Low RF-Complexity Massive MIMO Systems: Antenna Selection and Hybrid Analog-Digital Beamforming

Von der Fakultät für Ingenieurwissenschaften,
Abteilung Elektrotechnik und Informationstechnik
der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation

von

Yuan Gao

aus

China

Gutachter:
Prof. Dr.-Ing. Thomas Kaiser
Univ. Prof. Dr.-Ing. Klaus Solbach

Tag der mundlichen Prüfung: 09.10.2017

Acknowledgments

I would like to express my gratitude to my supervisor Prof. Dr.-Ing. Thomas Kaiser who gave me the chance for the PhD study in the institute of Digital Signal Processing (DSV), the University of Duisburg-Essen. He gave me a large degree of freedom to select the research topic of my interest, and his invaluable insight on wireless communications lightens and leads me on my research. Without his support and encouragement, I am not able to finish this thesis.

I am also thankful to Univ. Prof. Dr.-Ing. Klaus Solbach, the University of Duisburg-Essen, for being my second supervisor, and in particular for the invaluable discussions on the microwave circuit modeling of Rotman lens.

I would like also to thank all my colleagues in DSV for the highly friendly atmosphere, which, I believe, is quite important for my continuous research activities. Special thanks goes to

- Theo Kreul who is my group leader and gave me a lot of supervision and support during the past five years. Whenever I have problems, he can always spend his precious time to give me support. Thanks to his rich experience, he can even foresee many potential problems and challenges, which are of great benefit to me on the research road.

- Dr. Feng Zheng who spent a lot of precious time to review and comment my publications, including this thesis. He is also my lunch partner. I enjoy our lunch time very much, especially many interesting conversations from research experience to living experience.

- Heinz Schreiber who supported me a lot when I was building the testbeds. He can always find everything that I need, e.g., experiment tables, cables, electronic components, etc. He also support me on some mechanical works, such as the antenna mounter for the antenna selection system, etc. His rich experience on the measurement equipments like vector network analyzer bring me a lot of convenience.

- Sabine Jankowski who is our secretary and helped me a lot on consultations, travel arrangements, etc.
- Maher Kaliel for the research collaboration. He supported me on simulating and fabricating Rotman lens.
- And all the other colleagues in DSV who gave me supports, and shared a good time with me during the past five years.

My last acknowledgments go to those who are close to me: my parents who raised me and always encourage me to face any problems; my close friend Macey whose smile is a strong power for me; my close classmates, especially Donghang and Zhimiao, in Aachen who share a good time with me since we knew each other.

Abstract

Wireless data traffic has been increased dramatically in the last decades, and will continue to increase in the future. As a consequence, the infrastructure of wireless communication systems needs to advance on the data capacity. Massive Multiple-Input Multiple-Output (MIMO) is a promising candidate technology to meet the demand. By scaling up the conventional MIMO by orders of magnitude number of *active* antennas, a massive MIMO system can harvest considerable channel degrees of freedom to increase the spectral efficiency. However, increasing the number of *active* antennas needs to increase both the numbers of Radio Frequency (RF) transceivers and antenna elements *at the same rate*, which will increase the RF complexity and cost dramatically. It is known that the complexity and cost of antenna elements are usually much lower than that of RF transceivers, which motivates us to scale up MIMO by a lower increasing rate of the number of RF transceivers than that of antenna elements, resulting in so-called low RF-complexity massive MIMO systems. In this thesis, we study two types of low RF-complexity massive MIMO systems, i.e., massive MIMO antenna selection systems and massive MIMO hybrid analog-digital beamforming systems. Both systems use specific RF networks to bridge a massive number of antennas and a small number of RF transceivers, leading to signal dimension reduction from antennas to RF transceivers. The RF network used in antenna selection is referred to as RF switching network; while the RF network used in hybrid beamforming is referred to as Phase Shifting Network (PSN). Both RF networks have two types of architectures, i.e., full-array architecture and sub-array architecture. The latter has lower insertion loss, lower complexity and better scalability than the former, but at the price of performance degradation caused by connection constraint, which will be studied for both low RF-complexity systems in this thesis. In addition, a low RF-complexity PSN for the hybrid analog-digital beamforming system needs also to be studied to replace the conventional high-complexity-and-cost phase-shifter-based PSN.

In the antenna selection system, the upper bounds on the channel capacity using asymptotic theory on order statistics are derived at the large-scale limit. The optimal antenna selection algorithms are also developed, which are based on Branch And Bound (BAB) search algorithm. Through the theoretical and algorithm studies, it is found that the sub-array antenna selection has close performance to the full-array antenna selection. In the hybrid beamforming system, we propose to use Rotman lens as PSN, which is of lower complexity and cost than the conventional phase-shifter-based PSN. Two beam selection algorithms, i.e., sub-optimal greedy search and optimal BAB search, are also proposed. In addition, the Rotman lenses are designed, fabricated and measured. The measurement results together with the beam selection algorithms are used to perform Monte Carlo simulation. Simulation results show that the proposed Rotman-lens-based system with the sub-array architecture suffers noticeable performance degradation compared to the system with the full-array architecture when ideal Rotman lenses are used. But when practical non-ideal Rotman lens are used, the former outperforms the latter when the number of antennas is large enough. Most interestingly, with non-ideal hardware, the sub-array Rotman-lens-based system has close performance to the sub-array phase-shifter-based system, and also exhibits a wideband capability. To prove the advantage of the low RF-complexity massive MIMO, two testbeds are built up for the antenna selection and hybrid beamforming systems, respectively. The measurement results show the low RF-complexity massive MIMO systems have superior performance over the small-scale MIMO systems under the condition of the same number of RF transceivers.

The results in this thesis show that the low RF-complexity massive MIMO systems proposed in this thesis are feasible in technology and promising in performance, validating its potential usage for the future 5G wireless communication systems.

Zusammenfassung

Der drahtlose Datenverkehr ist in den letzten Jahrzehnten dramatisch gestiegen und wird auch in Zukunft weiter zunehmen. Infolgedessen muss die Datenkapazität der drahtlosen Infrastruktur erhöht werden. Mehrantennen Systeme mit einer sehr großen Anzahl an Antennen (engl. Massive Multiple-Input Multiple-Output (MIMO)) sind vielversprechende Technologiekandidaten, um diese Nachfrage zu erfüllen. Durch die Hochskalierung der Antennenanzahl eines konventionellen MIMO um mehrere Größenordnungen kann ein Massive MIMO-System erhebliche Kanalfreiheitsgrade erlangen, um die spektrale Effizienz zu verbessern. Allerdings muss mit der Anzahl der *aktiven* Antennen sowohl die Anzahl der Hochfrequenz (engl. Radio Frequency (RF)) Transceiver als auch die der Antennenelemente *im gleichen Maße* vergrössert werden, was die RF-Komplexität und Kosten dramatisch erhöht. Dabei ist bekannt, dass die Komplexität und die Kosten von Antennenelementen in der Regel viel niedriger sind als die von RF-Transceivern. Dies führt uns dazu dass wir das MIMO-System um eine im Verhältnis zur Antennenzahl geringere Anzahl von RF-Transceivern erweitern wollen, den so genannten Massive MIMO-Systemen mit geringer RF-Komplexität. In dieser Arbeit untersuchen wir zwei Arten von Massive MIMO-Systemen mit geringer RF-Komplexität, nämlich Massive MIMO-Antennenselektionssysteme und Massive MIMO-Hybrid-Analog-Digital-Strahlformungssysteme. Beide Systeme verwenden spezielle RF-Netzwerke, um eine größere Anzahl von Antennen von einer kleineren Anzahl von RF-Transceivern zu versorgen, was zu einer Signalraumreduktion von den Antennen zu den RF-Transceivern führt. Das bei der Antennenselektions verwendete RF-Netzwerk wird als RF-Koppelfeld bezeichnet, während das RF-Netzwerk, das bei der Hybrid-Strahlformung verwendet wird, als Phasenverschiebungsnetzwerk (engl. Phase Shifting Network, PSN) bezeichnet wird. Beide RF-Netzwerke können als Voll-Array-Architektur oder als Sub-Array-Architektur realisiert werden. Letztere hat eine geringere Einfügedämpfung, eine geringere Komplexität und eine bessere Skalierbarkeit als die erstere, aber zum Preis der Leistungsverschlechterung, die durch eine eingeschränkung Anzahl von Antennen-Transceiver-Verbindungen verursacht wird. Die vorliegende Arbeit untersucht dies für beide Systeme mit niedriger RF-Komplexität. Darüber hinaus wird auch ein PSN mit niedriger RF-Komplexität für das Hybride-Analog-Digital- Strahlformungssystem untersucht, das das herkömmliche hochkomplexe und kostenintensive PSN ersetzen soll.

Im Antennenselektionssystem werden die Obergrenzen der Kanalkapazität unter Verwendung der Asymptoten Theorie der Ordnungsstatistik im Grenzverhalten abgeleitet. Die optimalen Antennenselektions-Algorithmen, die auf dem Branch and Bound (BAB) Suchalgorithmus basieren, werden ebenfalls entwickelt. Die theoretischen und algorithmischen Untersuchungen zeigen, dass die Leistung der Sub-Array-Antennenauswahl dicht bei der der Voll-Array-Antennenselektions liegt. Im Hybrid-Strahlformungssystem schlagen wir vor, eine Rotman-Linse als PSN zu verwenden, die von geringerer Komplexität und Kosten ist als das herkömmliche auf Phasenverschiebung basierende PSN. Es werden zwei Strahlauswahlalgorithmen vorgeschlagen, eine suboptimale Greedy-Suche und eine optimale BAB-Suche. Darüber hinaus wird die Rotman-Linse entworfen, gefertigt und vermessen. Die Messergebnisse werden zusammen mit den Strahlselektionsalgorithmen zur Durchführung einer Monte-Carlo-Simulation verwendet. Simulationsergebnisse zeigen, dass das vorgeschlagene Rotman-Linsen-basierte System mit der Sub-Array-Architektur eine spürbare Leistungsverschlechterung im Vergleich zum System mit der Full-Array-Architektur erleidet, wenn ideale Rotman-Linsen verwendet werden. Aber wenn reale nicht-ideale Rotman-Linsen verwendet werden, übertrifft erstere die zweite, wenn die Anzahl der Antennen groß genug ist. Noch interessanter, mit nicht-idealer Hardware, zeigt das Sub-Array Rotman-Linsen-basierte System in etwa die gleiche Leistung wie das Sub-Array Phasenschieber-basierte System und weist auch Breitbandfähigkeiten auf. Um den Vorteil der Massive MIMO-Systeme mit geringer RF-Komplexität zu beweisen, werden zwei Testumgebungen für die Antennenauswahl- und Hybrid-Strahlformungssysteme aufgebaut. Die Messergebnisse zeigen, dass, unter der Bedingung einer gleichen Anzahl von RF-Transceivern, die Massive MIMO-Systeme mit geringer RF-Komplexität in der Leistung den normalen MIMO-Systemen überlegen sind.

Die Ergebnisse meiner Arbeit zeigen, dass die von mir vorgeschlagenen Massive MIMO-Systeme mit geringer RF-Komplexität technisch machbar und vielversprechend in der Leistung sind und bestätigen damit deren potentielle Nutzung für die zukünftigen 5G-Funkkommunikationssysteme.

Contents

List of Figures

List of acronyms

AoA	Angle of Arrival
AWG	Arbitrary Wave Generator
AWGN	Additive White Gaussian Noise
BAB	Branch And Bound
BER	Bit Error Rate
BS	Base Station
CDF	Cumulative Distribution Function
CP	Cyclic Prefix
CSI	Channel State Information
CST	Computer Simulation Technology
DC	Direct Current
DFT	Discrete Fourier Transform
DSA	Digital Signal Analyzer
EM	Electro-Magnetic
FAS	Full-Array Switching
FFT	Fast Fourier Transform
5G	Fifth Generation
4G	Forth Generation
GP	Guard Period
GPIO	General Purpose Input Output
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
OFDM	Orthogonal Frequency-Division Multiplexing

IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
LMMSE	Linear Minimum Mean Square Error
LTE	Long Term Evolution
MIMO	Multiple-Input Multiple-Output
MPC	Multi-Path Component
MSV	Minimum Singular Value
ML	Maximum Likelihood
mmWave	millimeter Wave
MSE	Mean Square Error
PAPR	Peak-to-Average Power Ratio
PC	Personal Computer
PDF	Probability Density Function
PSN	Phase Shifting Network
QAM	Quadrature Amplitude Modulation
RE	Resource Element
RF	Radio Frequency
RLD	Rotman Lens Designer
SAS	Sub-Array Switching
SIC	Successive Interference Cancellation
SISO	Single-Input Single-Output
SP8T	Single-Pole 8-Throw
SNR	Signal to Noise Ratio
SPMT	Single-Pole Multi-Throw
SVD	Singular Value Decomposition
TDD	Time Division Duplex
TTD	True Time Delay
UE	User Equipment
UHF	Ultra High Frequency
ULA	Uniform Linear Array
USB	Universal Serial Bus
VNA	Vector Network Analyzer
WLAN	Wireless Local Area Network

Notation

x	Scalar
$\mathbf{x}$	Vector
$\mathbf{X}$	Matrix
$\mathbf{X}^{-1}$	Inverse of matrix $\mathbf{X}$
$\mathbf{X}^{*}$	Complex conjugate of $\mathbf{X}$
$\mathbf{X}^{T}$	Transpose matrix of $\mathbf{X}$
$\mathbf{X}^{H}$	Conjugate transpose (Hermitian) matrix of $\mathbf{X}$
$\mathsf{tr}\left(\mathbf{X}\right)$	Trace of matrix $\mathbf{X}$
$\mathsf{det}\left(\mathbf{X}\right)$	Determinant of matrix $\mathbf{X}$
$\|\|\mathbf{X}\|\|$	L-2 norm of matrix $\mathbf{X}$
$\|\|\mathbf{X}\|\|_{\mathrm{F}}$	Frobenius norm of matrix $\mathbf{X}$
$\mathbb{E}(\mathbf{x})$	Mathematical expectation of random variable $\mathbf{x}$
$\mathbb{C}$	Complex field
$\mathbb{R}$	Real field
$\mathcal{P}$	A set
$\|\mathcal{P}\|$	Cardinality of the set $\mathcal{P}$
$[\mathbf{X}]_{\mathcal{P},:}$	A submatrix of matrix $\mathbf{X}$ indexed by the row indices in $\mathcal{P}$
$[\mathbf{X}]_{:,\mathcal{P}}$	A submatrix of matrix $\mathbf{X}$ indexed by the column indices in $\mathcal{P}$
$[\mathbf{X}]_{\mathcal{P},\mathcal{Q}}$	A submatrix of matrix $\mathbf{X}$ indexed by the row indices in $\mathcal{P}$ and column indices in $\mathcal{Q}$
$\odot$	Elementwise (Hadamard) product
$\setminus$	Set difference
$\otimes$	Kronecker product
$\mathcal{O}(\cdot)$	The order of computation steps

$\lvert x\rvert$	The Magnitude of the complex number x
$\angle x$	The angle of the complex number x
$x \bmod 2$	x modulo 2
$\mathbf{I}_N$	An $N \times N$ identity matrix
$\mathcal{CN}(\mathbf{0}, \mathbf{R})$	Complex Gaussian distribution with zero mean and covariance matrix $\mathbf{R}$

Chapter 1

Introduction

In the past decade, we witnessed a bloom of mobile Internet driven by smart phones, Tablets, etc. Along with numerous mobile applications, the mobile data traffic has been increased dramatically. For example, from 2011 to 2016, the global mobile data traffic has increased 18-fold [1], from 400 petabytes per month in 2011 to 7.2 exabytes per month at the end of 2016. In the forthcoming decade, the mobile data traffic will increase further, e.g., 7-fold growth from 2016 to 2021 as predicted by [1]. This is not only driven by the higher data volume demand with more smart devices, but also by the huge number of things that are going to be connected to the Internet which is also referred to as Internet of Things (IoT) [2, 3]. As a result, the current mobile networks need evolution or even revolution [4] to enhance the mobile data capacity. Taking the cellular network as an example, the current Forth Generation (4G) Long Term Evolution (LTE) network is going to be evolved and revolutionized under the umbrella of International Telecommunication Union Fifth Generation (5G) framework vision [5], which aims at three usage scenarios, i.e., enhanced mobile broadband, ultra-reliable and low latency communications, and massive machine type communications. These three scenarios, especially the enhanced mobile broadband, indicate that the enhancement of the data capacity will go on playing an important role in the future 5G cellular network.

In general, there are three ways to to increase the mobile network capacity:

- **Exploiting unused or underutilized spectrum** The straightforward way to increase the capacity is to allocate more spectrum. For example, there is a growing interest to allocate the millimeter Wave (mmWave) band, i.e., 30 - 300 GHz, for the future 5G cellular systems [6]. In addition to allocating more spectrum, the allocated spectrum can also be shared with multiple different mobile networks to fully utilize the spectrum, which can be realized by the technology of cognitive radio [7]. Note that the spectrum, especially the favorable Ultra High Frequency (UHF) band, is so scarce and expensive that we cannot always expect more spectrum to increase the network capacity.

- **Densifying networks** The method of network densification, particularly spatial densification, means that the geographic area is divided into smaller and smaller cells [8], which apparently increases the frequency reuse and consequently increases the network capacity per unit area. For example, a limited number of macro-cells accompanying with a large number of micro-cells, small-cells, pico-cells and femto-cells can be deployed geometrically to construct a heterogeneous network to share the same frequency band among these cells. Note that the network densification will introduce huge capital and operational expenditure for cell-site renting, backhaul laying, etc.

- **Improving spectral efficiency** Contrast to the scarce spectrum resource and the huge cost of the network densification, the method of improving spectral efficiency is usually affordable and attractive. There are many ways to improve the spectral efficiency, such as better capacity-approaching channel coding [9, 10], more advanced modulation schemes [10, 11], multi-cell interference management like coordinated multi-point transmission and reception [12], equipping more antennas [19], etc. Among these methods, the method of equipping more antennas is particularly attractive, because the spectral efficiency can be increased by simply using more antennas [19]. In this thesis, we mainly focus on the wireless systems with many antennas, especially with a massive number of antennas.

1.1 Motivation

1.1.1 Why do we need massive number of antennas?

Multi-antenna technology, also known as Multiple-Input Multiple-Output (MIMO), has drawn great attention in both academia and industry in the last three decades. Due to its advantages on the spectral efficiency and link reliability [19, 20], MIMO technology has been successfully used by many existing practical wireless communication systems, such as IEEE 802.11 Wireless Local Area Network (WLAN), 4G LTE-Advanced, etc. It should be noted that, in the conventional MIMO, the number of antennas is usually not very large. Taking the 4G LTE-Advanced system as an example, only up to 8×8 MIMO can be supported in the downlink data transmission [13].

In recent years, a great interest in MIMO technology is to scale up the conventional MIMO by orders of magnitude to massive MIMO [15, 16, 18]. In general, we can benefit a lot from scaling up MIMO, for example,

- **Higher spectral efficiency** Given a MIMO system with N_T transmit antennas and N_R receive antennas. Suppose a fast and frequency-flat fading MIMO channel and the receiver knows the perfect channel state knowledge. Then the spectral efficiency (measured by ergodic capacity) of this MIMO channel can be approximated by [14]

  $$C = \min(N_\mathrm{T}, N_\mathrm{R})c \tag{1.1}$$

  where the scaling factor c depends on the ratio of N_T and N_R and on the Signal to Noise Ratio (SNR) [14, 20]. Equ. (1.1) shows that the MIMO spectral efficiency increases *linearly* with $\min(N_\mathrm{T}, N_\mathrm{R})$ if we increase both N_T and N_R [20]. It should be noted that even just increasing N_R with fixed N_T and SNR, we can still benefit with the *logarithmic* growth in spectral efficiency [14].

- **More reliability** As aforementioned, one of the usage scenarios in future 5G is the ultra-reliable communications, such as in the car2X communication systems. MIMO technology is known to be able to enhance the link reliability via a so-called spatial diversity. It is known that wireless channels suffer from fading effect caused by multipath and movement. Deep fading potentially makes the communication links outage. The idea of spatial diversity is that, with more antennas the probability of all of the antenna channels suffering deep fading will be smaller (assuming independent fading). It is known that in MIMO systems, the link outage probability is proportional to $\rho^{-N_\mathrm{T}N_\mathrm{R}}$ [16] (ρ denotes SNR), which means that the link outage probability will be extremely small when the number of antennas is large.

- **Higher energy efficiency** When we increase the number of antenna in the BS, in principle, narrower beams can be formed since the beamwidth is inverse proportional to the array aperture size. The benefit of narrow beams is to reduce the unwanted energy waste, i.e., most of the useful radiated (or received) to (or from) the intended user, which can obviously increase the energy efficiency. It was reported in [21] that when BS have full *perfect* channel knowledge, the radiation power by users can be made inversely proportional to the number of base station antennas with no degradation performance. This means that with doubled number of BS antennas, the user transmit power can be halved without performance loss. Even with *estimated* channel knowledge, we can still achieve a law of square-root of the inverse of the number BS antennas [21].

- **Higher array gain** Like large antenna array in Radar systems, massive number of antennas in wireless communications can also harvest a high array gain, which is particularly necessary for the future wireless communications with frequency bands above UHF, such as mmWave. This is because when the carrier frequency goes higher, the larger path loss will be suffered with the smaller antenna aperture size under the higher carrier frequency. The large number of antennas constitutes an array with a large effective aperture size, which can compensate the path loss with effective beamforming [6].
- **Lower inter-cell interference** One of the impressive advantages of massive MIMO over the conventional MIMO lies in the reduction of inter-cell interference in multi-cell systems. The data rate at the cell boundaries suffers severe degradation due to the inter-cell interference. Solutions to the inter-cell interference in the current cellular systems require cooperation between cells, such as the aforementioned CoMP [12]. The inter-cell cooperation will require dedicated backhaul among Base Stations (BSs), and also introduce processing delay. When the concept of massive MIMO first arose in [15], the premier intention is to use unlimited number of antennas in each BS to eliminate the inter-cell interference. The idea is that, as the number of antennas tends to infinity, the channel vectors of different users on the cell boundaries tend to be orthogonal according to the central limit theorem. In the perspective of signal processing, one BS can form extreme narrow beam pointing the intended user and meanwhile reduce the energy leakage to other users that belong to other cells.

By summarizing the above advantages of massive MIMO, we can conclude that massive MIMO not only increases the spectral efficiency, but also addresses many fundamental limitations in the current wireless communication systems. Thus massive MIMO, as a recent breakthrough technology, is believed to be a promising technology in the future 5G cellular systems.

1.1.2 Why do we need low RF-complexity massive MIMO?

In the preceding discussions we assumed that each antenna has dedicated RF transceiver. Thus scaling up MIMO means scaling up both the number of RF transceivers and antennas. But if we relax the assumption and allow different number of RF transceivers and antennas, the picture will be different: we can make the scaling up rate of the number of RF transceivers lower than that of the number of antennas. Consequently, the number of RF transceivers will be much less than that of antennas, and we refer to this kind of massive MIMO system as low RF-complexity massive MIMO system. The question is: does the relaxed scaling up mechanism make sense in the future massive MIMO systems? We try to answer this question from the following aspects:

- **Hardware cost** Although the cost of RF transceivers is getting lower due to the development of semiconductor in the last decades, the cost of RF transceivers is still significantly higher than antennas, not to mention the cost-prohibitive RF transceivers in the frequency bands above UHF band, such as mmWave. Scaling up the number of antennas is obviously more cost-efficient than scaling up the number RF transceivers. This situation will remain in the near future. Thus from the perspective of hardware cost, we can say that the low RF-complexity massive MIMO is quite desirable for the future low-cost communication.

- **Energy efficiency** The energy efficiency analysis of massive MIMO [21] mentioned in the preceding subsection implies that an unbounded energy efficiency can be achieved by increasing the number of BS antennas. Obviously, this is not practical. The analysis in [21] only considered the useful radiation power, but ignored the device electronic circuit power consumption of the RF transceivers, such as local oscillators, filters, etc. Actually, when the number of RF transceivers is large, the circuit power consumption will be significant. The fact is that, when taking the circuit power consumption into account, we cannot always benefit from increasing the number of RF transceivers [22]. In contrast, allowing less number of RF transceivers than that of antennas can improve the energy efficiency, as shown by the studies in [42], [80], where antenna selection and hybrid analog-digital beamforming are used, respectively, to connect the less number of RF transceivers to the large number of antennas. Thus from the perspective of energy efficiency, the low RF-complexity massive MIMO is meaningful for the future green communications.

- **Channel matrix rank** One of the favorite massive MIMO configuration is that only one side (transmit or receive side) has massive number of antennas while the other side only has a small number of antennas. For example, in cellular systems, BS can have a massive number of antennas to serve *only a small number* of users [15]. The limitation on the number of antennas at one side comes from channel estimation and design feasibility, because we usually have limited physical resource, e.g., coherence bandwidth and coherence time, to allocate orthogonal pilots for users. Such a MIMO configuration will make the rank of channel matrix be limited by $\min(N_{\mathrm{T}}, N_{\mathrm{R}})$. In addition, the channel sparsity in the frequency band above UHF, such as mmWave band, will make the rank of the channel matrix even lower [61]. Like feature extraction or selection in machine learning [23], in communication systems, the low-rank channel matrix allows us to reduce the dimension of massive-antenna signals in the analog domain, and consequently requires only a small number of RF transceivers on analog-to-digital or digital-to-analog conversion. Thus from the perspective of signal processing, we can say that the low RF-complexity massive MIMO is feasible in principle.

By summarizing the above discussions, we can conclude that the low RF-cost massive MIMO is necessary to be studied for the future wireless communications.

1.2 Outline and Contributions

This thesis studies the low RF-complexity massive MIMO systems. In particular, two kinds of low RF-complexity systems are considered, i.e., massive MIMO antenna selection systems and massive MIMO hybrid analog-digital beamforming systems. The outline and the main contributions of this thesis is summarized as follows:

- **Chapter 2: System Model**
  This chapter mainly describes the general system model that is used in this thesis. We first give a narrowband multi-user massive MIMO for an uplink of a single-cell cellular system. Afterwards, we generally described the concept of the low RF-complexity massive MIMO systems that will be studied in this thesis. Some basic definitions and concepts, e.g., full-digital massive MIMO systems, balanced and unbalanced-configured massive MIMO systems, low RF-complexity massive MIMO, full-array and sub-array architectures, etc., are also given.

- **Chapter 3: Massive MIMO Antenna Selection Systems**
  In this chapter, we study the antenna-selection-based low RF-complexity massive MIMO systems. We first examine the massive MIMO antenna selection over the full-array and sub-array architectures. Analytical upper bounds on the antenna selection capacity are derived for both the full-array and sub-array architectures. In the derivation, we use asymptotic theories on order statistic to give relatively simple analytical forms of the

massive MIMO antenna selection systems in the large-scale limit. Afterwards, Channel State Information (CSI) acquisition is taken into account, and the upper bounds on the ergodic capacity is consequently derived. In addition to the theoretical work, we also develop optimal antenna selection algorithms based on Branch-And-Bound (BAB) search algorithm for both the full-array and sub-array architectures. Furthermore, some numerical and simulation results are presented and show that the low RF-complexity antenna selection system with sub-array architectures has performance very close to that of the high-complex antenna selection system with full-array architectures, which is validated by their exact capacities and their upper capacity bounds. Through the study in this chapter we can find that the low RF-complexity massive MIMO systems based on sub-array antenna selection is very cost-effective for the practical application. The content of this chapter has been submitted to [24].

- **Chapter 4: Massive MIMO Hybrid Digital-Analog Beamforming Systems Based on Rotman Lens**
  In this chapter, we study the hybrid-analog-digital-beamforming-based low RF-complexity massive MIMO systems. In particular, we propose a Rotman-lens-based system, which, to our best knowledge, is the first one reported in the literature. The full-array and sub-array architectures are compared, concerning the RF design feasibility, insertion loss and system scalability. Then the Rotman-lens-based system is designed with both algorithms and hardware. In particular, we design two algorithms, i.e., sub-optimal greedy search and optimal BAB search, to find a good subset of beam ports that leads to a small mean square error. In the hardware design, we design, fabricate and measure a sample of Rotman lens. With the measured results, we study the Rotman-lens-based system with Monte Carlo simulations. Our simulation results show that our low RF-complexity hybrid beamforming system based on the low-cost Rotman lens has performance comparable to that of the system using the high-cost phase shifters, and exhibits wideband capability and superior performance over the small-scale MIMO system under the same number of RF transceivers. The content of this chapter has been published in [98, 99].

- **Chapter 5: Testbed Performance Evaluation**
  This chapter shows the experiment studies on the low RF-complexity massive MIMO systems. In particular, we set up two massive MIMO testbeds, i.e., antenna selection system and hybrid analog-digital beamforming system, with the sub-array architecture.

The latter is based on the former by inserting the fabricated Rotman lenses in-between the antennas and RF switches. The details on these two testbeds, from the hardware and signal structure to the involved signal processing, are described. Especially, the selection algorithms designed for the narrowband systems in Chapter 3 and 4 are tailored for the wideband systems in this chapter. The methodology of the testbed development is also given. Finally, the measurement results of both testbeds are demonstrated.

- **Chapter 6: Conclusion**

  This chapter summarizes and concludes the low RF-complexity massive MIMO systems studied in this thesis. In addition, some constructive guidelines and recommendations are provided for the future extension work.

The structure and main contributions of this thesis with the relationship among the chapters is shown in Fig. 1.1.

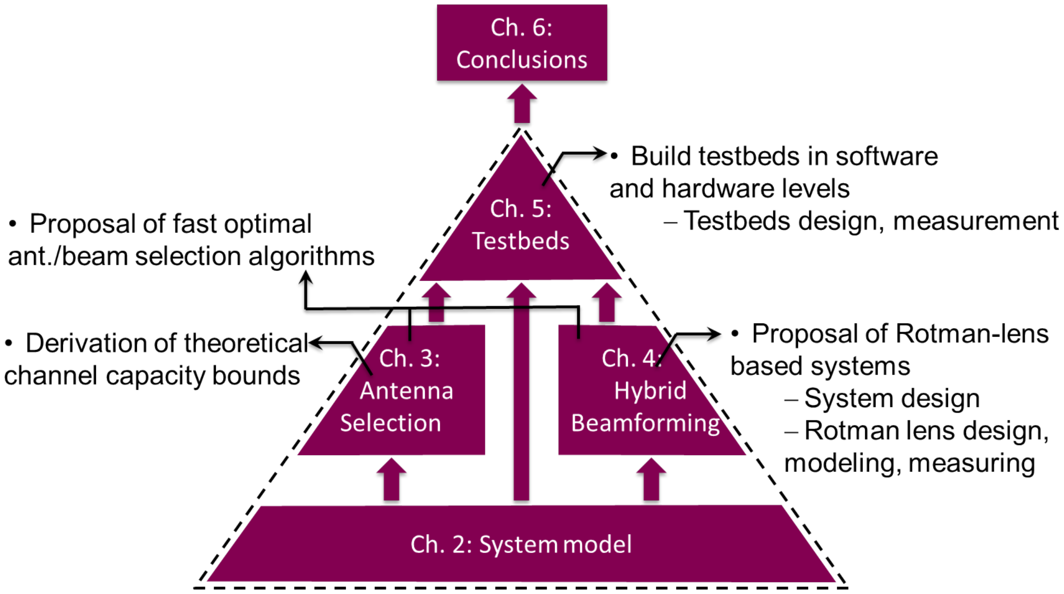


Figure 1.1: Schematic structure and main contributions of the thesis

Chapter 2

System Model

In this chapter, the general system model, i.e., a single-cell uplink multi-user massive MIMO system, used in this thesis is described. Based on the general system model, the specified system model used in each low RF-complexity massive MIMO system, i.e., antenna selection system and hybrid analog-digital beamforming system, can be specified further in Chapter 3 and 4, respectively. In addition, the general concept of the low RF-complexity massive MIMO systems and the associated array-architectures will also be described in this chapter.

2.1 Multi-user Massive MIMO

As shown in Fig. 2.1, we consider a single cell multi-user MIMO system with a single Base Station (BS) equipped with N_B antennas and N_U single-antenna User Equipments (UEs). For simplicity, we only consider the uplink, i.e., data communication link from UEs to BS, in this thesis. It should be noted that the derived theories, algorithms, and the designed hardware and testbeds in this thesis can be potentially used for the downlink as well. In the system, N_U UEs transmit N_U independent data streams (one data stream per UE) within the same frequency and time resource. The transmitted multi-user data streams will be spatially multiplexed through wireless channels and received by the BS. The BS will decouple the spatially multiplexed data streams based on the fact that the each user data steam undergoes independent or low-correlated spatial channels. This type of system is also called multi-user MIMO system, and the involved multiplexing technology is also referred to as spatial multiplexing.

In contrast to the conventional small-scale multi-user MIMO, where N_B and N_U are not too large, this thesis focuses on the multi-user MIMO system especially with a *massive* number

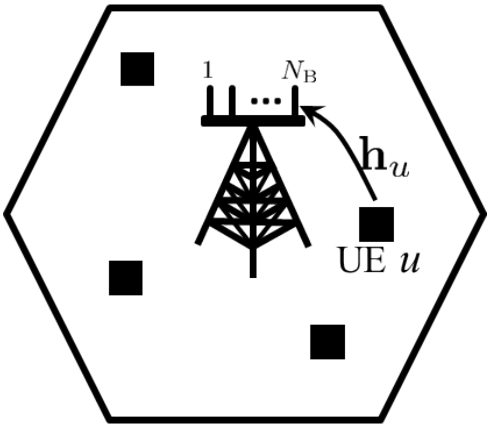


Figure 2.1: A single-cell uplink multi-user massive MIMO system. N_B antennas in BS and N_U single-antenna UEs.

of antennas. We assume a narrowband massive MIMO system with frequency-flat fading. It should be noted that it is straightforward to extend the narrowband system to the wideband system since the wideband system can be simplified to narrowband system with, for example, Orthogonal Frequency-Division Multiplexing (OFDM) technique. In this thesis, we will first use the narrowband system model to derive theories and to design systems in Chapter 3 and 4. The extension to the wideband case will be derived in Chapter 5 and studied experimentally.

The narrowband multi-user massive MIMO systems can be described by the following simple linear signal model [15] - [18]

$$\mathbf{y} = \sqrt{E_\mathrm{x}}\mathbf{H}\mathbf{x} + \mathbf{w}, \tag{2.1}$$

where $\mathbf{x} \in \mathcal{A}^{N_\mathrm{U}\times 1}$ is the transmitted symbol vector with the mean $\mathbb{E}\{\mathbf{x}\} = \mathbf{0}$ and the covariance matrix $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \frac{1}{N_\mathrm{U}}\mathbf{I}_{N_\mathrm{U}}$; $\mathcal{A}$ with the cardinality $|\mathcal{A}| = A$ is the set consisting of the constellation points of a modulation scheme; E_x is the total transmitted power of N_U UEs; $\mathbf{w}$ is the Additive White Gaussian Noise (AWGN) following the complex Gaussian distribution $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{N_\mathrm{B}})$ with the noise variance σ^2 and the Signal to Noise Ratio (SNR) is consequently defined as $\rho = E_\mathrm{x}/\sigma^2$; $\mathbf{H}$ denotes the channel matrix with $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_{N_\mathrm{U}}]$ and $\mathbf{h}_u \in \mathbb{C}^{N_\mathrm{B}\times 1}$ is the channel vector of the uth user. The channel matrix $\mathbf{H}$ plays a key role in communication systems, and will be discussed further in Chapter 3 and 4.

2.2 Low RF-Complexity Massive MIMO Systems

With the uplink model, the main task or workload lies in the BS receiver to estimate the transmitted symbol vector $\mathbf{x}$, given $\mathbf{y}$ and $\mathbf{H}$. Note that the channel matrix $\mathbf{H}$ is unknown in practice, which needs to be estimated before estimating $\mathbf{x}$ in coherent detection. In Chapter 3 and 4, we assume that $\mathbf{H}$ is perfectly known for the convenience of system design; while in Chapter 5, $\mathbf{H}$ will be estimated through channel estimation for experimental study. Obviously, with the large-scale signals $\mathbf{y}$ and $\mathbf{H}$, the detection of the discrete vector $\mathbf{x}$ is challenging in hardware. In particular, a large number of RF transceivers are required to down-convert and digitize the received antenna signals into the digital baseband, leading to very high RF-complexity and hardware cost. The RF complexity and RF hardware cost become even more challenging when scaling up MIMO under the same increasing rate of the number of RF transceivers and antennas in BS. On the contrary, the low RF-complexity massive MIMO systems intend to scale up the systems in a more economic way. The questions are: what are the low RF-complexity massive MIMO systems that this thesis focuses on, and to which scenarios the low RF-complexity massive MIMO systems are able to be applied? Before answering these questions, let us first give several definitions used in this thesis.

Definition 2.2.1 (Full-digital Multi-User Massive MIMO Systems) *Given an $N_U \times N_B$ multi-user massive MIMO system with N_U single-antenna UEs, and one BS with N_B antennas and L RF transceivers, the full-digital multi-user massive MIMO system is referred to as a system with $N_B = L \gg 1$.*

Definition 2.2.2 (Balance-configured Massive MIMO Systems) *A balanced-configured $N_U \times N_B$ multi-user massive MIMO system is a system that both N_B and N_U are large and comparable, tending to infinity at the same rate, which can be characterized by*

$$\lim_{N_B, N_U \to \infty} \frac{N_U}{N_B} = c, \tag{2.2}$$

where c is a positive constant.

Definition 2.2.3 (Unbalance-configured Massive MIMO Systems) *An unbalanced-configured $N_U \times N_B$ multi-user massive MIMO system is a system with $N_B \gg N_U$ and the increasing rate* *of N_B is faster than that of N_U, which can be characterized by*

$$\lim_{N_B, N_U \to \infty} \frac{N_U}{N_B} = 0. \tag{2.3}$$

Obviously, the low RF-complexity design is especially attractive for the unbalance-configured setup rather than the balance-configured setup in multi-user massive MIMO systems. In the unbalance-configured systems, the tall matrix $\mathbf{H} \in \mathbb{C}^{N_\mathrm{B} \times N_\mathrm{U}}$ implies that the useful information embedded in $\mathbf{y}$ has dimension of only N_U, which is much less than the original high dimension of N_B. Thus the high-dimensional receive signal vector $\mathbf{y}$ can be potentially compressed by simple analog signal processing network to low-dimensional signal vector, which allows a reduced number of RF transceivers to retain the transmit information. Besides the hardware-complexity reduction, another benefit of the dimension reduction is to reduce the computational complexity in the digital domain since only low-dimensional signal vector is now necessary to be used to estimate $\mathbf{x}$.

By summarizing the aforementioned definitions and explanations, the low RF-complexity massive MIMO system that is specifically studied in this thesis is defined as follows:

Definition 2.2.4 (Low RF-complexity Massive MIMO Systems) *Given an $N_U \times N_B$ multi-user massive MIMO system with L RF transceivers in BS. A low RF-complexity multi-user massive MIMO system is defined as*

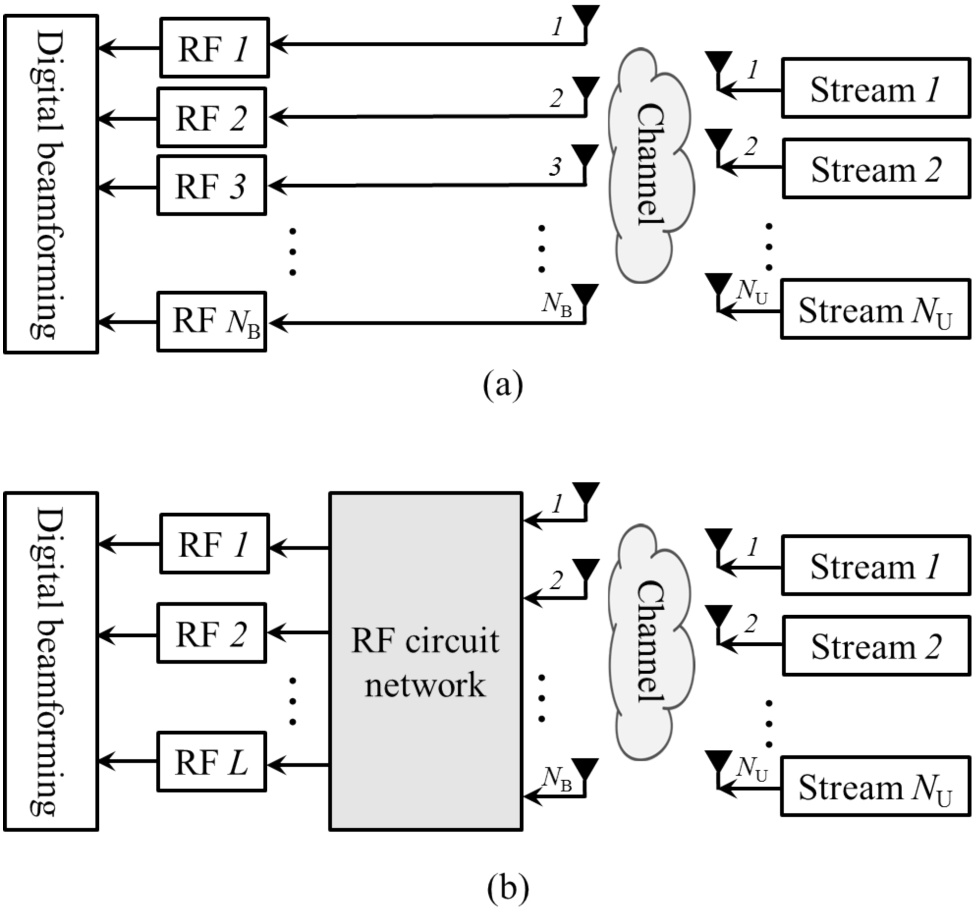
$$\lim_{N_B, N_U \to \infty} \frac{N_U}{N_B} = 0, \tag{2.4}$$

$$\lim_{N_B, L \to \infty} \frac{L}{N_B} = 0, \tag{2.5}$$

$$L \geq N_U. \tag{2.6}$$

Condition (2.4) specifies the application scenario of the low RF-complexity design in the multi-user MIMO systems to be unbalance-configured; condition (2.5) specifies that the increasing rate of the number of RF transceivers in BS should be less than that of BS antennas; condition (2.6) is necessary to allow BS to decouple N_U independent data streams (N_U unkonwn variables need at least N_U equations to estimate, which is able to retain the spatial multiplexing gain of the full-digital massive MIMO).

In this thesis, we fix $L = N_\mathrm{U}$ and let $N_\mathrm{B} \to \infty$, which is actually a special case of the general low RF-complexity massive MIMO systems given in Definition 2.2.4.



(a)

(b)

Figure 2.2: System diagrams of (a) full-digital massive MIMO, (b) low RF-complexity massive MIMO

2.3 Array Architectures

The system diagrams of both full-digital and low RF-complexity massive MIMO systems are shown in Fig. 2.2. Obviously, the RF circuit network is a critical component in the low RF-complexity massive systems, which is used to bridge the small number of RF transceivers and the large number of antennas in BS. In principle, the RF circuit network should be of low hardware complexity and feasible to implement; otherwise, we cannot thoroughly benefit from the low RF-complexity design. The complexity of the RF networks mainly depends on the array architectures. Generally, there are two most interesting types of array architectures, i.e., full-array architecture and sub-array architecture, with the definitions as follows:

Definition 2.3.1 (Full-array Architecture) *A full-array analog signal processing architecture is an architecture that all of the array elements, i.e., antennas, are interconnected to allow the*

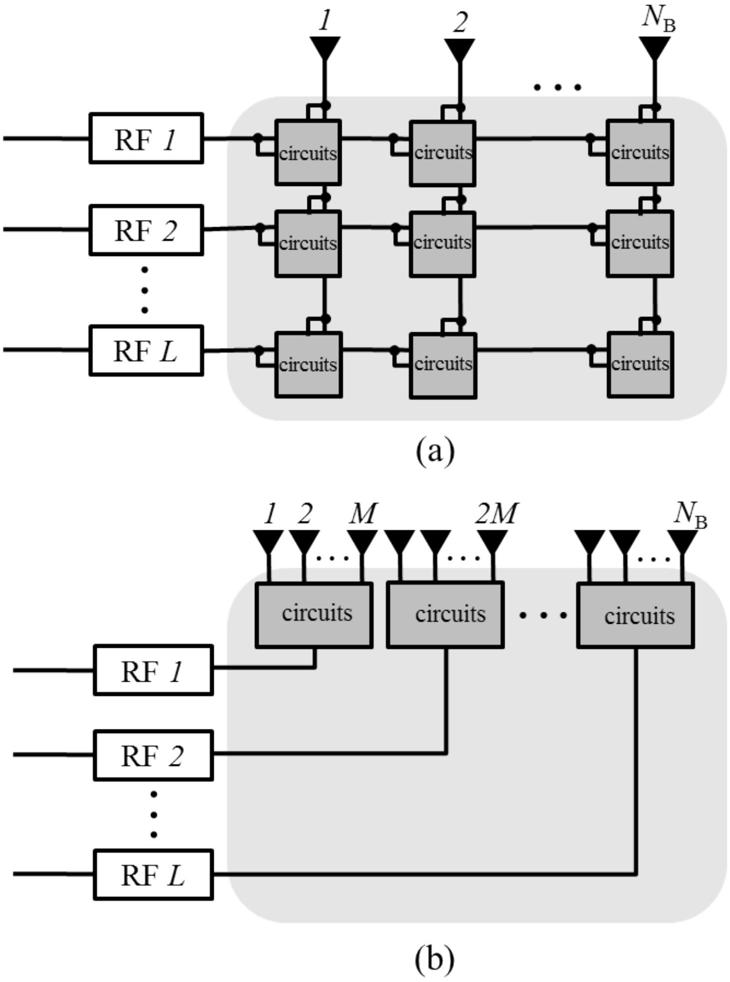


Figure 2.3: Diagrams of (a) full-array architecture, (b) sub-array architecture

joint signal processing to the all the array element signals.

Definition 2.3.2 (Sub-array Architecture) *In a sub-array analog signal processing architecture, the whole array is segmented into multiple disjoint sub-arrays, and the array elements belonging to the same sub-array are interconnected with each other. The analog signal processing is done in each sub-array separately.*

Fig. 2.3 shows the diagrams of the full-array and sub-array architectures. In the full-array architecture, each RF transceiver is connected with the whole antenna array with the basic analog circuits (denoted by solid boxes); while in the sub-array architecture, the whole antenna array is segmented into L sub-arrays of size M, and each RF transceiver is only connected to one sub-array antennas via the sub-RF-circuits (denoted by the rectangular boxes). Take the RF switching network as an example, which is the RF circuit network used in the antenna selection systems. In the full-array switching architecture, each output port can be switched to an arbitrary input port to allow being able to select an arbitrary subset of antennas; in contrast, in the sub-array switching architecture, the output port in one sub-array can only be switched to the input ports corresponding to that sub-array. Another example is the RF phase-shifting network that is the RF circuit network used in the hybrid beamforming systems. In the full-array phase-shifting network, all the input ports are interconnected to the output port via transmission lines, power dividers, power combiners and phase shifters, allowing the beamforming by using the whole array signals; in contrast, the sub-array phase-shifting network will constrains the interconnection and beamforming only within each sub-array. More detailed description on the full-array and sub-array architectures will be given in Chapter 3 and 4 for the RF switching and RF phase-shifting networks, respectively.

An intuition from Definition 2.3.1 and 2.3.2 is that the hardware complexity of the full-array architecture is higher than that of the sub-array architecture. In particular, as N_{B} increases, the full-array architecture tends to be infeasible to implement with low insertion loss and high isolation. In contrast, for the sub-array architecture, we can always segment the large-size full-array into many small-size sub-arrays that are sufficiently low-complex to implement. However, it should be noted that, when the RF networks are ideal without any loss, the system with the sub-array architecture has inferior performance compared to the system with the full-array architecture due to the connection constraint introduced by the array segmentation. Their performance gap will be especially studied in Chapter 3 and 4 for the antenna selection systems and the hybrid analog-digital beamforming systems, respectively.

