baseband processing is 37 bit/s/Hz, which is 86% of the MF capacity. The proposed ABS precoding scheme achieves a maximum of 32 bit/s/Hz capacity via 10 RF chains and only analogue control in the RF domain, which is 74% of the MF capacity. This implies that selecting the strongest estimated path by the proposed precoders contributes substantially to the achievable sum-rate with low-complexity, low-cost and energy-efficient signal processing.

Next, the impact of the total number of received paths L on the sum-rate capacity lower bound is evaluated. Fig. 4.5 demonstrates the simulated and numerical lower bounds for ABS and HBS exploiting the proposed channel deconstruction approach for SNR=15 dB,

0 dB and -15 dB, with $K = 10$ and $M = 128$. The performance of the MF precoder using unconstrained digital baseband signal processing and the phase-only precoding deploying analogue control using the phase of the CSI are also present for comparison. It can be seen that the lower bounds decrease when L increases for all of the precoders except the phase-only control. For ABS and HBS, this is mainly due to the reduction of the desired signal power. As the variance of the unsorted mmWave path $\frac{1}{L}$ decreases with the increase of L, the power of the selected path decreases that can be confirmed directly from (4.38) and (4.50). For MF precoding, the capacity deduction is slight. The ABS precoder has a larger performance reduction as L increases, and the HBS precoder has the largest. As L increases, the performance of the HBS precoder approximates the ABS precoder, e.g. the capacity disparity ΔC between HBS and ABS when SNR=0 dB and $L = 2$ is $\Delta C = 32 - 26 = 6$ but when $L = 6$, $\Delta C = 22 - 19 = 3$. The HBS outperforms the phase-onlly control when $L = 2$ and has approximately the same performance when $L = 3$. It is obvious that the number of received mmWave paths L largely limits the performance of the proposed ABS and HBS precoders. The proposed precoders are more effective when the channel is less scattered, i.e. L is not so large.

Next, the impact of the scale of the BS array on the lower bound is evaluated. Fig. 4.6 demonstrates the simulated and numerical lower bounds on the sum-rate capacities of ABS and HBS exploiting the proposed channel deconstruction approach versus the number of BS antennas M for SNR=15 dB, 0 dB, and -15 dB, when $K = 10$ and $L = 3$. The configurations of the UPA at the BS are shown in Table 3.1. The MF precoding is again simulated for comparison. As expected, the lower bounds of all precoders increase with M, which indicates the effectiveness of the proposed ABS and HBS precoding methods when the BS scales up.

The impact of the CSI accuracy on the proposed precoders' lower bounds is examined in Fig. 4.7. The simulated sum-rate lower bounds of ABS and HBS are presented versus the channel imperfection factor δ^2 for $L = 2, 3$, and 4, SNR=15 dB, and $K = 10$. The effective SINRs are calculated from (4.31) and (4.47). The proposed channel deconstruction approach is adopted for the path parameters' estimation considering that the BS has no knowledge of δ^2. It is shown that as δ^2 increases from 0 to 1, i.e. as the CSI becomes

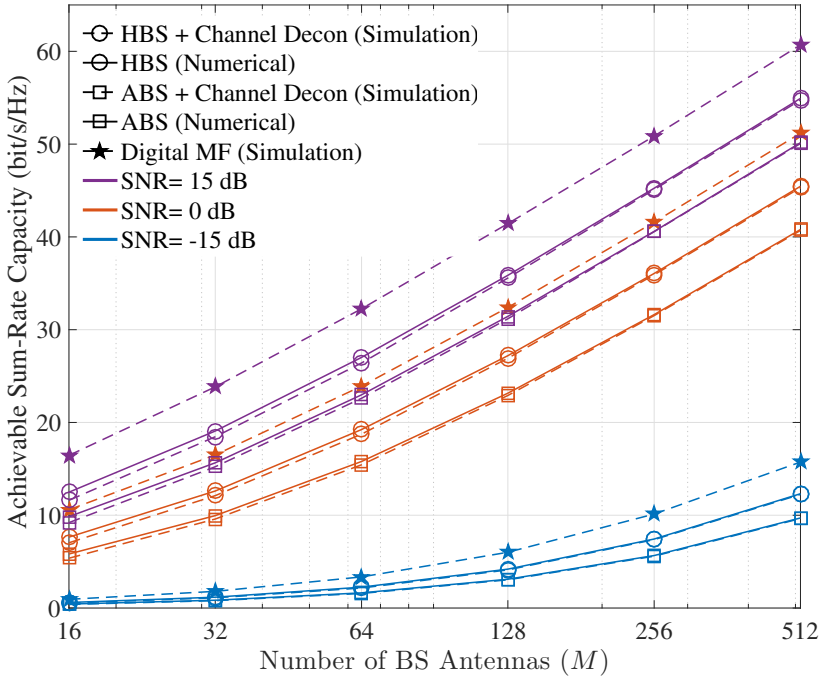


Fig. 4.6 Lower bounds on the sum-rate capacities vs M, for ABS, HBS implemented by the proposed hardware architectures, and MF achieved by the unconstrained digital baseband signal processing when SNR= -15 dB, 0 dB, and 15 dB. Perfect CSI is adopted. Total number of paths $L = 3$ and $K = 10$ UTs are served simultaneously.

more inaccurate, each lower bound remains stable at its maximum value at first, then decreases to zero as $\delta^2 \to 1$. The range of δ^2 over which the lower bound remains largely unaffected is determined by M. For example, when $M = 128$, $L = 2$, the lower bound of HBS is 40 bit/s/Hz for $0 \leq \delta^2 \leq 0.6$. As M increases, the range of δ^2 for maintaining the maximum lower bound becomes wider, e.g. when $M = 512$, 59 bit/s/Hz of the lower bound is kept over $0 \leq \delta^2 \leq 0.8$ that leads to a sharper degradation of the capacity when $\delta^2 \to 1$. This implies that the proposed ABS and HBS precoding methods in cooperation with the proposed channel deconstruction algorithm tolerate imperfect CSI over a wide range of δ^2 as M increases. Therefore, the CSI estimation accuracy requirement is relaxed to a great extent for the TDD-mode massive MIMO systems.

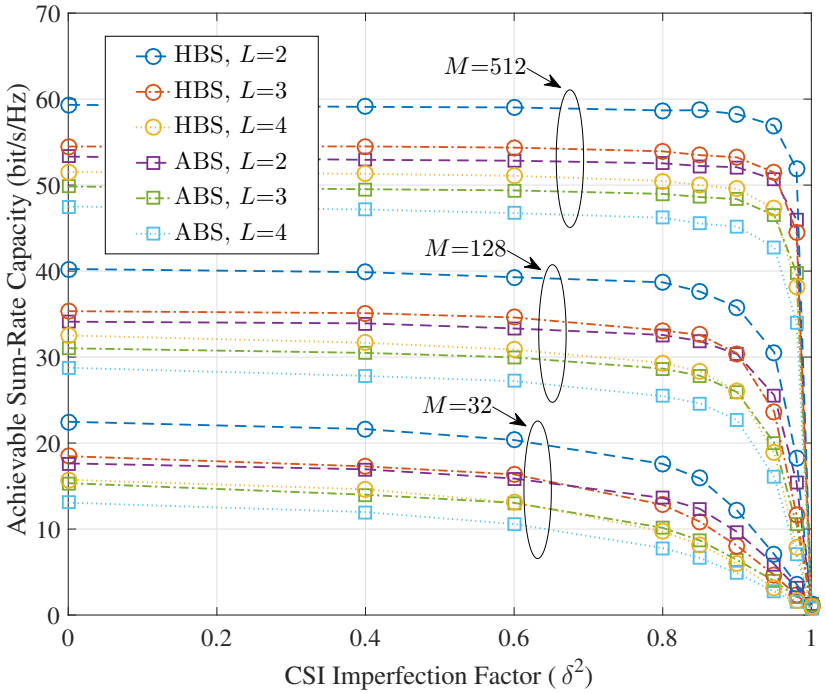


Fig. 4.7 Simulated lower bounds on the sum-rate capacities vs CSI imperfection factor δ^2 for ABS and HBS. δ^2 increases from $\delta^2 = 0$ (perfect CSI) to $\delta^2 = 1$ (no CSI). $M = 32, 128, 512$ and $L = 2, 3, 4$. $K = 10$ UTs are served simultaneously and SNR=15 dB.

4.8 Summary

Building on the findings of Chapter 2, which have shown the potential of mmWave massive MIMO in 5G networks, this chapter investigated the TDD-mode mmWave massive MIMO downlink transmissions.

The signal processing challenge of the large-scale mmWave transmissions and the complexity of the large-scale multi-user precoding are jointly considered. A channel deconstruction approach is first proposed for the estimation of the path parameters from the CSI. Exploiting the limited-scattering clustered property of mmWave channel and the asymptotic orthogonal property of the massive MIMO channel, it is proven that the proposed channel deconstruction approach is effective for massive MIMO systems even when the CSI is Gaussian-perturbed. The analogue and hybrid MSE performance of the estimation has been evaluated and compared over the number of BS antennas, the number of mmWave paths and the CSI imperfection factor. The estimation results show that the

strongest mmWave path parameters for each UT can be easily and accurately estimated with the perfect or Gaussian-perturbed CSI.

The challenge for the design of the low-complexity, low-cost, and energy-efficient mmWave massive MIMO precoding scheme is transformed into a mmWave path-selecting problem. The ABS and HBS precoders are proposed, that select the strongest path for linear precoding. Designs of the corresponding system hardware architectures are also demonstrated. A large-dimensional analogue control in RF domain is adopted for the ABS precoding architecture, based on which, an extra small-dimensional digital precoding in baseband is added, that forms the HBS precoding architecture. Both architectures deploy only K RF chains, and are superior to the conventional massive MIMO architecture, which requires M RF chains and a large-dimensional digital baseband signal processing, in terms of complexity, cost, and energy-efficiency of signal processing. The theoretical results of the effective SINRs and the corresponding sum-rate capacity lower bounds of the proposed precoders are derived using random matrix theory and order statistic principles. The system performance of the lower bound on the sum-rate capacity for the proposed precoders is examined through simulations and numerical analysis of the theoretical results, and is compared with the simulated performance of the MF precoder executed by the unconstrained digital baseband signal processing.

The results show that the HBS precoder outperforms the ABS precoder due to the extra digital control. Although the optimal system performance of MF adopting M RF chains cannot be achieved by the proposed precoders, the selection of the estimated strongest path by the proposed precoders with the deployment of only K RF chains contributes substantially to the system performance. Moreover, when the path parameters utilised for the precoding are estimated from the Gaussian-perturbed CSI, the lower bounds of the proposed ABS and HBS remain at the maximum value over a wide range of the CSI imperfection factor value. Therefore, the proposed ABS and HBS in cooperation with the channel deconstruction approach are effective in achieving a large proportion of the system capacity with low complexity, high energy efficiency, and low cost compared with the full digital baseband signal processing MF precoding. However, the performance of ABS and HBS are primarily limited by the number of received mmWave paths L.

In the upcoming chapter, an enhanced hybrid beam-selecting precoding algorithm is proposed along with its system architecture based on the proposed HBS precoder to achieve the optimal sum-rate capacity lower bound performance while adopting less RF chain and low-dimensional digital baseband precoding.

Chapter 5

MmWave Massive MIMO Enhanced Hybrid Precoding

5.1 Introduction

In Chapter 4, the low-complexity, low-cost and energy efficient TDD-mode mmWave massive MIMO downlink precoding was investigated. The limited-scattering clustered mmWave channel model was adopted, wherein a few clusters of paths contribute to the mmWave channel. The mmWave signal processing and large-scale multi-user precoding challenges were tackled and transformed into a multi-user mmWave beam-selecting problem. A mmWave channel deconstruction approach was proposed that enables the estimation of individual path parameters exploiting the asymptotic orthogonal property of the massive MIMO channel and the limited-scattering clustered characteristic of the mmWave channel. An analogue beam-selecting precoding approach and a hybrid beam-selecting precoding approach were proposed that utilise the estimated parameters of the strongest path by the proposed channel deconstruction algorithm. The corresponding hardware architectures were designed that employ phase only RF control of the transmitted signal and K RF chains for the multi-user transmissions. An additional low-dimensional digital baseband control is adopted in the hardware architecture of the HBS precoder. Compared with the conventional massive MIMO signal processing, in which full digital baseband

control and M RF chains are deployed, the proposed precoders reduce the signal processing complexity and fabrication cost while increasing the energy efficiency.

The theoretical and simulated system performance for the proposed precoders was validated. The results showed that although the optimal system performance obtained by MF precoder implementing the unconstrained digital baseband signal processing cannot be achieved by the proposed ABS and HBS precoders, the selection of the strongest estimated path by the proposed precoders substantially contributes to the system achievable rate. It is also revealed that the performance of the proposed precoders is primarily limited by the number of total received mmWave paths.

In this chapter, the objective is to achieve the optimal TDD-mode mmWave massive MIMO downlink precoding system performance through a low-cost, low-complexity, and energy-efficient approach. Recall that the proposed channel deconstruction approach is proven to be capable of obtaining parameters for all of the received paths by subtracting the estimated path from the CSI and repeating the estimation process. The selection of more estimated paths by the precoder is therefore enabled. To this end, an enhanced hybrid beam-selecting (E-HBS) precoder is proposed in this chapter that exploits the limited-scattering clustered property of the mmWave channel for the further enhancement of the system performance with respect to the ABS and HBS precoding schemes proposed in Chapter 4. Moreover, the corresponding hardware architecture for E-HBS can be constructed by connecting a few of the same substructures in parallel with each substructure being the duplication of the hardware architecture adopted by the proposed HBS precoder. With this hardware design, the number of RF chains and the dimensions of the baseband digital precoding adopted are not related to the number of BS antennas, but to the number of paths selected and the number of UTs served. This is vital for massive MIMO systems, as the signal processing complexity, hardware, and energy costs for the precoding no longer increase with the scaling up of the BS array while achieving the optimal system performance.

The theoretical effective SINR and the lower bound on the sum-rate capacity of the proposed E-HBS precoding for an arbitrary number of path selections are analysed assuming that perfect path parameters are present. It is noticed that the closed-form

expressions for the variances of the inter-cluster and inter-user interference terms in the effective SINR are intractable when the AoDs follow arbitrary distributions. However, the analysis reveals that it is the correlation of the path vector elements which impacts the strength of the interference. Hence, the interference analysis is transformed into the examination of the correlation between the elements of each path vector. Using random matrix theory, the uncorrelated and fully correlated cases are then examined, which yield the average effective SINR and lower bound on the effective SINR, respectively.

The performance of the proposed channel deconstruction algorithm on the estimation of each path's parameters is examined through Monte Carlo simulations, which validate the reliability of the estimated path parameters. The lower bound on the sum-rate capacity of the proposed E-HBS precoding is simulated and compared with the numerical result of the theoretical analysis under the uncorrelated assumption for the elements of each path vector. It is demonstrated that the optimal system performance of MF precoding using the unconstrained digital baseband signal processing can be achieved by the E-HBS precoding when all of the received paths are selected. Moreover, the near-optimal performance is obtained by the selection of only a few paths due to the limited-scattering clustered property of the mmWave signal. Therefore, the energy, fabrication costs, and complexity in massive MIMO signal processing can be significantly reduced. Simulations are also conducted to examine the effects of the AoD distributions on the system performance. The analysis of the average and lower bound on the effective SINR is also verified. The collective impact of AoD distributions and the number of selected paths on the system sum-rate are demonstrated through simulations. It is concluded that to achieve a certain capacity with high energy efficiency, low signal processing complexity and cost, the SNR level and AoD distributions must be jointly considered when determining the number of paths to select by the proposed E-HBS, such that the system resources' utilisation is maximised.

5.2 System model and channel model

The TDD-mode massive MIMO system model utilised in Chapter 4 is adopted again in this chapter, in which a BS equipped with an M-element antenna array is serving K single-antenna UTs simultaneously with $M \gg K$. The mmWave limited-scattering clustered channel model introduced in (4.1) is also considered, where the total number of received paths is denoted by L. Assume that the ordered individual path parameters, i.e. the ordered complex path gains $\alpha_k^{(l)}$ and the corresponding array response vectors $\mathbf{a}_k^{(l)}$ for $\forall l = \{1,\ldots,L\}$, $\forall k = \{1,\ldots,K\}$ are estimated by the channel deconstruction algorithm as proposed in (4.14), following the descending order of the norm of the complex path gain, i.e. $|\alpha_k^{(L)}| \geq \cdots \geq |\alpha_k^{(1)}|$. The estimated complex path gain of the l-th ordered path is denoted as $\hat{\alpha}_k^{(l)}$, and the corresponding estimated array response vector is denoted as $\hat{\mathbf{a}}_k^{(l)}(\hat{\theta}_k^{(l)}, \hat{\varphi}_k^{(l)})$.

5.3 Enhanced hybrid beam-selecting precoding

The objective is to develop a precoding approach that achieves the optimal system performance with respect to the MF precoding executed by the unconstrained digital baseband signal processing. Denote the number of selected paths for each UT as $\boldsymbol{\ell}$, with $1 \leq \boldsymbol{\ell} \leq L$. Consider $\boldsymbol{\ell}$ paths are estimated by the channel deconstruction method at the BS, which means that the BS obtains the estimated parameters of the L-th,…,$(L-\boldsymbol{\ell}+1)$-th ordered mmWave paths of each UT, i.e. the complex path gains $\hat{\alpha}_k^{(L)},\ldots,\hat{\alpha}_k^{(L-\boldsymbol{\ell}+1)}$ and the corresponding array response vectors $\hat{\mathbf{a}}_k^{(L)}(\hat{\theta}_k^{(L)}, \hat{\varphi}_k^{(L)})$,…,$\hat{\mathbf{a}}_k^{(L-\boldsymbol{\ell}+1)}(\hat{\theta}_k^{(L-\boldsymbol{\ell}+1)}, \hat{\varphi}_k^{(L-\boldsymbol{\ell}+1)})$, for $\forall k \in \{1,\ldots,K\}$. Note that the number of paths selected for each UT can be different. For simplicity, the same number of $\boldsymbol{\ell}$ is considered in this chapter as the same channel statistics are assumed at each UT.

The enhanced hybrid beam-selecting (E-HBS) precoding vector $\mathbf{e}_k \in \mathcal{C}^{M\times 1}$ for the k-th UT utilises the estimated parameters of $\boldsymbol{\ell}$ ordered paths and is expressed as

$$\mathbf{e}_k = \left(\sum_{l=L-\boldsymbol{\ell}+1}^{L} \hat{\alpha}_k^{(l)} \hat{\mathbf{a}}_k^{(l)}(\hat{\theta}_k^{(l)}, \hat{\varphi}_k^{(l)}) \right)^H. \tag{5.1}$$

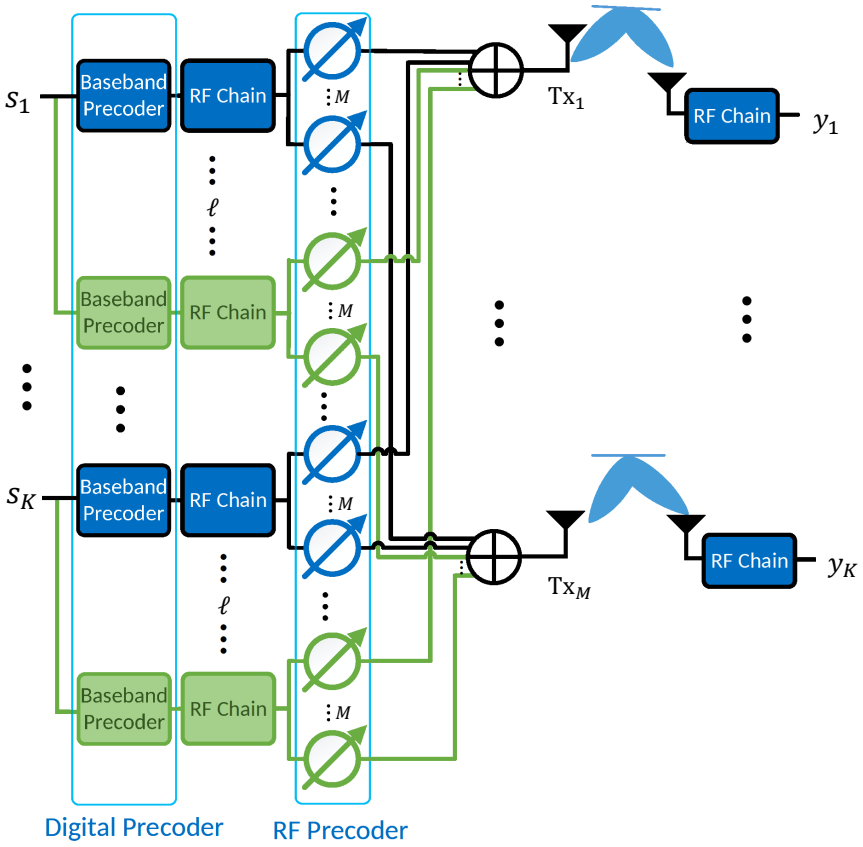


Fig. 5.1 Hardware architecture of the proposed E-HBS precoding for multi-user massive MIMO systems implemented by ℓ substructures including $K \times \ell$ RF chains, followed by $K \times M \times \ell$ RF phase shifters and M adders.

