Novel Antenna Selection And Dimensionality Reduction In Massive Mimo Communication With The Exploitation Of Multiple Antenna Radio Channels

by

Muhammad Tausif Afzal Rana

A Dissertation Submitted In Fulfilment Of The Requirements For The Degree Of

DOCTOR OF PHILOSOPHY

School of Engineering
Faculty of Science and Engineering
Macquarie University
Sydney, Australia

December 8, 2017

Produced in $\mathbf{L^AT_EX}$ $2_\in$.

ABSTRACT

Massive multiple-input-multiple-output (MIMO) has the potential to offer a high system capacity in the fifth generation of wireless communication systems, exploiting the large number of degrees of freedom gained by utilising many transmit antennas at the base station. For a signal to be transmitted from an antenna element at the base station, the element needs to be connected to a radio-frequency chain which comprises Digital to-Analogue converters, a power amplifier and mixers etc. The total number of radio-frequency chains equals the number of antenna elements used for active transmission. Therefore, the radio-frequency switching matrix represents the hardware components required in antenna selection for interconnection of the radio-frequency chains with their selected antennas.

However, having a large number of antennas there are numerous challenges, in terms of hardware system complexity, large matrix sizes and signal correlation due to less space between antennas. To confront these problems, novel antenna-selection algorithms and dimensionality-reduction algorithms are proposed and spatial-correlation based channel models are explored.

Antenna selection algorithms based on central Principal Component Analysis (PCA) are developed, and antenna selection algorithms using non-central Principal Component Analysis (NPCA) and Linear Dependence Avoidance System (LDAS) are proposed. These algorithms reduce the correlation between the signals received by users. The performance of the proposed schemes with different antenna-selection algorithms and sum-capacity is evaluated.

Signal correlation between the antennas is modelled by using spatially correlated channel models such as the Kronecker Model and the Weichselberger Model. The correlation matrices at both ends of the link are approximated by using Power Azimuth Spectrum (PAS) models. These models give the analytical signal modelling in correlative environments used to test the proposed antenna-selection algorithms to study the system behaviour in different realistic environments.

Finally, the analysis of PCA is used that can reduce the size of huge matrices significantly, which saves a large number of computations as compared with the full-dimensional system. This thesis analyses the dimensionality reduction of large matrices using Floating Point Operation (FLOP) methods. This method is also implemented in the above-mentioned spatially correlated channel models to evaluate the throughput.

The results in this thesis are presented mainly in four parts: a) novel antenna-selection algorithms, b) application of novel antenna selection algorithms to spatially correlated channels, c) large matrix computations are reduced using dimensionality reduction techniques, d) analysis of matrix dimension reduction for spatially correlated channel models. This dissertation therefore presents the methodologies of lowering hardware complexity, reducing large matrices' dimensions and modelling signal correlation by using spatially correlated channel models that affect the design in the fifth generation of MIMO broadcast wireless communication systems. method in selecting

STATEMENT OF CANDIDATE

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree to any other university or institution other than Macquarie University.

I confirm that the dissertation is an original piece of research and it has been written by me.

I affirm that all information sources and literature used are indicated in the thesis.

. . . . . .

Muhammad Tausif Afzal Rana

May 25, 2018

ACKNOWLEDGMENTS

I stoop to praise and extend my gratification to Allah almighty who bestowed upon me the wealth of health, cerebration, deliberation, talented advisers and co-operative friends who enabled me to contribute to the infinite realm of knowledge .

I offer my special and sincere thanks to my principal supervisor Dr Rein Vesilo for his long illuminating guidance, illustrious advice, beneficial suggestions and supervision throughout the research work. Working with him has been a true privilege and a great experience for me. I staunchly believe that the completion of my dissertation would never have happened without his special personal interest, sparkling suggestions and gleaming criticism. He always stood by me whenever I encountered any difficulties during my postgraduate studies.

I also benefited from a number of people during my graduate studies. I am exceedingly greatful to my associate supervisor, Dr Iain B. Collings, for his guidance and support during this study. I should also like to thank Dr Muhammad Masroor Ahmed for his research collaboration, guidance and valuable suggestions throughout this study. I owe my sincere thanks to Dr Keith Imrie for proofreading this thesis.

I sincerely appreciate the financial support from the International Macquarie University Research Excellence Scholarship (iMQRES), and the Macquarie University Higher Degree Research Fund, that made this study and my stay in Australia possible.

I am grateful to all my fellows and friends in the research group whose encouragement and inspiration propelled me in the right direction. I offer my earliest thanks to Mr Aziz-ul-Rehman and Mr Muhammad Zeeshan Baig for being my best house-mates, coffee and lunch mates and listening to all my ideas patiently and helping me in both technical and non-technical talks.

I would like to thank my affectionate parents, my mother Hafsa-un-Nisa and my father Dr Muhammad Afzal Rana, who have always been proud of me and were always there for continuing support and encouragement in my whole studies. I should also like to thank my siblings Dr Abida Siddiqa, Dr Sadia Siddiqa, Dr Bushra Siddiqa and Mr Muhammad Jawad Afzal Rana, for being my best friends. I also owe my deepest appreciation to my grandfather, my teacher Qari Abd-ur-Rauf (Late) and Mufti Sheikh Rehan, who always prayed for me in my hard times.

RELATED PUBLICATIONS

Conference Papers

[C01] **M. T. A. Rana**, Rein Vesilo, and Iain B. Collings "Antenna selection in massive MIMO using non-central principal component analysis", 26^{th} International Telecommunication Networks and Applications Conference (ITNACT), 2016.
(**Dunedin**, **NewZealand**).

[C02] **M. T. A. Rana** and Rein Vesilo "Channel Complexity Reduction in Massive MISO Using Principal Component Analysis", 17^{th} International Symposium on Communications and Information Technologies (ISCIT), 2017.
(**Cairns**, **Australia**).

[C03] **M. T. A. Rana**, Rein Vesilo, and Ahsan Saadat "Antenna Selection for Massive MIMO Kronecker Channel Models Using Non-Central Principal Component Analysis", 27^{th} International Telecommunication Networks and Applications Conference (ITNACT), 2017.
(**Melbourne**, **Australia**).

Journal Papers

[J01] **M. T. A. Rana** and Rein Vesilo "Channel Complexity Reduction Analysis in Spatially Correlated Massive MISO System Using Principal Component Analysis", ready to be submitted to ISI-index journal.

[J02] **M. T. A. Rana** and Rein Vesilo "Antenna Selection for Massive MIMO in Spatially Correlated Channel Models Using Non-Central Principal Component Analysis", ready to be submitted to ISI-index journal.

HOLY PROPHET (S. A. W.)

The Greatest Social Reformer

&

My Worthy Parents

TABLE OF CONTENTS

Abstract v

Acknowledgments xi

Related Publications xiii

Table of Contents xvii

List of Tables xxiii

List of Figures xxv

List of Acronyms xxix

Chapter 1: Introduction 1

1.1 Introduction 1

1.2 MIMO Systems 3

1.3 Types of Massive MIMO Operation 4

1.4 MIMO Broadcast Channels 5

1.5 Massive MIMO and Its Benefits 6

1.6 Motivation 8

1.7 Aims and Approaches 11

1.8 Thesis Overview 13

Chapter 2: Background and Related Work 17

2.1 Antenna Selection in Massive MIMO 18

2.1.1 Conventional Antenna Selection in Massive MIMO Broadcast Channels 18

2.2 MIMO Channel Models 23

2.2.1 Physical channel modelling 24

2.2.2 Non-physical channel modelling 24

2.3 Channel Modelling 25

2.3.1 Analytical Channel Models 26

2.3.2 Spatial-Correlation Based Models 28

2.3.2.1 The Independent Identically Distributed Rayleigh Fading Model 29

2.3.2.2 The Kronecker Model 30

2.3.2.3 The Weichselberger model 31

2.4 Precoding in MIMO 33

2.4.1 Non-Linear Precoding 35

2.4.2 Linear Precoding 35

2.4.3 MIMO Channel Decomposition 37

2.4.4 Power Allocation in MIMO System 39

2.5 Computational Complexity 41

2.6 Principal Component Analysis 42

2.7 Summary of the Chapter 43

Chapter 3: Antenna Selection in Massive MIMO Communication 45

3.1 Introduction 46

3.2 Related Work 48

3.3 System Model 49

3.4 Principal Components Analysis (PCA) 52

3.4.1 Population-Based Central PCA 52

3.4.2 Population-Based Non-Central PCA 54

3.4.3 Sample-Based PCA 54

3.5 Antenna Selection 55

3.5.1 Analysis using Non-Central PCA 58

3.6 Simulation Results and Discussion 59

3.7 Summary of the Chapter 69

TABLE OF CONTENTS

Chapter 4: Spatial Structure of Multiple Antenna Radio Channels **71**

4.1 Introduction 72

4.2 Related Work 73

4.3 Stochastic MIMO Channel Model 75

4.3.1 Kronecker Model 75

4.3.2 Weichselberger Channel Model 76

4.3.3 n^{th} power cosine PAS Model 79

4.3.4 Uniform PAS Model 80

4.4 Simulation Results and Discussion 82

4.5 Summary of the Chapter 90

Chapter 5: Dimensionality Reduction of Large Matrices **91**

5.1 Introduction 92

5.2 Linear Precoding 94

5.2.1 Zero Forcing Precoding 96

5.2.2 Minimum Mean Square Error Precoding 96

5.3 Channel Complexity Reduction via PCA 97

5.4 Computational Complexity 100

5.5 Simulation Results and Discussion 102

5.6 Summary of the Chapter 106

Chapter 6: Conclusions and Future Work 111

6.1 Conclusions 111

6.2 Future Directions 114

REFERENCES 116

LIST OF TABLES

5.1 Computational Complexity in FLOPs for multiple system dimensions 107

5.2 Computational time in seconds for multiple system dimensions 107

LIST OF FIGURES

1.1 Global mobile data traffic. 2

1.2 Global mobile device and connections growth. 2

1.3 Illustration of the Proposed Dissertation Methodology 16

2.1 MIMO channel and propagation models 22

2.2 Correlated MIMO Channel Capacity vs i.i.d. MIMO Channel 29

2.3 Sum-Capacity of Non-Linear (DPC) and Linear (ZF & MMSe) 34

2.4 MIMO Channel Split into Virtual Scalar and Independent Channels 37

2.5 SVD of MIMO Channel 38

2.6 Optimal Power Allocation Using Water-filling Principle 39

3.1 The MU-MISO System Model 48

3.2 Mean sum-capacity versus P ($N_s = 10$) ($\mathbf{H}_z$) using i.i.d. Rayleigh fading channel model 59

3.3 Mean sum-capacity versus P ($N_s = 6$) ($\mathbf{H}_z$) using i.i.d. Rayleigh fading channel model 60

3.4 (a) Mean sum-capacity versus P ($N_s = 10$) ($\mathbf{H}_z$) (b) Mean sum-capacity versus P ($N_s = 10$) ($\mathbf{H}_z$) (Conjugate beamforming) 62

3.5 (a) Mean sum-capacity versus P ($N_s = 6$) ($\mathbf{H}_z$) (b) Mean sum-capacity versus P ($N_s = 6$) ($\mathbf{H}_z$) (ZF) 63

3.6 Mean sum-capacity versus P ($N_s = 8$) ($\mathbf{H}_m$) using i.i.d. Rayleigh fading channel model 64

3.7 Mean sum-capacity versus N_s, $P = 10,000$, ($\mathbf{H}_z$) using i.i.d. Rayleigh fading channel model 65

3.8 Sum-capacity versus N_{PCA} ($N_s = 10$) ($\mathbf{H}_z$) using i.i.d. Rayleigh fading channel model 65

3.9 Mean sum-capacity versus N_{PCA} ($N_s = 8$) ($\mathbf{H}_m$) using i.i.d. Rayleigh fading channel model 66

3.10 Mean sum-capacity versus P ($N_s = 10$) ($\mathbf{H}_z$) using i.i.d. Rayleigh fading channel model 66

3.11 Channel estimation mean sum-capacity versus P 68

3.12 Estimated Error versus P 68

4.1 Two-antenna array in a scattering environment 74

4.2 Downlink MIMO system 78

4.3 Mean sum-capacity versus P ($N_s = 10$) using Kronecker channel model 81

4.4 Mean sum-capacity versus P ($N_s = 6$) using Kronecker channel model 82

4.5 Mean sum capacity versus N_s using Kronecker channel model 83

4.6 Spatial correlation coefficient using Kronecker channel model 84

4.7 Mean sum-capacity versus P ($N_s = 10$) using Weichselberger channel model . . 85

4.8 Mean sum-capacity versus P ($N_s = 10$) using Weichselberger channel model . . 86

4.9 Number of RF chains using Weichselberger channel model 87

4.10 Number of RF chains using Weichselberger channel model 87

4.11 Mean sum-capacity difference (bps/Hz) Versys Power $P = 10000$ (Weichselberger channel model Vs Kronecker channel model) 88

4.12 Mean sum-capacity difference (bps/Hz) Vs Power $P = 4000$ (Weichselberger channel model Vs Kronecker channel model) 88

5.1 The MU-MISO System Model 95

5.2 Typical decay of the ordered PCs 98

5.3 Sum rate simulations using PCA 101

5.4 Computational complexity in FLOPs for massive MU-MISO system 103

5.5 Decreasing-variance trend of PCA-based channel matrix $\hat{\mathbf{H}}_{\mathcal{J}_s}$ 104

5.6 Exponential decay of the largest PCs 105

5.7 (a) Sum-capacity simulations using PCA with 16×32 matrix dimension (b) Sum-capacity simulations using PCA with 32×64 matrix dimension 108

5.8 (a) Sum-capacity simulations using PCA with 64×128 matrix dimension (b) Sum-capacity simulations using PCA with 128×256 matrix dimension 109

LIST OF ACRONYMS

AWGN	Additive White Gaussian Noise
AS	Antenna Selection
BLAST	Bell Labs Layered Space-Time
BS	Base Station
BC	Broadcast Channels
CSI	Channel State Information
CSIT	Channel State Information at Transmitter
DOA	Direction of Arrival
DOD	Direction of Departure
DPC	Dirty Paper Coding
dB	Decibels
FLOP	Floating point operations
i.i.d.	Independent and Identically Distributed
LDAS	Linear Dependent Avoidance System
LTE	Long-term-evolution

LOS	Line of sight
MUI	Multi-user interference
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
MS	Mobile Station
MMSE	Minimum-Mean-Square-Error
NLOS	Non-Line-Of-Sight
$\mathbf{N_s}$	Number of radio-frequency chains
PAS	Power Azimuth Spectrum Model
PCA	Principal Component Analyis
PC's	Principal Components
NPCA	Non-Central Principal Component Analyis
RF	Radio Frequency
SDMA	Space Division Multiple Access
SISO	Single-Input Single-Output
SNR	Signal-to-Noise-Ratio
SINR	signal-to-Interference plus Noise Ratio
SC	Spatial Correlation

SVD	Singular Value Decomposition
WLAN	Wireless Local Area Network
ZMCSCG	Zero Mean Circularly Symmetric Complex Gaussian
ZF	Zero Forcing

Begin at the beginning" the King said, very gravely,"

and go on till you come to the end: then stop.

Lewis Carroll (1832-1898)

CHAPTER 1

INTRODUCTION

1.1 Introduction

In the last decade, a remarkable expansion of electronic devices such as smart phones, tablets, laptops, data traffic (both mobile and fixed) and various data-consuming wireless devices has occurred with an exponential increase. It is expected that in future the demand of wireless data traffic will be even more than at present [1–3]. By 2018, it is anticipated that Global mobile data traffic will accelerate to 17 exabytes per month, which is about a three-fold increase over 2016. The demand is for mobile data traffic, and the number of connected devices will grow to 11.6 billion by 2021 as shown in Figures 1.1 and 1.2.

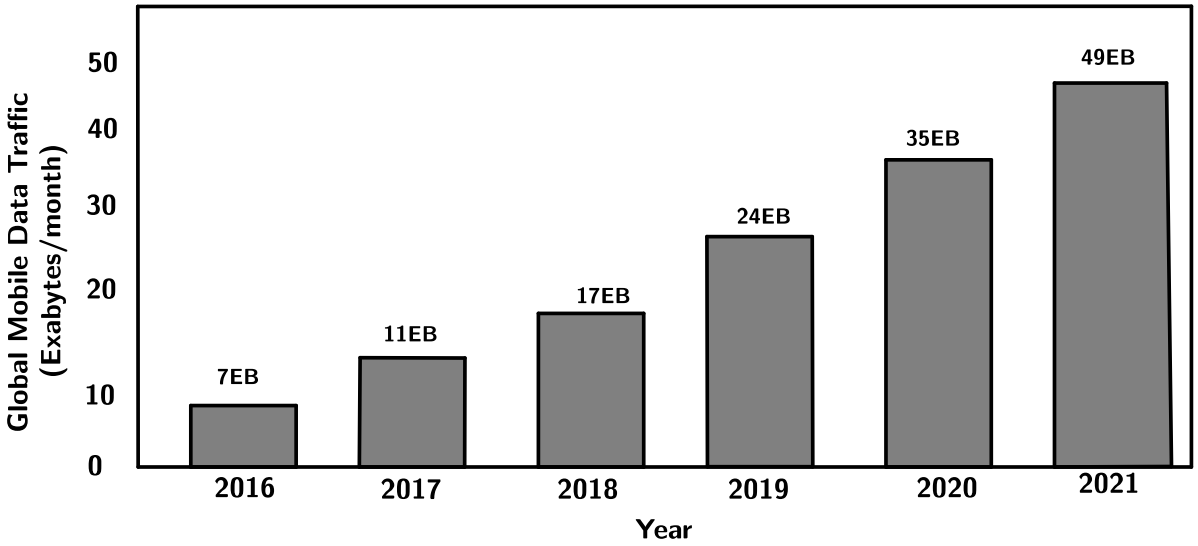


Figure 1.1: Global mobile data traffic.
(Source: Cisco [3])

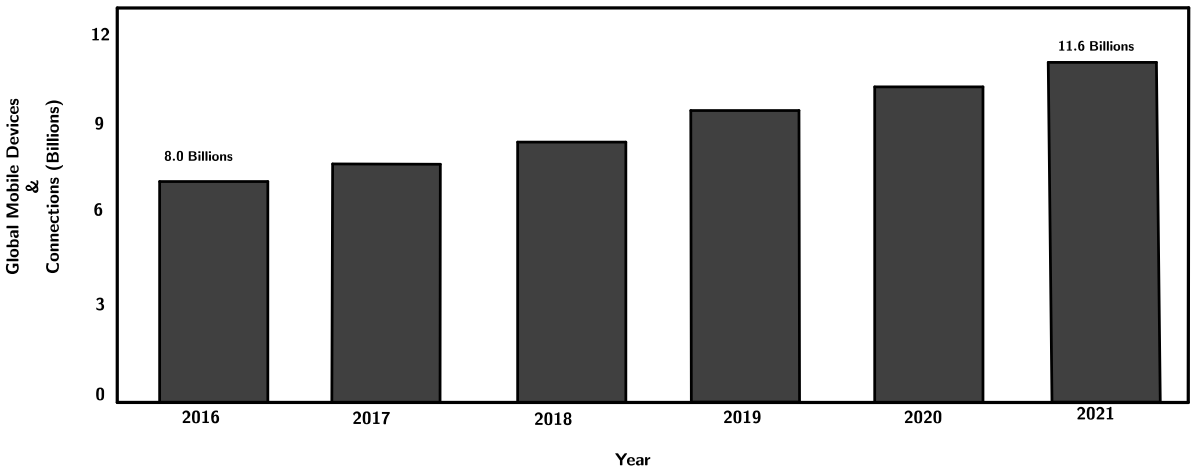


Figure 1.2: Global mobile device and connections growth.
(Source: Cisco [3])

It is clear from the figure that for the future wireless communication generations, new technologies are required to meet the high data traffic demands, and the important metric to accomplish this task is the wireless sum-rate/sum-capacity or throughput measured in bits/s and defined as:

$$Throughput = Bandwidth\ (Hz) \times Spectral\ efficiency\ (bits/s/Hz).$$

It is evident from the above equation that throughput will be enhanced by increasing the bandwidth or spectral efficiency or both to meet the high data rate in future wireless communication. In wireless communication systems, capacity enhancement has been always a hot topic. The capacity demand could be realised in a land-mobile radio telephone that uses a frequency of 2 MHz. This radiophone was installed for Police cars in 1921 by the Detroit Police Department, and the civilian use of wireless technology had begun. However, a low frequency creates a channel mismatch problem which was resolved by using high-frequency bands in 1933. The pioneering work of *Claude Elwood Shannon* in 1948 gave a birth to a new era in information and communication theory. He presented a research paper "*A mathematical theory of communication*" that opened a potential for modern wireless communication research.

1.2 MIMO Systems

Traditionally, there are two major components in wireless communication which can provide high system throughput or high data rates to meet current capacity requirements. These include the bandwidth measured in *hertz (Hz)* and/or the total transmission power measured in *Watts (W)*, but they are limited and expensive. However, by

using multiple antennas both at the base station (BS) and at the mobile station (MS), it is possible to increase the system throughput many fold regardless of any additional requirements in transmission power and expensive bandwidth. The technology with multiple numbers of transmit and receive antennas at BS and MS in a communication system is known as multiple-input multiple-output (MIMO) systems.

Communication between one transmitter and one receiver is referred as a *point-to-point* single-input single-output (SISO) wireless communication system. However, a MIMO system can yield a great capacity enhancement without increasing the channel bandwidth and power as compared to a conventional point-to-point SISO system [4–6]. Throughput is the key feature in the analysis of wireless systems, and the first MIMO schemes were analysed in point-to-point or single-user MIMO. In this case, a multiple number of antennas at base station and mobile station are used. A remarkable spectral efficiency for a wireless system in a rich scattering environment with multiple antennas is accurately reported in [4, 6, 7]. The performance of MIMO systems can be enhanced using an antenna selection technique in which signals from the chosen subset are processed further by the accessible radio-frequency chains. The key methodology in enhancing *Long Term Evolution* (LTE) and *Fifth Generation* (5G) is SU-MIMO, that achieves *Downlink* (DL) and *Uplink* (UL) data rates [8, 9]. In addition, LTE-advanced targets high DL and UL data rates, and the approach in this case is multi-user MIMO [10].

1.3 Types of Massive MIMO Operation

There are two main modes of operation for massive MIMO called spatial multiplexing based massive MIMO and beam-steering based massive MIMO. Spatial multiplexing massive MIMO operates in time division duplexing (TDD) mode, whereas

beam-steering massive MIMO operates in frequency division duplexing (FDD) mode. In a TDD multi-user massive MIMO system, multiple users are served in the same time-frequency slot by deploying a large number of antennas at the base station and using linear precoding and combining techniques to manage inter-user interference. The user nodes in a TDD massive MIMO are typically single-antenna terminals. Channel estimation in this case is not done at user nodes but is done at the base station through uplink pilots transmitted by user nodes. By exploiting the channel reciprocity in TDD, the downlink channel can be determined. The pilot sequence length is independent of the number of antennas at the base station. In the downlink, users do not have access to the exact channel state information (CSI) and rely on statistical CSI instead.