Chapter 3

Massive MIMO With Antenna Selection

In this chapter, we will study the low RF-complexity masssive MIMO systems with RF switches. As the RF switches are used to select a subset of antennas, this low RF-complexity massive MIMO systems is also known as massive MIMO antenna selection systems. We derive the upper bounds on channel capacity for the massive MIMO antenna selection systems with both full-array and sub-array architectures. In particular, the number of antennas gets large in massive MIMO antenna selection systems, which motivates us to use the asymptotic theory on order statistics to model the ordered random variables. The asymptotic distributions can give much simpler forms than the exact distribution, which consequently results in relatively simple forms of the upper bounds on capacity. Based on the derived analytical bounds, we also derive the upper bound of the ergodic capacity, to study the impact of the channel acquisition duration. Besides the analytical work, the optimal antenna selection algorithms based on BAB searching algorithms are developed for both full-array and sub-array architectures. Our numerical and simulation results show that, (1) the derived asymptotic bounds are tight and also apply to the finite-dimensional MIMO; (2) the CSI acquisition is one of the main limits for the massive MIMO antenna selection systems in the time-variant channels; (3) the proposed optimal antenna selection algorithms are much faster than the exhaustive-search-based antenna selection, e.g., three orders of magnitude speedup as observed in the large-scale system; (4) most importantly, the antenna selection system with sub-array architecture has very close performance to the antenna selection system with full-array architecture, which outperforms the small-scale MIMO under the condition of the same number RF transceivers. The results in this chapter show that the low RF-complexity massive MIMO systems based on sub-array antenna selection are very promising in performance for practical applications.

3.1 Introduction

In the past decades, many antenna selection criteria and algorithms were presented for the conventional small-scale MIMO. For example, the works in [26, 27] studied error-rate-oriented selection criteria with specific selection algorithms for practical receivers. Many works focused on the capacity-oriented selection criteria, such as the greedy search in [28, 29], convex optimization in [30], dominant-submatrix search in [31]. In recent years, some of them have also been studied and extended for the massive MIMO systems, such as the works in [32]- [37]. However, it should be noted that none of these massive MIMO antenna selection algorithms are optimal in the sense of maximum capacity. Exhaustive search, which is widely used in the conventional small-scale MIMO systems, becomes infeasible for the massive MIMO system due to the tremendous number of antenna subsets. An optimal antenna selection algorithm faster than the exhaustive search is thus necessary for the massive MIMO antenna selection systems.

In contrast to the rich results on the algorithm design, the theoretical results on the channel capacity with antenna selection were very limited. Gorokhov et al., in [38] gave a lower bound on the capacity achievable through antenna selection. Molisch et al., in [58] derived a statistical upper channel capacity bound by using the explicit joint distributions of the order statistics. The work in [40] analyzed the channel capacity of the antenna selection system based on the sub-optimal greedy search. The ergodic capacity of the MIMO multicasting channels with antenna selection was studied in [41]. For the systems with a large number of antennas, an approximated analytical form of the distributions of multiple-input single-output channel capacity with transmit antenna selection was derived in [42]. Single-antenna selection for large-scale MIMO was analyzed in [43]. Both [42] and [43] showed a channel hardening effect retained as the full-digital MIMO systems. To our knowledge, the capacity of massive MIMO channels with multiple-antenna selection is still missing.

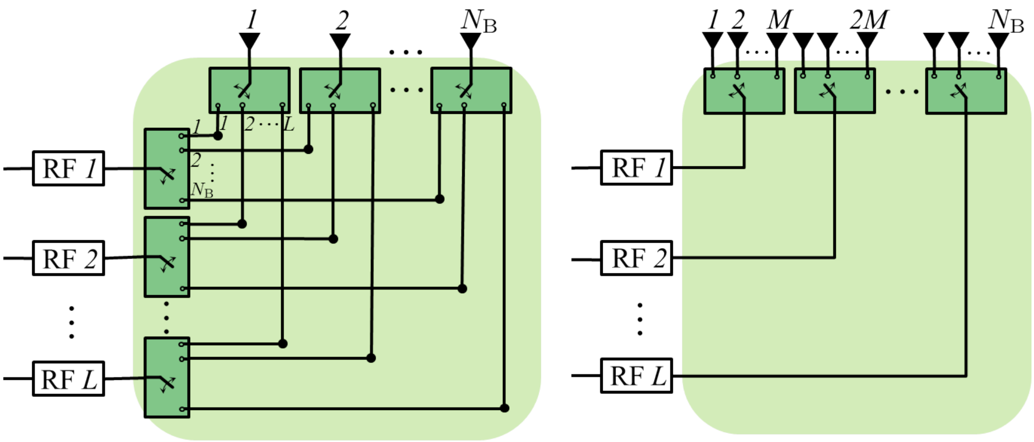
Another issue on the massive MIMO antenna selection is the array architecture. It is known that the RF switches are the critical RF-analog components in the MIMO antenna selection systems. In the conventional small-scale MIMO antenna selection systems, a fully-switching RF network is usually assumed, which supports to select any subset of antennas, as shown in Fig. 3.1 (a), to which we refer as Full-Array Switching (FAS). However, as the MIMO dimension increases, the fully-switching RF network becomes less power efficient [90] and even infeasible to implement. This issue dedicated for the massive MIMO antenna selection systems is getting more attention recently. For example, the works in [44, 45] considered a sub-array architecture, in which each RF chain is connected to one disjoint sub-array of size M, as shown in Fig. 3.1 (b), to which we refer as Sub-Array Switching (SAS). Similar architecture can also be found in [46], where each sub-array has $M = 2$ antennas. The authors in [90] generalized the sub-array architecture such that each antenna is allowed to be connected to multiple RF chains. For the convenience of analysis, we only focus on the FAS and SAS systems in this thesis.

In this chapter, we study the antenna selection at the receiver side that equips with a large number of antennas N_{B}, but with only a limited number of RF transceivers L, i.e., a typical low RF-complexity system defined in Chapter 2. By assuming independent and identically distributed (i.i.d.) Rayleigh-fading channels, we derive the statistical upper channel capacity bounds for both the SAS and FAS systems in the large-scale limit. In particular, the asymptotic theory on order statistics is used to model the distributions of the maximum and top-L random variables from a sequence of a large number of random variables. Numerical results show that the derived asymptotic bounds can also apply to the finite-dimensional MIMO systems where the number of antennas is not too large. By using the derived statistical upper bounds, the ergodic upper bound of the achievable spectral efficiency considering Channel State Information (CSI) acquisition is further derived. Our results show that the channel training is one of the main limits for the massive MIMO antenna selection systems with time-variant channels.

In addition, the optimal antenna selection algorithms are proposed for the SAS and FAS systems. In particular, we use Branch And Bound (BAB) searching algorithm and the associated search tree to find the antenna subset with the largest channel capacity. Since the objective function, i.e., the channel capacity, cannot be directly used for the BAB search, we address this issue by proposing an equivalent objective function. Due to the pruning operations in the BAB search tree, the BAB search is much more efficient than the exhaustive search, e.g., 3 orders of magnitude speedup observed in our simulation results when the number of antennas is large.

3.2 System Model and Problem Statement

The system diagrams with antenna selection are shown in Fig. 3.2 (a) and (b) for the full-array and sub-array antenna selection systems, respectively. Compared to the general low RF-complexity system model shown in Fig. 2.2 (b), the RF circuit network is specified to the RF switching network as shown in Fig. 3.1. Next we follow the signal model as given in Equ. (2.1) and can



(a) FAS

(b) SAS

Figure 3.1: RF switching network architecture to select L antennas from N_{B} antennas ($L < N_{\mathrm{B}}$): (a) full-array RF switching, (b) sub-array RF switching. For the sub-array switching architecture, each sub-array has M antennas.

directly get the signal model with receive antenna selection as

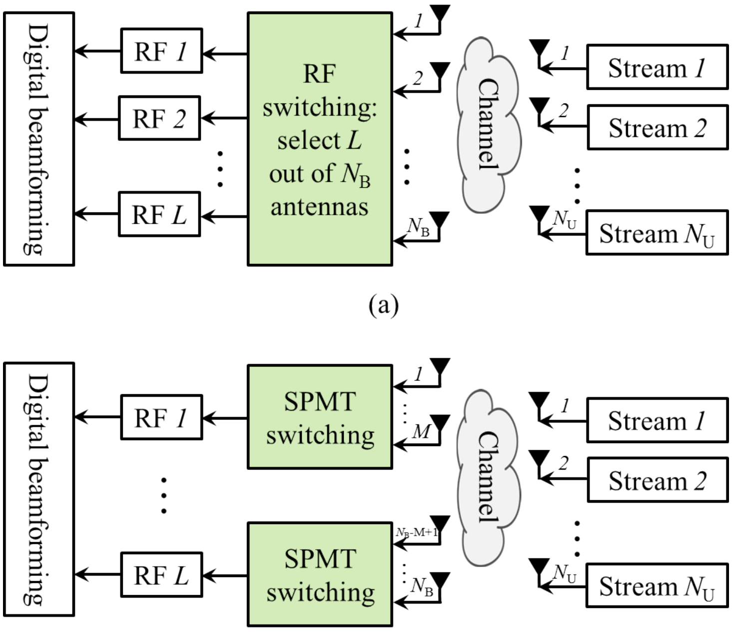
$$\begin{aligned} \mathbf{y}_{\mathrm{s}} &= \sqrt{E_{\mathrm{x}}}\mathbf{S}_{\mathrm{RF}}\mathbf{H}\mathbf{x} + \mathbf{w}_{\mathrm{s}} \\ &= \sqrt{E_{\mathrm{x}}}\mathbf{H}_{\mathrm{s}}\mathbf{x} + \mathbf{w}_{\mathrm{s}} \end{aligned} \tag{3.1}$$

where $\mathbf{S}_{\mathrm{RF}}$ is a selection matrix to model RF switches (an example of $\mathbf{S}_{\mathrm{RF}}$ is given by Equ. (4.3)); $\mathbf{H}_{\mathrm{s}} = \mathbf{S}_{\mathrm{RF}}\mathbf{H}$ is the selected channel submatrix from the complete channel matrix $\mathbf{H}$; $\mathbf{w}_{\mathrm{s}}$ is the additive complex Gaussian noise following $\mathcal{CN}(\mathbf{0}, \mathbf{I}_L)$. In the derivation of the upper bounds of the channel capacity in Section 3.3, we assume that the channel follows Rayleigh flat fading and the elements in $\mathbf{H}$ are i.i.d. complex Gaussian random variables with zero-mean and unit-variance. In the development of the optimal antenna selection algorithms in Section 3.4, the channel is general and has no limitations on the distributions.

We assume that the channel state information is only available at the receiver side. In addition, $\mathbf{x}$ is assumed to follow the independent Gaussian distribution with equal power components. The channel capacity of the full-digital system is then written as [19]

$$C_{\mathrm{fd}} = \log_2 \det(\mathbf{I}_{N_{\mathrm{U}}} + \bar{\rho}\mathbf{H}^H\mathbf{H}) \tag{3.2}$$

where $\bar{\rho} = \rho/N_{\mathrm{U}}$. With the channel state information at the receive side, the receive antenna selection can be performed, where the best L out of N_{B} receive antennas are selected, i.e.,



(a)

(b)

Figure 3.2: System models: (a) full-array antenna selection system; (b) sub-array antenna selection system

selecting a channel submatrix $\mathbf{H}_{\text{s}}$ of size $L \times N_{\text{U}}$ from the complete channel matrix $\mathbf{H}$. Let $\mathcal{H}_{\text{s,fas}}$ and $\mathcal{H}_{\text{s,sas}}$ denote the sets consisting of all possible $L \times N_{\text{U}}$ channel submatrices for the FAS and SAS systems, respectively. In order to optimize the system performance with respect to the channel capacity, the best L antennas are selected according to the optimization problem

$$\mathbf{H}_{\text{s, fas}|\text{sas}}^{\text{opt}} = \arg \max_{\mathbf{H}_{\text{s}} \in \mathcal{H}_{\text{s,fas}|\text{sas}}} \log_2 \det(\mathbf{I}_{N_{\text{U}}} + \bar{\rho}\mathbf{H}_{\text{s}}^{H}\mathbf{H}_{\text{s}}), \tag{3.3}$$

for the FAS and SAS systems, where $\bar{\rho} = \rho/N_{\text{U}}$, "|" denotes "or", and the superscript $(\cdot)^{H}$ denotes Hermitian transpose. The channel capacities of the FAS and SAS systems are thus $C_{\text{s, fas}|\text{sas}}^{\text{opt}} = \log_2 \det(\mathbf{I}_{N_{\text{U}}} + \bar{\rho}\mathbf{H}_{\text{s, fas}|\text{sas}}^{\text{opt},H}\mathbf{H}_{\text{s, fas}|\text{sas}}^{\text{opt}})$.

There are several interesting questions around the optimization problem given by Equ. (3.3). Given the random matrix $\mathbf{H}$, the channel capacities $C_{\text{s, fas}|\text{sas}}^{\text{opt}}$ are random variables as well. The distributions of $C_{\text{s, fas}|\text{sas}}^{\text{opt}}$ are of great interest to evaluate the performance of the massive MIMO antenna selection system. Since the closed-form solutions of $C_{\text{s, fas}|\text{sas}}^{\text{opt}}$ seems to be difficult to get, the statistical upper bounds are more practical. Thus the first question is: what are the theoretical upper capacity bounds of the FAS and SAS systems in massive MIMO?

Secondly, it should be noted that the optimization problem given by Equ. (3.3) can be solved only when the knowledge of $\mathbf{H}$ is available. Since only $L < N_{\text{B}}$ RF transceivers are available at the receiver, it is impractical to acquire the full channel knowledge at one time. This issue can be dealt with when the channel is static or slow-fading [58]. Specifically, the transmitter can send the pilot symbols with $M = N_{\text{B}}/L$ cycles in the training stage. In each cycle, each RF receiver will switch to one of the M receive antennas. Obviously, the extra $(M-1)$ training cycles will degrade the spectral efficiency, especially in massive MIMO systems where M can be very large. Thus the second question is: can we always benefit from the antenna selection by increasing the the number of antennas if we take the channel training into account?

Thirdly, to solve the optimization problem given by Equ. (3.3), antenna selection algorithms are required. The antenna selection problem given by Equ. (3.3) is essentially a submatrix selection problem to find the submatrix with the largest determinant, which is known to be NP-hard [54, 93]. In this case, exponential complexity is usually involved to find the optimal solution. One of the well-known optimal methods is naturally the exhaustive search, which has the exponential complexity of $|\mathcal{H}_{\text{s,fas}}| = \binom{N_{\text{B}}}{L}$ and $|\mathcal{H}_{\text{s,sas}}| = (\frac{N_{\text{B}}}{L})^{L}$ for the FAS and SAS systems, respectively, becoming infeasible when N_{B} is large. Thus the third question is: can we find a faster but also optimality-guaranteed antenna selection algorithm? In the following we will investigate the first two questions in Section 3.3; while the third question will be addressed in Section 3.4.

3.3 Upper Capacity Bounds

In this section, some preliminaries will be first given, followed by the derivation of the distribution of the upper bound on the antenna selection channel capacity. Afterwards, the mean capacity can be consequently obtained, to study the impact of the channel training period.

3.3.1 Preliminaries

Upper bounds on the MIMO channel capacity

The starting point to derive the upper capacity bounds is the general MIMO channel capacity bound used in [19, 58]. Given the channel matrix $\mathbf{H}$ with i.i.d. complex Gaussian entries, the upper bound on the capacity of a full-digital MIMO system can be defined as

$$C_{\text{ub,fd}} = \sum_{l=1}^{N_{\text{U}}} \log_2{(1 + \bar{\rho}\gamma_l)}, \tag{3.4}$$

where the random variables $\{\gamma_l\}_{l=1,2,\cdots,N_{\text{U}}}$, denote the chi-squared-distributed random variables with $2N_{\text{B}}$ degrees of freedom, corresponding to the transmit antennas $l = 1, 2, \cdots, N_{\text{U}}$. This capacity bound is obtained by assuming an impractical case that each transmit antenna has its own N_{B} receive antennas [19]. In a similar way, if we assume that each receive antenna has its own N_{U} transmit antennas, the upper bound on the capacity of the full-digital MIMO sysem can be given by

$$\tilde{C}_{\text{ub,fd}} = \sum_{l=1}^{N_{\text{B}}} \log_2{(1 + \bar{\rho}\tilde{\gamma}_l)}, \tag{3.5}$$

where the random variables $\{\tilde{\gamma}_l\}_{l=1,2,\cdots,N_{\text{B}}}$, denote the chi-squared-distributed random variables with $2N_{\text{U}}$ degrees of freedom, corresponding to the receive antennas $l = 1, 2, \cdots, N_{\text{B}}$. When the receive antenna selection is performed, the upper bound on the capacity for the FAS systems can be defined by [58]

$$C_{\text{ub,fas}} = \sum_{l=1}^{L} \log_2{(1 + \bar{\rho}\gamma_{(l)})}, \tag{3.6}$$

where the random variables $\{\gamma_{(l)}\}_{l=1,2,\cdots,N_{\text{B}}}$, are the *ordered* chi-square-distributed random variables, i.e., $\gamma_{(1)} > \gamma_{(2)} > \cdots > \gamma_{(N_{\text{B}})} > 0$, with $2N_{\text{U}}$ degrees of freedom. Obviously, only the first L-largest random variables, i.e., $\{\gamma_{(l)}\}_{l=1,2,\cdots,L}$, are selected due to the antenna selection. Following the principles of definitions in [19, 58], we can define the upper capacity bound for the SAS system as

$$C_{\text{ub,sas}} = \sum_{l=1}^{L} \log_2{(1 + \bar{\rho}\gamma_{\text{max}_M,l})}, \tag{3.7}$$

where $\gamma_{\text{max}_M,l}$ is the random variable to take the *maxima* of M realizations of γ_i in the lth sub-array, i.e., $\gamma_{\text{max}_M,l} = \max(\gamma_{(l-1)M+1}, \gamma_{(l-1)M+2}, \cdots, \gamma_{lM})$. In contrast to the non-identical and dependent random variables $\{\gamma_{(l)}\}$, the random variables $\{\gamma_{\text{max}_M,l}\}$, are i.i.d. instead, which implies that the analytical form of $C_{\text{ub,sas}}$ can be derived more easily than that of $C_{\text{ub,fas}}$. Thus in this chapter, we will first focus on the distribution of $C_{\text{ub,sas}}$, followed by that of $C_{\text{ub,fas}}$.

Asymptotic theory on order statistics

From Equ. (3.7) and (3.6) we can see that the distribution of the maximum random variable and the joint distribution of the top-L random variables are required to derive the analytical form of the upper capacity bounds of the SAS and FAS systems respectively. The exact distributions are usually computationally messy, especially in the cases that the number of antennas is large. Instead of the exact distributions, asymptotic distributions usually give much simpler forms, and thus become of special interest in the sense of large-scale behaviour. Meanwhile we also hope that the simple asymptotic distributions can approximate the exact distribution of the small-scale cases as well, since only limited number of antennas are feasible in the practical systems.

Let us first have a look at the asymptotic theory on the maximum value, also called extreme value theory. Suppose that $X_1, X_2, \ldots, X_n$ is an i.i.d. sequence of n random variables with common CDF $F(x)$. Let $X_{\max_n} = \max(X_1, X_2, \ldots, X_n)$. The asymptotic theory on the maximum value is to study the distribution of $X_{\max_n}$ given the common CDF $F(x)$ when $n \to \infty$, and has the following results:

Definition 3.3.1 *[49, 50] A CDF $F(x)$ is said to belong to the maximum domain of attraction of a nondegenerate CDF $G(x)$ if there exist a centring constant μ_n and a normalizing constant σ_n such that* $\lim_{n\to\infty} F^n(\mu_n + \sigma_n x) = G(x)$ *at all continuity points of $G(x)$.*

Lemma 3.3.2 *[49, 50] If $F(x)$ belongs to the maximum domain of attraction, then*

$$\frac{X_{max_n} - \mu_n}{\sigma_n} \xrightarrow{a.s.} Y, \tag{3.8}$$

where the symbol "$\xrightarrow{a.s.}$" denotes almost surely convergence. The random variable Y has the distribution $G(x)$, and $G(x)$ can only *be one of the following three distributions:*

Fréchet: $$G_1(x) = \begin{cases} 0, & x \leq 0, \alpha > 0, \\ \exp(-x^{-\alpha}), & x > 0, \alpha > 0 \end{cases}$$

Weibull: $$G_2(x) = \begin{cases} \exp(-(-x)^{\alpha}), & x \leq 0, \alpha > 0, \\ 1, & x > 0, \alpha > 0 \end{cases}$$

Gumbel: $$G_3(x) = \exp(-e^{-x}), x \in \mathbb{R}$$

In this thesis, we are only interested in the Gumbel distribution, because when X_i follows the chi-square distribution, the distribution of $X_{\max_n}$ belongs to $G_3(x)$ according to [49] (Table 3.4.4, Gamma distribution with $\alpha = K, \beta = 1$).

In the above extreme value theory, the centring constant μ_n and the normalizing constant σ_n are critical. Actually, the constants μ_n and σ_n are not unique in the extreme value theory. Different approximations methods for μ_n and σ_n can achieve very similar results when n is large, but noticeable difference when n is not too large. In this thesis, we use the following equations [49]:

$$\mu_n = F^{-1}(1 - \frac{1}{n}), \tag{3.9}$$

$$\sigma_n = \int_{\mu_n}^{\infty} \frac{1 - F(t)}{1 - F(\mu_n)} dt, \tag{3.10}$$

where $F^{-1}(x)$ is the inverse function of $F(x)$. Through our numerical experiments (not shown in this thesis), Equ. (3.9) and (3.10) give the best approximation when n is not too large.

Now let us have a look at the asymptotic joint distribution of the top-L random variables of the random sequence $X_1, X_2, \cdots, X_n$, as $n \to \infty$.

Lemma 3.3.3 *[50, 51] Given an i.i.d. random variable sequence $X_1, X_2, \cdots, X_n$ with the common CDF $F(x)$ that belongs to the maximum domain of attraction, i.e., $\frac{X_{max_n} - \mu_n}{\sigma_n}$ has the one of three distributions $G(x)$ in Lemma 3.3.2. Suppose that $X_{(1)} > X_{(2)} > \cdots > X_{(n)}$ is the ordered sequence of $X_1, X_2, \cdots, X_n$, then the L-dimensional vector $(\frac{X_{(1)} - \mu_n}{\sigma_n}, \frac{X_{(2)} - \mu_n}{\sigma_n}, \cdots, \frac{X_{(L)} - \mu_n}{\sigma_n})$ has the following asymptotic joint distribution*

$$g_{1,2,\cdots,L}(x_{(1)}, x_{(2)}, \cdots, x_{(L)}) = G(x_{(L)}) \prod_{l=1}^{L} \frac{g(x_{(l)})}{G(x_{(l)})}, \tag{3.11}$$

$$\textit{with}\, x_{(1)} > x_{(2)} > \cdots > x_{(L)},$$

where $g(x)$ is the Probability Density Function (PDF) of X_{max_n}.

3.3.2 Upper capacity bound of the SAS system

Let us start with the distribution of γ_i. Recall that the real and imaginary parts of the channel matrix element h_{ij} follows the independent Gaussian distribution with zero mean and variance of 1/2. Thus the random variable $\gamma_i = \sum_{j=1}^{N_U} = |h_{ij}|^2$ follows the chi-square distribution with the PDF

$$f_{\gamma_i}(x) = \frac{1}{\Gamma(N_\text{U})} x^{N_\text{U}-1} e^{-x}, x > 0, \tag{3.12}$$

and CDF

$$F_{\gamma_i}(x) = \frac{\gamma(N_\text{U}, x)}{\Gamma(N_\text{U})}, \tag{3.13}$$

where $\Gamma(\cdot)$ is Euler's Gamma function and $\gamma(\cdot, \cdot)$ is the lower incomplete Gamma function [48].

Now let us turn our attention to $\gamma_{\text{max}_M, l}$ in Equ. (3.7), i.e., the maxima of the lth sub-array. Since $\gamma_{\text{max}_M, l}$ has i.i.d. distribution regarding l, for convenience we drop the subscript l and just use γ_{max_M} to denote them in the rest of the chapter. Given γ_{max_M} as the maxima of M chi-square random variables $\gamma_1, \gamma_2, \cdots, \gamma_M$. The exact PDF of γ_{max_M} is known as [40]

$$f_{\gamma_{\text{max}_M}}(x) = \frac{M x^{N_\text{U}-1}}{(N_\text{U}-1)!} e^{-x} \left(1 - e^{-x} \sum_{k=0}^{N_\text{U}-1} \frac{x^k}{k!}\right)^{M-1}. \tag{3.14}$$

Obviously, the exact PDF given by Equ. (3.14) is too complicated to be used for the derivation of the analytical form of $C_\text{ub,sas}$. This motivates us to go for a simpler approximation form. Based on the extreme value theory in Lemma 3.3.2, γ_{max_M} asymptotically converges to Gumbel distribution as $M \to \infty$, i.e., $\gamma_{\text{max}_M} \xrightarrow{a.s.} \bar{\gamma}_{\text{max}_M}$, and $\bar{\gamma}_{\text{max}_M}$ is the Gumbel-distributed random variable with the PDF

$$f_{\bar{\gamma}_{\text{max}_M}}(x) = \frac{1}{\sigma_M} e^{-\frac{x-\mu_M}{\sigma_M}} e^{-e^{-\frac{x-\mu_M}{\sigma_M}}}, \tag{3.15}$$

where μ_M and σ_M are the centring and normalizing constants, respectively, calculated by Equ. (3.9) and (3.10) by replacing n with M.

Obviously, the asymptotic distribution given by Equ. (3.15) is much simpler than the exact distribution given by Equ. (3.14). But the question is: is the asymptotic distribution still sufficiently accurate when M is not too large? Fortunately, for the chi-squared-distributed random variables γ_i, the distribution of γ_{max_M} converges to Gumbel distribution very fast. In our simulation results shown in Section 3.5 (Fig. 3.5 with $N_\text{B} = 16$), $M = 4$ is already sufficiently large. Thus it is reasonable to use the Gumbel distribution to approximate the distribution of γ_{max_M} and derive the capacity bound.

Since $C_\text{ub,sas}$ is the sum of random variables shown in Equ. (3.7), the characteristic function is naturally a useful tool to derive the distribution. Let us define the characteristic function of

$C_{\text{ub,sas}}$ as $\Phi_{\text{ub,sas}}(\jmath\omega)$ which thus has the convergence

$$\Phi_{\text{ub,sas}}(\jmath\omega) \xrightarrow{a.s.} \tilde{\Phi}^{L}_{\text{ub,sas}}(\jmath\omega), \tag{3.16}$$

where $\tilde{\Phi}_{\text{ub,sas}}(\jmath\omega)$ is the characteristic function of $\log_2(1+\bar{\rho}\bar{\gamma}_{\max_M})$, and it is defined as

$$\begin{aligned}\tilde{\Phi}_{\text{ub,sas}}(\jmath\omega) &= \int_0^\infty e^{\jmath\omega \log_2(1+\bar{\rho}x)} f_{\bar{\gamma}_{\max_M}}(x)dx \\ &= \frac{1}{\sigma_M}\int_0^\infty (1+\bar{\rho}x)^{\frac{\jmath\omega}{\ln(2)}} e^{-\frac{x-\mu_M}{\sigma_M}} e^{-e^{-\frac{x-\mu_M}{\sigma_M}}} dx \end{aligned} \tag{3.17}$$

Inserting $t=(x-\mu_M)/\sigma_M$, $\nu=\jmath\omega/\ln(2)$, $a=(1+\bar{\rho}\mu_M)/(\bar{\rho}\sigma_M)$, and $b=\mu_M/\sigma_M$, into Equ. (3.17) gives

$$\begin{aligned}\tilde{\Phi}_{\text{ub,sas}}(\jmath\omega) &= (\bar{\rho}\sigma_M)^\nu \int_{-b}^\infty (t+a)^\nu e^{-t} e^{-e^{-t}} dt \\ &= (\bar{\rho}\sigma_M)^\nu F_{\nu,1,a,b}, \end{aligned} \tag{3.18}$$

where the function $F_{\nu,p,a,b} = \int_{-b}^\infty (t+a)^\nu e^{-pt} e^{-e^{-t}} dt$, $a,b,p>0$. In Equ. (3.18), we have $p=1$. Note that $F_{\nu,p,a,b}$ can be evaluated efficiently by numerical integration due to the fast attenuation of $e^{-pt}e^{-e^{-t}}$.

Finally, the characteristic function of the upper capacity bound, $\Phi_{\text{ub,sas}}$, asymptotically converges to

$$\Phi_{\text{ub,sas}}(\jmath\omega) \xrightarrow{a.s.} \bar{\Phi}_{\text{ub,sas}}(\jmath\omega), \tag{3.19}$$

where $\bar{\Phi}_{\text{ub,sas}}(\jmath\omega) = \tilde{\Phi}^{L}_{\text{ub,sas}}(\jmath\omega) = (\bar{\rho}\sigma_M)^{\nu L} F^{L}_{\nu,1,a,b}$. The PDF of the asymptotic upper capacity bound is obtained by performing a Fourier transform that can be approximated by discrete Fourier transform through sampling ω. Note that the sampling rate should be large enough to avoid aliasing.

3.3.3 Upper capacity bound of the FAS system

In order to derive analytical form of the upper capacity bound in Equ. (3.6), the joint distribution of the top-L random variables, i.e., $\gamma_{(1)}, \gamma_{(2)}, \cdots, \gamma_{(L)}$, should be known. The conventional way shown in [58] is to calculate the marginal distribution from the joint distribution of the total N_{B}

random variables, i.e.,

$$\begin{aligned}&f_{\gamma_{(1)},\gamma_{(2)},\cdots,\gamma_{(L)}}(x_{(1)},x_{(2)},\cdots,x_{(L)})\\&=\int_0^{x_{(L)}}dx_{(L+1)}\int_0^{x_{(L+1)}}dx_{(L+2)}\cdots\int_0^{x_{(N_\text{B}-1)}}dx_{(N_\text{B})}\\&\times f_{\gamma_{(1)},\gamma_{(2)},\cdots,\gamma_{(N_\text{B})}}(x_{(1)},x_{(2)},\cdots,x_{(N_\text{B})}),\end{aligned}\tag{3.20}$$

where the joint PDF $f_{\gamma_{(1)},\gamma_{(2)},\cdots,\gamma_{(N_\text{B})}}(x_{(1)},x_{(2)},\cdots,x_{(N_\text{B})})$ is known as [58]

$$\begin{aligned}&f_{\gamma_{(1)},\gamma_{(2)},\cdots,\gamma_{(N_\text{B})}}(x_{(1)},x_{(2)},\cdots,x_{(N_\text{B})})=N_\text{B}!\prod_{i=1}^{N_\text{B}}f(x_{(i)})\\&=\frac{N_\text{B}!}{\Gamma(N_\text{U})^{N_\text{B}}}\prod_{i=1}^{N_\text{B}}x_{(i)}^{N_\text{U}-1}e^{-x_{(i)}},\ \text{with}\, x_{(1)}>x_{(2)}>\cdots,x_{(N_\text{B})}>0.\end{aligned}\tag{3.21}$$

This method is only favorable for the conventional MIMO systems where N_B is not too large. For the massive MIMO systems, we turn to a simpler method based on the asymptotic joint distribution of the top-L random variables. According to Lemma 3.3.3, the top-L random variables $\gamma_{(1)},\gamma_{(2)},\cdots,\gamma_{(L)}$ have the asymptotic joint distribution with the joint PDF

$$\begin{aligned}&\bar{f}_{\gamma_{(1)},\gamma_{(2)},\cdots,\gamma_{(L)}}(x_{(1)},x_{(2)},\cdots,x_{(L)})\\&=F_{\bar{\gamma}_{\text{max}_{N_r}}}(x_{(L)})\prod_{l=1}^{L}\frac{f_{\bar{\gamma}_{\text{max}_{N_r}}}(x_{(l)})}{F_{\bar{\gamma}_{\text{max}_{N_r}}}(x_{(l)})},\ x_{(1)}>x_{(2)}>\cdots>x_{(L)},\end{aligned}\tag{3.22}$$

with the asymptotic PDF

$$f_{\bar{\gamma}_{\text{max}_{N_r}}}(x)=\frac{1}{\sigma_{N_\text{B}}}e^{-\frac{x-\mu_{N_\text{B}}}{\sigma_{N_\text{B}}}}e^{-e^{-\frac{x-\mu_{N_\text{B}}}{\sigma_{N_\text{B}}}}},\tag{3.23}$$

and the asymptotic CDF

$$F_{\bar{\gamma}_{\text{max}_{N_r}}}(x)=e^{-e^{-\frac{x-\mu_{N_\text{B}}}{\sigma_{N_\text{B}}}}},\tag{3.24}$$

where μ_{N_B} and σ_{N_B} can be decided by Equ. (3.9) and (3.10) by replacing n with N_B. Inserting Equ. (3.23) and (3.24) into Equ. (3.22) gives

$$\begin{aligned}
&\bar{f}_{\gamma_{(1)},\gamma_{(2)},\cdots,\gamma_{(L)}}(x_{(1)}, x_{(2)}, \cdots, x_{(L)}) \\
&= \frac{1}{\sigma_{N_\text{B}}^L} e^{-e^{-\frac{x_{(L)}-\mu_{N_\text{B}}}{\sigma_{N_\text{B}}}}} \prod_{l=1}^{L} e^{-\frac{x_{(l)}-\mu_{N_\text{B}}}{\sigma_{N_\text{B}}}}, \\
&\text{with}\, x_{(1)} > x_{(2)} > \cdots > x_{(L)} > 0. \qquad (3.25)
\end{aligned}$$

Now the characteristic function of $C_\text{ub,fas}$, denoted by $\Phi_\text{ub,fas}(\jmath\omega)$, has the asymptotic convergence

$$\Phi_\text{ub,fas}(\jmath\omega) \xrightarrow{a.s.} \bar{\Phi}_\text{ub,fas}(\jmath\omega), \qquad (3.26)$$

where $\bar{\Phi}_\text{ub,fas}(\jmath\omega)$ is expressed by

$$\begin{aligned}
\bar{\Phi}_\text{ub,fas}(\jmath\omega) = &\int_0^\infty dx_{(L)} \int_{x_{(L)}}^\infty dx_{(L-1)} \cdots \int_{x_{(2)}}^\infty dx_{(1)} \\
&\times e^{\jmath\omega \sum_{l=1}^{L} \log_2(1+\bar{\rho}x_{(l)})} \\
&\times \frac{1}{\sigma_{N_\text{B}}^L} e^{-e^{-\frac{x_{(L)}-\mu_{N_\text{B}}}{\sigma_{N_\text{B}}}}} \prod_{l=1}^{L} e^{-\frac{x_{(l)}-\mu_{N_\text{B}}}{\sigma_{N_\text{B}}}}. \qquad (3.27)
\end{aligned}$$

Substituting $t_{(l)} = (x_{(l)} - \mu_{N_\text{B}})/\sigma_{N_\text{B}}$, $\nu = \jmath\omega/\ln(2)$, $a = (1+\bar{\rho}\mu_{N_\text{B}})/(\bar{\rho}\sigma_{N_\text{B}})$, and $b = \mu_{N_\text{B}}/\sigma_{N_\text{B}}$, into Equ. (3.27) gives

$$\begin{aligned}
\bar{\Phi}_\text{ub,fas}(\jmath\omega) = (\bar{\rho}\sigma_{N_\text{B}})^{\nu L} &\int_{-b}^\infty dt_{(L)} \int_{t_{(L)}}^\infty dt_{(L-1)} \cdots \int_{t_{(2)}}^\infty dt_{(1)} \\
&\times e^{-e^{-t_{(L)}}} \prod_{l=1}^{L} (t_{(l)} + a)^\nu e^{-t_{(l)}}. \qquad (3.28)
\end{aligned}$$

As shown in Appendix 3.7.1, the lth-fold integration, $l = 0, 1, 2, \cdots, L-1$, of Equ. (3.28) has the following form

$$q_l = \sum_{i=0}^{l} e^{-it_{(l+1)}} \sum_{n=0}^{N} c_{l,i,n} (t_{(l+1)} + a)^{n+i(\nu+1)}, \qquad (3.29)$$

where the coefficients $c_{l,i,n}$ are determined with the following recursive formula

$$c_{l,i,n} = \begin{cases} 1, & \text{if } l,i,n = 0 \\ 0, & \text{if } i = 0, l,n > 0 \\ \sum\limits_{k=1}^{l}\sum\limits_{r=0}^{N} c_{l-1,k-1,r}\frac{e^{ak}\Gamma(r+k(\nu+1))}{k^{r+k(\nu+1)}}, \text{if } i,n = 0, l > 0 \\ -\sum\limits_{m=0}^{n} c_{l-1,i-1,n-m}\frac{i^m\Gamma(n-m+i(\nu+1))}{\Gamma(n+i(\nu+1)+1)}, & \text{else.} \end{cases} \tag{3.30}$$

Note that the number of terms N in Equ. (3.29) and (3.30) is ideally infinity. However, it is observed through our simulations that $N = 120$ terms are sufficient for $N_\text{B} = 256$, $N_\text{U} = L = 4$ MIMO systems. Finally, we have $q_{L-1} = \sum_{i=0}^{L-1} e^{-it_{(L)}} \sum_{n=0}^{N} c_{L-1,i,n}(t_{(L)} + a)^{n+i(\nu+1)}$ and the asymptotic characteristics function $\bar{\Phi}_\text{ub,fas}(\jmath\omega)$ is given by the Lth-fold integration

$$\begin{aligned} \bar{\Phi}_\text{ub,fas}(\jmath\omega) &= (\bar{\rho}\sigma_{N_\text{B}})^{\nu L}\int_{-b}^{\infty} e^{-e^{-t_{(L)}}}(t_{(L)} + a)^{\nu}e^{-t_{(L)}}q_{L-1}dt_{(L)} \\ &= (\bar{\rho}\sigma_{N_\text{B}})^{\nu L}\sum_{i=0}^{L-1}\sum_{n=0}^{N} c_{L-1,i,n} \\ &\times \int_{-b}^{\infty}(t_{(L)} + a)^{n+(i+1)\nu+i}e^{-(i+1)t_{(L)}}e^{-e^{-t_{(L)}}}dt_{(L)} \\ &= (\bar{\rho}\sigma_{N_\text{B}})^{\nu L}\sum_{i=0}^{L-1}\sum_{n=0}^{N} c_{L-1,i,n}F_{n+(i+1)\nu+i,i+1,a,b}. \end{aligned} \tag{3.31}$$

3.3.4 Achievable spectral efficiency considering CSI acquisition

In this subsection, we use the above derived asymptotic upper capacity bounds to study the upper bound of the achievable spectral efficiency of both the SAS and FAS systems considering the channel training stage. A Time Division Duplex (TDD) massive MIMO system is considered, in which the transmitters (users) send training pilots to the receivers (base station) to estimate the uplink channels, and the downlink channel is obtained as well with the assumption of channel reciprocity. The frame structure is shown in Fig. 3.3, where τ_coh and τ_tr are the channel coherence time and the training duration of the full-digital system, respectively. Obviously, for the massive MIMO antenna selection system, the training duration should be M times longer than the full-digital system, to cycle N_B antennas with only L RF transceivers.

In the preceding subsections, we derive the uplink upper capacity bounds for SAS and FAS systems. With the assumption of channel reciprocity, the uplink upper capacity bounds are also

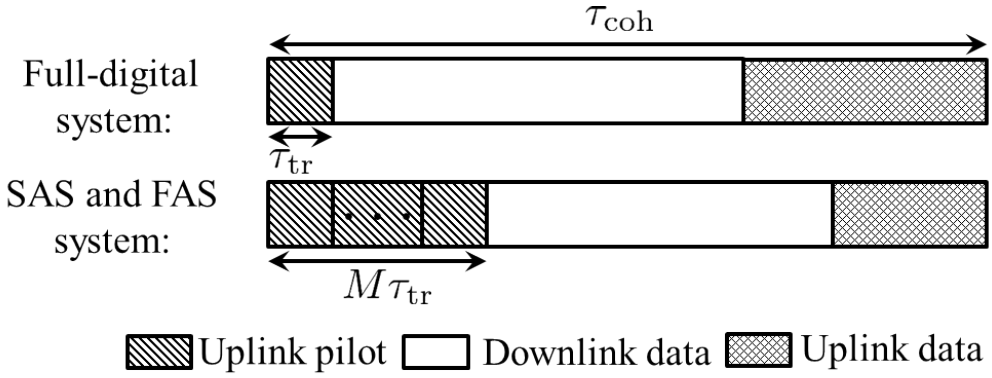


Figure 3.3: Frame structure of TDD systems

the downlink achievable spectral efficiencies when we use the equal power allocation scheme for both the uplink and downlink. Thus the upper bound of the overall achievable spectral efficiency considering the CSI acquisition is expressed as

$$\begin{aligned} R_{\text{ub,fas}|\text{sas}} &= \frac{\tau_{\text{coh}} - M\tau_{\text{tr}}}{\tau_{\text{coh}}} C_{\text{ub,fas}|\text{sas}} \\ &= (1 - M\eta) C_{\text{ub,fas}|\text{sas}}, \end{aligned} \tag{3.32}$$

where $\eta = \frac{\tau_{\text{tr}}}{\tau_{\text{coh}}} < \frac{1}{M}$. Obviously, as M grows, $C_{\text{ub,fas}|\text{sas}}$ increases, but the scaling factor $(1 - M\eta)$ decreases, which implies that there is probably an optimal value, M_{opt}, to give the largest $R_{\text{ub,fas}|\text{sas}}$. Since both $C_{\text{ub,fas}|\text{sas}}$ and $R_{\text{ub,fas}|\text{sas}}$ are random variables, we use their mean values to study M_{opt}.

Let us first study the SAS system. Recall the asymptotic characteristic function $\bar{\Phi}_{\text{ub,sas}}(\jmath\omega)$, whose first-order derivative decides the mean value of $C_{\text{ub,sas}}$, i.e.,

$$\begin{aligned} C_{\text{ub,sas,mean}} &= -\jmath \frac{d\bar{\Phi}_{\text{ub,sas}}(\jmath\omega)}{d\omega}\bigg|_{\omega=0} \\ &= \frac{1}{\log(2)} \Big((\bar{\rho}\sigma_M)^{\nu L} \log(\bar{\rho}\sigma_M) L F_{\nu,1,a,b}^{L} \\ &\quad + (\bar{\rho}\sigma_M)^{\nu L} L F_{\nu,1,a,b}^{L-1} \frac{dF_{\nu,1,a,b}}{d\nu} \Big)\bigg|_{\nu=0}. \end{aligned} \tag{3.33}$$

Define the function $G_{\nu,q,p,a,b} = \int_{-b}^{\infty} (t+a)^{\nu} (\log(t+a))^{q} e^{-pt} e^{-e^{-t}} dt$, $a, b, q, p > 0$. Then Equ. (3.33) can be written as

$$C_{\text{ub,sas,mean}} = \frac{L F_{0,1,a,b}^{L-1} \left(F_{0,1,a,b} \log(\bar{\rho}\sigma_M) - G_{0,1,1,a,b} \right)}{\log(2)}. \tag{3.34}$$

In the same way, we can derive the mean value of $C_{\text{ub,fas}}$ for the FAS system by calculating the first-order derivative of Equ. (3.31):

$$\begin{aligned} C_{\text{ub,fas,mean}} &= -\jmath \frac{d\bar{\Phi}_{\text{ub,fas}}(\jmath\omega)}{d\omega}\Big|_{\omega=0} \\ &= \frac{1}{\log(2)} \sum_{i=0}^{L-1} \sum_{n=0}^{N} \Big(c_{L-1,i,n} L \log(\bar{\rho}\sigma_{N_{\text{B}}}) F_{n+i,i+1,a,b} \\ &+ c'_{L-1,i,n}|_{\nu=0} \times F_{n+i,i+1,a,b} + (i+1) c_{L-1,i,n} G_{n+i,1,i+1,a,b} \Big), \end{aligned} \tag{3.35}$$

where $c'_{L-1,i,n}$ is the first-order derivative of the coefficients $c_{L-1,i,n}$ with respect to ν. With Equ. (3.30), we can calculate $c'_{L-1,i,n}$ in a recursive manner given in Equ. (3.36), where $\psi_0(\cdot)$ is zero-order polygamma function.

$$c'_{l,i,n} = \begin{cases} 0, \quad \text{if } i = 0 \text{ and } (l, n = 0 \text{ or } l, n > 0) \\ \displaystyle\sum_{k=1}^{l} \sum_{r=0}^{N} \Big(c'_{l-1,k-1,r} + c_{l-1,k-1,r} k \big(\psi_0(r + k(\nu+1)) \\ \qquad - \log(k)\big)\Big) \dfrac{e^{ak}\Gamma(r + k(\nu+1))}{k^{r+k(\nu+1)}}, \text{if } i, n = 0, l > 0 \\ \displaystyle -\sum_{m=0}^{n} \Big(c'_{l-1,i-1,n-m} + c_{l-1,i-1,n-m} \\ \qquad \times i(\psi_0(n - m + i(\nu+1)) - \psi_0(n + i(\nu+1) + 1)\Big) \\ \qquad \times \dfrac{i^m \Gamma(n - m + i(\nu+1))}{\Gamma(n + i(\nu+1) + 1)}, \quad \text{else.} \end{cases} \tag{3.36}$$

Finally, it has $M_{\text{opt}} = \arg\max_{M \in \mathbb{Z}^+} (1 - M\eta) C_{\text{ub,fas|sas,mean}}$, which will be studied via numerical methods in Section 3.5.

3.4 Optimal Antenna Selection Algorithm

In this section, we will first review a well-known *sub-optimal* antenna selection algorithm based on greedy search [29]. Afterwards, by using the results in the greedy-search-based antenna selection, we propose an *optimal* antenna selection algorithm based on the BAB search [55].

3.4.1 Sub-optimal antenna selection based on greedy search [29]

Let $\mathbf{H}_n$ denote the submatrix by selecting n rows from $\mathbf{H}$ after n selection steps and C_n denote the channel capacity $C_n = \log_2 \det(\mathbf{I}_{N_{\text{U}}} + \bar{\rho}\mathbf{H}_n^H \mathbf{H}_n)$, $n = 0, 1, \cdots, L-1$ (note that $C_0 = 0$).

Suppose that the jth row of $\mathbf{H}$, i.e., $\mathbf{h}_j^H$, is going to be selected from the rest candidate set in the $(n+1)$th step. The new $(n+1) \times N_\mathrm{U}$ channel submatrix is thus $\mathbf{H}_{n+1} = \begin{bmatrix} \mathbf{H}_n \\ \mathbf{h}_j^H \end{bmatrix}$. Then it has

$$\begin{aligned} C_{n+1} &= \log_2 \det \left(\mathbf{I}_{N_\mathrm{U}} + \bar{\rho}\mathbf{H}_{n+1}^H \mathbf{H}_{n+1} \right) \\ &= \log_2 \det \left(\mathbf{I}_{N_\mathrm{U}} + \bar{\rho}\mathbf{H}_n^H \mathbf{H}_n + \bar{\rho}\mathbf{h}_j \mathbf{h}_j^H \right) \\ &= C_n + \log_2 \det \left(\mathbf{I}_{N_\mathrm{U}} + \bar{\rho}(\mathbf{I}_{N_\mathrm{U}} + \bar{\rho}\mathbf{H}_n^H \mathbf{H}_n)^{-1} \mathbf{h}_j \mathbf{h}_j^H \right) \\ &\overset{(a)}{=} C_n + \underbrace{\log_2(1 + \bar{\rho}\mathbf{h}_j^H \mathbf{G}_n \mathbf{h}_j)}_{\Delta_{j,n}}, \end{aligned} \tag{3.37}$$

where the step "(a)" holds with the Sylvester's determinant identity and $\mathbf{G}_n = (\mathbf{I}_{N_\mathrm{U}} + \bar{\rho}\mathbf{H}_n^H \mathbf{H}_n)^{-1}$.

Lemma 3.4.1 *Given the channel submatrix* $\mathbf{H}_n \in \mathbb{C}^{n \times N_U}$, *the matrix* $\mathbf{G}_n = (\mathbf{I}_{N_U} + \bar{\rho}\mathbf{H}_n^H \mathbf{H}_n)^{-1}$ *is positive definite.*

Proof: First it is obvious that $\mathbf{G}_n$ is a Hermitian matrix. Then let $\mathbf{A}_n = \mathbf{H}_n^H \mathbf{H}_n = \mathbf{U}_n \Lambda_{\mathbf{A}_n} \mathbf{U}_n^H$, where $\mathbf{U}_n$ is the eigen-matrix of $\mathbf{A}_n$, and $\Lambda_{\mathbf{A}_n}$ is a diagonal matrix with the diagonal entries $\lambda_{\mathbf{A}_n,i}$, $i = 1, 2, \cdots, N_\mathrm{U}$ being the eigen-values of $\mathbf{A}_n$. Thus $\mathbf{G}_n$ can be written as $\mathbf{G}_n = \mathbf{U}_n(\mathbf{I}_{N_\mathrm{U}} + \Lambda_{\mathbf{A}_n})^{-1}\mathbf{U}_n^H = \mathbf{U}_n \Lambda_{\mathbf{G}_n} \mathbf{U}_n^H$, which is exactly the eigen-decomposition of $\mathbf{G}_n$ with eigen-values on the diagonal entries of $\Lambda_{\mathbf{G}_n}$ $\lambda_{\mathbf{G}_n,i} = 1/(1 + \lambda_{\mathbf{A}_n,i})$. Since $\mathbf{A}_n$ is a Gram matrix [94], we have $\lambda_{\mathbf{A}_n,i} \geq 0$ [94, Theorem 7.2.10]. Consequently we have $\lambda_{\mathbf{G}_n,i} > 0$, which means that $\mathbf{G}_n$ is a positive definite matrix. ∎

From Lemma 3.4.1 we know that the incremental value $\Delta_{j,n}$ given in Equ. (3.37) is a positive number when $\mathbf{h}_j \neq \mathbf{0}$, which means that the capacity function C_n is a monotonically increasing function as n increases. Thus heuristically, a greedy searching algorithm can be applied to find a sub-optimal solution. In detail, the incremental values of all of the antennas in the candidate set are examined and the maximum one to maximize the capacity function will be selected in the current step. The searching procedure goes forward until the last antenna is selected.