Hence, the E-HBS precoding matrix is formed following

$$\mathbf{W}_E = \frac{1}{\sqrt{\gamma_E}}\mathbf{E} = \frac{1}{\sqrt{\gamma_E}}[\mathbf{e}_1, \ldots, \mathbf{e}_K], \tag{5.2}$$

where

$$\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_K], \tag{5.3}$$

and $\mathbf{W}_E \in \mathcal{C}^{M \times K}$. γ_E denotes the E-HBS power normalisation factor which is given by

$$\gamma_E = \mathbb{E}[trace(\mathbf{E}\mathbf{E}^H)]. \tag{5.4}$$

From (5.1), the proposed E-HBS precoder utilises the limited-scattering clustered property of the mmWave channel that also facilitates the hardware architecture to be obtained from the proposed architecture of HBS precoding. As depicted in Fig. 5.1, ℓ

substructures are deployed in parallel for each transmit symbol by duplicating the HBS architecture as shown in Fig. 4.1. Each substructure implements the hybrid precoding for one of the selected paths. Therefore, each transmit symbol s_k passes through $\boldsymbol{\ell}$ baseband digital controls and $\boldsymbol{\ell}$ RF chains. Meanwhile, each transmit antenna is now connected to $K \times \boldsymbol{\ell}$ phase shifters via an exclusive adder. The resulting structure can be viewed as a spatial RAKE system.

5.4 E-HBS effective SINR and lower bound on the sum-rate capacity

In this section, the theoretical effective SINR and the lower bound on the sum-rate capacity of the proposed E-HBS precoder are derived for the selection of the $\boldsymbol{\ell}$ paths. In the analysis of the effective SINR, it is noticed that the closed-form expressions for the variances of interference terms are intractable when AoDs follow an arbitrary distribution. However, it is the correlation between the elements of path vectors that limits the strength of the received interference. Hence the uncorrelated and fully correlated cases, which respectively yield the average and lower bound on the effective SINR are analysed.

Consider that the perfect path parameters are estimated by the proposed channel deconstruction algorithm as shown in (4.14). Therefore, from (4.19) and (4.21), $\varepsilon_{k,m}^{A,(l)} = \varepsilon_{k,m}^{H,(l)} = 0$, $\alpha_k^{(l)} = \hat{\alpha}_k^{(l)}$, and $\mathbf{a}_k^{(l)} = \hat{\mathbf{a}}_k^{(l)}$.

From (2.6),(5.1),(5.2), and (5.4), with E-HBS precoding, the received signal of the k-th UT can be denoted as

$$
\begin{aligned}
y_k = & \underbrace{\sqrt{\frac{\rho}{\gamma_E}}\Big(\sum_{l=L-\boldsymbol{\ell}+1}^{L} \alpha_k^{(l)}\mathbf{a}_k^{(l)}\Big)\Big(\sum_{l=L-\boldsymbol{\ell}+1}^{L} \alpha_k^{(l)}\mathbf{a}_k^{(l)}\Big)^H s_k}_{\text{Desired Signal}} \\
& + \underbrace{\sqrt{\frac{\rho}{\gamma_E}}\Big(\sum_{i=1}^{L-\boldsymbol{\ell}} \alpha_k^{(i)}\mathbf{a}_k^{(i)}\Big)\Big(\sum_{l=L-\boldsymbol{\ell}+1}^{L} \alpha_k^{(l)}\mathbf{a}_k^{(l)}\Big)^H s_k}_{\text{Inter - cluster Interference}} \\
& + \underbrace{\sqrt{\frac{\rho}{\gamma_E}}\sum_{\substack{q=1,\\ q\neq k}}^{K}\Big(\sum_{i=1}^{L} \alpha_k^{(i)}\mathbf{a}_k^{(i)}\Big)\Big(\sum_{l=L-\boldsymbol{\ell}+1}^{L} \alpha_q^{(l)}\mathbf{a}_q^{(l)}\Big)^H s_q}_{\text{Inter - user Interference}} + \underbrace{n_k}_{Noise}.
\end{aligned}
\tag{5.5}
$$

Following the same principle of the lower bound on the system sum-rate capacity analysed in Chapter 2, the effective SINR of the k-th UT when $\boldsymbol{\ell}$ paths are selected by the E-HBS can be denoted from (5.5), and given by

$$\text{SINR}_k^{E\text{-}HBS} = \frac{\frac{\rho}{\gamma_E}\mathbb{E}[|(\sum_{l=L-\boldsymbol{\ell}+1}^{L}\alpha_k^{(l)}\mathbf{a}_k^{(l)})(\sum_{l=L-\boldsymbol{\ell}+1}^{L}\alpha_k^{(l)}\mathbf{a}_k^{(l)})^H|^2]}{\frac{\rho}{\gamma_E}\mathbb{E}[|(\sum_{i=1}^{L-\boldsymbol{\ell}}\alpha_k^{(i)}\mathbf{a}_k^{(i)})(\sum_{l=L-\boldsymbol{\ell}+1}^{L}\alpha_k^{(l)}\mathbf{a}_k^{(l)})^H|^2] + \frac{\rho}{\gamma_E}\mathbb{E}[|\sum_{\substack{q=1\\q\neq k}}^{K}(\sum_{i=1}^{L}\alpha_k^{(i)}\mathbf{a}_k^{(i)})(\sum_{l=L-\boldsymbol{\ell}+1}^{L}\alpha_q^{(l)}\mathbf{a}_q^{(l)})^H|^2] + 1}. \tag{5.6}$$

- From (5.6), the expected value of the desired receive signal power of the k-UT can be denoted as

$$\mathbb{E}[|(\sum_{l=L-\boldsymbol{\ell}+1}^{L}\alpha_k^{(l)}\mathbf{a}_k^{(l)})(\sum_{l=L-\boldsymbol{\ell}+1}^{L}\alpha_k^{(l)}\mathbf{a}_k^{(l)})^H|^2] = M^2\mathbb{E}[(\sum_{l=L-\boldsymbol{\ell}+1}^{L}|\alpha_k^{(l)}|^2)^2] \tag{5.7}$$

  Let random variable $X = |\alpha_k^l|^2$, where the i.i.d. complex path gain is $\alpha_k^l \sim \mathcal{CN}(0, \frac{1}{L})$. Hence, X follows the gamma distribution with $X \sim \Gamma(1, \frac{1}{L})$, as illustrated in Chapter 4 section 4.6.1. For L samples of X, the p.d.f. on the l-th ordered $X_{(l)}$ can be denoted as [99]

$$f(x_{(l)}) = \frac{L!}{(l-1)!\,(L-l)!}[F_\Gamma(x)]^{l-1}[1-F_\Gamma(x)]^{L-l}f_\Gamma(x), \tag{5.8}$$

  where the p.d.f. $f_\Gamma(x)$ and c.d.f. $F_\Gamma(x)$ on X is denoted in (4.34) and (4.35). The joint p.d.f. $f(x_{L-\boldsymbol{\ell}+1},\ldots,x_L)$ of $\boldsymbol{\ell}$ samples of ordered random variables $(X_{(L-\boldsymbol{\ell}+1)},\ldots,X_{(L)})$ that correspond to the variances of complex paths gains for $\boldsymbol{\ell}$ selected paths can be denoted as [99, 101]

$$f(x_{(L-\boldsymbol{\ell}+1)},\ldots,x_{(L)}) = \frac{L!}{(L-\boldsymbol{\ell})!}F_\Gamma^{L-\boldsymbol{\ell}}(x_{(L-\boldsymbol{\ell}+1)})f_\Gamma(x_{(L-\boldsymbol{\ell}+1)})\ldots f_\Gamma(x_{(L)}). \tag{5.9}$$

  For $\boldsymbol{\ell} = 1,\ldots,L$, the above joint p.d.f. is given as follows:

  - When the number of selected paths $\boldsymbol{\ell} = 1$,

$$f\left(x_{(L)}\right) = L^2\left(1-e^{-Lx_{(L)}}\right)^{L-1}e^{-Lx_{(L)}}. \tag{5.10}$$

    This case corresponds to the HBS precoding, i.e. the strongest estimated path is selected.

– When the number of selected paths $\boldsymbol{\ell} = 2$, the L-th and $(L-1)$-th strongest estimated paths are selected,

$$f\left(x_{(L-1)}, x_{(L)}\right) = L^3 (L-1) (1 - e^{-Lx_{(L-1)}})^{L-2} e^{-Lx_{(L-1)}} e^{-Lx_{(L)}}. \tag{5.11}$$

– .....

– When the number of selected paths $\boldsymbol{\ell} = L$, i.e. all of the L estimated paths are selected by the E-HBS precoder,

$$f\left(x_{(1)}, \ldots, x_{(L)}\right) = L^L L! e^{-L\left(\sum_{l=1}^{L} x_{(l)}\right)}. \tag{5.12}$$

Further, the expected value of the power of the desired signal in (5.7) can be calculated by the joint p.d.f. on the order statistics of the variances of $\boldsymbol{\ell}$ selected complex path gains as shown in (5.9), and given by

$$\begin{aligned} &M^2 \mathbb{E}[(\sum_{l=L-\boldsymbol{\ell}+1}^{L} |\alpha_k^{(l)}|^2)^2] \\ =&M^2 \int \ldots \int \left(x_{(L-\boldsymbol{\ell}+1)} + \ldots + x_{(L)}\right)^2 f\left(x_{(L-\boldsymbol{\ell}+1)}, \ldots, x_{(L)}\right) dx_{(L-\boldsymbol{\ell}+1)} \ldots dx_{(L)}. \end{aligned} \tag{5.13}$$

- The power normalisation factor γ_E can be drived from (5.4) and given by

$$\begin{aligned} \gamma_E =&\mathbb{E}[trace(\mathbf{E}\mathbf{E}^H)] \\ =&MK\mathbb{E}[\sum_{l=L-\boldsymbol{\ell}+1}^{L} |\alpha_k^{(l)}|^2] \\ =&MK \int \ldots \int \left(x_{(L-\boldsymbol{\ell}+1)} + \ldots + x_{(L)}\right) f\left(x_{(L-\boldsymbol{\ell}+1)}, \ldots, x_{(L)}\right) dx_{(L-\boldsymbol{\ell}+1)} \ldots dx_{(L)}. \end{aligned} \tag{5.14}$$

- From the analysis of the interference of massive MIMO LoS transmissions in Chapter 3, and the ABS and HBS mmWave massive MIMO transmissions in 4, the closed-form expressions of the expected values of the power of the inter-cluster and inter-user interference in (5.5) are intractable when AoDs follow the uniform distribution or other common distributions such as Gaussian or Laplacian [42, 102]. However, it is the phase shifts $\mathrm{a}_{k,1}^{(l)}, \ldots, \mathrm{a}_{k,M}^{(l)}$ in the array response vector $\mathbf{a}_k^{(l)}$ that impact on the correlation between elements of each path vector $\alpha_k^{(l)} \mathbf{a}_k^{(l)}$ for $\forall m \in \{1, \ldots, M\}$ and hence limit the strength of the received interference [103, 104].

The actual correlation between elements $\alpha_k^{(l)}\mathrm{a}_{k,m}^{(l)}$ and $\alpha_k^{(l)}\mathrm{a}_{k,n}^{(l)}$ of the path vector for $\forall m,n \in \{1,\ldots,M\}$, $m \neq n$ is limited by many factors, such as the AoDs, carrier frequency f, the separation between the antennas d, the number of antennas M, and the distribution of $\alpha_k^{(l)}$ as shown in (4.2) and (4.3). This results in enormous difficulties in interference analysis for the multiple selection of the paths by the proposed E-HBS precoder. However, as analysed in Chapter 3 section 3.5.1, when the elements of each path vector are i.i.d. Gaussian distributed, i.e. the elements of each path are uncorrelated following the same Gaussian distribution, which lead to the average interference level. Also, when the elements of each path vector are identical, i.e. fully correlated with the complex path gain of each path following Gaussian distribution as a result of LoS propagation with $\theta_k^l = 90^\circ$ and $\varphi_k^l = 0^\circ$ for all k and l, the maximum cross-correlation, which produces the maximum interference is obtained.

Define the m-th element of the inter-cluster interference from the unselected paths on the selected paths for the k-th UT as

$$v_{k,m} \triangleq (\sum_{i=1}^{L-\ell} \alpha_{k,m}^{(i)} a_{k,m}^{(i)})(\sum_{l=L-\ell+1}^{L} \alpha_{k,m}^{(l)} a_{k,m}^{(l)})^*, \tag{5.15}$$

and denote the inter-cluster interference for the k-th UT in (5.5) as

$$I_k^c \triangleq \sum_{m=1}^{M} v_{k,m}. \tag{5.16}$$

Hence, the expected value of the power of I_k^c is given from (5.15) by

$$\mathbb{E}[|I_k^c|^2] = \mathbb{E}[|(\sum_{i=1}^{L-\ell} \alpha_k^{(i)} \mathbf{a}_k^{(i)})(\sum_{l=L-\ell+1}^{L} \alpha_k^{(l)} \mathbf{a}_k^{(l)})^H|^2] = \mathbb{E}[|\sum_{m=1}^{M} v_{k,m}|^2]. \tag{5.17}$$

Note that in (5.15), the m-th element of the i-th and l-th ordered path vectors $\alpha_{k,m}^{(i)}\mathrm{a}_{k,m}^{(i)}$ and $\alpha_{k,m}^{(l)}\mathrm{a}_{k,m}^{(l)}$ are independent (due to independent AoDs) with mean zero and variance $\sigma_{(i)}^2$ and $\sigma_{(l)}^2$, for $\forall i,l \in \{1,\ldots,L\}, i \neq l$. Note also that the unit magnitude element $\mathrm{a}_{k,m}^{(l)}$ of the array response vector $\mathbf{a}_k^{(l)}$ for all m does not change the variance or the mean of the corresponding complex path gain $\alpha_k^{(l)}$ for $\forall l \in \{1,\ldots,L\}$. Hence, $v_{k,m}$ for $\forall m \in \{1,...,M\}, \forall k \in \{1,...,K\}$ is an i.i.d. random variable with mean zero. Since the expected value of $v_{k,m}$ is zero, the

variance of $v_{k,m}$ is denoted from (5.15), and given by

$$\begin{aligned}
&variance(v_{k,m}) = \sigma_c^2 = \mathbb{E}[|v_{k,m}|^2] - \mathbb{E}[v_{k,m}]^2 \overset{(a)}{=} \mathbb{E}[|v_{k,m}|^2] \\
&\overset{(b)}{=} (\sum_{i=1}^{L-\ell} \sigma_{(i)}^2)(\sum_{l=L-\ell+1}^{L} \sigma_{(l)}^2) \\
&= \left(\sum_{i=1}^{L-\ell} X_{(i)}\right)\left(\sum_{l=L-\ell+1}^{L} X_{(l)}\right) \\
&= \int ... \int \left(x_{(1)} + ... + x_{(L-\ell)}\right)\left(x_{(L-\ell+1)} + ... + x_{(L)}\right) f\left(x_{(1)},...,x_{(L)}\right) dx_{(1)} \ldots dx_{(L)},
\end{aligned} \tag{5.18}$$

where (a) holds since $\mathbb{E}[v_{k,m}]^2 = 0$ and (b) is from (5.15) with $\sigma_{(l)}^2$ denotes the variance of the complex path gain $\alpha_k^{(l)}$ of the l-th ordered paths, and $(X_{(1)},\ldots,X_{(L)})$ follow the joint p.d.f. of L samples of X with $f\left(x_{(1)},...,x_{(L)}\right)$ given by (5.12).

– The uncorrelated case for the elements of each path vector

  Assume $\alpha_k^{(l)} \mathrm{a}_{k,m}^{(l)}$ for the l-th ordered path of the k-th UT is i.i.d. $\sim \mathcal{CN}(0, \sigma_{(l)}^2)$ for $\forall m \in \{1,...,M\}$, $\forall l \in \{1,\ldots,L\}$. Hence from (5.16), the variance of I_k^c is given by

$$variance(I_k^c) = M variance(v_{k,m}). \tag{5.19}$$

  As the variances of the complex path gains of the ordered paths are jointly distributed, the expected value of the power of the average inter-cluster interference in (5.17) is then denoted from (5.19), and given by

$$\begin{aligned}
\mathbb{E}[|I_k^c|^2] &= \mathbb{E}[|(\sum_{i=1}^{L-\ell} \alpha_k^{(i)} \mathbf{a}_k^{(i)})(\sum_{l=L-\ell+1}^{L} \alpha_k^{(l)} \mathbf{a}_k^{(l)})^H|^2] \\
&= variance(I_k^c) + (\mathbb{E}[I_k^c])^2 \\
&= M\sigma_c^2,
\end{aligned} \tag{5.20}$$

  where σ_c^2 is given in (5.18) and $\mathbb{E}[I_k^c] = 0$. Note that when $\ell = 1$, the uncorrelated case considered here corresponds to the analysis utilised for HBS inter-cluster interference, see (4.52) and (4.53).

– The fully correlated case for the elements of each path vector

  Assume $\alpha_k^{(l)} \mathrm{a}_{k,m}^{(l)} = \alpha_k^{(l)} \mathrm{a}_{k,n}^{(l)} = \alpha_k^{(l)}$ for all $\forall m,n\{1,\ldots,M\}$. This case is the result of the LoS transmission with $\theta_k^l = 90^\circ$ and $\varphi_k^l = 0^\circ$ for all of the UTs and all of the mmWave

paths. Hence, from (5.16)

$$variance(I_k^c) = variance(Mv_{k,m}) = M^2 variance(v_{k,m}) \tag{5.21}$$

The array response vector $\mathbf{a}_k^{(l)}$ for $\forall l = \{1,\ldots,L\}$ and $\forall k = \{1,\ldots,K\}$ becomes a unit vector. Therefore, the maximum expected value of the power of the inter-cluster interference in (5.17) is given from (5.21), and expressed as

$$\begin{aligned}\mathbb{E}[|I_k^c|^2] &= \mathbb{E}[|(\sum_{i=1}^{L-\ell} \alpha_k^{(i)}\mathbf{a}_k^{(i)})(\sum_{l=L-\ell+1}^{L} \alpha_k^{(l)}\mathbf{a}_k^{(l)})^H|^2] \\ &= variance(I_k^c) + (\mathbb{E}[I_k^c])^2 \\ &= M^2\sigma_c^2.\end{aligned} \tag{5.22}$$

where σ_c^2 is given in (5.18) and $\mathbb{E}[I_k^c] = 0$.