In FDD beam-steering massive MIMO, a single user is served in a given time-frequency slot via beamforming using the large antenna array. This approach is more suited for millimetre-wave systems. The downlink channel estimation at the user nodes can be obtained by exploiting channel sparsity and feeding the estimates back to the base station. The focus in this thesis is on spatial multiplexing based massive MIMO.

1.4 MIMO Broadcast Channels

Multi-user multiple-input multiple-output (MU-MIMO) is a system where an array of antennas at a base station (BS) communicates with multiple mobile stations (MS) equipped with single or multiple antennas. The MIMO broadcast channel (BC) generally falls into a class of non-degraded Gaussian broadcast channels [11]. In a non-degraded MIMO-BC, the capacity region is an unsolved problem, however, an achievable capacity by region is defined in [12]. The authors in [13, 14] investigate

the Gaussian MIMO-BC capacity region using the dirty paper coding (DPC) algorithm in [15], where the known interference of the users has been pre-subtracted at the base station. It is demonstrated in [13, 14] that with two users and two antennas at transmitter then the MIMO-BC throughput is equal to the maximum achievable sum-rate of the DPC capacity region.

Following the theory in [5,6], the MIMO systems concept leads to a technique known as a MU-MIMO broadcast (BC) system. This scheme has been of significant research interest over the past decade [16], for example, the throughput of MIMO-BC, both in capacity region and sum-rate is larger than for a point-to-point single antenna system. However, the disadvantages associated with this system include hardware complexity, receiver complexity and space limitation. A possible solution to alleviate these issues in point-to-multipoint MIMO-BC systems is the use of a smaller number of radio-frequency chains than of available antennas. A low-cost switching circuit can be used to connect the selected subset of transmit antennas to the available radio-frequency chains.

1.5 Massive MIMO and Its Benefits

The use of a high number of antennas at the base station gives more degrees of freedom to communicate with users simultaneously with the same time-frequency resources to achieve a high system capacity. However, due to the high signal dimensions in large antenna arrays, the conventional signal-processing techniques become complex. To achieve a high multiplexing gain with reduced signal-processing complexity, and a cost-effective hardware implementation, the authors in [17] showed that the deployment of a high number of base station antennas makes linear processing simple and almost optimum. More precisely, as the number of BS antennas grows

large the effect of fast signal attenuation, interference between the cells and uncorrelated noise disappears by using an appropriate scheme. As an illustration, in an uplink scenario maximum-ratio combining (MRC) can be used, and maximum-ratio transmission (MRT) can resolve problems in the downlink system [18]. A MU-MIMO system serving ten or more users with the same time-frequency resources, with a hundred or more antennas at the transmitter, is known as a Massive MIMO system. The major benefits of Massive MIMO systems are:

(1) **High spectral efficiency**: The massive MIMO scheme inherits all the benefits of conventional MU-MIMO technology, meaning that with N_t transmit antennas at the BS and N_r receive antennas, a diversity of order N_t and a multiplexing gain of $min(N_r, N_t)$ can be achieved. A huge spectral efficiency can be obtained by increasing both N_r and N_t.

(2) **Simple signal processing**: The first key feature in a massive MIMO system is the propagation environment, when the channel vector between the users and the BS is nearly orthogonal then the large number of base station antennas over the number of users yields favourable propagation. In such propagation environments simple linear signal processing, i.e. linear precoding in the downlink and linear decoding in the uplink, can be used to remove the effect of interference and noise. Therefore, a system becomes nearly optimal by using a simple linear processing approach.

The second important property in a massive MIMO systems is that, with the high number of antennas at the base station, the spatial diversity phenomenon leads toward channel hardening, meaning that the channels become, nearly deterministic, and thus the effect of small-scale fading is averaged out.

The third useful factor in massive MIMO systems is scheduling diversity, power

control that mitigates the need for downlink pilots and simplifies the signal processing significantly.

1.6 Motivation

Antenna Selection and Radio Frequency Chains

Massive multiple-input-multiple-output (MIMO) has the potential to offer high system capacity in the fifth generation of wireless communication systems, exploiting the large number of degrees of freedom gained by utilizing many transmit antennas at the base station. However, the full-digital implementation of the massive MIMO systems, i.e., each antenna connected to a dedicated Radio Frequency (RF) transceiver, is very cost-prohibitive and power-hungry. To improve both cost and power efficiency, analog signal processing can be additionally introduced to reduce the signal dimension in the RF-analog-domain, thus allowing a reduced number of RF transceivers. This paradigm of hybrid analog-digital signal processing is being widely studied in recent years for massive MIMO systems, such as hybrid analog-digital beamforming, antenna selection, etc, see [19] and references therein. In the RF-analog processing, specific types of RF-analog components are required. In contrast to the phase shifters that are usually used in the hybrid analog-digital beamforming, simpler RF switches are used in the antenna selection system, which have less cost and power consumption than the phase shifters [19].

Recent studies [20, 21] show several approaches to alleviate the radio-frequency complexity in massive MIMO system. Besides the benefits in [20, 21], the deployment of a large number of transmit antennas at the base station incurs the problem of hardware complexity, large power consumption, and thus the cost of the communication system. The hardware complexity is due to the use of multiple radio-

frequency chains, which include low-noise amplifiers, down-converters, digital-to-analogue converters, analogue-to-digital converters and so on [22, 23]. The hardware complexity can be minimised by using reduced connectivity of the radio-frequency chains associated with the antenna elements. The authors in [24, 25] show that the optimal selection of a subset of the available transmit antennas reduces the hardware complexity in MIMO systems.

Antenna selection improves system performance by exploiting spatial selectivity, that is, with more antennas than radio-frequency chains in the system, the best antennas could be chosen and switched to a limited number of radio-frequency chains. Indeed, this thesis will demonstrate that close to optimal sum throughput can be achieved using appropriate antenna selection techniques. However, this comes with a price of increased implementation overhead and the complexity of the transceiver design for massive MIMO systems [26, 27]. To alleviate the radio-frequency complexity in massive MIMO, it is argued in [26] that cost-efficient antenna selection strategies can be employed.

Furthermore, a system becomes nearly optimal by using a simple linear processing approach, that is linear precoding in the downlink scenario. One way to achieve a large sum-capacity of MIMO broadcast channels is to employ a precoder at the transmitter side to reduce the multi-user interference (MUI). Precoders exploit the channel state information at the transmitter (CSIT) to adapt the transmission strategies or variables such as the 'direction' and 'magnitude' (or power) of the transmission for each user's data symbols. Perfect CSIT will be an ideal case and is often considered in the performance analysis of MIMO communication systems as a benchmark for practical scenarios.

Signal Modelling and Antenna Selection in Correlated Channel Models

In massive MIMO systems, a base station is equipped with a large number of transmit antennas serving several simultaneous single antenna users. Thus, by using the multiple antennas, the base station transmits independent data streams to serve multiple users simultaneously. However, it is difficult to place an increasing number of antennas in a limited space, and the antenna space limitation causes a high spatial correlation between the antennas, further resulting in systems performance degradation [28]. Therefore it is important to study the impact of correlation in massive MIMO systems.

Researchers in [29] show a linear increase in the throughput of narrow-band MIMO systems under the assumption of an idealised channel model. Such models [29] represent a rich scattering environment that assumes independent and identically distributed (i.i.d.) channel coefficients of a Rayleigh-fading channel model. Such idealised channel models oversimplify the throughput problems. Thus, in a real propagation environment, it is shown in various analytical and experimental findings that, in the presence of spatial correlation (SC) in MIMO channels, the system throughput is expected to be less than that of i.i.d. channel models [30, 31].

To simulate correlated MIMO channels it is important to quantify the effect of correlation on system capacity. Thus, spatial structures are used to determine the Direction of Departure (DOD) and Direction of Arrival (DOA) of receive and transmit correlation matrices. These correlation matrices are approximated at the receiver and transmitter by using Power Azimuth Spectrum (PAS) models [32] . Once the MIMO channel matrix is modelled in different realistic environments, the proposed antenna-selection algorithms are implemented and the system throughput is analysed.

Dimensionality Reduction in Massive MIMO

In massive MIMO, as the number of antennas increases, so does the matrix size. The problem here is the computation of large precoding matrices in massive MIMO which becomes a bottleneck in the transmission of data. A possible solution is the use of a data analysis tool, namely "Principal Component Analysis" (PCA) [33], that reduces a large matrix's size without sacrificing its effectiveness. In fact, PCA transforms high-dimensional data into lower-dimensional data, which reduces a high system computational complexity in terms of Floating Point Operations. In addition, spatial-correlation models are explored to analyse the computational complexity involved in large-dimensional matrices.

1.7 Aims and Approaches

The aims of this thesis are simple yet very practical. The dissertation explores questions which deal with the cost-performance tradeoff in implementation: What is the benefit of doing antenna selection with the precoding in massive MIMO at the base station? How can we do an effective and efficient antenna selection in a correlated environment? Do reduced dimensions of large matrices in both uncorrelated and correlated environments result in the reduction of computation complexity in massive MIMO? This dissertation will provide answers to the above questions, and it turns out the solution is closely related to enhance the throughput and spectral efficiency of the system.

Optimal antenna selection through a brute-force search algorithm involves high computations, and the computational load increases rapidly with the number of antennas. So the examination of sub-optimal algorithms for selecting the best antennas is thus of great theoretical as well as practical interest. A sub-optimal scheme limits

the number of radio-frequency chains, that is, the connection of the best subset of transmit antennas to the available radio-frequency chains.

PCA is a general method that reduces the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible the variation present in the data set. This is achieved by transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables. In particular algorithms based on central and non-central principal component analysis are developed in the thesis to select the best transmit antennas to maximise the sum-capacity. The proposed antenna-selection schemes are compared with the conventional antenna-selection techniques to evaluate the system performance. Simulation results show that the proposed antenna-selection algorithms reduced the correlation between signals received at users.

To model the signal correlation between the antennas, spatially correlated channel models, such as the Kronecker Model and the Weichselberger Model can be used to study system behaviour in different realistic environments. The Direction of Departure (DOD) and Direction of Arrival (DOA) of the received and transmitted signals can be modelled by using Power Azimuth Spectrum models. The outcome of this scheme is in terms of throughput analysis and hardware complexity reduction. The performance of the proposed antenna-selection algorithms in approximated real channel matrices is optimal compared with the exhaustive-search antenna-selection algorithms.

Dimensionality reduction of matrices using the aforementioned algorithms is examined by using Floating Point Operation (FLOP) methods. This technique use PCA analysis to calculate the number of operations involved in computations. In a nut-

shell, the large matrix size is reduced and a sum-capacity is analysed in both full-dimensional and optimal-dimensional system. Simulations show that a significant reduction in computational complexity can be achieved using the PCA technique in both un-correlated and correlated MIMO systems.

Consequently, this dissertation proposes a comprehensive study of the methodologies for lowering hardware complexity, reducing large matrices dimensions and spatially correlated channel models that affect fifth generation of MIMO broadcast systems. Current antenna-selection algorithms are investigated and, using PCA, two semi-heuristic antenna-selection algorithms are proposed. The eigenvalues are decomposed into two components in a PCA analysis which shows how antenna selection depends on the channel matrix structure. Numerical results suggest that the proposed scheme outperforms the sub-optimal antenna selection and approaches the performance of exhaustive search scheme but with a much lower complexity. The outcome of the dissertation supports the improvement of high-speed wireless communications and influences the development of the fifth generation of MIMO broadcast wireless communication systems.

1.8 Thesis Overview

Following the methodology depicted in Figure 1.3, we have presented most of our theoretical and practical findings in various international conferences, at different stages of my research tenure. The papers listed under the title of "Related Publications" show such presentations. This dissertation thus presents the systematised and combined version of all those papers. However, some important additions have made it more vigorous and comprehensive.

This dissertation is organised into six Chapters listed as follows:

- In **Chapter 2** the background and related work is discussed in the field of massive MU-MIMO, including, antenna selection algorithms, MIMO channel models, precoding in MIMO and complexity analysis of the wireless system. Previous antenna-selection algorithms used to reduce radio-frequency (RF) chains are studied. Spatial-correlation based channel models to study the system performance in different realistic environments are explored. Precoding techniques such as ZF, MMSE and DPC, followed by our proposed methodologies that will be exploited in later chapters, provide a technical context for this thesis.

- In **Chapter 3**, the focus is on reducing the number of RF chains resulting in reduced hardware system complexity. Under the assumption of the Rayleigh fading channel model and using the analysis of Principal Component Analysis (PCA), antenna-selection algorithms, such as Non-Central Principal Component Analysis (NCPCA) and Linear Dependence Avoidance System (LDAS), are proposed. These techniques remove the antennas that contribute least to the sum capacity.

- In **Chapter 4**, analytical channel models are explored to study the system behaviour in different realistic environments. Specifically, spatial-correlation models such as the Kronecker model and the Weichselberger model, are exploited. The signal correlation of these correlated channel models is approximated using the Power Azimuth Spectrum (PAS) method. Once the channel matrix is constructed under practical channel models, the proposed antenna-selection algorithms are then applied to select the best antennas.

- In **Chapter 5**, a complexity analysis of a wireless system in terms of Floating Point Operations (FLOP) is examined. In massive MIMO as the number of an-

tennas increases so does the matrix size. By using the analysis of PCA, the matrix dimension in massive MU-MIMO is reduced significantly. This lower-dimensional matrix reduces the computational complexity without sacrificing its effectiveness. In addition, spatial-correlation models are explored to analyse the computational complexity involved in linear precoding techniques.

- Finally, in **Chapter 6**, conclusions and discussions on future research work based on the results presented in the dissertation are presented.

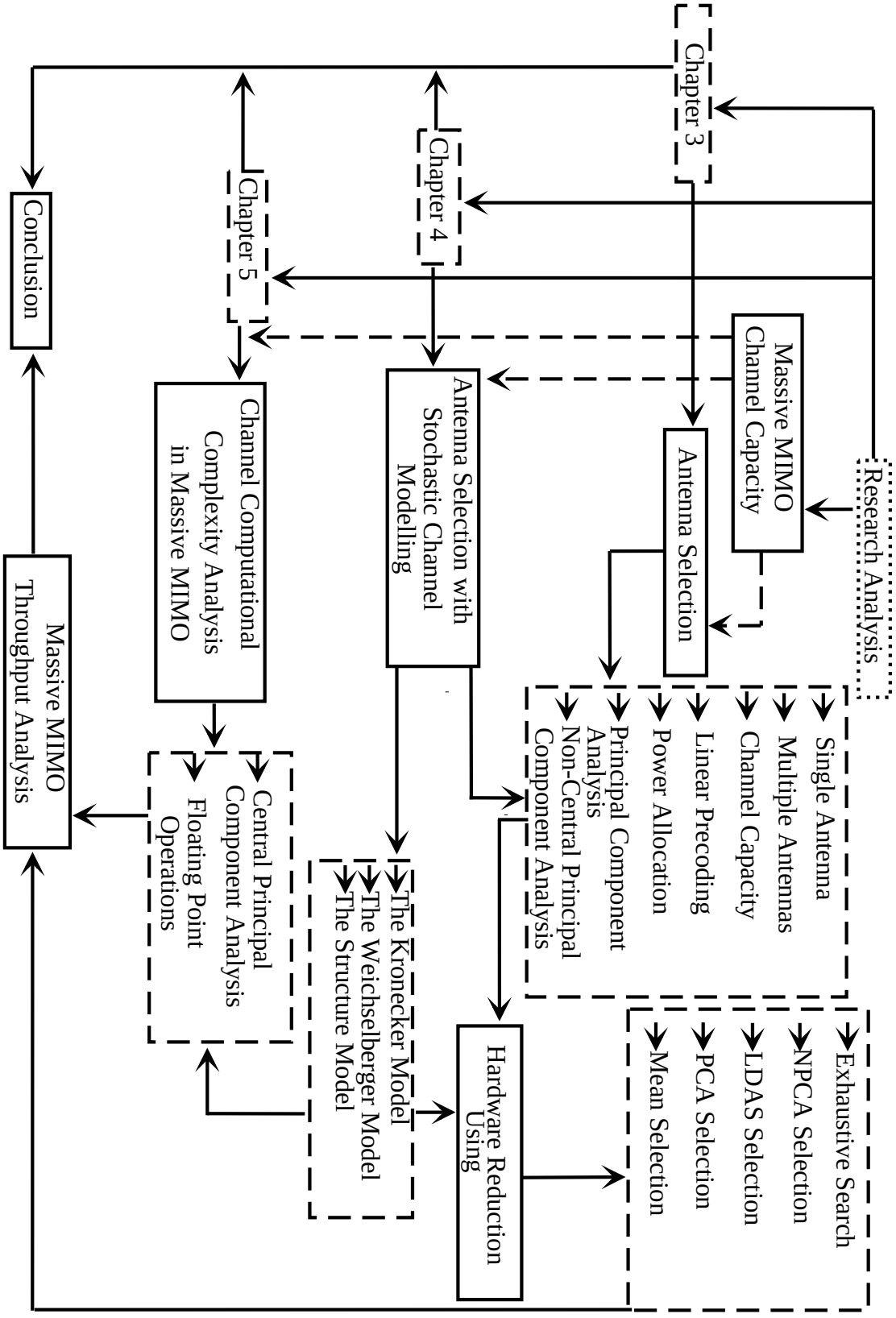


Figure 1.3: Illustration of the Proposed Dissertation Methodology

The difficulty of literature is not to write but to write what you mean;

not to affect your reader, but to affect him precisely as you wish.

Robert Louis Stevenson (1850 - 1894)

CHAPTER 2

BACKGROUND AND RELATED WORK

This chapter first gives a detailed overview of antenna selection techniques in Section 2.1. The two different categories of MIMO channel models are discussed in Section 2.2, and in Section 2.3 a detailed analysis of channel modelling is explained. Section 2.4 presents different types of precoding with more emphasis on the linear precoding techniques that are used in this thesis. The computational complexity in terms of floating point operations is explored in Section 2.5, and Section 2.6, presents the proposed technique of Principal Component Analysis that is used in this dissertation. Finally Section 2.7 concludes this chapter.

2.1 Antenna Selection in Massive MIMO

Massive MIMO is an emerging technology which is considered to be a potential candidate for fifth-generation (5G) wireless communications [34, 35]. The anticipated capacity requirements for 5G can be met by using the concept of massive MIMO, i.e. deploying a high number (say tens to hundreds) of base-station transmit antennas [17, 36]. Thus by employing massive MIMO a more than 10-fold throughput improvement has been suggested in massive MIMO as compared to Long Term Evolution (LTE) [21, 37]. However, the deployment of large numbers of antennas in massive MIMO at a base station (BS) brings hardware complexity and thus high cost challenges [38].

A massive MIMO system with N_t transmit and N_r receive antennas requires $N_t(N_r)$ complete RF chains at the transmitter and receiver respectively to achieve the anticipated throughput for 5G wireless communication. To overcome the aforementioned challenges, antenna selection in massive MIMO is proposed in the literature in which the best L out of N antenna signals are chosen and processed for transmission. Thus the reduced number of radio-frequency chains leads to a significant saving in system complexity. However, radio-frequency savings comes at the price of a small loss in system performance compared with the full complex system.

2.1.1 Conventional Antenna Selection in Massive MIMO Broadcast Channels

In massive MIMO, a large number of transmit antennas and RF chains are used, which results in increased spectral and transmit-energy efficiency. However, this presents challenges of large hardware energy consumption and high system complexity. As, in massive MIMO, all antennas do not contribute equally [39], this leads

us towards antenna selection algorithms in which the base station has to select an appropriate subset of transmit antennas with low complexity and high performance.

Since the advantages of diversity and multiplexing gain in MIMO systems include improved system throughput, the conventional antenna selection under the constraint of transmission power in a massive MIMO system focuses mainly on the optimisation of the signal-to-noise ratio (SNR) or capacity maximisation in massive MIMO. To grasp the advantages of a massive MIMO system, transmit antenna selection techniques with less-complex and low-cost methodologies are presented in the literature. The previous work on transmit-antenna selection for MIMO broadcast wireless channels is presented in [25, 40–46].

In [40], the authors present two antenna-selection schemes. In the first scheme, the base station selects the antenna with the maximum channel Frobenius norm. In the second scheme, assuming a statistical knowledge of the channels at the base station, antenna selection is based on maximising the determinant of the covariance of the channel vector. The signal transmission is done using orthogonal space-time block codes (OSTBC). The antenna selection that maximises the channel norm provides better results.

A tutorial paper on antenna selection in MIMO broadcast channels is presented in [25]. It is also shown in the companion paper [42] that antenna selection provides a considerable capacity gain in the system. In this scheme, the antenna with the highest signal-to-noise ratio (SNR) is selected in a single antenna system. For a multiple-antenna system, the authors used hybrid maximal ratio transmission, where a subset of the transmit antennas that provides a combined maximum channel gain is selected. The base station requires to know the antenna indices along with the channel gains. The authors also present a space-time code-based joint transmit/receive

antenna selection scheme that successively selects a pair of receive and transmit antennas that maximises the channel Frobenius norm.

In [43], the authors present a capacity-maximising antenna-selection scheme. In this scheme antenna-selection is based on the covariance matrix of the channel matrix. This antenna selection scheme works iteratively, selecting one antenna at a time.

Antenna selection for correlated MIMO fading channels is discussed in [44]. This work is extended in [45] and an antenna selection scheme based on the outage probability of a multi-user spatial multiplexing system is presented for correlated MIMO channels.

However, recently transmit-antenna selection techniques shown in [47–52], are based on the optimisation of the SNR or capacity maximisation under a transmitted power constraint.

In [47], the authors propose two fast antenna-selection algorithms for a wireless large-scale MIMO system. The algorithms start with a full set of selected antennas and then remove one antenna per step from this set until there is one user that does not meet the minimum sum-rate. The first algorithm aims to maximise the sum-capacity of a downlink MIMO with low computational complexity as compared with the optimal exhaustive-search scheme. In the second scheme, by maintaining a predefined full set of selected antennas, the second proposed algorithm then removes one antenna per step until reaching the minimum criterion, which leads to the lowest reduction of the sum rate and a significant reduction in the hardware cost of the system.

In [48], the authors use a convex optimisation technique to achieve the best compromise between the achievable capacity and system complexity. The interior-point algorithm is used to solve the constrained convex optimisation problem with relax-

ation. The interior-point algorithm finds the optimum number of transmit antennas and, in each step, the antenna which gives the lowest contribution to the capacity is removed. This leads to the reduction of the signalling overhead and the cost of the RF front end.