Note that the matrix inverse $\mathbf{G}_n = (\mathbf{I}_{N_\mathrm{U}} + \bar{\rho}\mathbf{H}_n^H \mathbf{H}_n)^{-1}$ can be updated using the Sherman-Morrison formula [94]. Suppose that the antenna index J_n was found in the nth step. Then we

Algorithm 3.1 Greedy-search-based antenna selection [29]

```
$\mathcal{S} = \emptyset$
$\mathbf{G} := \mathbf{I}_{N_{\mathrm{U}}}$
$\mathcal{J} := \{1, 2, \cdots, N_{\mathrm{B}}\}$
$\nu_j := \mathbf{h}_j^H \mathbf{h}_j, \forall j \in \mathcal{J}$                ▷ $\mathcal{O}(N_{\mathrm{U}} N_{\mathrm{B}})$
for $n = 1 : L$ do
    $J = \arg\max_{j \in \mathcal{J}} \nu_j$
    $\mathcal{S} := \mathcal{S} \cup J$
    $\mathcal{J} := \mathcal{J} \setminus J$, or                ▷ for FAS
    $\mathcal{J} := \mathcal{J} \setminus \mathcal{I}_{f(J)}^{\mathrm{SAS}}$                ▷ for SAS
    if $n < L$ then
        $\mathbf{g} = \frac{\mathbf{G}\mathbf{h}_J}{\sqrt{\bar{\rho}^{-1} + \nu_J}}$                ▷ $\mathcal{O}(N_{\mathrm{U}}^2 L)$
        $\xi_j = \mathbf{h}_j^H \mathbf{g}, \forall j \in \mathcal{J}$                ▷ $\mathcal{O}(N_{\mathrm{U}} N_{\mathrm{B}} L)$
        $\nu_j := \nu_j - |\xi_j|^2, \forall j \in \mathcal{J}$                ▷ $\mathcal{O}(N_{\mathrm{B}} L)$
        $\mathbf{G} := \mathbf{G} - \mathbf{g}\mathbf{g}^H$                ▷ $\mathcal{O}(N_{\mathrm{U}}^2 L)$
    end if
end for
return the final set $\mathcal{S}$
```

have

$$\begin{aligned}\mathbf{G}_{n+1} &= (\mathbf{I}_{N_{\mathrm{U}}} + \bar{\rho}\mathbf{H}_{n+1}^H \mathbf{H}_{n+1})^{-1} \\ &= (\mathbf{I}_{N_{\mathrm{U}}} + \bar{\rho}\mathbf{H}_n^H \mathbf{H}_n + \bar{\rho}\mathbf{h}_{J_n}\mathbf{h}_{J_n}^H)^{-1} \\ &= \mathbf{G}_n - \mathbf{g}_{n+1}\mathbf{g}_{n+1}^H,\ n = 0, 1, \cdots, L-1,\end{aligned} \tag{3.38}$$

where $\mathbf{g}_{n+1} = \frac{\mathbf{G}_n \mathbf{h}_{J_n}}{\sqrt{\bar{\rho}^{-1} + \mathbf{h}_{J_n}^H \mathbf{G}_n \mathbf{h}_{J_n}}}$. Define $\nu_{j,n+1} = \mathbf{h}_j^H \mathbf{G}_{n+1} \mathbf{h}_j$, which can be updated with the following formula:

$$\begin{aligned}\nu_{j,n+1} &= \mathbf{h}_j^H \mathbf{G}_{n+1} \mathbf{h}_j \\ &= \mathbf{h}_j^H (\mathbf{G}_n - \mathbf{g}_{n+1}\mathbf{g}_{n+1}^H)\mathbf{h}_j \\ &= \nu_{j,n} - |\xi_{J_n,j,n+1}|^2,\end{aligned} \tag{3.39}$$

where $\xi_{J_n,j,n+1} = \mathbf{h}_j^H \mathbf{g}_{n+1}$. Then the incremental values of all the candidate antenna indices are simply calculated as $\Delta_{j,n+1} = \log_2(1 + \bar{\rho}\nu_{j,n+1})$. The antenna index in the (n+1)th step is found by $J_{n+1} = \arg\max_{j \in \mathcal{J}_{n+1}} \nu_{j,n+1}$, where $\mathcal{J}_{n+1}$ is the candidate set of the antennas to be selected in the $(n+1)$th step. Note that the greedy-search-based antenna selection presented in [29] was originally for the FAS system. However, it actually applies to the SAS system by specifying the candidate set in each selection step. Specifically, for the FAS system, we have $\mathcal{J}_{n+1} = \mathcal{J}_n \setminus J_n$;

while for the SAS system, we have $\mathcal{J}_{n+1} = \mathcal{J}_n \setminus \mathcal{I}^{\mathrm{SAS}}_{f(J_n)}$, where $\mathcal{I}^{\mathrm{SAS}}_{f(J_n)}$ is the antenna set consisting of all of the antennas in the $f(J_n)$th sub-array in which the antenna J_n lies. The symbol "\" denotes set difference. The low-complexity greedy antenna selection algorithm is summarized in Alg. 3.1 with the required overall computational complexity $\mathcal{O}(\max(N_{\mathrm{U}}, N_{\mathrm{B}})N_{\mathrm{U}}L)$. Since $N_{\mathrm{B}} > N_{\mathrm{U}}$ in the system model we considered, the complexity becomes $\mathcal{O}(N_{\mathrm{B}}N_{\mathrm{U}}L)$.

3.4.2 Optimal antenna selection based on BAB search

The greedy-search-based antenna selection is fast but not optimal. The authors in [29] showed that the performance is near to the optimal selection based on exhaustive search in the small-scale FAS MIMO systems. The question is: is it still near-optimal for the massive MIMO systems and SAS MIMO systems? An optimal antenna selection algorithm is necessary to answer this question. Obviously, the exhaustive search is not feasible for the massive MIMO systems due to the tremendous searching space. In contrast to the exhaustive search, BAB is a much faster optimal searching method [55], whose higher searching efficiency comes from structuring a search tree with pruning operations. The subset selection problems satisfying the condition of monotonic relationship between subset size and selection objective function can be solved efficiently by the BAB method.

The BAB method is classified into upwards and downwards ones according to the searching direction [56]. Upwards BAB starts the searching procedure from an empty set and increases the subset size gradually until a full subset is reached. On the contrary, downwards BAB has a reverse manner in which it decreases the subset size gradually until an empty set is reached. Obviously, when only a small subset of elements is selected, upwards BAB is more attractive. Thus we focus on upwards BAB for antenna selection in this thesis.

As shown in Lemma 3.4.1, the objective function C_n is a monotonically increasing function with respect to the subset size. This means that we can only use the upwards BAB to find the minimum value rather than the maximum value of the objective function in the antenna selection problem. Instead, we need a monotonic decreasing objective function for upwards BAB to find the maximum value. In the following, we will give an equivalent objective function which is monotonically decreasing with respect to the subset size.

A new objective function

Suppose that the new objective function has the form

$$\tilde{C}_n = C_n - \sum_{m=0}^{n-1} Z_m, \tag{3.40}$$

where Z_m is defined as $Z_m = \log_2(1 + \bar{\rho}\zeta_m^2)$, and $\zeta_m = \max_{j \in \mathcal{I}_m^{\text{FAS}} \text{ or } \mathcal{I}_m^{\text{SAS}}} ||\mathbf{h}_j||_{\text{F}}$ for the full-array and sub-array selections. The sets $\mathcal{I}_m^{\text{FAS}}$ and $\mathcal{I}_m^{\text{SAS}}$ consists of all of the candidate antenna indices in the mth selection step for the full-array and sub-array selections, respectively. With the above definitions, we will prove that $\tilde{C}_n$ is a monotonically decreasing function of n.

With Equ. (3.37) and (3.40), $\tilde{C}_n$ can be written into the following recursive manner

$$\tilde{C}_{n+1} = \tilde{C}_n + \Delta_{j,n} - Z_n, \; n = 0, 1, \cdots, L-1. \tag{3.41}$$

To ensure $\tilde{C}_{n+1}$ is a monotonically decreasing function, we need $\Delta_{j,n} \le Z_n$, which means that Z_n should be an upper bound of $\Delta_{j,n}$.

Lemma 3.4.2 *Given $\Delta_{j,n} = \log_2(1 + \bar{\rho}\mathbf{h}_j^H \mathbf{G}_n \mathbf{h}_j)$ in Equ. (3.37) and the definition of Z_n. Then we have $\Delta_{j,n} \le Z_n$, $n = 0, 1, \cdots, L-1$.*

Proof: First, from Equ. (3.38) we know that

$$\mathbf{G}_n = \begin{cases} \mathbf{I}_{N_{\text{U}}}, & n = 0, \\ \mathbf{I}_{N_{\text{U}}} - \sum_{m=1}^{n} \mathbf{g}_m \mathbf{g}_m^H, & n = 1, 2, \cdots, L-1 \end{cases} \tag{3.42}$$

Case 1: $n = 0$. We simply have

$$\begin{aligned} \Delta_{j,0} &= \log_2(1 + \bar{\rho}\mathbf{h}_j^H \mathbf{h}_j) \\ &\overset{\text{(b)}}{\le} \log_2(1 + \bar{\rho}\zeta_0^2) = Z_0, \end{aligned} \tag{3.43}$$

where the inequality "(b)" holds when the candidate antenna index $j \in \mathcal{I}_0^{\text{FAS}}$, or $\mathcal{I}_0^{\text{SAS}}$ in the 0th selection step.

Case 2: $n = 1, 2, \cdots, L-1$. Let us define $\bar{\mathbf{G}}_n = \sum_{m=1}^{n} \mathbf{g}_m \mathbf{g}_m^H$. Since the Gram matrix $\mathbf{g}_m \mathbf{g}_m^H$ is positive semi-definite, the sum of positive semi-definite matrices, i.e., $\bar{\mathbf{G}}_n$, is also a positive semi-definite matrix [94, Observation 7.1.3]. Then it follows that

$$\mathbf{h}_j^H \mathbf{G}_n \mathbf{h}_j - \mathbf{h}_j^H \mathbf{h}_j = -\mathbf{h}_j^H \bar{\mathbf{G}}_n \mathbf{h}_j \leq 0. \tag{3.44}$$

Thus we have

$$\begin{aligned}\Delta_{j,n} &= \log_2(1 + \bar{\rho}\mathbf{h}_j^H \mathbf{G}_n \mathbf{h}_j) \\ &\leq \log_2(1 + \bar{\rho}\mathbf{h}_j^H \mathbf{h}_j) \\ &\overset{\text{(c)}}{\leq} \log_2(1 + \bar{\rho}\zeta_m^2) = Z_n,\end{aligned} \tag{3.45}$$

where the inequality "(c)" holds when the candidate antenna index $j \in \mathcal{I}_n^{\text{FAS}}$ or $\mathcal{I}_n^{\text{SAS}}$ in the nth selection step. ■

Since the new objective function $\tilde{C}_n$ is a monotonically decreasing function with respect to n, we can use the upwards BAB search method to find the antenna subset that gives the largest value of $\tilde{C}_L$ with branching and bounding procedure efficiently. Note that the new objective function in the last selection step $\tilde{C}_L$ equals the original objective function C_L minus the offset $\sum_{m=0}^{L-1} Z_m$ according to Equ. (3.40). Because the offset $\sum_{m=0}^{L-1} Z_m$ is a fixed number in each channel realization, and irrelevant to the selected antenna subset, the antenna subset leading to largest value of $\tilde{C}_L$ also leads to the largest value of C_L. Thus the BAB algorithm with the new objective function finds the largest value of C_L.

BAB searching procedure

Now let us show how to use upwards BAB to find the optimal antenna subset. As an illustration, an example of search trees by selecting 2 out of 6 antennas for the full-array and sub-array selection is shown in Figure 3.4(a) and 3.4(b) respectively. For both the full-array and sub-array antenna selections, the search trees have $L + 1 = 3$ levels to select a subset of size $L = 2$. In each level, an antenna node is selected until we reach the leaf nodes. The searching procedure is thus to find the path from the root node to the leaf node which has the minimum $\tilde{C}_L$ in the tree.

The search trees are different between full-array and sub-array antenna selections. For the full-array antenna selection we have an asymmetric tree structure shown in Figure 3.4(a). Suppose that the parent node J_n in the nth level has the child-node set $\mathcal{I}_{n,J_n}^{\text{FAS}}$ in the $(n+1)$th level. For example, in Figure 3.4(a) the parent node $J_1 = 3$ in level 1 has the child-node set

$\mathcal{I}_{1,3}^{\text{FAS}} = \{4, 5, 6\}$. Thus the candidate set consisting all of the candidates in the $(n+1)$th level is $\mathcal{I}_{n}^{\text{FAS}} = \bigcup_{J_n \in \mathcal{I}_{n-1}^{\text{FAS}}} \mathcal{I}_{n,J_n}^{\text{FAS}}$. With the labeling scheme shown in Figure 3.4(a), we simply have $\mathcal{I}_{n,J_n}^{\text{FAS}} = \{J_n+1, J_n+2, \cdots, N_{\text{B}}+n-L+1\}$ and $\mathcal{I}_{n}^{\text{FAS}} = \{n+1, n+2, \cdots, N_{\text{B}}+n-L+1\}$. In contrast to the asymmetric tree of the full-array selection, a symmetric tree shown in Figure 3.4(b) is constructed for the sub-array antenna selection, where each parent node in the nth level generates equal number of child nodes, i.e., M, in the $(n+1)$th level until the Lth level is reached. Note that in the sub-array search tree, all the parent nodes in the same level have the same child-node set. With the labeling scheme shown in Figure 3.4(b), we simply have $\mathcal{I}_{n}^{\text{SAS}} = \{Mn+1, Mn+2, \cdots, M(n+1)\}$.

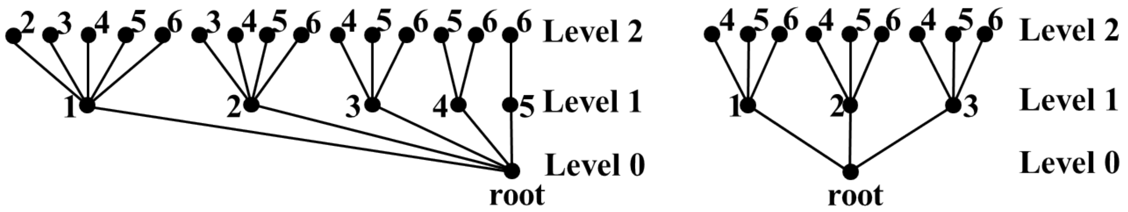
With the aforementioned search trees, the BAB search algorithm will explore the trees with branching and pruning procedures as follows. The search procedure will start from the root node of an empty set. One antenna element will be added from nth to $(n+1)$th tree level. If the cost value $\tilde{C}_{n+1}$ of the currently visited node is less than the global lower bound B, according to the decreasing monotonicity of the cost function $\tilde{C}_n$, all its child nodes have less cost than the global lower bound, and thus all these child nodes will not be necessarily evaluated and will consequently be pruned (pruning procedure); otherwise the child nodes will be explored from the best one which has the largest $\tilde{C}_{n+1}$ to the least one (branching with the best-first strategy). When all of the subtree nodes of the currently visited node are either evaluated or pruned, the searching will go back to its parent nodes to evaluate other subtrees. This searching procedure will be repeated until all the nodes are either evaluated or pruned, thus giving a global optimal solution. It should be noted that the global lower bound B used for the pruning procedure is unknown in advance. Instead, we can first initialize it as $B = -\infty$, and then update B when leaf nodes are reached and evaluated in the searching procedure. Specifically, if the cost value $\tilde{C}_L$ of the currently visited leaf node is larger than the current global lower bound, the new cost value $\tilde{C}_L$ will be the new global lower bound; otherwise B is not changed. The optimal antenna selection algorithms based on the described BAB searching algorithm are summarized in Alg. 3.2 for the FAS and SAS systems.

Complexity of the BAB based antenna selection

As shown in Alg. 3.2, the overall complexity of $\mathcal{O}(N_{\text{U}}N_{\text{B}}N_{\text{nodes}})$ is required, where N_{nodes} denotes the number of non-leaf nodes that will be branched in the search tree, which requires updating calculations shown from line 20 to 24 in Alg. 3.2. Note that since the number of nodes to be branched in the search tree depends on channel realization like the depth-first sphere

Algorithm 3.2 BAB-search-based antenna selection

```
 1: Initialization: $\mathbf{G} = \mathbf{I}_{N_{\text{U}}}$, $B = -\infty$, $\tilde{C} = N_{\text{U}}$, $n = 0$, $J = 0$, and the selected antenna index vector $\mathbf{s} = \mathbf{0}_L$, $\mathcal{J} = \{1, 2, \cdots, N_{\text{B}}\}$, $\mathcal{L} = \{1, 2, \cdots, L\}$
 2: $\nu_j = ||\mathbf{h}_j||_{\text{F}}^2, \forall j \in \mathcal{J}$    ▷ $\mathcal{O}(N_{\text{U}}N_{\text{B}})$
 3: $\zeta_m = \max_{j \in \mathcal{I}_m^{\text{FAS}} \text{ or } \mathcal{I}_m^{\text{SAS}}} \nu_j, \forall m \in \mathcal{L}$
 4: $Z_m = \log_2(\zeta_m), \forall m \in \mathcal{L}$    ▷ $\mathcal{O}(L)$
 5: if $n = L - 1$ then
 6:     $c_j := \tilde{C} + \Delta_j - Z_{L-1}, \forall j \in \mathcal{I}_{L-1,J}^{\text{FAS}}$ or $\mathcal{I}_{L-1}^{\text{SAS}}$
 7:     if $\min_{m \in \mathcal{I}_{L-1,J}^{\text{FAS}} \text{ or } \mathcal{I}_{L-1}^{\text{SAS}}} c_m < B$ then    ▷ Updating bound
 8:         update $[\mathbf{s}]_L = \arg\max_{m \in \mathcal{I}_{L-1,J}^{\text{FAS}} \text{ or } \mathcal{I}_{L-1}^{\text{SAS}}} c_m$, $B := \max_{m \in \mathcal{I}_{L-1,J}^{\text{FAS}} \text{ or } \mathcal{I}_{L-1}^{\text{SAS}}} c_m$, and $\bar{\mathbf{s}} := \mathbf{s}$
 9:     end if
10: else
11:     $c_j := \tilde{C} + \Delta_j - Z_n, \forall j \in \mathcal{I}_{n,J}^{\text{FAS}}$ or $\mathcal{I}_n^{\text{SAS}}$
12:     sort $c_j, \forall j \in \mathcal{I}_{n,J}^{\text{FAS}}$ or $\mathcal{I}_n^{\text{SAS}}$ in a descend order to get an ordered index vector $\mathbf{j}$
13:     $\mathbf{G}_{\text{tmp}} := \mathbf{G}$, $\nu_{\text{tmp},j} := \nu_j, \forall j \in \mathcal{J}$
14:     for $i$ from 1 to $|\mathcal{I}_{n,J}^{\text{FAS}}|$ or $M$ do
15:         $J = [\mathbf{j}]_i$
16:         if $c_J > B$ then    ▷ Branching
17:             $\mathcal{K} := \{J + 1, J + 2, \cdots, N_{\text{B}}\}$, or    ▷ for FAS
18:             $\mathcal{K} := \mathcal{I}_{n+1}^{\text{SAS}} \cup \mathcal{I}_{n+2}^{\text{SAS}} \cup \cdots \cup \mathcal{I}_{L-1}^{\text{SAS}}$    ▷ for SAS
19:             update the index vector $[\mathbf{s}]_{n+1} = J$
20:             $\mathbf{g} := \frac{1}{\sqrt{\bar{\rho}^{-1} + \nu_{\text{tmp},J}}} \mathbf{G}\mathbf{h}_J$, and $\mathbf{G} := \mathbf{G}_{\text{tmp}} - \mathbf{g}\mathbf{g}^H$    ▷ $\mathcal{O}(N_{\text{U}}^2 N_{\text{nodes}})$
21:             $\tilde{C} := c_J$
22:             $\xi_m := \mathbf{h}_m^H \mathbf{g}, \forall m \in \mathcal{K}$,    ▷ $\mathcal{O}(N_{\text{U}}N_{\text{B}}N_{\text{nodes}})$
23:             $\nu_m := \nu_{\text{tmp},m} - |\xi_m|^2, \forall m \in \mathcal{K}$    ▷ $\mathcal{O}(N_{\text{B}}N_{\text{nodes}})$
24:             $\Delta_m := \log_2(1 + \bar{\rho}\nu_m), \forall m \in \mathcal{K}$    ▷ $\mathcal{O}(N_{\text{U}}N_{\text{B}}N_{\text{nodes}})$
25:             $n := n + 1$, go to line 6
26:         else    ▷ Pruning
27:             break the loop
28:         end if
29:     end for
30: end if
31: return the final antenna index vector $\bar{\mathbf{s}}$
```



(a) Search tree of FAS (b) Search tree of SAS

Figure 3.4: Examples of search trees to select 2 out of 6 antennas, (a) for FAS systems, (b) for SAS systems ($M = 6/2 = 3$). The label besides each node denotes the index of the selected antenna by this node. A path from the root to a leaf node corresponds to one of $\binom{N_{\mathrm{B}}}{L}$ and $(\frac{N_{\mathrm{B}}}{L})^L$ selected antenna subsets for the FAS and SAS systems, respectively.

decoding in MIMO detection, the run-time of the optimal BAB-search-based antenna selection algorithm is unpredictable, which will be studied through simulations in Section 3.5.

3.4.3 Comparison between greedy search and BAB search

In this section, we have proposed the BAB-based antenna selection algorithm from the known greedy-search-based antenna selection. It is interesting to note that the greedy search is actually a special case of the BAB search when only the first path is explored in the searching tree. Since the optimal solution cannot be guaranteed in the first path, the greedy search has sub-optimal performance compared to the optimal performance of the BAB search. For the exhaustive search, it is not necessary to directly use the high-complexity matrix inversion to calculate the capacity value of each antenna subset, but use the searching tree shown in Fig. 3.4 and the updating method given by Equ. (3.41) to calculate the cost value. The difference between the BAB searching and the exhaustive search is that the exhaustive search has no pruning operations so that all the intermediate nodes in the tree are involved with the updating operations.

By using the searching tree shown in Fig. 3.4, it is also fair to compare their computational complexities, namely, we can just compare the number of visited nodes asking for updating operations in the tree. Obviously, the number of visited nodes asking for updating operations is fixed for both exhaustive search and greedy search. For the exhaustive search, we have $N^{\mathrm{FAS}}_{\mathrm{nodes,es}} = \sum_{n=1}^{L-1} \left(\binom{N_{\mathrm{B}}+n-L}{n}\right)$ and $N^{\mathrm{SAS}}_{\mathrm{nodes,es}} = \sum_{n=1}^{L-1} M^n$ for the full-array and sub-array selection, respectively. For the greedy search, the number is fixed to be L, no matter with the full-array or sub-array architectures. In contrast to the fixed run-time of the greedy and exhaustive search, the BAB search has varying run-time, because the number of nodes to be branched in

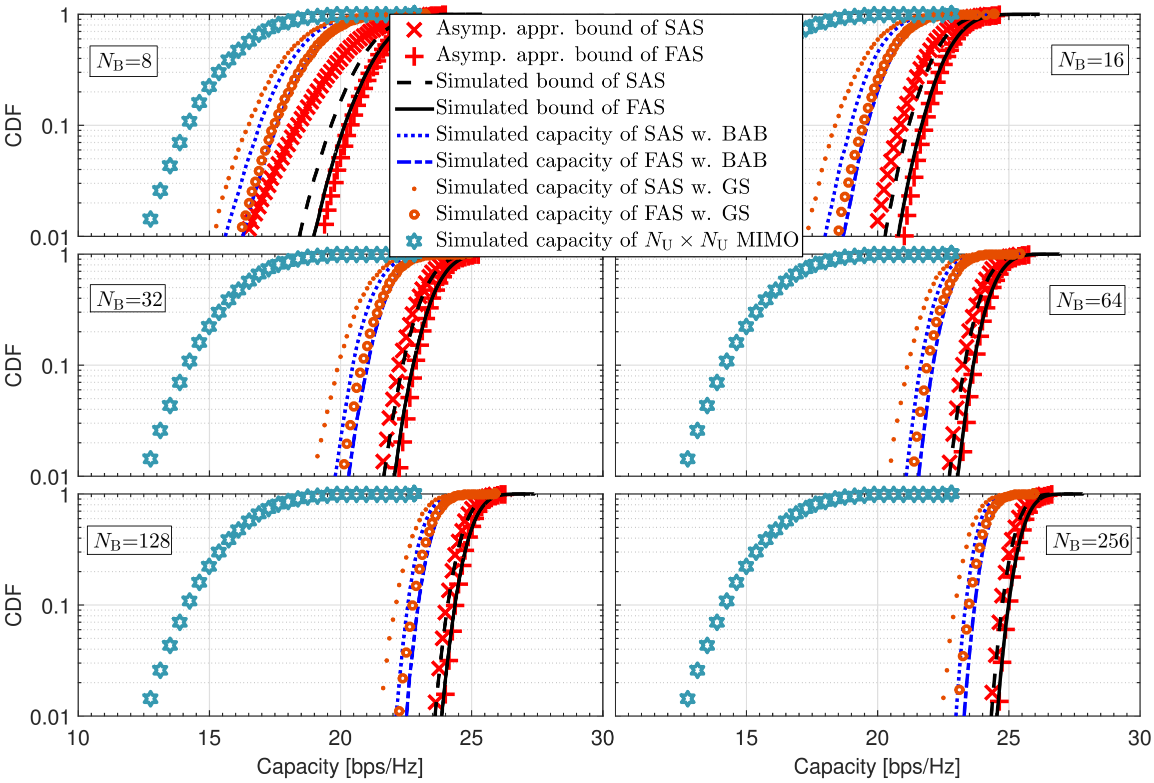


Figure 3.5: CDF of the asymptotic approximated upper capacity bounds, exact upper capacity bounds and exact capacities with BAB search and Greedy Search (GS) for SAS and FAS systems with $L = N_U = 4$ and $\rho = 15$ dB.

the searching tree, N_{nodes}, depends on channel realization like the depth-first sphere decoding in MIMO detection. Obviously, the greedy search is preferred for the practical implementation; while the BAB search is suitable for benchmark purpose, as an alternative of exhaustive search. Their performance and complexity comparison will be given in Section 3.5.

3.5 Results

In this section, we give the numerical and simulation results derived in the preceding sections. We assume that the channel matrix $\mathbf{H}$ is estimated perfectly by the receiver. In Section 3.3, we derived the asymptotic upper capacity bounds for both FAS and SAS systems, whose approximation to the exact upper bounds is of great interest especially when N_B is not so large. Since the analytic form of the exact upper bounds is unknown, we use Monte-Carlo simulations with a large number of trials to approach the exact upper bounds. In each trial we randomly generate an i.i.d. $\mathbf{H}$, and calculate the upper capacity bounds with Equ. (3.6) and (3.7). It was observed that 5×10^4 trials are sufficient for the N_B up to 256 of our interest. In the same way, we can get good approximations of the exact capacities of the SAS and FAS systems with the BAB and greedy search antenna selection algorithms, which will be used to examine the tightness of the derived upper capacity bounds.

Let us fix $L = N_{\mathrm{U}} = 4$ and increase N_{B} from 8 to 256. Note that $M = N_{\mathrm{B}}/L$ will also increase as N_{B}. Fig. 3.5 shows the CDF of the derived asymptotic upper capacity bounds, the simulated upper capacity bounds, and the simulated capacities with BAB and greedy search algorithms, for both SAS and FAS systems. As a benchmark purpose we also simulate the small-scale $N_{\mathrm{U}} \times N_{\mathrm{U}}$ MIMO systems which uses the same number of RF transceivers as the antenna selection systems to examine the advantages of the excessive antennas in the antenna selection systems. It can be observed that the larger the N_{B}, the more capacity gain the antenna selection systems can achieve over the small-scale $N_{\mathrm{U}} \times N_{\mathrm{U}}$ MIMO system. In addition, the distributions of the capacities of the antenna systems become more concentrated as N_{B} increases, which is similar to the channel hardening effect [42]. Most importantly, we can see that the larger the N_{B}, the better agreement between the derived asymptotic bounds and the simulated bounds for both SAS and FAS systems. For the FAS system the derived bound agrees well with the simulated bound starting from $N_{\mathrm{B}} = 8$; while for the SAS system it starts from $N_{\mathrm{B}} = 16$, which implies that at least 4 samples are necessary to use the Gumbel distribution to approximate the maxima of chi-squared samples. When $N_{\mathrm{B}} \geq 32$, the mismatch between the derived and simulated bounds is negligible for both SAS and FAS systems. Since the mismatches in the cases of $N_{\mathrm{B}} = 8, 16$ are not very large, the derived upper capacity bounds are believed to be still usable when N_{B} is not too large. Another important observation is that the derived upper capacity bounds are tight compared to the simulated capacities with the optimal selection algorithms. For example, with $N_{\mathrm{B}} = 32$, their capacity gap is only about 1.8 bps/Hz at the CDF $= 0.1$, i.e., 10% outage capacity, for both the SAS and FAS systems. This tightness can also be observed when increasing N_{B} further. Next, let us compare the performance of the antenna selection algorithms. For the FAS systems, the sub-optimal greedy algorithm has very close performance to the optimal BAB algorithm from $N_{\mathrm{B}} = 8$ to 256; while for the SAS systems, they have a noticeable performance gap, e.g., around 0.5 bps/Hz at CDF=0.1. Thus we can conclude that the conventional greedy antenna selection still works well in massive MIMO. At last, let us compare the simulated capacities between the SAS and FAS systems. It is clear to see that the performance of the SAS and FAS systems is very close to each other from $N_{\mathrm{B}} = 8$ to 256. Interestingly, this closeness can also be observed by comparing their asymptotic upper capacity bounds. Thus it can be foreseen that the SAS system is more favorable in practice due to its

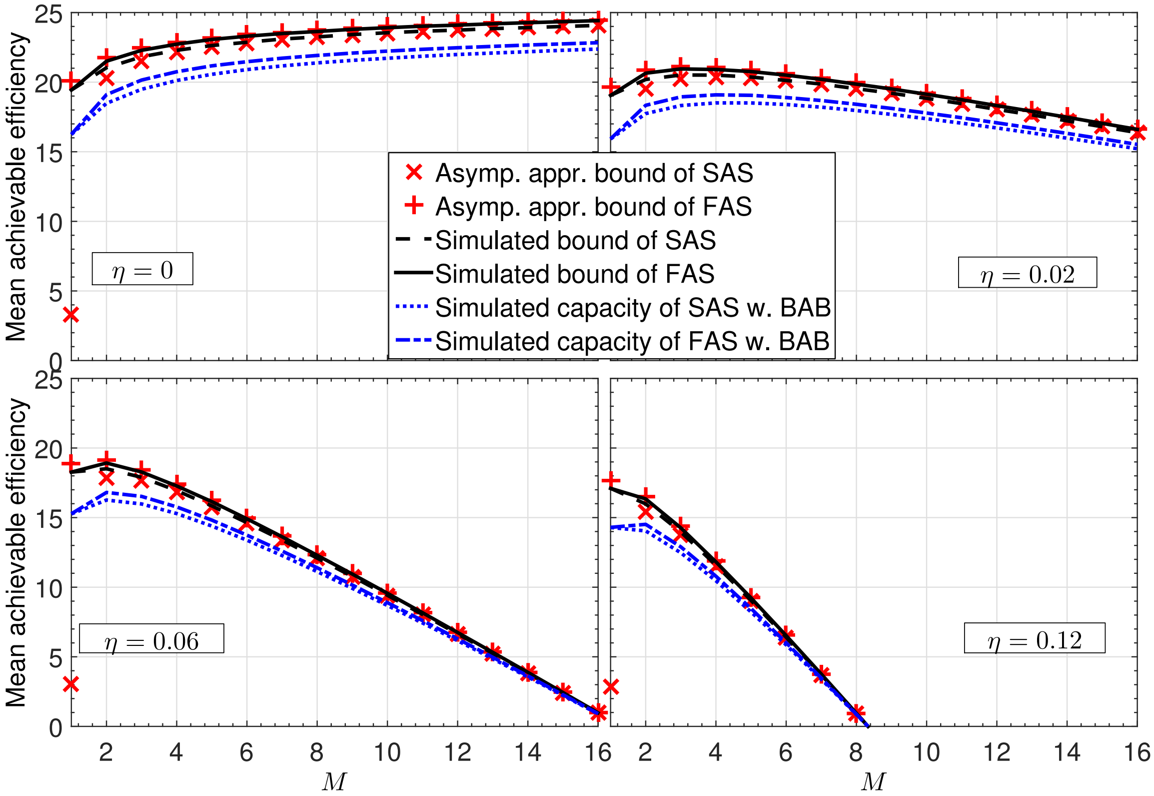


Figure 3.6: Achievable spectral efficiency considering channel training, $L = N_{\rm U} = 4$ and $\rho = 15$ dB.

greatly simplified RF switching network.

Now let us turn to the impact of the CSI acquisition on the mean achievable spectral efficiency. Equ. (3.34) and (3.35) are used to upper-bound the mean achievable spectral efficiency. The simulated bound and the simulated mean achievable spectral efficiencies are obtained similarly to the preceding subsection via Monte-Carlo simulations. The results of the mean achievable spectral efficiencies with respect to M are shown in Fig. 3.6 with various η values. We can see that the derived bounds agree well with the simulated bounds except the cases of $M = 1$ for the SAS system. The reason of this large error is due to the fact that the Gumbel approximation only applies when there are multiple samples. In addition, we can see that the derived bounds are very informative, since they exhibit very similar behaviours to the simulated achievable spectral efficiency for both the SAS and FAS systems when $M > 1$. Let us go further with the behaviours of the mean achievable spectral efficiency with respect to M with different η values. When $\eta = 0$, which indicates the ideal case that the channel training duration is negligible compared to the channel coherence time, e.g., in quasi-static channels, the mean achievable spectral efficiency increases monotonically as M increases. In contrast, when

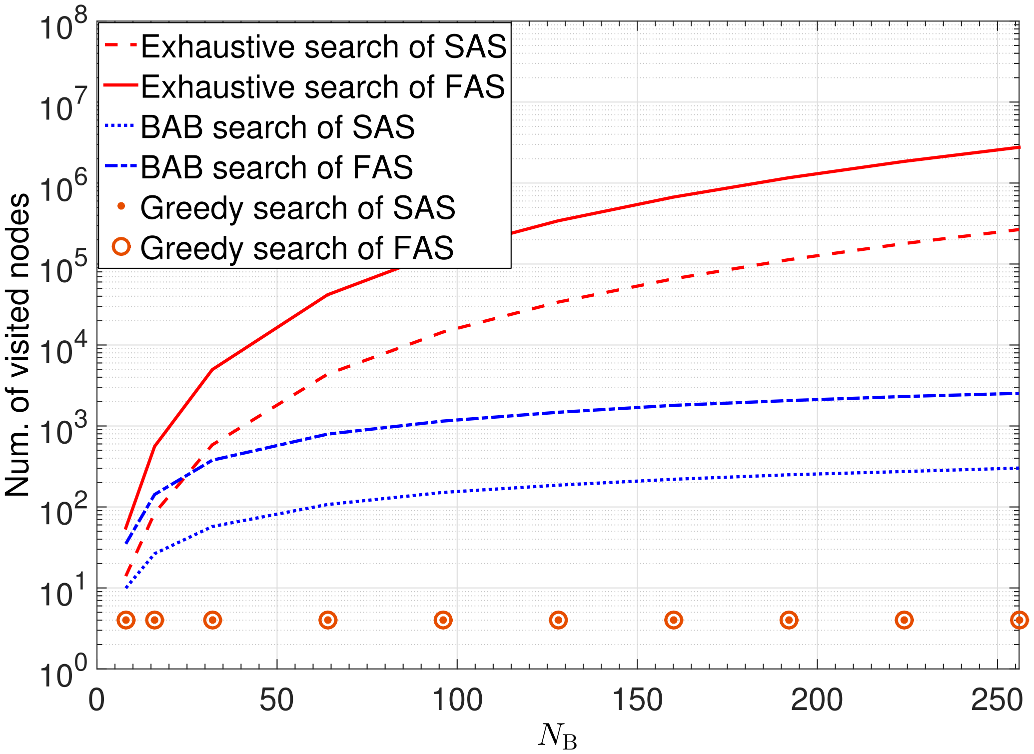


Figure 3.7: Complexity comparison, $L = N_\mathrm{U} = 4$ and $\rho = 15$ dB.

$\eta = 0.02$ or 0.06, the mean achievable spectral efficiencies first increase and then decrease as M increases. Obviously, there exists an M_opt, which is 3 and 2 for $\eta = 0.02$ and 0.06, respectively. The larger the η, the less the M_opt. When η is large enough, for example, $\eta = 0.12$, M_opt will be 1, which means that excessive antennas over RF transceivers will even lead to spectral efficiency loss due to the cost of the CSI acquisition. Thus we can conclude that the CSI acquisition is one of the main limits in the massive MIMO antenna selection systems with time-variant channels, and more advanced CSI acquisition schemes are necessary. For example, we can utilize the antenna correlation of practical limited-scattering channels to reduce the training cycles, e.g., only a subset of antennas are trained and the remaining channels are estimated via spatial interpolation.

At last let us study the computational complexity of the BAB-based antenna selection algorithms. As a benchmark purpose, the complexities of the exhaustive search and greedy search are also studied. For fair comparison, we define the complexity as the number of visited nodes asking for updating operations in the search tree. Obviously, for the exhaustive search, all of nodes should be evaluated due to the lack of pruning operations; while for the greedy search, only L nodes are evaluated. Fig. 3.7 shows the results. We can see that the BAB search has a complexity in-between the greedy search and exhaustive search. For both exhaustive search and BAB search, the SAS system has lower complexity than the FAS system, due to the less number of candidate antenna subsets. Most interestingly, the increasing rate of the BAB search is much slower than that of the exhaustive search. The larger the N_{B}, the more advantageous the BAB. For example, with $N_{\mathrm{B}} = 256$, the complexity of the BAB search is reduced by 3 orders of magnitude compared to the exhaustive search for both the FAS and SAS systems.

3.6 Conclusions

In this chapter, we derive the tight asymptotic upper capacity bounds for the SAS and FAS systems in the limit of large-scale MIMO systems, which also applies to the finite-dimensional MIMO systems. We use the derived bounds to study the achievable spectral efficiency considering the CSI acquisition. The results show that the CSI acquisition is one of the main limit in the massive MIMO antenna selection systems with time-variant channels. In addition, the proposed optimal antenna selection based on BAB is much more efficient than the exhaustive search, serving as a benchmark in the future work on the massive MIMO antenna selection algorithm design. At last, the SAS system has very close performance to the FAS system, which is validated by their exact capacities and their upper capacity bounds.

3.7 Appendix

3.7.1 Derivation of q_l and $c_{l,i,n}$

Let us start with the first-fold integral in Equ. (3.28), i.e., $q_1 = \int_{t_{(2)}}^{\infty} (t_{(1)} + a)^{\nu} e^{-t_{(1)}} dt_{(1)}$. Define a function $f(\nu, u) = \int_{u}^{\infty} (x + a)^{\nu} e^{-\beta x} dx$. Substituting $y = \beta(x + a)$ gives

$$\begin{aligned} f(\nu, u) &= \frac{e^{a\beta}}{\beta^{\nu+1}} \int_{\beta(u+a)}^{\infty} y^{\nu} e^{-y} dy \\ &= \frac{e^{a\beta}}{\beta^{\nu+1}} \Gamma(\nu + 1, \beta(u + a)) \\ &= \frac{e^{a\beta}}{\beta^{\nu+1}} (\Gamma(\nu + 1) - \gamma(\nu + 1, \beta(u + a))), \end{aligned} \tag{3.46}$$

where $\Gamma(\cdot,\cdot)$ is the upper incomplete gamma function [48]. Note that the lower incomplete Gamma function $\gamma(a,z)$ has the power series [52]

$$\gamma(a,z)=z^{a}\Gamma(a)e^{-z}\sum_{n=0}^{\infty}\frac{z^{n}}{\Gamma(a+n+1)}. \tag{3.47}$$

With Equ. (3.46) and (3.47) we can rewrite q_1 as

$$\begin{aligned}q_1&=\int_{t_{(2)}}^{\infty}(t_{(1)}+a)^{\nu}e^{-t_{(1)}}dt_{(1)}\\&=e^{a}\Gamma(\nu+1)-e^{-t_{(2)}}\sum_{n=0}^{\infty}\frac{\Gamma(\nu+1)}{\Gamma(n+\nu+2)}(t_{(2)}+a)^{n+\nu+1}\end{aligned} \tag{3.48}$$

In the same way we can derive q_2 as

$$\begin{aligned}q_2&=\int_{t_{(3)}}^{\infty}(t_{(2)}+a)^{\nu}e^{-t_{(2)}}q_1dt_{(2)}\\&=e^{a}\Gamma(\nu+1)\int_{t_{(3)}}^{\infty}(t_{(2)}+a)^{\nu}e^{-t_{(2)}}dt_{(2)}\\&-\sum_{n=0}^{\infty}\int_{t_{(3)}}^{\infty}e^{-2t_{(2)}}\frac{\Gamma(\nu+1)}{\Gamma(n+\nu+2)}(t_{(2)}+a)^{n+2\nu+1}dt_{(2)}\\&=(e^{a}\Gamma(\nu+1))^{2}+\sum_{n=0}^{\infty}\frac{e^{2a}\Gamma(\nu+1)\Gamma(n+2\nu+2)}{2^{n+2\nu+2}\Gamma(n+\nu+2)}\\&-e^{-t_{(2)}}\sum_{n=0}^{\infty}\frac{e^{a}\Gamma^{2}(\nu+1)}{\Gamma(n+\nu+1)}(t_{(2)}+a)^{n+\nu+1}\\&-e^{-2t_{(2)}}\underbrace{\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{2^{m}\Gamma(\nu+1)\Gamma(n+2\nu+2)}{\Gamma(m+n+\nu+3)}(t_{(2)}+a)^{m+n+2\nu+2}}_{C}.\end{aligned} \tag{3.49}$$

Substituting $r=m+n$ into the Term C in Equ. (3.49), we have $C=\sum_{r=0}^{\infty}\sum_{m=0}^{r}\frac{2^{m}\Gamma(\nu+1)\Gamma(r-m+2\nu+2)}{\Gamma(r+\nu+3)}(t_{(2)}+a)^{r+2\nu+2}$. By the induction from Equ. (3.48) and (3.49), we can generalize the result of the lth-fold integration as

$$q_l=\sum_{i=0}^{l}e^{-it_{(l+1)}}\sum_{n=0}^{\infty}c_{l,i,n}(t_{(l+1)}+a)^{n+i(\nu+1)}. \tag{3.50}$$

Following this form, we have

$$q_{l-1} = \sum_{i=0}^{l-1} e^{-it_{(l)}} \sum_{n=0}^{\infty} c_{l-1,i,n}(t_{(l)} + a)^{n+i(\nu+1)}. \tag{3.51}$$

Inserting Equ. (3.51) into $q_l = \int_{t_{(l+1)}}^{\infty} (t_{(l)} + a)^{\nu} e^{-t_{(l)}} q_{l-1} dt_{(l)}$, we have Equ. (3.52), where the step "(a)" holds with $k = i + 1, r = m + n$. Comparing the coefficients between Equ. (3.50) and (3.52), we can get $c_{l,i,n}$ in Equ. (3.30).

$$\begin{aligned}
q_l &= \sum_{i=0}^{l-1}\sum_{n=0}^{\infty} c_{l-1,i,n} \int_{t_{(l+1)}}^{\infty} (t_{(l)} + a)^{n+(i+1)(\nu+1)-1} e^{-(i+1)t_{(l)}} dt_{(l)} \\
&= \sum_{i=0}^{l-1}\sum_{n=0}^{\infty} \frac{c_{l-1,i,n} e^{a(i+1)}}{(i+1)^{n+(i+1)(\nu+1)}} \Gamma\left(n + (i+1)(\nu+1), (i+1)(t_{(l+1)} + a)\right) \\
&= \sum_{i=0}^{l-1}\sum_{n=0}^{\infty} \frac{c_{l-1,i,n} e^{a(i+1)} \Gamma(n + (i+1)(\nu+1))}{(i+1)^{n+(i+1)(\nu+1)}} \\
&\quad - \sum_{i=0}^{l-1} e^{-(i+1)t_{(l+1)}} \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{c_{l-1,i,n}(i+1)^m \Gamma(n + (i+1)(\nu+1))}{\Gamma(m + n + (i+1)(\nu+1) + 1)} (t_{(l+1)} + a)^{m+n+(i+1)(\nu+1)} \\
&\overset{(a)}{=} \sum_{k=1}^{l}\sum_{n=0}^{\infty} \frac{c_{l-1,k-1,n} e^{ak} \Gamma(n + k(\nu+1))}{k^{n+k(\nu+1)}} \\
&\quad - \sum_{k=1}^{l} e^{-kt_{(l+1)}} \sum_{r=0}^{\infty}\sum_{m=0}^{r} \frac{c_{l-1,k-1,r-m} k^m \Gamma(r - m + k(\nu+1))}{\Gamma(r + k(\nu+1) + 1)} (t_{(l+1)} + a)^{r+k(\nu+1)}
\end{aligned} \tag{3.52}$$

Chapter 4

Hybrid Analog-Digital Beamforming Massive MIMO Based on Rotman Lens

Similar to the antenna selection in Chapter 3, hybrid analog-digital beamforming is another well-known low RF-complexity signal processing method for massive MIMO systems. In this chapter, we propose a low RF-complexity hybrid beamforming massive MIMO system based on Rotman lens analog beamforming with beam selection and digital beamforming. Rotman lens is a low-RF-complexity low-cost true-time-delay analog beamforming network supporting wide bandwidth signals, which is thus more attractive than the analog beamforming networks based on conventional high-cost phase shifters. To study our Rotman-lens-based system, two potential array architectures, i.e., full-array and sub-array architectures, are first examined, concerning the RF design feasibility, insertion loss and system scalability. Since the beam selection is required in our system, we propose two beam selection algorithms, i.e., greedy search based method and branch and bound based method, aiming to optimize the hybrid beamforming system in terms of error-rate performance for both the full-array and sub-array architectures. To validate our system in practice, the proposed system is also investigated and verified experimentally. In particular, we design and fabricate a sample of Rotman lens operating in the 5-GHz band, whose measured results agrees well with the computer simulation results. The measurement results are incorporated into the Monte-Carlo simulation with the proposed beam selection algorithms to study the error-rate performance. The results in this chapter show that our low RF-complexity hybrid beamforming system based on the low-cost Rotman lens has comparable performance to that of the system using the high-cost phase shifters, and exhibits wideband capability and superior performance over the small-scale MIMO system under the same number of RF transceivers.

4.1 Introduction

The hybrid analog-digital beamforming was initially studied as the soft antenna selection for general MIMO systems, e.g., [58] - [60]. The hybrid analog-digital beamforming particularly with massive number of antennas was pioneered by the work [61] for the millimeter wave (mmWave) systems. Substantial amount of successive works can be found in the literature, e.g., [62] - [74]. By summarizing the works [58]- [74], we can find that the hybrid analog-digital beamforming can be categorized into two types according to the beamspace used in the analog beamforming, i.e. static beamspace, such as the works [58], [61] - [64], [68], and dynamic beamspace, such as the works [59] - [60], [65] - [67], [69] - [74]. We refer to the beamspace given by the fixed phase RF network as static beamspace, e.g., the beamspace spanned by the steering vectors of antenna arrays; while we refer to the beamspace given by the adaptive phase RF network as dynamic beamspace which expands the fixed beamspace by tuning the phase shifters to an arbitrary optimal beamspace dynamically. The analog beamforming with the static beamspace is also referred to as beam steering and the associated MIMO is also called beamspace MIMO [75] - [77].