Compare the expected values of the power of the average and maximum inter-cluster interference denoted in (5.20) and (5.22), the fully correlated elements of each path vector yields an M times higher expected value of the inter-cluster interference than the uncorrelated case, due to the identical elements of each path vector.

- For the inter-user interference, similar correlation analysis can be implemented following the inter-cluster interference analysis. Define the inter-user interference on the k-th UT as

$$I_k^u \triangleq \sum_{\substack{q=1,\\ q\neq k}}^{K} \sum_{m=1}^{M} u_{k,m}, \tag{5.23}$$

and also define the m-th element of the inter-user interference from the q-th UT on the k-th UT as

$$u_{k,m} \triangleq (\sum_{i=1}^{L} \alpha_k^{(i)} \mathrm{a}_{k,m}^{(i)})(\sum_{l=L-\ell+1}^{L} \alpha_q^{(l)} \mathrm{a}_{q,m}^{(l)})^*. \tag{5.24}$$

The expected value of the power of the inter-user interference in (5.5) is hence obtained from (5.23) and (5.24), and further given by

$$\begin{aligned}\mathbb{E}[|I_k^u|^2] &= \mathbb{E}[|(\sum_{i=1}^{L} \alpha_k^{(i)}\mathbf{a}_k^{(i)})(\sum_{\substack{q=1\\ q\neq k}}^{K} (\sum_{l=L-\ell+1}^{L} \alpha_q^{(l)}\mathbf{a}_q^{(l)})^H)|^2] \\ &= (K-1)\mathbb{E}[|\sum_{m=1}^{M} u_{k,m}|^2].\end{aligned} \tag{5.25}$$

In (5.24), since $\alpha_k^l \sim \mathcal{CN}(0,\frac{1}{L})$, therefore, $\sum_{i=1}^{L} \alpha_k^{(i)} a_{k,m}^{(i)} \sim \mathcal{CN}(0,1)$ for $\forall m = \{1,...,M\}, \forall l = \{1,...,L\}, \forall k = \{1,...,K\}$. $u_{k,m}$ for $\forall m \in \{1,...,M\}, \forall k \in \{1,...,K\}$ is a i.i.d. random variable with mean zero. The variance of $u_{k,m}$ is obtained from (5.24), and given by

$$\begin{aligned} & variance(u_{k,m}) = \sigma_u^2 \\ &= \sum_{l=L-\ell+1}^{L} \sigma_{(l)}^2 \\ &= \sum_{l=L-\ell+1}^{L} X_{(l)} \\ &= \int ... \int \left(x_{(L-\ell+1)} + ... + x_{(L)}\right) f\left(x_{(L-\ell+1)},...,x_{(L)}\right) dx_{(L-\ell+1)} ... dx_{(L)}. \end{aligned} \tag{5.26}$$

where $(X_{(L-\ell+1)},\ldots,X_{(L)})$ follow the joint p.d.f. on ℓ samples of X denoted in (5.9).

– The uncorrelated case for the elements of each path vector

  Assume the m-th element $\alpha_k^{(l)} a_{k,m}^{(l)}$ for the l-th ordered path of the k-th UT is i.i.d. $\sim \mathcal{CN}(0,\sigma_{(l)}^2)$ for $\forall m \in \{1,...,M\}$. Since the variances of the complex path gains of the ordered paths are jointly distributed, the variance of I_k^u is obtained from (5.23) and given by

$$variance(I_k^u) = M(K-1)variance(u_{k,m}). \tag{5.27}$$

  From (5.25), the expected value of the power of the average inter-user interference is hence given by

$$\begin{aligned} \mathbb{E}[|I_k^u|^2] &= variance(I_k^u) + (\mathbb{E}[I_k^u])^2 \\ &= M(K-1)\sigma_u^2, \end{aligned} \tag{5.28}$$

  where σ_u^2 is given by (5.26) and $\mathbb{E}[I_k^u] = 0$. Note that when $\ell = 1$, this case corresponds to the HBS inter-user interference analysis in (4.55).

– The fully correlated case for the elements of each path vector

  Using the same assumption made in the fully correlated analysis for the inter-cluster interference, the variance of the inter-user interference is obtained from (5.23) and given by

$$variance(I_k^u) = (K-1)variance(Mu_{k,m}) = M^2(K-1)variance(u_{k,m}). \tag{5.29}$$

Hence, the expected value of the power of the maximum inter-user interference in (5.25) is given by

$$\begin{aligned}\mathbb{E}[|I_k^u|^2] &= M^2 variance(u_{k,m}) \\ &= M^2(K-1)\sigma_u^2.\end{aligned} \tag{5.30}$$

where σ_u^2 is given by (5.26).

Again, from (5.28) and (5.30), the full correlated elements of each path vector yield an M times higher inter-user interference compared with the uncorrelated case.

So far, the uncorrelated and fully correlated cases for the elements of the path vectors are demonstrated that respectively yield the average and maximum inter-cluster and inter-user interference. Hence the average effective SINR can be obtained when both interference terms in the effective SINR have their average values. From (5.13),(5.9),(5.18),(5.26),(5.20) and (5.28), the average effective SINR is given by

$$\text{SINR}_{avg}^{E\text{-}HBS} = \frac{\frac{\rho}{\gamma_E} M^2 \int \ldots \int \left(x_{(L-\ell+1)} + \ldots + x_{(L)}\right)^2 f\left(x_{(L-\ell+1)}, \ldots, x_{(L)}\right) dx_{(L-\ell+1)} \ldots dx_{(L)}}{\frac{\rho}{\gamma_E} M\sigma_c^2 + \frac{\rho}{\gamma_E} M(K-1)\sigma_u^2 + 1}. \tag{5.31}$$

The lower bound on the effective SINR can be obtained when both interferences have their maximum values. From (5.13),(5.9),(5.18),(5.26),(5.22), and (5.30), the lower bound on the effective SINR is given by

$$\text{SINR}_{low}^{E\text{-}HBS} = \frac{\frac{\rho}{\gamma_E} M^2 \int \ldots \int \left(x_{(L-\ell+1)} + \ldots + x_{(L)}\right)^2 f\left(x_{(L-\ell+1)}, \ldots, x_{(L)}\right) dx_{(L-\ell+1)} \ldots dx_{(L)}}{\frac{\rho}{\gamma_E} M^2\sigma_c^2 + \frac{\rho}{\gamma_E} M^2(K-1)\sigma_u^2 + 1}, \tag{5.32}$$

where the joint p.d.f. $f\left(x_{(L-\ell+1)}, \ldots, x_{(L)}\right)$ is given by (5.9), σ_c^2 and σ_u^2 are given by (5.18) and (5.26), and γ_E is in (5.14).

The above average effective SINR and lower bound on the effective SINR correspond to the average and minimum lower bounds on the system sum-rate that can be obtained by substituting (5.31) and (5.32) into (2.19).

5.5 Simulation results for channel deconstruction algorithm and E-HBS precoding

In this section, the effectiveness of the proposed channel deconstruction approach on the estimation of all of the ordered mmWave paths is validated through Monte Carlo simulations. The simulation results of the performance of A-MSE and H-MSE are present for comparison. Such validation is essential since the estimated parameters of mmWave paths will be exploited by the E-HBS precoder and hence limit the precoding performance.

The effectiveness of the proposed E-HBS precoder on achieving the optimal system performance obtained by MF precoding executed by the unconstrained digital baseband signal processing is evaluated through simulations for the lower bound on the sum-rate capacity exploiting the path parameters estimated by the channel deconstruction algorithm. The impact of the number of selected paths is hence examined.

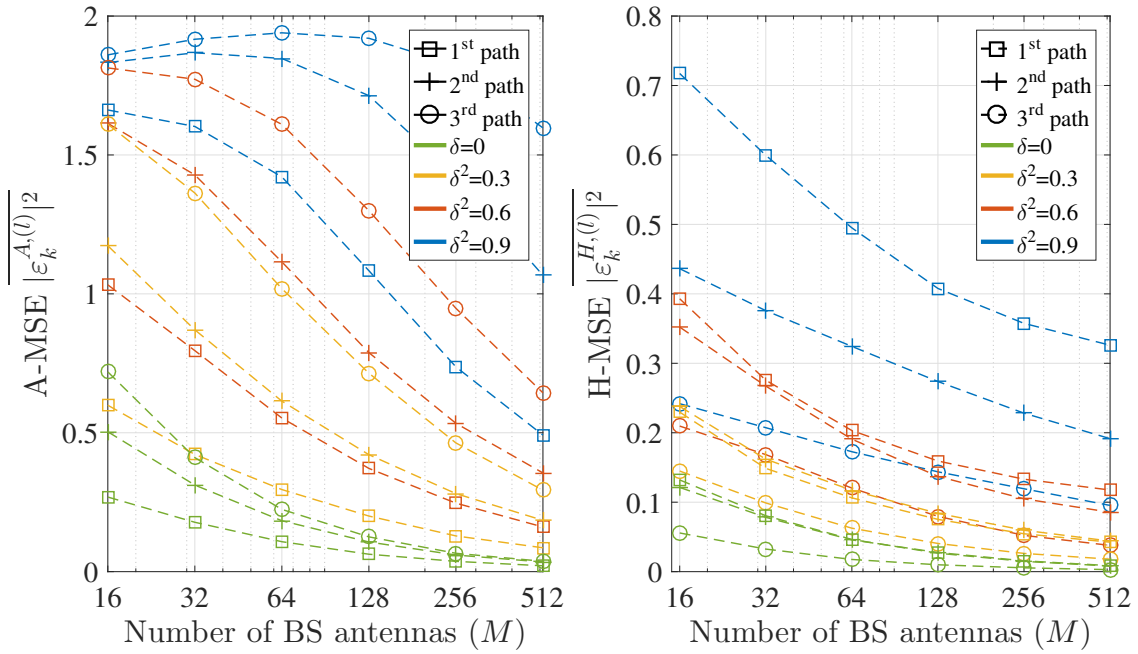
The average effective SINR and lower bound on the effective SINR are then evaluated and verified by the evaluation of the system performance for different AoDs distributions. The collective impact of the AoD distributions and the number of selected paths on the system performance is then simulated and discussed.

In the following experiments, single-cell mmWave massive MIMO downlink transmissions from a BS equipped with an M-element UPA to K single-antenna UTs are considered. The limited-scattering clustered mmWave channel model in (4.1) is adopted with a carrier frequency of f=28 GHz and $d = 0.5\lambda$. L clusters of paths are received by each UT with a single path in each cluster and the large-scale fading for each UT $\beta_1 = \cdots = \beta_k$ are assumed to be normalised to unity. The number of selected paths of the E-HBS precoder is ℓ with $1 \leq \ell \leq L$. The complex path gain of each path follows $\alpha_k^l \sim$i.i.d.$\mathcal{CN}(0, \frac{1}{L})$ for $\forall l \in \{1\ldots,L\}$, $\forall k \in \{1\ldots,K\}$. The AoDs of each path are modelled as i.i.d. random variables following uniform distributions. For the mmWave channel deconstruction algorithm, it is considered that the number of training beam spatial directions N is large enough that the training beam spatial directions cover the exact actual AoDs of each path. The simulated result of the lower bound on the sum-rate capacity is calculated by (2.19) with the simulated effective SINR calculated by (5.6).

5.5.1 MSE performance of channel deconstruction approach on estimating parameters for each mmWave path

Fig. 5.2a and Fig. 5.2b demonstrate the estimation MSEs for all of the ordered paths versus the number of antennas M when $L = 3$ under different CSI quality levels, i.e. $\delta^2 = 0, 0.3, 0.6$, and 0.9. The UPA configurations are taken from Table.3.1. The channel deconstruction is performed using (4.14) assuming the BS has no knowledge of δ^2. The A-MSE and H-MSE are calculated using (4.20) and (4.22). It can be observed from both figures that the analogue and hybrid MSEs on the estimation of each path decrease when M increases. Moreover, when imperfect CSI is adopted, as δ^2 decreases, i.e. the CSI becomes more accurate, the estimation MSEs of each path decreases, which results in a more precise estimation of paths' parameters.

For the analogue MSE performance, under a certain level of CSI quality δ^2 and M, the order of the estimation MSEs follows the order of the expected value of the path power, i.e. the estimation of the path with a higher $\mathbb{E}[|\alpha_k^{(l)}|^2]$ has a lower estimation MSE. However, for hybrid MSE, an opposite trend is shown. The overall H-MSE is lower than the A-MSE. Note that it is the path coefficients that will be utilised in the E-HBS precoding, the proposed channel deconstruction



(a) A-MSE (b) H-MSE

Fig. 5.2 Estimation MSEs for each ordered path vs M. The total number of paths $L = 3$. The perfect CSI $\delta^2 = 0$ and imperfect CSI $\delta^2 = 0.3, 0.6$, and 0.9 are considered . $K = 10$ UTs are served simultaneously.

algorithm is therefore reliable for the estimation of all mmWave paths' parameters with the scaling up of the BS array.

5.5.2 System performance of the lower bound on the sum-rate for E-HBS precoding

In this section, the lower bounds on the system sum-rate capacities for the proposed E-HBS precoding utilising the estimated results of the channel deconstruction approach are evaluated considering $M = M_x \times M_y = 16 \times 8 = 128$, $K = 10$, $L = 5$ and the number of selected path is ℓ, with $1 \leq \ell \leq L$. In the first part of the evaluation, the effectiveness of the E-HBS precoder on achieving the optimal system performance is evaluated by comparing the system performance of E-HBS with the MF precoding implemented by the unconstrained digital baseband signal processing when ℓ paths are selected. The uncorrelated elements, i.e. i.i.d. Gaussian distributed elements for each path vector are considered and simulated. Simulation results are also compared with numerical results from the theoretical average effective SINR denoted in (5.31). The second part further evaluates the proposed average effective SINR and lower bound on the effective SINR and demonstrates the impact of AoDs' distribution on the system performance of E-HBS. Lower bounds on the sum-rate for the presented average and lower bound on effective SINRs obtained from the uncorrelated and fully correlated path vector elements are presented and compared with the simulation results for different AoDs distributions.

For Fig. 5.3 and Fig. 5.4, consider that the number of selected paths is ℓ with $\ell = 1, \ldots, L$, and the total number of mmWave paths is $L = 5$. The paths' parameters used by E-HBS are estimated by the channel deconstruction approach with perfect CSI. The uncorrelated elements, i.e. i.i.d. Gaussian distributed elements for each path vector is considered and simulated. The numerical results are calculated by (5.31) and (2.19). The AoDs of each path are modelled with $\theta_k^{(l)} \sim$ i.i.d.$\mathbf{U}[0, 90^\circ]$ and $\varphi_k^{(l)} \sim$ i.i.d.$\mathbf{U}[-180^\circ, 180^\circ]$ for $\forall k \in \{1, \ldots, K\}$ and $\forall l \in \{1, \ldots, L\}$.

Fig. 5.3 shows the lower bounds on the sum-rate capacities for E-HBS versus the SNR (dB) when the number of selected paths is $\ell = 1, \ldots, 5$. The system performance of MF is illustrated for comparison. It is evident that the simulated results of E-HBS are consistent with the numerical results. For each case of ℓ from 1 to L, the lower bound increases with SNR to different non-zero limit due to the presence of the inter-cluster and inter-user interference. Note also that under the same level of SNR (dB), when more paths are selected, i.e. with a higher ℓ, a higher lower bound

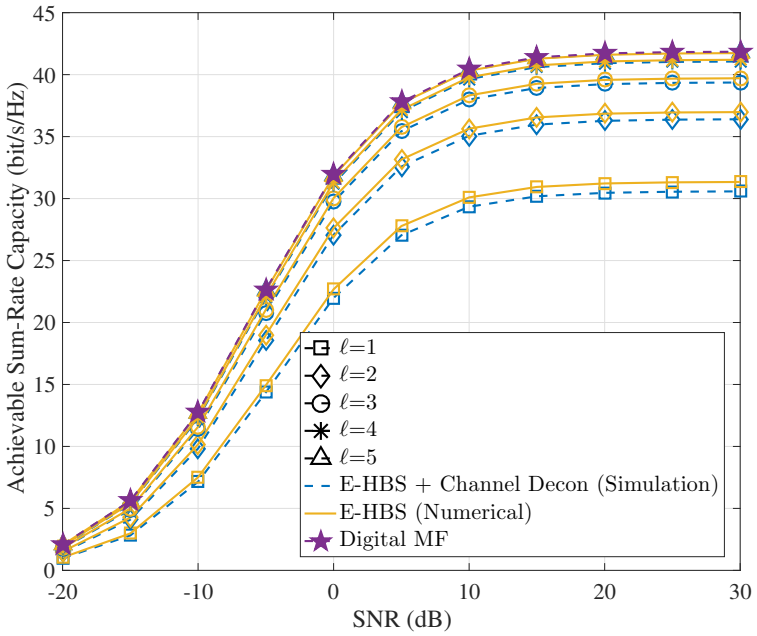


Fig. 5.3 Lower bounds on the sum-rate capacities achieved by E-HBS vs SNR (dB) for the number of selected paths $\ell = 1,2,3,4,5$. $M = M_x \times M_y = 16 \times 8 = 128, L = 5, K = 10$ are simulated. The uncorrelated elements, i.e. i.i.d. Gaussian distributed elements for each path vector is considered.

is achieved. Moreover, when $\ell = L = 5$, i.e. all of the paths are selected, the E-HBS precoding achieves the optimal system performance that is achieved by MF precoding. In this case, the number of RF chains adopted by E-HBS is only $K \times L = 50$, however, $M = 128$ RF chains are required for the unconstrained digital baseband signal processing of MF. Therefore, regardless of the hardware imperfection, the E-HBS precoding scheme facilitates the optimal system performance with low complexity, substantial energy efficiency improvement and reduced fabrication cost by selecting all of the estimated mmWave paths.

Next, the impact of the number of selected paths on the system performance is further evaluated. Fig. 5.4 shows the lower bound on the sum-rate capacity for the E-HBS precoder versus the number of selected paths ℓ when SNR= -15 dB, 0 dB, and 15 dB. It is obvious that the lower bound increases when more paths are selected. The increase rate against ℓ is not linear. A higher increase rate is presented when ℓ is relatively small, while a lower increase rate is shown when ℓ is relatively high, especially for high SNR level. For example, for SNR=15 dB, when the number of selected paths ℓ increases from 1 to 2, the lower bound increase is $\Delta C = 36 - 30 = 6$ bit/s/Hz, however when ℓ increases from 4 to 5, $\Delta C = 41.5 - 41 = 0.5$ bit/s/Hz. This is due to the fact that, the paths are

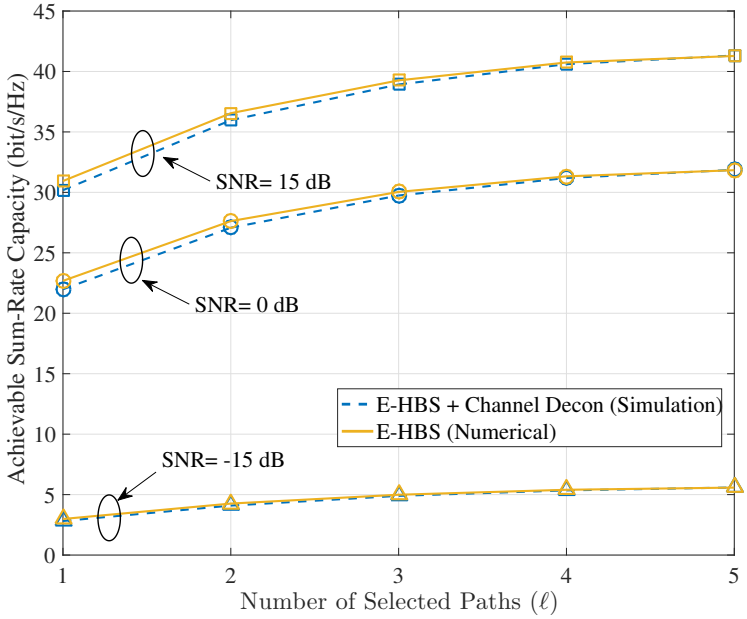


Fig. 5.4 Lower bounds on the sum-rate capacities vs the number of selected paths ℓ for SNR=15 dB, 0 dB, -15 dB. $M = 128, L = 5, K = 10$ are simulated. The uncorrelated elements, i.e. i.i.d. Gaussian distributed elements for each path vector is considered.