In [49], an anti-interference method for massive MIMO based on antenna selection is presented. The transmitter and receiver select a set of antennas under the assumption of real-time channel-state information. Once the antennas are selected according to the optimisation criterion, the interference from different spatial flows is aligned in one direction or in one subspace, leaving the spatial degrees of freedom for the most useful signals, and hence the system capacity is maximised, especially with high-SNR regimes.

In [50], three sub-optimal antenna-subset-selection schemes are proposed. These selection techniques are based on SNR maximisation or minimisation, of the trace of a matrix at the scheduled users. The signals are precoded using zero forcing at the base stations. The antenna that gives least to the matrix trace is carried out in the selection process. However, the highest channel-vector norm in the selection may not result in minimising the matrix trace. The authors extend the conventional single-cell multiuser MIMO into a multiple-point transmission known as network MIMO in the literature.

In [51], a joint transmit/receive antenna-selection algorithm is proposed in which a channel matrix is partitioned in such a way that the computational complexity of the exhaustive search algorithm is reduced significantly. The key idea in the proposed technique is calculating the capacity increase in each step instead of the whole subsystem capacity when selecting a new candidate pair of antennas in each iteration.

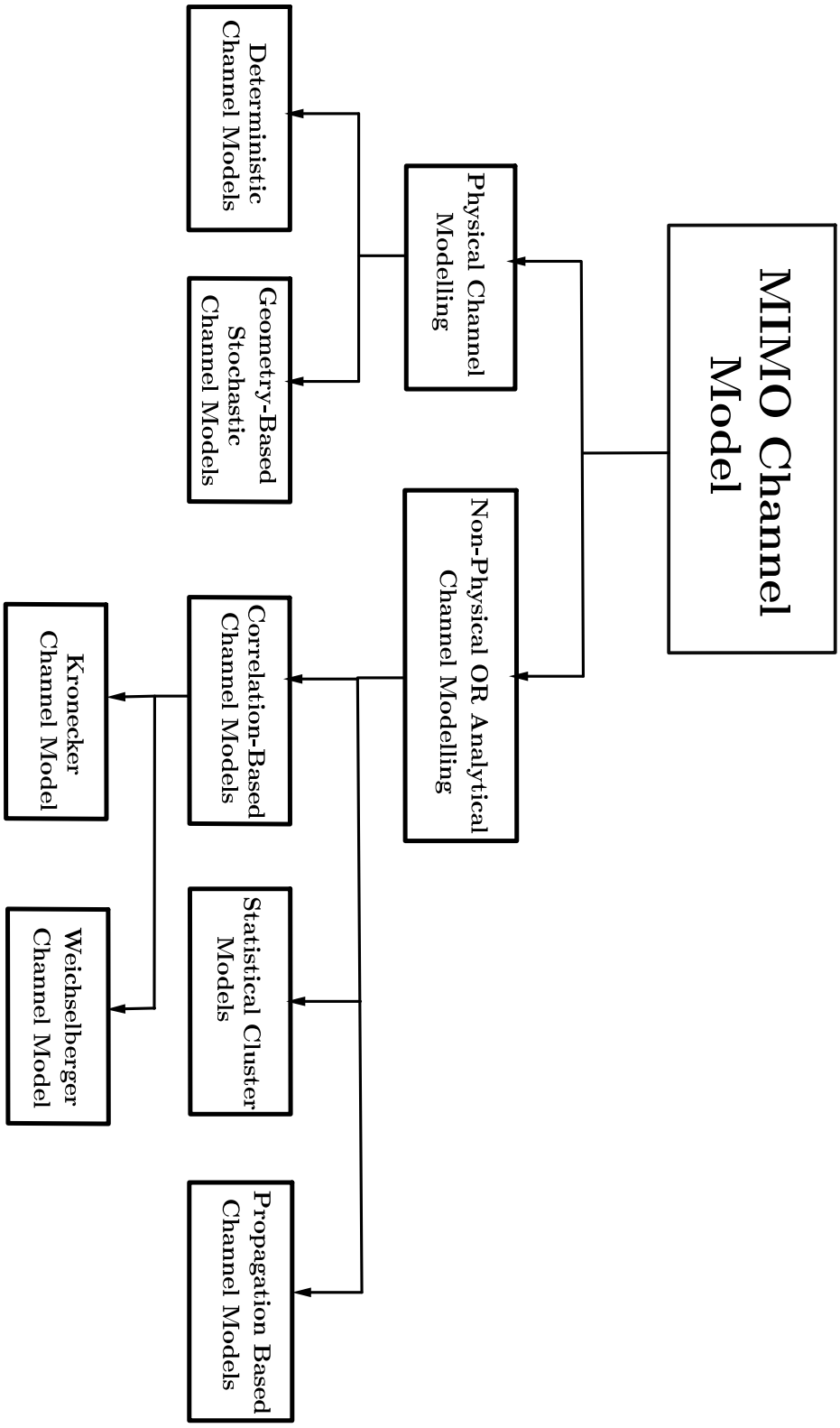


Figure 2.1: MIMO channel and propagation models

2.2 MIMO Channel Models

In wireless communication systems the design of channel models plays an important role to analyse the MIMO system. A detailed analysis of propagation models is shown in Figure 2.1 and correlation-based channel models under non-physical channel modelling are discussed in this section.

To evaluate the system performance, antenna arrays in MIMO systems are used to enhance system capacity, signal detection, and interference cancellation. A high information-theoretic capacity is shown in [5, 7], but a realistic evaluation of the capabilities of different MIMO schemes and system architectures requires realistic channel models [53]. In the development of wireless standards a lot of effort has been directed towards the measurement of MIMO channel models [54–57]. However, with the large number of antennas the channel measurements in massive MIMO are expected to yield accurate and realistic models in outdoor and indoor settings. Therefore, to better study the behaviour of realistic models in outdoor and indoor settings, it is important to understand the spatial properties of the channel matrix used in MIMO arrays [58]. Theoretically, in such channel models, the performance analysis is carried out using computer simulations and/or mathematical analysis. However, in practice a good channel model accurately captures a real channel behaviour, helping in the design, analysis and development of MIMO systems. Although it is not easy to classify the previous channel modelling approaches, especially in the earlier work of [59–61], Molisch in [53] categorises channel modelling approaches in MIMO systems into **Physical channel modelling** and **Non-physical channel modelling**.

2.2.1 Physical channel modelling

In physical channel modelling the properties of the physical environment, such as the location of the scatterers/reflectors or the direction of multipath components between the transmitter and receiver antenna arrays is modelled. This approach is used in many standard models such as COST 259 DCM [62], COST273 [63], IEEE 802.11n [64], HIPERLAN/2 [65], Stanford University Interim (SUI) [66] and IEEE 802.16 [67] and 3GPP [68]. The physical channel-modelling approaches for the location of scatterers/reflectors are known as *spatial channel models* or *scattering models*. The spatial models are important in evaluating the performance of a wireless communication system as the scatterers locations provide the angular and temporal information of the multipath signal in antenna arrays with respect to the MS and BS. A significant contribution in [69–75] is made in modelling spatial channels. The important parameters in modelling a spatial channel involve the scatter density and the shape of the scattering region to model the scattering phenomenon for macrocell, microcell and picocell cellular environments [69, 74, 76].

2.2.2 Non-physical channel modelling

In non-physical channel modelling the correlation of the fading of the signal at the antenna element is modelled [77]. In fact, space is an additional dimension that needs to be modelled on its own when dealing with MIMO channels, and physical channel modelling does not allow the explicit design of a space-time coding technique. In contrast with physical channels, non-physical/analytical channel modelling is used to describe the end-to-end transfer function between transmitting and receiving antenna arrays. The analytical channel models are useful in analysing the impact of correlation on system performance, and mathematically an MIMO channel matrix is

modelled as a function of a random-Gaussian-fading matrix. The traditional analytical MIMO channel models are expected to continue to model recent massive MIMO channel models.

2.3 Channel Modelling

A summary of current correlation-based analytical massive MIMO channel modelling in today's wireless standards is now presented. First a brief analysis of a system model is discussed which then leads to different correlation analytical channel models.

Channel Models

The structure of the received signal is affected by various parameters associated with a wireless channel. The frequency component of the signal is shifted relatively between the transmitter and the receiver due to the Doppler effect. However, with a large number of antenna arrays in massive MIMO, the users in a broadcast wireless channel are assumed to be in fixed positions. In a broadcast wireless system, a multi-path propagation model for the received signal is the tap-delay model [78]. The impulse response of the channel can be defined as

$$h(t) = \sum_{i=1}^{p} a_i \delta(t - \tau_i) \qquad \textit{Eq. (2.1)}$$

The transfer function of the channel $h(t)$ can be obtained by taking the Fourier transform, which is generally not constant over frequency and so it is referred to as a frequency selective channels in the literature. Now suppose that the bandwidth of the message signal is small, then in the frequency domain the channel is simply multiplied by a constant factor, commonly known as small-scale fading. Consider a

MIMO system consisting of multiple transmit and receive antenna elements, where each pair of transmit-receive antennas is associated with a channel gain. The received signal in a MIMO channel under the assumption of background noise can be modelled as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \qquad \textit{Eq. (2.2)}$$

where **y** denotes a received vector, the channel matrix **H** describes the channel gain between each pair of transmitter-receiver antenna elements, the transmitted signal vector is referred to as **x** and **n** is the noise vector at the receiver side, modelled as additive white Gaussian noise (AWGN). In a non-line-of-sight rich scattering environment, suppose that each element of **H** is independently and identically distributed (i.i.d.) with mean 0 and unit variance, denoted by $\mathcal{CN}(0,1)$. The maximum number of streams of N_R, N_T can be sent to achieve the anticipated system capacity under the assumption of the full rank of a channel matrix **H**. But, practically, poor channel gains of one or more antenna elements results in correlation [79].

2.3.1 Analytical Channel Models

The spatially independent identically distributed *i.i.d.* is the most common frequency non-selective flat-fading channel model which is used to observe system performance [80]. In this narrowband channel model, the antenna transmit-receive channel gain is modelled as a complex Gaussian random variable. This model assumes that the antenna elements between transmitter and receiver links are well separated, and that the channel gain between the links in the presence of a rich scattering environment can be approximated by a Gaussian random variable [81]. Let the line-of-sight

(LOS) component of the signal be the one that travels along the direct path between transmitter and receiver. This LOS signal component has non-zero real and imaginary mean values $h_{LOS,re}$ and $h_{LOS,im}$ respectively. The baseband channel model in small scale fading with the LOS component can be written as

$$h_{small} = (h_{re} + h_{LOS,re}) + j(h_{im} + h_{LOS,im}), \qquad \textit{Eq. (2.3)}$$

where h_{re} and h_{im} are the real and imaginary components of the channel model in small scale fading. The magnitude is therefore $|h_{small}| = \sqrt{(h_{re} + h_{LOS,re})^2 + (h_{im} + h_{LOS,im})^2}$, and the *Rician K-factor K* measures the dominance of the significant path. It is defined as the ratio of the power of the constant part of the signal to the average power of the random part of the signal, or

$$K = \frac{|h_{LOS,re}|^2 + |h_{LOS,im}|^2}{\boldsymbol{\sigma}^2},$$

where $\boldsymbol{\sigma}^2$ is the variance for the Rician distribution and the probability density function in this case is given by

$$P_r\{|h_{small}|\} = \left(\frac{|h_{small}|}{\boldsymbol{\sigma}^2/2}\right) e^{-\frac{|h_{small}|^2}{\sigma^2}} e^{-K} I_\circ\left(\frac{|h_{small}|\sqrt{2K}}{\boldsymbol{\sigma}/2}\right), \qquad \textit{Eq. (2.4)}$$

$I_\circ(.)$ is the modified $zero^{th}$ order Bessel function of the first kind. For a scattering channel matrix $\mathbf{H}$, with dimensions $N_r \times N_t$, a stochastic part H_s and a deterministic part H_d with rice factor K, is defined as [82]

$$\mathbf{H} = \sqrt{\frac{1}{1+K}}\mathbf{H}_s + \sqrt{\frac{K}{K+1}}\mathbf{H}_d. \qquad \textit{Eq. (2.5)}$$

To obtain a Rayleigh distribution, that is in a pure multipath environment without an LOS component and with zero mean gain put $K = 0$, and $K = \infty$ corresponds to unfaded AWGN channel.

2.3.2 Spatial-Correlation Based Models

In spatial correlation based analytical channel modelling, the MIMO channel matrix is statistically characterised in terms of the correlation between the entries of the channel matrix. The MIMO capacity and system performance are affected by the spatial correlation between the transmit and receive antenna elements, as shown in Figure 2.2, where capacity is plotted both for i.i.d. and correlated MIMO scenario. Therefore to model a MIMO channel in such a pure multipath environment without any LOS component, consider $K = 0$ in (2.5) giving $\mathbf{H} = \mathbf{H}_s$. Suppose that $\mathbf{h} = vec(\mathbf{H})$ and that $\mathbf{R}_H \triangleq \mathbb{E}[\mathbf{h}\mathbf{h}^H]$ is a correlation matrix of dimensions $N_t N_r \times N_t N_r$. The complex zero mean multivariate Gaussian distribution of $\mathbf{h}$ can be written as

$$f(\mathbf{h}) = \frac{1}{\pi^{N_t N_r} \det\{\mathbf{R}_H\}} e^{(-\mathbf{h}\mathbf{R}_H^{-1}\mathbf{h})}. \qquad \textit{Eq. (2.6)}$$

The elements in $\mathbf{R}_H$ contain the correlation of the channel matrix that describes the spatial statistics. The distribution in (2.6) gives the realisation of vector $\mathbf{h}$, and hence the realisation of the channel matrix $\mathbf{H}$ can be written as

$$\mathbf{h} = \mathbf{R}_H^{1/2}\mathbf{g}, \qquad \textit{Eq. (2.7)}$$

$\mathbf{g}$ is an $N_t N_r \times 1$ zero mean, unity-variance i.i.d. Gaussian random vector and $\mathbf{R}_H^{1/2}$ is any matrix that satisfies $\mathbf{R}_H^{1/2}\mathbf{R}_H^{H1/2} = \mathbf{R}_H$. For large N_t, N_r (massive MIMO), the number of real-valued parameters required to fully specify $\mathbf{R}_H$, is $N_t^2 N_r^2$. So,

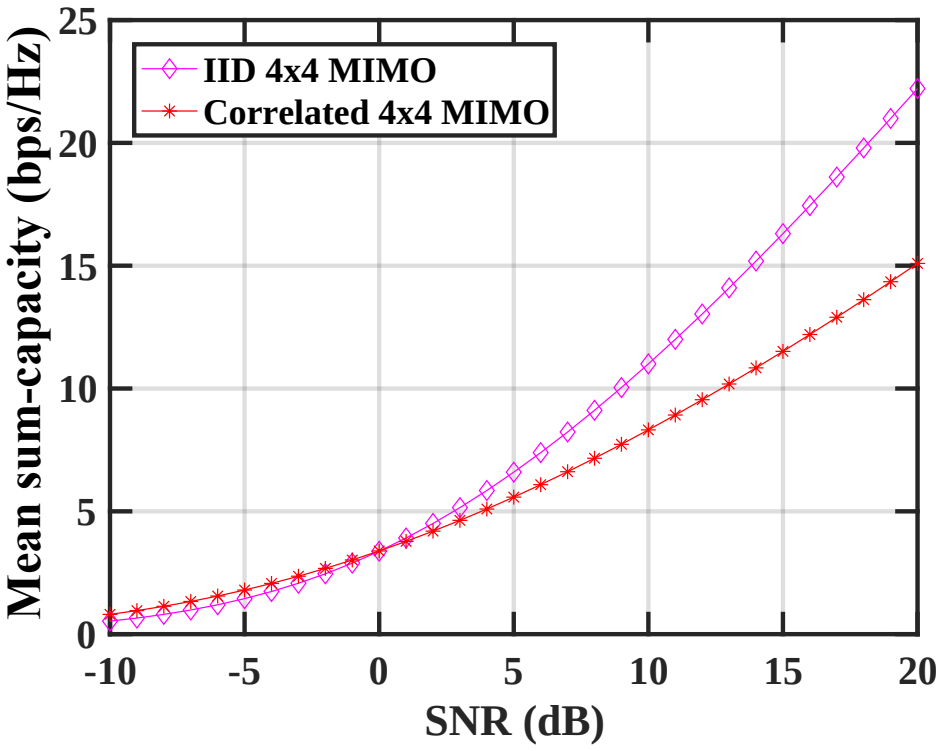


Figure 2.2: Correlated MIMO Channel Capacity vs i.i.d. MIMO Channel

this requirement can be reduced by imposing a certain structure on the correlation matrix. There are different methods to construct a structure of a channel matrix given in Equation (2.2) that leads to the construction of different correlation based channel models, which are now presented.

2.3.2.1 The Independent Identically Distributed Rayleigh Fading Model

The *i.i.d.* Rayleigh fading model is a very commonly used MIMO analytical channel model. In this model all the elements of the channel matrix **H** are independent, and identically distributed uniformly in all directions in a rich scattering environment that corresponds to a spatially white MIMO channel. The elements in the channel matrix **H** are uncorrelated and have an equal variance of ρ^2, i.e.

$$\mathbf{R}_H = \rho^2 \mathbf{I}_{N_t N_r} \qquad \textit{Eq. (2.8)}$$

This model is used in MIMO performance evaluation, information-theoretic analysis and simulating MIMO algorithms [78] due to the simplicity that attracts researchers to use it for analysis in massive MIMO studies.

2.3.2.2 The Kronecker Model

The existence of channel correlation is due to an inter-element structure and antenna array structure which degrades the throughput performance of the wireless system. To mitigate channel correlation, a simple uniform linear array structure using the Kronecker model is proposed that incorporates spatial correlation between antenna elements. This model assumes that the distance between the transmitter and receiver is sufficient such that the correlation on the transmitter side has no effect on the receiver side and vice versa. It is also assumed that the transmitter and receiver spatial-correlation matrices are separable. Let $R_{Tx} = \mathbb{E}[\mathbf{H}^H\mathbf{H}]$ and $R_{Rx} = \mathbb{E}[\mathbf{H}\mathbf{H}^H]$ denote the transmit correlation matrix and the receive correlation matrix, respectively. The Kronecker channel model in product form under these assumptions can be written as

$$\mathbf{R}_H = \mathbf{R}_{Tx} \otimes \mathbf{R}_{Rx} \qquad \textit{Eq. (2.9)}$$

The vector $\mathbf{h}$ in the Kronecker model becomes

$$\mathbf{h} = (\mathbf{R}_{Tx} \otimes \mathbf{R}_{Rx})^{1/2}\mathbf{g}, \qquad \textit{Eq. (2.10)}$$

and

$$\mathbf{H} = \mathbf{R}_{Tx}^{(1/2)}\mathbf{G}\mathbf{R}_{Rx}^{(1/2)^T} \qquad \textit{Eq. (2.11)}$$

where, as before, $\mathbf{g}$ is an $N_t N_r \times 1$ vector with i.i.d. Gaussian entries with zero mean and unit-variance, and $\mathbf{G}$ is an i.i.d. unity variance matrix obtained by performing an inverse vec(.) operation on $\mathbf{g}$. The number of parameters that characterise this model is N_t^2 (parameters in $\mathbf{R}_{Tx}$) plus N_r^2 (parameters in $\mathbf{R}_{Rx}$), unlike the $N_T^2 N_R^2$ parameters in the full correlation matrix. The Kronecker model has also been shown to underestimate the throughput in adaptive modulation in MIMO compared to the throughput obtained using measured channel matrices [83]. Despite the limitation of ignoring the coupling between the *Direction of departure* (DOD) and *Direction of arrival* (DOA) at the transmit and receive ends, the Kronecker model has been popularly used in information-theoretic capacity analysis and simulation studies in [84–88]. The model in (2.11) is called the Kronecker channel model and is popular for its simplicity.

2.3.2.3 The Weichselberger model

A simple correlation-based stochastic channel model is the Kronecker model, that can be implemented with a few driving parameters, i.e. the signal correlation at each end of a link. An extension of the Kronecker model is the Weichselberger model. This model does not assume the separability of the correlation matrices [55, 89]. A special structure of channel is used to determine the (DOD) and (DOA) at the transmitter and receiver ends respectively. These special correlation matrices are obtained by using eigenvalue decomposition of the transmitter and receiver matrices as follows:

$$\mathbf{R}_{Tx} = \mathbf{U}_{Tx} \mathbf{\Lambda}_{Tx} \mathbf{U}_{Tx}^H \qquad \textit{Eq. (2.12)}$$

$$\mathbf{R}_{Rx} = \mathbf{U}_{Rx} \mathbf{\Lambda}_{Rx} \mathbf{U}_{Rx}^H, \qquad \textit{Eq. (2.13)}$$

where $\mathbf{\Lambda}_{Tx}$ and $\mathbf{\Lambda}_{Rx}$ are the eigenvalue matrices with corresponding orthonormal eigenvectors $\mathbf{U}_{Tx}$ and $\mathbf{U}_{Rx}$ at transmitter and receiver respectively. The matrices $\mathbf{R}_{Tx}^{1/2}$ and $\mathbf{R}_{Rx}^{(1/2)^T}$ are not unique and can be chosen as:

$$\mathbf{R}_{Tx}^{(1/2)^T} = \sqrt{\mathbf{\Lambda}_{Tx}}\mathbf{U}_{Tx}^H \qquad \textit{Eq. (2.14)}$$

$$\mathbf{R}_{Rx}^{1/2} = \mathbf{U}_{Rx}\sqrt{\mathbf{\Lambda}_{Rx}}, \qquad \textit{Eq. (2.15)}$$

where

$$\sqrt{\mathbf{\Lambda}_{Tx}} = \mathrm{diag}(\lambda_{Tx,1}, \lambda_{Tx,2}, \ldots, \lambda_{Tx,n}) \qquad \textit{Eq. (2.16)}$$

$$\sqrt{\mathbf{\Lambda}_{Rx}} = \mathrm{diag}(\lambda_{Rx,1}, \lambda_{Rx,2}, \ldots, \lambda_{Rx,n}). \qquad \textit{Eq. (2.17)}$$

By substitution of these into (2.11) we get

$$H_{\text{Weich}} = \mathbf{U}_{Rx}\sqrt{\mathbf{\Lambda}_{Rx}}\mathbf{G}\sqrt{\mathbf{\Lambda}_{Tx}}\mathbf{U}_{Tx}^H \qquad \textit{Eq. (2.18)}$$

The coupling between DOD-DOA in a Weichselberger channel is taken into account through a coupling matrix $\mathbf{\Omega}$, defined as

$$\mathbf{\Omega}_{\text{w}} = \begin{pmatrix} \lambda_{Rx,1} \\ \lambda_{Rx,2} \\ \vdots \\ \lambda_{Rx,Nnx} \end{pmatrix} \begin{pmatrix} \lambda_{Tx,1}, \lambda_{Tx,2}, \ldots, \lambda_{Tx,n} \end{pmatrix} \qquad \textit{Eq. (2.19)}$$

Note that the Equation (2.19) can be written in terms of the Weichselberger model as

$$\mathbf{H} = \mathbf{U}_{Rx}\mathbf{\Omega}_w \odot \mathbf{G}\mathbf{U}_{Tx}^T, \qquad \textit{Eq. (2.20)}$$

where $\odot$ represents an element-wise multiplication of the matrices. The elements of $\mathbf{\Omega}$ are real-valued and non-negative, and they determine the average power coupling between the transmit and receive eigenmodes. It can be observed that the Weichselberger model in (2.20) becomes the Kronecker model when the coupling matrix is a rank-1 matrix given by [55,89]. To design signal processing algorithms to achieve the anticipated capacity, diversity and beamforming the structure of $\mathbf{\Omega}$ can be exploited. For instance, the performance metrics such as the diversity order and the capacity of the MIMO channel in the Weichselberger model only depend on the coupling matrix $\mathbf{\Omega}$, and are independent of the transmit and receive eigenmodes.