Although the dynamic beamspace system has better performance than the static beamspace system because of more freedoms provided by the dynamic beamspace [80], in this chapter, we focus on the static beamspace spanned by the steering vectors of the antenna arrays, rather than the dynamic beamspace. The reason is three-fold: firstly it is more feasible for practical implementation due to its fixed nature; secondly it is more plausible for the wideband system since the beamspace only depends on the antenna array instead of the channel realizations; thirdly the static beamspace system can be realized with simple passive RF structures, such as lens antenna, which, in some scenarios, is more energy-efficient than the dynamic system usually realized by the active tunable phase shifters that require power to excite (an example study is shown in [80]). To realize the beamspace in the analog beamforming, the RF Phase-Shifting Network (PSN) is critical. The PSN can be realized either by a network of discrete phase shifters together with power dividers and combiners, or simply by a time-delay-based network. It is known that the discrete phase shifters compromise among cost, bandwidth and phase resolution. The phase shifters with wide bandwidth and high phase precision are costly, not very attractive for the hybrid beamforming systems. In contrast to the discrete phase shifters, the time-delay-based PSN is of great interest, since it is of low cost and inherently wideband, and has no quantization issue of the phase. The time-delay-based PSN can be realized by the lens antennas [89]. However, in [89] only the concept of the hybrid systems based on lens antennas was presented without specifying any lens antennas and experimental verification. In this chapter, we specify Rotman lens [81, 82] with Uniform Linear Array (ULA) to realize the static beamspace. As a True-Time Delay (TTD) beamformer, the Rotman lens is broadband, being attractive for the future wideband systems, e.g., mmWave systems. Compared to the other lens antennas, the Rotman lens is very low-cost and simple to realize and fabricate with printed circuit boards due to its simple microwave structure [98].

It is observed in [87] that the insertion loss of the phase-shifter-based PSN has significant impact on the hybrid systems, and the system gain can diminish in the worst case. Thus the insertion loss of the Rotman lens should be considered carefully in the system design. The Rotman lens has significant insertion loss [98], especially when its dimension is large. This is because for a fully-connected structure, i.e., all the antennas are connected to a single Rotman lens, the dimension of the Rotman lens becomes large, leading to a long analog routing distance. To make the Rotman lens more suitable for the massive MIMO hybrid system, we divide the single large array into multiple disjoint sub-arrays, each of which is connected with a small-size Rotman lens individually. We refer to the hybrid system with partially-connected structure as sub-array hybrid beamforming system; while the hybrid system with fully-connected structure as full-array hybrid beamforming system.

The contributions of this chapter are summarized as follows:

- A hybrid analog-digital beamforming system based on Rotman lens is presented for massive MIMO systems. To our best knowledge, this is the first reported hybrid massive MIMO system based on Rotman lens in the literature. We examine and simulate the full-array and sub-array architectures, and find that the latter is more competitive than the former regarding the RF design feasibility, insertion loss, system scalability and performance.

- Given the fixed nature of the Rotman lens and the well-known Linear Minimum Mean-Squared Error (LMMSE) estimator as the digital beamformer, we formulate the system optimization as beam selection to select an optimal beam subset with the minimum mean-squared error. Furthermore, we propose two beam selection algorithms, i.e., greedy-search-based and Branch-And-Bound (BAB)-search-based beam selection, for both full-array and sub-array systems. The former has very low complexity but sub-optimal performance while the latter has high complexity but globally optimal performance.

- We give a mathematical model of Rotman lens by considering the inherent material dissipation loss and spill-over loss, to study the relationship between insertion loss and the number of antennas.
- A sample of Rotman lens in the 5-GHz radio band is designed and fabricated to study the hybrid system experimentally. We measure the scattering coefficients of the fabricated Rotman lens, and use them to simulate the system error-rate performance in the wide-band range.

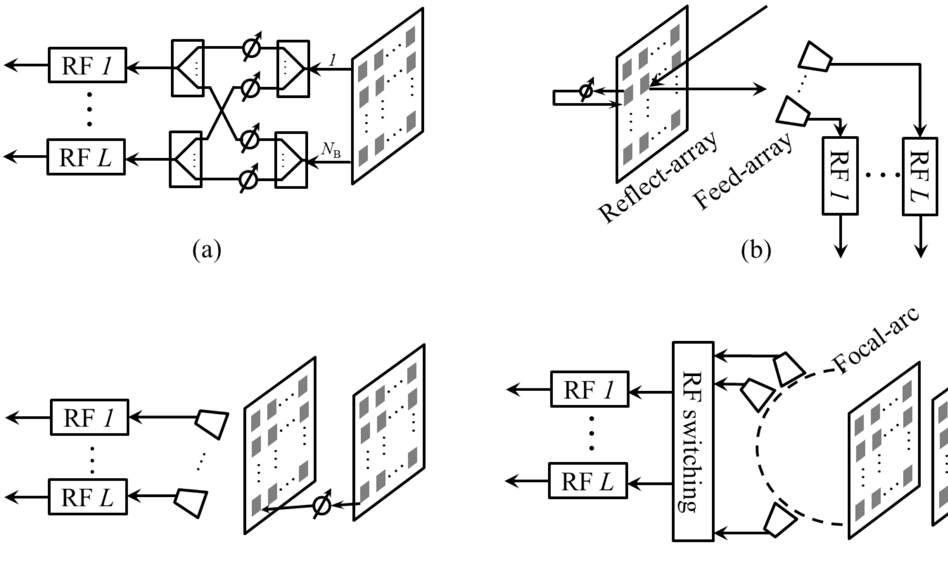
4.2 System Model and Problem Statement

A typical hybrid analog-digital beamforming system generally consists of analog beamforming and digital beamforming. The analog beamforming is to map the array signal vector of dimension-N_{B} to the RF transceiver input signal of dimension-L ($L \ll N_{\mathrm{B}}$). In the view of pure signal processing, the optimal analog beamforming is the eigen-beamforming [61], which is to extract most significant eigen-channels of the full channel matrix $\mathbf{H}$. The eigen-beamforming requires both amplitude and phase variations to the array signals. However, in the analog domain it is only efficient to adapt the phase rather than the magnitude. Thus the analog beamforming is generally an amplitude-constraint problem and the analog beamforming networks are critical for the design of massive MIMO hybrid analog-digital beamforming systems.

In this section, several analog beamforming networks will be reviewed briefly by taking the array architectures into account. By comparing different analog beamforming networks, the Rotman lens shows many advantages over the other analog beamforming networks, thus being used to design our massive MIMO hybrid beamforming systems. The system model, i.e., signal input-output model, will also be given in this section. Finally, the problem of the hybrid beamforming system based on Rotman lens is formulated in mathematics.

4.2.1 A brief review on the analog beamforming networks and array architectures

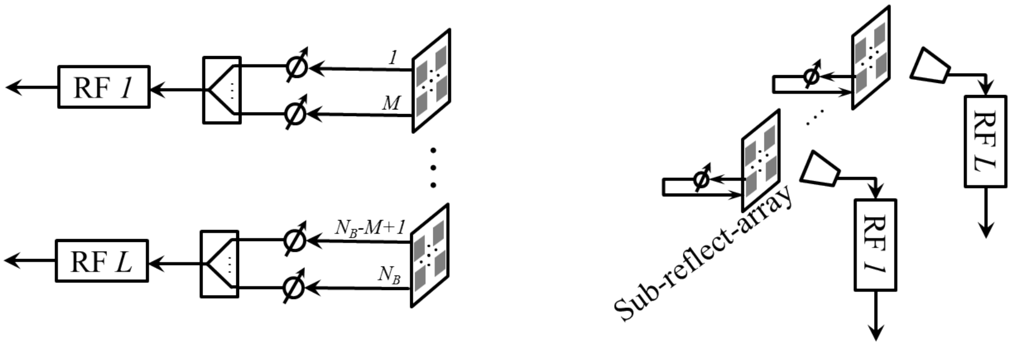
There are several types of analog circuit networks that can realize the analog beamforming, such as phased-array as shown in Fig. 4.1 (a) [61, 73], reflect-array with tunable phase shifters as



(a) (b)

(c) (d)

Figure 4.1: Several analog beamforming networks with the full-array architecture: (a) phased-array; (b) reconfigurable reflect-array with tunable phase shifters; (c) reconfigurable lens-array with tunable phase shifters; (d) fixed-multi-beam lens-array



(a) (b)

Figure 4.2: Phase-array and reflect-array with the sub-array architecture: (a) phased-array; (b) reconfigurable reflect-array with tunable phase shifters

shown in Fig. 4.1 (b) [78], lens-array with tunable phase shifter as shown in Fig. 4.1 (c) [78], etc. Note that all the analog circuit networks shown in Fig. 4.1 follow the full-array architecture. The sub-array architecture could be directly derived by segmenting the full-array into multiple sub-arrays. For example, Fig. 4.2 shows the phased-array and reflect-array with the sub-array architecture [72]. Obviously, the sub-array architecture has lower circuit complexity than the full-array architecture.

Now let us briefly compare these analog beamforming networks and the array architectures.

In the signal processing point of view, the phased-array is most favorite compared to other analog beamforming networks, since arbitrary beam patterns can be synthesized by tuning the phase shifters, which can be optimized to reduce the inter-user interference in the wireless communication systems. However, the advantage of the phased-array on the large number of degrees of freedom for beamforming is at the price of high complexity and high cost. A large number of phase shifters can be foreseen in the massive MIMO hybrid systems. For example, the numbers of $N_{\mathrm{B}}L$ and N_{B} phase shifters are required in the full-array and sub-array architectures, respectively. In addition, the large number of antenna elements will also reduce the power feeding efficiency dramatically, especially in the high frequency [78]. In contrast, the reflect-array and lens-array have the advantage of efficient power feeding, which is realized by the inherent spatial feeding. In both reflect-array and lens-array, the phase of each antenna element is tunable, thus referring to such arrays as reconfigurable reflect-array and reconfigurable lens-array, respectively. The concept behind the reconfigurable reflect-array and lens-array is actually very similar that both arrays focus the plane waves received by the antenna elements to the feeders. The focusing/beamforming is achieved by collimating the plane waves at the antenna elements with phase compensation through the phase shifters attached to each antenna element [78]. The difference between the reflect-array and the lens-array lies in the different positions of the feeders. For the reflect-array, the feeders are in front of the antenna array; while for the lens array, the feeders are in the back of the antenna array. Note that in the reflect-array, the blockage issue will become significant if there are many feeders in front of the antenna array. Thus from the perspective of blockage loss, the sub-array architecture is superior to the full-array architecture, because of only one feeder in front of the antenna array in the sub-array architecture as shown in Fig. 4.2 (b). In both the reconfigurable reflect-array and lens-array, the variable phase-shifters are required for each antenna element, which is actually the same to the phased-array with the scenario of sub-array architecture. Tuning the large number of phase shifters will bring in high-complexity control circuits, which is not very flexible when N_{B} is large. An alternative way is fixing the phase shifting at each antenna element and placing a fixed number of N_{B} feeders at the focal arc as shown in Fig. 4.1 (d), which is referred to as fixed-multi-beam lens array. L out of N_{B} feeders can be selected through the RF switching network discussed in Chapter 3. Rotman lens is an example of the fixed-multi-beam lens-array. Besides the advantages on the power feeding and RF control, the Rotman lens is also low-cost and low-complexity to implementation, thus being studied in this thesis particularly.

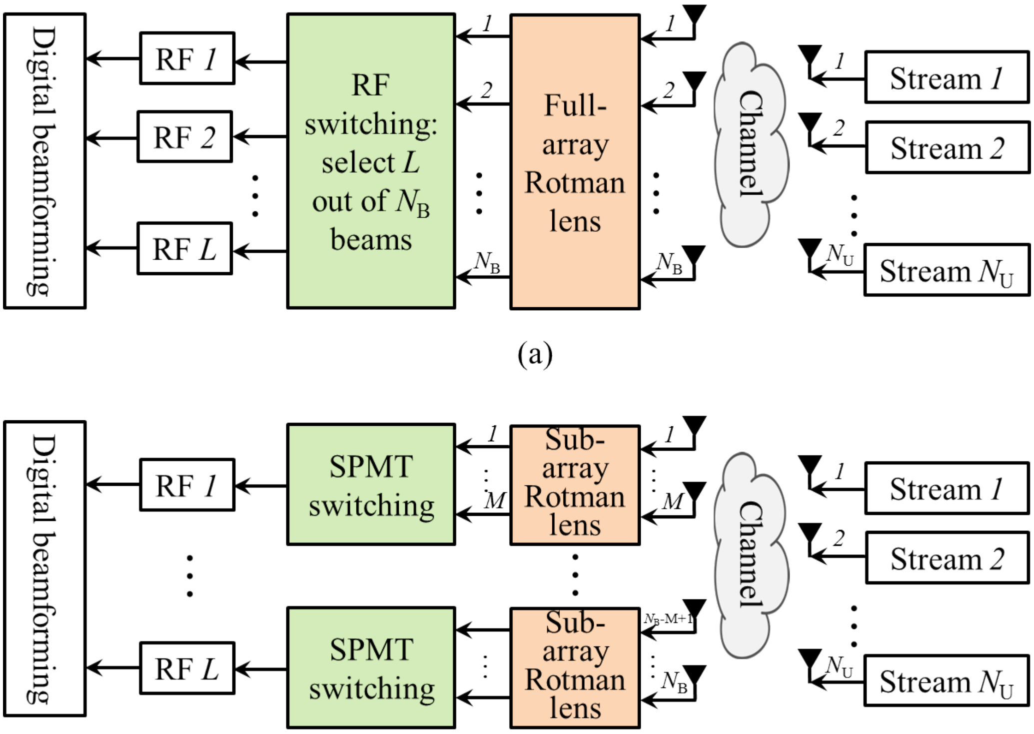
4.2.2 Rotman-lens-based hybrid beamforming

The system model with a single large-size Rotman lens is shown in Fig. 4.3 (a) for the full-array architecture. As aforementioned, the full-array architecture has the advantage on the analog beamforming by jointly processing the large-array signal and synthesizing very narrow beams. However, there are several critical challenges on the practical implementations on the system architecture as shown in Fig. 4.3 (a). The first challenge is the feasibility of the full-array RF selection network that is able to select an arbitrary subset from N_{B} beam ports, which is the same to the RF switching network in Chapter 3. As mentioned in Chapter 3, a low-loss and also high-isolation RF switching network is challenging to design when N_{B} is large. Secondly it is also difficult to design a single low-loss large-size Rotman lens, because the Rotman lens is a passive network, and the insertion loss will increase as the number of antennas increases due to the longer analog signal transmission distance in the Rotman lens. Thirdly the full-array architecture is lack of scalability since larger-size Rotman lens should be redesigned when the number of antennas increases.

A more practical architecture is the sub-array architecture as shown in Fig. 4.3 (b). The analog beamforming will be performed independently in each sub-array with a small-size Rotman lens. Without loss of generality we assume that the full-array are divided into L equal-size sub-arrays, each of which has $M = N_{\text{B}}/L$ antenna ports and M beam ports. In each sub-array, only one beam port is selected through the SPMT RF switch, which is the same to the sub-array antenna selection systems as discussed in Chpater 3. Obviously, the sub-array architecture has advantages on the RF switching and insertion loss over the full-array architecture [90]. In addition, the sub-array architecture has better scalability than the full-array architecture by simply adding more sub-arrays into the system, making it particularly suitable for the massive MIMO systems. Therefore, the sub-array architecture is particularly attractive for the Rotman-lens-based hybrid beamforming systems.

4.2.3 Basic concepts of Rotman lens

Rotman lens is a kind of TTD beamforming network [82] by artificially introducing time delays to focus the EM wave transmitted or received by an antenna array (here ULA is used). Fig. 4.4 shows a schematic representation of a general Rotman lens with N_{BP} beam ports and N_{AP} antenna ports. Without loss of generality, we assume that $N_{\text{BP}} = N_{\text{AP}} = N_{\text{P}}$. Note that for the full-array system as shown in Fig. 4.3 (a), we have $N_{\text{P}} = N_{\text{B}}$; while for the sub-array system as shown



(a)

(b)

Figure 4.3: System model and array architectures of the hybrid beamforming: (a) full-array architecture with a single large-size Rotman lens; (b) sub-array architecture with multiple small-size Rotman lenses

in Fig. 4.3 (b), we have $N_\mathrm{P} = M$. Now let us have a look at the principle behind Rotman lens. As shown in Fig. 4.4, the ith and jth wave transmission paths from the ith and jth antenna ports to the mth beam port have the distance $r_{mi} = d_{mi} + w_i$ and $r_{mj} = d_{mj} + w_j$, respectively, where w_i is the transmission line length from the ith antenna port to the ith linear taper in the array-port contour and d_{mi} is the wave transmission distance from the ith linear taper to the mth beam port (w_j and d_{mj} can be defined similarly for the jth antenna port). Suppose that the distance between neighbouring antennas is Δ_ant. The wave delay distance b_{ji} can be calculated as $b_{ji} = (j - i)\Delta_\mathrm{ant}\sin(\phi_m)$. In order to ensure that the incident plane wave with Angle of Arrival (AoA) ϕ_m can be focused at the mth beam port, the transmission distance difference of the ith and jth paths should compensate the phase difference introduced by b_{ji}, so that the signals can be superimposed constructively. Thus, the geometric shape of the Rotman lens, e.g., beam-port and array-port contours, transmission line length, etc., should be designed aiming to satisfy the following equation

$$k_\mathrm{s}(r_{mi} - r_{mj}) = k_0 b_{ji}, \tag{4.1}$$

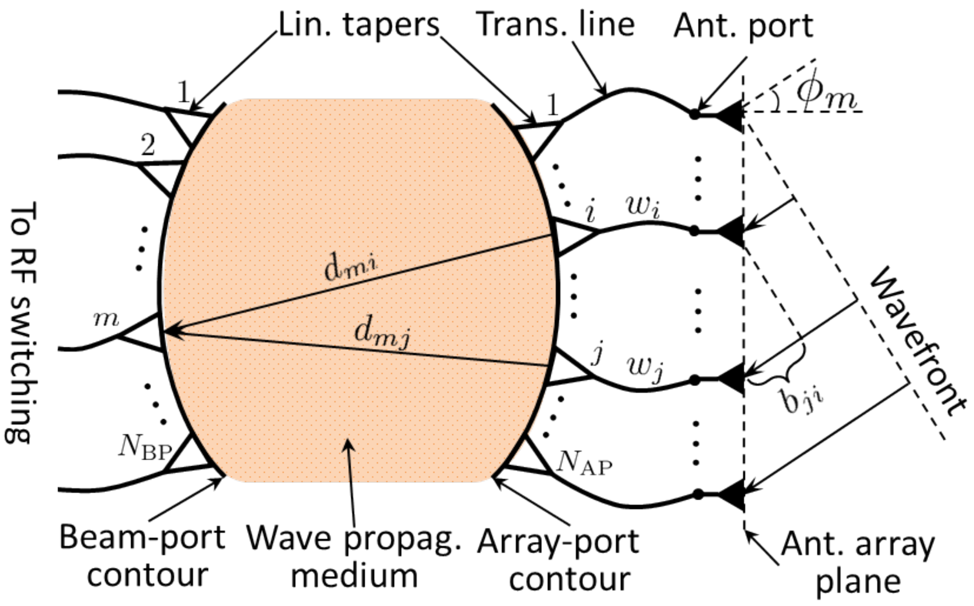


Figure 4.4: Schematic representation of Rotman lens with antenna elements and incident wave (top view)

where k_{s} and k_0 are the wavenumbers of the Rotman lens substrate dielectric material and air, respectively. The point that beam port m lies at the beam-port contour is also called focal point and the beam-port contour is also called focal arc. Actually, it is impossible to fulfill Equ. (4.1) for *all of* the beam ports at the same time [81]. The practical way to design Rotman lens is usually to pick up only three perfect focal points on the beam-port contour and the beam-port contour becomes a circular focal arc which can guarantee limited phase errors of the rest beam ports [81]- [86]. We will illustrate this in Section 4.4.

Generally speaking, the $N_{\mathrm{P}} \times N_{\mathrm{P}}$ Rotman lens collects the incident RF signals from N_{P} antennas to form N_{P} beams corresponding to N_{P} AoAs. The transmit case is similar to the receive case, where the RF signals from the mth beam port will be transmitted along the direction ϕ_m. In order to characterize the Rotman lens on the beamforming behaviour used for the following signal model, we need the mathematical model of it. Since the Rotman lens is essentially a multi-port RF network, we can describe the Rotman lens with a $N_{\mathrm{B}} \times N_{\mathrm{B}}$ matrix $\mathbf{F}_{\mathrm{RM}}$ with the entry $[\mathbf{F}_{\mathrm{RM}}]_{mn}$ as the transmission coefficient, e.g., the scattering parameter, from the nth antenna port to the mth beam port. If the Rotman lens is ideal without any phase errors and insertion loss, we simply have $\mathbf{F}_{\mathrm{RM}} = \mathbf{F}_{\mathrm{RM,full\text{-}array}} = [\mathbf{a}_{N_{\mathrm{P}}}(\phi_1^{\mathrm{FAS}}), \mathbf{a}_{N_{\mathrm{P}}}(\phi_2^{\mathrm{FAS}}), \cdots, \mathbf{a}_{N_{\mathrm{P}}}(\phi_{N_{\mathrm{P}}}^{\mathrm{FAS}})]^H$ and $\mathbf{F}_{\mathrm{RM}} = \mathbf{I}_L \otimes \mathbf{F}_{\mathrm{RM,sub\text{-}array}} = \mathbf{I}_L \otimes [\mathbf{a}_M(\phi_1^{\mathrm{SAS}}), \mathbf{a}_M(\phi_2^{\mathrm{SAS}}), \cdots, \mathbf{a}_M(\phi_M^{\mathrm{SAS}})]^H$ for the full-array and sub-array systems, respectively, where $\mathbf{a}_N(\phi)$ is the steering vector of an N-element ULA, i.e.,

$$\mathbf{a}_N(\phi) = \frac{1}{\sqrt{N}} \left[1, e^{\jmath \frac{2\pi}{\lambda_\text{c}} \Delta_\text{ant} \sin(\phi)}, \cdots, e^{\jmath (N-1) \frac{2\pi}{\lambda_\text{c}} \Delta_\text{ant} \sin(\phi)}\right]^T, \tag{4.2}$$

where $\jmath = \sqrt{-1}$, and λ_c is the wave-length in the air. Note that $\{\phi_1^\text{FAS}, \phi_2^\text{FAS}, \cdots, \phi_{N_\text{B}}^\text{FAS}\}$ and $\{\phi_1^\text{SAS}, \phi_2^\text{SAS}, \cdots, \phi_M^\text{SAS}\}$ are the sets of sweeping angles corresponding to the beam ports of the full-array and sub-array Rotman lenses, respectively. The static beamspace specified by the Rotman lens is thus spanned by the row vectors of $\mathbf{F}_\text{RM}$. However, in practice the Rotman lens is a passive network with loss and phase errors, which implies that the trace $\text{tr}(\mathbf{F}_\text{RM}\mathbf{F}_\text{RM}^H) < N_\text{P}$. The mathematical model of the Rotman with power loss and phase errors will be considered in Section 4.4, in which a sample of Rotman lens will also be designed, fabricated and measured, and the measurement results of the practical matrix $\mathbf{F}_\text{RM}$ will be used to study the impacts of its impairments on the hybrid system performance.

In the view of signal processing, $\mathbf{F}_\text{RM}$ is to sample the multi-antenna signals in the angular domain with fixed angular bins [82]. Given the sparse nature of the wireless channels, e.g., 5-GHz band channel used in this thesis, the information only dominates in a subset of angular bins which will be selected by RF swithces. In mathematics, this RF selection functionality can be described by the selection matrix $\mathbf{S}_\text{RF}$ to select a subset of rows from $\mathbf{F}_\text{RM}$ (note that each row of $\mathbf{F}_\text{RM}$ corresponds to each beam port of the Rotman lens). Thus $\mathbf{S}_\text{RF}$ is an $L \times N_\text{B}$ matrix, and each row of it has one and only one nonzero value (unity value) and the column indices of these nonzero values are the selected beam indices. For example, suppose that the Rotman lens has eight beamports indexed from 1 to 8. If the 3th and 5th beam ports are going to be selected, we have

$$\mathbf{S}_\text{RF, example} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

4.2.4 Noise model

Given a noise-limited communication system, the receive noise has a major impact on the receiver performance. Since our receiver performs two-stage signal processing from the analog domain to the digital domain, the noise model should be considered carefully. It is commonly known that the total receive noise $\tilde{\mathbf{w}}$ usually comprises the external noise $\mathbf{w}_\text{ext}$ and the internal noise $\mathbf{w}_\text{int}$, i.e.,

$$\tilde{\mathbf{w}} = \mathbf{S}_\text{RF}\mathbf{F}_\text{RM}\mathbf{w}_\text{ext} + \mathbf{w}_\text{int}, \tag{4.3}$$

The external noise $\mathbf{w}_{\text{ext}}$ comes from radio environments, e.g., environment thermal noise, environment man-made noise, etc.; while the internal noise $\mathbf{w}_{\text{int}}$ comes from the receiver noisy electronic components, e.g., amplifiers, mixers, etc. Note that the white external noise filtered by the Rotman lens will become colorful, which will make the system optimization difficult. Fortunately, in practice with the carrier frequency of 5 GHz of our interest, the internal noise dominates according to [91, Fig. 3.1]. Hence the receive noise can be simply modeled as zero-mean Additive While Gaussian Noise (AWGN) following $\mathcal{CN}(\mathbf{0}, \sigma_{\tilde{\mathbf{w}}}^2 \mathbf{I}_L)$, where $\sigma_{\tilde{\mathbf{w}}}^2$ is the noise variance.

4.2.5 System model with equivalent channel

Given the aforementioned analog beamforming and noise models, and the general signal model in Chapter 2, the digital baseband signal vector after the beam selection reads

$$\tilde{\mathbf{y}} = \sqrt{E_{\text{x}}}\mathbf{S}_{\text{RF}}\mathbf{F}_{\text{RM}}\mathbf{H}\mathbf{x} + \tilde{\mathbf{w}}, \tag{4.4}$$

where $\mathbf{x} \in \mathcal{A}^{N_{\text{U}} \times 1}$ is the transmitted symbol vector with the mean $\mathbb{E}\{\mathbf{x}\} = \mathbf{0}$ and the covariance matrix $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \frac{1}{N_{\text{U}}}\mathbf{I}_{N_{\text{U}}}$; $\mathcal{A}$ is the modulation constellation space; E_{x} is the total transmitted power over the N_{U} antennas. The Signal to Noise Ratio (SNR) is thus defined as $\rho = E_{\text{x}}/\sigma_{\tilde{\mathbf{w}}}^2$. The $N_{\text{B}} \times N_{\text{U}}$ matrix $\mathbf{H}$ is the channel matrix with $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_{N_{\text{U}}}]$ and $\mathbf{h}_u \in \mathbb{C}^{N_{\text{B}} \times 1}$ is the channel vector for the uth user. With the sparse nature of the 5 GHz wireless channels, the channel response vector $\mathbf{h}_u$ in the spatial domain can be modeled as the superposition of the channel impulse response in the angular domain [92]:

$$\mathbf{h}_u = \sum_{p=1}^{P_u} \sqrt{\frac{N_{\text{B}}}{P_u}} g_{up} \mathbf{a}_{N_{\text{B}}}(\phi_{up}) \tag{4.5}$$

where P_u is the total number of Multi-Path Components (MPCs) for the uth user; g_{up} and ϕ_{up} denote the complex MPC gain and the AoA of the pth MPC in the uth user, respectively.

By defining the equivalent channel matrix $\tilde{\mathbf{H}} = \mathbf{F}_{\text{RM}}\mathbf{H}$, we can rewrite Equ. (4.4) and get the system model under the equivalent channel

$$\tilde{\mathbf{y}} = \sqrt{E_{\text{x}}}\mathbf{S}_{\text{RF}}\tilde{\mathbf{H}}\mathbf{x} + \tilde{\mathbf{w}}. \tag{4.6}$$

4.2.6 Problem statement

The linear Equ. (4.6) indicates that selecting a high-quality submatrix from the equivalent channel matrix $\tilde{\mathbf{H}}$ is the key of the receiver, i.e., the selected submatrix should be well-conditioned rather than ill-conditioned, to ensure high performance of the digital beamforming. In this thesis, we focus on the linear receiver LMMSE due to its simplicity and high performance. With the LMMSE filter, the digital baseband beamforming matrix has the well-known form $\mathbf{F}_{\text{BB}} = (\tilde{\mathbf{H}}_{\text{s}}^{H}\tilde{\mathbf{H}}_{\text{s}} + N_{\text{U}}/\rho\mathbf{I}_{N_{\text{U}}})^{-1}\tilde{\mathbf{H}}_{\text{s}}^{H}$, where $\tilde{\mathbf{H}}_{\text{s}} = \mathbf{S}_{\text{RF}}\tilde{\mathbf{H}}$ is the equivalent channel submatrix after the beam selection, i.e., selecting L rows from $\tilde{\mathbf{H}}$. The total Mean-Squared Error (MSE) of estimating N_{U} data streams is known as the trace of the MSE matrix $(\tilde{\mathbf{H}}_{\text{s}}^{H}\tilde{\mathbf{H}}_{\text{s}} + N_{\text{U}}/\rho\mathbf{I}_{N_{\text{U}}})^{-1}$. Let us define $\mathbf{G} = (\mathbf{I}_{N_{\text{U}}} + \bar{\rho}\tilde{\mathbf{H}}_{\text{s}}^{H}\tilde{\mathbf{H}}_{\text{s}})^{-1}$ and $\bar{\rho} = \rho/N_{\text{U}}$. In order to optimize the error-rate performance of the system with the LMMSE digital beamformer, we consequently have the following optimization problems for the full-array selection and sub-array selection systems:

$$\tilde{\mathbf{H}}_{\text{s}}^{\text{opt}} = \arg \min_{\tilde{\mathbf{H}}_{\text{s}} \in \tilde{\mathcal{H}}_{\text{FAS}} \text{ or } \tilde{\mathcal{H}}_{\text{SAS}}} \operatorname{tr}(\mathbf{G}), \tag{4.7}$$

where $\tilde{\mathcal{H}}_{\text{FAS}}$ and $\tilde{\mathcal{H}}_{\text{SAS}}$ are the sets consisting of all of the $L \times N_{\text{U}}$ equivalent channel submatrices for the full-array selection system and sub-array selection system, respectively. Since each submatrix corresponds to one beam subset of size L from the complete beam set of size N_{B}, the optimization given by Equ. (4.7) is thus to select the optimal beam subset to minimize the MSE.

4.3 Beam Selection Algorithms

The beam selection problem given in Equ. (4.7) is actually similar to the antenna selection problem [79] discussed in Chapter 3, where the beams are now selected instead of antennas. The antenna selection algorithms presented in Chapter 3 can also be used to perform beam selection. Similar algorithms can also be found in [79]. However, it should be noted that selection algorithms presented in Chapter 3 and [79] are to maximize the channel capacity while the beam selection problem given by Equ. (4.7) is to minimize the MSE. Thus different selection algorithm are necessary to be studied. Some antenna selection algorithms with MSE-oriented selection criteria can be applied to solve our beam selection problem. For example, the works in [95, 96] used a greedy search to find the sub-optimal antenna subsets. Given the nonlinear LMMSE successive interference cancellation receiver, the antenna selection criterion in [95, 96] is based on min-max-MSE, i.e., minimization of the maximum MSE of the multiple detection stages. In contrast, our system uses the simple LMMSE receiver, following the minimum-MSE-based selection criterion as shown in Equ. (4.7). Similar to the greedy search in [79, 95, 96], in this section we propose a greedy-search-based beam selection algorithm with a recursive updating method aiming to minimize the MSE in low-complexity. As the NP-hardness of the submatrix selection problems [93], the greedy search with polynomial-complexity is non-optimal. The conventional optimal method, i.e., exhaustive search, has the complexity of $\mathcal{O}\left(\binom{N_{\mathrm{B}}}{L}\right)$ and $\mathcal{O}(M^L)$ for the full-array and sub-array selection, respectively, becoming infeasible when N_{B} is large. By using the updating method studied in the greedy search, we then propose an optimal beam selection algorithm which is based on the well-known BAB search algorithm [55], which has much lower complexity than the exhaustive search.

4.3.1 Sub-optimal beam selection based on greedy search

Let $\tilde{\mathbf{H}}_n$ denote the submatrix by selecting n rows from $\tilde{\mathbf{H}}$ after n selection steps and T_n denote the trace of the matrix $\mathbf{G}_n = (\mathbf{I}_{N_{\mathrm{U}}} + \bar{\rho}\tilde{\mathbf{H}}_{\mathrm{s}}^H\tilde{\mathbf{H}}_{\mathrm{s}})^{-1}$, $n = 0, 1, \cdots, L-1$ (note that $\mathbf{G}_0 = \mathbf{I}_{N_{\mathrm{U}}}$). Suppose that the jth row of $\tilde{\mathbf{H}}$, i.e., $\tilde{\mathbf{h}}_j^H$, is going to be selected from the rest candidate set in the $(n+1)$th step. The new $(n+1) \times N_{\mathrm{U}}$ channel submatrix is thus

$$\tilde{\mathbf{H}}_{n+1} = \begin{bmatrix} \tilde{\mathbf{H}}_n \\ \tilde{\mathbf{h}}_j^H \end{bmatrix}. \tag{4.8}$$

Then T_{n+1} is calculated as

$$\begin{aligned} T_{n+1} &= \mathrm{tr}\left(\mathbf{G}_{n+1}\right) \\ &= \mathrm{tr}\left((\mathbf{I}_{N_{\mathrm{U}}} + \bar{\rho}\tilde{\mathbf{H}}_{n+1}^H\tilde{\mathbf{H}}_{n+1})^{-1}\right) \\ &= \mathrm{tr}\left((\mathbf{I}_{N_{\mathrm{U}}} + \bar{\rho}\tilde{\mathbf{H}}_n^H\tilde{\mathbf{H}}_n + \bar{\rho}\tilde{\mathbf{h}}_j\tilde{\mathbf{h}}_j^H)^{-1}\right). \end{aligned} \tag{4.9}$$

By using the Sherman-Morrison formula [94], Equ. (4.9) can be rewritten into the following recursive form

$$T_{n+1} = \mathrm{tr}\left(\mathbf{G}_n - \frac{\bar{\rho}\mathbf{G}_n\tilde{\mathbf{h}}_j\tilde{\mathbf{h}}_j^H\mathbf{G}_n}{1 + \bar{\rho}\tilde{\mathbf{h}}_j^H\mathbf{G}_n\tilde{\mathbf{h}}_j}\right) \overset{\text{(a)}}{=} T_n - \underbrace{\frac{\tilde{\mathbf{h}}_j^H\mathbf{G}_n^2\tilde{\mathbf{h}}_j}{\bar{\rho}^{-1} + \tilde{\mathbf{h}}_j^H\mathbf{G}_n\tilde{\mathbf{h}}_j}}_{\Delta_{j,n}}, \tag{4.10}$$

where the equality (a) holds with the linear and cyclic properties of the matrix trace.

Algorithm 4.1 Greedy-search-based beam selection

```
Initialization: $\mathcal{S} = \emptyset$, $\mathbf{G} := \mathbf{I}_{N_\mathrm{U}}$, $\mathcal{J} := \{1, 2, \cdots, N_\mathrm{B}\}$
$\mathbf{a}_j := \tilde{\mathbf{h}}_j^H, \forall j \in \mathcal{J}$
$\nu_j := \tilde{\mathbf{h}}_j^H \tilde{\mathbf{h}}_j, \forall j \in \mathcal{J}$    ▷ $\mathcal{O}(N_\mathrm{U} N_\mathrm{B})$
$\Delta_j = \frac{\nu_j}{\bar{\rho}^{-1} + \nu_j}, \forall j \in \mathcal{J}$    ▷ $\mathcal{O}(N_\mathrm{B})$
for $n = 1 : L$ do
    $J = \arg\max_{j \in \mathcal{J}} \Delta_j$
    $\mathcal{S} := \mathcal{S} \cup J$
    $\mathcal{J} := \mathcal{J} \setminus J$, or    ▷ full-array sel.
    $\mathcal{J} := \mathcal{J} \setminus \mathcal{I}_{f(J)}^{\mathrm{SAS}}$    ▷ sub-array sel.
    if $n < L$ then
        $\mathbf{g} = \frac{\mathbf{G}\tilde{\mathbf{h}}_J}{\sqrt{\bar{\rho}^{-1} + \nu_J}}$    ▷ $\mathcal{O}(N_\mathrm{U}^2 L)$
        $\xi_j = \tilde{\mathbf{h}}_j^H \mathbf{g}, \forall j \in \mathcal{J}$    ▷ $\mathcal{O}(N_\mathrm{U} N_\mathrm{B} L)$
        $\mathbf{a}_j := \mathbf{a}_j - \xi_j \mathbf{g}, \forall j \in \mathcal{J}$    ▷ $\mathcal{O}(N_\mathrm{U} N_\mathrm{B} L)$
        $\nu_j := \nu_j - |\xi_j|^2, \forall j \in \mathcal{J}$    ▷ $\mathcal{O}(N_\mathrm{B} L)$
        $\Delta_j = \frac{||\mathbf{a}_j||_\mathrm{F}^2}{\bar{\rho}^{-1} + \nu_j}, \forall j \in \mathcal{J}$    ▷ $\mathcal{O}(N_\mathrm{U} N_\mathrm{B} L)$
        $\mathbf{G} := \mathbf{G} - \mathbf{g}\mathbf{g}^H$    ▷ $\mathcal{O}(N_\mathrm{U}^2 L)$
    end if
end for
return the selected beam indices are the elements of $\mathcal{S}$
```

Obviously, both $\mathbf{G}_n$ and $\mathbf{G}_n^2$ are positive definite matrices. Thus the decremental value $\Delta_{j,n}$ given in Equ. (4.10) is a positive number when $\tilde{\mathbf{h}}_j \neq \mathbf{0}$, which means that the MSE function T_n is a monotonically decreasing function as n increases. Thus heuristically, a greedy searching algorithm can be applied to find a sub-optimal solution. In detail, we examine the decremental value of each beam in the candidate set and select the maximum one to minimize the MSE in the current step. The searching procedure goes forward until the last beam is selected. It should be noted that the greedy beam select algorithm by directly using Equ. (4.10) to find the desired beam indices with L selection steps involves a computational complexity of $\mathcal{O}(N_\mathrm{U}^2 N_\mathrm{B} L)$ [100] for both full-array and sub-array selections. In order to reduce the complexity, we propose the following updating method. Suppose that the beam index J_n was found in the nth step. Then the matrix $\mathbf{G}_{n+1}$ can be calculated with the updating

$$\mathbf{G}_{n+1} = \mathbf{G}_n - \mathbf{g}_{n+1}\mathbf{g}_{n+1}^H, \; n = 0, 1, \cdots, L-1 \tag{4.11}$$

where $\mathbf{g}_{n+1} = \frac{\mathbf{G}_n \tilde{\mathbf{h}}_{J_n}}{\sqrt{\bar{\rho}^{-1} + \tilde{\mathbf{h}}_{J_n}^H \mathbf{G}_n \tilde{\mathbf{h}}_{J_n}}}$. Let us define $\mathbf{a}_{j,n+1} = \tilde{\mathbf{h}}_j^H \mathbf{G}_{n+1}$ and $\nu_{j,n+1} = \tilde{\mathbf{h}}_j^H \mathbf{G}_{n+1} \tilde{\mathbf{h}}_j$, which can be updated with the following formulas:

$$\begin{aligned}
\mathbf{a}_{j,n+1} &= \tilde{\mathbf{h}}_j^H \mathbf{G}_{n+1} \\
&= \tilde{\mathbf{h}}_j^H (\mathbf{G}_n - \mathbf{g}_{n+1}\mathbf{g}_{n+1}^H) \\
&= \mathbf{a}_{j,n} - \xi_{J_n,j,n+1}\mathbf{g}_{n+1}^H, \qquad (4.12) \\
\nu_{j,n+1} &= \tilde{\mathbf{h}}_j^H \mathbf{G}_{n+1} \tilde{\mathbf{h}}_j \\
&= \tilde{\mathbf{h}}_j^H (\mathbf{G}_n - \mathbf{g}_{n+1}\mathbf{g}_{n+1}^H) \tilde{\mathbf{h}}_j \\
&= \nu_{j,n} - |\xi_{J_n,j,n+1}|^2, \qquad (4.13)
\end{aligned}$$

where $\xi_{J_n,j,n+1} = \tilde{\mathbf{h}}_j^H \mathbf{g}_{n+1}$. Then the decremental values of all the candidate beam indices are simply calculated as $\Delta_{j,n+1} = \frac{||\mathbf{a}_{j,n+1}||_\mathrm{F}^2}{\bar{\rho}^{-1}+\nu_{j,n+1}}$. The beam index in the $(n+1)$th step is found by $J_{n+1} = \arg\max_{j \in \mathcal{J}_{n+1}} \Delta_{j,n+1}$, where $\mathcal{J}_{n+1}$ is the candidate set of the beams to be selected in the $(n+1)$th step. For the full-array selection, we have $\mathcal{J}_{n+1} = \mathcal{J}_n \setminus J_n$; while for the sub-array selection, we have $\mathcal{J}_{n+1} = \mathcal{J}_n \setminus \mathcal{I}_{f(J_n)}^{\mathrm{SAS}}$, where $\mathcal{I}_{f(J_n)}^{\mathrm{SAS}}$ is the beam set consisting of all of the beams in the $f(J_n)$th sub-array in which the beam J_n lies. The low-complexity greedy beam selection algorithm is summarized in Alg. 4.1 with the required computational complexity. From Alg. 4.1 we can get the overall complexity as $\mathcal{O}(\max(N_\mathrm{U}, N_\mathrm{B})N_\mathrm{U}L)$. Since we have $N_\mathrm{B} \gg N_\mathrm{U}$ in our system, the complexity becomes $\mathcal{O}(N_\mathrm{B}N_\mathrm{U}L)$, which indeed has lower complexity than directly using the updating method in Equ. (4.10).

4.3.2 Optimal beam selection based on BAB

Similar to the BAB based antenna selection proposed in Chapter 3, the beam selection problem formulated by Equ. 4.7 can also be solved by BAB algorithm, because the cost function T_n is also a monotonically decreasing function with respect to the subset size. But it also means that we can only use the upwards BAB to find the maximum value rather than the minimum value of the objective function in our beam selection. Instead, we need a monotonic increasing objective function for upwards BAB to find the minimum value. In the following, we will give an equivalent objective function which is monotonically increasing with respect to the subset size.

A new objective function

Suppose that the new objective function has the form

$$\tilde{T}_n = T_n + \sum_{m=0}^{n-1} Z_m, \tag{4.14}$$

where Z_m is defined as $Z_m = \frac{\zeta_m^2}{\bar{\rho}^{-1}+\zeta_m^2}$, and $\zeta_m = \max_{j\in\mathcal{I}_m^{\text{FAS}} \text{ or } \mathcal{I}_m^{\text{SAS}}} ||\tilde{\mathbf{h}}_j||_{\text{F}}$ for the full-array and sub-array selections. The sets $\mathcal{I}_m^{\text{FAS}}$ and $\mathcal{I}_m^{\text{SAS}}$ consists of all of the candidate beam indices in the mth selection step for the full-array and sub-array selections, respectively. With the above definitions, we will prove that $\tilde{T}_n$ is a monotonically increasing function of n.

With Equ. (4.10) and (4.14), $\tilde{T}_n$ can be written into the following recursive form

$$\tilde{T}_{n+1} = \tilde{T}_n - \Delta_{j,n} + Z_n,\ n = 0, 1, \cdots, L-1. \tag{4.15}$$

To ensure $\tilde{T}_{n+1}$ is a monotonically increasing function, we need $\Delta_{j,n} \le Z_n$, which means that Z_n should be an upper bound of $\Delta_{j,n}$.

Lemma 4.3.1 *Given* $\Delta_{j,n} = \frac{\tilde{\mathbf{h}}_j^H \mathbf{G}_n^2 \tilde{\mathbf{h}}_j}{\bar{\rho}^{-1}+\tilde{\mathbf{h}}_j^H \mathbf{G}_n \tilde{\mathbf{h}}_j}$ *in Equ. (4.10) and the definition of* Z_n*. Then we have* $\Delta_{j,n} \le Z_n$*,* $n = 0, 1, \cdots, L-1$*.*

Proof: First, from Equ. (4.11) we know that

$$\mathbf{G}_n = \begin{cases} \mathbf{I}_{N_\text{U}}, & n = 0, \\ \mathbf{I}_{N_\text{U}} - \sum_{m=1}^{n} \mathbf{g}_m \mathbf{g}_m^H, & n = 1, 2, \cdots, L-1 \end{cases} \tag{4.16}$$

Case 1: $n = 0$. We simply have

$$\begin{aligned} \Delta_{j,0} &= \frac{\tilde{\mathbf{h}}_j^H \tilde{\mathbf{h}}_j}{\bar{\rho}^{-1} + \tilde{\mathbf{h}}_j^H \tilde{\mathbf{h}}_j} \\ &= 1 - \frac{\bar{\rho}^{-1}}{\bar{\rho}^{-1} + \tilde{\mathbf{h}}_j^H \tilde{\mathbf{h}}_j} \\ &\overset{\text{(b)}}{\le} \frac{\zeta_0^2}{\bar{\rho}^{-1} + \zeta_0^2} \\ &= Z_0, \end{aligned} \tag{4.17}$$

where the inequality (b) holds when the candidate beam index $j \in \mathcal{I}_0^{\text{FAS}}$, or $\mathcal{I}_0^{\text{SAS}}$ in the 0th selection step.

Case 2: $n = 1, 2, \cdots, L-1$. Let us define $\bar{\mathbf{G}}_n = \sum_{m=1}^{n} \mathbf{g}_m \mathbf{g}_m^H$. Since the gram matrix $\mathbf{g}_m \mathbf{g}_m^H$ is positive semi-definite, the sum of positive semi-definite matrices, i.e., $\bar{\mathbf{G}}_n$, is also a positive semi-definite matrix [94, Observation 7.1.3]. Then it follows that

$$\tilde{\mathbf{h}}_j^H \mathbf{G}_n \tilde{\mathbf{h}}_j - \tilde{\mathbf{h}}_j^H \tilde{\mathbf{h}}_j = -\tilde{\mathbf{h}}_j^H \bar{\mathbf{G}}_n \tilde{\mathbf{h}}_j \leq 0. \tag{4.18}$$

Now let us study $\tilde{\mathbf{h}}_j^H \mathbf{G}_n^2 \tilde{\mathbf{h}}_j - \tilde{\mathbf{h}}_j^H \mathbf{G}_n \tilde{\mathbf{h}}_j = -\tilde{\mathbf{h}}_j^H \mathbf{G}_n \bar{\mathbf{G}}_n \tilde{\mathbf{h}}_j$. Define $\mathbf{B}_n = \mathbf{G}_n \bar{\mathbf{G}}_n$. Obviously, $\mathbf{B}_n$ is a hermitian matrix. With the eigen-decomposition $\mathbf{G}_n = \mathbf{U}_n \Lambda_{\mathbf{G}_n} \mathbf{U}_n^H$, $\bar{\mathbf{G}}_n$ consequently has the eigen-decomposition $\bar{\mathbf{G}}_n = \mathbf{I}_{N_\mathrm{U}} - \mathbf{G}_n = \mathbf{U}_n \Lambda_{\bar{\mathbf{G}}_n} \mathbf{U}_n^H$ with $\Lambda_{\bar{\mathbf{G}}_n} = \mathbf{I}_{N_\mathrm{U}} - \Lambda_{\mathbf{G}_n}$. Thus the eigen-decomposition of $\mathbf{B}_n$ can be written as $\mathbf{B}_n = \mathbf{U}_n \Lambda_{\mathbf{G}_n} \Lambda_{\bar{\mathbf{G}}_n} \mathbf{U}_n^H$. Since $\mathbf{G}_n$ and $\bar{\mathbf{G}}_n$ is positive definite and positive semi-definite matrices, respectively, the eigen-values of $\mathbf{B}_n$ are non-negative, which means that $\mathbf{B}_n$ is a positive semi-definite matrix. Thus we have

$$\tilde{\mathbf{h}}_j^H \mathbf{G}_n^2 \tilde{\mathbf{h}}_j - \tilde{\mathbf{h}}_j^H \mathbf{G}_n \tilde{\mathbf{h}}_j = -\tilde{\mathbf{h}}_j^H \mathbf{B}_n \tilde{\mathbf{h}}_j \leq 0. \tag{4.19}$$

Finally with inequalities (4.18) and (4.19), we have

$$\begin{aligned}
\Delta_{j,n} &= \frac{\tilde{\mathbf{h}}_j^H \mathbf{G}_n^2 \tilde{\mathbf{h}}_j}{\bar{\rho}^{-1} + \tilde{\mathbf{h}}_j^H \mathbf{G}_n \tilde{\mathbf{h}}_j} \\
&\leq \frac{\tilde{\mathbf{h}}_j^H \mathbf{G}_n \tilde{\mathbf{h}}_j}{\bar{\rho}^{-1} + \tilde{\mathbf{h}}_j^H \mathbf{G}_n \tilde{\mathbf{h}}_j} \\
&= 1 - \frac{\bar{\rho}^{-1}}{\bar{\rho}^{-1} + \tilde{\mathbf{h}}_j^H \mathbf{G}_n \tilde{\mathbf{h}}_j} \\
&\overset{\text{(c)}}{\leq} 1 - \frac{\bar{\rho}^{-1}}{\bar{\rho}^{-1} + \tilde{\mathbf{h}}_j^H \tilde{\mathbf{h}}_j} \\
&\overset{\text{(d)}}{\leq} \frac{\zeta_n^2}{\bar{\rho}^{-1} + \zeta_n^2} \\
&= Z_n,
\end{aligned} \tag{4.20}$$

where we have used Equ. (4.18) in the inequality (c) and the inequality (d) holds when the candidate beam index $j \in \mathcal{I}_n^{\text{FAS}}$ or $\mathcal{I}_n^{\text{SAS}}$ in the nth selection step. ■

Since the new objective function $\tilde{T}_n$ is a monotonically increasing function with respect to n, we can use upwards BAB search method to find the beam set that gives the minimum value of $\tilde{T}_L$ with branching and bounding procedure efficiently. Note that the new objective function in the last selection step $\tilde{T}_L$ equals the original objective function T_L plus the offset $\sum_{m=0}^{L-1} Z_m$ according to Equ. (4.14). Because the offset $\sum_{m=0}^{L-1} Z_m$ is a fixed number in each channel realization, and irrelevant to the selected beam subset, the beam subset leading to minimum value of $\tilde{T}_L$ also leads to the minimum value of T_L. Thus the BAB algorithm with the new objective function finds the minimum value of T_L.