Table 5.1 The sum-rate lower bound achieved by the E-HBS over the optimal performance of MF precoding implemented by the unconstrained digital baseband signal processing and the corresponding E-HBS hardware configuration for SNR=15 dB, $L = 5$, $f = 28$ GHz.

	E-HBS					Digital MF
Number of paths selected	$\ell = 1$	$\ell = 2$	$\ell = 3$	$\ell = 4$	$\ell = 5$	N/A
% of capacity achieved over digital MF	76%	90%	96%	99%	100%	100%
Number of RF chains	10	20	30	40	50	128
Number of analogue phase shifters	1280	2560	3840	5120	6400	Digital

selected following the decreasing order of the path power, and the smaller the power is, the less sum-rate contribution that path provides.

Table 5.1 presents the percentage of the lower bound on the sum-rate capacity achieved by E-HBS over the optimal performance obtained by the MF precoding using the unconstrained digital baseband signal processing taken from the SNR=15 dB curve in Fig. 5.4. The signal processing hardware configuration of the number of RF chains and analogue phase shifters is also indicated in the table. It can be seen that 99% of the optimal system performance can be achieved by only selecting $\ell = 4$ paths for E-HBS obtained with only 40 RF chains. Compared with the digital baseband signal processing that requires 128 RF chains, 88 RF chains are saved by the E-HBS

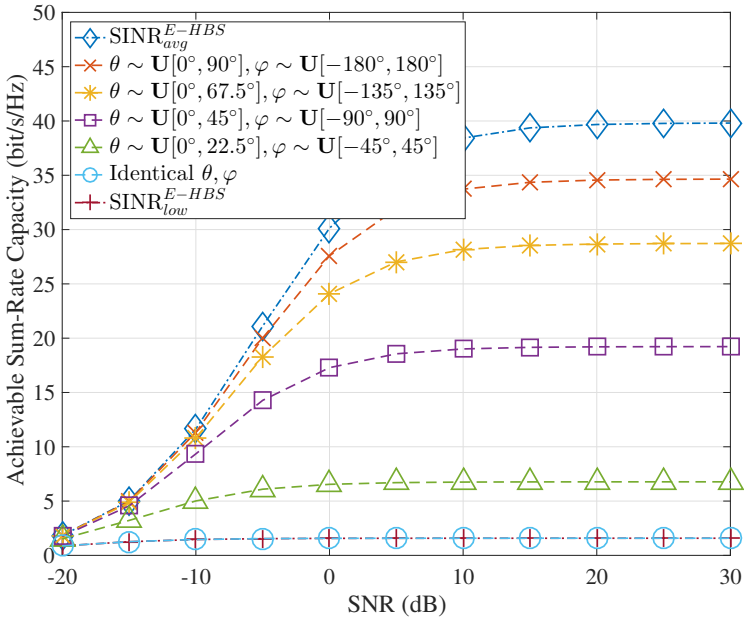


Fig. 5.5 Simulation results of the lower bounds on the system sum-rate capacities of E-HBS precoding versus SNR (dB) for different AoDs distributions, average effective SINR and lower bound on the effective SINR. $M = M_x \times M_y = 16 \times 8 = 128, L = 5, K = 10, \ell = 3$ are considered with the perfect path selection. The actual interference characterised by different AoDs distributions is utilised in the effective SINR.

precoder. This reveals that there's a substantial opportunity to further reduce the number of RF chains adopted and the dimensions of the digital and analogue controls for E-HBS precoding to achieve a near-optimal system performance.

Next, the impact of AoDs distribution on the lower bound performance of E-HBS precoding is investigated alone with the validation of the average effective SINR and lower bound on the effective SINR analysed in section 5.4. Consider $M = M_x \times M_y = 128, L = 5$ and $K = 10$. AoDs follow i.i.d. uniform distributions with $\theta_k^{(l)} \sim \mathbf{U}[0, \theta_{max}]$ and $\varphi_k^{(l)} \sim \mathbf{U}[\varphi_{min}, \varphi_{max}]$. Therefore, the ranges of AoDs distribution is varied by θ_{max}, φ_{min} and φ_{max}. The results are obtained using the actual interference simulated for different AoDs distributions as the simulations for mean-square cross-correlation in Chapter 3 Fig. 3.1 and Fig. 3.2.

Fig. 5.5 illustrates the system performance achieved by the E-HBS when $\ell = 3$ versus SNR (dB). It is shown that the performance of the lower bound on the sum-rate when AoDs are uniformly distributed remains within the average (from SINR$_{avg}^{E-HBS}$) and minimum (from SINR$_{low}^{E-HBS}$) lower bounds on the sum-rate capacity. When AoDs have identical values with $\theta_1^{(l)} = \cdots = \theta_K^{(l)} = 90°$ and $\varphi_1^{(l)} = \cdots = \varphi_K^{(l)} = 0°$ for all k and l, the performance matches the performance of the lower bound on the effective SINR, i.e. SINR$_{low}^{E-HBS}$ that comes from the fully correlated identical elements of

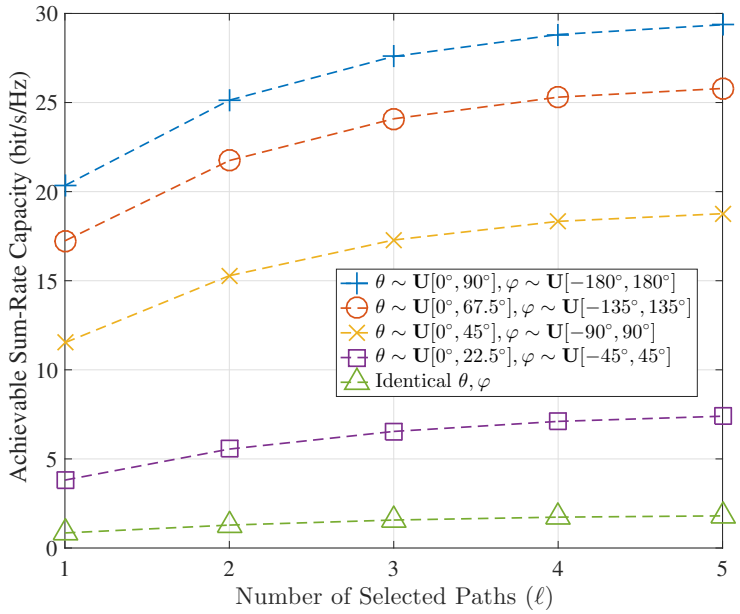


Fig. 5.6 Simulation results of the lower bounds on the system sum-rate capacities vs the number of selected paths ℓ by E-HBS for different AoDs distributions. $M = M_x \times M_y = 16 \times 8 = 128, L = 5, K = 10$, SNR=0 dB are considered with the perfect path selection. The actual interference characterised by different AoDs distributions is utilised in the effective SINR.

each path vector. With the enlargement of the AoDs distribution range, i.e. θ_{max} and $(\varphi_{max} - \varphi_{min})$, the system performance improves at the same level of SNR. However, for sparse channels, even when the range of AoDs distribution covers the whole geometric angle range, i.e. $\theta_{max} = 90°$ and $\varphi_{max} - \varphi_{min} = 360°$, the performance cannot reach the performance of $\text{SINR}_{avg}^{E-HBS}$. This is also justified in [31] as i.i.d. path vector elements are incompatible with the wave equation when AoDs are random variables. It is evident that the system performance of the mmWave massive MIMO system exploiting E-HBS precoding is limited by the distributions of AoDs. The system performance enhances when the distribution of AoDs is within a wider range, as the correlation between elements of the path vectors decreases.

Fig. 5.6 shows the collective effect of the AoDs distribution and the number of paths selected on the system performance of E-HBS. The simulated lower bounds on the sum-rate capacities for E-HBS versus the number of selected paths ℓ for different AoDs' distributions are shown. Explicit relationships between the number of selected paths ℓ, the AoDs distribution ranges θ_{max}, $(\varphi_{max} - \varphi_{min})$, and the system performance are illustrated. It is shown that not only the increase of ℓ but also a wider range of the AoD distribution enables a higher system performance. For instance, to achieve a capacity lower bound of 25 bit/s/Hz, 4 paths have to be selected out of 5 received paths

when $\theta_k^{(l)} \sim \mathbf{U}[0^\circ, 67.5^\circ]$, $\varphi_k^{(l)} \sim \mathbf{U}[-135^\circ, 135^\circ]$, however, only two paths need to be selected when AoDs have a wider distribution range with $\theta_k^{(l)} \sim \mathbf{U}[0^\circ, 90^\circ]$, $\varphi_k^{(l)} \sim \mathbf{U}[-180^\circ, 180^\circ]$.

Combined with the relationship between SNR, ℓ, and the system performance shown in Fig. 5.4, it is vital for the BS to jointly consider the AoDs distribution, SNR level, and the objective system performance when choosing the number of paths to select for E-HBS precoding in mmWave massive MIMO transmissions, such that the desired performance can be achieved without wasting extra system resources on estimating extra paths by the channel deconstruction approach or wasting costs on the fabrication and energy by transmitting with more RF chains, analogue phase shifters, and digital controls.

5.6 Summary

In this chapter, the objective is to design a low-complexity, low-cost linear precoding approach that achieves the near optimal system performance for mmWave massive MIMO transmissions with higher energy efficiency. An E-HBS precoder is proposed which enables the selection of ℓ mmWave paths estimated by the proposed channel deconstruction approach from L received paths. The advanced hardware architecture design is proposed for E-HBS precoding, wherein the number of RF chains adopted is decided by the number of paths selected and the number of UTs served, which significantly reduces the number of RF chains required compared with the conventional massive MIMO digital baseband signal processing. Using random matrix theory and order statistic principles, the theoretical results of the effective SINRs and the lower bounds on the sum-rate capacities for the proposed E-HBS precoding are presented for different ℓ, where the average effective SINR and lower bound on the effective SINR are demonstrated for the interference analysis.

The effectiveness of the proposed channel deconstruction approach on the estimation of each mmWave path's parameters is validated. It is illustrated that such an approach is especially reliable for E-HBS precoding. The system performance for E-HBS precoding when $\ell = 1, \ldots, L$ is then simulated. The results suggest that the system performance increases with more selected paths. Moreover, the optimal system performance can be achieved when all of the received paths are selected. Therefore, high energy efficiency and low fabrication cost are provided by the proposed linear precoder of E-HBS, especially when M increases compared with the conventional digital precoding method.

The impact of the number of selected paths on the system performance of E-HBS is presented. It is shown that a near-optimal performance can be achieved by E-HBS without selecting all of the received paths that yields the opportunity to further improve the signal processing energy efficiency and reduce the complexity and cost.

The impact of the AoDs distribution on the system performance is also examined. It is demonstrated that a wider distribution range of AoDs facilitates an enhanced system performance. It is suggested that in order to achieve the desired system performance, the AoDs distribution and SNR level should be jointly considered when determining the number of paths to select for the E-HBS precoding, such that system resources will not be wasted on estimating unnecessary parameters or transmitting with redundant RF chains, and digital or phase controls.

Chapter 6

Spatial User Scheduling for Sub-6 GHz Massive MIMO LoS Transmission

6.1 Introduction

The arguments in Chapter 3 revealed that, for LoS transmissions in sub-6 GHz TDD-mode massive MIMO systems, although the MF and ZF precoding schemes using the statistical LoS CSI are promising with low-complexity, a severe system performance degradation may occur when transmitting to randomly located UTs with spatial AoDs of the UTs distributed within a small range. The analysis has shown that the LoS channel cross-correlation primarily limits the system performance especially when the BS scale is moderately large. To enhance the system performance of LoS transmissions, a user scheduling scheme is imperative that allows the BS to transmit to selected UTs following a scheduling policy.

In principle, a user scheduling policy aiming to maximise the system sum-rate capacity can be achieved by exhaustively searching for the optimal group of UTs from all possible combinations of the active UTs. However, with the scaling up of the system, such an exhaustive search is time and resource consuming, along with having a high overhead and is therefore not applicable.

Recent research has been focused on user scheduling algorithms to enable performance improvement for over-loaded multi-user MIMO and massive MIMO systems in which the number of active UTs is much larger than the number of antennas at the BS. In [105], a multi-user MIMO system where the UTs are equipped with multiple antennas is considered. By representing the

mutual interference as an inter-user subspace correlation, user scheduling schemes based on the greedy algorithm are proposed, which take into account the fairness among users and the maximisation of the system sum-rate. The proposed algorithms achieve high performance with low computational complexity. However, the channel is modelled by the Kronecker Product Form model for the correlated Rayleigh fading channel. Such a model is infeasible in LoS transmissions since it fails to address the LoS channel characteristic regarding to the array response. Furthermore, the proposed algorithms are based on the metrics of mutual interference among users, which is measured in angular or subspace domains. Such metrics are not suitable for massive MIMO systems in which each UT is equipped with a single antenna, and the complexity will become extremely high in large-scale systems. In [106], the beam-space beam selection and user scheduling are jointly considered to implement the Lyapunov-drift optimisation for maximising the system sum-rate in the limited-scattering clustered channel. Although near-optimal performance can be achieved by the proposed user scheduling algorithm, the adopted channel model again does not capture the sub-6 GHz LoS transmission feature, and the algorithm uses a highly complex iterative method, which increases the costs of the system resources. In [63, 107], the authors assume that UTs can be divided into groups and within each group, the UTs share similar slow-fading characteristics. The authors proposed an algorithm that first groups the UTs by the eigenvectors of the channel covariance matrix such that the linear subspace of the channel matrix is distinguishable among groups. The pre-beamforming is implemented using the channel covariance matrices. The ZF precoding is then executed for the interference elimination within each group. However, the UTs must feedback the effective CSI after the pre-beamforming, and the performance relies on the number of UTs per group.

The existing research on the user scheduling scheme for LoS transmissions is very limited. In this chapter, the objective is to develop a low-complexity user scheduling scheme to improve the system performance for downlink LoS transmissions in TDD-mode massive MIMO systems. To this end, first, the LoS channel cross-correlations for 2-D transmissions from a ULA and 3-D transmissions from a UPA are examined assuming the active UTs are spatially continuously distributed. The correlation-free channel constraints are then demonstrated, which reveal the fact that with the knowledge of the UTs' spatial AoDs, for a base UT, a group of UTs with orthogonal LoS channels can be discovered in 2-D LoS transmissions from a ULA and two groups of UTs with orthogonal LoS channels within each group can be found in 3-D LoS transmissions from a UPA. It is also demonstrated that the number of UTs within each group is restricted by the space between

antenna elements, the carrier frequency, the configuration of the antenna array at the BS, and the spatial AoDs of the base UT. For realistic LoS transmissions, in which the active UTs have discrete distribution, two novel spatial user scheduling schemes are then proposed for each transmission scenario based on the knowledge of the statistical CSI of the UTs' spatial AoDs. The proposed spatial user scheduling schemes enhance the system performance by reducing the LoS channel cross-correlation with low complexity and no feedback. Due to the slow-varying characteristic of the statistical CSI, the knowledge of AoDs can be easily acquired from statistical CSI of the LoS channel, which can be obtained from the channel averaging or long-term feedback as introduced in Chapter 3 section 3.1. Since such schemes do not eliminate the channel cross-correlation, to further improve the system performance, a capacity-enhancement check scheme is proposed to cooperate with the proposed spatial user scheduling schemes, by which the UTs that are harmful to the system sum-rate are excluded from the scheduled group. The system performance, the average number of scheduled UTs per group, and the system performance gain of the proposed schemes over the random user scheduling scheme are investigated through simulations in 3-D LoS transmissions from a UPA. The simulation results show that the proposed spatial user scheduling scheme outperforms the random user scheduling scheme with a high gain. When executed with the proposed capacity-enhancement check scheme, the spatial user scheduling scheme achieves an even higher performance improvement.

6.2 System model

Consider a sub-6 GHz TDD-mode massive MIMO system, in which a BS is equipped with an M-element array and communicating with $\mathcal{K}$ active single-antenna UTs. The channels between the BS and the UTs are LoS[†]. The system is overloaded, which means $M \ll \mathcal{K}$. Therefore, the BS is unable to serve all $\mathcal{K}$ active UTs simultaneously, and user scheduling is required, such that only a subset of UTs will be served each time the BS performs LoS transmissions. Assume that, $\mathcal{T}$ scheduling instances are implemented during a coherent interval.

The lower bound on the system sum-rate capacity introduced in Chapter 3 section 3.4.1 is again considered to be the system performance metric. The objective is to improve the system performance through the scheduling of a subset of the UTs from $\mathcal{K}$ active UTs. To this end, an exhaustive search can be implemented for the optimal solution. However, in a massive MIMO system with $\mathcal{K} \gg M$,

[†]The channels can be Rician fading, in this case, since the power of the LoS component is stronger than the NLoS component, only LoS is considered for simplicity.

the computational demand is extremely high. For example, if $M = 128$ and $\mathcal{K} = 2M = 256$, the maximum number of UTs that can be served each time is M for independent information transmissions. Therefore, the number of possible combinations of the UTs is $\frac{(2M)!}{(M!)^2} = 5.77 \times 10^{75}$. Clearly, it is unrealistic to perform such an exhaustive search.

6.3 LoS channel cross-correlation and LoS orthogonal users' group

The analysis of MF and ZF precoding presented in Chapter 3 section 3.4 indicates that the LoS channel cross-correlation primarily limits the system performance of LoS transmissions. To gain further insights into the LoS channel cross-correlation, in this section, both ULA and UPA LoS transmission scenarios are investigated, that is 2-D transmissions from a BS equipped with a ULA and 3-D transmissions from a UPA. The LoS orthogonal users' group, wherein the LoS channels of the UTs have no cross-correlation is found for both scenarios.

6.3.1 LoS cross-correlation for 2-D transmissions from a ULA

For a BS equipped with an M-element ULA, define the LoS channel cross-correlation between the k-th UT and the j-th UT as:

$$R \triangleq \mathbf{a}_k \mathbf{a}_j^H = \sum_{m=1}^{M} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m-1)(\cos\varphi_k - \cos\varphi_j)}, \tag{6.1}$$

where $\mathbf{a}_k$ is the LoS channel vector of the k-th UT denoted in (2.9).