2.4 Precoding in MIMO

In the literature, the term precoding or pre-equalisation can be implemented on the transmitter side to represent any transmit pre-processing apart from channel coding. The precoding techniques (see Section 2.4.2) are used in MIMO wireless systems to improve the system performance as shown in Figure 2.3. By using Monte-Carlo method, mean sum-capacity is plotted against SNR with 4×4 transmit and receive antennas. However, the channel knowledge is important in achieving the best capacity performance. The channel knowledge known at the transmitter is referred to as the channel state information at the transmitter (CSIT). It can be obtained by using a feedback network. It is not easy to get up-to-date feedback all the time, especially in rapidly varying channels [80]. In addition, a large bandwidth is required in sending channel information via feedback in large MIMO systems, therefore partial channel-

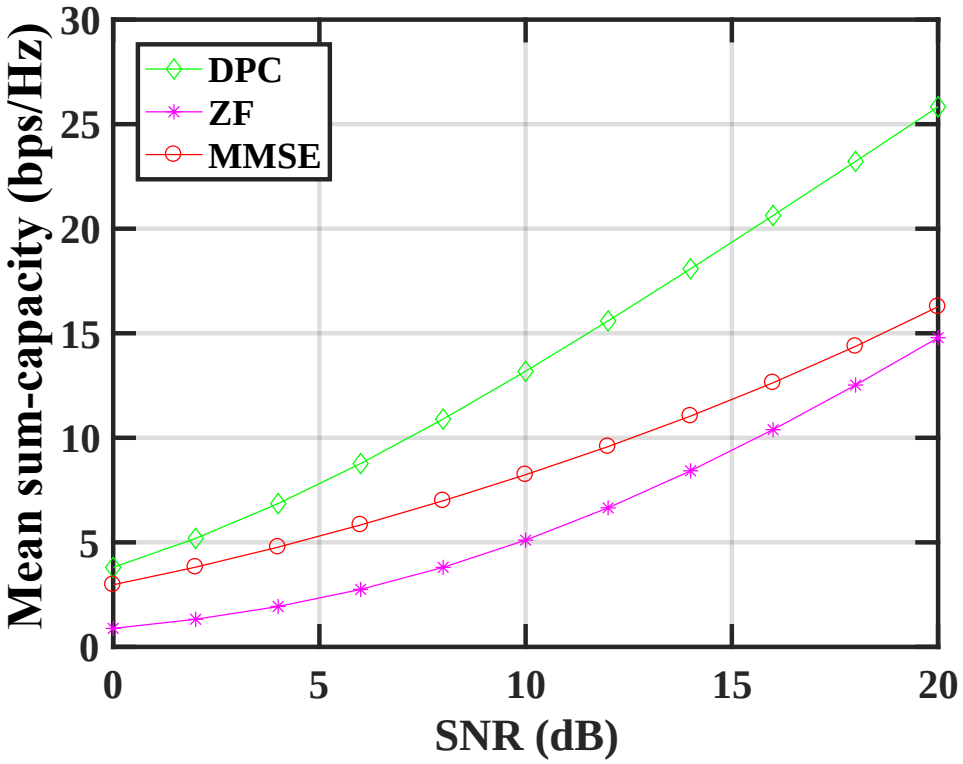


Figure 2.3: Sum-Capacity of Non-Linear (DPC) and Linear (ZF & MMSe)

state information is used instead [90]. In multiuser MIMO, the capacity gain from CSIT is different at low and high signal-to-noise ratios (SNR) [91]. At low SNR, the ergodic capacity increases significantly with the help of CSIT. In fact, the power is transmitted to only strong channel modes as the transmitter uses channel knowledge, whereas the transmitter distributes equal power in every direction if CSIT is not available. At high SNR, the anticipated capacity is dependent on the antenna configuration. For example, CSIT helps to achieve an incremental capacity with more transmit than receive antennas [92]. In general, the throughput of the channel can be maximised using precoding designs that are studied for various scenarios. Perfect CSIT is explored in [93], mean CSIT is explained in [94], transmit covariance CSIT is studied in [95]. So, based on the transmit-side pre-processing, the precoding techniques fall into *non-linear precoding* and *linear precoding*.

2.4.1 Non-Linear Precoding

The optimal scheme based on *writing on dirty paper* given by Coast is known as Dirty Paper Coding (DPC), categorised as a non-linear precoding technique [15]. The authors in [13,14] show that a MIMO broadcast channel achieves a maximum system capacity by using a DPC precoding technique. However, the associated complexity makes it impractical in real systems [96]. To overcome the shortcomings of DPC, a sub-optimal non-linear precoding with reduced achievable rates is proposed. This technique is known as *Tomlinson-Harashima precoding (THP)*, that successively pre-subtracts the known interference at the transmitter [97].

2.4.2 Linear Precoding

This scheme involves a linear transformation of the data in a precoding matrix. The benefit of linear precoding is its low complexity and its simple implementation in practical systems [98,99], however, we make a sacrifice in the achievable throughput as shown in Figure 2.3. The interference is pre-cancelled in linear precoding due to the knowledge of CSI at the transmitter. Popular linear precoding techniques are the zero forcing (ZF) and minimum mean square error (MMSE) precoding techniques.
In ZF precoding, a constraint is set that forces all the interference from other users to zero with the help of channel state information available at the transmitter. Consider a transformation $\mathbf{x} = \mathbf{W}\mathbf{u}$, where, an information symbol vector $\mathbf{u}$ is encoded into a transmit vector $\mathbf{x}$ and a precoding matrix $\mathbf{W}$. This precoding matrix is designed to achieve zero inter-symbol interference (ISI) from other users. Using (2.2) the received vector is $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, and by putting in the transformation of $\mathbf{x}$ we get

$$\mathbf{y} = \mathbf{H}\mathbf{W}\mathbf{u} + \mathbf{n} \qquad \textit{Eq. (2.21)}$$

A precoding matrix $\mathbf{W}$ in a case of conjugate beamforming also known as maximum-ratio transmission (MRT) can be defined as $\mathbf{W} = \mathbf{H}*$. However, To get interference-free communication, a precoding matrix

$$\mathbf{W} = \beta_{ZF}\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1} \qquad \textit{Eq. (2.22)}$$

is chosen [100], where β_{ZF} is a normalisation factor introduced in order to meet the transmitted power constraint after precoding and is defined as [101]

$$\beta = \sqrt{\frac{N_t}{tr(\mathbf{H}\mathbf{H}^H)^{-1}}} \qquad \textit{Eq. (2.23)}$$

The total power constraint in this suboptimal precoding technique is

$$\mathbb{E}\{||\mathbf{x}||^2\} = tr\{\mathbf{W}\mathbf{W}^H\} = ||\mathbf{W}||_F^2 \leq P, \qquad \textit{Eq. (2.24)}$$

where, P is the total transmitted power. In addition, the performance of the ZF precoder is poor when β_{ZF} in (2.23) become smaller in an ill-conditioned matrix. In fact, in such a matrix, the complete nulling of other users creates a power penalty, i.e. enhanced noise. To alleviate this issue, a well-known MMSE constraint and its solution is given by

$$\mathbf{W} = \beta_{MMSE}\mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \alpha\mathbf{I}_k)^{-1}, \qquad \textit{Eq. (2.25)}$$

where $\alpha > 0$ determines the amount of interference; $\alpha = 0$ gives (2.22). So no matter how poorly conditioned channel matrix $\mathbf{H}$ is, by choosing α large enough, the inverse in (2.25) can behave as well as desired. The normalized factor β_{MMSE} is

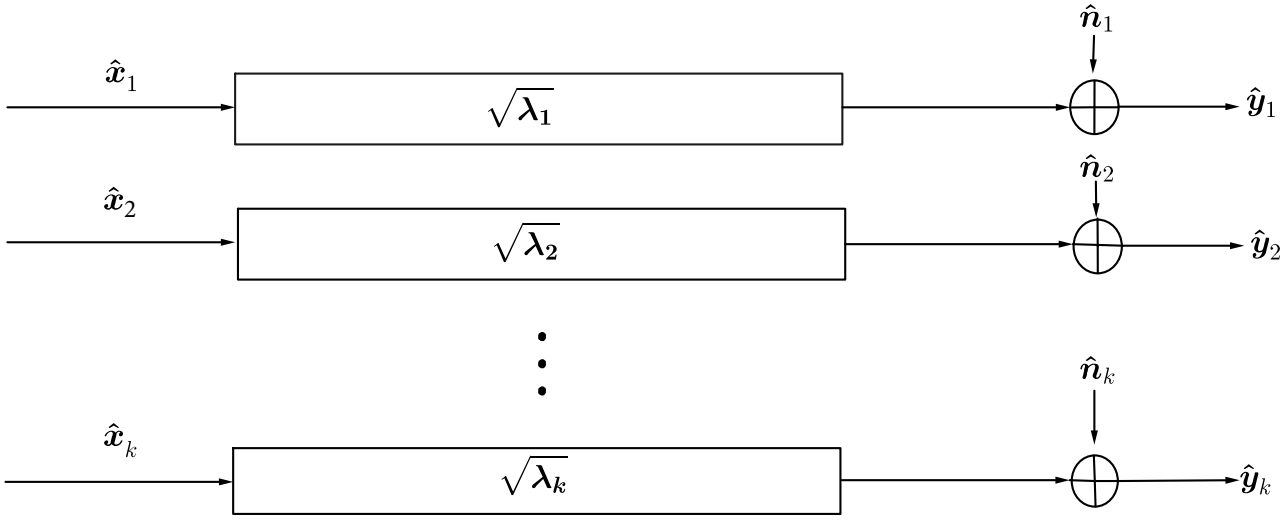


Figure 2.4: MIMO Channel Split into Virtual Scalar and Independent Channels

defined as [102]

$$\beta_{MMSE} = \sqrt{\frac{N_t}{tr(\mathbf{H}\mathbf{H}^H + \alpha \mathbf{I}_k)^{-1}}}, \qquad \textit{Eq. (2.26)}$$

2.4.3 MIMO Channel Decomposition

The MIMO channel capacity can be computed by splitting the channel into parallel and independent scalar sub-channels in the presence of channel knowledge at transmitter and receiver [103]. So, the throughput of an MIMO channel with available CSIT can be computed. The capacity in this case is equivalent to the sum-rate of independent parallel scalar sub-channels, with equal power allocation to each sub-channel as shown in Figure 2.4 [32, 103].

A linear transformation, such as singular-value decomposition (SVD), is used to decompose a channel matrix into sub-channels. Consider a channel matrix $\mathbf{H}_{N_r \times N_t}$, with rank r and a SVD decomposition $\mathbf{H} = \mathbf{U}\Lambda\mathbf{V}^H$, shown in Figure 2.5. The figure shows a pre-processed transmitted signal with a $\mathbf{U}$ matrix, and a receive signal is

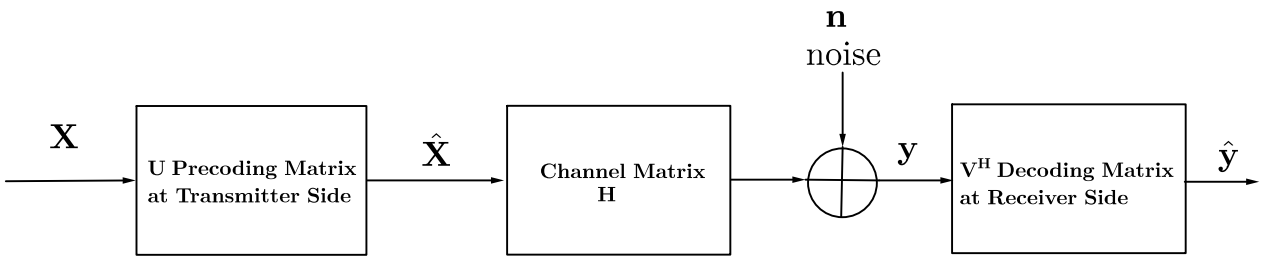


Figure 2.5: SVD of MIMO Channel

processed at the receiver side with the $\mathbf{V}^H$ matrix. The dimensions of these unitary matrices are $\mathbf{U}_{N_r \times N_r}$ and $\mathbf{V}_{N_t \times N_t}$ respectively, and $\Lambda_{N_r \times N_t}$ is a diagonal matrix, whose diagonal elements are the singular values of matrix $\mathbf{H}$ and can be written as

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_r \end{bmatrix}. \qquad \textit{Eq. (2.27)}$$

The MIMO channel capacity after decomposition of the channel matrix can be written as

$$\mathbf{C} = \log_2\{\det(\mathbf{I}_{N_r} + \frac{1}{N_0}\mathbf{U}\Lambda\mathbf{V}^H\mathbf{R}_{xx}\mathbf{V}\Lambda^H\mathbf{V}^H)\}. \qquad \textit{Eq. (2.28)}$$

Using the identity in [104], $\det(\mathbf{I}+\mathbf{AB}) = \det(\mathbf{I}+\mathbf{BA})$ in (2.28), the capacity equation can be re-written as

$$\mathbf{C} = \log_2\{\det(\mathbf{I}_{N_r} + \frac{1}{N_0}\mathbf{R}_{xx}\Lambda)\}, \qquad \textit{Eq. (2.29)}$$

where N_0 is the average noise power.

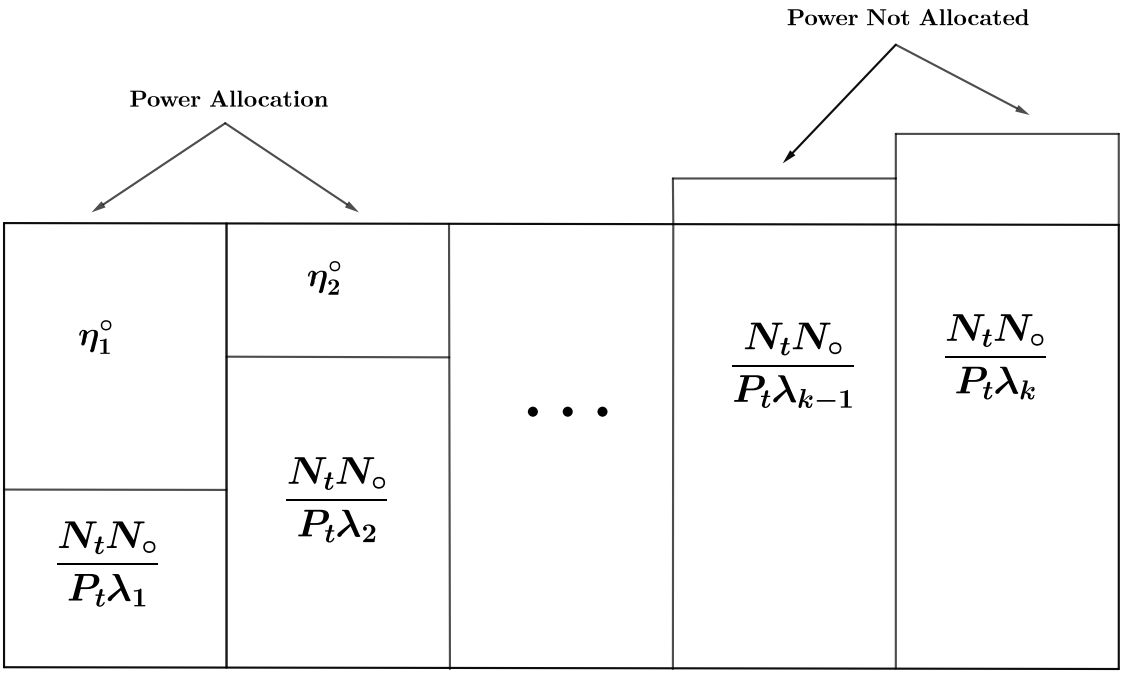


Figure 2.6: Optimal Power Allocation Using Water-filling Principle

Since the total transmitter power P_T is $P_T = tr\{\mathbf{R}_{xx}\} = \sum_{i=1}^{N_t} E|\mathbf{x}_i|^2$ and if $\eta_i = \frac{P_t}{N_t} E|\mathbf{x}_i|^2$ is the i^{th} transmit antenna power where $(i = 1, \cdots, N_t)$, then the MIMO channel is converted into a virtual SISO channel as shown in Figure 2.4, and mathematically the capacity can be expressed as

$$\mathbf{C}(\eta) = \sum_{i=1}^{r} \log_2 \left(1 + \frac{P_t}{N_t N_0} \eta_i \lambda_i\right). \qquad \textit{Eq. (2.30)}$$

2.4.4 Power Allocation in MIMO System

To enhance the system throughput an optimum power can be assigned in the presence of CSIT to each transmit antenna [12, 78]. The most popular technique is the water-filling scheme that allocates power among different sub-channels by exploiting channel state information at the transmitter. A schematic diagram of the water-filling

algorithm is shown in Figure 2.6. The figure clarifies the power allocation according to the channel strength i.e. the power is not distributed to the channels with small eigenvalues, but between the rest of the channels. In terms of the water-filling algorithm, the throughput of the system can mathematically expressed as [12]

$$\mathbf{C} = \max_{\eta} \sum_{i=1}^{r} \log_2 \left\{ 1 + \frac{SNR}{N_t} \eta_i \lambda_i \right\}, \qquad \textit{Eq. (2.31)}$$

where $SNR = \frac{P_t}{N_t}$ and the constraint over η is $\sum_{i=1}^{r} \eta_i = P_T$, then the water-filling power optimisation problem can be defined as

$$\eta_i^o \;=\; \left\{ \mu - \frac{N_t}{SNR\lambda_i} \right\}^+,$$

where $i = 1, \cdots, r$, $\sum_{i=1}^{r} \eta_i{}^o = P_T$, μ is a constant and $(x)^+$ is defined as $\max(x, 0)$. Using the water-filling technique, the power-optimisation MIMO capacity equation with average transmit power $\eta^o(i)$ can be written as

$$\mathbf{C} = \sum_{i=1}^{r} \log_2 \left\{ 1 + \frac{SNR}{N_t} \eta_i^o \lambda_i \right\}, \qquad \textit{Eq. (2.32)}$$

Let us consider a case when channel knowledge is not available at the transmitter, then power is distributed among all the transmitter antennas regardless of channel strength. This method is known as equal power allocation, and the capacity formula in this case is

$$\mathbf{C} = \log_2 \left\{ \det \left(\mathbf{I}_{N_r} + \frac{SNR}{N_t} \mathbf{H}\mathbf{H}^H \right) \right\}. \qquad \textit{Eq. (2.33)}$$

In this thesis, we pursue the antenna-selection issue with both the equal power allocation and the water filling techniques in Chapter 3 and 4.

2.5 Computational Complexity

There are multiple signalling schemes in MIMO wireless communication systems to enhance link quality, range, stability and system throughput. In massive MIMO a base station is equipped with a very large number of antennas and, when doing linear precoding at the base station, there is a need to compute the pseudo-inverse of large matrices. So, matrix inversion in massive MIMO multiplies the computational complexity of a system manyfold [105]. The computational complexity here is in terms of floating-point operations (Flops), and a possible solution is to reduce the matrix dimension which results in reduced computational complexity.

There have been numerous complexity-reduction techniques proposed in recent years [106–109]. In [106], a quasi-Newton algorithm is applied to reduce the complexity in downlink massive MIMO systems. The authors claim that the proposed algorithm exhibits faster convergence than conventional methods. In [107], a precoding technique based on a truncated polynomial expansion is used to mitigate the complexity of massive MIMO systems. In [108], the Neumann series expansion is chosen for inversion of matrices over the traditional exact computation in linear ZF precoding. The results in [108] show that, with the help of the Neumann series expansion, a significant reduction in computational complexity can be achieved. In [109], a low-complexity method to compute the sum rate using linear MMSE is discussed, where instead of using the conventional technique of matrix inversion in linear MMSE precoding, a matrix polynomial method is proposed in which only a

few terms are sufficient to closely approach the sum rate when compared with a classical MMSE precoder.

This dissertation gives analysis of computational complexity reduction in Chapter 5.

2.6 Principal Component Analysis

Principal component analysis is a statistical technique developed by Karl Pearson in the early nineteenth century. This method is a very useful data analysis tool in today's research [110] . For instance, a high-dimensional data matrix can be transformed into a lower-dimensional matrix which results in numerous advantages. Firstly, the reduced dimension keep most of the useful information and reduces noise and other undesirable artefacts. Secondly, the time and memory used in data processing are smaller. Thirdly, it provides a way to understand and visualise the structure of complex data sets. The first objective of PCA used in this thesis is *feature selection*, discussed in Chapter 3 and Chapter 4. The second purpose of using PCA is *reducing high-dimensional data into lower-dimensional data*, and is presented in Chapter 5. However, other distinguishing characteristics of PCA are discussed in [110].

Let us assume a vector $\mathbf{x}$ containing p random variables. The point of interest is the variance and the structure of the correlation or covariance in p random variables. In large-system analysis, it is not easy to look at the p variances and all of the $\frac{1}{2}p(p-1)$ correlations or covariances. However, an alternative choice is to look for a $(\ll p)$ derived variables. These derived variables retain almost all of the information of the variance and correlation or covariance structure. The derived variables can be obtained by looking at a linear function $\boldsymbol{\alpha}'\mathbf{x}$ of the elements of $\mathbf{x}$ having maximum

variance. Mathematically, it can be expressed as

$$\boldsymbol{\alpha}_1'\mathbf{x} = \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1p}x_p = \sum_{j=1}^{p} \alpha_{1j}x_j, \qquad \textit{Eq. (2.34)}$$

where $\boldsymbol{\alpha}_1'$ is a vector of p constants $\alpha_{11}, \alpha_{12}, \ldots, \alpha_p$ and $'$ denotes transpose. Similarly, the linear combination $\boldsymbol{\alpha}_2'\mathbf{x}$ can be written with a second maximum variance, however, the second linear function should be uncorrelated with $\boldsymbol{\alpha}_1'\mathbf{x}$. Therefore, the k^{th} linear function $\boldsymbol{\alpha}_k'\mathbf{x}$ has a maximum variance subject to being uncorrelated with $\boldsymbol{\alpha}_1'\mathbf{x}, \boldsymbol{\alpha}_2'\mathbf{x}, \ldots, \boldsymbol{\alpha}_{k-1}'\mathbf{x}$. Theoretically, the computations of the principal components reduces to the solution of an eigenvalue-eigenvector problem for a positive semi-definite symmetric matrix [33] and only the first k principal components account for most of the variability of the original data.