Algorithm 4.2 BAB-search-based beam selection

```
1:  Initialization: $\mathbf{G} = \mathbf{I}_{N_\text{U}}$, $B = +\infty$, $\tilde{T} = N_\text{U}$, $n = 0$, $J = 0$, and the selected beam index
    vector $\text{s} = \mathbf{0}_L$, $\mathcal{J} = \{1, 2, \cdots, N_\text{B}\}$, $\mathcal{L} = \{1, 2, \cdots, L\}$, $\mathbf{a}_j = \tilde{\mathbf{h}}_j^H, \forall j \in \mathcal{J}$
2:  $\nu_j = ||\tilde{\mathbf{h}}_j||_\text{F}^2, \forall j \in \mathcal{J}$                                ▷ $\mathcal{O}(N_\text{U} N_\text{B})$
3:  $\Delta_j = \frac{\nu_j}{\bar{\rho}^{-1} + \nu_j}, \forall j \in \mathcal{J}$                                  ▷ $\mathcal{O}(N_\text{B})$
4:  $\zeta_m = \max_{j \in \mathcal{I}_m^\text{FAS} \text{ or } \mathcal{I}_m^\text{SAS}} \nu_j, \forall m \in \mathcal{L}$
5:  $Z_m = \zeta_m / (\bar{\rho}^{-1} + \zeta_m), \forall m \in \mathcal{L}$                              ▷ $\mathcal{O}(L)$
6:  if $n = L - 1$ then
7:      $t_j := \tilde{T} - \Delta_j + Z_{L-1}, \forall j \in \mathcal{I}_{L-1,J}^\text{FAS} \text{ or } \mathcal{I}_{L-1}^\text{SAS}$
8:      if $\min_{m \in \mathcal{I}_{L-1,J}^\text{FAS} \text{ or } \mathcal{I}_{L-1}^\text{SAS}} t_m < B$ then                        ▷ Updating bound
9:          update $[\text{s}]_L = \arg\min_{m \in \mathcal{I}_{L-1,J}^\text{FAS} \text{ or } \mathcal{I}_{L-1}^\text{SAS}} t_m$, $B := \min_{m \in \mathcal{I}_{L-1,J}^\text{FAS} \text{ or } \mathcal{I}_{L-1}^\text{SAS}} t_m$, and $\bar{\text{s}} := \text{s}$
10:     end if
11: else
12:     $t_j := \tilde{T} - \Delta_j + Z_n, \forall j \in \mathcal{I}_{n,J}^\text{FAS} \text{ or } \mathcal{I}_n^\text{SAS}$
13:     sort $t_j, \forall j \in \mathcal{I}_{n,J}^\text{FAS} \text{ or } \mathcal{I}_n^\text{SAS}$ in an ascend order to get an ordered index vector $\text{j}$
14:     $\mathbf{G}_\text{tmp} := \mathbf{G}$, $\mathbf{a}_{\text{tmp},j} := \mathbf{a}_j, \nu_{\text{tmp},j} := \nu_j, \forall j \in \mathcal{J}$
15:     for $i$ from 1 to $|\mathcal{I}_{n,J}^\text{FAS}|$ or $M$ do
16:         $J = [\text{j}]_i$
17:         if $t_J < B$ then                                                        ▷ Branching
18:             $\mathcal{K} := \{J + 1, J + 2, \cdots, N_\text{B}\}$, or                  ▷ for the full-array selection
19:             $\mathcal{K} := \mathcal{I}_{n+1}^\text{SAS} \cup \mathcal{I}_{n+2}^\text{SAS} \cup \cdots \cup \mathcal{I}_{L-1}^\text{SAS}$                 ▷ for the sub-array selection
20:             update the index vector $[\text{s}]_{n+1} = J$
21:             $\mathbf{g} := \frac{1}{\sqrt{\bar{\rho}^{-1} + \nu_{\text{tmp},J}}} \mathbf{G} \tilde{\mathbf{h}}_J$, and $\mathbf{G} := \mathbf{G}_\text{tmp} - \mathbf{g}\mathbf{g}^H$                  ▷ $\mathcal{O}(N_\text{U}^2 N_\text{nodes})$
22:             $\tilde{T} := t_J$
23:             $\xi_m := \tilde{\mathbf{h}}_m^H \mathbf{g}$, and $\mathbf{a}_m := \mathbf{a}_{\text{tmp},m} - \xi_m \mathbf{g}, \forall m \in \mathcal{K}$          ▷ $\mathcal{O}(N_\text{U} N_\text{B} N_\text{nodes})$
24:             $\nu_m := \nu_{\text{tmp},m} - |\xi_m|^2, \forall m \in \mathcal{K}$                          ▷ $\mathcal{O}(N_\text{B} N_\text{nodes})$
25:             $\Delta_m := \frac{||\mathbf{a}_m||_\text{F}^2}{\bar{\rho}^{-1} + \nu_m}, \forall m \in \mathcal{K}$                              ▷ $\mathcal{O}(N_\text{U} N_\text{B} N_\text{nodes})$
26:             $n := n + 1$, go to line 6
27:         else                                                                   ▷ Pruning
28:             break the loop
29:         end if
30:     end for
31: end if
32: return the final selected beam indices are the entries of $\bar{\text{s}}$
```

BAB searching procedure

The main BAB searching procedure for the minimum-MSE-based selection in hybrid beamforming is similar to that for the maximum-capacity-based selection in antenna selection proposed in Chapter 3. There are several differences between them due to the different monotonicity of the objective functions. Firstly, the bound B is upper bound and initialized as $B = +\infty$ for the former while lower bound initialized as $B = -\infty$ for the latter; secondly, the pruning condition is that the cost value is larger than the bound for the former while less than the bound for the latter; thirdly, the bound is updated when the new cost value at the leaf node is less than the old bound for the former while larger for the latter. The full-array and sub-array beam selection algorithms based on the described BAB searching algorithm are summarized in Alg. 4.2.

Complexity of the BAB based beam selection

The computational complexity of the BAB-search-based beam selection algorithm is also summarized in Alg. 4.2 and the overall complexity of $\mathcal{O}(N_{\text{U}}N_{\text{B}}N_{\text{nodes}})$ is required, where N_{nodes} denotes the number of non-leaf nodes that will be branched in the searching tree. The overall complexity oder is the same to that of BAB based antenna selection algorithms.

4.3.3 Comparisons among greedy, BAB and exhaustive search

The comparison among greedy, BAB and exhaustive search is the same to the comparison of antenna selection algorithms in Subsection 3.4.3. Their performance and complexity comparison for beam selection will be given in Section 4.5.

4.4 Rotman Lens Design and Measurement Results

In the preceding sections, the principles of hybrid beamforming based on the Rotman lens analog beamforming was illustrated. As a proof-of-concept, a realistic Rotman lens should be designed and validated for a practical system. In addition, due to the imperfect hardware characteristics of the practical Rotman lens, e.g., insertion loss, phase errors, etc., it is of great importance to evaluate their impacts on the practical hybrid beamforming performance. Thus we designed and fabricated one sample of Rotman lens. In this section, we will first study the insertion loss of the Rotman lens, and then describe the design of an 8×8 Rotman lens in 5 GHz band, whose measurement results will be also given in this section.

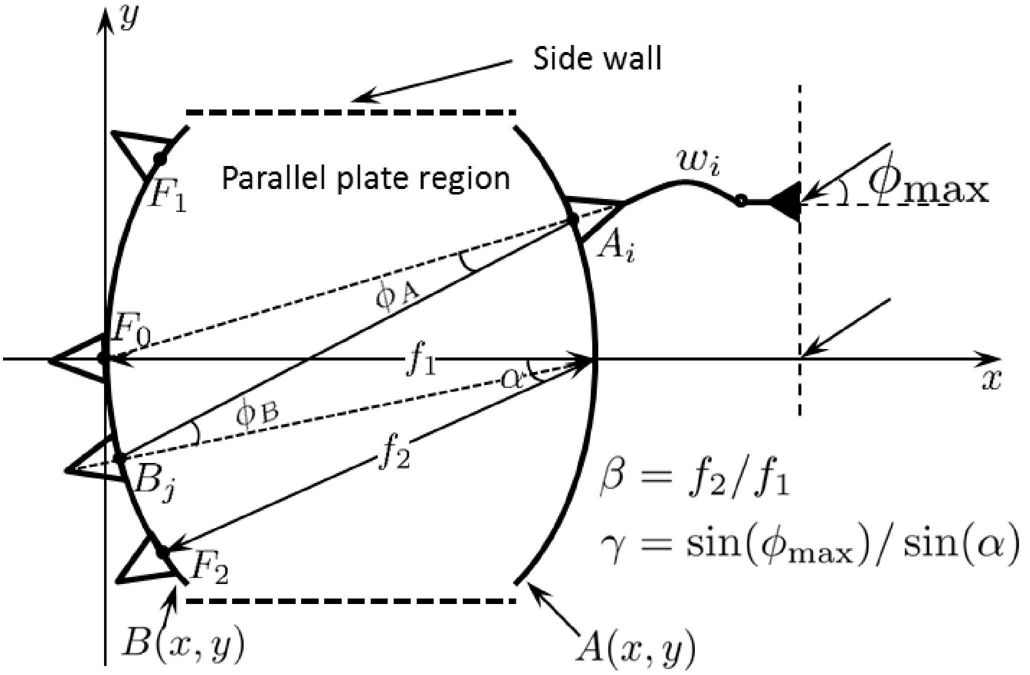


Figure 4.5: Rotman lens design parameters and port coupling. f_1 and f_2 are the on-axis and off-axis focal length, respectively. α and ϕ_{max} are the focal angle and the corresponding maximum sweeping angle, respectively.

4.4.1 Rotman lens design parameters and insertion loss

As aforementioned in Subsection 4.2.3, the determination of the Rotman lens geometry starts from selecting three focal points without phase errors, i.e., the on-axis focal point F_0, and two off-axis symmetric focal points F_1 and F_2, as shown in Fig. 4.5. The three perfect focal points result in three equations according to Equ. (4.1) [81]. By solving this equation set, the parallel plate region, i.e., beam-port contour $B(x,y)$ and array-port contour $A(x,y)$, and the transmission line length w_i, will be determined [81]. The solutions were given in [81,83,84]. Due to the limited space, we will not show them here. Briefly speaking, the solutions are functions of four basic design parameters ($f_1, \beta, \alpha, \gamma$) [83], as shown in Fig. 4.5.

Since the Rotman lens is an analog RF network in the front-end, its insertion loss has a critical impact on the system performance, which should be studied carefully, especially in the MIMO systems with massive number of antennas. Generally speaking, there are two contributors to the insertion loss, i.e., dissipation loss and spill-over loss. The dissipation loss comes from the conductor and the dielectric substrate material of the parallel plate region and transmission lines. For simplicity, we only consider the loss from the parallel plate region. Given transverse electromagnetic wave mode, the conductor loss α_c is given by [88]

$$\alpha_\mathrm{c} = R_\mathrm{s}/(\eta_\mathrm{s} d_\mathrm{s}) \quad \text{Np/m}, \tag{4.21}$$

where R_s is the surface resistance of the conductor; η_s and d_s are the intrinsic impedance and thickness of the substrate material, respectively. The dielectric loss α_d is given by [88]

$$\alpha_\mathrm{d} = k_\mathrm{s}\delta_\mathrm{s}/2 \quad \text{Np/m}, \tag{4.22}$$

where k_s and δ_s are the wavenumber and loss tangent, respectively. Thus given a Rotman lens with the design parameters ($f_1, \beta, \alpha, \gamma$), the transmission distance d_{ji} between the beam port B_j and array port A_i can be determined, and the corresponding signal attenuation coefficient will be calculated by $e^{-(\alpha_\mathrm{c}+\alpha_\mathrm{d})d_{ji}}$. Now let us look at the spill-over loss, which comes from the broad illumination angular range of the beam and array linear tapers, namely, part of the power leaked to the side walls as shown in Fig. 4.5. This loss can be estimated through evaluating the amount of power coupled from the beam ports to the array ports. The signal coupling coefficient from port B_j to A_i was given by [84]

$$\begin{aligned} S_{A_iB_j} =& f(\phi_{A_i}) g\left(\frac{k_\mathrm{s} w_{A_i}}{2}\sin(\phi_{A_i})\right) f(\phi_{B_j}) g\left(\frac{k_\mathrm{s} w_{B_j}}{2}\sin(\phi_{B_j})\right) \\ &\times \sqrt{\frac{w_{A_i} w_{B_j}}{\lambda_\mathrm{s} d_{ji}}} e^{-\jmath(k_\mathrm{s} d_{ji}+\pi/4)}, \end{aligned} \tag{4.23}$$

where λ_s is the wavelength in the substrate material; w_{A_i} is the array-port width; ϕ_{A_i} is the angle between the boresight of the port A and the segment connecting the phase centers of the ports A_i and B_j as shown in Fig. 4.5 (w_{B_j} and ϕ_{B_j} can be defined similarly.); the space factor $g(x)$ is defined as $g(x) = \sin(x)/x$. Different to the coupling formula in [84], we additionally take the element factor $f(x) = (1 + \cos(x))/2$ into account for the aperture size less than one wavelength [101]. Finally, the coupling coefficient from the $\jmath$th beam port to the $\imath$th antenna considering the dissipation loss and the transmission line delay reads

$$\tilde{S}_{A_iB_j} = e^{-(\alpha_\mathrm{c}+\alpha_\mathrm{d})d_{ji} - \jmath k_\mathrm{s} w_i} S_{A_iB_j}, \tag{4.24}$$

and the insertion loss of the beam port B_j can be calculated by summing up the power to all of

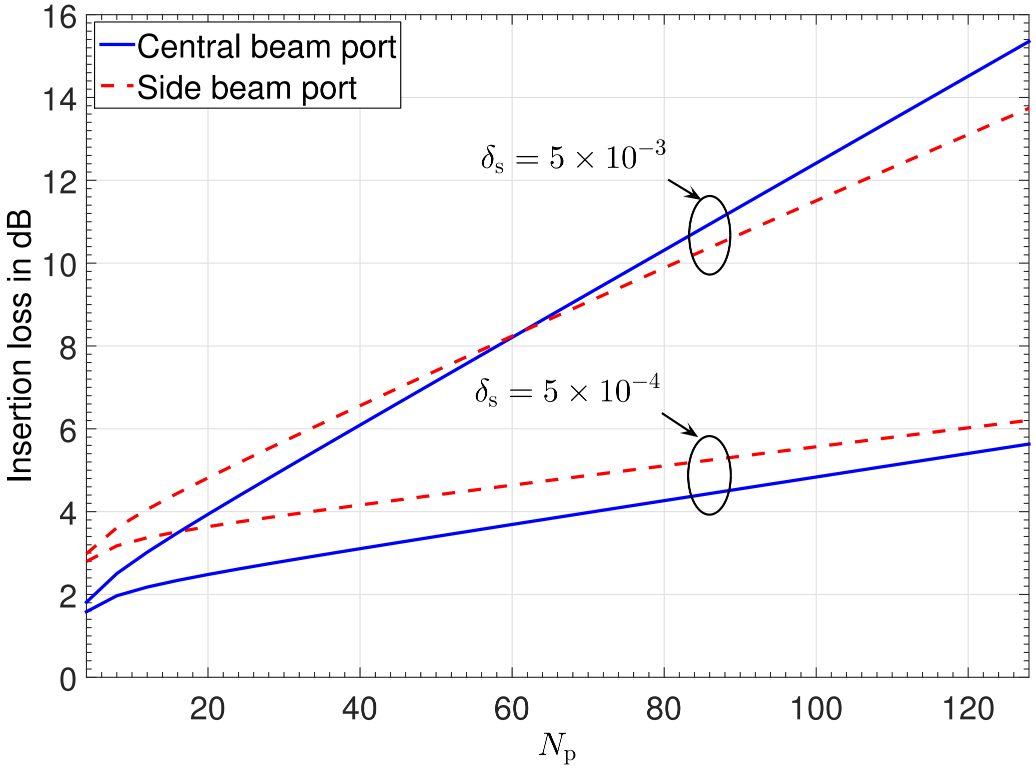


Figure 4.6: Insertion loss versus N_{P}. We fix $\beta = 1/1.2$, $\alpha = 37.35°$, $\gamma = 1.16$ as a rule-of-thumb.

the array ports [86], i.e.,

$$\mathrm{IL}_{B_j} = \sum_{i=1}^{N_{\mathrm{P}}} |\tilde{S}_{A_i B_j}|^2. \tag{4.25}$$

Given the number of antenna elements N_{P}, the array aperture size is given by $\kappa = (N_{\mathrm{P}} - 1)\Delta_{\mathrm{ant}}$. Since f_1 is usually proportional to κ to limit the phase and amplitude errors [83], the physical size of the Rotman lens increases with N_{P} by fixing the antenna spacing Δ_{ant}. Obviously, the analog signal transmission distance d_{ji} also increases with N_{P}, which means that the insertion loss of the Rotman lens increases with N_{P}. To study the increasing rate of the insertion loss with N_{P}, we fix (β, α, γ) and let $f_1 = 0.82\kappa$ as a rule of thumb. The insertion loss of the central beam port lying at the point F_0 and the side beam port lying at the point F_1 can be calculated according to Equ. (4.25), and the numerical results are shown in Fig. 4.6. We can see that the insertion loss in decibel increases *linearly* with N_{P}. The larger the loss tangent δ_{s}, the larger the increasing rate. Given δ_{s}, the increasing rate of the central beam is larger than that of the side beam, which will make the side beam port more advantageous than the central beam port when N_{P} is large. For example, with $\delta_{\mathrm{s}} = 5 \times 10^{-3}$, the side beam port outperforms the central beam port when $N_{\mathrm{P}} > 62$. This is because when the size of Rotman lens is small, the spill-over loss dominates and the central beam port outperforms the side beam port due to the better illumination; when the size of Rotman lens is large, the dissipation loss dominates and the side beam port outperforms the central beam port due to the less distance to the array contour. Interestingly, when N_{P} is not too large, e.g., $N_{\mathrm{P}} = 8$, the insertion loss is relatively small with respect to different beam port locations and δ_{s} values. Thus in the following we will design an 8×8 Rotman lens. It should be noted that the ideal array gain, i.e., $10\log_{10}(N_{\mathrm{P}})$, only increases *logarithmically*, which means that we *cannot* always benefit from increasing N_{P} in the view of practical array gain. A straightforward conclusion from Fig. 4.6 is that the sub-array architecture is more competitive than the full-array architecture for the Rotman lens-based massive MIMO systems, since the increasing rate of the insertion loss of the former is only one-Lth of the latter. Take 8×8 ($N_{\mathrm{P}} = 8$) and 32×32 ($N_{\mathrm{P}} = 32$) Rotman lenses with $\delta_{\mathrm{s}} = 5 \times 10^{-3}$ as a comparison example. From Fig. 4.6 we can see that the former has the insertion losses of 2.5 and 3.6 dB for the central and side beam ports respectively; while 5.2 and 5.9 dB of the latter.

4.4.2 Design

We specify the operating frequency of the designed Rotman lens in the 5 GHz band with the center carrier frequency of 5.5 GHz. The reason to specify this frequency band is three-fold. First, this band is of great importance for the WLAN systems, e.g., IEEE 802.11n, etc., and also for the future cellular communication systems, e.g., LTE-U, etc; second, the channel of this frequency band is relatively sparse, making the hybrid beamforming system based on Rotman lens feasible and effective; third, the low-cost RF switches in this frequency band are off-the-shelf. Another important design issue is on the number of beam and antenna ports of the Rotman lens. Considering the physical size and also the insertion loss as shown in Fig. 4.6, we fix the Rotman lens with 8 beam ports and 8 antenna ports for each sub-array. The 8 beams are evenly distributed in the scanning range $[-45°,\ 45°]$ with approximately $11°$ angle resolution. This configuration of the scanning range corresponds to the scenario that the coverage of the BS is sectorized into 4 sectors of $90°$ each. Note that the Rotman lens with larger scanning range is possible to be designed, but the insertion loss will be increased dramatically, which will severely degrades the performance.

The design of the Rotman lens is carried out by the software Rotman Lens Designer (RLD), in which the Rotman lens design parameters, e.g., $(f_1, \beta, \alpha, \gamma)$, can be tuned to optimize the network performance in terms of beam pattern, insertion loss, side lobe level, phase and

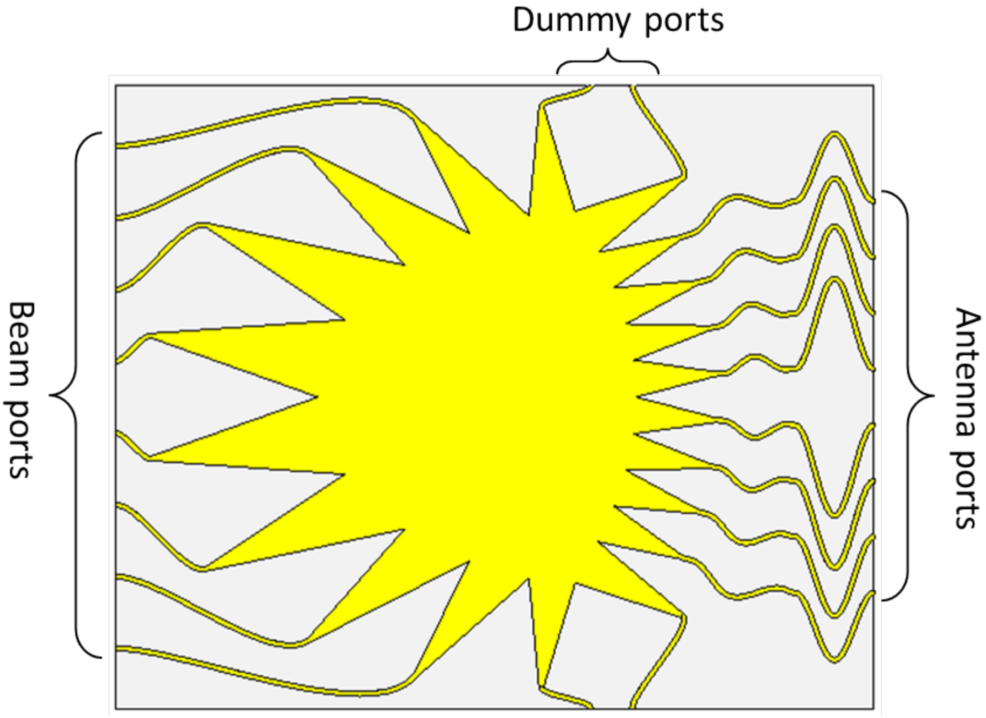


Figure 4.7: The layout of the designed 8×8 Rotman lens

amplitude errors. In our design, the beam ports are co-located on the *circular* beam port contour and the flaring angles are kept less than 15° to minimize the reflections from the open boundary beam-port contour [85]. The array contour and all the connecting microstrip transmission lines are calculated automatically by RLD. The transmission lines are routed with the appropriate curvature and separation to guarantee smoothness and non-overlapping. The boresight of each linear taper points to the center of the opposite contour to decrease the insertion loss. The layout of the designed 8×8 Rotman lens is shown in Fig. 4.7. Note that we use dummy ports connected with dummy loads to absorb the reflected waves on the side walls.

It should be noted that RLD is based on geometrical optics, which analyzes the EM wave distribution in the Rotman lens layout with ray tracing technology, resulting in very rapid synthesis and analysis of the design, but at the cost of accuracy. Thus the designed Rotman lens should be validated via more accurate EM simulation tools. In this thesis, we utilized CST MICROWAVE STUDIO with the Frequency Domain Solver, whose simulation results will be shown and compared with the RLD results in the next subsection.

After the layout design and validation with RLD and CST, we fabricated one sample of Rotman lens, whose substrate, e.g., permittivity, thickness, etc., should be chosen carefully to compromise the physical size, phase errors and manufacturing errors. Generally speaking,

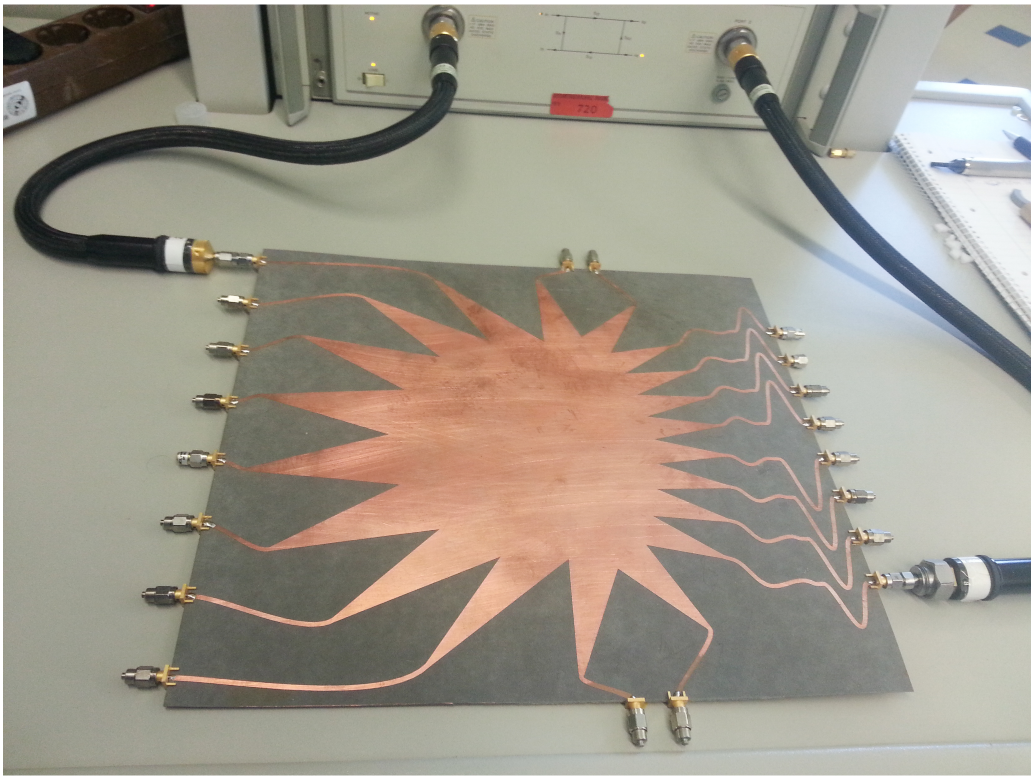
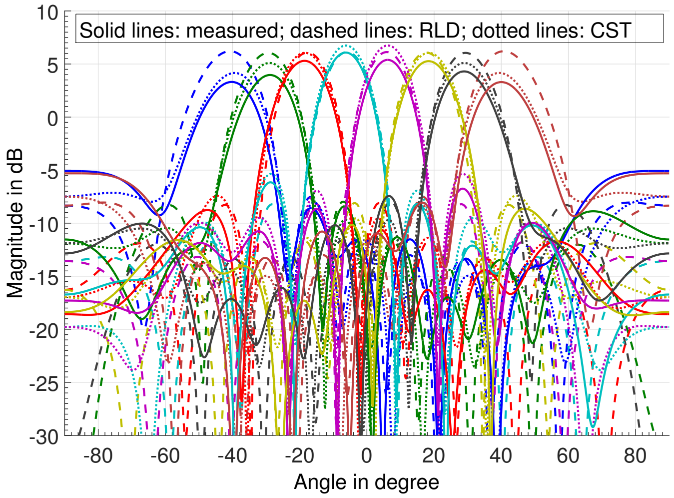


Figure 4.8: Fabricated Rotman lens and measurement with VNA

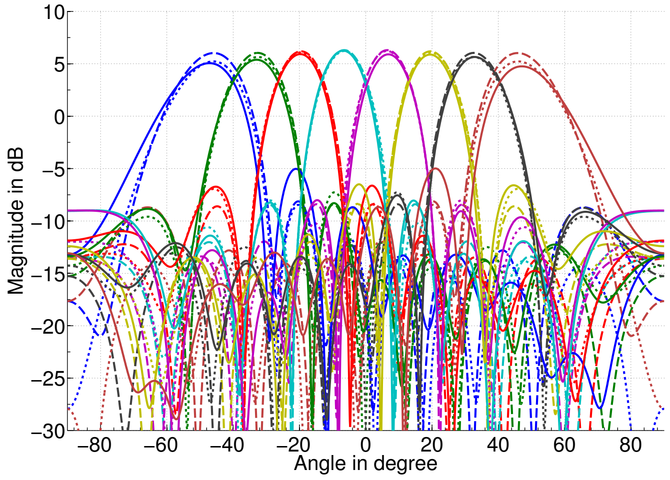
the higher the substrate permittivity, the smaller the Rotman lens physical size. However, the high substrate permittivity generally results in large phase and manufacturing errors. For the thickness of the substrate, it is known that the lens size decreases as the substrate thickness decreases. However, the thinner the substrates, the narrower the connecting transmission lines, and consequently the more sensitive to the manufacturing errors. Finally, the Rogers RO5870 substrate with the relative permittivity $\epsilon_{\mathrm{r}} = 2.33$ and thickness $d_{\mathrm{s}} = 0.768$ mm is utilized for the fabrication. The fabricated Rotman lens is of size 310×344 mm shown in Fig. 4.8.

4.4.3 Measurement results

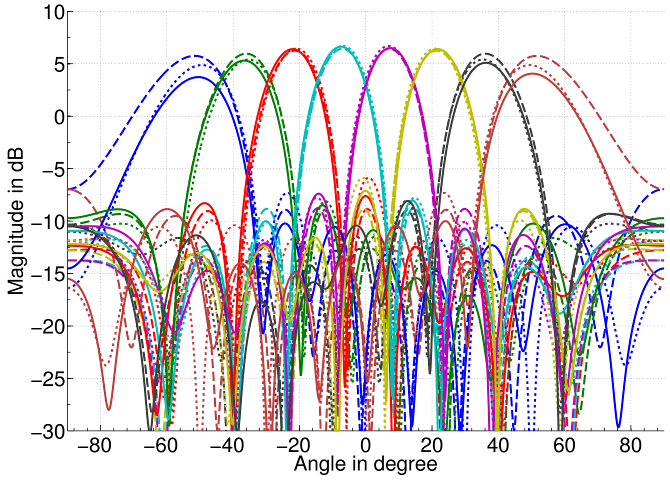
We measured the scattering parameters, specifically the transmission coefficients $[\mathbf{F}_{\text{RM,sub-array}}]_{mn}$ between the mth beam port and the nth antenna port, $m, n = 1, 2, \cdots, 8$, of the fabricated Rotman lens with Vector Network Analyser (VNA) as shown in Fig. 4.8. The dummy loads are connected to all the ports except the two ports under measurement. Supposing that the ideal isotropic antenna elements are used, the beam patterns of the Rotman lens can thus be directly calculated from $\mathbf{F}_{\text{RM,sub-array}}$, and shown in Fig. 4.9. For comparison, the beam patterns obtained by the RLD and CST are also shown in the figure. We can see that the $M = 8$ beams are evenly



(a) 5.0 GHz

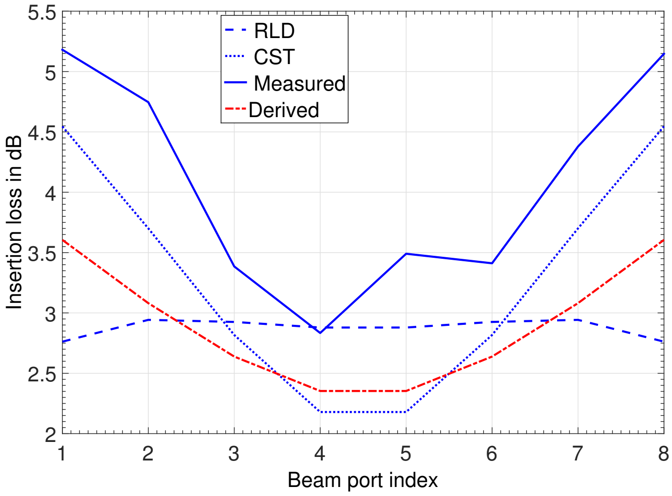


(b) 5.5 GHz

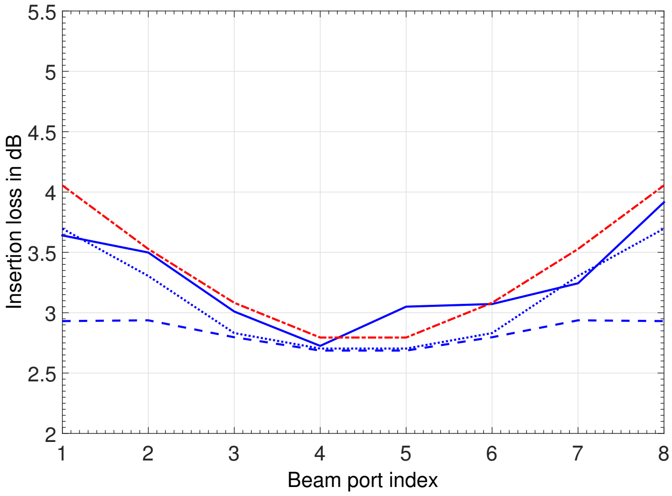


(c) 6.0 GHz

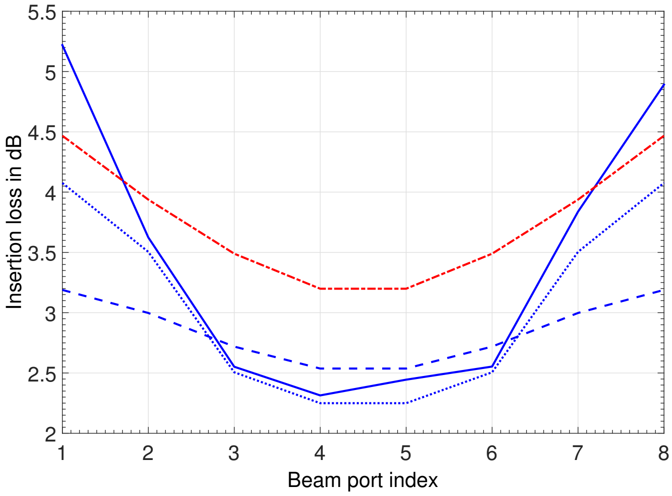
Figure 4.9: Beam patterns of the designed Rotman lens with different carrier frequencies.



(a) 5.0 GHz



(b) 5.5 GHz



(c) 6.0 GHz

Figure 4.10: Insertion loss of the designed Rotman lens with different carrier frequencies.

distributed in the range $[-45°, 45°]$. More importantly, the beam patterns generated by RLD, CST, and the measurements agree with each other very well, especially for the central beams. Note that at the boundary frequencies, i.e., 5 and 6 GHz, the measured side beam gains are significantly less than RLD simulated ones. The array gains are distributed in the range $[3, 6]$ dB depending on the beam port indices and the carrier frequencies. The side beams have less gains than the central beams due to the worse illumination.

Following Equ. (4.25), we can also calculate the insertion loss the fabricated Rotman lens, as shown in Fig. 4.10. In addition, we also show the results obtained by RLD, CST and the derivation given in Section 4.4.1. Note that for the derived insertion loss, we also take the transmission line loss into account. We can observe that RLD gives relatively flat insertion loss across the beam ports, which has large errors especially for the side beam ports. In contrast, CST gives the insertion loss nearer to the measurement results. The derived insertion loss agrees well with the measured results at the design frequency of 5.5 GHz, within 0.5 dB error, but has significant mismatch to the measurement results at the boundary frequencies. The measured insertion loss ranges from 2.7 to 3.9 dB in our design frequency of 5.5 GHz, becoming worse at the boundary frequencies, i.e., 5 and 6 GHz, ranging from 2.25 to 5.25 dB. Note that the side beams have larger insertion loss than the central beams, corresponding to the observation of the lower array gain of the side beams than the central beams shown in Fig. 4.9.

4.5 Simulation Results

In this section, we evaluate the Bit Error Rate (BER) performance of the hybrid beamforming systems. Given an $N_{\mathrm{U}} \times N_{\mathrm{B}}$ massive MIMO system, we assume $N_{\mathrm{U}} = 4$ independent uncoded data streams with 16-QAM modulation are transmitted from $N_{\mathrm{U}} = 4$ single-antenna users. The transmitted signals go through the channels with multipath fading. We assume that each user channel has $P_u = 8$ non-line-of-sight paths with independent Rayleigh fading, whose AoAs follow the uniform distribution within $[-45°, 45°]$, corresponding to the scenario of the roof-top BS antennas responsible for one sector. In the receiver, an ULA of N_{B} antennas with half wave-length spacing is assumed. For the full-digital $N_{\mathrm{U}} \times N_{\mathrm{B}}$ large-scale system, the signals from N_{B} antennas will be directly processed with an LMMSE digital beamformer. For the hybrid beamforming systems, we simulate our systems based on Rotman lens with both full-array and sub-array architectures. For the full-array architecture, we assume a single $N_{\mathrm{B}} \times N_{\mathrm{B}}$ large-size Rotman lens is used; while for the sub-array architecture, we divide the N_{B} antennas into

$L = N_{\mathrm{U}} = 4$ sub-arrays and each sub-array connected with one $M \times M$ small-size Rotman lens. As a benchmark purpose, we also simulate a full-digital $N_{\mathrm{U}} \times L$ small-scale MIMO system, which has the same number of RF transceivers to the hybrid beamforming system. Note that in the full-digital small-scale MIMO system, the L receive antenna signals are picked up from the whole N_{B} antennas in every M antennas. In all of the simulations, we assume that the channel state information and the noise variance are perfectly known by the receiver.

To compare the performance of the static beamspace system based on Rotman lens to that of the dynamic beamspace system, we also simulate a state-of-the-art hybrid beamforming system using dynamic beamspace [73], which is based on a so-called Successive Interference Cancellation (SIC) method, in which the phase values of the phase shifters in each sub-array are calculated successively aiming to maximize the system achievable date rates. Note that the original work in [73] mainly focuses on the analog beamforming to decouple the data streams, while the digital beamforming is simply a diagonal matrix for power allocation. We follow the analog beamforming method of [73], but additionally use the LMMSE digital beamformer to decouple the data streams further. Also note that the SIC method in [73] was specifically proposed for the sub-array hybrid beamforming system, but actually it is also applicable to the full-array system with minor modifications. In addition, different to the non-ordered sub-array SIC in [73], we use optimal SIC ordering of sub-arrays via exhaustive search, since the optimally ordered SIC outperforms the non-ordered SIC with $1 \sim 2$ dB gain through our simulation.

4.5.1 Comparison study with ideal RF hardware

In this subsection we assume that all the RF components, e.g., Rotman lens, phase shifters, etc., are ideal without loss and phase errors. The simulation results on BER performance is shown in Fig. 4.11 with $N_{\mathrm{B}} = 32, 128$. In general, the performance of hybrid beamforming systems is in-between the full-digital large-scale and small-scale systems. To compare the performance of the hybrid beamforming systems, we conduct three comparison studies, i.e., dynamic versus static beamspaces, full-array versus sub-array architectures, and comparison among different beam selection algorithms.

Comparing the two systems, i.e., the static beamspace (using Rotman lens) based system with the optimal BAB-search-based beam selection algorithms and the dynamic beamspace (using phase shifters) based system with SIC algorithms, the former has nearly 4 dB performance loss compared to the latter at the BER of 10^{-3} for the full-array architecture with $N_{\mathrm{B}} = 32$.

The performance gap increases slightly up to less than 5 dB as N_B increases from 32 to 128. In contrast, for the sub-array architecture, the performance gap between these two systems is 2 dB for the case of $N_\mathrm{B} = 32$, but increases noticeably as N_B increases (e.g., 4 dB gap when $N_\mathrm{B} = 128$). Interestingly, the slopes of the BER curves of the two systems are very close to each other, both of which are close to the slope of the full-digital massive MIMO system. This implies that the hybrid system can almost retain the diversity gain of the full-digital system.

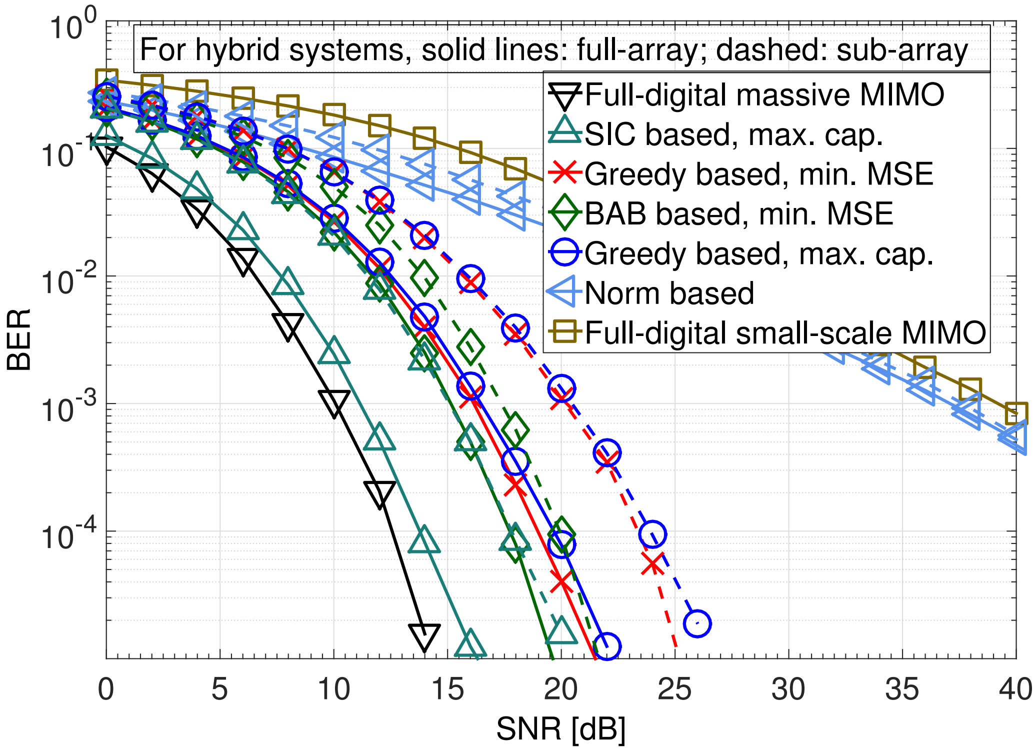
Comparing the full-array and sub-array architectures, we can observe that the sub-array architecture system has approximately $4 \sim 5$ dB loss in SNR compared to the full-array one for the dynamic beamspace systems. For the static beamspace systems, the sub-array architecture system still performs worse than the full-array system, but the performance loss is less than that with dynamic beamspace system when $N_\mathrm{B} = 32$ (i.e., 2.3 dB versus 4 dB), and the loss is comparable when $N_\mathrm{B} = 128$ (i.e., 4 dB versus 5 dB). This implies that it is very cost-effective to split the full-array into sub-array architectures for the Rotman-lens-based hybrid beamforming systems, especially when N_B is not too large.

Now let us compare the beam selection algorithms used in the static beamspace system. In addition to the proposed greedy and BAB-search-based beam selection algorithms, we also simulate the conventional norm-based method and a state-of-the art beam selection based on greedy search aiming to maximize the channel capacity [79]. Note that in the full-array norm-based method, after selecting the strongest beam port, we exclude its two neighbouring beam ports in the candidate set for next round beam-port selection. This procedure will repeat until L beam ports are selected. The reason of excluding neighbouring beam ports is due to the beam port power leakage issue [102]. We can observe from Fig. 4.11 that the norm-based method has very poor performance. The proposed greedy method aiming to minimize the MSE has superior performance over the norm-based method, and also has slightly better performance than the method of [79]. The performance of the proposed greedy method is very near to the optimality for the full-array architecture systems, and is sub-optimal for the sub-array architecture systems. Interestingly, as N_B increases, the performance gap between the greedy and the BAB-search-based method decreases. For example, the performance gap is approximately 1.6 dB when $N_\mathrm{B} = 128$, which is nearly half of that when $N_\mathrm{B} = 32$. In addition, we can also observe from Fig. 4.11 that all the hybrid beamforming systems have superior performance over the full-digital small-scale MIMO system. For example, the sub-array beamforming system based on Rotman lens with the greedy search beam selection, which has the worst performance in all the hybrid beamforming systems, has almost 20 dB gain over the small-scale MIMO. This impressive gain comes from the analog beamforming gain and the beam selection diversity gain.

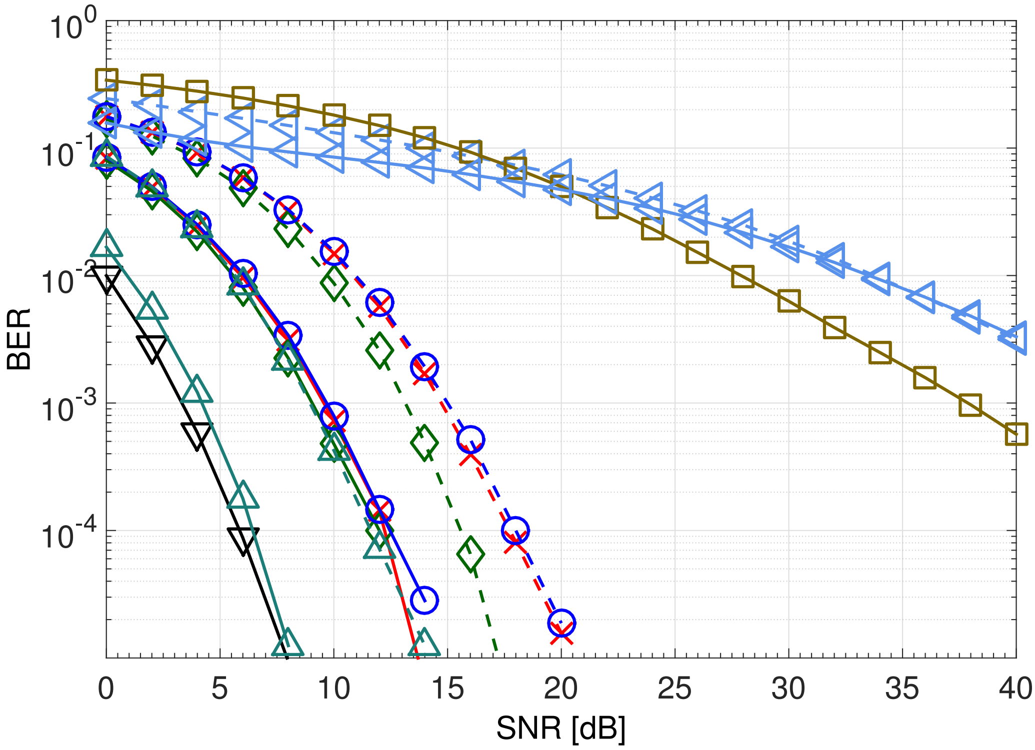
Regarding the complexity of the beam selection algorithms, the norm based method only requires the calculation the norm of each element of the channel matrix, and hence it has the complexity of $\mathcal{O}(N_{\mathrm{U}}N_{\mathrm{B}})$. In contrast our greedy method has the complexity of $\mathcal{O}(N_{\mathrm{U}}N_{\mathrm{B}}L)$, which is only slightly higher than the norm based method when L is not too large. At last, the greedy method in [79] has the complexity of $\mathcal{O}(N_{\mathrm{B}}^{3})$, which is obviously much higher than ours. From the comparison on the performance and complexity, we can see that our proposed low-complexity greedy-search-based beam selection algorithm is very effective for the practical applications.