Note that the mean-square cross-correlation defined in (3.40) is the expected value of the squared norm of the channel cross-correlation. Hence, if the cross-correlation equals zero, i.e. the LoS channels of the UTs are correlation-free, the mean-square cross-correlation is zero. (6.1) is the geometric series of a geometric progression. The 1^{st} term is unity, the m-th term is $e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m-1)(\cos\varphi_k + \cos\varphi_j)}$, and there are M terms in total. Therefore, the LoS channel cross-correlation can be further derived

as

$$
\begin{aligned}
R = \mathbf{a}_k \mathbf{a}_j^H &= \sum_{m=1}^{M} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m-1)\left(\cos\varphi_k - \cos\varphi_j\right)} \\
&\underline{\underline{\text{geometric series}}} \ \frac{1 - e^{-\mathbf{j}\frac{2\pi d}{\lambda}M\left(\cos\varphi_k - \cos\varphi_j\right)}}{1 - e^{-\mathbf{j}\frac{2\pi d}{\lambda}\left(\cos\varphi_k - \cos\varphi_j\right)}} \\
&= \frac{e^{-\mathbf{j}\frac{\pi d}{\lambda}M\left(\cos\varphi_k - \cos\varphi_j\right)} e^{\mathbf{j}\frac{\pi d}{\lambda}M\left(\cos\varphi_k - \cos\varphi_j\right)} - e^{-\mathbf{j}\frac{\pi d}{\lambda}M\left(\cos\varphi_k - \cos\varphi_j\right)} e^{-\mathbf{j}\frac{\pi d}{\lambda}M\left(\cos\varphi_k - \cos\varphi_j\right)}}{e^{-\mathbf{j}\frac{\pi d}{\lambda}\left(\cos\varphi_k - \cos\varphi_j\right)} e^{\mathbf{j}\frac{\pi d}{\lambda}\left(\cos\varphi_k - \cos\varphi_j\right)} - e^{-\mathbf{j}\frac{\pi d}{\lambda}\left(\cos\varphi_k - \cos\varphi_j\right)} e^{-\mathbf{j}\frac{\pi d}{\lambda}\left(\cos\varphi_k - \cos\varphi_j\right)}} \\
&= \frac{e^{-\mathbf{j}\frac{\pi d}{\lambda}M\left(\cos\varphi_k - \cos\varphi_j\right)} \left(e^{\mathbf{j}\frac{\pi d}{\lambda}M\left(\cos\varphi_k - \cos\varphi_j\right)} - e^{-\mathbf{j}\frac{\pi d}{\lambda}M\left(\cos\varphi_k - \cos\varphi_j\right)}\right)}{e^{-\mathbf{j}\frac{\pi d}{\lambda}\left(\cos\varphi_k - \cos\varphi_j\right)} \left(e^{\mathbf{j}\frac{\pi d}{\lambda}\left(\cos\varphi_k - \cos\varphi_j\right)} - e^{-\mathbf{j}\frac{\pi d}{\lambda}\left(\cos\varphi_k - \cos\varphi_j\right)}\right)} \\
&= e^{-\mathbf{j}\frac{\pi d}{\lambda}(M-1)\left(\cos\varphi_k - \cos\varphi_j\right)} \frac{\sin\left(\frac{\pi d}{\lambda}M\left(\cos\varphi_k - \cos\varphi_j\right)\right)}{\sin\left(\frac{\pi d}{\lambda}\left(\cos\varphi_k - \cos\varphi_j\right)\right)}.
\end{aligned}
\tag{6.2}
$$

Define

$$
\Phi \triangleq \cos\varphi_k - \cos\varphi_j, \tag{6.3}
$$

by substituting (6.3) into (6.2), the LoS channel cross-correlation is further expressed as

$$
R = \mathbf{a}_k \mathbf{a}_j^H = e^{-\mathbf{j}\frac{\pi d}{\lambda}(M-1)\Phi} \frac{\sin\left(\frac{\pi d}{\lambda}M\Phi\right)}{\sin\left(\frac{\pi d}{\lambda}\Phi\right)}. \tag{6.4}
$$

Hence, the LoS channel cross-correlation between the LoS channels of the k-th UT and the j-th UT becomes a function of Φ.

6.3.2 LoS orthogonal users' group for 2-D transmissions from a ULA

In this section, the LoS correlation-free constraints on the AoDs of the UTs in 2-D LoS transmissions from a ULA are analysed through Φ as defined in (6.3). The argument illustrates that if the active UTs are spatially continuously distributed in 2-D space, with respect to the k-th UT, there is a group of UTs that have orthogonal LoS channels within the group. This enables the search for LoS orthogonal UTs to be implemented with information about the spatial AoDs of the active UTs.

To gain the relationship between the AoDs of the UTs and the LoS correlation-free constraints on Φ from (6.4), define the composition sinusoidal function and its numerator and denominator functions as

$$
\mathcal{S} \triangleq \sin(\frac{\pi d}{\lambda}M\Phi)/\sin(\frac{\pi d}{\lambda}\Phi), \tag{6.5}
$$

$$\mathcal{S}^n \triangleq \sin(\frac{\pi d}{\lambda} M\Phi), \tag{6.6}$$

$$\mathcal{S}^d \triangleq \sin(\frac{\pi d}{\lambda} \Phi). \tag{6.7}$$

Consider φ_k and φ_j are continuously distributed from -180° to 180°, then $\cos\varphi_k$ and $\cos\varphi_j$ are continuously distributed within ± 1, hence $\Phi = \cos\varphi_k - \cos\varphi_j \leq \pm 2$. Fig. 6.1 shows the relationship between the sinusoidal functions $\mathcal{S}^d$, $\mathcal{S}^n$, and $\mathcal{S}$ as the functions of Φ assuming $M = 16$ for instance. From (6.4), the LoS channel cross-correlation is eliminated, i.e. $\mathcal{S} = 0$ when the following constraints on the sinusoidal functions are satisfied:

$$\mathcal{S}^n = \sin\left(\frac{\pi d}{\lambda} M\Phi\right) = 0\,, \mathcal{S}^d = \sin\left(\frac{\pi d}{\lambda}\Phi\right) \neq 0. \tag{6.8}$$

From Fig. 6.1, the correlation-free constraints on Φ in (6.8) can be translated to the value of Φ that corresponds to one of the nulls of the composition sinusoidal function of $\mathcal{S}$. Also note that the set of nulls of $\mathcal{S}$ are the set of nulls of $\mathcal{S}^n$ excluding the set of nulls of $\mathcal{S}^d$. Define $N_{\langle k,j\rangle} \in \mathbb{R}$ corresponding to the k-th UT and the j-th UT to be

$$N_{\langle k,j\rangle} \triangleq \frac{d}{\lambda} M(\cos\varphi_k - \cos\varphi_j). \tag{6.9}$$

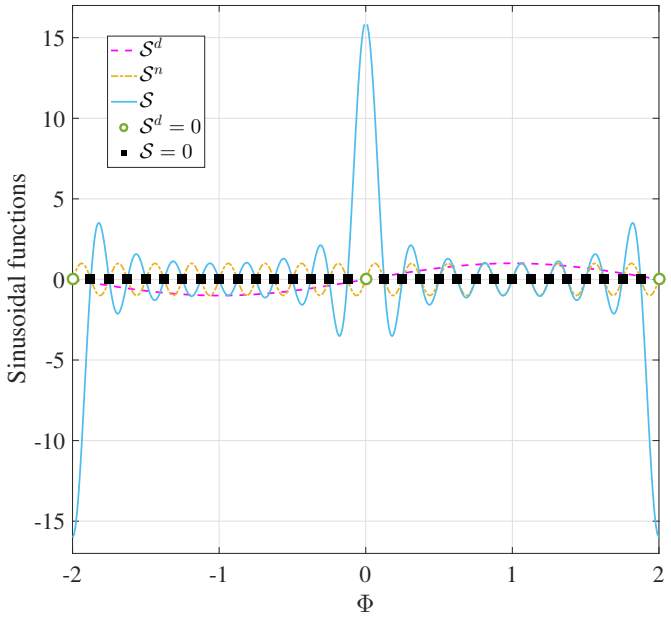


Fig. 6.1 The relationship between the sinusoidal functions $\mathcal{S}^d$, $\mathcal{S}^n$, $\mathcal{S}$ and nulls vs Φ. An ULA is considered with $M = 16, d = 0.5\lambda, f = 2.6$ GHz.

The following theorem demonstrates the LoS correlation-free constraints on AoDs of the UTs and the group of UTs that are LoS orthogonal with respect to the k-th UT.

Theorem 4. *In 2-D LoS transmissions from a BS equipped with a ULA, assume the active UTs are spatially continuously distributed. For the k-th UTs, the LoS channels for a group of UTs: k-th, j-th, …, i-th, are orthogonal when $N_{\langle k,j\rangle}$, …, $N_{\langle k,i\rangle}$ defined in (6.9) correspond to a unique element in set $\mathfrak{N}$, which is given by*

$$\mathfrak{N} = \{\mathbb{Z} : \left((\cos\varphi_k - 1)\frac{d}{\lambda}M, (\cos\varphi_k + 1)\frac{d}{\lambda}M\right)\} - \{0\} - \{\pm 2dM/\lambda\}, \tag{6.10}$$

where $-\{\cdot\}$ denotes the exclusion of the element from a set.

The maximum number of the UTs P in the group (including the k-th UT), i.e. the maximum number of the elements in $\mathfrak{N}$ plus 1, is $|\{\mathbb{Z} : [-\frac{d}{\lambda}M, \frac{d}{\lambda}M]\}|$ when $\cos\varphi_k \neq \pm 1$ or $|\{\mathbb{Z} : [-\frac{d}{\lambda}M, \frac{d}{\lambda}M]\}| - 1$ when $\cos\varphi_k = \pm 1$, where $\mathbb{Z}$ denotes the integer set, $|\{\cdot\}|$ denotes the cardinality of a set, i.e. the number of elements in a set.

Proof. The UTs with zero LoS channel cross-correlation against the k-th UT correspond to the nulls of $\mathcal{S}$ as defined in (6.5). From the periods of sinusoidal functions $\mathcal{S}, \mathcal{S}^n$ and $\mathcal{S}^d$, the correlation-free constraints in (6.8) on Φ with respect to the k-th UT are transferred to the conditions that the AoDs of the k-th UT and the j-th UT must satisfy

$$\frac{\pi d}{\lambda}M\Phi = N\pi, \Phi \neq 0, \Phi \neq \pm 2, \tag{6.11}$$

where $N \in \mathbb{Z}$ is an integer and Φ is defined in (6.3). Therefore, if the j-th UT satisfies (6.11) with respect to the k-th UT, it gives

$$\frac{\pi d}{\lambda}M(\cos\varphi_k - \cos\varphi_j) = N_{\langle k,j\rangle}\pi, (\cos\varphi_k - \cos\varphi_j) \neq 0, (\cos\varphi_k - \cos\varphi_j) \neq \pm 2, \tag{6.12}$$

where $N_{\langle k,j\rangle} \in \mathbb{Z}$ and $N_{\langle k,j\rangle} \neq 0$.

Note that in Fig. 6.1, the range of $\Phi \leq \pm 2$ is from the possible values of $\Phi = \cos\varphi_k - \cos\varphi_j$. For the k-th UT, the range of Φ is actually bounded by $\cos\varphi_j$. Note the fact that $\cos\varphi_j = \cos\varphi_k - \frac{N_{\langle k,j\rangle}\lambda}{dM} \leq \pm 1$ [†]. Therefore, to obtain the total number of nulls of $\mathcal{S}$, substitute the range of $\cos\varphi_j$ into (6.11), it gives the range of $N_{\langle k,j\rangle}$ with respect to the AoD of the base UT φ_k, which can

[†]The range of $\cos\varphi_j$ depends on the distribution range of φ of all active UTs, therefore can be different with respect to the transmission scenarios. Here, $\varphi \leq \pm 180^\circ$ is considered.

be denoted as

$$(\cos\varphi_k - 1)\frac{d}{\lambda}M \leq N_{\langle k,j\rangle} \leq (\cos\varphi_k + 1)\frac{d}{\lambda}M, \tag{6.13}$$

where $N_{\langle k,j\rangle} \in \mathfrak{N}$ is an integer.

From (6.13), the number of possible values for $N_{\langle k,j\rangle}$ is the number of integers within the range $[-\frac{d}{\lambda}M, \frac{d}{\lambda}M]$. Also, from (6.12), $\cos\varphi_k - \cos\varphi_j \neq 0$ is equivalent to $N_{\langle k,j\rangle} \neq 0$. $\cos\varphi_k - \cos\varphi_j \neq \pm 2$ is equivalent to $N_{\langle k,j\rangle} \neq \pm 2dM/\lambda$, which only constrains $N_{\langle k,j\rangle}$ when $\cos\varphi_k = \pm 1$.

Therefore the maximum cardinality of the set $\mathfrak{N}$ that contains all possible values of $N_{\langle k,j\rangle}$ is $|\{\mathbb{Z} : [-\frac{d}{\lambda}M, \frac{d}{\lambda}M]\}| - 1$ when $\cos\varphi \neq \pm 1$ or $|\{\mathbb{Z} : [-\frac{d}{\lambda}M, \frac{d}{\lambda}M]\}| - 2$ when $\cos\varphi = \pm 1$. Including the k-th UT, the number of the elements in $\mathfrak{N}$ plus 1 is the maximum number of UTs P in this correlation-free UTs' group. Note that the cardinality of $\mathfrak{N}$ is restricted by d, λ, M and $\cos\varphi_k$, hence P is also determined by these parameters.

For the j-th and the i-th UTs that both satisfy (6.11), and correspond to different $N_{\langle k,j\rangle}$ and $N_{\langle k,i\rangle}$, i.e. different nulls of $\mathcal{S}$, using (6.12) gives

$$\cos\varphi_j - \cos\varphi_i = \cos\varphi_k - \frac{N_{\langle k,j\rangle}\lambda}{dM} - (\cos\varphi_k - \frac{N_{\langle k,i\rangle}\lambda}{dM}) = \frac{(N_{\langle k,i\rangle} - N_{\langle k,j\rangle})\lambda}{dM}, \tag{6.14}$$

which is equivalent to

$$(\cos\varphi_j - \cos\varphi_i)\frac{\pi d}{\lambda}M = \pi(N_{\langle k,i\rangle} - N_{\langle k,j\rangle}). \tag{6.15}$$

Since $N_{\langle k,j\rangle}$ and $N_{\langle k,i\rangle}$ are integers and $N_{\langle k,j\rangle} \neq N_{\langle k,i\rangle}$, $(N_{\langle k,i\rangle} - N_{\langle k,j\rangle})$ is a non-zero integer and $\neq \pm 2dM/\lambda$. Therefore, the LoS channels of the j-th and the i-th UTs are correlation-free.

The LoS channel matrix of the UTs' group can be denoted as $\mathbf{H}_{LoS} = [\mathbf{a}_k^T, \mathbf{a}_i^T, \ldots, \mathbf{a}_j^T]^T$ and hence $\mathbf{H}_{LoS}\mathbf{H}_{LoS}^H = M|\{\mathbb{Z} : [-\frac{d}{\lambda}M, \frac{d}{\lambda}M]\}|\mathbf{I}_{|\{\mathbb{Z}:[-\frac{d}{\lambda}M, \frac{d}{\lambda}M]\}|}$, where $\mathbf{I}_{|\{\mathbb{Z}:[-\frac{d}{\lambda}M, \frac{d}{\lambda}M]\}|}$ denotes $|\{\mathbb{Z} : [-\frac{d}{\lambda}M, \frac{d}{\lambda}M]\}|$-dimensional identity matrix. Therefore, the LoS channels of the UTs are orthogonal. □

For 2-D transmissions from a BS equipped with a ULA, assume the AoDs of the $\mathcal{K}$ active UTs are continuously distributed within $[-180°, 180°]$, for the k-th UT, a group of UTs with orthogonal LoS channels can be found following Theorem 4 with the knowledge of their spatial AoDs. Note that if there is more than one UT that correspond to the same element in $\mathfrak{N}$, only one can be counted at a time.

6.3.3 LoS cross-correlation for 3-D transmissions from a UPA

Consider 3-D LoS transmissions from a BS equipped with an $M_x \times M_y$-element UPA, from the LoS channel vector denoted in (2.10), the LoS channel cross-correlation between the k-th and j-th UTs is denoted as

$$R = \mathbf{a}_k \mathbf{a}_j^H = \sum_{m_x=1}^{M_x} \sum_{m_y=1}^{M_y} e^{-\mathbf{j}\frac{2\pi d}{\lambda}((m_x-1)\sin\theta_k\cos\varphi_k+(m_y-1)\sin\theta_k\sin\varphi_k)} e^{\mathbf{j}\frac{2\pi d}{\lambda}\left((m_x-1)\sin\theta_j\cos\varphi_j+(m_y-1)\sin\theta_j\sin\varphi_j\right)}. \tag{6.16}$$

Define

$$\Theta \triangleq \sin\theta_k\cos\varphi_k - \sin\theta_j\cos\varphi_j, \tag{6.17}$$

and

$$\Xi \triangleq \sin\theta_k\sin\varphi_k - \sin\theta_j\sin\varphi_j. \tag{6.18}$$

Note that with respect to m_x in cross-correlation R, the terms in (6.16) containing $m_y = 1,\ldots,M_y$ can be derived as

$$\begin{aligned}
&\sum_{m_y=1}^{M_y} e^{-\mathbf{j}\frac{2\pi d}{\lambda}\left((m_x-1)\left(\sin\theta_k\cos\varphi_k-\sin\theta_j\cos\varphi_j\right)+(m_y-1)\left(\sin\theta_k\sin\varphi_k-\sin\theta_j\sin\varphi_j\right)\right)} \\
&= e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m_x-1)\left(\sin\theta_k\cos\varphi_k-\sin\theta_j\cos\varphi_j\right)} \sum_{m_y=1}^{M_y} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m_y-1)\left(\sin\theta_k\sin\varphi_k-\sin\theta_j\sin\varphi_j\right)} \\
&= e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m_x-1)\Theta} \sum_{m_y=1}^{M_y} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m_y-1)\Xi} \\
&= e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m_x-1)\Theta} e^{-\mathbf{j}\frac{\pi d}{\lambda}(M_y-1)\Xi} \frac{\sin\left(\frac{\pi d}{\lambda}M_y\Xi\right)}{\sin\left(\frac{\pi d}{\lambda}\Xi\right)}.
\end{aligned} \tag{6.19}$$

Hence, by substituting (6.19) into (6.16), and applying the similar derivation to terms containing $m_x = 1,\ldots,M_x$, the LoS channel cross-correlation between the k-th UT and the j-th UT in 3-D LoS

transmissions from a BS equipped with a UPA is given by

$$\begin{aligned} R &= \mathbf{a}_k \mathbf{a}_j^H \\ &= e^{-\mathbf{j}\frac{\pi d}{\lambda}(M_y-1)\Xi} \frac{\sin\left(\frac{\pi d}{\lambda}M_y\Xi\right)}{\sin\left(\frac{\pi d}{\lambda}\Xi\right)} \sum_{m_x=1}^{M_x} e^{-\mathbf{j}\frac{2\pi d}{\lambda}(m_x-1)\Theta} \\ &= e^{-\mathbf{j}\frac{\pi d}{\lambda}((M_x-1)\Theta+(M_y-1)\Xi)} \frac{\sin\left(\frac{\pi d}{\lambda}M_x\Theta\right)}{\sin\left(\frac{\pi d}{\lambda}\Theta\right)} \frac{\sin\left(\frac{\pi d}{\lambda}M_y\Xi\right)}{\sin\left(\frac{\pi d}{\lambda}\Xi\right)} \\ &= e^{-\mathbf{j}\frac{\pi d}{\lambda}((M_x-1)\Theta+(M_y-1)\Xi)} \mathcal{F}\mathcal{J}. \end{aligned} \tag{6.20}$$

where

$$\mathcal{F} \triangleq \frac{\sin\left(\frac{\pi d}{\lambda}M_x\Theta\right)}{\sin\left(\frac{\pi d}{\lambda}\Theta\right)}, \tag{6.21}$$

$$\mathcal{J} \triangleq \frac{\sin\left(\frac{\pi d}{\lambda}M_y\Xi\right)}{\sin\left(\frac{\pi d}{\lambda}\Xi\right)}. \tag{6.22}$$

Therefore, the LoS channel cross-correlation between the k-th and j-th UTs becomes a function of Θ and Ξ.