The detailed analysis of principal component analysis is given in Chapter 3.

2.7 Summary of the Chapter

In this chapter, we have addressed the issue of hardware complexity in 5G wireless communication. We have studied intensively previous approaches used in antenna selection for wireless communication. We will confront the problem of antenna-selection with our proposed algorithm in the following chapter. We have developed the necessary channel models that can be applied to the antenna selection research problem with our proposed methodology. We will also address the computational-complexity issue that appears in inverting large matrices in massive MIMO when doing linear precoding with our proposed scheme.

I cannot teach anybody anything. I can only make them think.

Socrates (469 - 399 BC)

CHAPTER 3

ANTENNA SELECTION IN MASSIVE MIMO COMMUNICATION

In this chapter, Section 3.1 gives a brief description of an antenna selection problem. The related work is discussed in Section 3.2. We describe the system model used in this chapter in Section 3.3. The theory of PCA is explored in Section 3.4. In Section 3.5, the proposed antenna-selection techniques are presented followed by the simulation results, and a discussion of the results is given in Section 3.6. In the end, Section 3.7 concludes the chapter and presents final remarks.

3.1 Introduction

This chapter examines the antenna selection algorithms for massive multi-user MIMO (MU-MIMO) broadcast wireless channels. We consider that the base station is equipped with a large number of transmit antennas transmitting signals to multiple users with each user equipped with a single receive antenna. Thus, by using the multiple antennas the base station transmits independent data streams to serve multiple users simultaneously. However, in multi-user communication, signals for different users in the same frequency band create interference to other users, which results in reduced system capacity. Therefore, in downlink MU-MIMO broadcast channels the transmitted signal is pre-coded before transmission to mitigate the interference from other users. In this case, the number of transmit antennas (N_t) needs to be at least equal to the number of receive antennas over all users for broadcast communication to occur. For receivers, with a single antenna, the total number of receive antennas equals the number of selected users (K). If $N_t > K$ then there are unused degrees of freedom that can be used to improve the sum capacity of the system.

In order to achieve high capacity, one method is to use antenna selection techniques that select a subset of antennas for active transmission. For a signal to be transmitted from an antenna element at the BS, the element needs to be connected to a radio-frequency (RF) chain which comprises Analogue-to-digital/Digital-to-Analogue converters, a power amplifier and mixers etc. The total number of RF chains in this thesis is assumed to be N_s, equal to the number of antenna elements used for active transmission. In [111], it has been shown that the main power consumption and hardware cost of a cellular network comes from the radio-access network. Therefore, using a reduced number of expensive RF chains at the BS can significantly reduce

both the power consumption, the system complexity and the hardware cost for the operators.

In this chapter we presents techniques for selecting antennas to reduce the number of RF chains required, sacrificing only a small amount of sum capacity. It is not easy to obtain the optimal subset of antennas to maximise the downlink sum-rate. In fact, it is a combinational problem which involves matching N_s RF chains to N_t antenna elements. In addition, it is assumed that the computational load cannot be negligible, and N_t can potentially be large. Other researchers have addressed this problem as either a convex optimisation problem or using sub-optimal approaches [112, 113].

This chapter proposes two semi-heuristic antenna-selection techniques based on the use of principal components analysis. In general, PCA is used in statistics to determine the dominant factors in a set of data. In our approach the transmit antennas are considered as a vector of random variables and the signals received at a user are considered as a random observation of that vector. Our first technique, called Linear Dependence Avoidance Selection (LDAS), uses PCA analysis to minimise the correlation between the remaining antennas after an antenna has been removed. Our second technique, called PCA-based antenna selection, improves the selection of antennas by choosing the subset of antennas that LDAS selects from, taking into consideration the smallest positive eigenvalue and associated eigenvector obtained in the PCA analysis.

The PCA technique also allows us to determine how the channel matrix structure impacts the antenna-selection process. In particular, the effect of the mean transmit power from an antenna element can be separated from the effect of correlations between antennas. We apply our technique to the ZF precoding scheme, under the assumption of equal transmit power and unequal power allocation for each data stream

using a water-filling algorithm. Simulation results are used to compare the proposed algorithms with exhaustive-search and mean-power antenna selection.

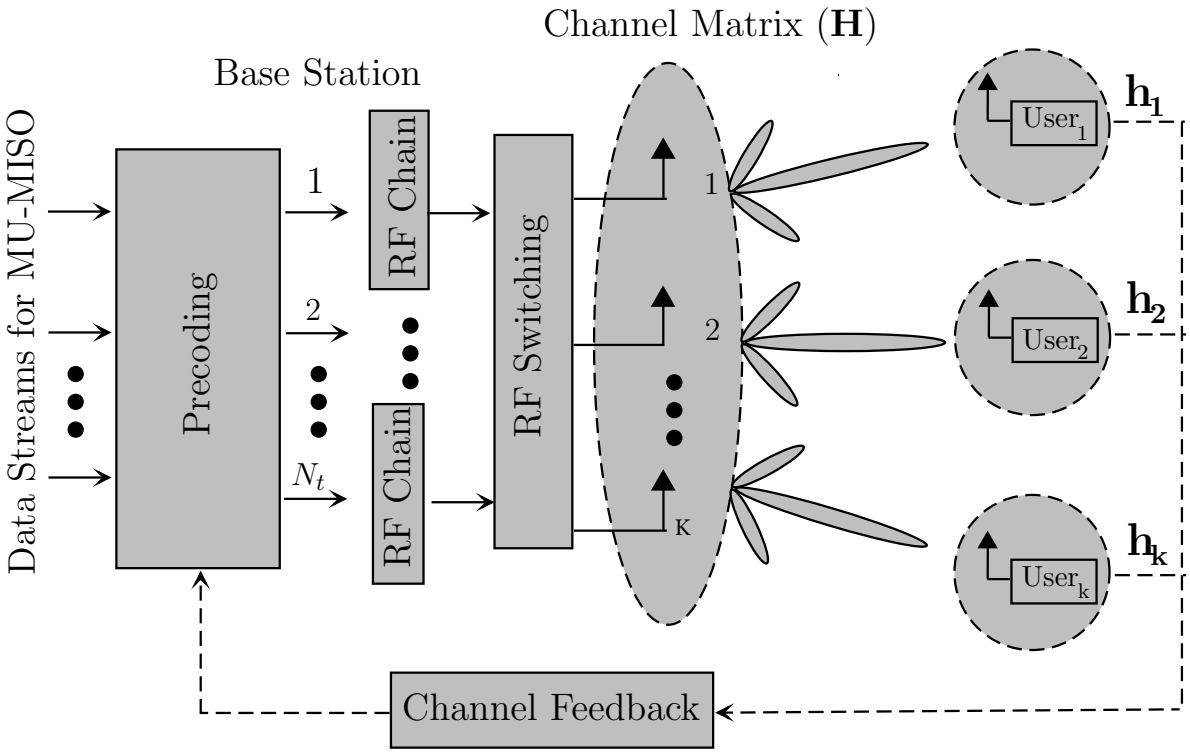


Figure 3.1: The MU-MISO System Model

3.2 Related Work

In this section, we review some current antenna-selection algorithms for MIMO broadcast wireless channels.

Various antenna-selection algorithms have been presented and analysed in recent years [26, 114]. In [114], the authors propose transmit antenna-selection to improve spectral efficiency in large-scale MIMO. To address channel hardening, in [115] the authors obtain a good approximation of large-scale MIMO distribution from which an antenna-selection algorithm is derived. In [116], a power-allocation algorithm is proposed for large-scale MIMO. The relationships between the selected number of

antennas, spectral efficiency, and better capacity trade-off are analysed in [26]. Several antenna-selection techniques have been proposed in [117–119], for multiuser MIMO networks. In [117], an antenna group-scheduling algorithm is proposed combining antenna selection and user scheduling to reduce feedback overhead. In [118], a joint antenna-selection and user-scheduling algorithm is proposed to maximise the achievable sum-rate with lower complexity. In [119], the authors rank antenna elements according to their channel gains for all users to derive a selection technique.

3.3 System Model

A MU-MIMO broadcast channel is considered with a single BS that supports K users, each equipped with one antenna (Figure 3.1). The BS has N_t transmit antennas that are used to transmit data to the K users such that $N_t \geq K$. Each selected antenna in the BS is supported by a separate RF chain. Assuming narrow-band communication, the $K \times 1$ vector of signals received by the users, denoted by $\mathbf{y}$, is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \qquad \textit{Eq. (3.1)}$$

where $\mathbf{H} = [h_{ij}]_{K \times N_t}$ is the complex channel matrix, h_{ij} is the channel gain between the j^{th} transmit antenna and the i^{th} user's receive antenna; the transmitted signal $\mathbf{x}$ is an $N_t \times 1$ vector and $\mathbf{n}$ is a $K \times 1$ vector denoting noise. The $K \times 1$ vector $\mathbf{u} = (u_1, \ldots, u_k)^T$ is the user data to be transmitted and it is assumed to be precoded by the matrix $\mathbf{G} \in \mathbb{C}_{N_t \times K}$, so that $\mathbf{x} = \mathbf{G}\mathbf{u}$. Column k of $\mathbf{G}$, denoted by $\mathbf{g}_k$, is the $N_t \times 1$ beam-forming vector for user k. The signal-to-interference-plus-noise (SINR)

ratio of the signal received at user k is given by

$$\Gamma_k = \frac{|\mathbf{h}_k \mathbf{g}_k|^2 v_k^2}{\sigma_k^2 + \sum_{j \neq k} |\mathbf{h}_k \mathbf{g}_j|^2 v_j^2}, \qquad k = 1, \ldots, K, \qquad \textit{Eq. (3.2)}$$

where $v_k = E(|u_k|^2)$ is the power allocated to the k^{th} user and σ_k^2 is the noise power at user k and $\mathbf{h}_k$ is the k^{th} row of $\mathbf{H}$. The total transmit power is given by $trE(\mathbf{x}\mathbf{x}^H) = \sum_k ||\mathbf{g}_k||^2 v_k$. The precoding problem is to maximise the weighted sum-rate of the system and obtain the beam-forming vectors $\mathbf{g}_k$ and user powers $v_k (k = 1, \ldots, K)$ under the constraint that the total transmitted power is P. This can be expressed as the following optimisation problem (e.g. see [120]):

$$\begin{aligned} &\underset{\mathbf{g}_k, v_k}{\text{maximise}} \sum_{k=1}^{K} q_k \log_2 \left(1 + \frac{|\mathbf{h}_k \mathbf{g}_k|^2 v_k^2}{\sigma^2 + \sum_{j \neq k} |\mathbf{h}_k \mathbf{g}_j|^2 v_j}\right) \qquad \textit{Eq. (3.3)} \\ &\text{subject to} \quad v_k \geq 0, \quad k = 1, \ldots, K \\ &\sum_k ||\mathbf{g}_k||^2 v_k \leq P, \end{aligned}$$

where q_k is the k^{th} quality-of-service weight. In the case of zero forcing, the transmit precoding matrix is given by $\mathbf{G} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1}$. With this precoding matrix $\mathbf{h}_k \mathbf{g}_j = 0, j \neq k$ and all interfering signals at a receiver are nulled. It can be shown in this case that the optimisation problem (3.3) can be rewritten as (e.g. see [120])

$$\begin{aligned} &\underset{v_k}{\text{maximise}} \sum_{k=1}^{K} q_k \log_2 \left(1 + \frac{v_k}{\sigma^2}\right) \qquad \textit{Eq. (3.4)} \\ &\text{subject to} \quad v_k \geq 0, \quad k = 1, \ldots, K \\ &\sum_k [(\mathbf{H}\mathbf{H}^H)]^{-1}_{(k,k)} \, v_k \leq P. \end{aligned}$$

Using Singular Value Decomposition (SVD) this problem can be expressed in terms of the eigenvalues of $\mathbf{H}\mathbf{H}^H$. To begin, suppose that the SVD of $\mathbf{H}$ is given by

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Xi}\mathbf{V}, \qquad \textit{Eq. (3.5)}$$

where $\boldsymbol{\Xi} = [y_{ij}]$ is a $K \times N_t$ matrix with singular values ρ_i on the diagonal (i.e. $y_{ii} = \rho_i$ $(i = 1, \cdots, K)$) and y_{ij} are zero elsewhere, $\mathbf{U} = [u_{ij}]$ is a $K \times K$ unitary matrix and $\mathbf{V}$ is an $N_t \times N_t$ unitary matrix. Therefore,

$$\mathbf{H}\mathbf{H}^H = (\mathbf{U}\boldsymbol{\Xi}\mathbf{V})(\mathbf{U}\boldsymbol{\Xi}\mathbf{V})^H = \mathbf{U}\boldsymbol{\Xi}\boldsymbol{\Xi}^H\mathbf{U}^H = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H. \qquad \textit{Eq. (3.6)}$$

so that $\boldsymbol{\Lambda} = \boldsymbol{\Xi}\boldsymbol{\Xi}^H$ is a diagonal matrix of the eigenvalues of $\mathbf{H}\mathbf{H}^H$ with diagonal elements $\lambda_i = |\rho_i|^2$ and the columns of $\mathbf{U}$ are the eigenvectors of $\mathbf{H}\mathbf{H}^H$ (it can also be shown that the columns of $\mathbf{V}$ are the eigenvectors of $\mathbf{H}^H\mathbf{H}$). The eigenvalues are assumed to be sorted in decreasing order: $\lambda_1 \geq \cdots \geq \lambda_K$. Considering the constraints in Equation (3.4), the following is obtained:

$$\begin{aligned}(\mathbf{H}\mathbf{H}^H)^{-1} &= (\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H)^{-1} = \mathbf{U}(\boldsymbol{\Lambda}^H)^{-1}\mathbf{U}^H \\ &= (\lambda_1)^{-1}\mathbf{u}_1\mathbf{u}_1^H + \ldots + (\lambda_K)^{-1}\mathbf{u}_K\mathbf{u}_K^H, \qquad \textit{Eq. (3.7)}\end{aligned}$$

where $\mathbf{u}_k$ is the k^{th} column of $\mathbf{U}$. For the diagonal elements

$$[(\mathbf{H}\mathbf{H}^H)^{-1}]_{ii} = \sum_{j=1}^{K} |u_{ij}|^2 \lambda_j^{-1}. \qquad \textit{Eq. (3.8)}$$

3.4 Principal Components Analysis (PCA)

Principal component analysis is a statistical technique first described by Pearson and Hotelling [33]. It uses a unitary transformation to linearly transform a set of possibly correlated random variables to uncorrelated new variables called Principal Components (PCs).

3.4.1 Population-Based Central PCA

In population-based PCA, assume that $\mathbf{x} = [x_1, \ldots, x_p]^T$ is a p-dimensional zero mean random vector with covariance matrix $\mathbf{\Sigma}$ given by $\mathbf{\Sigma} = E[\mathbf{x}\mathbf{x}^T]$ [33]. The aim of PCA is to transform $\mathbf{x}$ into a set of uncorrelated variables, say $\mathbf{z} = [z_1, \ldots, z_p]^T$ called the population-based Principal Components, that are linear combinations of the original variables $x_1, \ldots, x_p$. The linear combination for the k^{th} PC is given by

$$z_k = \mathbf{a}_k^H \mathbf{x} = \sum_{j=1}^{p} a_{kj} x_j, \quad k = 1, \ldots p, \qquad \textit{Eq. (3.9)}$$

where $\mathbf{a}_k = [a_{1k}, \ldots, a_{pk}]^T$ is a vector of constants. The vectors $\mathbf{a}_k$ are chosen to have unit norm. The first PC is obtained by choosing $\mathbf{a}_1$ to maximise the variance of $z_1 = \mathbf{a}_1^H \mathbf{x}$:

$$\begin{aligned} \underset{\mathbf{a}_1}{\text{maximise}} \quad & Var(\mathbf{a}_1^H \mathbf{x}) = \mathbf{a}_1^H \mathbf{\Sigma} \mathbf{a}_1 \\ \text{subject to} \quad & \mathbf{a}_1^H \mathbf{a}_1 = 1 \end{aligned}$$

Subsequent PCs are obtained by maximising the variance of $z_k = \mathbf{a}_k^H \mathbf{x}$ subject to them being uncorrelated with $z_1, \cdots, z_{k-1}$:

$$\begin{aligned} &\underset{\mathbf{a}_k}{\text{maximise}} \quad Var(\mathbf{a}_k^H \mathbf{x}) = \mathbf{a}_k^H \boldsymbol{\Sigma} \mathbf{a}_k \\ &\text{subject to} \quad \mathbf{a}_k^H \mathbf{a}_k = 1 \\ &\text{and} \quad Cov(z_j, z_k) = 0, \qquad\qquad j = 1, \ldots, k-1, . \end{aligned}$$

The above maximisation problems can be solved using Lagrange-multiplier techniques to show that the $\mathbf{a}_k$ are eigenvectors of $\boldsymbol{\Sigma}$, and if the k^{th} largest eigenvalue of $\boldsymbol{\Sigma}$ is λ_k then $\lambda_k = Var(\mathbf{a}_k x) = Var(z_k)$. The variables z_k have decreasing variance with k (see [33]). Dominant PCs can be interpreted as determining which linear combinations of the original random variables explain most of the variability. It can be shown that $\mathbf{a}_i^H \mathbf{a}_j = \mathbf{a}_i^H \boldsymbol{\Sigma} \mathbf{a}_j = 0$ for $i \neq j$. Defining the matrix $\mathbf{A}$ such that its k^{th} column is given by $\mathbf{a}_k$ then the transformation between $\mathbf{x}$ and $\mathbf{z}$ is given by $\mathbf{z} = \mathbf{A}^H \mathbf{x}$. Using the properties of $\mathbf{a}_k$, $\mathbf{A}$ is a unitary matrix. If $\boldsymbol{\Lambda}$ is the diagonal matrix of the eigenvalues of $\boldsymbol{\Sigma}$ then it can be shown that $\mathbf{A}^H \boldsymbol{\Sigma} \mathbf{A} = \boldsymbol{\Lambda}$. It can also be shown that $\boldsymbol{\Sigma}$ has the spectral representation [33]

$$\boldsymbol{\Sigma} = \sum_{k=1}^{p} \boldsymbol{\lambda}_k \mathbf{a}_k \mathbf{a}_k^H . \qquad \textit{Eq. (3.10)}$$

In the case that $\mathbf{x}$ does not have a zero mean, suppose that $E[\mathbf{x}] = \boldsymbol{\mu} = [\mu_1, \ldots, \mu_p]^T$ and $\boldsymbol{\Sigma} = E[\mathbf{x}\mathbf{x}^T] - \boldsymbol{\mu}\boldsymbol{\mu}^T$. After subtracting the mean to centre the random variables and determining $\mathbf{A}$ then $\mathbf{z} = \mathbf{A}^H(\mathbf{x} - \boldsymbol{\mu})$ is obtained.

3.4.2 Population-Based Non-Central PCA

Non-central population PCA is similar to central PCA except that the mean $\boldsymbol{\mu}$ is not subtracted from $\mathbf{x}$ before conducting the PCA analysis. Given random variables $x_1,\ldots,x_p$ the k^{th} non-central principal component is given as

$$z_k = \mathbf{a}_k^T\mathbf{x} = \sum_{j=1}^{n} a_{kj}x_j, \quad k = 1,\ldots p, \qquad \textit{Eq. (3.11)}$$

where $\mathbf{a}_k = (a_{k1},\cdots,a_{kp})$. (For ease of exposition the same notation will be used as in the central PCA case, relying on the context to determine which form of PCA is being used.) The maximisation problem becomes

$$\begin{aligned} &\underset{\mathbf{a}_k}{\text{maximise}} \quad E((\mathbf{a}_k^H\mathbf{x})^2) \\ &\text{subject to} \quad \mathbf{a}_k^H\mathbf{a}_k = 1, \\ &\text{and} \quad Cov(z_j,z_k) = 0,\ j = 1,\ldots,k-1,. \end{aligned}$$

Suppose that $\boldsymbol{\Sigma} = E(\mathbf{x}\mathbf{x}^H)$ and that the eigenvalues of $\boldsymbol{\Sigma}$, $\{\lambda_k\}$, are sorted in decreasing order, then $\mathbf{a}_k$ can be determined to be the eigenvector of $\boldsymbol{\Sigma}$ belonging to thek^{th} eigenvalue. This gives $\lambda_k = E((\mathbf{a}_k^T\mathbf{x})^2) = E(z_k^2)$. Letting $\mathbf{A}$ be the matrix whose k^{th} column is $\mathbf{a}_k$ the transformation $\mathbf{z} = \mathbf{A}^H\mathbf{x}$ is obtained.

3.4.3 Sample-Based PCA

In sample-based PCA it is assumed that associated with the vector random variables $\mathbf{x} = [x_1,\ldots,x_p]$, there is a set of N data samples where $\hat{\mathbf{x}}_n = [\hat{x}_{n1},\ldots,\hat{x}_{np}]^T$ is the n^{th} data sample of $\mathbf{x}$. All the observations can be expressed in a single matrix $\mathbf{X}$ where row n is the n^{th} data sample $\hat{\mathbf{x}}_n$. Denote by $\bar{\mathbf{X}}$ the matrix where each row is equal to

the row vector of column means of $\mathbf{X}$. The matrix $\mathbf{X}_0 = \mathbf{X} - \bar{\mathbf{X}}$ is the matrix of data sample values where the column mean has been subtracted from each column. The sample covariance matrix, $\widehat{\Sigma}$, is given by $\hat{\mathbf{\Sigma}} = \frac{1}{n-1}\mathbf{X}_0^H\mathbf{X}_0$. In order to perform central PCA analysis, $\widehat{\Sigma}$ is used as an estimate for Σ. However, to conduct non-central PCA analysis the sample matrix $\hat{\Sigma} = \frac{1}{n-1}\mathbf{X}^H\mathbf{X}$ is used as an estimate for Σ.