4.5.2 Comparison study with non-ideal RF hardware

In the last subsection, we study our system performance with ideal Rotman lens. But the question is: how much gain can we still achieve if the practical Rotman lens is used rather than the ideal one? In this subsection, we turn our attention to the hybrid beamforming systems with non-ideal RF hardware. In this study we only focus on the sub-array architecture with $N_{\mathrm{B}} = 32$, $L = 4$, and $N_{\mathrm{U}} = 4$. For the static beamspace hybrid beamforming system based on Rotman lens, we use the measurement results of the fabricated Rotman lens, i.e., transmission coefficient matrix $\mathbf{F}_{\text{RM,sub-array}}$, to perform the simulation (we assume that the 4 Rotman lenses are identical, and all of them can be described by our measured matrix.); while for the dynamic beamspace hybrid beamforming system based on phase shifters, we model the state-of-the-art 5-GHz-band phase shifters MAPS-011008 [97] with 4 dB insertion loss, 6 bits phase resolution, 2.2° and 1.0° root-mean-square-phase-error at 5 and 6 GHz, respectively. To study the wideband characteristics, we perform the simulations with two boundary carrier frequencies, i.e., 5 GHz, and 6 GHz. Fig. 4.12 shows the simulation results. We can see that at the BER of 10^{-3}, the system with the practical Rotman lens has approximately 4 dB performance loss compared to the system with the ideal Rotman lens, which is consistent with the measured insertion loss shown in Section 4.4. The hybrid beamforming system with practical phase shifters has also 4 dB loss compared to the system with the ideal phase shifters, which is due to the 4 dB insertion loss of the phase shifters. The system based on Rotman lens with the optimal BAB search has about 2 dB loss compared to the system based on phase shifters at the BER of 10^{-3}. In addition, the system based on Rotman lens has superior performance over the small-scale MIMO system, e.g., about $15 \sim 16$ dB gain for the system with greedy search, validating the advantage of our



(a) $N_\mathrm{B} = 32$



(b) $N_\mathrm{B} = 128$

Figure 4.11: BER performance with ideal RF components with respect to different N_B. N_U is fixed to be 4.

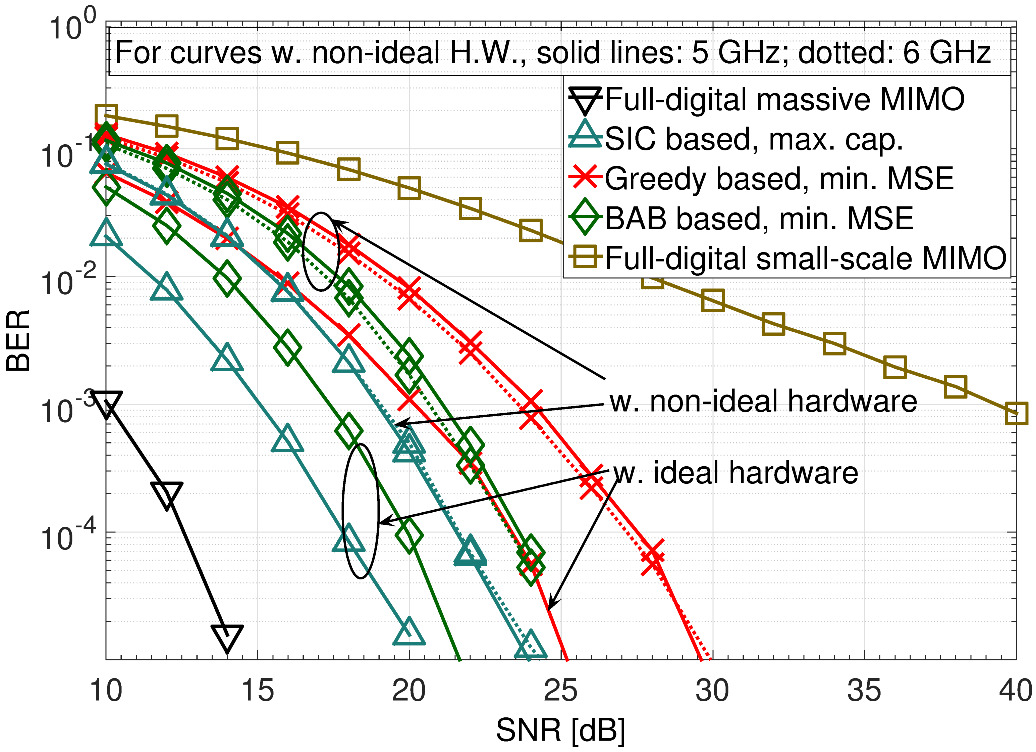


Figure 4.12: BER performance of the sub-array hybrid beamforming system: ideal versus non-ideal RF hardwares. For the curves with non-ideal RF hardwares, two carrier frequencies are used, i.e., 5 and 6 GHz

hybrid system. More importantly, like the system based on phase shifters, our system also has very similar error-rate performance at the two boundary carrier frequencies, implying the flat performance in the designed 5-GHz band and validating the wideband feature of our hybrid beamforming system.

Remarks. From the simulation results in the subsections 4.5.1 and 4.5.2 we can observe that the performance of the hybrid beamforming system based on Rotman lens is comparable to that based on phase shifters. However, the cost of the Rotman lenses is much lower than that of the phase shifters. Take the fabricated 8×8 Rotman lens and MAPS-011008 phase shifters as a comparison example. The cost of the former is only approximately 1/3 of the latter. Therefore, it can be concluded that the system based on Rotman lens is more competitive than the system based on the conventional phase shifters.

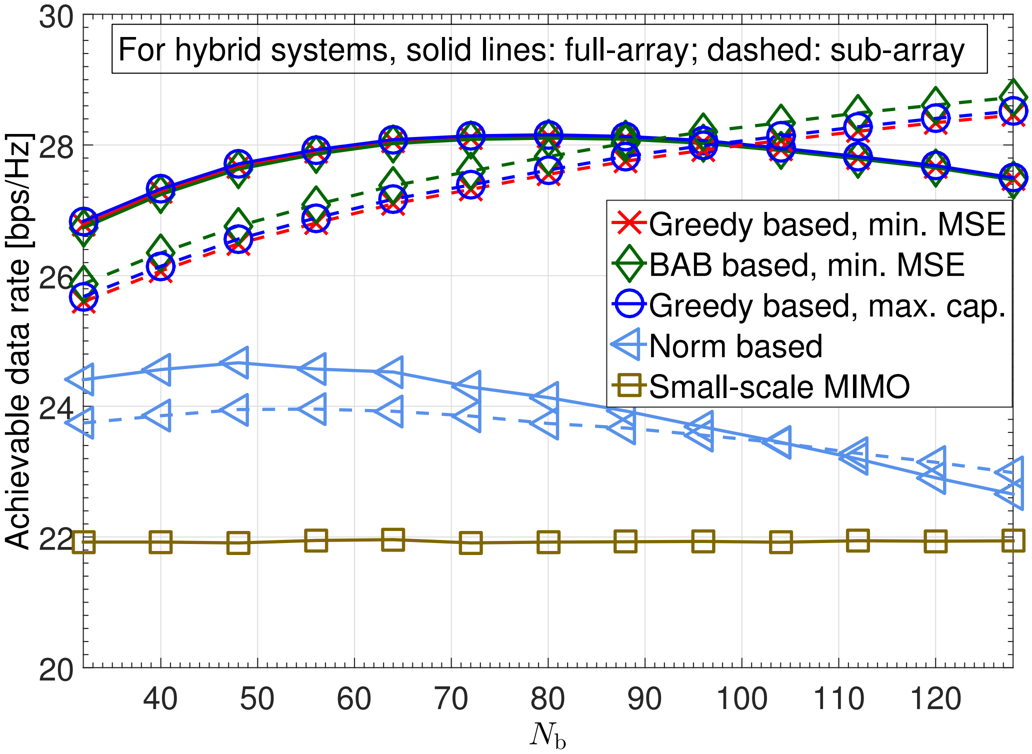


Figure 4.13: Achievable data rate versus N_B. $N_\mathrm{U} = 4$, $\rho = 20$ dB.

4.5.3 Data rate comparison between full-array and sub-array considering insertion loss

As aforementioned in Section 4.4.1, the the sub-array architecture potentially performs better than the full-array on the array gain when considering insertion loss with large N_B. Let us use the derived S-parameters in Equ. (4.24) with a moderate loss tangent $\delta_\mathrm{s} = 2 \times 10^{-3}$ to model the Rotman lens. Since this model is most accurate at the carrier frequency of 5.5 GHz as shown in Fig. 4.10 (b), we fix the carrier frequency to be 5.5 GHz. Fig. 4.13 shows the achievable data rate versus N_B. The data rate of the full-array systems first increase and then decrease due to the fast growth of the insertion loss. In contrast, the sub-array systems with greedy and BAB search methods increase monotonically in the range of N_B of our interest. And indeed, the sub-array performs better than the full-array when N_B is large. Another interesting observation is that, for the full-array systems, both greedy search methods has indistinguishable performance from the BAB search; while for the sub-array systems the greedy methods have noticeable degradation to the BAB search. In addition, the Rotman lens system with greedy or BAB search has superior performance over the norm-based hybrid system and the small-scale MIMO system.

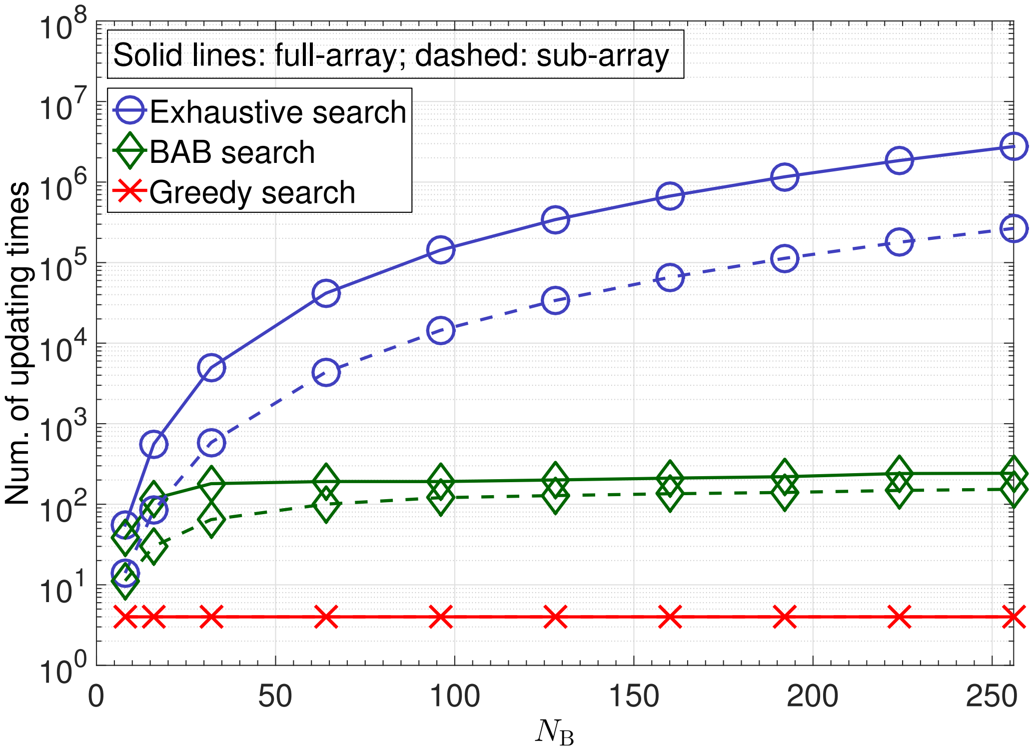
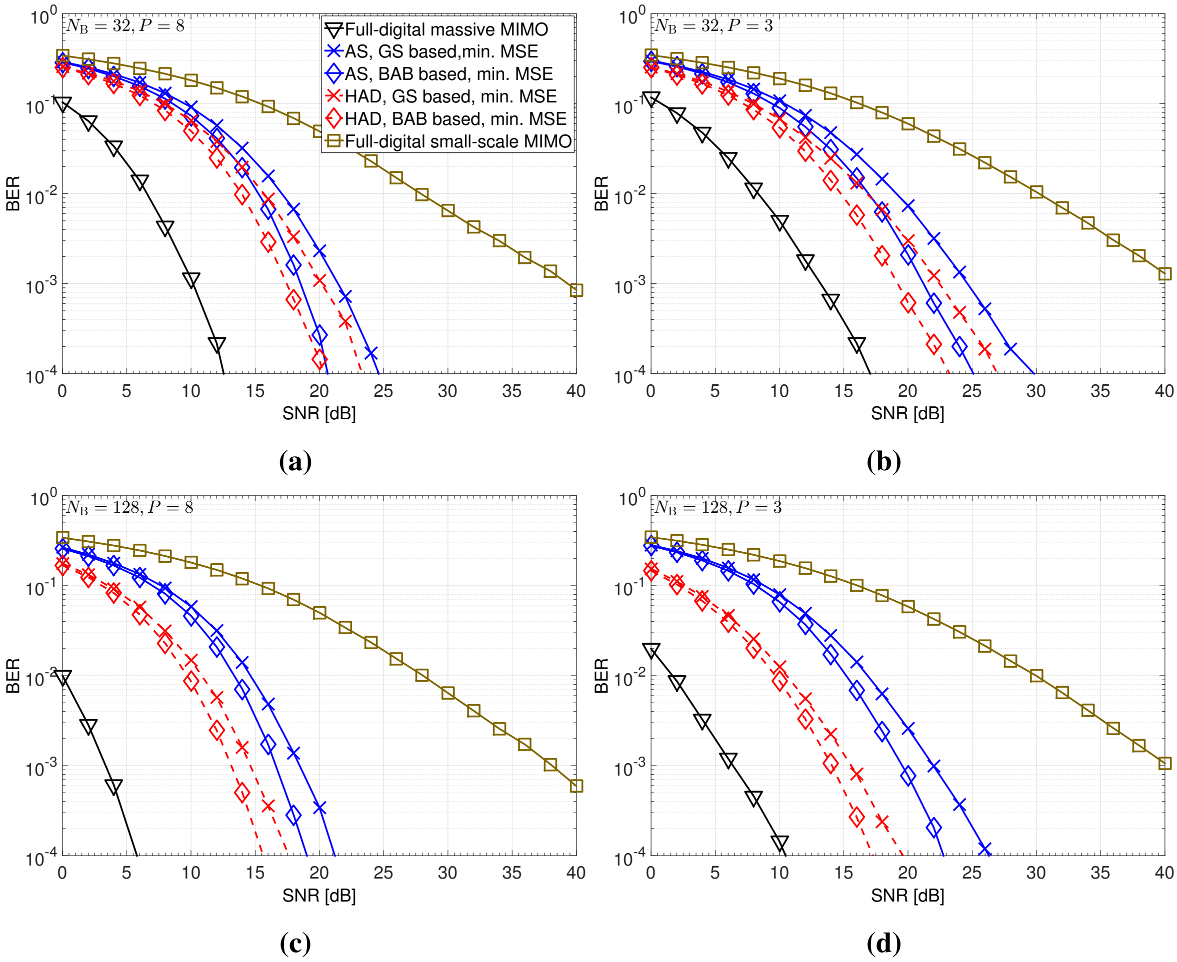


Figure 4.14: The number of visited nodes requiring update operations versus N_B. We fix $N_\mathrm{U} = 4$ and $\rho = 20$ dB. Note that the complexity of the BAB search at other high SNRs is similar to the case of $\rho = 20$ dB, thus not necessarily shown here.

4.5.4 Computational complexity comparison

Now let us study the computational complexity of the proposed beam selection algorithms. We assume that the ideal Rotman lenses are used for the simulation. Fig. 4.14 shows the simulation results of the BAB search on the number of visited nodes asking for updating operations versus N_B. The numerical complexity given in Section 4.3.3 for the exhaustive search and greedy search is also shown in the figure. We can see that the BAB search has a complexity in-between the greedy search and exhaustive search. For both exhaustive search and BAB search, the sub-array system has lower complexity than the full-array system, due to the less number of candidate beam subsets of the sub-array system. But it should be noted that the gap between the full-array and sub-array system for the BAB search is much smaller than that of the exhaustive search. Most interestingly, the increasing rate of the BAB search is much slower than that for the exhaustive search. As N_B increases, the BAB search becomes more advantageous than the exhaustive search. For example, with $N_\mathrm{B} = 256$, the complexity of the BAB search is reduced by 4 and 3 orders of magnitude compared to the exhaustive search for the full-array and sub-array systems, respectively.



(a) (b) (c) (d)

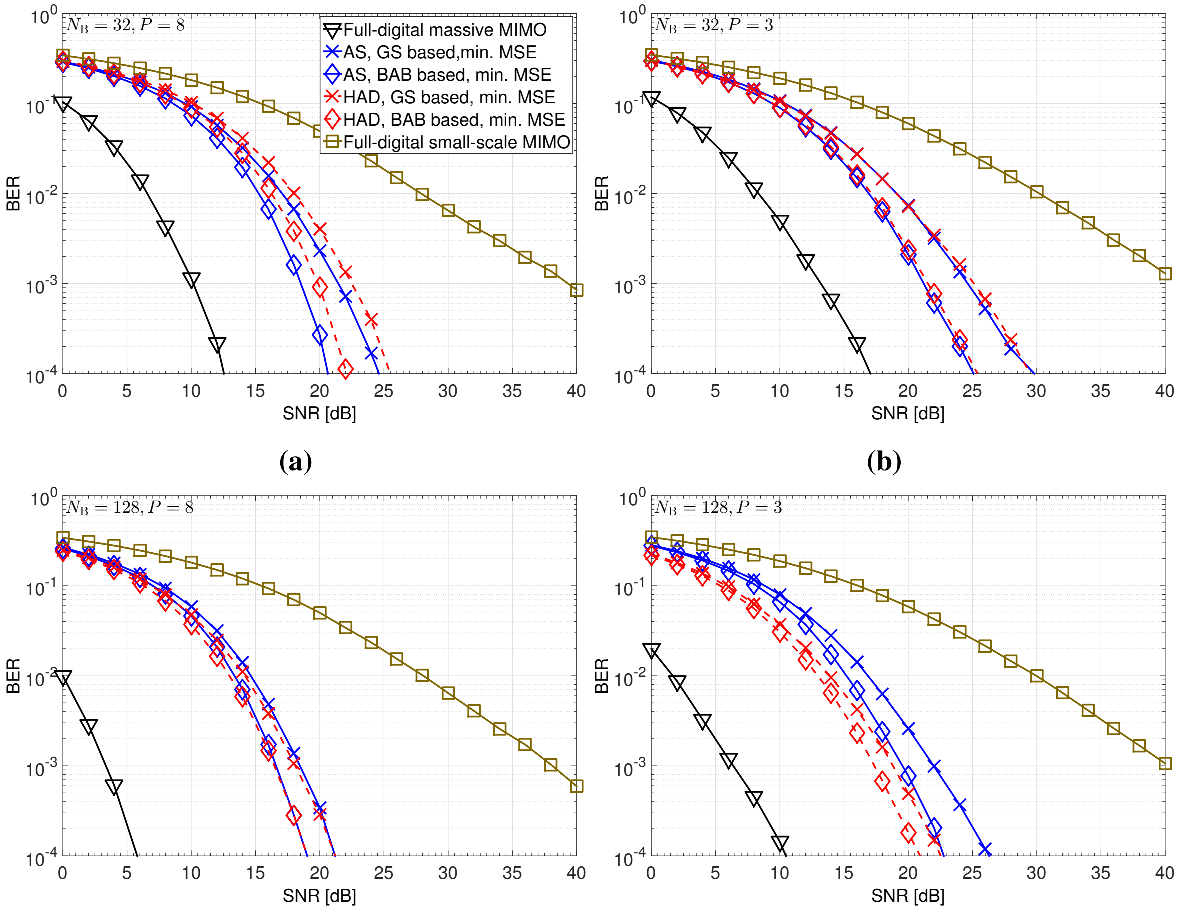
Figure 4.15: BER performance comparison of Antenna Selection (AS) and Hybrid Analog-Digital (HAD) beamforming with *ideal* hardware. $N_U = 4$, and 4 sub-arrays.

4.5.5 Comparison of antenna selection system and Rotman-lens-based hybrid beamforming system

As aforementioned, the Rotman-lens-based hybrid beamforming system is conceptually similar to the antenna selection system. The former has additional Rotman lenses in-between the antennas and RF switches to allow selecting beams instead of selecting antennas. The interesting questions are: which system performs better and can we benefit from the extra Rotman lenses? This subsection is to compare their BER performance with Monte-Carlo simulations to answer these questions. For the following comparison, we only focus on the sub-array systems due to their advantage on the practical feasibility over the full-array systems.

Let us first have a look at the systems with the ideal hardware that no loss exists in the RF switches and Rotman lenses. Fig. 4.15 shows the simulation results. The selection criterion based on the minimum MSE is used for the both the antenna and beam selection algorithms. The results with the maximum-capacity-based criterion are similar, thus not necessarily shown here. Two types of radio environment are considered, i.e., relatively rich-scattering channels and very sparse channels. The number of physical propagation paths P_u in the channel model given by Equ. (4.5) is used to measure the channel sparsity. $P_u = 8$ denotes relatively rich-scattering channels; while $P_u = 3$ denotes very sparse channels. From Fig. 4.15 (a) we can see that, the hybrid beamforming system has minor performance advantage over the antenna selection system when $N_\mathrm{B} = 32$ and $P_u = 8$ for both selection algorithms. By comparing Fig. 4.15 (a) and (b) we can see that the advantage of the hybrid beamforming system over the antenna selection system is increased with less number of channel paths. By comparing Fig. (a) and (c) we can see that the hybrid beamforming system is obviously more advantageous over the antenna selection system as N_B increases. The largest performance gain of the hybrid beamforming system over the antenna selection system can be observed when increasing N_B and meanwhile decreasing P_u, which is shown in Fig. 4.15 (d). About 6 dB SNR gain can be observed for both BAB search and greedy search. Thus we can conclude that the Rotman-lens-based hybrid beamforming system is more favorable in the large-scale array with very sparse channels. Such condition can be foreseen in the future mmWave systems. In contrast, the antenna selection system is more economic for the rich scattering channels and MIMO systems without too many antennas, such as the conventional cellular systems with carrier bands below 3 GHz.

The Rotman lenses and RF switches are lossy RF components in the front-end, which will degrade the performance of the whole systems. Thus it is also necessary to compare the antenna selection system and the Rotman-lens-based system with non-ideal hardware. Since both systems have the same RF switches, we drop the impact of RF switches and only assume non-ideal Rotman lenses. The non-ideal Rotman lens model in Equ. (4.24) with a low loss tangent $\delta_\mathrm{s} = 5 \times 10^{-4}$ is used to model the non-ideal Rotman lens. Fig. 4.16 shows the simulation results with the same configurations to those in Fig. 4.15. We can see that, for all of the four configurations, the SNR gain of the Rotman lens systems over the antenna selection systems is decreased as compared to the cases of the ideal Rotman lenses shown in Fig. 4.15. For example, in Fig. 4.16 (a) the hybrid beamforming system has even worse performance than the antenna selection system. But when the scale of MIMO is large enough and the channels are sparse enough, we can still observe some performance gain as shown in Fig. 4.15 (d). Thus we can conclude that the insertion loss of the Rotman lenses is critical in the Rotman-lens-based system. To exhibit the performance advantage over the antenna selection systems, very low-loss Rotman



(a) **(b)**

(c) **(d)**

Figure 4.16: BER performance comparison of Antenna Selection (AS) and Hybrid Analog-Digital (HAD) beamforming with *non-ideal* hardware. $N_\mathrm{U} = 4$, and 4 sub-arrays.

lenses are necessary. One way is to consider very low-loss conductors and substrate materials for fabrication; another way is that we can try to optimize the geometry of the Rotman lenses further.

4.6 Conclusion

In this chapter, we propose a low-cost wideband hybrid analog-digital beamforming system based on Rotman lens. We use the low-cost Rotman lens to replace the conventional high-cost phase shifters, and study the system with both full-array and sub-array architectures. With the Rotman lens analog beamforming and LMMSE digital beamforming, the system optimization problem is formulated as a beam selection problem to minimize MSE. The two proposed beam selection algorithms, i.e., greedy search and BAB search, for both full-array and sub-array architectures exhibit trade-off between the performance and the complexity. The greedy search method has low-complexity but sub-optimal performance, which is preferred for practical implementations; while the BAB search method has optimal performance but higher complexity than the greedy search. Compared to another well-known optimal method, i.e., exhaustive search, the BAB search can reduce the complexity several orders of magnitude, which is of great importance for the large-scale MIMO. To validate the system, we have designed and fabricated a sample of Rotman lens for the sub-array architecture system, whose measurement results were used to simulate the system error-rate performance. The measurement and simulation results show that the system based on low-cost Rotman lens can achieve comparable performance to the system based on the high-cost phase shifters. In addition, our system also shows wideband capability and superior performance over the small-scale MIMO systems with the same number of RF transceivers. When considering realistic hardware loss, the Rotman-lens-based hybrid beamforming performs better than the antenna selection only when N_{B} is large and channel is very sparse. As a conclusion, the hybrid beamforming system based on Rotman lens is more competitive than the system based on phase shifters and the small-scale MIMO systems, and it is more competitive than the antenna selection only in the very large-scale systems and sparse channels.

Chapter 5

Testbed Performance Evaluation

In Chapter 3 and 4, the antenna selection and Rotman-lens-based hybrid analog-digital beamforming systems are validated via Monte-Carlo simulations. Whether the systems can work in practice is unknown. Additionally, in the simulation study we make many assumptions for the convenience of system design. For example, we assume that the knowledge of channel matrix, SNR, timing offset, etc., is perfectly known, which is actually unknown in practice, and needs to be estimated by implementing some estimation algorithms. Whether the implementation error has severe impact on the system performance is a great concern for practical applications. Another concern on the simulation study is the channel model. For the convenience of study, simple statistical channel models are used. It should be noted that the statistical channel model cannot fully characterized the realistic radio environment. Since the channel characteristic has a great impact on the performance of wireless systems, the practical channels get very necessary to verify the antenna selection and Rotman-lens-based systems. More interestingly, it is also necessary to verify both systems with realistic wideband signals, since many practical systems, such as WLAN, LTE, etc., are wideband, which means that the selection algorithms for the narrowband system used in Chapter 3 and 4 have to be tailored for the wideband system. Thus in order to verify the both low RF-complexity systems under the scenario of the existing estimation errors, realistic channels and wideband signals, testbeds are necessary to be built up for experimental study. In this chapter, the details of the built-up testbeds of the both systems are described. The full structure of the testbeds, including hardware structure, signal structure, etc., is presented in Section 5.1. The signal processing chain and several key signal processing modules, e.g., synchronization, channel estimation, beam selection in the wideband systems, etc., involved in the testbeds, are described in Section 5.2. The testbed integration is given in

Section 5.3. The measurement layout setup in the laboratory is shown in Section 5.4. Afterwards, the measurement results, such as BER performance, etc., are given in Section 5.5. Finally, the conclusions will be drawn in Section 5.6.

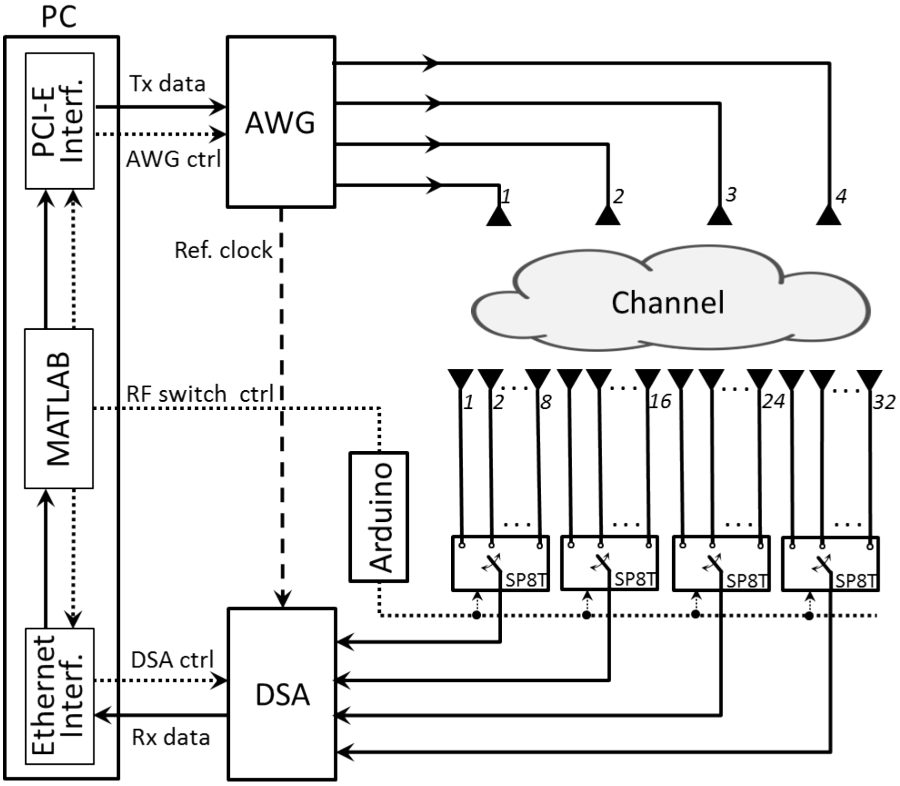
5.1 Testbed Setup

This section describes the structure of the testbeds, including the hardware structure and the signal structure. The flow chart of the experiment measurement, i.e., the process of signal transmission and reception, is also given.

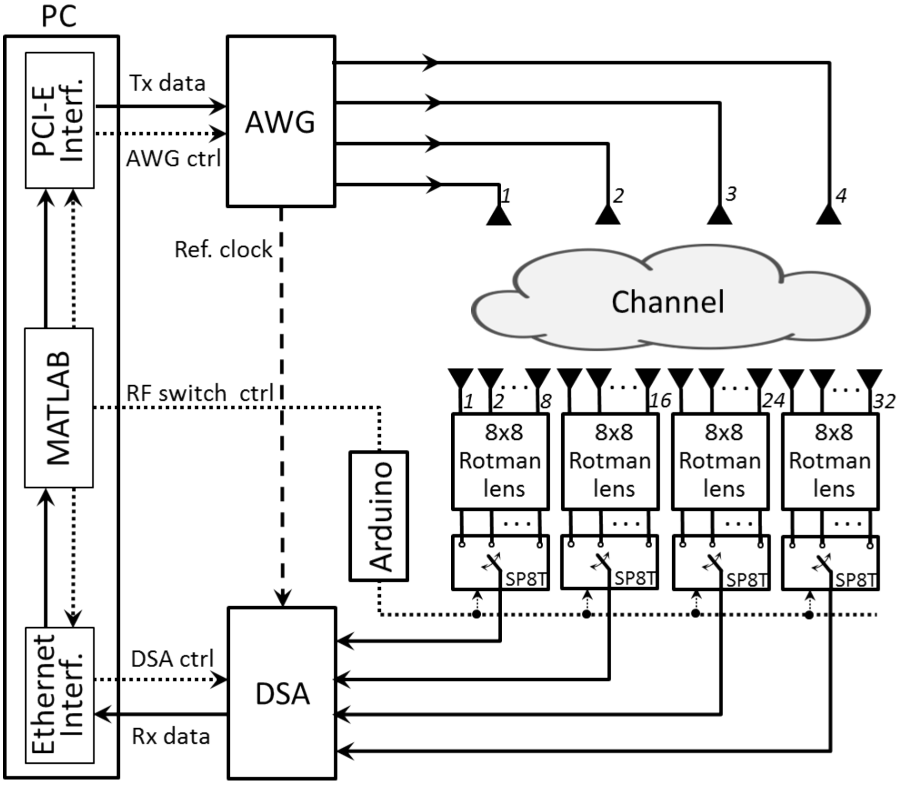
5.1.1 Hardware structure

The uplink multi-user massive MIMO system model given in Chapter 2 is followed for the testbed setup, in which there are $N_{\mathrm{U}} = 4$ single-antenna users and one BS with $N_{\mathrm{B}} = 32$ receive antennas. The hardware setup of the testbeds is shown in Fig. 5.1. By comparing Fig. 5.1 (a) and (b) we can see that the Rotman-lens-based system is based on the antenna selection system by inserting Rotman lenses in-between the receive antennas and the RF switches. Generally, both testbeds comprise three types of links, i.e., communication-data links (denoted by solid lines), device-control links (denoted by dotted lines), and reference-clock link (denoted by dashed lines).

The communication-data links are dedicated to conveying data streams for transmission and reception. At the transmitter side, four independent data streams are generated via the digital signal processing modules with MATLAB in PC, which then go through the PCI-E interface to the 4-channel Arbitrary Wave Generator (AWG) for digital-to-analog conversion. Afterwards, the generated RF analog signals are radiated by $N_{\mathrm{U}} = 4$ transmit antennas, going through realistic channels, and received by the $N_{\mathrm{B}} = 32$ antennas at the BS side. The full-array of 32 BS antennas are segmented into $L = 4$ disjoint sub-arrays of size $M = 8$. Each sub-array antennas are connected to one 8×8 Rotman lens and one followed SP8T RF switch for the Rotman-lens-based system; while the Rotman lenses are excluded for the antenna selection system. The 4 RF switches associated with 4 selected antenna/beam signals only require 4 digitizers instead of 32 digitizers. Clearly, it is the RF switches that realize the signal dimension reduction and low RF-complexity design. The digitization of 4-beam data streams is achieved by the 4-channel Digital Signal Analyzer (DSA). The digitized signals by the DSA are collected



(a) Antenna selection system



(b) Rotman-lens-based hybrid beamforming system

Figure 5.1: Hardware setup of the antenna selection and Rotman-lens-based massive MIMO hybrid beamforming testbeds

and transfered back to the PC via Ethernet interface for the further digital signal processing, e.g., demodulation, channel estimation, MIMO detection, etc. in MATLAB.

The device-control links are used to cover the controlling functionality of the mentioned hardware equipments, e.g., AWG, DSA, RF switches. Various control commands, such as sampling rate, transmit power level, data collection period, etc., are sent from MATLAB in PC to the AWG and DSA. In addition, the antenna/beam selection command, i.e., the indices of the selected antennas/beams calculated by the antenna/beam selection algorithms in MATLAB, is sent to the Arduino board via Universal Serial Bus (USB) interface, and the micro-controller in the Arduino board resolves the control command to generate “HIGH” (+5 volts) and “LOW” (0 volt) levels to control SP8T switches via its General Purpose Input Output (GPIO) pins. Eight beam ports in each SP8T correspond to three control pins from the Arduino board. In addition, the power-supply pins and ground pins of the four RF switches are shared with the power-supply pin and ground pin of the Arduino board, respectively.

The third type of links in the testbed is the reference-clock link between AWG and DSA. In this testbed setup, the AWG internal 100 MHz reference clock is shared to DSA via the reference clock link for frequency synchronization. The frequency synchronization between the transmitter and receiver is very critical, especially for the OFDM waveform used in the testbed, because the OFDM waveform is very sensitive to the frequency offset, and even minor frequency offset will cause severe inter-subcarrier-interference and consequently severe performance degradation. The reference clock shared between AWG and DSA can eliminate a large portion of the frequency offset between the transmitter and receiver, and the residual frequency offset can be simply estimated and compensated in the digital baseband signal processing (see Section 5.2).

The antenna selection system shown in Fig. 5.1 (a) is very similar to the Rotman-lens-based hybrid beamforming shown in Fig. 5.1 (b), where the RF switches are directly connected to the antennas by excluding the Rotman lenses.

5.1.2 Signal Structure

As aforementioned, the most important motivations in this experiment are to study the performance of both the low RF-complexity systems with practical implementation errors, *realistic* channels and *wideband* signals. Thus the signal structure used in the testbed should be carefully designed, including frame structure, resource grid structure of the channel training pilots and user payload data, which will be described in the following.

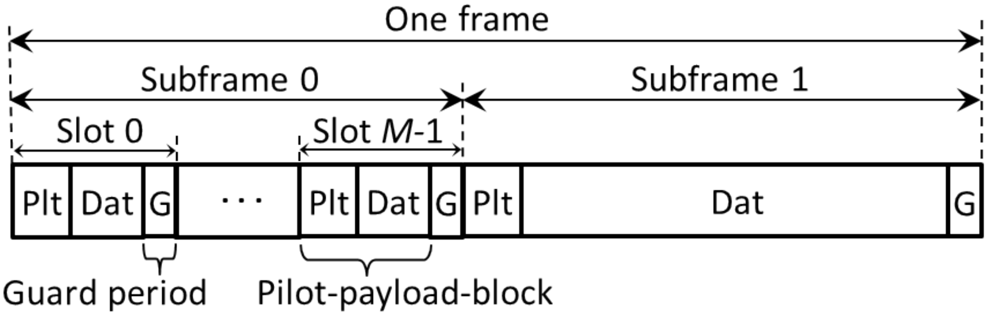


Figure 5.2: Frame structure

Frame Structure

The frame structure designed for the testbed is shown in Fig. 5.2. One frame consists of two subframes, i.e., subframe 0 and subframe 1. Let us first have a look at subframe 0. In subframe 0, there are M slots, corresponding to the number of antennas/beams in each sub-array. Each slot consisting of one pilot OFDM symbol, possible payload data OFDM symbols, and a Guard Period (GP). The pilot and payload data symbols constitute a pilot-payload-block, which is the basic resource block for the pilot and payload data transmission. The pilot symbol in each slot is used to train one of M antenna/beams in each sub-array. Obviously, subframe 0 comprises M pilot-payload-blocks. The GP in each slot is especially reserved for RF switching activities, to make sure that the RF switching activities do not interrupt the transmission/reception of pilot-payload-blocks. Subframe 1 has similar structure to subframe 0, but only one pilot-payload-block and one GP is comprised. It should be noted that although both subframe 0 and subframe 1 consist of pilot-payload-block(s), their intentions are different: subframe 0 is mainly used for channel training and short payload data transmission/reception; while subframe 1 is mainly used for long payload data transmission/reception, because subframe 1 is transmitted/received by using the *high reliable* selected antennas/beams. In addition, since in the antenna/beam selection system the antennas/beams are switched through sequentially in subframe 0 and no dedicated antenna/beam selection algorithms are performed, the payload data transmission and reception in subframe 0 can emulate both systems without specific selection functionality, which serves as benchmark purpose for our antenna/beam selection algorithms. More importantly, this sequential switching of subframe 0 in the antenna selection system can also emulate a 4×4 small-scale MIMO, which serves as a benchmark for comparing the low RF-complexity massive MIMO and the conventional small-scale MIMO.

Obviously, the channel training is the challenge for this kind of switching-based system. Much time-frequency resources have to be spent on training the channel, i.e., M-beam sub-array requires M OFDM symbols to send the pilots. When M is very large, the overall spectral efficiency will be degraded severely. Thus in our testbed, a moderate value $M = 8$ is used. In addition, to ensure the selected antenna/beam channels based on the channel knowledge measured in subframe 0 are still valid for the subframe 1, the frame duration should be shorter than the channel coherence time. Thus the switching-based system fit the quasi-static or slow-fading radio environment rather than the fast-fading radio environment. And for simplicity, the quasi-static radio environment is used for the testbed measurement in Section 5.5.

OFDM Waveform Configuration

In the testbed, the wideband OFDM waveform is used due to its popularity in many practical communication standards, such as 802.11 WLAN, LTE, etc. The number of subcarriers or the FFT size in OFDM modulation is $N_{\text{fft}} = 2048$, in which only the central $N_{\text{used}} = 300$ subcarriers centered by the DC subcarrier are active and the high-frequency subcarriers are inactive with zero values. Note that the DC subcarrier is also zero to avoid the DC offset that reduces the effective dynamic range of the AWG DACs and DSA ADCs. The subcarrier spacing is set to $\Delta_{\text{F}} = 15$ KHz, i.e., the same as in the LTE standard. Thus the bandwidth of the waveform is $B = N_{\text{used}}\Delta_{\text{F}} = 4.5$ MHz. The baseband sampling rate of the waveform is $F_{\text{s}} = 30.72$ MHz with the sampling period $T_{\text{s}} \approx 32.55$ ns. The number of samples for CP is $N_{\text{cp}} = 256$, resulting in the time duration of $\tau_{\text{cp}} = N_{\text{cp}}T_{\text{s}} = 8.33$ µs. Thus the time duration of one OFDM symbol including CP is $\tau_{\text{sym}} = 1/\Delta_{\text{F}} + \tau_{\text{cp}} = 75$ µs.

Pilot-payload-block Structure

The structure of one pilot-payload-block in the time-frequency domain is shown in Fig. 5.3. Each square box denotes one OFDM subcarrier, and one column consisting of N_{fft} subcarriers denote one OFDM symbol in the frequency domain. Note that in each OFDM symbol only the central N_{used} subcarriers are active to carry pilots or payload data and the other subcarriers are inactive with zeros. The active subcarrier is also called Resource Element (RE) in this experiment. The REs in the first OFDM symbol carry only pilots and the REs in the rest OFDM symbols carry payload data.

The pilots in the first OFDM symbol of one pilot-payload-block is arranged with comb-type

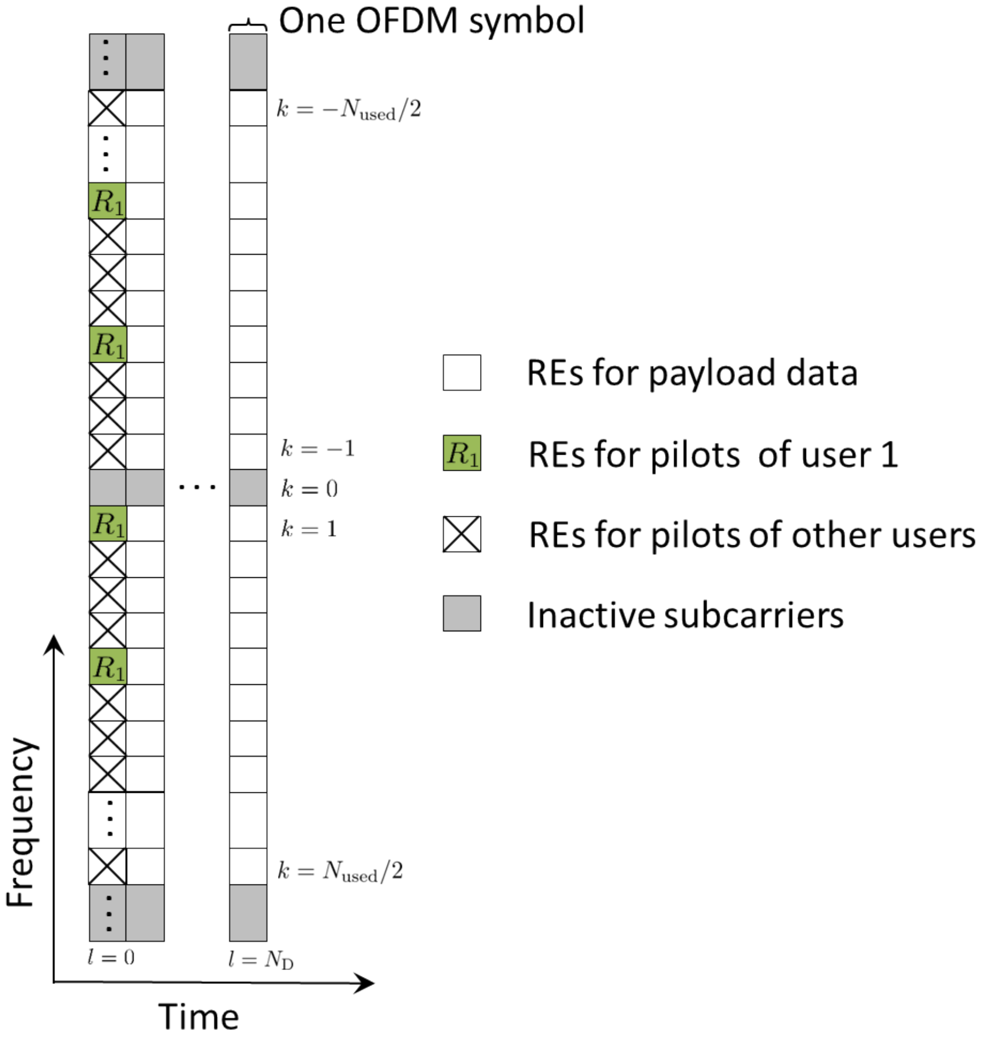


Figure 5.3: Pilot-payload-block structure and the REs allocation for user-1 pilots

with the period of P_{plt}, and the pilots of different users are allocated with different REs to keep orthogonality among them as shown in Fig. 5.3. The number of OFDM symbols for the payload data is $N_{\text{D,blk}}$. Suppose that there are totally $N_{\text{D}}^{(0)}$ and $N_{\text{D}}^{(1)}$ OFDM symbols for the payload data in subframe 0 and subframe 1, respectively. Then we have $N_{\text{D,blk}} = N_{\text{D}}^{(0)}/M$ for the pilot-payload-block in subframe 0; while $N_{\text{D,blk}} = N_{\text{D}}^{(1)}$ for the pilot-payload-block in subframe 1. This configuration is for fair performance comparison between subframe 0 and 1.

5.1.3 Flow Chart

One of the challenges in the switching-based systems is the timing synchronization between the frame receiving and RF switching in the BS. As aforementioned, the RF switching activity

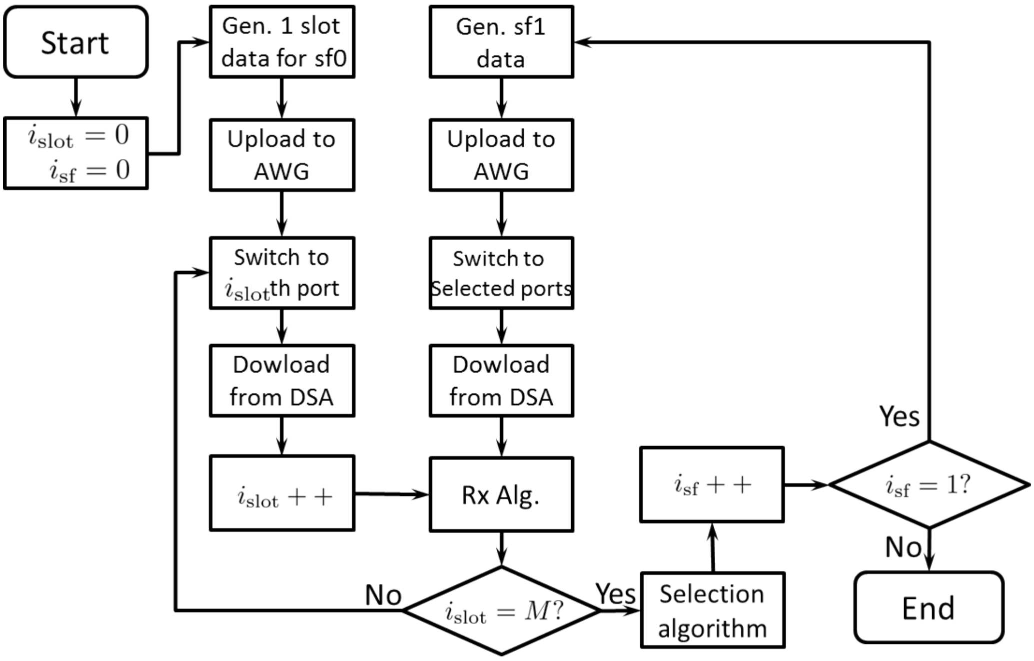


Figure 5.4: Flow chart of the data transmission, RF switching and data reception in the experiment

should occur within the GP in the frame, which requires highly precise timing synchronization between frame receiving and RF switching. In the testbeds, the RF switches are controlled by the Arduino board, and it is impossible to synchronize the Arduino board and the DSA (note that there is no reference clock input to the Arduino board). An alternative way to align the frame receiving and RF switching should be considered. In our experiment, the frame receiving and RF switching are arranged in a sequential order. The arrangement can be illustrated by the flow chart shown in Fig. 5.4. Initially, both the slot index variable i_{slot} and subframe index variable i_{sf} are set to zeros. For simplicity, the payload data transmitted by all the slots in subframe 0 are assumed to be identical. Thus just one slot data is necessary to be generated and uploaded to AWG for subframe 0. The AWG will be configured as *continuous* mode, which will repeat the data in its memory buffer continuously. Before the data collection by DSA, the RF switch in each sub-array will be switched to the i_{slot}th antenna/beam. The collected one-slot data by DSA will be downloaded to PC for the processing of receiving algorithms, e.g., time synchronization to determine the first sample of the slot data, etc. (see Section 5.2 for more details of the involved receiving algorithms). If the number of the collected slots is less than M, it will go back to switch to the succeeding antenna/beam in each sub-array and then follow the same receiving routing aforementioned. On the other hand, if M slots have been received, which means that all of the antennas/beams in each sub-array have been completely cycled, the antenna/beam selection algorithm will be performed to find the desired antenna/beam subsets, and then it goes back to the transmission of subframe 1, RF switching to the selected antenna/beams, and data reception by DSA. After the transmission and reception of the subframe 1, the trial of one frame transmission and reception is finished.