6.3.4 LoS orthogonal users' group for 3-D transmissions from a UPA

From (6.20), the LoS channel cross-correlation for 3-D transmissions from a UPA is eliminated when either $\mathcal{F}$ or $\mathcal{J}$ equals zero, i.e. either Θ or Ξ follows the following constraints:

$$\sin(\frac{\pi d}{\lambda}M_x\Theta) = 0\,, \sin(\frac{\pi d}{\lambda}\Theta) \neq 0, \tag{6.23}$$

or

$$\sin(\frac{\pi d}{\lambda}M_y\Xi) = 0\,, \sin(\frac{\pi d}{\lambda}\Xi) \neq 0. \tag{6.24}$$

Note that (6.23) and (6.24) correspond to the elements of the UPA configuration along the x-axis and y-axis of the 3-D space as shown in Fig. 2.4b. Define $X_{\langle k,j\rangle}$ with respect to the k-th UT and the j-th UT as

$$X_{\langle k,j\rangle} \triangleq \frac{d}{\lambda}M_x(\sin\theta_k\cos\varphi_k - \sin\theta_j\cos\varphi_j), \tag{6.25}$$

and $Y_{\langle k,j\rangle}$ with respect to the k-th UT and the j-th UT as

$$Y_{\langle k,j\rangle} \triangleq \frac{d}{\lambda}M_y(\sin\theta_k\sin\varphi_k - \sin\theta_j\sin\varphi_j). \tag{6.26}$$

By separately analysing (6.23) and (6.24) using the similar method denoted in section 6.3.2, the following theorem is obtained:

Theorem 5. *In 3-D LoS transmissions from a BS equipped with a UPA, assume the active UTs are spatially, continuously distributed in 3-D space. For the k-th UT, a group of UTs including the k-th, j-th, ..., i-th UTs can be found with orthogonal LoS channels when* $X_{\langle k,j\rangle},\ldots,X_{\langle k,i\rangle}$ *defined in (6.25) each corresponds to a unique element in set* $\mathfrak{X}$ *that is given by*

$$\mathfrak{X}=\{\mathbb{Z}:\left((\sin\theta_k\cos\varphi_k-1)\frac{d}{\lambda}M_x,(\sin\theta_k\cos\varphi_k+1)\frac{d}{\lambda}M_x\right)\}-\{0\}-\{\pm 2dM_x/\lambda\}. \tag{6.27}$$

The maximum number of UTs P in this group (including the k-th UT) is the maximum number of the elements in $\mathfrak{X}$ *plus 1, which is* $|\{\mathbb{Z}:[-\frac{d}{\lambda}M_x,\frac{d}{\lambda}M_x]\}|$ *when* $\sin\theta_k\cos\varphi_k\neq\pm1$ *or* $|\{\mathbb{Z}:[-\frac{d}{\lambda}M_x,\frac{d}{\lambda}M_x]\}|-1$ *when* $\sin\theta_k\cos\varphi_k=\pm1$.

Meanwhile, for the k-th UT, the LoS channels of another group of UTs: k-th, a-th, ..., b-th, are orthogonal when $Y_{\langle k,a\rangle},\ldots,Y_{\langle k,b\rangle}$ *defined in (6.26) each corresponds to a unique element in set* $\mathfrak{Y}$ *that is given by*

$$\mathfrak{Y}=\{\mathbb{Z}:\left((\sin\theta_k\sin\varphi_k-1)\frac{d}{\lambda}M_y,(\sin\theta_k\sin\varphi_k+1)\frac{d}{\lambda}M_y\right)\}-\{0\}-\{\pm 2dM_y/\lambda\}. \tag{6.28}$$

The maximum number of UTs Q in this group (including the k-th UT), i.e. the maximum number of the elements in $\mathfrak{Y}$ *plus 1, is* $|\{\mathbb{Z}:[-\frac{d}{\lambda}M_y,\frac{d}{\lambda}M_y]\}|$ *when* $\sin\theta_k\sin\varphi_k\neq\pm1$ *or* $|\{\mathbb{Z}:[-\frac{d}{\lambda}M_y,\frac{d}{\lambda}M_y]\}|-1$ *when* $\sin\theta_k\sin\varphi_k=\pm1$.

Proof. Similar to Theorem 4. Note that the LoS channel cross-correlation among the UTs found from $\mathfrak{X}$ and the UTs found from $\mathfrak{Y}$ are not eliminated, for example, the cross-correlation between the j-th UT and the a-th UT is not eliminated. □

For 3-D transmissions from a UPA with continuously distributed active UTs, from Theorem 5, for the k-th UT, two groups of UTs can be found with orthogonal LoS channels with respect to the UPA configuration of M_x and M_y. The number of UTs in each group is constrained by d,λ,M_x or M_y and AoDs of the base UT. However, the channels between the UTs from the two groups are not orthogonal. From Theorem 5, the advantage of 3-D LoS transmissions from a UPA is revealed which allows more degrees of freedom for the searching for the correlation-free UTs from either $\mathfrak{X}$ or $\mathfrak{Y}$.

6.4 Spatial user scheduling schemes

The objective is to design a user scheduling scheme for sub-6 GHz TDD-mode massive MIMO LoS transmissions to cooperate with the precoding scheme and improve the system performance of the lower bound on the sum-rate capacity. In this section, two spatial user scheduling (SUS) schemes are proposed respectively for 2-D LoS transmissions from a ULA and 3-D LoS transmissions from a UPA. The proposed SUS schemes enable the scheduling of UTs with mitigated LoS cross-correlation with respect to the k-th UT with the presence of the knowledge of UTs' spatial AoDs.

6.4.1 SUS scheme for 2-D LoS transmissions from a ULA

Consider a sub-6 GHz TDD-mode massive MIMO system. The BS is equipped with an M-element ULA, and there are $\mathcal{K}$ active UTs to be served with $\mathcal{K} \gg M$. Assume that, the channels between the UTs and the BS are LoS, and $\mathcal{T}$ instances of user scheduling can be performed during a coherence interval. For each instance of user scheduling at t, the BS selects a group of UTs for LoS transmissions based on the scheduling scheme.

From the arguments in Chapter 3, for LoS transmissions with MF and ZF precoding, the LoS channel mean-square cross-correlation restricts the lower bound on the system sum-rate capacity. Therefore, scheduling the UTs with low LoS channel cross-correlation enables the improvement in system performance.

The LoS orthogonal users' group arguments in section 6.3.2 and section 6.3.4 illustrate that, with respect to the k-th UT, the BS can find a group of UTs with orthogonal LoS channels, with the knowledge of their spatial AoDs, if the $\mathcal{K}$ active UTs are spatially continuously distributed. However, in realistic systems, the UTs have discrete spatial distribution, and the correlation-free constraints on the AoDs may not be satisfied by any active UT. Therefore, the search for the UTs with minimum LoS cross-correlation against the k-th UT can be implemented for the performance improvement. With Theorem 4, the search for the group of UTs with minimum LoS cross-correlation against the k-th UT can be implemented by a sequential search for the UT that satisfies the following condition corresponding to different elements in set $\mathfrak{N}$:

$$\cos\varphi_j = \underset{\forall j=1,\ldots,\mathcal{K}}{\operatorname{argmin}} \{|R|^2\} = \underset{\forall j=1,\ldots,\mathcal{K}}{\operatorname{argmin}} \{|\mathbf{a}_k \mathbf{a}_j^H|^2\} = \operatorname{argmin}\{|\mathcal{S}|^2\} \tag{6.29}$$

$$= \underset{\forall j=1,\ldots,\mathcal{K}, p\in\{1,\ldots,P-1\}, N_p\in\mathfrak{N}}{\operatorname{argmin}} |\frac{d}{\lambda} M(\cos\varphi_k - \cos\varphi_j) - N_p|, \tag{6.30}$$

where $\mathcal{S}$ is defined in (6.5), P is the number of UTs in this group with respect to the k-th UT, i.e. the number of elements in set $\mathfrak{N}$ plus 1, and N_p denotes the p-th element in $\mathfrak{N}$ as shown in (6.10).

Therefore, with the knowledge of AoDs for all $\mathcal{K}$ active UTs at the BS, for 2-D LoS transmissions from a ULA, a SUS scheme is obtained and summarised as follows: at the t-th scheduling time instance, the BS chooses the k-th UT as the base UT and adds this UT to the scheduling group $\mathcal{G}^t$. The BS calculates the set $\mathfrak{N}$ from (6.10). For each element $N_p \in \mathfrak{N}$ for $p = \{1,\ldots,P-1\}$, the BS finds the UT from the rest of the active UTs ($\mathcal{K}$ UTs exclude the base UT) that satisfies (6.30). Add the UTs to the scheduled group and merge the same UT added for different N_p if any.

Note that the selection of the base UT can be random. However, the fairness of user's accumulative capacity can be taken into account when selecting the base UT at t. To this end, let $C_j^{a,t} = \sum_{t=0}^{t} C_j^t$ to be the accumulate capacity for the j-th UT at the t-th scheduling time instance with initial $C_j^{a,t=0} = C_j^{t=0} = 0$, for $\forall j = \{1,\ldots,\mathcal{K}\}$, where the capacity for the j-th UT at t is denoted as

$$C_j^t = log2(1 + \text{SINR}_j^t). \tag{6.31}$$

where SINR_j^t denotes the SINR for the j-th UT at scheduling instance t that can be calculated by (2.15) for MF or ZF precoding. The BS updates the accumulated capacity for each active UT every time after scheduling. Apparently, if the UT is not scheduled at the t-th instance, $C_j^t = 0$. Hence, the UT with the lowest accumulated capacity at t can be chosen as the base UT.

6.4.2 SUS scheme for 3-D transmissions from a UPA

For 3-D transmissions from a UPA, the LoS cross-correlation can be mitigated with respect to either the x-axis or the y-axis configurations of the UPA. Corresponding to each axis of the array configuration, a group of UTs can be found using the scheme similar to the proposed SUS scheme for 2-D transmissions. Since the cross-correlation between UTs from each group is not mitigated, the BS must choose only one group to schedule.

With Theorem 5, the search for the group of UTs with minimum LoS cross-correlation against the k-th UT can be implemented by a sequential search for the UT that satisfies the following

condition corresponding to different elements in set $\mathfrak{X}$:

$$(\sin\theta_j, \cos\varphi_j) = \underset{\forall j=1,\ldots,\mathcal{K}}{\operatorname{argmin}} \{|R|^2\} = \underset{\forall j=1,\ldots,\mathcal{K}}{\operatorname{argmin}} \{|\mathbf{a}_k \mathbf{a}_j^H|^2\} = \underset{\forall j=1,\ldots,\mathcal{K}}{\operatorname{argmin}} \{|\mathcal{F}|^2\} \tag{6.32}$$

$$= \underset{\forall j=1,\ldots,\mathcal{K}, p\in\{1,\ldots,P-1\}, X_p\in\mathfrak{X}}{\operatorname{argmin}} |\frac{d}{\lambda} M_x(\sin\theta_k\cos\varphi_k - \sin\theta_j\cos\varphi_j) - X_p| \tag{6.33}$$

where $\mathcal{F}$ is denoted in (6.21), P denotes the number of UTs in this group, i.e. the number of elements in $\mathfrak{X}$ plus 1, and X_p denotes the p-th element in $\mathfrak{X}$.

Similarly, the search for the group of UTs with minimum LoS cross-correlation against the k-th UT can be executed following:

$$(\sin\theta_j, \sin\varphi_j) = \underset{\forall j=1,\ldots,\mathcal{K}}{\operatorname{argmin}} \{|R|^2\} = \underset{\forall j=1,\ldots,\mathcal{K}}{\operatorname{argmin}} \{|\mathbf{a}_k \mathbf{a}_j^H|^2\} = \underset{\forall j=1,\ldots,\mathcal{K}}{\operatorname{argmin}} \{|\mathcal{J}|^2\} \tag{6.34}$$

$$= \underset{\forall j=1,\ldots,\mathcal{K}, q\in\{1,\ldots,Q-1\}, Y_q\in\mathfrak{Y}}{\operatorname{argmin}} |\frac{d}{\lambda} M_y(\sin\theta_k\sin\varphi_k - \sin\theta_j\sin\varphi_j) - Y_q| \tag{6.35}$$

where $\mathcal{J}$ is denoted in (6.22), Q denotes the number of UTs in this group, i.e. the number of elements in $\mathfrak{Y}$ plus 1, and Y_q denotes the q-th element in $\mathfrak{Y}$.

Based on the proposed scheduling scheme for 2-D transmissions from a ULA, the spatial user scheduling scheme for 3-D transmissions from a UPA can be obtained and summarised as follows: at the t-th scheduling time instance, the BS chooses the k-th UT as the base UT based on the accumulative capacity fairness consideration and adds this UT to scheduling groups $\mathcal{G}_x^t$ and $\mathcal{G}_y^t$. The BS calculates the set $\mathfrak{X}$ and $\mathfrak{Y}$ from (6.27) and (6.28). Assume the number of elements in $\mathfrak{X}$ is $P-1$ and in $\mathfrak{Y}$ is $Q-1$. For each element $X_p \in \mathfrak{X}$ and $Y_q \in \mathfrak{Y}$ for $p = 1,\ldots,P-1$, $q = 1,\ldots,Q-1$, the BS searches for the UTs from the active UTs which satisfy (6.33) and (6.35), and adds the UTs to the corresponding group of $\mathcal{G}_x^t$ and $\mathcal{G}_y^t$. Merge the same UT added in each group if any. Compare the numbers of UTs in $\mathcal{G}_x^t$ and $\mathcal{G}_y^t$. The final decision of the scheduled group $\mathcal{G}^t$ is the group with more UTs.

6.4.3 Capacity enhancement check algorithm

Note that the proposed SUS schemes schedule the UTs with minimum LoS cross-correlation against the base UT. Therefore, the adding of each UT into the group may increase LoS channel cross-correlation of the UTs that are already in the group, i.e. may be harmful to their user capacities. When the total deduction of other UTs' user capacities becomes greater than the user capacity of

the new added UT, the enrolment of this UT in the group is harmful to the system sum-rate capacity at the t-th scheduling time instance. To tackle this issue, the capacity-enhancement check (CEC) can be implemented at every scheduling time instance after the group of UTs $\mathcal{G}^t$ is found by the proposed SUS schemes. To be specific, the system sum-rate C^t for a UTs group $\mathcal{G}^t$ obtained by a SUS scheme can be calculated by the summation of the UTs capacities denoted in (6.31) for all the grouped UTs. The BS temporarily removes the i-th UT that has the lowest user capacity from the group, followed by a recalculation of the system sum-rate $C^{t\prime}$. If $C^{t\prime} \geq C^t$, this UT is harmful to the system sum-rate. Hence, the BS removes the i-th UT from the scheduled group. The BS repeats such a process until the removal of the UT with the lowest user capacity no longer increases the system sum-rate, i.e. $C^{t\prime} < C^t$. The remaining UTs in $\mathcal{G}^t$ are the final decision of the scheduled group at the t-th scheduling time instance. The outlines of the SUS schemes for 2-D LoS transmissions from a ULA and 3-D LoS transmissions from a UPA in cooperation with the sum-rate enhancement check scheme are given in Algorithm 1 and Algorithm 2.

Algorithm 1 Spatial user scheduling algorithm for 2-D LoS transmissions from a ULA at the BS in cooperation with the sum-rate enhancement check algorithm

Initialisation AoDs of all $\mathcal{K}$ UTs, $C_k^{a,t=0} = 0$ for $k = 1, \ldots, \mathcal{K}$;

for $t = 1 : \mathcal{T}$ **do**
  if $t = 0$ **then**
    Randomly select a UT to be the base UT;
  elseSelect the UT with $\min C_k^{a,t}$ to be the base UT
  end if
  Add the base UT to the scheduling UT group $\mathcal{G}^t$;
  Compute set $\mathfrak{N}$ which contains $P-1$ elements, the p-th element N_p, for the base UT from (6.10);
  for $p = 1 : P-1$ **do**
    Find the j-th UT which from the active users group that satisfies $|\frac{d}{\lambda} M(\cos\varphi_k - \cos\varphi_j) - N_p|_{min}$, with $j = 1, \ldots, \mathcal{K}$, $j \neq k$;
    Add the j-th UT to the scheduled user group $\mathcal{G}^t$;
  end for
  if Repeated UTs in the group **then**
    Merge the repeated choice of UTs in $\mathcal{G}^t$;
  end if
  Compute the system sum-rate C^t for group $\mathcal{G}^t$;
  if Sum-rate enhancement check **then**
    $\mathcal{G}^{t\prime} = \mathcal{G}^t$;
    Find the i-th UTs with the lowest user capacity in $\mathcal{G}^{t\prime}$;
    Exclude this UTs from $\mathcal{G}^{t\prime}$;
    Calculate the system sum-rate $C^{t\prime}$ for $\mathcal{G}^{t\prime}$;
    if $C^{t\prime} > C^t$ **then**
      $\mathcal{G}^t = \mathcal{G}^{t\prime}$;
    elseBreak
    end if
  end if
  Update the actual number of UTs (K^t for SUS or K_c^t for SUS + CEC) in $\mathcal{G}^t$;
  Compute the C_k^t for $k = 1, \ldots, K^t$ or K_c^t;
  Update user accumulate capacity $C_k^{a,t} = C_k^{a,t-1} + C_k^t$;
  OutputScheduled user group $\mathcal{G}^t$ at t;
end for

Algorithm 2 Spatial user scheduling algorithm for 3-D LoS transmissions from a UPA at the BS in cooperation with the sum-rate enhancement check algorithm

Initialisation AoDs of all $\mathcal{K}$ UTs, $C_k^{a,t=0} = 0$ for $k = 1, \ldots, \mathcal{K}$;

for $t = 1 : \mathcal{T}$ **do**
  if $t = 0$ **then**
    Randomly select a UT to be the base UT;
  elseSelect the UT with $\min C_k^{a,t}$ to be the base UT;
  end if
  Add the base UT to the scheduling UT group $\mathcal{G}^t$;
  Compute set $\mathfrak{X}$, with $P-1 = |\mathfrak{X}|$, the p-th element X_p, for the base UT from (6.27);
  Compute set $\mathfrak{Y}$, with $Q-1 = |\mathfrak{Y}|$, with the q-th element Y_q, for the base UT from (6.28);
  for $p = 1 : P-1$ **do**
    Find the j-th UT which satisfies $|\frac{d}{\lambda} M_x(\sin\theta_k \cos\varphi_k - \sin\theta_j \cos\varphi_j) - X_p|_{min}$, with $j = 1, \ldots, \mathcal{K}$, $j \neq k$;
    Add the j-th UT to the scheduling user group $\mathcal{G}_x^t$;
  end for
  if Repeated UTs in $\mathcal{G}_x^t$ **then**
    Merge the repeated choice of UTs in $\mathcal{G}_x^t$;
  end if
  Update the actual number of UTs in $\mathcal{G}_x^t$;
  for $q = 1 : Q-1$ **do**
    Find the i-th UT which satisfies $|\frac{d}{\lambda} M_y(\sin\theta_k \sin\varphi_k - \sin\theta_i \sin\varphi_i) - Y_q|_{min}$, with $i = 1, \ldots, \mathcal{K}$, $j \neq k$;
    Add the i-th UT to the scheduling user group $\mathcal{G}_y^t$;
  end for
  if repeated UTs in the group **then**
    Merge the repeated choice of UTs in $\mathcal{G}_y^t$;
  end if
  Update the actual number of UTs in $\mathcal{G}_y^t$;
  Compare the number of UTs in $\mathcal{G}_x^t$ and $\mathcal{G}_y^t$. The group with more UTs is the scheduling user group $\mathcal{G}^t$;
  if Sum-rate enhancement check **then**
    $\mathcal{G}^{t\prime} = \mathcal{G}^t$
    Find the UT with the lowest user capacity in $\mathcal{G}^{t\prime}$;
    Exclude this UTs from $\mathcal{G}^{t\prime}$;
    Calculate the system sum-rate $C^{t\prime}$ for $\mathcal{G}^{t\prime}$;
    if $C^{t\prime} > C^t$ **then**
      $\mathcal{G}^t = \mathcal{G}^{t\prime}$;
    elseBreak
    end if
  end if
  Update the actual number of UTs in $\mathcal{G}^t$ (K^t for SUS or K_c^t for SUS + CEC);
  Compute the C_k^t for $k = 1, \ldots, K^t$ or K_c^t;
  Update user accumulate capacity $C_k^{a,t} = C_k^{a,t-1} + C_k^t$;
  OutputScheduled user group $\mathcal{G}^t$ at t;
end for

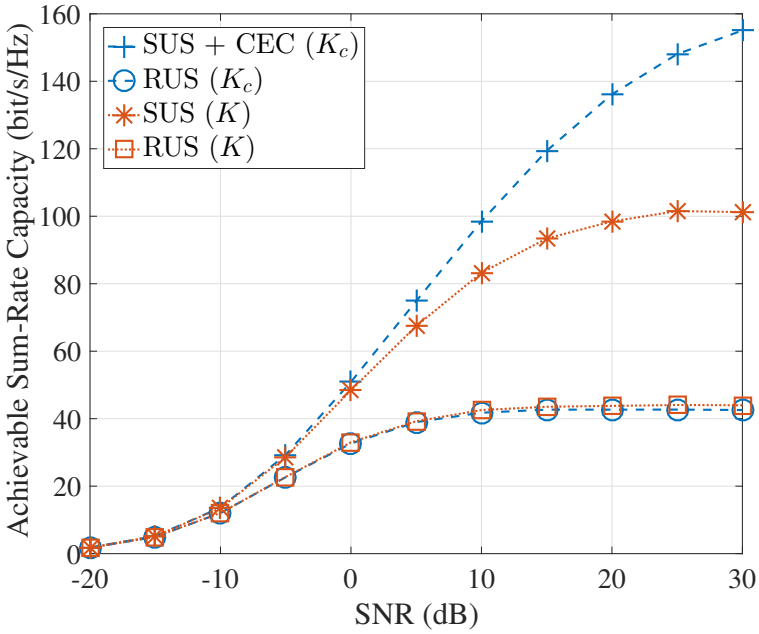
6.5 Spatial user scheduling simulation results

In this section, the system performance of the lower bound on the sum-rate capacity of the proposed SUS scheme for 3-D LoS transmissions from a UPA is evaluated by Monte Carlo simulations for sub-6 GHz massive MIMO systems since 3-D transmissions from a UPA are more realistic and provide more degrees of freedom for LoS transmissions compared with 2-D transmissions from a ULA. The effectiveness of the proposed CEC scheme on the performance enhancement for the proposed SUS is also evaluated. The random user scheduling (RUS) scheme that is based on the scheme adopted in LoS transmissions investigation in Chapter 3 is simulated for comparison. With RUS scheme, random UTs are selected by the BS for transmissions.