3.5 Antenna Selection

Initially, assume there are N_t transmit antennas and the aim is to remove one antenna in such a way that it results in the least reduction in sum capacity. The antenna selection problem can be formalized as follows. Let $\mathbf{H}_{-j}$ be the channel matrix with the j^{th} column (antenna) removed and let C_{-j} be the weighted sum capacity for such a channel. The problem is to determine $j^* = \underset{j}{\text{argmax}}\, C_{-j}$.

An exhaustive search would involve computing C_{-j} for all possible values of j but this would be of high computational complexity so alternatives of lower complexity are presented. A simple method of selecting antennas is mean selection. Express $\mathbf{H}$ as $\mathbf{H} = \mathbf{H}_0 + \overline{\mathbf{H}}$ where $\overline{\mathbf{H}} = [\overline{h}_{ij}]$ is the matrix where the rows are all equal to the row vector of column means of $\mathbf{H}$ and $\mathbf{H}_0 = \mathbf{H} - \overline{\mathbf{H}}$. In mean antenna selection the aim is to remove the antenna which has the weakest average signals, giving the selection rule: $j^* = \underset{j}{\text{argmin}}\, \overline{h}_{1,j}$.

An antenna selection technique called linear dependent avoidance selection (LDAS) is now presented. It is based on the following interpretation of $\mathbf{H}$. In $\mathbf{H}$, each column (antenna) of a matrix can be considered as representing a random variable and the signals received by user k (row k of $\mathbf{H}$) ($k = 1, \cdots, K$) as a data observation of those random variables. Next, note that the non-zero eigenvalues of $\mathbf{H}\mathbf{H}^H$ are the same as those of $\mathbf{H}^H\mathbf{H}$ (this follows from the general result that if $\mathbf{C}$ is an $m \times n$ matrix and $\mathbf{D}$

is an $n \times m$ matrix then the non-zero eigenvalue of **CD** and **DC** are identical [121]). The matrix $\mathbf{H}^H\mathbf{H}$ is of size $N_t \times N_t$, allowing antennas to be identified by columns in that matrix. It is assumed that **H** has a maximum rank of K. Using the eigenvalue decomposition $\mathbf{H}^H\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ then **U** is the same as **V** in (3.5) and it is assumed that the eigenvalues on the diagonal in $\mathbf{\Lambda}$ are ordered to satisfy $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_K > 0$ and $\lambda_{K+1} = \cdots = \lambda_{N_t} = 0$. The non-zero eigenvalues are the same as in (3.6). Letting $\widehat{\mathbf{\Sigma}} = \mathbf{H}^H\mathbf{H}/(K-1)$, the sample covariance matrix of the random variables representing the antennas, it can be seen that matrix of eigenvectors **U** is the same as the **A** matrix in the non-central PCA technique so that $\mathbf{a}_k = \mathbf{u}_k$ $(k = 1, \cdots, N_t)$. For $K+1 \leq k \leq N_t$, $E((\mathbf{u}_k^H\mathbf{x})^2) = \lambda_k = 0$. This implies

$$\sum_{i=1}^{N_t} u_{ki}x_i = 0\,. \qquad \textit{Eq. (3.12)}$$

To be more precise consider a particular antenna, say j, and suppose for the moment that $u_{kj} = 0$. This implies that

$$\sum_{i=1, i\neq j}^{N_t} u_{ki}x_i = 0 \qquad \textit{Eq. (3.13)}$$

and so the signals from the antennas $1, \cdots, j-1, j+1, \cdots, N_t$ would be linearly dependent. This would reduce the channel capacity significantly. The LDAS method of antenna selection is designed to avoid the signals from the remaining antennas being close to linearly dependent. Since the columns $K+1, \cdots, N_t$ of **U** all have a zero eigenvalue the aim is to maximise the norms of the row vectors belonging to those

columns, giving the selection rule:

$$j^* = \text{argmax} \parallel (u_{j,K+1}, \cdots, u_{j,N_t}) \parallel_2 . \qquad \textit{Eq. (3.14)}$$

The PCA-based antenna selection is built on LDAS by using non-central PCA analysis to reduce the number of antennas that LDAS makes a selection from. Examining (3.4), it is observed that to maximise the weighted sum-capacity, v_k should be as large as possible. In order to do this $[(\mathbf{H}\mathbf{H}^H)]_{kk}^{-1}$ should be as small as possible. Examining (3.8), the dominant term in the equation is that one associated with the smallest eigenvalue and the following approximation can be made:

$$[(\mathbf{H}\mathbf{H}^H)]_{ii}^{-1} \approx |\mathbf{u}_{iK}|^2 |\lambda_K|^{-1} . \qquad \textit{Eq. (3.15)}$$

To make this as small as possible, and making the approximation of ignoring the factor $|\mathbf{u}_{iK}|^2$, it is necessary to make λ_K as large as possible. That is, the smallest eigenvalue is being maximised. The eigenvector of $\mathbf{H}^H\mathbf{H}$ corresponding to the smallest non-zero eigenvalue is $\mathbf{u}_K$. The subset $\Omega_{PCA}(N_{PCA})$ of the N_{PCA} antennas selected consists of those antennas that make the smallest contributions to the eigenvalue $\mathbf{u}_K$ i.e. the N_{PCA} smallest element values in magnitude in $\mathbf{u}_K$:

$$\Omega_{PCA}(N_{PCA}) = \underset{N_{PCA}\text{smallest}}{\text{argmin}} \; |u_{Kj}| . \qquad \textit{Eq. (3.16)}$$

LDAS is then applied to the antennas in $\Omega_{PCA}(N_{PCA})$ to determine which antenna should be removed. To remove more than one antenna using either LDAS, PCA-based selection or mean selection an iterative approach is used, removing one antenna in each iteration.

3.5.1 Analysis using Non-Central PCA

This section examines in more detail the impact of the mean per-antenna channel gains and the correlation structure between antennas on the eigenvalues in PCA-based selection. Assuming that $\mathbf{H}$ has the maximum rank of K then $\mathbf{H}_0$ has a rank of $K-1$ and $\overline{\mathbf{H}}$ has a rank of 1. The relationship $\mathbf{H}_0^H\overline{\mathbf{H}} = \overline{\mathbf{H}}^H\mathbf{H}_0 = 0$, gives the decomposition

$$\mathbf{H}^H\mathbf{H} \;=\; (\mathbf{H}_o+\overline{\mathbf{H}})^H(\mathbf{H}_o+\overline{\mathbf{H}}) = \mathbf{H}_0^H\mathbf{H}_0+\overline{\mathbf{H}}^H\overline{\mathbf{H}}.$$

Suppose that the following SVD decompositions exist: $\mathbf{H}^H\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, $\mathbf{H}_0^H\overline{\mathbf{H}} = \mathbf{U}_0\mathbf{\Lambda}_0\mathbf{U}_0^H$ and $\overline{\mathbf{H}}^H\overline{\mathbf{H}} = \overline{\mathbf{U}}\overline{\mathbf{\Lambda}}\overline{\mathbf{U}}^H$ Inserting these into the above gives $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H = \mathbf{U}_0\mathbf{\Lambda}_0\mathbf{U}_0^H + \overline{\mathbf{U}}\overline{\mathbf{\Lambda}}\overline{\mathbf{U}}^H$. Since $\mathbf{U}$ is unitary then

$$\mathbf{\Lambda} \;=\; \mathbf{U}^H\mathbf{U}_o\Lambda_o\mathbf{U}_o^H\mathbf{U}+\mathbf{U}^H\overline{\mathbf{U}\Lambda\mathbf{U}}^H\mathbf{U}. \qquad \textit{Eq. (3.17)}$$

The first term on the right gives the contribution due to antenna correlation and the second gives the contribution due to the channel means.

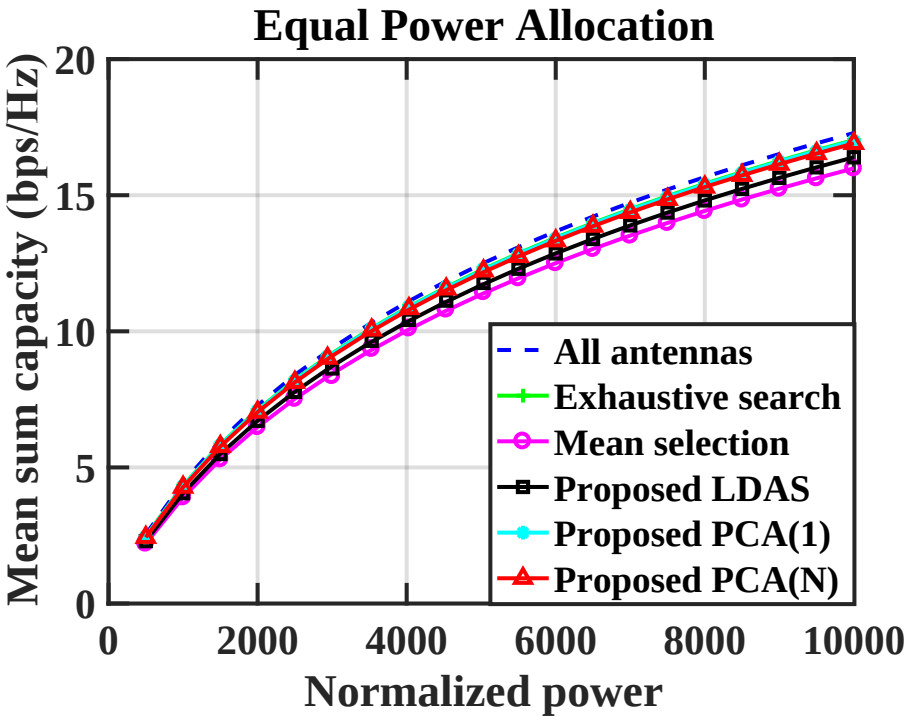


Figure 3.2: Mean sum-capacity versus P ($N_s = 10$) ($\mathbf{H}_z$) using i.i.d. Rayleigh fading channel model

3.6 Simulation Results and Discussion

In this section the simulation results using the equal power allocation method are provided to demonstrate the performance of the proposed algorithms. The random channel matrix $\mathbf{H}$ was generated in such a way so as to obtain some control over the column means and fading structure of the channels. The procedure is as follows. First, the matrix $\widetilde{\mathbf{H}}$ is generated so that each element is a Circularly Symmetric Complex Gaussian (CSCG) random variable with mean zero and variance 1. The next step is to subtract the column means from $\widetilde{\mathbf{H}}$ to produce the matrices $\widetilde{\mathbf{H}}_0$ and $\widetilde{\overline{\mathbf{H}}}$ so that $\widetilde{\mathbf{H}} = \widetilde{\mathbf{H}}_0 + \widetilde{\overline{\mathbf{H}}}$ where $\widetilde{\mathbf{H}}_0$ has zero column means and $\widetilde{\overline{\mathbf{H}}}$ is constructed from the column means. As the ranks of $\widetilde{\mathbf{H}}_0$ and $\widetilde{\overline{\mathbf{H}}}$ are less than K a perturbation matrix $\widetilde{\mathbf{H}}_\Delta$ of CSCG random variables is added.

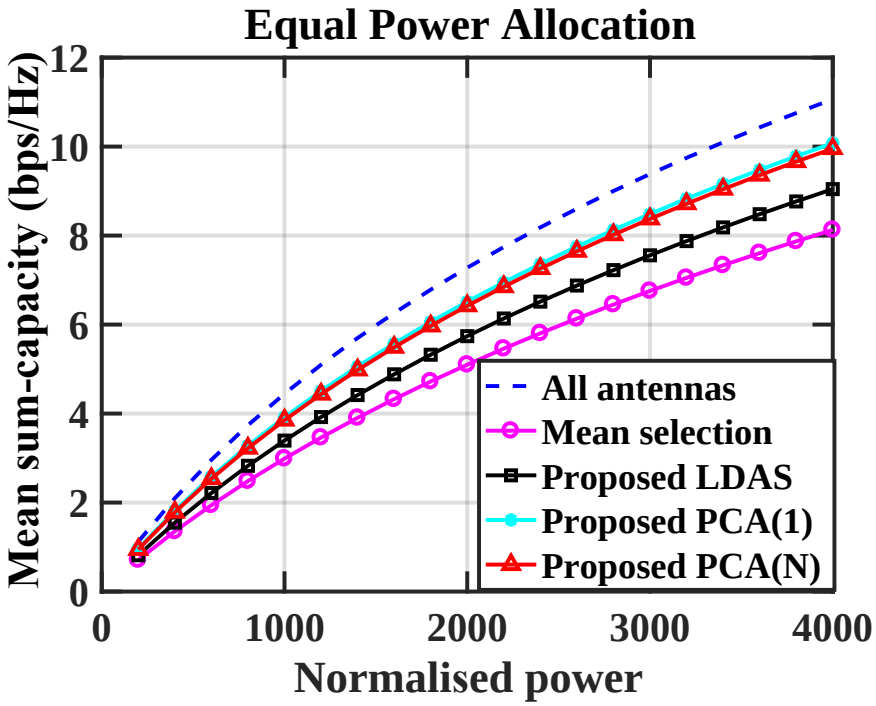


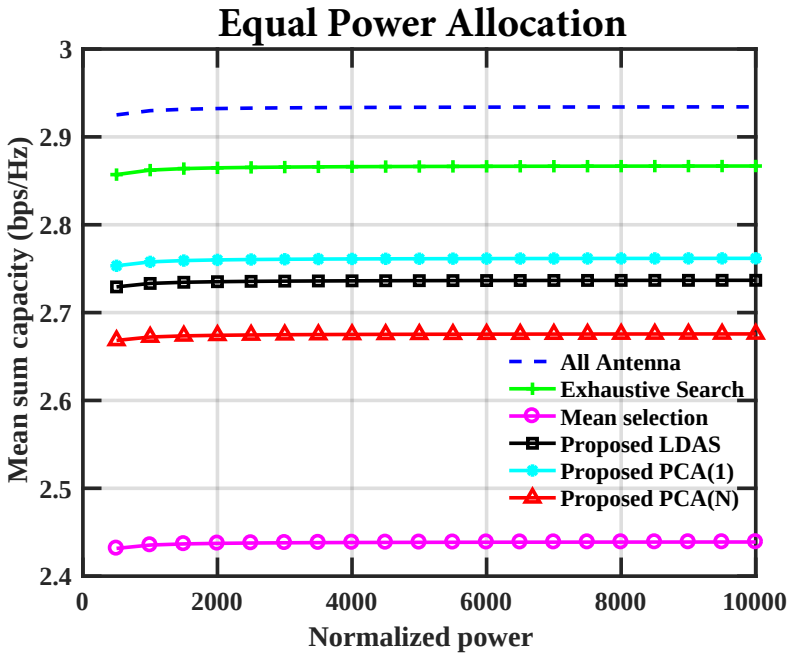
Figure 3.3: Mean sum-capacity versus P ($N_s = 6$) ($\mathbf{H}_z$) using i.i.d. Rayleigh fading channel model

The perturbation matrix has mean 0 and a variance σ_Δ^2 that can be adjusted. Using these two channel matrices are created: $\mathbf{H}_z = \widetilde{\mathbf{H}}_0 + \widetilde{\mathbf{H}}_\Delta$ and $\mathbf{H}_m = \widetilde{\overline{\mathbf{H}}} + \widetilde{\mathbf{H}}_\Delta$. For small values of σ_Δ^2 the matrix $\mathbf{H}_z$ has columns whose means are close to zero. Such a matrix can be used to model massive MIMO channels, where the law of large numbers applies. The matrix $\mathbf{H}_m$ has rows that are almost equal. Such a matrix can be used to model a distributed MIMO network where there is a cluster of users and a poor scattering environment. In all the results presented the parameters $N_t = 12$ and $K = 6$ were used.

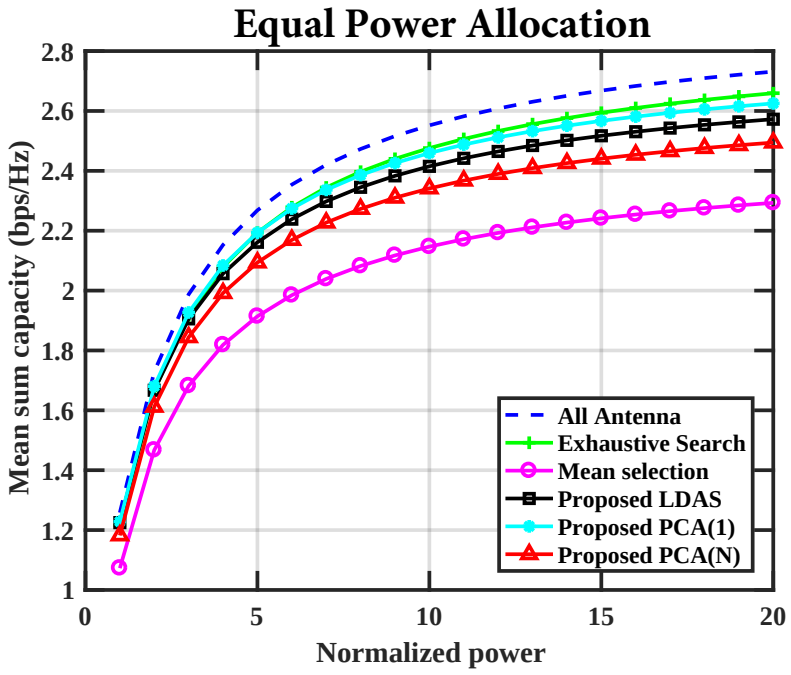
The first set of results plots the mean sum-capacity versus the available normalised transmit power (P) when different numbers of antennas, N_s, are selected. An ensemble of 5000 channel matrices was simulated by using the $\mathbf{H}_z$ method of generating channel matrices, with $\sigma_\Delta = 0.01$, when the channel correlations dominate, and the

mean sum-capacity was computed for different total power values with equal power allocation. The weights in the sum capacity formula were set to $q_k = 1$ so that sum-capacity was in normalised units.

Figures 3.2 and 3.3 show plots of the mean sum-capacity versus total power for $N_s = 10, 6$, respectively, using LDAS, PCA-based selection with $N_{PCA} = 1$ (PCA(1)), PCA-based selection with $N_{PCA} = 3$ (PCA(N)), and mean antenna selection as discussed in Section 3.5. The sum-capacity using mean antenna selection is shown in the plots when no antennas were removed. For $N_s = 10$, the mean sum capacity using exhaustive search is also shown. The plots show that LDAS, PCA(1) and PCA(N) all perform better than mean selection. The differences between the selection methods increase as N_s is reduced (results for $N_s = 10$ and $N_s = 6$ are shown). PCA(N) and PCA(1) are comparable, with PCA(N) being slightly better. Both are better than LDAS. The second set of results plots the mean sum-capacity versus the available normalised transmit power (P) using conjugate beamforming. Under the aforementioned simulation parameters conjugate beamforming is applied in proposed antenna selection techniques. A sum-capacity with normalized power $P = 10,000$ and $P = 4000$ using equal power algorithm is shown in Figure 3.4(a) (conjugate beamforming) and 3.5(a) (ZF). It is noticed that the proposed techniques of PCA(N), PCA(1) and LDAS performed better than mean selection. The optimal antenna selection is also shown in Figure 3.4(a). In addition, all the curves follows a straight line path for $P = 10,000$ and 4000 in Figure 3.4(a) and 3.5(a). However, when the power reduces as shown in Figure 3.4(b) (conjugate beamforming) and 3.5(b) (ZF), it does not affect mean-sum capacity much. Therefore, it can be deduced that conjugate beamforming gives less sum-capacity compared with ZF beamforming under similar simulation parameters of power, number of antennas at transmitter



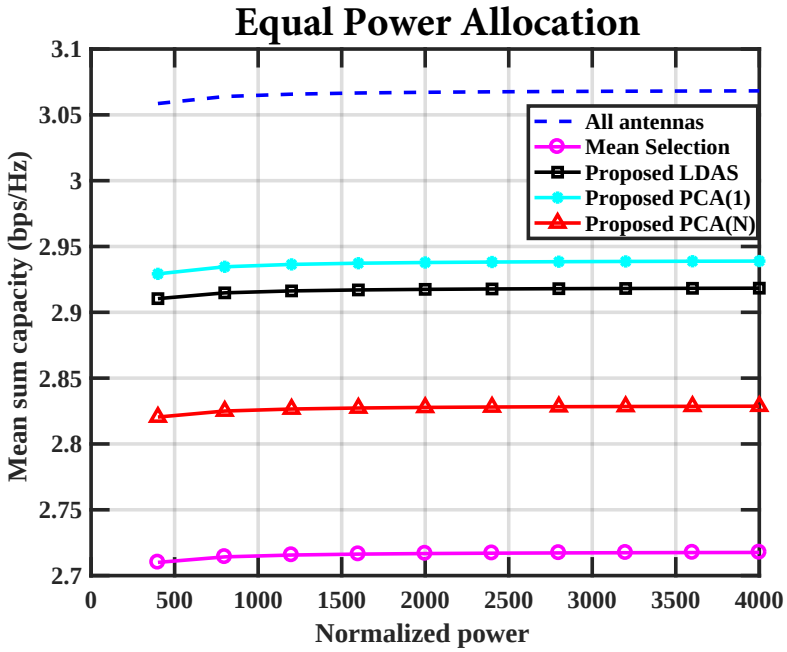
(a)



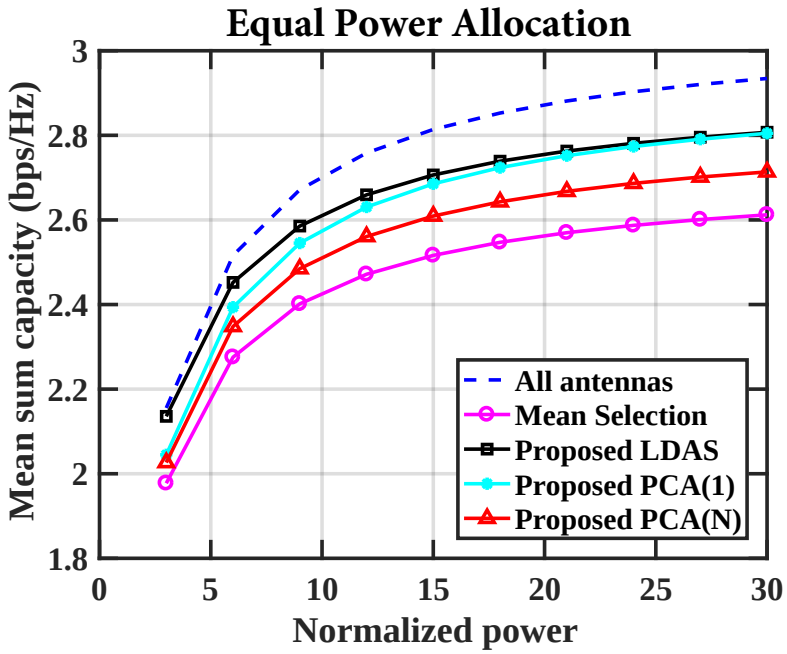
(b)

Figure 3.4: (a) Mean sum-capacity versus P ($N_s = 10$) ($\mathbf{H}_z$)
(b) Mean sum-capacity versus P ($N_s = 10$) ($\mathbf{H}_z$) (Conjugate beamforming)

and receiver and number of selected antennas as shown in Figure 3.2 and 3.3.