5.2 Signal Processing in the Testbeds

In the preceding section, the general hardware and signal structure of the testbed are described. In this section, the involved signal processing methods implemented for the experiment will be described, including the overall signal processing chain to be described in Subsection 5.2.1 and several key signal processing modules to be illustrated in Subsection 5.2.2.

5.2.1 Signal Processing Chain

The signal processing chain is shown in Fig. 5.5. The transmit payload bits are first digitally modulated and mapped to complex symbols with the Gray coding and specified constellation such as 4-QAM. In parallel, the pilot symbols are generated and multiplexed together with the payload symbols to constitute the pilot-payload-blocks in the frame structure. Afterwards, the baseband OFDM waveform is generated and digitally up-sampled with the sampling rate $F_{\text{s,awg}} = 512F_{\text{s}} = 15.73$ Gsa/s required by the AWG and up-converted to the carrier frequency F_{c} in the 5 GHz band. The up-sampled and up-converted OFDM signals are further processed by reducing the Peak-to-Average Power Ratio (PAPR) and then uploaded to AWG. At the receiver side, the received signals by $N_{\text{B}} = 32$ antennas are first processed by the analog processing modules, i.e., Rotman lenses or/and RF switches. The result $L = 4$ RF analog signals are directly digitized by the DSA with the sampling rate of $F_{\text{s,dsa}} = 20$ Gsa/s. The digitized signals are down-converted to the baseband and then down-sampled to the digital basedband with the sampling rate of $F_{\text{s}} = 30.72$ MHz. The time synchronization is followed to find the first sample of the OFDM symbols. The frequency synchronization can also be done to compensate the residual frequency offset. After the time and frequency synchronization, the OFDM demodulation is performed to extract the pilot-payload-blocks. The pilot REs are used for channel estimation and SNR estimation; while the data REs go to the MIMO detector to decouple the spatially-multiplexed data. The estimated channel matrix is used for MIMO detection since we are considering the coherent demodulation, and is also be used for the beam selection to find the desired beam subset which is feedback to the analog signal processing module. The received payload bits are

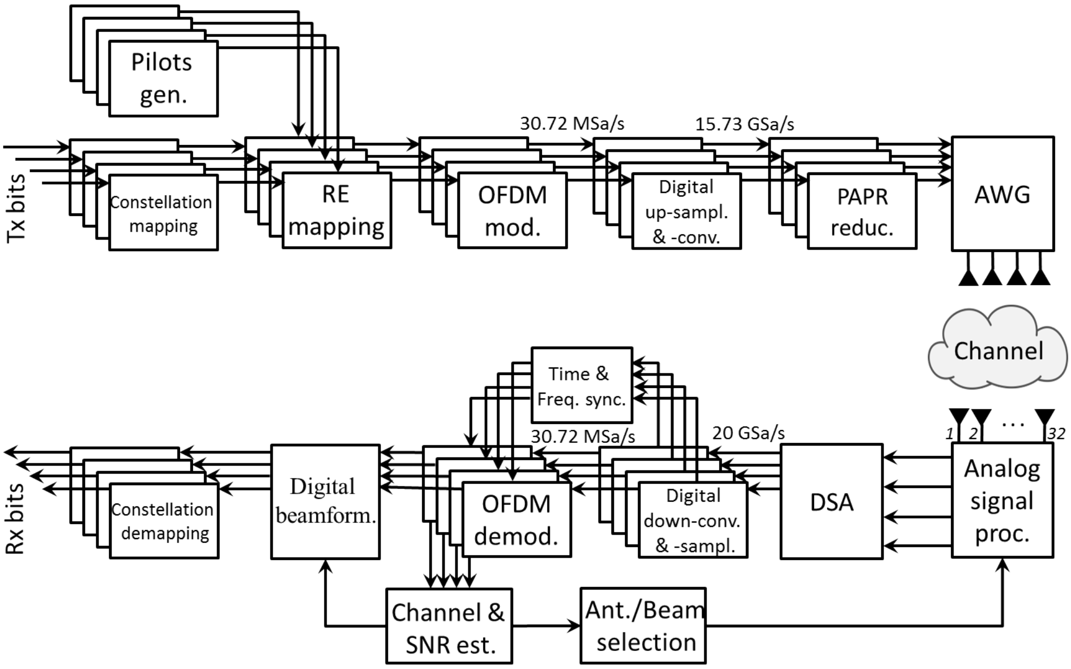


Figure 5.5: Signal processing chain in the experiment

obtained by demodulating the decoupled data streams from the MIMO detector, e.g., 4-QAM demodulator, and finally compared with the transmit payload bits to get the number of error bits.

5.2.2 Signal Processing Modules

In this thesis, the completed signal processing chain shown in Fig. 5.5 is implemented by using MATLAB. The implementation of some key modules will be described in detail in this subsection.

Pilot generation

Given N_{used} REs in the first column of the pilot OFDM symbol, i.e., the first column of the pilot-payload-block in Fig. 5.3 and the period of pilot REs P_{plt}, the number of pilot symbols for each user reads $N_{\text{plt}} = N_{\text{used}}/P_{\text{plt}}$. The pilot symbols for each user are generated referring to the LTE standard [13] as

$$x_{\text{plt}}(m) = \sqrt{\frac{\delta_{\text{boost}}}{2}}\big(1 - 2c_{\text{gld}}(2m)\big) + \jmath\sqrt{\frac{\delta_{\text{boost}}}{2}}\big(1 - 2c_{\text{gld}}(2m+1)\big), m = 0, 1, \cdots, N_{\text{plt}} - 1, \tag{5.1}$$

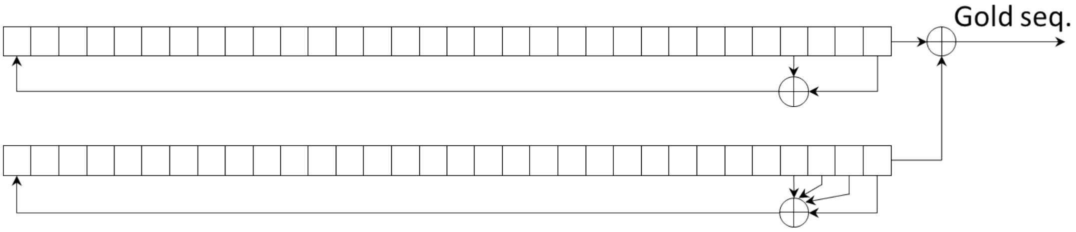


Figure 5.6: Gold code generation with shifting registers

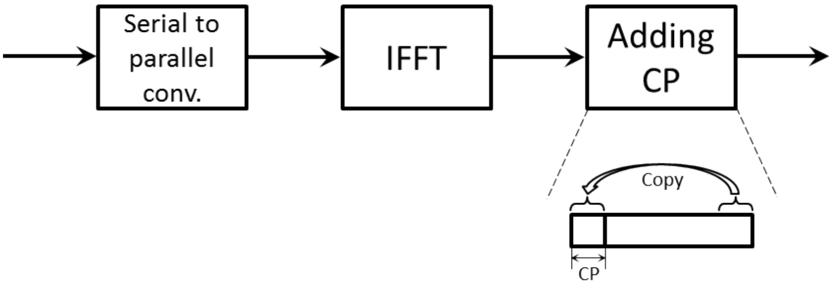


Figure 5.7: OFDM modulation

where δ_{boost} is the pilot power boosting factor over the payload data REs to improve the performance of channel estimation and synchronization; $c_{\text{gld}}(\cdot)$ is a Gold code with order-31, generated by shifting registers shown in Fig. 5.6. The Gold code is derived from the modulo-2 addition of two maximum-length sequences as $c_{\text{gld}}(n) = \big(x_{\text{mseq1}}(n+N_{\text{shft}}) + x_{\text{mseq2}}(n+N_{\text{shft}})\big) \bmod 2$, where N_{shft} is fast-forward value used to reduce the correlation between sequences, and the two maximum-length sequences are calculated by

$$\begin{aligned} x_{\text{mseq1}}(n+31) &= \big(x_{\text{mseq1}}(n+3) + x_{\text{mseq1}}(n)\big) \bmod 2 \\ x_{\text{mseq2}}(n+31) &= \big(x_{\text{mseq2}}(n+3) + x_{\text{mseq2}}(n+2) + x_{\text{mseq2}}(n+1) + x_{\text{mseq2}}(n)\big) \bmod 2, \end{aligned} \tag{5.2}$$

as shown in Fig. 5.6.

OFDM modulation/demodulation

Fig. 5.7 shows the diagram of OFDM modulation [103]. The constellation modulated symbols will be first rearranged by serial to parallel conversion. Let $x_{\text{fd}}(u,k)$ denote the constellation symbol of the uth user and kth subcarrier in the frequency domain. The key processing in OFDM modulation is the IFFT processing, i.e., transforming the frequency-domain data $x_{\text{fd}}(u,k)$ to the time-domain data $x_{\text{td}}(u,n)$, which is expressed by

$$x_{\text{td}}(u,n) = \frac{1}{N_{\text{fft}}} \sum_{k\in\mathcal{U}} x_{\text{fd}}(u,k) e^{\frac{\jmath 2\pi kn}{N_{\text{fft}}}},\ n = 0,1,\cdots,N_{\text{fft}}-1 \tag{5.3}$$

where $\mathcal{U}$ is the set of the active subcarriers in one OFDM symbol. After transforming the data from the frequency domain to the time domain, the tail part of the time-domain samples will be copied to the beginning of the OFDM symbol as shown in Fig. 5.7, which is called Cyclic Prefix (CP). The reason to introduce CP is to remove the inter-symbol interference caused by multipath channels. In addition, the redundancy introduced by CP can also be utilized by the following signal processing, e.g., timing synchronization, etc. OFDM with CP is also called CP-OFDM. The OFDM demodulation is just the reverse of OFDM modulation, i.e. removing CP, FFT and parallel to serial conversion.

Digital up/down-sampling/conversion

The digital up/down-sampling is done by an anti-aliasing Finite Impulse Response (FIR) low-pass filter, which is realized by the MATLAB function "resample($\mathbf{x}, p, q, N_{\text{fir}}$)" to resample the input sequence $\mathbf{x}$ at p/q times of the sampling rate of $\mathbf{x}$, where $\mathbf{x}$ is the output of the OFDM modulatioin and the samples from the DSA for the up/down-sampling, respectively; $p/q = F_{\text{s,awg}}/F_{\text{s}}$ and $p/q = F_{\text{s}}/F_{\text{s,dsa}}$ for the up/down-sampling, respectively; N_{fir} is the order of the anti-aliasing low-pass filter.

After the up/down-sampling, the signals will be up/down-converted from baseband/RF-band to the RF-band/baseband at the transmit side and receive side, respectively. Take the transmit side as an example. Suppose that the nth sample of the uth up-sampled sequence, $u = 1,2,\cdots,N_{\text{U}}$, is denoted by $x_{\text{upsam}}(u,n)$. The nth sample of the up-converted sequence $x_{\text{upcon}}(u,n)$ can be obtained by

$$x_{\text{upcon}}(u,n) = x_{\text{upsam}}(u,n) e^{\jmath 2\pi F_{\text{c}} n T_{\text{s,awg}}}. \tag{5.4}$$

The down-conversion is similar to the up-conversion. The down-converted samples are obtained by

$$x_{\text{downcon}}(l,n) = x_{\text{dsa}}(l,n) e^{-\jmath 2\pi F_{\text{c}} n T_{\text{s,dsa}}}, \tag{5.5}$$

where $x_{\text{dsa}}(l,n)$ is the nth sample of the lth DSA output data stream, $l = 1,2,\cdots,L$.

PAPR reduction

It is known that the OFDM samples in the time domain is Gaussian-like with a large PAPR. In order to fully use the dynamic range of the AWG DACs, i.e., increasing the average transmit power with the fixed full-scale AWG DACs (± 1 volt), the PAPR should be reduced before the digital-to-analog-conversion in AWG. The PAPR reduction in this experiment is based on the conventional clipping method that the samples exceeding threshold will be clipped to the threshold value [104]. Suppose that the maximum magnitude of the up-converted transmitted sequence $x_{\text{upcon}}(u,n)$ is denoted by $x_{\text{max},u}$. Then the threshold can be defined by $\eta_u = \alpha x_{\text{max},u}$, where α is a fixed value, set to 0.8 in the experiment. The clipped signal is thus

$$x_{\text{clip}}(u,n) = \begin{cases} x_{\text{upcon}}(u,n), & |x_{\text{upcon}}(u,n)| \le \eta_u, \\ \frac{x_{\text{upcon}}(u,n)}{|x_{\text{upcon}}(u,n)|}\eta_u, & |x_{\text{upcon}}(u,n)| > \eta_u. \end{cases} \tag{5.6}$$

Synchronization

Synchronization, including time and frequency synchronization, is critical in the receiver. The time synchronization is to find the first sample of the OFDM symbols for the OFDM demodulation, i.e., the FFT window. The frequency synchronization is to compensate the residual frequency offset between the AWG and DSA (note that although the AWG and DSA share the same reference clock, there is still noticeable residual frequency offset, i.e., around 100 Hz by measurement).

The time synchronization is based on the CP auto-correlation method with Maximum Likelihood (ML) estimation [105], which utilizes the redundant information introduced by the CP in OFDM symbols. First, the observation interval of DSA should be decided. Since the transmit data is sent slot-by-slot, it seems that the observation interval should be at least twice of the slot duration to guarantee a completed slot can be extracted. However, recall that in our experiment the AWG is configured to be the "continuous" mode and repeats the slot data continuously, which allows the observation interval to be only one slot, and allows the cyclic auto-correlation within the observation interval to find the start sample, and finally performs cyclic shift to get the required slot data. As the four channels of DSA have already been synchronized within the device, it is only necessary to perform the time synchronization on one DSA-channel. Considering the robustness to the noise, the channel with strongest power is selected for the time synchronization. Suppose that the l_0th channel has the strongest power, and then the cyclic auto-correlation can be calculated by

$$r(l_0, n) = \sum_{n=m}^{N_{\text{sam1slot}}-1} x_{\text{downcon}}(l_0, n) x^*_{\text{downcon}}(l_0, (n + N_{\text{fft}}) \text{mod}\, N_{\text{sam1slot}}), \tag{5.7}$$

where N_{sam1slot} is the number of samples in one slot. Since there are multiple OFDM symbols in one slot, multiple peaks can be expected in $|r(l_0, n)|$, which causes ambiguity to decide the start sample index of the slot. Fortunately, with the large pilot power boosting factor δ_{boost}, the start sample index can be consequently determined by the maximum peak of $r(l_0, n)$, i.e.,

$$n_0 = \arg\min_{n} |r(l_0, n)|. \tag{5.8}$$

Then following the ML method in [105], the frequency offset can be estimated by

$$\Delta f_{\text{off}} = -\frac{1}{2\pi} \angle r(l_0, n_0). \tag{5.9}$$

Note that if $n_0 + N_{\text{fft}} + N_{\text{cp}} > N_{\text{sam1slot}}$, the former smaller peak, i.e., with the index of $n_0 - N_{\text{fft}} - N_{\text{cp}}$, should be used to estimate the frequency offset, which guarantees the constant N_{fft} samples offset between CP samples and the OFDM tail samples in the *sequential* time.

Channel estimation

With the orthogonal pilot symbols among users, the MIMO channel estimation can be decoupled into Single-Input Single-Output (SISO) channel estimation. For the antenna selection system, we estimate the MIMO channel from the transmit antennas to the receive antennas; while for the Rotman-lens-based system, we directly estimate the MIMO channel from the transmit antennas to the beam ports of the Rotman lenses. We refer to the former channel as the antenna channel and the latter channel as the beam channel. The time-domain Least Square (LS) based method in [106] is used for channel estimation. Take the uth user and lth receive antenna as an example. The observations on the pilot subcarriers (REs) in the frequency domain are expressed by

$$[\mathbf{y}_l]_{\mathcal{P}_u} = [\mathbf{h}_{lu}]_{\mathcal{P}_u} \odot [\mathbf{x}_u]_{\mathcal{P}_u} + [\mathbf{w}_l]_{\mathcal{P}_u}, \tag{5.10}$$

where $\mathbf{x}_l$, $\mathbf{y}_l$ and $\mathbf{w}_l$ are the uth transmit vector, lth receive vector and lth noise vector, respectively; $\mathbf{h}_{lu}$ is the channel response vector of the uth user and lth receive antenna in the frequency domain; $\mathcal{P}_u$ is the pilot index set of the uth user. The Hadamard multiplication form in Equ.

(5.10) can also be written into the matrix multiplication form

$$[\mathbf{y}_l]_{\mathcal{P}_u} = [\mathbf{X}_u]_{\mathcal{P}_u,\mathcal{P}_u}[\mathbf{h}_{lu}]_{\mathcal{P}_u} + [\mathbf{w}_l]_{\mathcal{P}_u}, \tag{5.11}$$

where $\mathbf{X}_u$ is a diagonal matrix with the elements in $\mathbf{x}_u$ as diagonal entries. The channel response $\mathbf{h}_{lu}$ in the frequency domain is the Fourier transform of the channel impulse response in the time domain. In OFDM modulation, the Fourier transform is approximated by DFT, i.e.,

$$\mathbf{h}_{lu} = \mathbf{F}\mathbf{g}_{lu}, \tag{5.12}$$

where $\mathbf{F}$ is the $N_{\text{fft}} \times N_{\text{fft}}$ DFT matrix; $\mathbf{g}_{lu} \in \mathbb{C}^{N_{\text{fft}}\times 1}$ is the discrete channel impulse response in the time domain. Thus Equ. (5.11) can be rewritten as

$$[\mathbf{y}_l]_{\mathcal{P}_u} = [\mathbf{X}_u]_{\mathcal{P}_u,\mathcal{P}_u}[\mathbf{F}]_{\mathcal{P}_u,:}\mathbf{g}_{lu} + [\mathbf{w}_l]_{\mathcal{P}_u}. \tag{5.13}$$

Obviously, the Equ. (5.13) is underdetermined to estimate $\mathbf{g}_{lu}$: there are more variables (N_{fft}, i.e., the length of $\mathbf{g}_{lu}$) than the number of equations (N_{plt}, i.e., the number of pilots). Fortunately, as the channel path power only concentrates in the first multiple taps, supposing K_g taps, in $\mathbf{g}_{lu}$, accurate least square estimation of the channel impulse response can be obtained if $K_g \ll N_{\text{plt}}$, which is expressed by

$$[\hat{\mathbf{g}}_{lu}]_{\mathcal{Q}} = ([\mathbf{F}]^H_{\mathcal{P}_u,\mathcal{Q}}[\mathbf{F}]_{\mathcal{P}_u,\mathcal{Q}})^{-1}[\mathbf{F}]^H_{\mathcal{P}_u,\mathcal{Q}}[\mathbf{X}_u]^{-1}_{\mathcal{P}_u,\mathcal{P}_u}[\mathbf{y}_l]_{\mathcal{P}_u}, \tag{5.14}$$

where the set $\mathcal{Q} = \{1, 2, \cdots, K_g\}$. Note that if $K_g > N_{\text{plt}}$, the matrix $[\mathbf{F}]^H_{\mathcal{P}_u,\mathcal{Q}}[\mathbf{F}]_{\mathcal{P}_u,\mathcal{Q}}$ in Equ. (5.15) is non-invertible, which can be addressed by a regularization factor δ_{reg}, resulting in

$$[\hat{\mathbf{g}}_{lu}]_{\mathcal{Q}} = ([\mathbf{F}]^H_{\mathcal{P}_u,\mathcal{Q}}[\mathbf{F}]_{\mathcal{P}_u,\mathcal{Q}} + \delta_{\text{reg}}\mathbf{I}_{K_g})^{-1}[\mathbf{F}]^H_{\mathcal{P}_u,\mathcal{Q}}[\mathbf{X}_u]^{-1}_{\mathcal{P}_u,\mathcal{P}_u}[\mathbf{y}_l]_{\mathcal{P}_u}. \tag{5.15}$$

In the experiment, the regularization factor $\delta_{\text{reg}} = 0.01$ is validated to be able to achieve stable channel estimation results. Finally, the channel response on all of the used subcarriers can be obtained by applying DFT to the estimated channel impulse response $\hat{\mathbf{g}}_{lu}$, i.e.,

$$\hat{\mathbf{h}}_{lu} = [\mathbf{F}]_{\mathcal{U},\mathcal{Q}}[\hat{\mathbf{g}}_{lu}]_{\mathcal{Q}}, \tag{5.16}$$

where $\mathcal{U}$ is the set of the active subcarriers in one OFDM symbol.

SNR estimation

To estimate SNR, the signal power and the noise power should be first estimated, and then the SNR is estimated by the ratio of the estimated signal power and noise power. Let us first study the signal power estimation. With the linear system model in the frequency domain, the observation vector on the kth subcarrier is expressed by

$$\mathbf{y}_k = \hat{\mathbf{H}}_k \mathbf{x}_k + \tilde{\mathbf{w}}_k, \tag{5.17}$$

where $\hat{\mathbf{H}}_k \in \mathbb{C}^{N_\text{B} \times N_\text{U}}$ is the estimated channel matrix; $\tilde{\mathbf{w}}_k$ is the noise plus channel estimation error. The total signal power by N_B receive antennas can be estimated by

$$\begin{aligned} P_\text{sig} &= \mathbb{E}\{\text{tr}\{\hat{\mathbf{H}}_k \mathbf{x}_k \mathbf{x}_k^H \hat{\mathbf{H}}_k^H\}\} \\ &= \text{tr}\{\hat{\mathbf{H}}_k \mathbb{E}\{\mathbf{x}_k \mathbf{x}_k^H\} \hat{\mathbf{H}}_k^H\}. \end{aligned} \tag{5.18}$$

Note that $\mathbb{E}\{\mathbf{x}_k \mathbf{x}_k^H\} = \mathbf{I}_{N_\text{U}}$ assumed in the system model. Thus we have

$$\begin{aligned} P_\text{sig} &= \text{tr}\{\hat{\mathbf{H}}_k \hat{\mathbf{H}}_k^H\} \\ &= ||\hat{\mathbf{H}}_k||_\text{F}^2. \end{aligned} \tag{5.19}$$

As a result, the average power per subcarrier and per receive antenna is

$$\bar{P}_\text{sig} = \frac{1}{N_\text{B} N_\text{used}} \sum_{k=1}^{N_\text{used}} ||\hat{\mathbf{H}}_k||_\text{F}^2. \tag{5.20}$$

If ignoring the channel estimation error, the noise power can be roughly calculated by

$$\bar{P}_\text{noise} = \bar{P}_\mathbf{y} - \bar{P}_\text{sig}, \tag{5.21}$$

where $\bar{P}_\mathbf{y}$ is the average receive power per subcarrier and per receive antenna. Obviously, the accuracy of the noise power estimation given by Equ. (5.21) depends on the accuracy of channel estimation. When the channel estimation error is large, the estimation result given by Equ. (5.21) will be dominated by the channel estimation error. The dominance of the channel estimation error is significant especially when SNR is high, e.g., 18 dB observed in the experiment, which will make the SNR estimation be underestimated. Thus the noise power should be estimated independently to the signal power. In the experiment, the CP data in the time domain is used instead to estimated the noise power [107]. Recall that the maximum channel delay is K_g taps, $K_g < N_{\text{cp}}$, and thus the CP samples with indices $\mathcal{P}_1 = \{K_g + 1, K_g + 2, \cdots, N_{\text{cp}}\}$ in one OFDM symbol will be free of inter-symbol-interference. As a result, the noise power in the time domain can be estimated by

$$\bar{P}_{\text{noise,td}} = \frac{1}{N_{\text{U}}(N_{\text{cp}} - K_g)} \sum_{l=1}^{N_{\text{B}}} \sum_{n \in \mathcal{P}_1} |x_{\text{downcon}}(l, n) - x_{\text{downcon}}(l, n + N_{\text{fft}})|^2. \tag{5.22}$$

Then the noise power per subcarrier in the *frequency* domain can be directly obtained by

$$\bar{P}_{\text{noise}} = \bar{P}_{\text{noise,td}} \frac{||\mathbf{F}||_{\text{F}}^2}{N_{\text{fft}}}. \tag{5.23}$$

Note that if the DFT matrix $\mathbf{F}$ is normalized, we simply have $\bar{P}_{\text{noise}} = \bar{P}_{\text{noise,td}}$.

Finally, the average SNR on each subcarrier is obtained by

$$\rho_{\text{est}} = \frac{\bar{P}_{\text{sig}}}{\bar{P}_{\text{noise}}}. \tag{5.24}$$

Note that, in order to make fair comparison between the antenna selection system and Rotman-lens-based hybrid beamforming system, the SNR of both systems should be defined at the receive antenna side, rather than in the digital baseband. Thus the above estimated SNR in the digital baseband should be calibrated by taking the insertion loss of the RF switches and Rotman lens into account. For the antenna selection system, we have

$$\rho_{\text{est,as,dB}} = \rho_{\text{est,dB}} + \beta_{\text{IL,switch}}, \tag{5.25}$$

and for the hybrid beamforming system, we have

$$\rho_{\text{est,had,dB}} = \rho_{\text{est,dB}} + \beta_{\text{IL,switch}} + \beta_{\text{IL,rotman}}, \tag{5.26}$$

where $\rho_{\text{est,dB}} = 10 \log_{10}(\rho_{\text{est}})$; $\beta_{\text{IL,switch}}$ and $\beta_{\text{IL,rotman}}$ are the insertion losses of the RF switches and Rotman lenses in decibel, respectively. Through measurement, we have $\beta_{\text{IL,switch}} = 3$ dB and $\beta_{\text{IL,rotman}} = 4.5$ dB. Note that we use Rogers RO4350 substrate for the fabrication of the four Rotman lenses, which has larger insertion loss than the RO5870 used in Chapter 4.

Antenna/beam selection in wideband systems

The antenna/beam selection module shown in Fig. 5.5 is to calculate the desired subset of antennas/beams from the estimated channel matrix, and feedback the indices of the selected antennas/beams to the analog signal processing module, i.e., RF switches. The algorithms to select antenna/beams follow the antenna/beam selection algorithms described in Chapter 3 and 4. Note that the selection algorithms developed in Chapter 3 and 4 are based on the narrowband model; while the signal model used in the testbeds are wideband. A straightforward method to address this issue is that we can rely on only one subchannel's channel matrix (i.e., one subcarrier in OFDM systems, 15 KHz bandwidth, thus narrowband) to perform the selection algorithms. Obviously, the selected beam subset is a good option for the chosen subchannel, but probably not good for the other subchannels due to the fact that the channel information of the other subchannel is ignored. In order to address this issue, the narrowband selection algorithms should be tailored for the wideband systems.

With the estimated antenna/beam channel matrix $\hat{\mathbf{H}}_{sk} \in \mathbb{C}^{L\times N_{\text{U}}}$ of the kth subcarrier and sth submatrix, the MSE matrix of the LMMSE digital beamformer is $(\hat{\mathbf{H}}_{sk}^H\hat{\mathbf{H}}_{sk} + N_{\text{U}}/\rho\mathbf{I}_{N_{\text{U}}})^{-1}$. Let $\hat{\mathbf{G}}_{sk} = (\mathbf{I}_{N_{\text{U}}} + \bar{\rho}\hat{\mathbf{H}}_{sk}^H\hat{\mathbf{H}}_{sk})^{-1}$, which is the scaled version of the MSE matrix. The optimization problem aiming to minimize the MSE of all of subcarriers is formulated as

$$\mathbf{s}^{\text{opt}} = \arg \min_{\mathbf{s}\in\mathcal{S}^{\text{FAS}} \text{ or } \mathcal{S}^{\text{SAS}}} \sum_{k=1}^{N_{\text{used}}} \operatorname{tr}(\hat{\mathbf{G}}_{sk}), \tag{5.27}$$

where $\mathcal{S}^{\text{FAS}}$ and $\mathcal{S}^{\text{SAS}}$ are the sets of all antenna/beam subsets for the full-array and sub-array selection. The optimization problem given by Equ. (5.27) is very similar to the optimization problem in the narrowband system given by Equ. (4.7), except that there is an additional summation operation in Equ. (5.27) to count for the wideband characteristics. The problem given by Equ. (5.27) is difficult to solve, and an simplified or approximated alternative is necessary.

It is known from [108] that the function $\operatorname{tr}(\mathbf{A})$ is convex for the positive definite matrix $\mathbf{A}$. In our problem (5.27), $\hat{\mathbf{G}}_{sk}$ is obviously a positive definite matrix, and thus the objective function $\operatorname{tr}(\hat{\mathbf{G}}_{sk})$ is convex. Consequently, we have

$$\begin{aligned}\frac{1}{N_{\text{used}}}\sum_{k\in\mathcal{U}} \operatorname{tr}(\hat{\mathbf{G}}_{sk}) &\geq \operatorname{tr}\left((\mathbf{I}_{N_{\text{U}}} + \frac{\bar{\rho}}{N_{\text{used}}}\sum_{k\in\mathcal{U}}\hat{\mathbf{H}}_{sk}^H\hat{\mathbf{H}}_{sk})^{-1}\right) \\ &= \operatorname{tr}\left((\mathbf{I}_{N_{\text{U}}} + \bar{\rho}\mathbf{R}_s)^{-1}\right) \end{aligned} \tag{5.28}$$

where $\mathcal{U}$ is the set of the active subcarriers in one OFDM symbol; $\mathbf{R}_s = \frac{1}{N_{\text{used}}} \sum_{k \in \mathcal{U}} \hat{\mathbf{H}}_{sk}^H \hat{\mathbf{H}}_{sk}$ is actually the sample covariance matrix, which is the principle submatrix of the complete sample covariance matrix $\mathbf{R}$, i.e.,

$$\mathbf{R}_s = \mathbf{S}\mathbf{R}\mathbf{S}^T, \tag{5.29}$$

where $\mathbf{S}$ is the selection matrix corresponding to the the selected subset of antenna/beam ports. From Equ. (5.28) we can see that optimizing the original objective function can be approximately replaced by optimizing its lower bound, which is much simpler than the original problem. Let the eigen-value decomposition of $\mathbf{R}$ be denoted by $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, then we have

$$\begin{aligned} \operatorname{tr}\left((\mathbf{I}_{N_{\text{U}}} + \bar{\rho}\mathbf{R}_s)^{-1}\right) &= \operatorname{tr}\left((\mathbf{I}_{N_{\text{U}}} + \bar{\rho}\mathbf{S}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H\mathbf{S}^T)^{-1}\right) \\ &= \operatorname{tr}\left((\mathbf{I}_{N_{\text{U}}} + \bar{\rho}\mathbf{S}\tilde{\mathbf{U}}\tilde{\mathbf{U}}^H\mathbf{S}^T)^{-1}\right), \end{aligned} \tag{5.30}$$

where $\tilde{\mathbf{U}} = \mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}$. Equ. (5.30) shows that the original optimization problem can be simplified to selecting a good submatrix from the square-root beam correlation matrix $\tilde{\mathbf{U}}$. In contrast, for the narrowband system we are aiming to selecting a good submatrix from the antenna/beam channel matrix $\tilde{\mathbf{H}}$ as shown by Equ. (4.7). All the submatrix selection algorithms in Chapter 3 and 4 can be used to select a good submatrix from $\tilde{\mathbf{U}}$.

We can see that the wideband selection is based on the channel co-variance matrix rather than the instantaneous channel realizations. It is known that the channel co-variance is a long-term statistics, which is much more stable than the instantaneous. Thus the channel training can be done in a long-term way, which is very favorite for the switching based systems.

Digital beamformer

The digital beamforming is simply the LMMSE estimator as shown in Chapter 4. Note that the digital beamforming should be performed on each active subcarrier of the payload OFDM symbols.

5.3 Testbed Integration

As discussed in the preceding sections, the testbeds consist of various hardware modules and digital signal processing modules, which makes the development procedure quite challenging, and the failure of even a single module will fail the whole testbed. Thus an appropriate integration

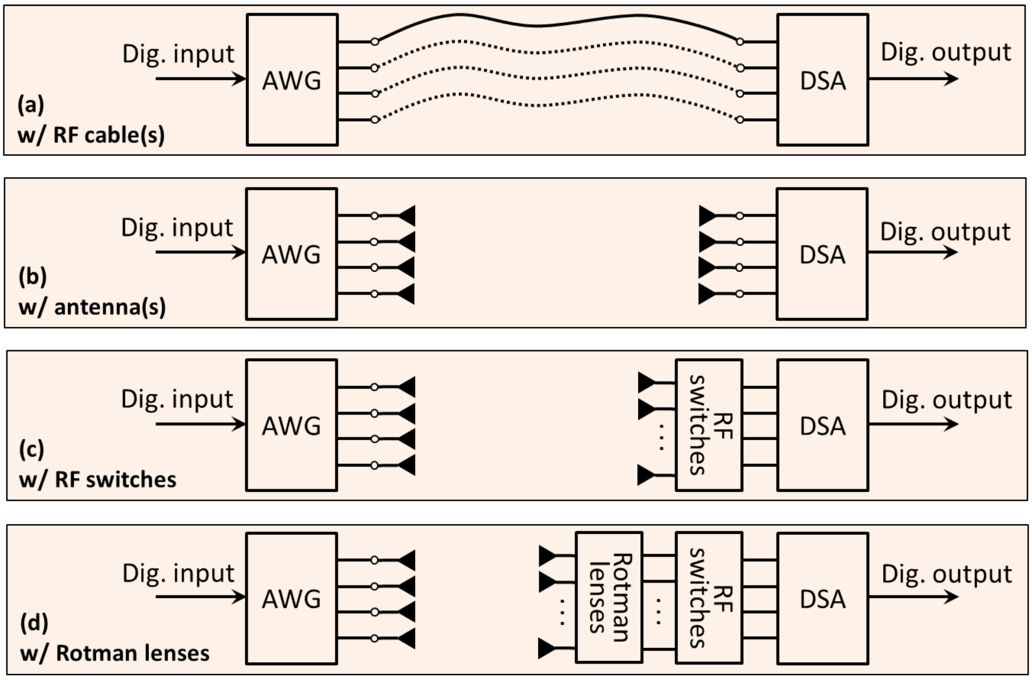


Figure 5.8: Testbed development methodology

strategy is necessary to guide the development procedure. Fig. 5.8 gives the integration strategy used in the development of the testbeds. The principle is to develop the testbeds from the very basic setup to more complicated setup. The whole development procedure is segmented into four steps. In step (a), the RF cable(s) are used to directly connect the transmitter and the receiver. Initially, a SISO system with a single RF cable is set up, which aims to enable the communication-data and device-control links between PC and the AWG and DSA. Another task is to implement and verify many digital signal processing modules described in Section 5.2, e.g., pilot generation, OFDM (de)modulation, channel estimation, synchronization, etc. When the SISO system is up and running, e.g., very low BER, another three channels are taken into account to set up four parallel RF-cable-based communication links, which is used to verify the functionality of an ideal MIMO system without inter-antenna-interference. Replacing the RF cables with antennas is the main focus of step (b), in which the involved digital signal processing modules, especially the time synchronization, channel estimation modules, etc., are verified with the realistic 4×4 MIMO channels, to ensure the robustness of these algorithms over multipath-fading channels. The massive number of receive antennas are considered in step (c), where the RF switches are introduced at the receive side to set up the antenna selection system. The associated control functionality from PC to RF switches with a single hop by the Arduino board is implemented and verified. In addition, the completed flow chart described in Subsection

5.1.3 is realized in this step. Afterwards, in step (d) the Rotman lenses are inserted in-between the receive antennas and the RF switches to finalize the testbed setup and verification.

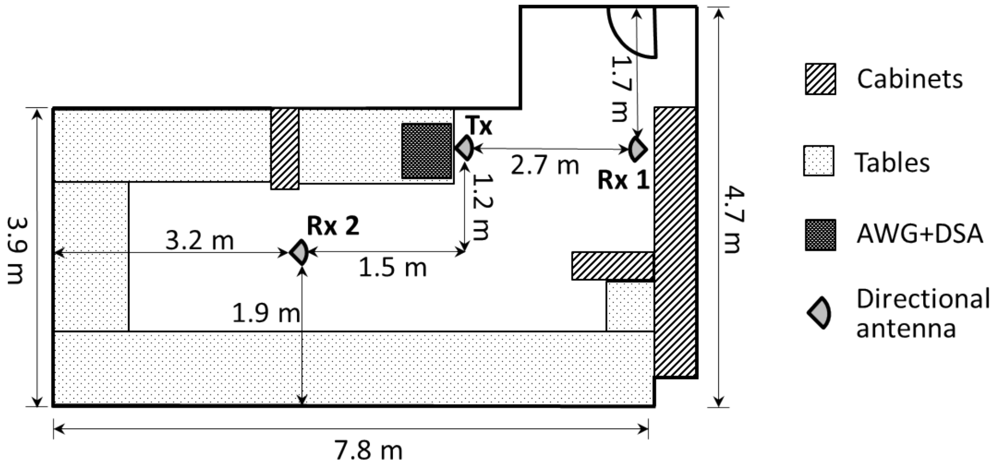


Figure 5.9: Measurement layout

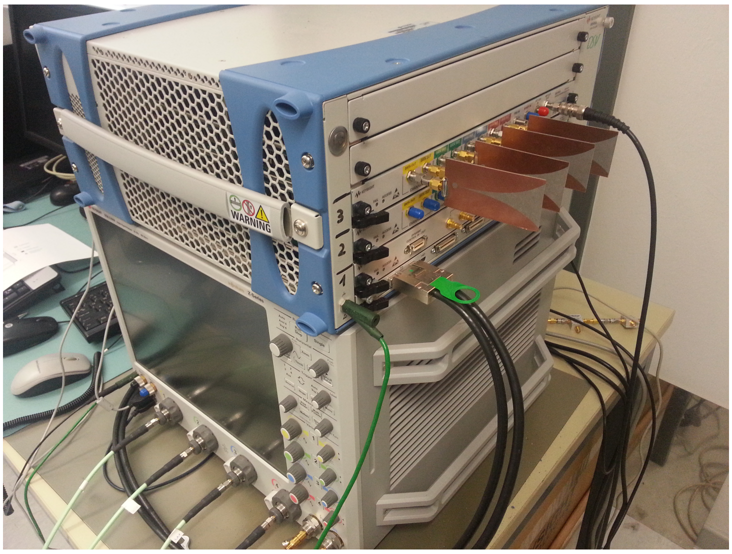
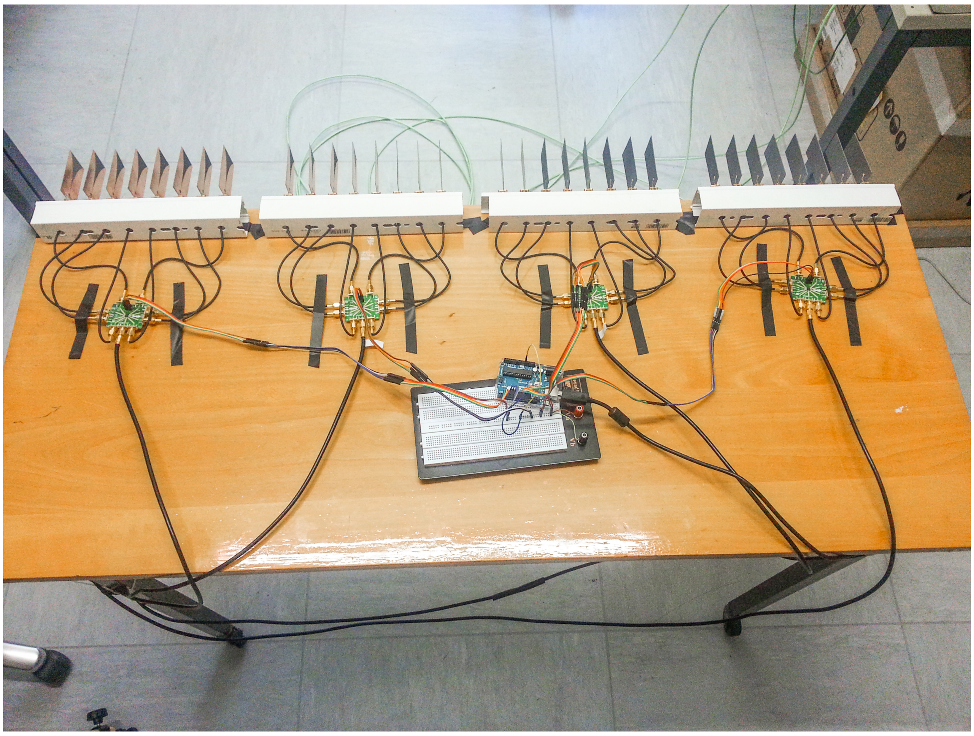


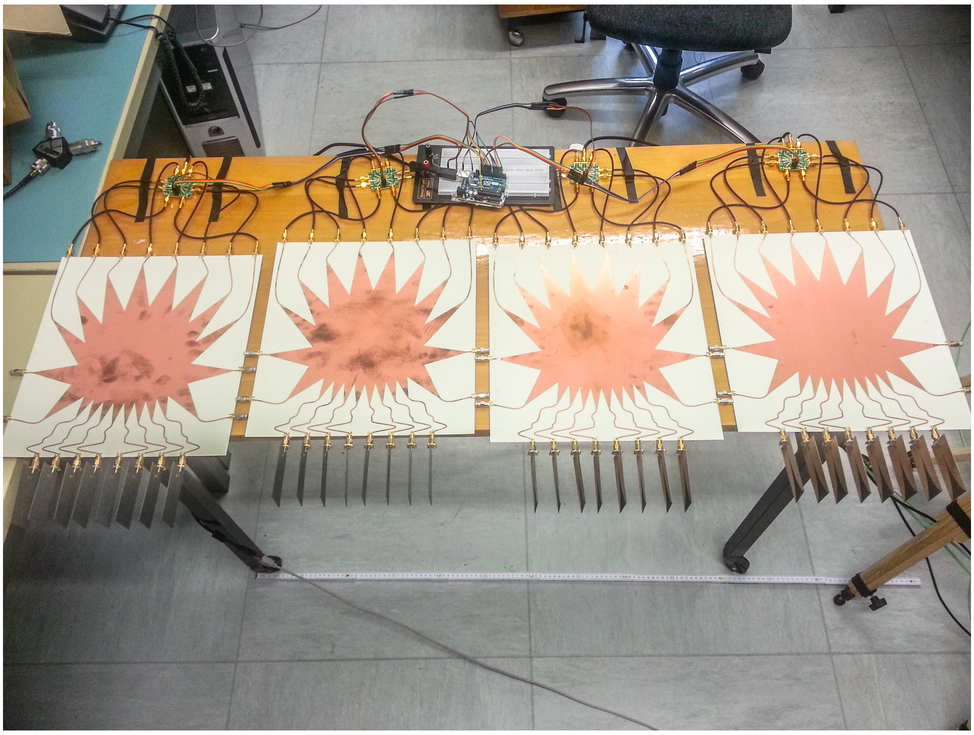
Figure 5.10: AWG, DSA, and transmit antennas

5.4 Layout of the Measurement Environment

Fig. 5.9 shows layout of the laboratory with one transmit location and two receive locations. The $N_\mathrm{U} = 4$ transmit antennas are arranged by an ULA with an antenna spacing of 4.5 cm and 1.2 m height over the floor, which is shown in Fig. 5.10. The $N_\mathrm{B} = 32$ receive antennas are arranged with an ULA and 0.8 m height over the floor as shown in Fig. 5.11. Note that both



(a) Receiver with antenna selection



(b) Receiver with Rotman-lens-based hybrid beamforming

Figure 5.11: BS receiver

AWG and DSA are placed at the location Tx as shown in Fig. 5.10. The receive antenna array can be placed at two receive locations, i.e., Rx1 and Rx2 in Fig. 5.9, which are for the line of sight and non-line of sight scenarios, respectively. The received signals at the location Rx1 or Rx2 are transferred to DSA with RF cables. The distance between the transmit and the receive locations is about two to three meters, which is limited by the low transmit power of AWG and the RF cable length.

Table 5.1: Testbed measurement configurations

Parameters	Value	Description
F_{c}	5.2 GHz	Carrier frequency
F_{s}	30.72 MSa/s	Baseband sampling rate
$F_{\text{s,awg}}$	15.73 GSa/s	Sampling rate of AWG
$F_{\text{s,dsa}}$	20 GSa/s	Sampling rate of DSA
N_{fft}	2048	FFT size of OFDM modulation/demodulation
N_{used}	300	The number of active subcarriers
Δ_{F}	15 KHz	Subcarrier spacing
B	4.5 MHz	Bandwidth
N_{cp}	256	The number of samples for CP
K_g	128	The number of channel delay taps
P_{plt}	4	Pilot period in frequency domain for each user
τ_{sym}	75 µs	Time duration of one OFDM symbol
τ_{cp}	8.33 µs	Time duration of CP
N_{B}	32	The number of receive antennas
N_{U}	4	The number of users
M	8	The number of antennas in one sub-array
L	4	The number of receive RF transceivers
$N_{\text{D}}^{(0)}$,$N_{\text{D}}^{(1)}$	8	The number of payload OFDM symbols in subframe 0 and 1
$\lvert\mathcal{A}\rvert$	4	Alphabet size of constellation mapping, 4-QAM

5.5 Testbed Measurement Results

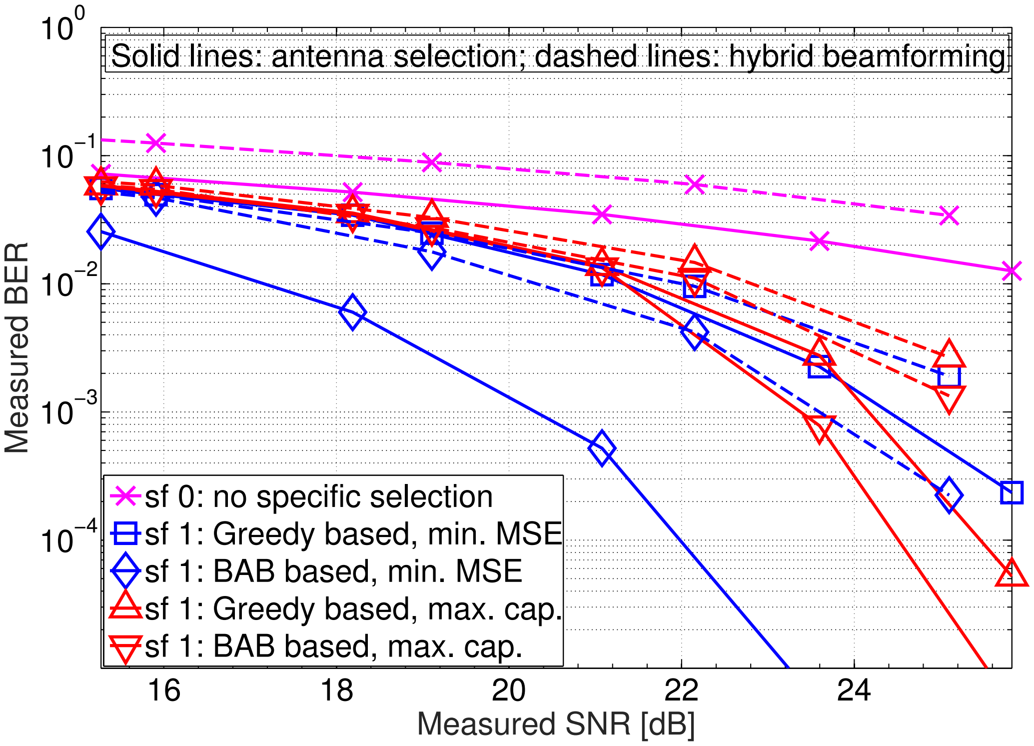
The configuration of the testbed measurement is summarized in Tab. 5.1. Note that the sampling rates of AWG and DSA are chosen according to the limited step size of the AWG and DSA, respectively. Each subframe transmits $N_{\text{used}} \times M \times \log_2(|\mathcal{A}|) = 4800$ payload bits per user, corresponding to $4800 \times 4 = 19200$ bits for four users. In addition, we tune the AWG with 5 different transmit power levels to achieve five different SNRs at the receiver. For each transmit power level, we measure $N_{\text{trials}} = 10$ trials (frames) to get the averaged BER.