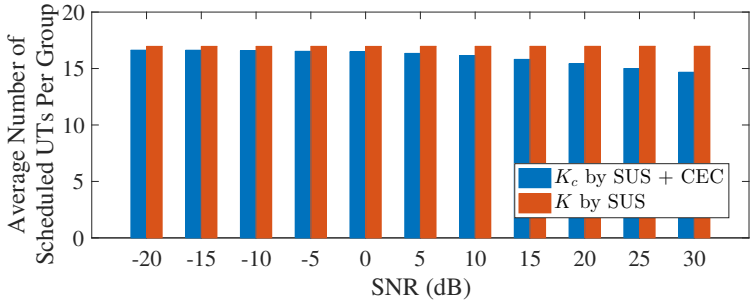
Consider that the BS is equipped with an $M_x \times M_y = M$-element UPA with carrier frequency of 2.6 GHz. The number of active UTs during a coherence interval is $\mathcal{K} = 2M$, and the AoDs of UTs are uniformly distributed with $\theta \sim$i.i.d.$\mathbf{U}(0^\circ, 90^\circ)$ and $\varphi \sim$i.i.d.$\mathbf{U}(-180^\circ, 180^\circ)$. Assume $\mathcal{T} = 200$ that is longer than the typical value of 196 adopted for 4G Long-Term Evolution (LTE) orthogonal frequency-division multiplexing (OFDM) [52] systems, as the LoS transmissions have a longer coherence time compared with fading channel transmissions. Therefore, 200 instances of scheduling can be implemented every time a new group of active UTs is generated. Assume the number of scheduled UTs per group at t by SUS is K^t and for CEC in cooperation with SUS is K_c^t. To compare the system performance, the RUS scheme is simulated considering the number of scheduled UTs per group equals K^t and K_c^t. MF is adopted for the precoding.

Fig. 6.2 illustrates the impact of SNR on the performance of the proposed schemes. $M = M_x \times M_y = 16 \times 8 = 128$ is considered. The system performance of the lower bound on the sum-rate is shown in Fig. 6.2a against the SNR. The average number of scheduled UTs per group is presented in Fig. 6.2b that is calculated by $K_c = \mathbb{E}[K_c^t]$ when CEC is adopted in cooperation with SUS, and $K = \mathbb{E}[K^t]$ when only SUS is implemented. The lower bound gain of the proposed schemes over the RUS scheme is shown in Fig. 6.2c.

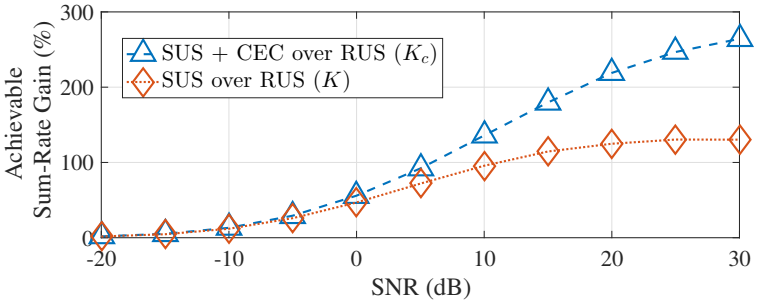
Fig. 6.2a depicts that the lower bounds on the sum-rate capacities for all of the scheduling schemes increase with the SNR to different limits due to the presence of the inter-user interference. The proposed SUS scheme outperforms the RUS scheme when serving the same number of UTs (K), thanks to the reduction of the LoS channel cross-correlation that reduces the MF precoding inter-user interference. With the help of the CEC scheme, SUS enables the best performance



(a) Lower bounds on the system sum-rate capacities vs SNR (dB) for the proposed schemes and the RUS scheme.



(b) Average number of scheduled UTs per group vs SNR (dB).



(c) Sum-rate capacities gains vs SNR (dB) for the proposed schemes over the RUS.

Fig. 6.2 Impact of SNR (dB) on the performance of the scheduling schemes. $M_x \times M_y = 16 \times 8 = 128$. The number of active UTs per coherence interval is $\mathcal{K} = 2M$. The perfect AoDs knowledge is adopted. $d = 0.5\lambda$, f =2.6 GHz. For CEC in cooperation with SUS, K_c UTs are scheduled per group. For SUS, K UTs are scheduled per group. MF precoding is adopted.

especially in the high SNR regime, due to the reduction of the inter-user interference by SUS with respect to the RUS scheduling and the exclusion of the sum-rate harmful UTs by CEC.

Note from Fig. 6.2b that the average number of scheduled UTs K by the SUS scheme is independent of SNR and is higher than the average number of scheduled UTs K_c by SUS when CEC is implemented. This confirms that there are sum-rate harmful UTs scheduled by the proposed SUS scheme. With the help of CEC, the average number of scheduled UTs per group by SUS decreases when SNR increases, while the lower bound of the sum-rate increases (see Fig. 6.2a). This indicates that for MF precoding, as SNR increases, the inter-user interference increases which causes the decrease of the system sum-rate. The proposed CEC scheme enables the SUS scheme to achieve a higher system performance while serving fewer UTs by excluding the sum-rate harmfull UTs.

Fig. 6.2c shows the system performance gain achieved by the proposed SUS scheme over the RUS scheme. It is evident that the performance gain of SUS over RUS increases with SNR to a limit. The proposed CEC scheme enables the SUS to achieve a higher gain over the RUS scheme when serving the same number of UTs per group, especially in the high SNR region. When SNR=10 dB, the CEC helps the SUS to achieve a 135% sum-rate enhancement over the RUS when $K_c \approx 16$ UTs are served per group, and the SUS itself enables a 95% sum-rate increment over the RUS when $K \approx 17$ UTs are served per group. Therefore, both SUS and CEC schemes are preferable in the high SNR region.

In Fig. 6.3, the system lower bound on the sum-rate (left y-axis) and the corresponding average number of scheduled UTs per group (right y-axis) are shown against the number of antennas M at the BS. $M_x \times M_y = 4 \times 4, 8 \times 4, 8 \times 8, 16 \times 8, 16 \times 16$, and 32×16 are simulated, considering SNR=10 dB. The figure illustrates that all of the scheduling schemes enable an improvement of the achievable rate when the BS scales up. However, for different M, with the help of CEC, the SUS scheme achieves a 130% of sum-rate gain over the RUS scheme when scheduling the same number of UTs per group. The SUS scheme itself allows an 87% gain. This illustrates that the number of antennas at the BS has little impact on the sum-rate performance gain for the proposed SUS scheme even when CEC scheme is jointly exploited. The disparity between K_c and K is larger when M is smaller. This is due to the constraint of $\mathcal{K} = 2M$. Under such constraint, when there are not many active UTs available as M is small, the UTs selected by the SUS scheme are more likely to be highly correlated that facilities the CEC scheme to perform the exclusion of the UTs.

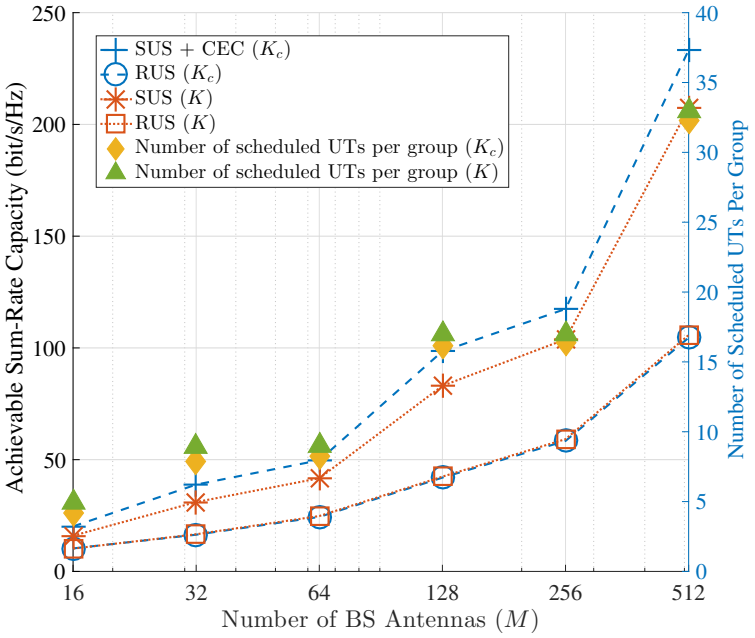


Fig. 6.3 Lower bounds on the system sum-rate capacities (left y-axis) and the average number of UTs per scheduled group (right y-axis) for the proposed schemes and the RUS scheme vs the number of antennas M at the BS, for $M = M_x \times M_y = 4 \times 4, 8 \times 4, 8 \times 8, 16 \times 8, 16 \times 16, 32 \times 16$. The number of active UTs per coherence interval is $\mathcal{K} = 2M$. SNR=10 dB is considered. Perfect AoDs knowledge is adopted. $d = 0.5\lambda$, f =2.6 GHz. For CEC in cooperation with SUS, K_c UTs are scheduled per group. For SUS, K UTs are scheduled per group. MF precoding is adopted. MF precoding is adopted.

Hence, the proposed CEC scheme helps the proposed SUS to contribute substantially to the system performance even when M is not so large.

6.6 Summary

In this chapter, spatial user scheduling schemes for LoS transmissions in sub-6 GHz TDD-mode massive MIMO systems are investigated. The LoS channel cross-correlation for 2-D transmissions from a ULA and 3-D transmissions from a UPA is investigated. By examining the correlation-free channel constraints, it is demonstrated that if the UTs are spatially continuously distributed within the space, with the knowledge of the AoDs, for each UT, a group of UTs can be found with orthogonal LoS channels within the group in 2-D LoS transmissions from a ULA, whereas two groups of UTs can be found in 3-D LoS transmissions from a UPA. The number of UTs in the group is determined by the separation between the adjacent antennas at the BS, carrier frequency, antenna array configuration, and AoDs of the base UT. Combined with the argument obtained from

Chapter 3 that the enhancement of the system performance for LoS transmissions can be achieved by reducing the LoS channel cross-correlation, two SUS schemes are proposed respectively for 2-D LoS transmissions from a ULA and 3-D transmissions from a UPA. With the proposed schemes, the BS schedules a group of UTs based on their AoD knowledge with mitigated LoS cross-correlation against the selected base UT. To further increase the system performance, a capacity-enhancement check scheme is proposed that excludes the UTs that are harmful to the system sum-rate and can be jointly implemented with the proposed SUS scheme. It is demonstrated through simulations for 3-D transmissions from a UPA that the proposed low-complexity SUS schemes achieve a substantial system performance enhancement compared with the RUS scheme. The cooperation of the proposed CEC scheme improves the performance of SUS scheme even further. No feedback is required by the proposed schemes.

The performance analysis also suggests that the SNR constrains the performance gain of the proposed SUS schemes over the RUS scheme. A higher SNR leads to a higher gain, especially when the proposed CEC scheme is jointly exploited. However, the number of antenna elements at the BS has less impact on the gain. Even when M and the number of active UTs are not so large, almost the same gain can be achieved by the proposed SUS schemes compared with the case when M and $\mathcal{K}$ are large. For example, when SNR=10 dB, a 130% of the system performance gain over the RUS scheme can be obtained by CEC in cooperation with the SUS. The SUS itself enables an 87% of the gain. The performance enhancement achieved by the proposed schemes is vital for massive MIMO LoS transmissions in both the LoS channel and Rician fading channel.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This research started with an aim to study and extend transmission capabilities of physical layer techniques for sub-6 GHz and mmWave massive MIMO systems that enable the use of massive MIMO to support the data demand for future communications networks. Throughout the thesis, comprehensive studies on massive MIMO techniques and mmWave transmissions have been achieved, and their key performance was evaluated along with the identification of performance limitations and the proposal of novel solutions.

Being the most compelling sub-6 GHz physical layer technique for 5G networks, massive MIMO achieves enormous spectral efficiency and energy efficiency by low-complexity linear processing. Meanwhile, the mmWave spectrum ranging from 28 GHz-300 GHz offers orders of magnitude throughput improvement compared with the sub-6 GHz spectrum and is also considered to be the essential essence of 5G networks. The thesis has focused on the performance of TDD-mode massive MIMO systems in sub-6 GHz LoS transmissions and mmWave transmissions. The investigations include the potential of being applied to different scenarios such as LoS transmissions for small-cell back-haul, satellite communications, and stadium transmissions, and mmWave massive MIMO transmissions for broadband high-quality, low-latency video, and multimedia applications.

The focus has been on the enhancement of the spectral efficiency for TDD-mode massive MIMO downlink transmissions in LoS or limited-scattering clustered transmission environments. A lower bound on the system sum-rate capacity is utilised for the system performance metric.

After the presentation of the overview of massive MIMO systems, representative sub-6 GHz channel models, and the conventional massive MIMO performance in the Rayleigh fading channel, investigations began with the study of the performance of massive MIMO in LoS transmissions where the statistical LoS CSI was adopted by the low-complexity linear precoding schemes of MF and ZF. Compared with the instantaneous CSI of fading channel, the statistical CSI of LoS channels is relatively easy to obtain and yields a longer coherence time for the system. It was shown that for a realistic system configuration of the deployment of a moderately large-scale antenna array at the BS, the system performance of LoS transmissions in massive MIMO systems is robust in the presence of NLoS interference. A key outcome from Chapter 3 is the expressions of the effective SINR for MF and ZF precoding for LoS 2-D transmissions from a ULA and 3-D transmissions from a UPA from which the key factors that impact on the system performance were revealed and evaluated through simulations. It is demonstrated that for both LoS and fading channel transmissions, the system performance increases with the scaling up of the BS array and the increase of SNR. For LoS transmissions, both a wider range of the AoDs distribution and higher power of the LoS channel component can enhance the system performance.

To achieve mmWave transmissions in TDD-mode massive MIMO systems, signal processing challenges for mmWave transmissions were first studied which indicate that the conventional massive MIMO digital baseband signal processing is infeasible with the scaling up of the system degrees of freedom and the increase of carrier frequency. Consequently, a low-cost and energy-efficient hardware architecture is imperative and can be achieved by separating the processing into a low-dimensional digital baseband control followed by a large-scale RF control. The mmWave massive MIMO multi-user precoding problem was then transformed into a mmWave beam-selecting problem.

Considering the limited-scattering clustered mmWave channel, note that in the TDD-mode, it is the channel CSI, which includes the contributions of all received paths that is obtained by the pilot estimation. A channel deconstruction scheme was proposed that enables the estimation of the parameters for each path with the perfect CSI or Gaussian-perturbed CSI exploiting the limited-scattering clustered property of mmWave channel and the asymptotic property of the massive MIMO channel. It was shown through analytical formulations and simulations that the proposed

channel deconstruction scheme achieves the mitigation of estimation MSEs with the scaling up of the BS array. Therefore, accurate estimation of the paths' parameters is obtained by the proposed channel deconstruction scheme for mmWave massive MIMO systems.

The low-complexity analogue and hybrid beam-selecting precoding schemes were proposed that exploit the estimated strongest path's parameters by the channel deconstruction. The hardware architectures were also designed, which deploy the low-cost and energy-efficient signal processing. System performance of the proposed precoding schemes was analysed using random matrix theory and order statistics principles and evaluated through simulations. It was shown that the selection of the estimated strongest path by the proposed analogue beam-selecting (ABS) and hybrid beam-selecting (HBS) precoders substantially contributes to the system performance. With only 10 RF chains adopted, the HBS and ABS precoders provide 86% and 74% of the capacity achieved by the MF precoding using the unconstrained digital baseband signal processing with 128 RF chains. It was also shown that in cooperation with the proposed channel deconstruction approach, the ABS and HBS precoders maintain the system performance over a wide range of the CSI imperfection factor value. Therefore, the accuracy requirement of the CSI estimation by the BS can be considerably relaxed. However, the performance of both precoders is primarily limited by the total number of received mmWave paths. The fewer paths the UTs receive, the better the performance the precoders achieve.

To obtain the optimal system performance in mmWave massive MIMO, the enhanced hybrid beam-selecting precoding (E-HBS) scheme was proposed. Exploiting the mmWave channel structure and the ability of channel deconstruction approach on the estimation of each path's parameters, the proposed E-HBS precoder enables multiple paths to be selected. More importantly, the hardware architecture of E-HBS can be formed by deploying multiple duplications of the HBS architecture in parallel as substructures. Such a configuration is vital with the scaling up of the system since the number of RF chains and baseband digital controls adopted are no longer related to the number of antennas at the BS but the number of paths selected instead and the number of UTs served. The system performance of E-HBS precoding was analysed and evaluated through simulations. It was shown that the optimal system performance is achieved by selecting all of the estimated paths. Moreover, the near-optimal performance is obtained by selecting most of the estimated strong paths. This indicates the potential of E-HBS for further reduction of the signal processing complexity, costs of fabrication and energy consumption in massive MIMO systems. It was also suggested that the AoDs distribution and SNR level should be jointly considered when

determining the number of paths to select for E-HBS precoding to fulfil the maximum utilisation of the system resources while achieving an objective system performance.