(a)



(b)

Figure 3.5: (a) Mean sum-capacity versus P ($N_s = 6$) ($\mathbf{H}_z$)
(b) Mean sum-capacity versus P ($N_s = 6$) ($\mathbf{H}_z$) (ZF)

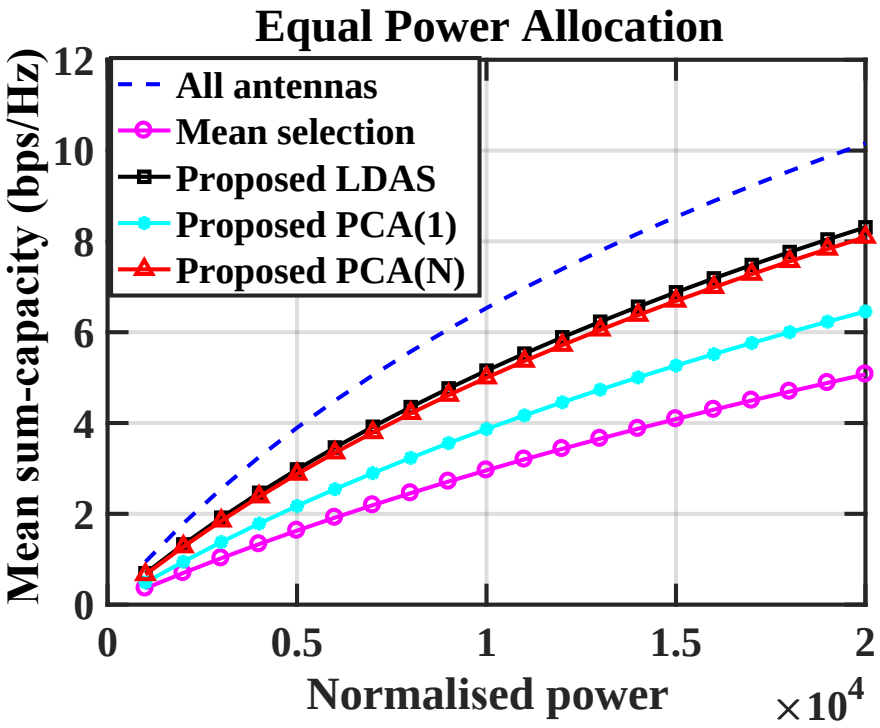


Figure 3.6: Mean sum-capacity versus P ($N_s = 8$) ($\mathbf{H}_m$) using i.i.d. Rayleigh fading channel model

Figure 3.6 shows a plot of the mean sum-capacity against power for $N_s = 8$ when the channel matrices were simulated using $\mathbf{H}_m$, when channel means dominate. In this case LDAS is just slighter better than PCA(N) and both are better than PCA(1).

Figure 3.7 shows the impact of varying N_s ($\mathbf{H}_z$) for $N_{PCA} = 3$ and $P = 10000$. For example, to achieve a sum-capacity of 15 using either PCA(1) or PCA(N) $N_s = 7$ RF chains are required. Using LDAS $N_s = 8$, and using mean selection $N_s = 9$, are required to achieve the same sum-capacity.

Figure 3.8 shows the impact of varying the size (N_{PCA}) of the subset used in PCA-based selection for the case of using $\mathbf{H}_z$ and $\sigma_\Delta = 0.01$. We set $N_s = 10$ and varied N_{PCA} from 1 to 10 (10 corresponds to LDAS).

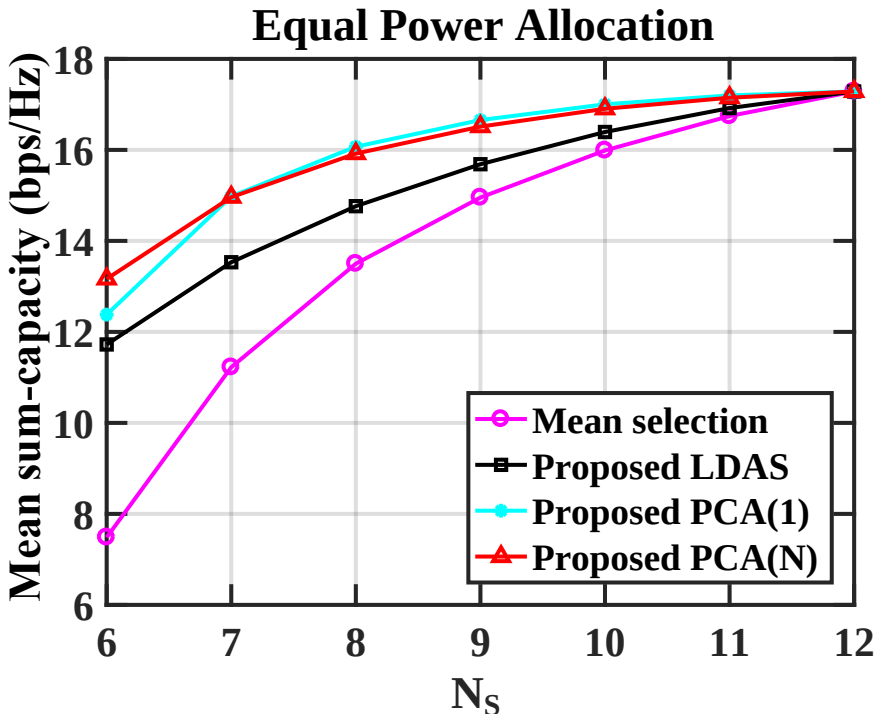


Figure 3.7: Mean sum-capacity versus N_s, $P = 10,000$, $(\mathbf{H}_z)$ using i.i.d. Rayleigh fading channel model

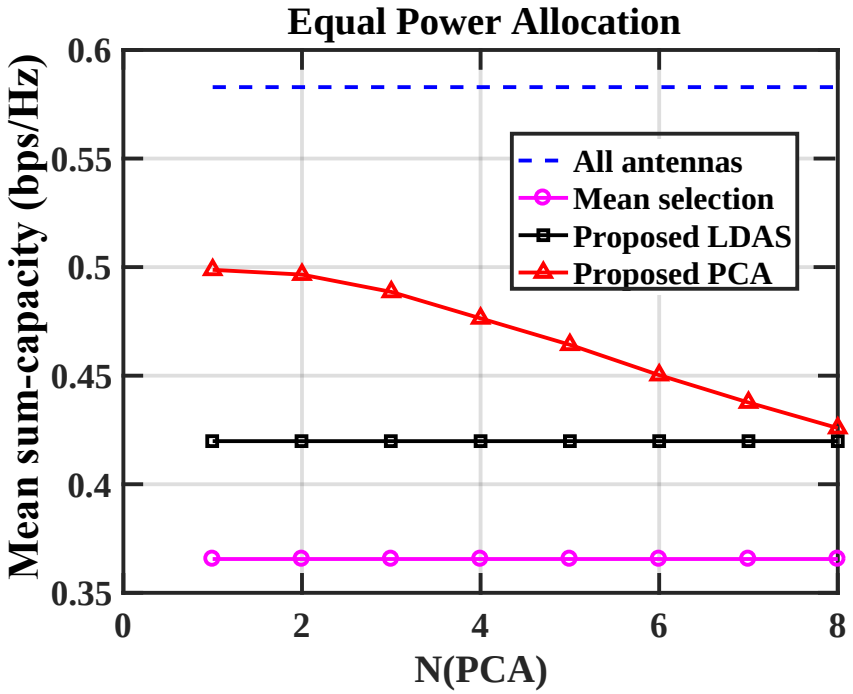


Figure 3.8: Sum-capacity versus N_{PCA} ($N_s = 10$) $(\mathbf{H}_z)$ using i.i.d. Rayleigh fading channel model

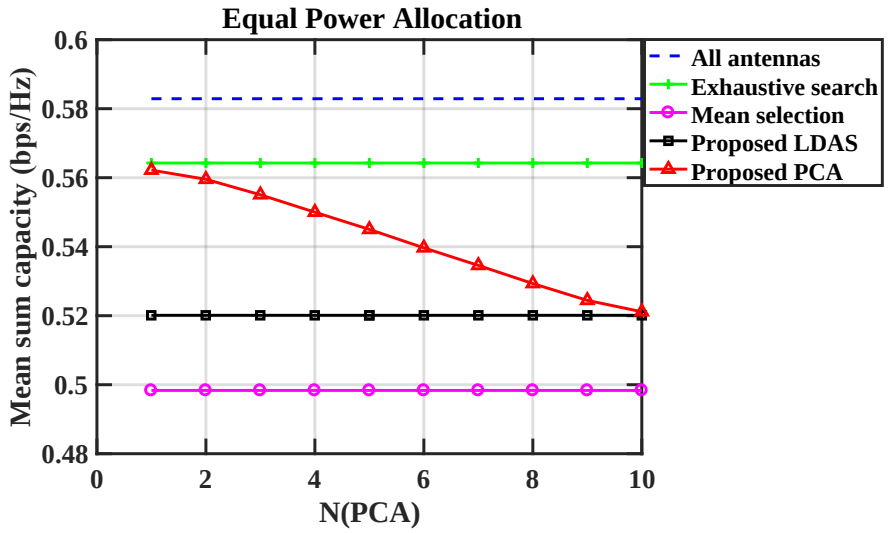


Figure 3.9: Mean sum-capacity versus N_{PCA} ($N_s = 8$) ($\mathbf{H}_m$) using i.i.d. Rayleigh fading channel model

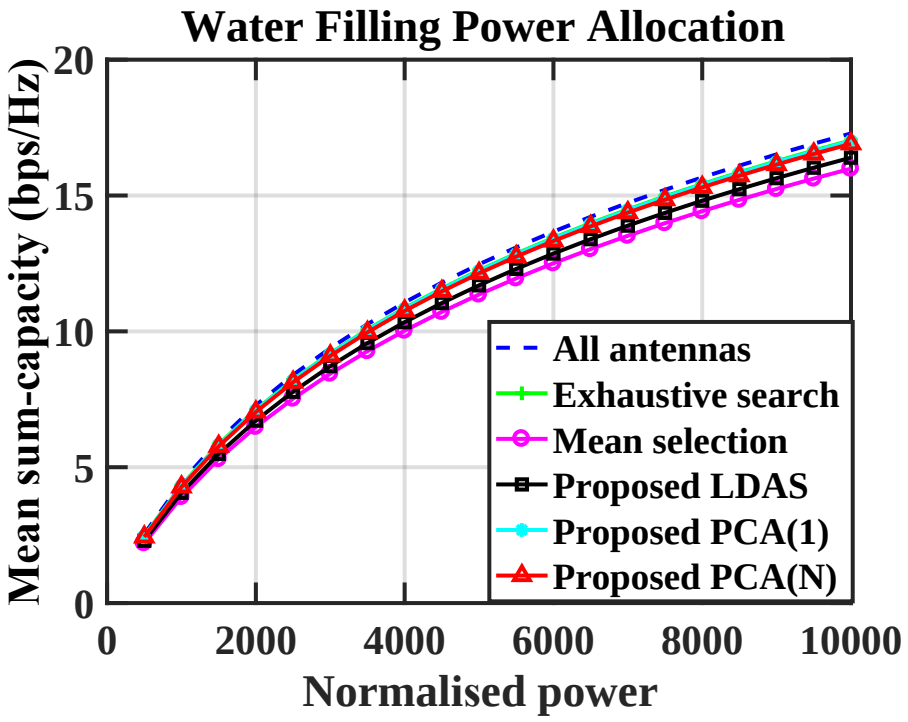


Figure 3.10: Mean sum-capacity versus P ($N_s = 10$) ($\mathbf{H}_z$) using i.i.d. Rayleigh fading channel model

An ensemble of 5000 channel matrices was generated and the same ensemble was used for each value of N_{PCA}. Choosing $N_{PCA} = 1$ gave the best performance. For $N_s = 6$ (not shown) choosing $N_{PCA} = 3$ was optimal. The results of using $\mathbf{H}_m$ with $N_s = 8$ are shown in Figure 3.9. Choosing $N_{PCA} = 1$ gives the worst result. However, using $N_{PCA} = 6$ is optimal for PCA selection (and better than LDAS).

The third set of results use the water-filling power-allocation method to evaluate the performance of the proposed algorithms. An ensemble of 5000 matrices was generated by using the $\mathbf{H}_z$ method of generating channel matrices and the mean sum-capacity is shown in Figure 3.10 using LDAS, PCA-based selection with $N_{PCA} = 1$ (PCA(1)), PCA-based selection with $N_{PCA} = 3$ (PCA(N)), and mean antenna selection. This figure also shows that the performance for different antenna selections is similar when water-filling power allocation method is used. The sum-capacity in this case is comparatively high as compared with other types of power allocation (see Figure 3.2). The figure also shows that LDAS, PCA(1) and PCA(N) all performed better than mean selection.

Figures 3.11 shows the channel estimation with imperfect channel state information. In this figure, the parameters $N_t = 12$ and $K = 10$ were used. An ensemble of 5000 channel matrices was simulated and a sum-capacity is plotted using ZF precoder with out σ_Δ. In addition, a sum-capacity is also plotted using ZF precoder with the addition of $\sigma_\Delta = 0.1$ and $\sigma_\Delta = 0.01$ under the normalized power $P = 10000$. The estimated error is shown in Figure 3.12. It can be noticed that the error increases as we increase the power and vice versa.

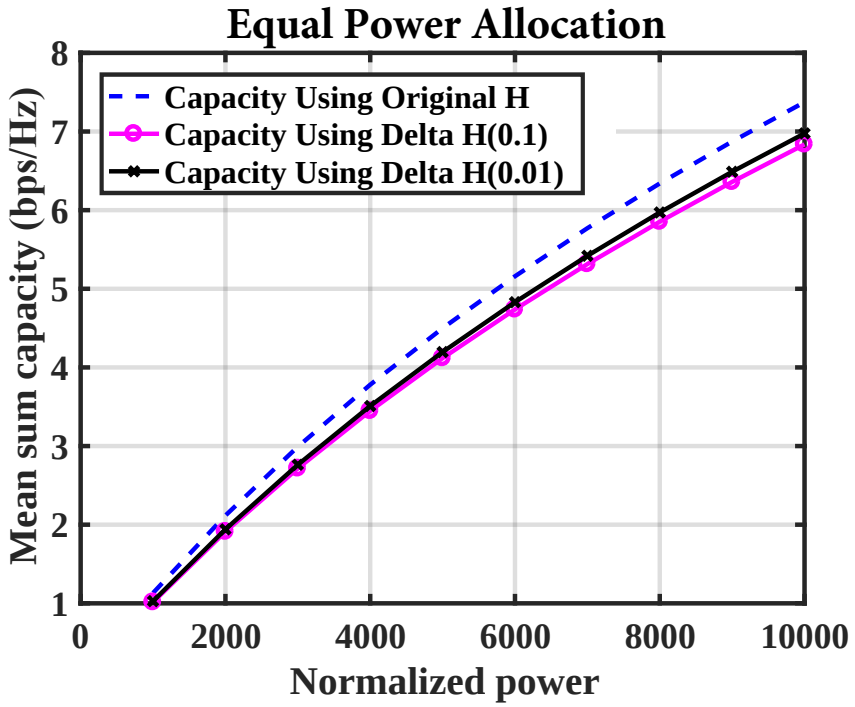


Figure 3.11: Channel estimation mean sum-capacity versus P

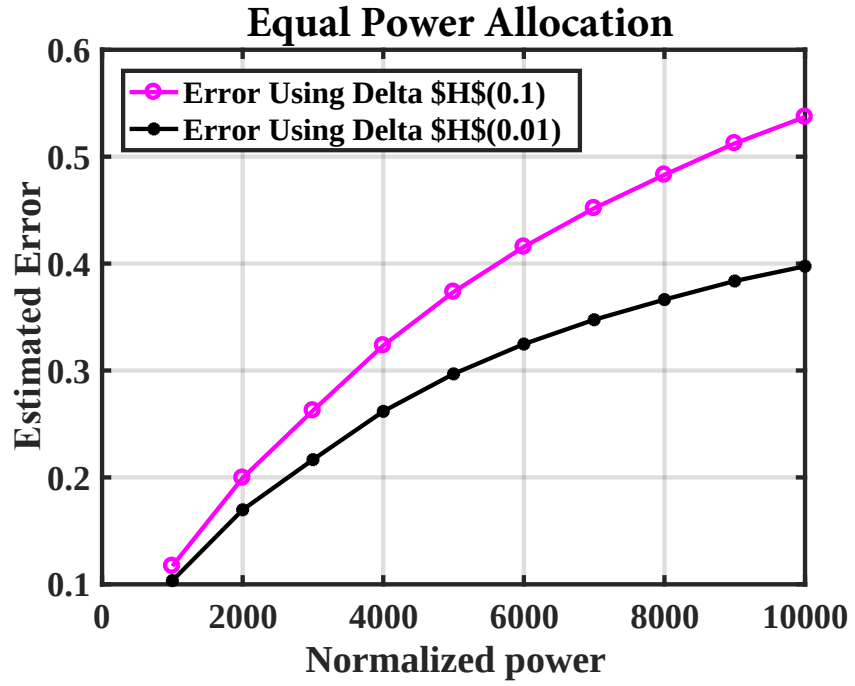


Figure 3.12: Estimated Error versus P

3.7 Summary of the Chapter

The selection of antennas at a MU-MIMO broadcast BS is an effective technique for the efficient use of RF units. In this chapter, we presented two methods of antenna selection based on non-central PCA (LDAS and PCA-based antenna selection) to reduce the number of RF chains required for use with zero forcing precoded MU-MIMO systems, and they performed much better than mean antenna selection. We showed how the mean sum-capacity varied with different channel conditions and which antenna selection schemes were preferred for those conditions.

All models are wrong, but some are useful.

George Edward Pelham Box (1919 - 2013)

CHAPTER 4

SPATIAL STRUCTURE OF MULTIPLE ANTENNA RADIO CHANNELS

This chapter briefly explains spatially correlated channels in Section 4.1, and previous related work is Section 4.2. A stochastic MIMO channel model is presented in Section 4.3 The design of transmit and receive correlation matrices using Weichselberger and Kronecker channel models has been described in details. Simulation results are discussed in Section 4.4 and this chapter concludes in Section 4.5.

4.1 Introduction

In the previous chapter we examined different antenna-selection algorithms and their performance in terms of the system total throughput. It was shown that the maximum system performance could be achieved by selecting a suitable subset of transmitting antennas at the base station.

In this chapter, we investigate more closely the effect of signal correlation between the antennas by using spatially correlated channel models. In particular, correlated channel models such as the Kronecker channel model and the Weichselberger channel model are used to observe the system behaviour in different realistic environments with the selection of the best transmit antennas at the base station.

In massive MIMO, channel modelling can give us a significant gain in terms of spectral efficiency by exploiting the multi-path richness of the channel. The spectral efficiency is an important parameter to analyse the throughput for a 5G system. Initial studies indicate a linear increase in the capacity of narrow-band MIMO systems with the number of antennas [31], [20]. However, the large capacity gain depends on the orthogonality of the sub-channels that constitute a MIMO system. To ensure the orthogonality, in massive MIMO the number of transmit antennas at the BS is significantly larger than the number of users served. Once the channel of each user to/from the BS is nearly orthogonal to that of any other user then applying linear precoding at the transmitter can possibly give the best link capacity [122].

However, the most recent studies are based on an idealised channel model, representing a rich scattering environment, that assumes the independent and identically distributed (i.i.d.) channel coefficients of a Rayleigh-fading channel model. Such idealized channel models oversimplify the throughput problems. Thus, in a real

propagation environment, it is shown in several experimental and analytical studies that, in the presence of spatial correlation (SC) in MIMO channels [30], the system throughput is expected to be less than that of i.i.d. channel models [31]. To simulate the correlated MIMO channels it is important to quantify the effect of correlation on system capacity.

The most popular channel model used to analyse throughput is the Kronecker channel model, that captures correlation at both receiver and transmitter sides in a real propagation environment [30]. The second correlated spatial model discussed in this chapter is the Weichselberger channel model.

In this chapter we propose two semi-heuristic antenna-selection techniques to investigate the ergodic capacity of downlink massive MU-MIMO systems under a spatially correlated channel model. The best link capacity is examined in spatially correlated channel models using antenna-selection algorithms that reduce the required number of radio-frequency chains.

4.2 Related Work

Most of the existing work devoted to analysing the spectral efficiency of MIMO system with respect to channel modelling is reported in [123–127]. In [123], a generalised analysis for the spectral efficiency in massive MIMO was performed analytically. The Kronecker channel model was reformulated by exploiting the Weichselbeger method to analyse the closed-form performance evaluation. In [124] the asymptotic behaviour of the spectral efficiency of massive MIMO systems is analysed by using the Kronecker channel model. Kamaga et al. [124] develop a comprehensive analytical channel model accounting for channel correlation, channel polarisation, antenna cross-polarisation discrimination and environmental cross-polar

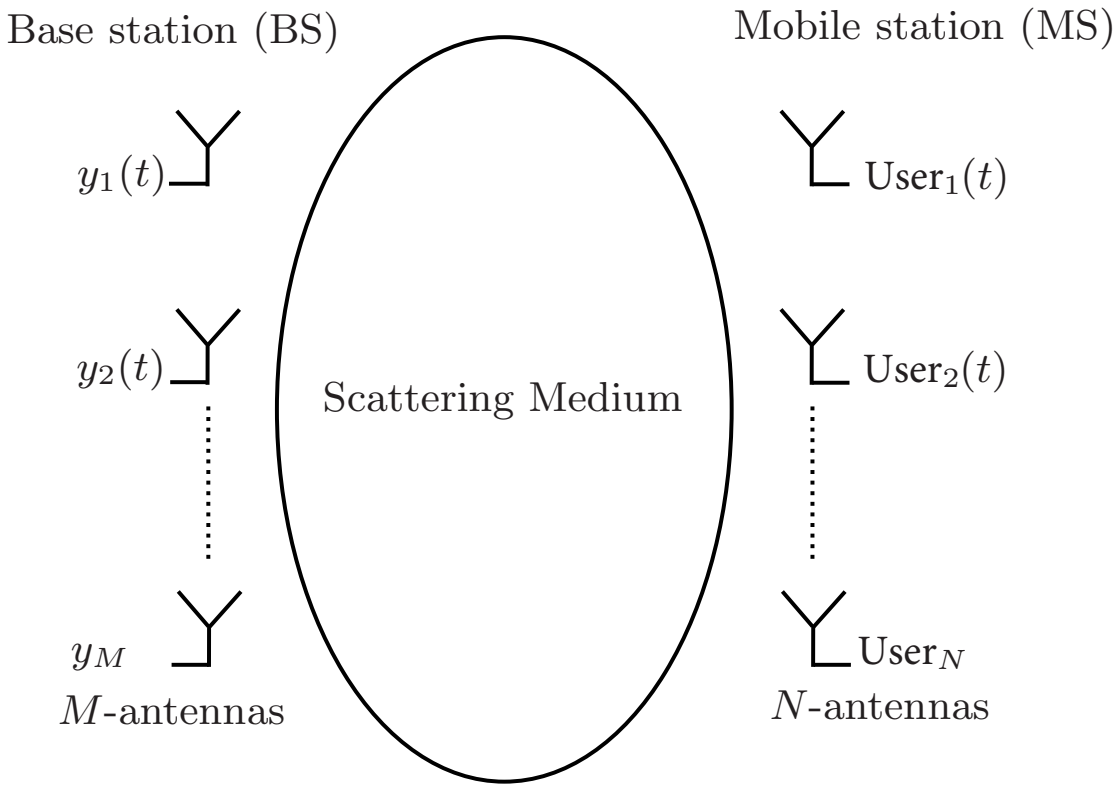


Figure 4.1: Two-antenna array in a scattering environment

coupling, path loss, shadowing effects, and multi-path fading, in a mathematically tractable way. In [125], lower and upper bounds on the SNR are derived to investigate the impact of SC by using the Kronecker model with known transmitter and receiver correlation matrices.