The measured BER at the locations Rx1 and Rx 2 is shown in Fig. 5.12. We can see that, for both line of sight and non-line of sight, the subframe 1 has superior performance over the subframe 0, which means that the antenna/beam selection are very effective in practical indoor environment. Since the subframe 0 in antenna selection testbed can emulate the conventional 4 small-scale MIMO system, the measurement results shown in Fig. 5.12 also indicate that the antenna/beam selection system has superior advantage over the conventional small-scale MIMO under the same number of RF transceivers. Most interestingly, the BER of the selection systems is much steeper than the small-scale MIMO system, which is due to the selection diversity gain.

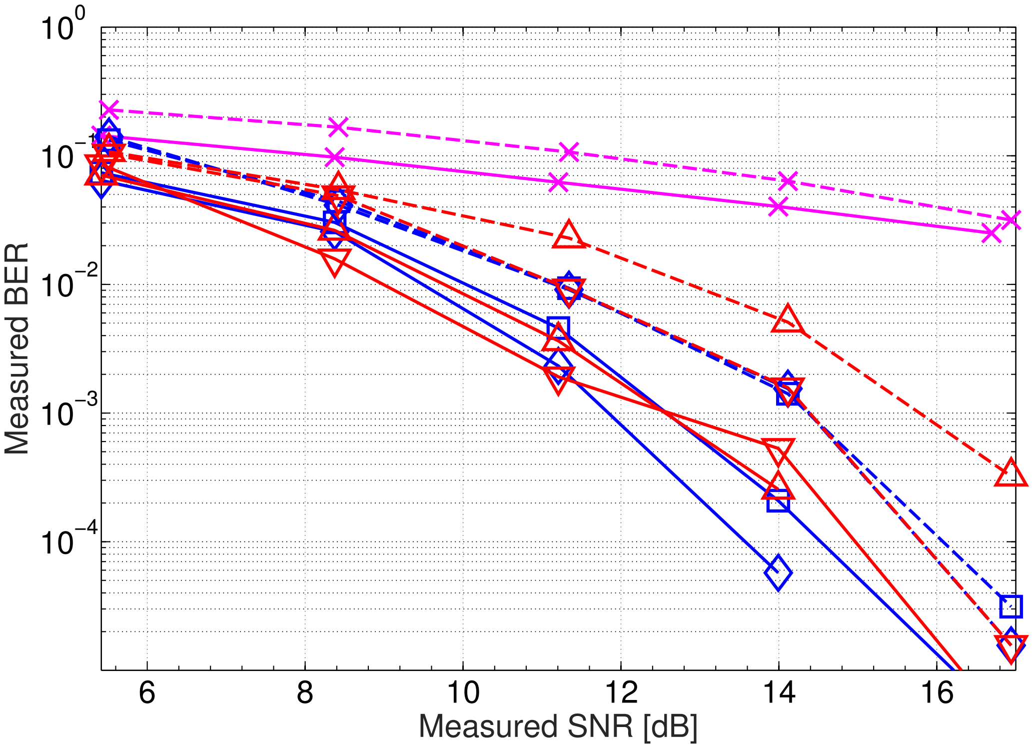
Comparing different selection algorithms we can see that the BAB-based method with minimum MSE criterion has the overall best performance in both testbeds. The greedy methods, i.e., min-MSE-oriented and max-cap-oriented, have very similar performance for both antenna and beam selection testbeds. For the non-line of sight scenario, the non-optimal greedy methods have close performance to the optimal BAB methods.

Comparing the antenna and beam selection systems we can see that the antenna selection systems outperform the beam selection systems, which is due to the large insertion loss introduced by the Rotman lenses. Take the non-line of sight measurements and BAB-min-MSE method as an example, the antenna selection system has 2 dB gain over the beam selection system. Recall that the Rotman lenses have approximate 4.5 dB insertion loss. If the Rotman lenses are ideal without insertion loss, the beam selection system should be more than 2 dB advantageous over the antenna selection system. This indicates that designing very low loss Rotman lens is very critical in the future study. Note that in our testbed, the substrate material RO4350 is used for fabricating the four Rotman lenses due to its much lower cost compared to the RO5870 used in Chapter 4, but the at the price of larger insersion loss. Fig. 5.13 compares the measured insertion loss of Rotman lenses based on RO5870 and RO4350. We can see that the better material, i.e., RO5870, indeed has the advantage on insertion loss. If the Rotman lens cost is not a big issue in the future mass production, we can use even better substrate material for the fabrication. And the performance advantage of the Rotman-lens-based system over the antenna selection system can be foreseen.

Fig. 5.12 shows that the selection systems have superior performance over the small-scale MIMO system, which can be actually explained by the channel characteristics. Since the antenna selection systems have better performance than the beam selection systems, we just focus on the antenna selection systems in the following. It is known that the performance of the MIMO system with the linear receiver depends on the Minimum Singular Values (MSVs) of the channel matrices [26, 109]. The larger the MSV, the better the performance. This is intuitively the same

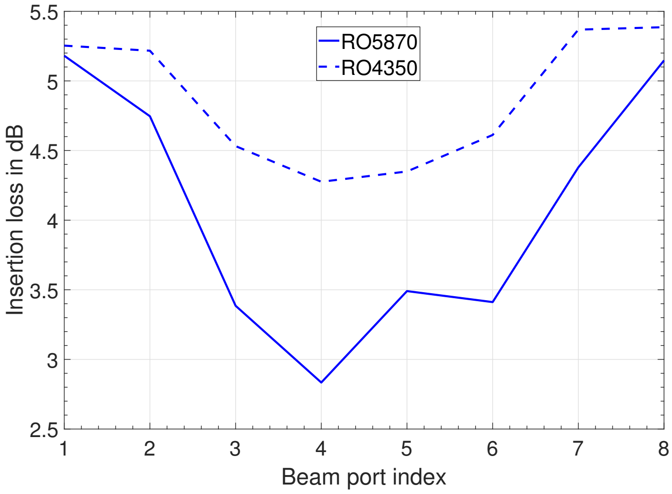


(a) Line of sight

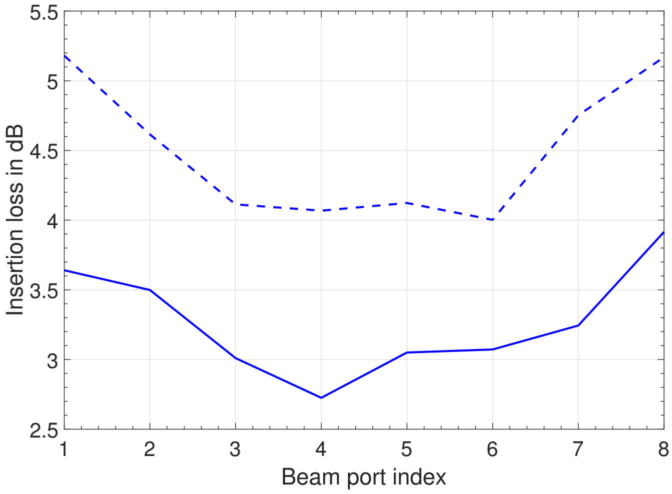


(b) Non-line of sight

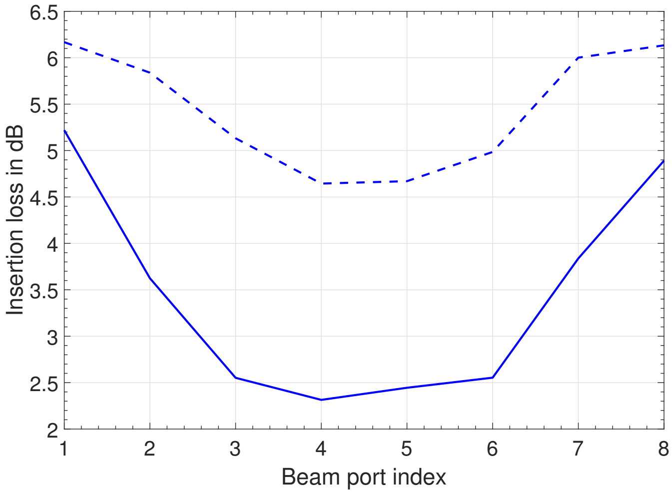
Figure 5.12: Measured BER versus measured SNR



(a) 5.0 GHz

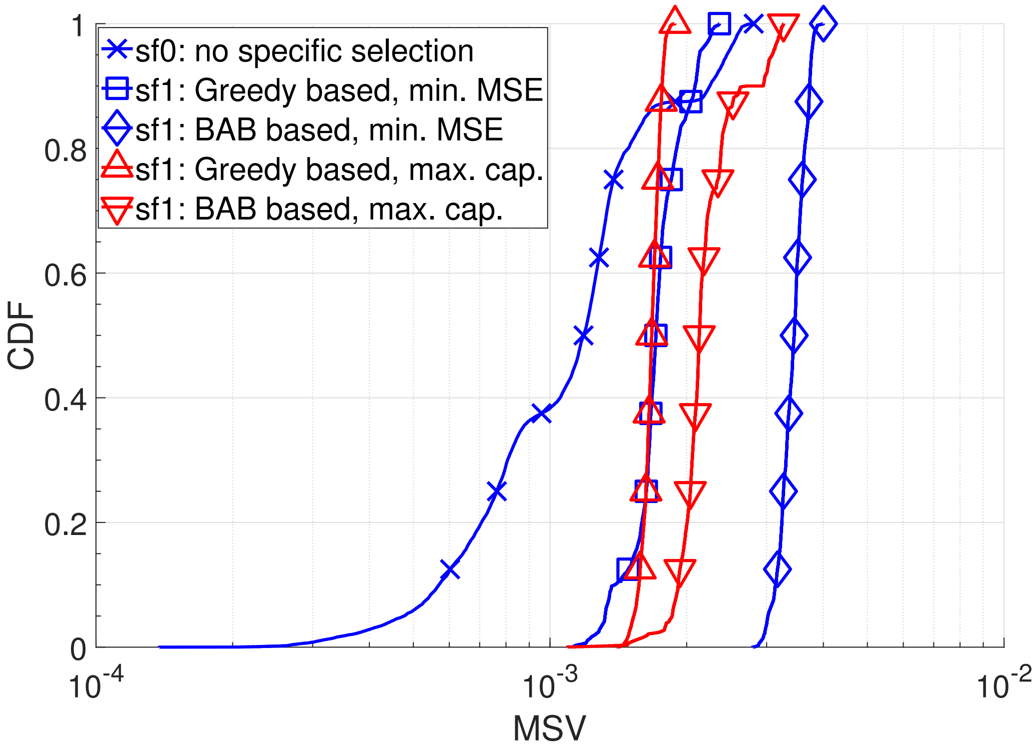


(b) 5.5 GHz

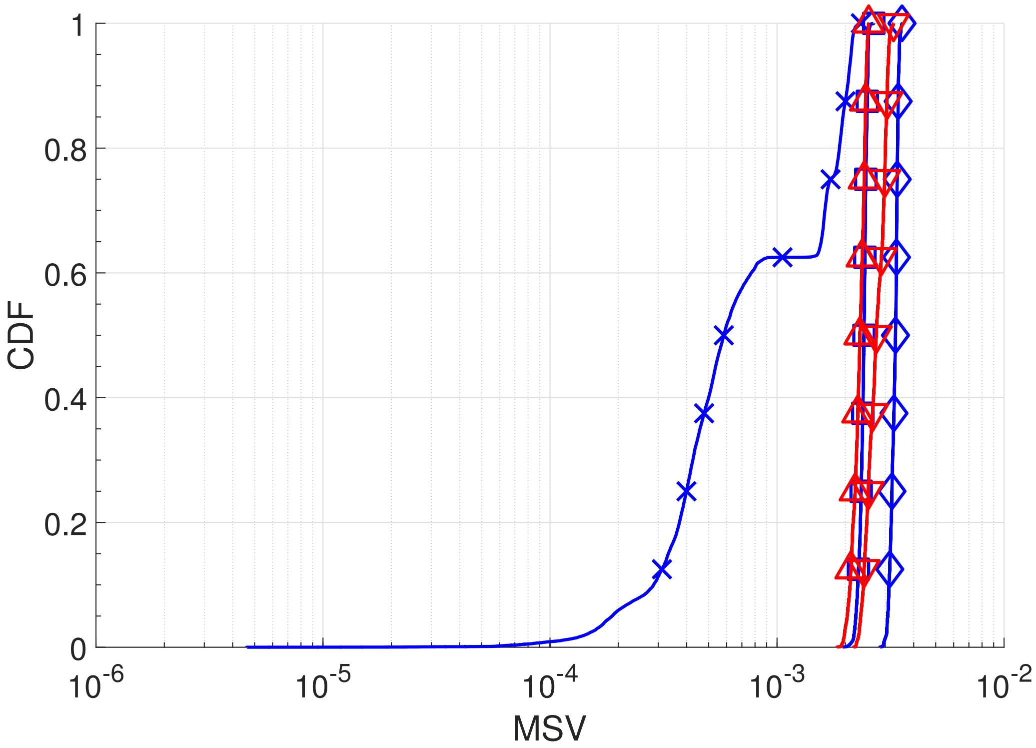


(c) 6.0 GHz

Figure 5.13: Measured insertion losses of the designed Rotman lenses with different substrate materials.



(a) Line of sight, SNR = 25.83 dB



(b) Non-line of sight, SNR = 16.69 dB

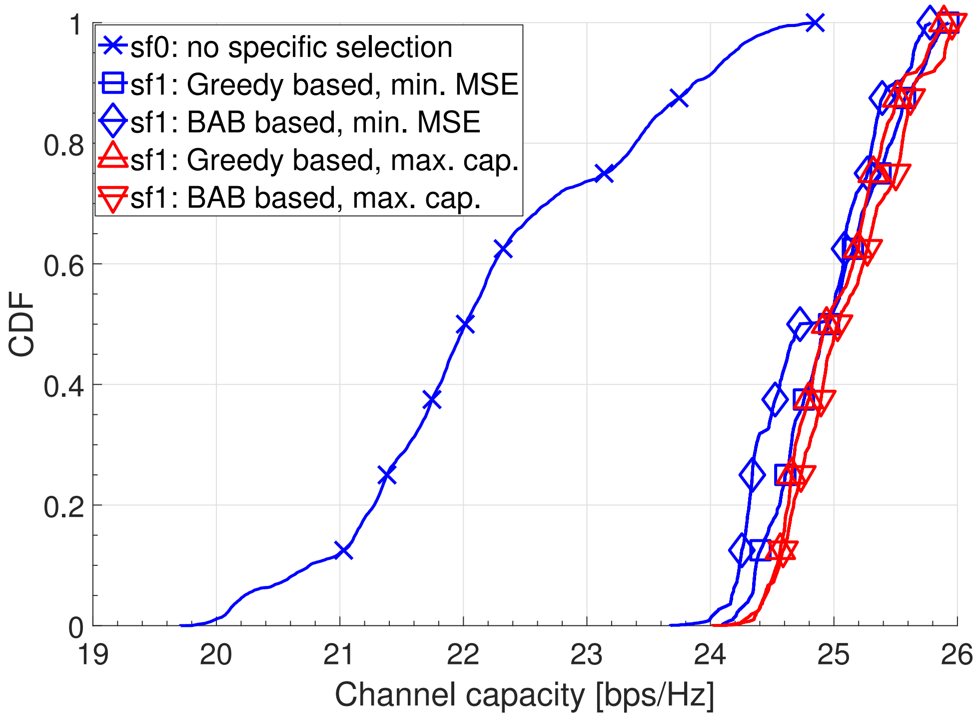
Figure 5.14: Distribution of the measured MSVs

as solving linear equation systems that independent linear equations are usually preferred. In the testbed measurement, we can get the measured channel matrix for each subcarrier, which is decomposed with Singular Value Decomposition (SVD) to get the MSV. Finally we have $N_{\text{used}} \times N_{\text{trials}} = 3000$ MSVs for the antenna selection system and $N_{\text{used}} \times M \times N_{\text{trials}} = 24000$ MSVs for the small-scale MIMO system. Fig. 5.14 shows the MSV distributions of the measured channel matrices. We can see that the MSV distributions of the antenna selection system is much steeper than the small-scale MIMO, especially have much shorter left tails than that of the small-scale MIMO. This mean that, with the antenna selection, the probability of the extreme small MSVs is very low, which is the main reason why the BER of the antenna selection system is much better than the small-scale MIMO in Fig. 5.12.

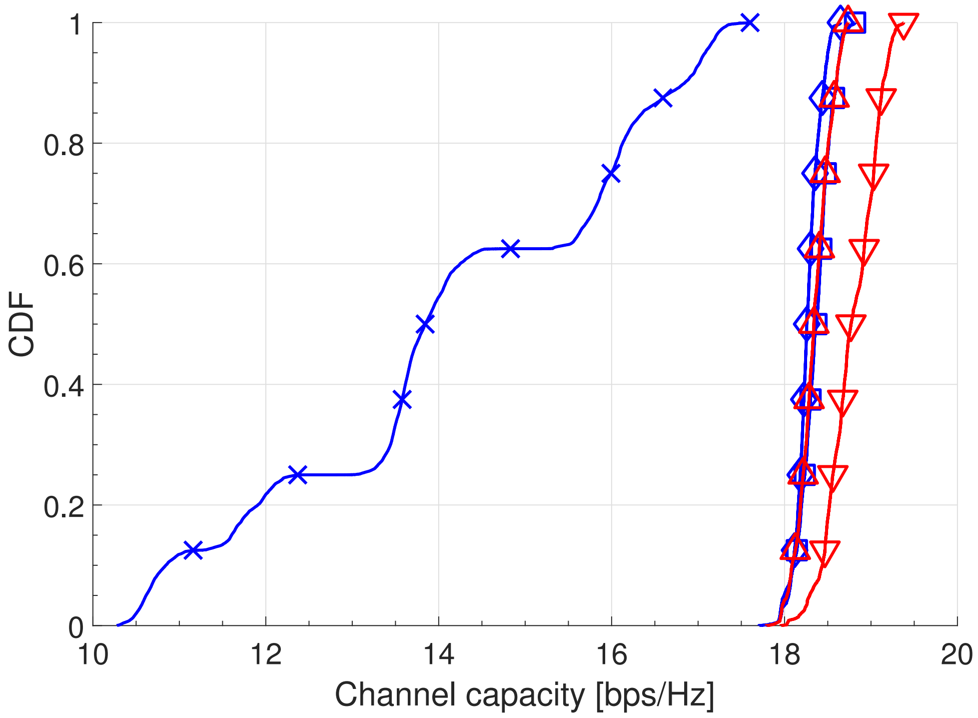
In the preceding measurement results, we show the BER-based metric to compare the system performance. Now let us have a look at another important system performance metric, i.e., channel capacity, which can be calculated by using the MIMO channel capacity equation (3.2). Similar to the derivations of the distribution of MSVs, we can get the distribution of the channel capacity of each subcarrier channel matrix. Fig. 5.15 shows the results. We can see that the BAB based method with the max-cap-oriented selection has the best performance, which is reasonable because it is optimal with respect to the maximum channel capacity. In contrast, the BAB based method with the min-MSE-oriented selection has worst performance in the antenna selection systems, which means that minimizing the MSE cannot guarantee the optimal performance on the channel capacity. Both greedy methods has the performance in-between the two BAB methods. Most importantly, all of the antenna selection systems has superior performance over the small-scale MIMO systems, which can be observed by comparing their distribution values and the slopes.

5.6 Conclusion

In this chapter, the testbed implementation of antenna selection and Rotman-lens-based hybrid beamforming systems are described for massive MIMO systems. Measurements are performed with the wideband signals under the 5-GHz indoor radio environment. Both line of sight and non-line of sight scenarios are considered. Measurement results show that the antenna selection and Rotman-lens-based hybrid beamforming systems work well in practice. Specifically, the MSE-oriented selection algorithms have better error rate performance than the the max-capacity-oriented selection algorithms under the LMMSE digital beamformer, and the MSE-oriented



(a) Line of sight, SNR = 25.83 dB



(b) Non-line of sight, SNR = 16.69 dB

Figure 5.15: Distribution of the channel capacity

selection algorithm based on BAB search has the best error rate performance. The antenna selection system outperforms the Rotman-lens-based hybrid beamforming system on the error rate performance, which is due to the high insertion loss introduced by the Rotman lenses with the lossy RO4350 substrate material. Both antenna selection and Rotman-lens-based hybrid beamforming systems have superior performance over the small-scale MIMO system under the same number of RF transceivers. We can conclude that antenna selection technology is very promising for the future 5-GHz band indoor massive MIMO systems. The Rotman-lens-based system potentially outperforms the antenna selection system when better material or better geometry design are used for the Rotman lens.

Chapter 6

Conclusions and Outlook

The low RF-complexity massive MIMO systems are studied in this thesis. In particular, two kinds of low RF-complexity massive MIMO systems are studied, i.e., massive MIMO antenna selection systems and massive MIMO hybrid analog-digital beamforming systems. Due to the unbalanced number of RF transceivers and antennas, both systems require specified RF networks to bridge the RF transceivers and antennas. Two RF network architectures, i.e., full-array and sub-array architecture, are considered during the study.

In the massive MIMO antenna selection systems, a tight upper bound on capacity using asymptotic theory on order statistic is derived at the large-scale limit. In addition, optimal antenna selection algorithms based on BAB algorithm are proposed. Both the theoretical and simulation results show that the sub-array antenna selection has performance close to that of the full-array antenna selection when the ideal RF network is assumed, and exhibits superior performance over the small-scale MIMO under the condition of the same number of RF transceivers. Considering the advantages of the low-loss and low-complexity sub-array switching network in practical implementation, the low RF-complexity massive MIMO based on sub-array antenna selection is promising for future wireless communication systems.

In the massive MIMO hybrid analog-digital beamforming systems, it is proposed to use Rotman-lens to perform analog beamforming, which is lower-complexity and more cost-efficient than the phase-shifter-based analog beamforming. In the full-array architecture, a single large-size Rotman lens is necessary; while in the sub-array architecture, multiple small-size Rotman lenses are required. The Rotman lenses are designed, fabricated and measured to validate the Rotman-lens-based hybrid beamforming system. In the digital domain, an LMMSE based digital beamformer is used to minimize the mean square error. Two beam selection algorithms, i.e., sub-optimal greedy search and optimal BAB search, are proposed to select a good subset of beams for both full-array and sub-array architectures. The simulation results show that the Rotman-lens-based system has close performance to the phase-shifter-based system with sub-array architecture when both ideal and non-ideal hardware is considered. In addition, the Rotman-lens-based system shows superior performance over the small-scale MIMO under the condition of the same number of RF transceivers. By comparing the Monte-Carlo simulations of antenna selection and Rotman-lens-based hybrid beamforming with realistic hardware, the former is advantageous in the rich-scattering channels with not too many antennas; while the latter advantageous in the sparse channels with a large number of antennas. It can be foreseen that the Rotman-lens-based hybrid beamforming systems are promising in the future mmWave systems.

To validate the low RF-complexity massive MIMO systems, experimental study is performed by setting up two testbeds with each consisting of 4×32 massive MIMO for both the antenna selection and hybrid beamforming systems. In the testbeds, a completed communication link from the transmitter to the receiver is designed, including hardware structure, signal structure, and the involved signal processing modules. The structures and the signal processing methods are illustrated in detail in this thesis. The testbed development methodology is also described. The measurement results validate both the antenna selection and hybrid beamforming systems under the realistic channels and imperfect channel estimation, synchronization errors, etc. Both systems have superior performance over the small-scale MIMO under the condition of the same number of RF transceivers. Impressively, such advantages are observed in the wideband case, although the systems are designed based on the narrowband system model. The Rotman-lens-based systems have noticeably performance loss compared to the antenna selection system, due to the large insertion loss of the used substrate material. The work in this thesis also shows that the Rotman lens can potentially outperform the antenna selection system when better substrate material is used. The experimental study in this thesis confirms the advantages of the low RF-complexity massive MIMO systems for practical applications.

Some further interesting and promising research topics are summarized as follows:

- In this thesis, the testbeds are measured with only the indoor radio environment. In the future, it will also be interesting to measure the testbeds with outdoor radio environments, since outdoor channels have very different characteristics from indoor channels. For example, outdoor channels are more frequency-selective, which causes more challenges to the wideband antenna/beam selection.

- The channel capacity of massive MIMO antenna selection systems is studied in this thesis; while the channel capacity of massive MIMO hybrid beamforming based on Rotman lens is not addressed, which can be one of the interesting topics in the future, especially considering the sparse nature of the channel, such as mmWave channel.

- In this thesis, especially for the hybrid beamforming systems based on Rotman lens, the 5-GHz-band is mainly focused on, due to its flexibility on hardware design and testbed setup. In fact, the Rotman-lens-based system will become more advantageous as the carrier frequency increases to higher frequency band, such as mmWave band which is of great interest to be studied due to its promising application in the future 5G cellular systems. The design a low-loss mmWave Rotman lens is a very interesting topic.

- In this thesis, the data channel is mainly focused on. In practical systems, some signaling channels, such as broadcast channels, control channels, etc., are indispensable. Unlike in the data channel where we need massive MIMO to form sharp and narrow beams to improve the link quality, in the signaling channels we usually need omni-directional radiation to cover all the position-unknown users in the cell. The design of the signaling channels in the low RF-complexity massive MIMO systems is a very important topic for the practical cellular systems.

List of Publications and Awards

Journal Papers

[1] Y. Gao, M. Khaliel, F. Zheng, and T. Kaiser, ”Rotman Lens Based Hybrid Analog-Digital Beamforming in Massive MIMO Systems: Array Architectures, Beam Selection Algorithms and Experiments,” accepted by IEEE Trans. Veh. Technol., Jun. 2017

[2] Y. Gao, H. Vinck, and T. Kaiser, ”Massive MIMO Antenna Selection: Array Architectures, Capacity Bounds and Optimal Antenna Selection Algorithms,” Submitted to IEEE Trans. Signal Process., in Major Revision, Jun. 2017

Conference Papers

[1] Y. Gao, H. Niu, and T. Kaiser, ”Massive MIMO Detection Based on Belief Propagation in Spatially Correlated Channels,” International ITG Conference on Systems, Communications and Coding (SCC), Hamburg, Feb. 2017

[2] Y. Gao, M. Khaliel, and T. Kaiser, ”Wideband Hybrid Analog-Digital Beamforming Massive MIMO Systems Based on Rotman Lens,” IEEE International Conference on Communication Systems (ICCS), Shenzhen, China, Dec. 2016 (**Best Paper Award**)

[3] Y. Gao, and T. Kaiser, ”Antenna Selection in Massive MIMO Systems: Full-array Selection or Subarray Selection?,” IEEE SAM conference, Rio, Brazil, Jun. 2016

[4] Y. Gao, W. Jiang, and T. Kaiser, ”Bidirectional Branch and Bound Based Antenna Selection in Massive MIMO Systems,” IEEE PIMRC conference, Hongkong, China, Sept. 2015

[5] Y. Gao, T. Kreul, and T. Kaiser, ”A Simplified RF Model to Investigate Impacts of RF Impairments on OFDM,” 18th International OFDM Workshop, Essen, Germany, Aug. 2014

[6] F. Sheikh, M. El-Absi, Y. Gao, and T. Kaiser, ”Terahertz Band: Analysis of Sounding Bandwidth and Center Frequency on Power Delay Profile Model,” Antennas and Propagation Conference (LAPC), Loughborough, UK, Nov. 2015

[7] A. Kabbani, A. Ali, H. Cao, A. B. Güven ,Y. Gao, S. Peethala, and T. Kaiser, ”Implementation Aspects of a DSP-Based LTE Cognitive Radio Testbed,” 10th International Conference CROWNCOM , Doha, Qatar, Apr. 2015

[8] T. T. Nguyen, H. Cao, A. B. Güven, Y. Gao, A. Kabbani, T. Kreul, and T. Kaiser, ”Robust Spectrum Sensing of DVB-T2 Signal Using the First Preamble Symbol,” International Conference on Advanced Technologies for Communications (ATC), Vietnam, Oct. 2013

[9] W. Jiang, H. Cao, T. T. Nguyen, A. B. Güven, Y. Wang, Y. Gao, A. Kabbani, M. Wiemeler, T. Kreul, F. Zheng, T. Kaiser, ”Key issues towards beyond LTE-Advanced systems with cognitive radio,” IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Darmstadt, Germany, Jun. 2013

[10] H. Cao, W. Jiang, T. T. Nguyen, A. B. Güven, Y. Wang, Y. Gao, A. Kabbani, M. Wiemeler, T. Kreul, F. Zheng, and T. Kaiser, ”The design of an LTE-A system enhanced with cognitive radio,” 2013 Proceedings of the 21st European Signal Processing Conference (EUSIPCO), Marrakech, Morocco, Sept. 2013

Patent

1. Y. Gao, “Method with low complexity for synchronization of wireless communication networks ”, German patent DE102015106292, filed

Awards

1. Best Paper Award of the 15th IEEE International Conference on Communication Systems (ICCS), Shenzhen, China, 2016

Bibliography

[1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021", *Cisco White Paper*, 2016

[2] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions, Future Generation Computer Systems", *Future Generation Computer Systems*, vol. 27, no. 7, pp. 1645-1660, 2013

[3] D. Miorandi, S. Sicari, F. Pellegrini, and I. Chlamtac, "Internet of things: Vision, applications and research challenges", *Ad Hoc Networks*, vol. 10, no. 7, pp. 1497-1516, 2012

[4] K. Mallinson, "The path to 5G: as much evolution as revolution", *3GPP Article*, May 2016. Availabe at *http://www.3gpp.org/news-events/3gpp-news/1774-5g_wiseharbour*

[5] ITU, "IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond ", *Recommendation ITU-R M.2083-0*, Sept. 2015. Availabe at *http://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf*

[6] W. Roh, J. Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Noncooperative Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results", *IEEE Commun. Magazine*, vol. 52, no. 2, pp. 106-113, Feb. 2014

[7] J. Mitola, and G. Q. Maguire, "Cognitive radio: making software radios more personal", *IEEE Personal Commun.*, vol. 6, no. 4, pp. 13-18, Aug. 1999

[8] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel, and S. Geirhofer, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas", *IEEE Commun. Magazine*, vol. 52, no. 2, pp. 82-89, Feb. 2014

[9] D. J. Costello, and G. D. Forney, "Channel coding: The road to channel capacity", *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1150-1177, Jun. 2007

[10] G. D. Forney, and G. Ungerboeck, "Modulation and coding for linear Gaussian channels", *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2384-2415, Oct. 1998

[11] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, "Modulation Formats and Waveforms for 5G Networks: Who Will Be the Heir of OFDM?: An overview of alternative modulation schemes for improved spectral efficiency", *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 80-93, Nov. 2014

[12] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H. P. Maye, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results", *IEEE Commun. Magazine*, vol. 49, no. 2, pp. 102-111, Feb. 2011

[13] S. Sesia, I. Toufik, and M. Baker, "LTE - The UMTS Long Term Evolution: From Theory to Practice, 2nd Edition", *Wiley*, Jul. 2011

[14] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels", *IEEE Trans. Sel. Area in Commun.*, vol. 21, no. 5, pp. 684-702, Jun. 2003

[15] 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

[16] F. Rusek, D. Persson, K.L. Buon, E.G. Larsson, T.L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays", *IEEE Signal Processing Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013

[17] J. Hoydis, S.T. 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. 32, no. 2, pp. 160-171, Feb. 2013

[18] A. Chockalingam, and B. S. Rajan, "MIMO detection", in *Large MIMO Systems*, Cambridge, UK, Cambridge Uni. Press, Feb. 2014

[19] G. J. Foschini, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas", *IEEE Personal Commun.*, vol. 6, no. 3, pp. 311-335, 1998

[20] I. E. Telatar, “Capacity of Multi-antenna Gaussian Channels”, *European Trans. Telecommun.*, vol. 10, pp. 583-595, 1999

[21] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems”, *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436-1449, Apr. 2013

[22] E. Bjornson, 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

[23] E. Alpaydin, “Introduction to Machine Learning”, *The MIT Press*, 2010

[24] Y. Gao, H. Vinck, and T. Kaiser “Massive MIMO Antenna Selection: Array Architectures, Capacity Bounds and Optimal Antenna Selection Algorithms, *submitted to IEEE Trans. Signal Process., in Major Revision*, Jun. 2017

[25] 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

[26] R.W. Heath, S. Sandhu, and A. Paulraj, “Antenna selection for spatial multiplexing systems with linear receivers”, *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 142-144, Apr. 2001

[27] D. Lu, and D. K. C. So, “Performance based receive antenna selection for V-BLAST systems”, *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 214-225, Jan. 2009

[28] Gorokhov Alexei, “Antenna Selection Algorithms for MEA Transmission Systems, in *IEEE International Conf. on Acoustics, Speech, and Signal Process (ICASSP)*, Orlando, USA, May 2002, pp 2857-2860

[29] M. Gharavi-Alkhansari, and A.B.Gershman, “Fast Antenna Subset Selection in MIMO Systems”, *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 339-347, Feb. 2004

[30] A. Dua, K. Medepalli, and A.J. Paulraj, “Receive Antenna Selection in MIMO Systems Using Convex Optimization”, *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2353-2357, Sept. 2006

[31] B.H. Wang, H.T. Hui, and M.S. Leong, “Global and Fast Receiver Antenna Selection for MIMO Systems”, *IEEE Trans. Commun.*, vol. 58, no. 9, pp. 2505-2510, Sept. 2006

[32] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, “Massive MIMO in Real Propagation Environments: Do All Antennas Contribute Equally?”, *IEEE Trans. Commun.*, vol. 63, no.11, pp. 3917-3928, Nov. 2015

[33] B. Fang, Z.P. Qian, W.S., and W. Zhong, “RAISE: A New Fast Transmit Antenna Selection Algorithm for Massive MIMO Systems”, *Wireless Personal Commun.*, vol. 80, no. 3, pp. 1147-1157, Feb. 2015

[34] K. Dong, N. Prasad, X.D. Wang, and S.H. Zhu, “Adaptive antenna selection and Tx/Rx beamforming for large-scale MIMO systems in 60 GHz channels”, *EUSIP J. Wireless Commun. and Network.*, vol. 2011, no. 1, pp. 59, 2011

[35] Gkizeli M. and Karystinos G.N., “Maximum-SNR Antenna Selection Among a Large Number of Transmit Antennas”, *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 891-901, Oct. 2014

[36] D. Mi, M. Dianati, S. Muhaidat, and Y. Chen, “A Novel Antenna Selection Scheme for Spatially Correlated Massive MIMO Uplinks with Imperfect Channel Estimation”, *IEEE 81st Vehicular Technology Conference (VTC Spring)*, Glasgow, Scotland, Mar. 2015

[37] Y. Gao, W. Jiang, and T. Kaiser, “Bidirectional Branch and Bound Based Antenna Selection in Massive MIMO Systems”, *IEEE Personal, Indoor and Mobile Radio Communications Conference (PIMRC)*, Hong Kong, China, Sept. 2015.

[38] A. Gorokhov, D. Gore, and A. Paulraj, “Receive antenna selection for MIMO flat-fading channels: theory and algorithms”, *IEEE Trans. Inf. Theory*, vol. 49, no. 11, pp. 2687-2696, Oct. 2003

[39] A.F. Molisch, M.Z. Win, and J.H. Winters, “Capacity of MIMO Systems With Antenna Selection”, *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1759-1772, Jul. 2005

[40] S. Sanayei, and A. Nosratinia, “Capacity of MIMO Channels With Antenna Selection”, *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4356-4362, Nov. 2007

[41] S. Y. Park, and D. J. Love, "Capacity Limits of Multiple Antenna Multicasting Using Antenna Subset Selection", *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2524-2534, Jun. 2008

[42] H. Li, L. Song, and M. Debbah, "Energy Efficiency of Large-Scale Multiple Antenna Systems with Transmit Antenna Selection", *IEEE Trans. Commun.*, vol. 62, no. 2, pp. 638-647, Feb. 2014

[43] D. Bai, P. Mitran, S. S. Ghassemzadeh, R. R. Miller, and V. Tarokh, "Rate of Channel Hardening of Antenna Selection Diversity Schemes and Its Implication on Scheduling", *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4353-4365, Oct. 2009

[44] R. Médez-Rial, C. Rusu, N. Gonzáez-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO Architectures for Millimeter Wave Communications: Phase Shifters or Switches?", *IEEE Access*, vol. 4, pp. 247-267, 2016

[45] Y. Gao, and T. Kaiser, "Antenna selection in massive MIMO systems: full-array selection or subarray selection?", *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Rio, Brasil, Jul. 2016.

[46] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Multi-Switch for Antenna Selection in Massive MIMO", *IEEE Global Communications Conference (GLOBECOM)*, San Diego, USA, Dec. 2015.

[47] A. Garcia-Rodriguez, C. Masouros, and P. Rulikowski, "Reduced Switching Connectivity for Energy-Efficient Large Scale Antenna Selection", [Online]. Available: *arXiv:1605.01549*, May 2016

[48] I.S. Gradshteyn, and I.M. Ryzhik, "Table of Integrals, Series, and Products", Boston, US, Academic Press, 7th Edition, 2007

[49] P. Embrechts, C. Klueppelberg, and T. Mikosch, "Modelling Extremal Events for Insurance and Finance", Berlin, Germany, Springer-Verlag Berlin Heidelberg, 1997, ch. 3, pp. 113-179

[50] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja, "A First Course in Order Statistics", New York: Wiley, 1992, ch. 8, pp. 205-222

[51] H. A. David, and H.N. Nagaraja, "Order Statistics, 3rd Edition", John Wiley & Sons, 2003, ch. 10, pp. 306-307

[52] M. Abramowitz, and I. A. Stegun, "Handbook of Mathematical Functions", National Bureau of Standards (NBS), United States, 1964, ch. 6, pp. 260-262

[53] J. J. Bartholdi III, "A good submatrix is hard to find", *Operations Research Lett.*, vol. 1, no. 5, pp. 190-193, 1982

[54] C. Ali, and M.I. Malik, "On selecting a maximum volume sub-matrix of a matrix and related problems", *J. Theoretical Comput. Sci.*, vol. 410, no. 47-49, 2009

[55] M. Narendra Patrenahalli, and K. Fukunaga, "A Branch and Bound Algorithm for Feature Subset Selection", *IEEE Trans. Comput.*, vol. C-26, no. 9, pp. 917-922, Sept. 1977

[56] Y. Cao, and V. Kariwala, "Bidirectional Branch and Bound for Controlled Variable Selection Part I. Principles and Minimum Singular Value Criterion", *J. Comput. and Chemical Eng.*, vol. 32, no. 10, pp. 2306-2319, 2008

[57] R. A. Horn, "Matrix analysis", *Cambridge University Press*, New York, NY, USA, 1986

[58] A.F. Molisch, and X.Y. Zhang, "FFT-based hybrid antenna selection schemes for spatially correlated MIMO channels", *IEEE Commun. Lett.*, vol. 8, no. 1, pp. 36-38, Jan. 2004

[59] X.Y. Zhang, A.F. Molisch, and S.Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection", *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091-4103, Nov. 2005

[60] J. Ahmadi-Shokouh, S. H. Jamali, and S. Safavi-Naeini, "Optimal Receive Soft Antenna Selection for MIMO Interference Channels", *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5893-5903, Dec. 2009

[61] O.E. Ayach, S. Rajagopal, S. Abu-Surra, Z.Y. Pi, and R.W. Heath, "Spatially Sparse Precoding in Millimeter Wave MIMO Systems", *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499-1513, Mar. 2014

[62] 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, pp. 831-846, Oct. 2014

[63] M. Kim, and Y.H. Lee, "MSE-based Hybrid RF/Baseband Processing for Millimeter Wave Communication Systems in MIMO Interference Channels", *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, Jun. 2014, pp. 2714-2720,

[64] T.E. Bogale, and L.B. Le, “Beamforming for Multiuser Massive MIMO Systems: Digital versus Hybrid Analog-Digital”, *IEEE Global Communications Conf.*, Austin, USA, Dec. 2014, pp. 4066-4071

[65] C. Chen, “An Iterative Hybrid Transceiver Design Algorithm for Millimeter Wave MIMO Systems”, *IEEE Wireless Commun. Lett.*, vol. 4, no. 3, pp. 285-288, Jun. 2015

[66] Weiheng Ni, Xiaodai Dong, and Wu-Sheng Lu, “Near-Optimal Hybrid Processing for Massive MIMO Systems via Matrix Decomposition”, *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 3922-3933, Aug. 2017

[67] E. Zhang, and H. Huang, “On Achieving Optimal Rate of Digital Precoder by RF-Baseband Codesign for MIMO Systems, *IEEE Veh. Technol. Conf.*, Vancouver, BC, Sept. 2014, pp. 1-5

[68] 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, Sept. 2014

[69] D. Kim, G. Lee, and Y. Sung, “Two-Stage Beamformer Design for Massive MIMO Downlink By Trace Quotient Formulation”, *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2200-2211, Jun. 2015

[70] J. Chen, and V.K.N. Lau, “Two-Tier Precoding for FDD Multi-Cell Massive MIMO Time-Varying Interference Networks”, *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1230-1238, Jun. 2014

[71] L. Liang, W. Xu, and X.D. Dong, “Low-Complexity Hybrid Precoding in Massive Multiuser MIMO Systems”, *IEEE Wireless Commun. Lett*, vol. 3, no. 6, pp. 653-656, Dec. 2014

[72] A. Li, and C. Masouros, “Hybrid Analog-Digital Millimeter-Wave MU-MIMO Transmission with Virtual Path Selection”, *IEEE Commun. Lett.*, vol. PP, no. 99, pp. 1-1, 2016

[73] X. Gao, L. Dai, S. Han, C. L. I, and R. W. Heath, “Energy-Efficient Hybrid Analog and Digital Precoding for MmWave MIMO Systems With Large Antenna Arrays”, *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998-1009, Apr. 2016

[74] F. Sohrabi, and W. Yu “Hybrid digital and analog beamforming design for large-scale MIMO systems, *IEEE Int. Conf. Acoust., Speech and Signal Process.*, South Brisbane, Queensland, Australia, Apr. 2015

[75] J. Wang, Z. Lan, C.-W. Pyo, T. Baykas, C.-S. Sum, M. A. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada et al., "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems ", *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390-1399, 2009.

[76] A. Sayeed, and N. Behdad, "Continuous aperture phased MIMO: Basic theory and applications ", *in Proc. of 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2010

[77] J. Brady, N. Behdad, and A. Sayeed, "Beamspace MIMO for Millimeter-Wave Communications: System Architecture, Modeling, Analysis, and Measurements", *IEEE Trans. Antennas Propag.*, vol. 6, no. 7, pp. 3814-3827, Jul. 2013

[78] S. V. Hum, and J. Perruisseau-Carrier, "Reconfigurable Reflectarrays and Array Lenses for Dynamic Antenna Beam Control: A Review", *IEEE Trans. Antennas Propag.*, vol. 62, no. 1, pp. 183-198, Jan. 2014

[79] P.V. Amadori, and C. Masouros, "Low RF-Complexity Millimeter-Wave Beamspace-MIMO Systems by Beam Selection", *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2212-2223, Jun. 2015

[80] X. Gao, L. Dai, and A. M. Sayeed, "Low RF-Complexity Technologies for 5G Millimeter-Wave MIMO Systems with Large Antenna Arrays ", [Online]. Available: *arXiv:1607.04559*, 2016

[81] W. Rotman, and R. F. Turner, "Wide-Angle Microwave Lens for Line Source", *IEEE Trans. Antennas Propag.*, vol. 11, no. 6, pp. 623-632, Nov. 1963

[82] R. Rotman, M. Tur, and L. Yaron, "True Time Delay in Phased Arrays", *IEEE Proc.*, vol. 104, no. 3, pp. 504-518, Mar. 2016

[83] R.C. Hansen, "Design Trades for Rotman Lenses", *IEEE Trans. Antennas Propag.*, vol. 39, no. 4, pp. 464-472, Apr. 1991

[84] P. S. Simon, "Analysis and synthesis of Rotman lenses", *AIAA International Communications Satellite Systems Conference*, Monterey, US, May 2004

[85] G. Tudosie, "Analysis and Design of Compact Passive Distribution Networks for Microwave Applications ", Dissertation, ETH, 2009

[86] L. Schulwitz, and A. Mortazawi, “A New Low Loss Rotman Lens Design Using a Graded Dielectric Substrate”, *IEEE Trans. Mirow. Theory Tech.*, vol. 56, no. 12, pp. 2734-2741, Dec. 2008

[87] A. Garcia-Rodriguez, V. Venkateswaran, P. Rulikowski, and C. Masouros, “Hybrid Analog-Digital Precoding Revisited Under Realistic RF Modeling”, *IEEE Wireless Commun. Lett*, vol. 5, no. 5, Oct. 2016

[88] D. M. Pozar, “Microwave Engineering”, *John Wiley & Sons*, 2012

[89] Y. Zeng, and R. Zhang, “Millimeter Wave MIMO With Lens Antenna Array: A New Path Division Multiplexing Paradigm”, *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1557-1571, Apr. 2016

[90] A. Garcia-Rodriguez, C. Masouros, and P. Rulikowski, “Reduced Switching Connectivity for Energy-Efficient Large Scale Antenna Selection”, *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2250-2263, May 2017

[91] A.F. Molisch, “Wireless Communications, 2nd edition”, *John Wiley & Sons*, Chichester, 2011

[92] A.M. Sayeed, “Deconstructing multiantenna fading channels”, *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563-2579, Oct. 2006

[93] J. J. Bartholdi III, “A good submatrix is hard to find”, *Operations Research Lett.*, vol. 1, no. 5, pp. 190-193, 1982

[94] R. A. Horn, “Matrix analysis”, *Cambridge University Press*, New York, NY, USA, 1986

[95] D. Lu, and D. K. C. So, “Performance based receive antenna selection for V-BLAST systems”, *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 214-225, Jan. 2009

[96] C.E. Chen, and W.H. Chung, “Computationally efficient near-optimal combined antenna selection algorithms for V-BLAST systems”, *Elsevier Digital Signal Process.*, vol. 23, no. 1, pp. 375¨C381, Jan. 2013

[97] M/A-Com Technology Solutions, “MAPS-011008-TR0500, Rev. V2”, [online]. Available: *http://cdn.macom.com/datasheets/MAPS-011008.pdf*

[98] Y. Gao, M. Khaliel, and T. Kaiser “Wideband hybrid analog-digital beamforming massive MIMO systems based on Rotman lens, *IEEE Int. Conf. on Communication Systems (ICCS)*, Shenzhen, China, Dec. 2016

[99] Y. Gao, M. Khaliel, F Zheng, and T. Kaiser “Rotman Lens Based Hybrid Analog-Digital Beamforming in Massive MIMO Systems: Array Architectures, Beam Selection Algorithms and Experiments, *accepted by IEEE Trans. Veh. Technol.*, Jun. 2017

[100] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein “Introduction to Algorithms”, *MIT Press*, 2009

[101] C. A. Balanis, “Antenna Theory: Analysis and Design”, *John Wiley & Sons*, Hoboken, 2005

[102] P. Schniter, and A. Sayeed, “Channel estimation and precoder design for millimeter-wave communications: The sparse way”, *48th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, 2014

[103] S. B. Weinstein, “The history of orthogonal frequency-division multiplexing ”, *IEEE Commun. Magazine*, vol. 47, no.11, pp. 26-35, Nov. 2009

[104] R. O’Neill, and L. B. Lopes “Envelope variations and spectral splatter in clipped multicarrier signals, *IEEE Personal, Indoor and Mobile Radio Communications Conference (PIMRC)*, Toronto, Canada, Sept. 1995

[105] J. J. van de Beek, M. Sandell, and P. O. Borjesson, “ML estimation of time and frequency offset in OFDM systems”, *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1800-1805, Jul. 1997

[106] L. Deneire, P. Vandenameele, L. van der Perre, B. Gyselinckx, and M. Engels, “A low-complexity ML channel estimator for OFDM”, *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 135-140, Feb. 2003

[107] F. X. Socheleau, A. Aissa-El-Bey, and S. Houcke, “Non data-aided SNR estimation of OFDM signals”, *IEEE Commun. Letters*, vol. 12, no. 11, pp. 813-815, Nov. 2008

[108] S. Boyed, and L. Vandenberghe “Convex Optimization”, *Cambridge University Press*, Mar. 2004

[109] G. Burel, “Statistical analysis of the smallest singular value in MIMO transmission systems”, *In Proc. of the WSEAS Int. Conf. on Signal, Speech and Image Processing (ICOSSIP)*, Greece, Sept. 2002