Finally, spatial user scheduling algorithms were proposed for sub-6 GHz LoS transmissions in massive MIMO systems. It has been shown that the cross-correlation of LoS channels limits the performance of LoS transmissions. By examining the correlation-free constraints for both 2-D transmissions from a ULA and 3-D transmissions from a UPA, the LoS orthogonal users' group can be found based on the knowledge of the statistical CSI of the UTs' spatial AoDs when AoDs are continuously distributed. For realistic LoS transmissions wherein the active UTs have discrete distribution, the spatial user scheduling schemes for each transmission scenario were then proposed to mitigate the LoS channel cross-correlation among the scheduled UTs using the AoD knowledge of all of the active UTs. The capacity-enhancement check scheme was proposed for further system performance improvement, whereby the UTs selected by the spatial user scheme that are harmful to the sum-rate capacity are excluded. The performance of the proposed schemes was evaluated through simulations for 3-D transmissions from a UPA and compared with the random user scheduling scheme. The results illustrated that the proposed schemes provide substantial performance gain with no feedback required. The performance gain of the proposed schemes is mainly related to the SNR level while the scale of the array at the BS has less impact. When SNR=10 dB, a 130% system performance gain can be achieved by the proposed spatial user scheduling scheme for 3-D transmissions from a UPA when operating with the proposed capacity-enhancement check scheme, and an 87% gain is obtained when only the spatial user scheduling scheme is adopted. These gains are vital for massive MIMO LoS transmissions in both the LoS channel and Rician fading channel.

Overall, novel precoding schemes and user scheduling schemes have been proposed to enhance the capabilities of massive MIMO systems in different transmission scenarios, such that 5G deployment can be accelerated. It has been demonstrated that the channel cross-correlation is the most significant limiting factor in the system performance of massive MIMO sub-6 GHz and mmWave transmissions. To improve the system performance of massive MIMO in sub-6 GHz LoS transmissions with a moderately large-scale system configuration, user scheduling is imperative. The optimal system performance implemented by digital baseband signal processing in mmWave massive MIMO transmissions can be achieved with low-complexity, low-cost, and energy-efficient precoding scheme, which separates the processing into analogue RF-domain control and digital baseband control.

7.2 Future work

- **Antenna array configuration**

  Although the comprehensive investigation for massive MIMO transmissions when the BS is equipped with a ULA or UPA has been accomplished, other practical antenna array configurations require a thorough investigation and performance comparison. Non-uniform and cylindrical array adoptions are possible for massive MIMO systems to provide higher BS compaction and radiation directivity [108]. The array response will be different in the cases of ULA and UPA. Therefore, the system performance for such BS configurations should be evaluated and compared.

- **The effectiveness of the proposed schemes for uplink transmissions**

  In this thesis, precoding schemes and spatial user scheduling schemes were analysed and evaluated for downlink massive MIMO transmissions. In uplink transmissions, UTs transmit signals to the BS, and the BS implements detection to the received signals. The received signal at each antenna element is a combination of the transmitted signals from all UTs. The effectiveness of the proposed precoding schemes in sub-6 GHz LoS and mmWave uplink transmissions must be evaluated, as they could be used for signal detection. The effectiveness of the spatial user scheduling in uplink transmissions also must be investigated.

- **The effectiveness of the proposed schemes in cellular networks**

  As introduced, ultra-dense networks are essential for 5G deployment. In multi-cell systems, transmission activities proceed simultaneously, except for the pilot assignment and uplink power control. Data transmissions in cellular systems are identical to the single-cell case. However, the inter-cell interference occurs from both the contaminating cells (due to pilot reuse) and non-contaminating cells. The performance of the proposed low-complexity non-cooperative schemes must be investigated considering cellular systems to evaluate their effectiveness against cell-to-cell cooperative schemes.

- **Hardware imperfections and implementations**

  Although massive MIMO transmissions in rich scattering environments are shown to be capable of averaging out hardware imperfections, such as non-linearity of the power amplifier, I/Q imbalance from mismatch between the in-phase and quadrature receiver components,

and quantisation noise from the analogue-to-digital converters, the impact of such hardware impairments on sub-6 GHz LoS and mmWave massive MIMO transmissions is unclear and must be investigated along with the modelling of the hardware impairments. Meanwhile, the implementation of directional antenna modulation techniques for the proposed precoding schemes is promising, whereby the energy efficiency of the hardware architecture can be further improved through the adoption of RF-domain analogue phase-only control. The investigations of such implementations are required.

- **User scheduling for mmWave transmissions**

  The proposed spatial user scheduling schemes provide an enormous system performance gain with respect to random user scheduling. Note that the array response vector is a common component in the LoS and mmWave channels. The author is particularly interested in the development of the cooperation of the proposed spatial user scheduling scheme and the proposed mmWave precoding schemes for further system performance enhancement.

References

[1] ITU-R M.2083-0, "IMT vision-framework and overall objectives of the future development of IMT for 2020 and beyond," 2015.

[2] Cisco, "Cisco visual networking index: Forecast and methodology, 2016–2021," Tech. Rep., 2018. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf

[3] IEEE, "IEEE 5G and beyond technology roadmap white paper," Tech. Rep., 2018. [Online]. Available: https://5g.ieee.org/roadmap

[4] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5g mobile and wireless communications: the vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26 – 35, 2014.

[5] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and A.-L. P. Popovski, "Five disruptive technology directions for 5g," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, 2014.

[6] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. D. Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201 – 1221, 2017.

[7] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: the next wireless revolution?" *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, 2014.

[8] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2522 – 2545, 2016.

[9] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[10] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[11] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2012.

[12] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[13] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[14] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, 2011.

[15] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. zar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Guiterrez, "Millimeter wave mobilecommunications for 5G cellular: it will work!" vol. 1, no. 1, pp. 335–349, May 2013.

[16] T. Bai, A. Alkhateeb, and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 70–77, Sep. 2014.

[17] Q. Zhao and J. Li, "Rain attenuation in millimeter wave ranges," in *IEEE Int. Symp. Antennas, Propag. EM Theory*, Guilin, China, Oct. 2007, pp. 1–4.

[18] J. N. Murdock, E. Ben-Dor, Y. Qiao, J. I. Tamir, and T. S. Rappaport, "A 38 GHz cellular outage study for an urban outdoor campus environment," in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, China, Apr. 2012.

[19] B. Biglarbegian, M. Fakharzadeh, D. Busuioc, M.-R. Nezhad-Ahmadi, and S. Safavi-Naeini, "Optimized microstrip antenna arrays for emerging millimeter-wave wireless applications," *IEEE Trans. Antennas Propag.*, vol. 59, no. 5, pp. 1742 – 1747, May 2011.

[20] S. Mumtaz, J. Rodriguez, and L. Dai, *MmWave massive MIMO: a paradigm for 5G*. London, U.K.: Elsevier science technology books, 2016.

[21] F.-L. Luo and C. Zhang, *Signal processing for 5G: algorithms and implementations*. West Sussex, U.K.: John Wiley & Sons, 2016.

[22] E. M. Report, "Future mobile data usage and traffic growth," Tech. Rep., 2018. [Online]. Available: https://www.ericsson.com/en/mobility-report/future-mobile-data-usage-and-traffic-growth

[23] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617 – 1655, 2016.

[24] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *40th Asilomar Conference on Signals, Systems and Computers ACSSC*, Pacific Grove, CA, USA, Nov. 2006.

[25] E. Bjornson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[26] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.

[27] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Lliya, "A comprehensive survey of pilot contamination in massive MIMO - 5G system," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 905–923, 2016.

[28] H. Yin, D. Gesbert, M. C. Filippou, and Y. Liu, "Decontaminating pilots in massive MIMO systems," in *IEEE Int. Conf. Commun. (ICC)*, Budapest, Hungary, Jun. 2013.

[29] K. Appaiah, A. Ashikhmin, and T. L. Marzetta, "Pilot contamination reduction in multi-user tdd systems," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 905–923, 2016.

[30] R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, 2014.

[31] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of massive MIMO*. Cambridge, U.K.: Cambridge university press, 2016.

[32] J. Zhang, B. Zhang, S. Chen, X. Mu, M. El-Hajjar, and L. Hanzo, "Pilot contamination elimination for large-scale multiple-antenna aided ofdm systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 759–772, Oct. 2014.

[33] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[34] K. Appaiah, A. Ashikhmin, and T. L. Marzetta, "Pilot contamination reduction in multi-user tdd systems," in *IEEE Int. Conf. Commun (ICC)*, Cape Town, South Africa, May 2010, pp. 1–5.

[35] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie, "Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1152–1163, 2014.

[36] "Wireless HD specification overview," Tech. Rep., 2010. [Online]. Available: http://www.wirelesshd.org/pdfs/WirelessHD-Specification-Overview-v1.1May2010.pdf

[37] IEEE, *IEEE Standard – Part 15.3: Wireless MAC and PHY Specifications for High Rate WPANs Amendment 2: Millimeter-wave-based Alternative Physical Layer Extension*, IEEE Std. 802.15.3c Std., 2009.

[38] *Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for High Rate Wireless Personal Area Networks (WPANs) Amendment 2: Millimeter-Wave-Based Alternative Physical Layer Extension*, IEEE Std. 802.15.3c Std., 2010.

[39] ISO/IEC/IEEE, *IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band (adoption of IEEE Std 802.11ad-2012)*, IEEE Std. 802.11ad Spec Std., Dec. 2010.

[40] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391 – 4403, Oct. 2013.

[41] T. S. Rappaport, R. W. H. Jr., R. C. Daniels, and J. Murdock, *Millimeter Wave Wireless Communications*. NJ, USA: Prentice-Hall.: Englewood Cliffs, 2014.

[42] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, p. 1499–1513, Mar. 2014.

[43] abriel M Rebeiz, S.-Y. Kim, O. Inac, W. Shin, O. Gurbuz, Y.-C. Ou, T. K. Fatih. Golcuk, and B.-H. Ku, "Millimeter-wave large-scale phased-arrays for 5G systems," in *IEEE MTT-S Int. Microw. Symp. (IMS)*, Barcelona, Spain, May 2015, pp. 1–3.

[44] S. Han, C.-L. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–195, Jan. 2015.

[45] A. Sayeed and J. Brady, "Beamspace MIMO for high-dimensional multiuser communication at millimeter-wave frequencies," in *Global Communications Conference (GLOBECOM), 2013 IEEE*, Atlanta, GA, USA, Dec. 2013.

[46] A. M. Sayeed and V. Raghavan, "Maximizing MIMO capacity in sparse multipath with reconfigurable antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 1, pp. 156–166, May 2007.

[47] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, p. 3814–3827, Jun. 2013.

[48] A. M. Sayeed and N. Behdad, "Continuous aperture phased MIMO: Basic theory and applications," in *Annu. Allerton Conf. Commun. Control Comput.*, Sep. 2010, p. 1196–1203.

[49] C. A. Balanis, *Antenna Theory: Analysis and Design, 4th Edition*. Wiley, 2015.

[50] H. Ji, Y. Kim, J. Lee, E. Onggosanusi, Y. Nam, J. Zhang, B. Lee, and B. Shim, "Overview of full-dimension MIMO in lte-advanced pro," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 176 – 184, Feb. 2017.

[51] Q. Nadeem, A. Kammoun, M. Debbah, and M.-S. Alouini, "Design of 5G full dimension massive MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 726 – 740, Feb. 2018.

[52] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[53] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, 2003.

[54] H. Yang and T. L. Marzetta, "Massive MIMO with max-min power control in line-of-sight propagation environment," vol. 65, no. 11, pp. 4685–4693, Nov. 2017.

[55] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 0733–8716, Feb. 2013.

[56] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[57] X. Mestre, "Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5113–5129, Nov. 2008.

[58] M. Dai, B. Clerckx, D. Gesbert, and G. Caire, "A rate splitting strategy for massive MIMO with imperfect CSIT," vol. 15, no. 7, pp. 4611–4624, Jul. 2016.

[59] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Acquiring partial CSI for spatially selective transmission by instantaneous channel norm feedback," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 4611–4624, Feb. 2008.

[60] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[61] A. Liu and V. Lau, "Phase only RF precoding for massive MIMO systems with limited RF chains," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4505–4515, Sep. 2014.

[62] A. Liu and V. K. N. Lau, "Impact of CSI knowledge on the codebook-based hybrid beamforming in massive MIMO," *IEEE Trans. Signal Process.*, vol. 64, no. 24, pp. 6545–6556, Dec. 2016.

[63] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint spatial division and multiplexing – the large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Aug. 2013.

[64] A. Adhikary, E. A. Safadi, M. K. Samimi, R. Wang, G. Caire, T. S. Rappaport, and A. F. Molisch, "Joint spatial division and multiplexing for mm-Wave channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1239–1255, Jun. 2014.

[65] S. Jin, W. Tan, M. Matthaiou, J. Wang, and K. Wong, "Statistical eigenmode transmission for the MU-MIMO downlink in Rician fading," vol. 14, no. 22, pp. 6650–6663, Dec. 2015.

[66] X. Li, S. Jin, and H. A. Suraweera, "Statistical 3-d beamforming for large-scale MIMO downlink systems over rician fading channels," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1529–1543, Feb. 2016.

[67] M. Coldrey, J.-E. Berg, L. Manholm, C. Larsson, and J. Hansryd, "Non-line-of-sight small cell backhauling using microwave technology," *IEEE Commun. Mag.*, vol. 51, no. 96, pp. 78–84, Sep. 2013.

[68] S. Chia, M. Gasparroni, and P. Brick, "The next challenge for cellular networks: backhaul," *IEEE Microw. Mag.*, vol. 10, no. 5, pp. 1527–3342, Aug. 2009.

[69] L. Liu, C. Tao, D. W. Matolak, Y. Lu, B. Ai, and H. Chen, "Stationarity investigation of a LOS massive MIMO channel in stadium scenarios," in *Vehicular Technology Conference (VTC Fall), 2015 IEEE 82nd*, Boston, MA, USA, Sep. 2015.

[70] B. R. Elbert, *The Satellite Communication Applications Handbook*. Artech house Inc, 2004.

[71] D. Roddy, *Satellite communications*. McGraw-Hill, 2006.

[72] A. Babakhani, D. B. Rutledge, and A. Hajimiri, "Near-field direct antenna modulation," *IEEE Microw. Mag.*, vol. 10, no. 1, pp. 36–46, Feb. 2009.

[73] A. B. Babakhani, D. B. Rutledge, and A. Hajimiri, "Transmitter architectures based on near-field direct antenna modulation," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 2674–2692, Dec. 2008.

[74] M. P. Daly, E. L. Daly, and J. T. Bernhard, "Demonstration of directional modulation using a phased array," *IEEE Trans. Antennas Propag.*, vol. 58, no. 5, pp. 1545–1550, May 2010.

[75] M. P. Daly and J. T. Bernhard, "Directional modulation technique for phased arrays," *IEEE Commun. Mag.*, vol. 57, no. 9, pp. 2633–2640, Sep. 2009.

[76] H. Shi and A. Tennant, "Characteristics of a two element direction dependent antenna array," in *Loughborough Conf. on Antennas and Propagation (LAPC)*, Loughborough, U.K, Jan. 2011, pp. 1–4.

[77] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of favorable propagation in massive MIMO," in *European Signal Processing Conf. (EUSIPCO)*, Lisbon, Portugal, Sep. 2014.

[78] M. Chiani, "On the probability that all eigenvalues of gaussian, wishart, and double wishart random matrices lie within an interval," *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4521 – 4531, Jul. 2017.

[79] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge, 2011.

[80] F. J. Lopez-Martinez, E. Martos-Naya, J. F. Paris, and A. Goldsmith, "Eigenvalue dynamics of a central wishart matrix with application to mimo systems," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2693 – 2707, May 2015.

[81] T. O'Farrell, "Code-division multiple-access (CDMA) techniques in optical fibre local area networks," Ph.D. dissertation, Univ. of Manchester, 1989.

[82] P. E. Omiyi and T. O'Farrell, "Throughput analysis of novel CDMA-based MAC protocol for wireless LANs."

[83] T. O'Farrell and M. Beale, "Code-division multiple-access (CDMA) techniques in optical fibre LANs," in *Telecommunications, 1989. Second IEE National Conference on*, York, UK, Apr. 1989.

[84] H. Donelan and T. O'Farrell, "Families of ternary sequences with aperiodic zero correlation zones for MC-DS-CDMA."

[85] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: system architecture modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.

[86] Y. Han, H. Zhang, S. Jin, X. Li, R. Yu, and Y. Zhang, "Investigation of transmission schemes for millimeter-wave massive MU-MIMO systems," *IEEE Syst. J.*, vol. 11, no. 1, pp. 72–83, Mar. 2017.

[87] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[88] J. Wang, Z. Lan, C.-W. Pyo, T. Baykas, C.-S. Sum, M. A. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.

[89] Y. M. Tsang, A. S. Y. Poon, and S. Addepalli, "Coding the beams: Improving beamforming training in mmwave communication system," in *IEEE Global Telecommunications Conference (GLOBECOM)*, Kathmandu, Nepal, Dec. 2011.

[90] K. Hosoya, N. Prasad, K. Ramachandran, N. Orihashi, S. K. S. Rangarajan, and K. Maruhashi, "Multiple sector ID capture (MIDC): A novel beamforming technique for 60-GHz band multi-Gbps WLAN/PAN systems," *IEEE Trans. Antennas Propag.*, vol. 63, no. 1, pp. 81–96, Jan. 2015.

[91] O. E. Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "The capacity optimality of beam steering in large millimeter wave MIMO systems," in *IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Cesme, Turkey, Jun. 2012.

[92] P. Liu, M. D. Renzo, and A. Springer, "IEEE 802.11ad: directional 60 GHz communication for multi-gigabit-per-second Wi-Fi," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 132–141, Dec. 2014.

[93] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.

[94] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, p. 831–846, Oct. 2014.

[95] J. Choi, "Beam selection in mm-Wave multiuser MIMO systems using compressive sensing," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2936–2947, Aug. 2015.

[96] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, p. 1164–1513, Jun. 2014.

[97] A. D. Dabbagh and D. J. Love, "Multiple antenna mmse based downlink precoding with quantized feedback or channel mismatch," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1859–1868, 2008.

[98] R. V. Hogg, J. W. McKean, and A. T. Craig, *Introduction to mathematical statistics, 7th edition*. Pearson, 2013.

[99] H. A. David and H. N. Nagaraja, *Order Statistics (3rd ed.)*. Wiley, 2003.

[100] N. Balakrishnan and A. P. Basu, *Exponential Distribution: Theory, Methods and Applications*. CRC Press, 1996.

[101] N. Balakrishnan and C. R. Rao, *Order Statistics: Theory and Method, Volume 16*. Elsevier Science, 1998.

[102] A. Forenza, D. J. Love, S. Abu-Surra, and R. W. Heath, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, p. 1924–1934, Jul. 2007.

[103] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Commun. Lett.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[104] A. M. Tulino, A. Lozano, S. Zhou, and S. Verdú, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.

[105] X. Yi and E. K. S. Au, "User scheduling for heterogeneous multiuser MIMO systems: A subspace viewpoint," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, p. 4004–4013, Oct. 2011.

[106] Z. Jiang, S. Chen, S. Zhou, and Z. Niu, "Joint user scheduling and beam selection optimization for beam-based massive MIMO downlinks," *IEEE Trans. Wireless Commun.*, vol. PP, no. 99, pp. 1–1, 2018.

[107] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 876–890, Oct. 2014.

[108] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang, and L. Hanzo, "Survey of large-scale mimo systems," *IEEE Commun. Surveys Tuts.*, vol. 17, pp. 1738 – 1760, Apr. 2015.