In [126, 127], to address the challenges related to antenna correlation, the system performance was studied by means of bounds on spectral efficiency.

The ZF precoding scheme is applied under the assumption of equal transmit power and unequal power allocation for each data stream using a water-filling algorithm. The simulation results compare the proposed algorithms with exhaustive-search and mean-power antenna-selection. These results verify that our novel proposed antenna selection algorithm requires less complexity, achieves higher channel

capacity and approaches the optimum method in its efficiency.

4.3 Stochastic MIMO Channel Model

A multi-user (MU) MIMO broadcast channel with a single BS that supports K users is considered. Each user is equipped with one antenna as shown in Figure 4.1. The BS uses N_t transmit antennas to transmit data to the K users such that $N_t \geq K$. In the BS each selected antenna is supported by a separate RF chain. The $K \times 1$ signal vector received by a users under the assumption of narrow-band communication is denoted by $\mathbf{y}$ and is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \qquad \textit{Eq. (4.1)}$$

where $\mathbf{H} = [\mathbf{h}_{ij}]_{K \times N_t}$ is the complex channel matrix, $\mathbf{h}_{ij}$ is the channel gain between the j^{th} transmit antenna and the i^{th} receive antenna of the user. The transmitted signal $\mathbf{x}$ is an $N_t \times 1$ vector and $\mathbf{n}$ is a $K \times 1$ vector denoting noise. This system model is discussed in detail in Chapter 3, Section 3.3.

The channel vectors in rich scattering environments are correlated due to the spatial correlation (SC) present among different users. This affects the system throughput and the effect of fading correlation can be modelled by using the Kronecker channel model and the Weichselberger channel model.

4.3.1 Kronecker Model

Kronecker channel modelling is widely used in wireless communication due to its main advantage, its simplicity [128]. Consider a correlated channel modelled by using the Kronecker model. In this model the channel matrix in (4.1) using the Kro-

necker model can be expressed as

$$\mathbf{H} = \mathbf{R}_{MS}^{\frac{1}{2}} \tilde{\mathbf{H}} \mathbf{R}_{BS}^{\frac{1}{2}H}, \qquad \textit{Eq. (4.2)}$$

where the elements of the channel matrix $\tilde{\mathbf{H}}$ are Zero Mean Circularly Symmetric Complex Gaussian (ZMCSCG) random variables with independent identically distributed (i.i.d.) elements, i.e. $\tilde{\mathbf{H}} \in \mathbf{C}^{N_r \times N_t} \sim N(0,1)$. Further $\mathbf{R}_{MS}$ and $\mathbf{R}_{BS}$ are the receive and transmit correlation matrices. The full covariance matrix in the Kronecker model is assumed as

$$\mathbf{R} = \mathbf{R}_{MS} \otimes \mathbf{R}_{BS}^{T}.$$

The eigenvalue decomposition of the transmit and receive covariance matrices are $\mathbf{R}_{MS} = \mathbf{U}_{MS} \Lambda_{MS} \mathbf{U}_{MS}^{H}$ and $\mathbf{R}_{BS} = \mathbf{U}_{BS} \Lambda_{BS} \mathbf{U}_{BS}^{H}$ respectively and then

$$\mathbf{R} = (\mathbf{U}_{BS} \otimes \mathbf{U}_{MS})(\Lambda_{BS} \otimes \Lambda_{MS})(\mathbf{U}_{BS} \otimes \mathbf{U}_{MS})^{H}.$$

The basic assumption in (4.2) is that the correlation between the transmitter and receiver can be separated, which holds when the antenna spacing in the transmitter and receiver is sufficiently smaller than the distance between the transmitter and receiver.

4.3.2 Weichselberger Channel Model

The mathematical equation for a Weichselberger channel as discussed in section 2.3.2.3 can be written as

$$\mathbf{H} = \mathbf{U}_{Rx} \Omega_{w} \odot \mathbf{G} \mathbf{U}_{Tx}^{T}, \qquad \textit{Eq. (4.3)}$$

where $\mathbf{U}_{Rx}$ and $\mathbf{U}_{Tx}$ are the received and transmit bases of eigenvectors. These eigenvectors are obtained from the singular-value decomposition of the received and transmit correlation matrices $\mathbf{R}_{Rx}$ and $\mathbf{R}_{Tx}$ respectively. Here, $\mathbf{G}$ represents an i.i.d. random matrix whose entries are zero mean with a complex-normal distribution, and Ω is a power coupling matrix. The power coupling matrix specifies the average energy coupled between an eigenvector at the receiver and transmitter. $(\odot)$ is the element-wise multiplication operator. The spatial correlation matrices $\mathbf{R}_{Rx}$ and $\mathbf{R}_{Tx}$ can be written as

$$\mathbf{R}_{Tx} = \mathbf{U}_{Tx}\mathbf{\Lambda}_{Tx}\mathbf{U}_{Tx}^{H}$$
$$\mathbf{R}_{Rx} = \mathbf{U}_{Rx}\mathbf{\Lambda}_{Rx}\mathbf{U}_{Rx}^{H},$$

The coupling matrix in a Weichselberger channel can be written as

$$\mathbf{\Omega}_{\mathrm{w}} = \left(\lambda_{Rx,1}, \lambda_{Rx,2} \ldots \lambda_{Rx,Nnx}\right)^{T} \left(\lambda_{Tx,1}, \lambda_{Tx,2}, \ldots, \lambda_{Tx,Nnx}\right) \qquad \textit{Eq. (4.4)}$$

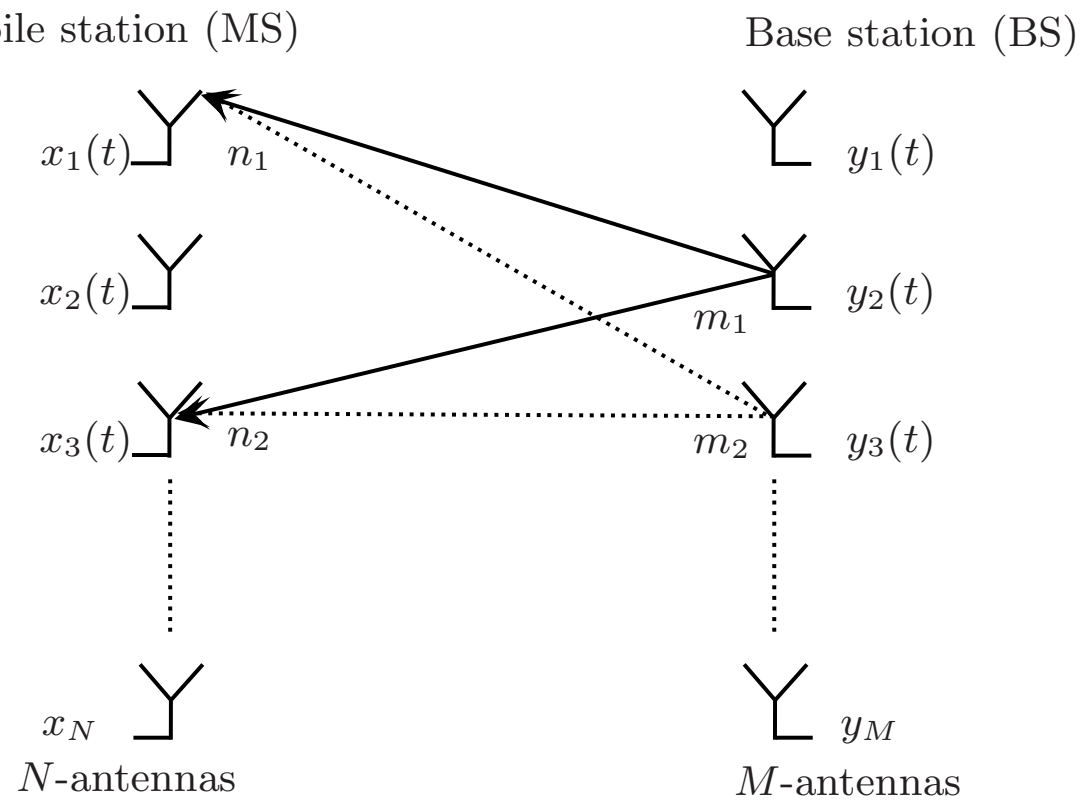


Figure 4.2: Downlink MIMO system

Referring to Figure 4.2, the correlation coefficient for a downlink MIMO system for two different MS antennas n_1 and n_2 can be expressed as

$$\rho_{n_1 n_2}^{MS} = \langle \alpha_{mn_1}, \alpha_{mn_2} \rangle,\ m = 1,2,\ldots,M. \qquad \textit{Eq. (4.5)}$$

Similarly the coefficient for two different BS antennas (uplink) m_1 and m_2 can be written as

$$\rho_{m_1 m_2}^{BS} = \langle \alpha_{m_1 n}, \alpha_{m_2 n} \rangle,\ n = 1,2,\ldots,N. \qquad \textit{Eq. (4.6)}$$

Using (4.5) and (4.6), spatial correlation (SC) matrices for receiver and transmitter

can be defined as

$$\mathbf{R}_{MS} = \begin{pmatrix} \rho_{11}^{MS} & \rho_{12}^{MS} & \cdots & \rho_{1N}^{MS} \\ \rho_{21}^{MS} & \rho_{22}^{MS} & \cdots & \rho_{2N}^{MS} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N1}^{MS} & \rho_{N2}^{MS} & \cdots & \rho_{NN}^{MS} \end{pmatrix}$$

and

$$\mathbf{R}_{BS} = \begin{pmatrix} \rho_{11}^{BS} & \rho_{12}^{BS} & \cdots & \rho_{1M}^{BS} \\ \rho_{21}^{BS} & \rho_{22}^{BS} & \cdots & \rho_{2M}^{BS} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{M1}^{BS} & \rho_{M2}^{BS} & \cdots & \rho_{MM}^{BS} \end{pmatrix}.$$

The correlation coefficients $\rho_{n_1 n_2}^{MS}$ and $\rho_{m_1 m_2}^{BS}$ can be generated for given Power Azimuth Spectrum (PAS) models [87]. In this chapter we use the n^{th} power of a cosine function PAS model and a Uniform PAS model to analyse the system throughput.

4.3.3 n^{th} power cosine PAS Model

The n^{th} power of a cosine PAS $p(\phi)$ is defined as

$$p(\phi) = \frac{Q}{n}\cos^n(\phi), \quad \frac{-\pi}{2} + \phi_\circ \leq \phi \leq \frac{\pi}{2} + \phi_\circ$$

where n is an even integer related to the beamwidth, and Q is a factor used to normalise $p(\phi)$, so that it integrates into a numerical value of one [61]. For simplification $n = 2$ is used in simulations [87].

$$
\begin{aligned}
R_{xx}(D,\phi_\circ) &= \int_{-\pi/2}^{\pi/2} \cos(D\sin\phi).\frac{Q}{n}\cos^n(\phi-\phi_\circ)d\phi \\
R_{xy}(D,\phi_\circ) &= \int_{-\pi/2}^{\pi/2} \sin(D\sin\phi).\frac{Q}{n}\cos^n(\phi-\phi_\circ)d\phi,
\end{aligned}
$$

where $R_{xx}(D,\phi_\circ)$ and $R_{xy}(D,\phi_\circ)$ represent correlations between the real parts of two received signals, and between the real and imaginary parts, respectively [129].

4.3.4 Uniform PAS Model

This model is suited for modelling a rich scattering environment, such as an indoor environment. It represents a situation with a uniform power distribution over the specified range of angle and the constants Q are derived such that $p(\phi)$ fulfils the requirements of a probability distribution function [129]:

$$
\int_{-\pi}^{\pi} p(\phi)\, d(\phi) = \sum_{k=1}^{n} \int_{\phi_{\circ,k}-\Delta\phi_k}^{\phi_{\circ,k}+\Delta\phi_k} Q_k\, d\phi = 1,
$$

which leads to $2\sum_{k=1}^{n} Q_k\Delta\phi_k = 1$ [129] and $\Delta\phi = \sqrt{3}\sigma_A$ [32]. Here σ_A represents azimuth spread and is defined by the central moment of PAS, that is

$$
\sigma_A = \sqrt{\left(\int (\phi-\phi_\circ)^2\, P_A(\phi)\, d\phi\right)},
$$

where $\phi_\circ$ is the mean DOA i.e., $\phi_\circ = \int(\phi P_A(\phi)\, d\phi)$ and $P_A(\phi) = \int(P(\phi,\tau)d\tau)$ [32].

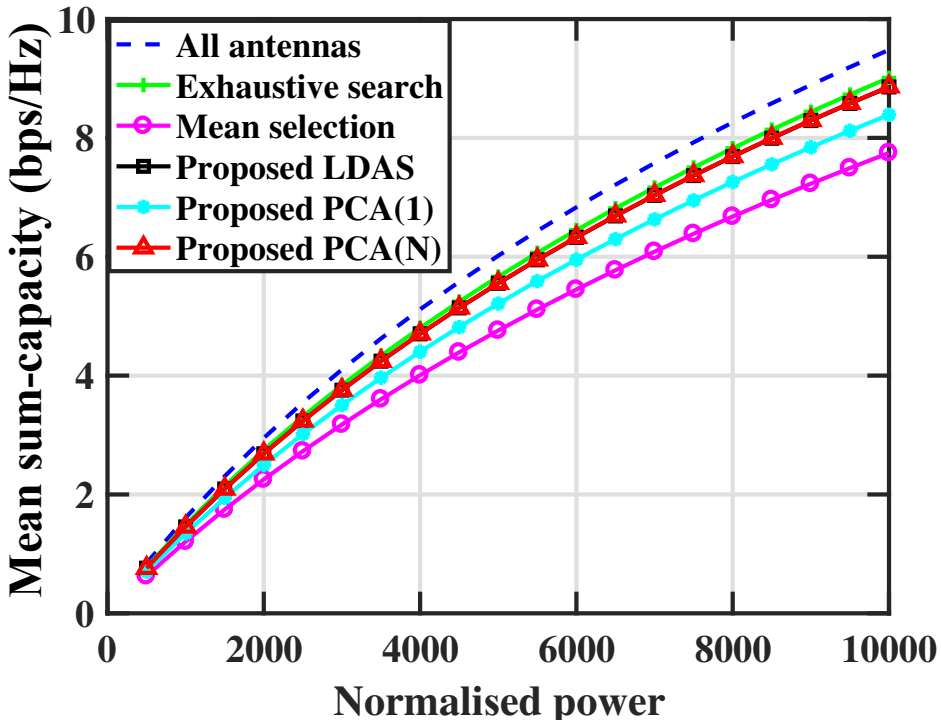


Figure 4.3: Mean sum-capacity versus P ($N_s = 10$) using Kronecker channel model

The SC function can be written as

$$R_{xx}(D,\phi_\circ) = J_\circ(D) + \frac{4Q\sum_{m=1}^{\infty} J_{2m}(D)\cos(2m\phi_\circ)\sin(2m.\Delta\phi_\circ)}{2m},$$

$$R_{xy}(D) = \frac{4Q\sum_{m=1}^{\infty} J_{2m+1}(D,\phi_\circ)\sin((2m+1)\phi_\circ)\sin((2m+1).\Delta\phi_\circ)}{2m+1}$$

where $J_m(\cdot)$ is the first-kind m^{th}-order Bessel function.

The theory of central and non-central PCA is used as discussed in Chapter 3 , Section 3.4. The formalization of antenna-selection problem is discussed in detail in Chapter 3, Section 3.5. However, the components of channel matrix used in Chapter 4 is stochastically modelled by using Weichselberger and Kronecker channel model as discussed in the relative sections.

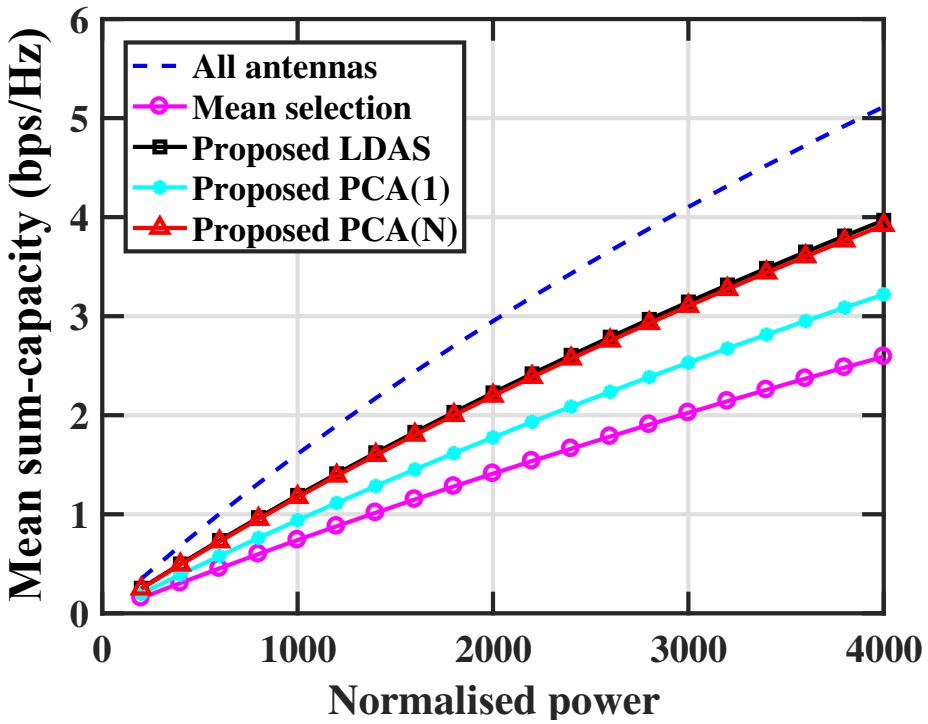


Figure 4.4: Mean sum-capacity versus P ($N_s = 6$) using Kronecker channel model

4.4 Simulation Results and Discussion

This section first presents the simulation results using equal power allocation in a Kronecker-based channel model. The simulated results in this chapter are extended by using the Weichselberger channel model that uses both the equal power allocation and the water-filling power allocation algorithms.

The channel matrix $\mathbf{H}$ was generated by using (4.2). In Figures 4.3 and 4.4 , $N_t = 12$ and $K = 6$ were used. An ensemble of 5000 channel matrices was simulated and the mean sum capacity versus the available normalised transmit power (P) with different numbers of N_s selected antennas is plotted. An equal power allocation is used to compute the mean sum capacity for different total power values by setting the weights $q_k = 1$ in the sum capacity formula to normalise the units. The mean sum

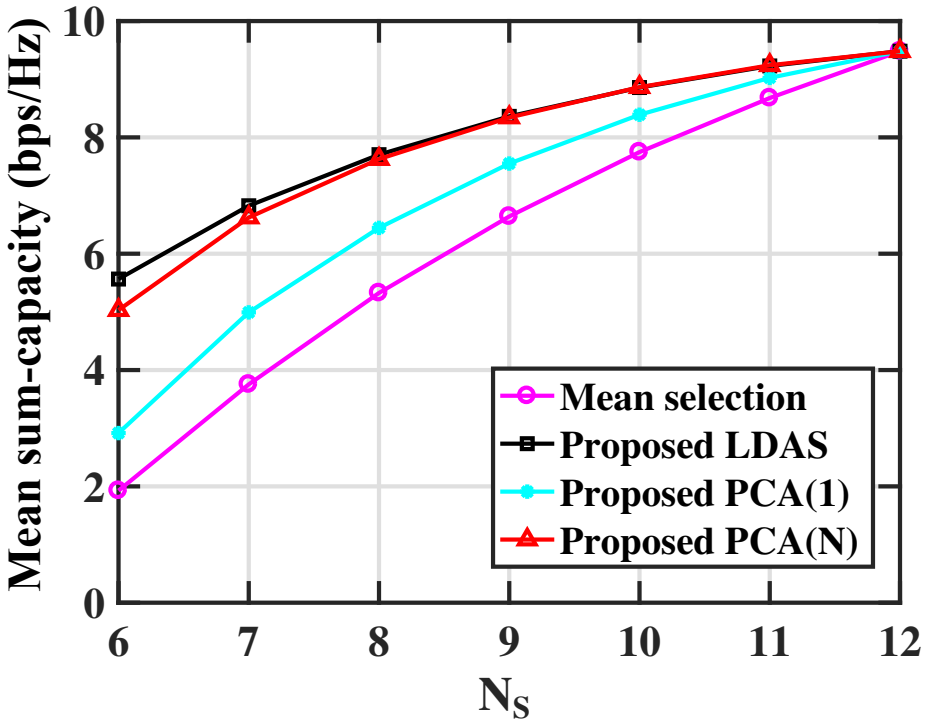


Figure 4.5: Mean sum capacity versus N_s using Kronecker channel model

capacity versus total power for $N_s = 10, 6$ is shown in Figures 4.3 and 4.4 respectively. These figures show the throughput performance when no antennas were removed and antenna-selection methods using LDAS, PCA-based selection with $N_{PCA} = 1$ (PCA(1)), PCA-based selection with $N_{PCA} = 3$ (PCA(N)). The mean sum capacity using exhaustive search is also shown for $N_s = 10$. The plots show that the LDAS, PCA(1) and PCA(N) selection techniques performed better than mean selection, and the difference between the selection methods increases as N_s is reduced (results for $N_s = 10$ and $N_s = 6$ are shown). PCA(N) and PCA(1) are comparable with PCA(N), being slightly better, however both are better than LDAS.

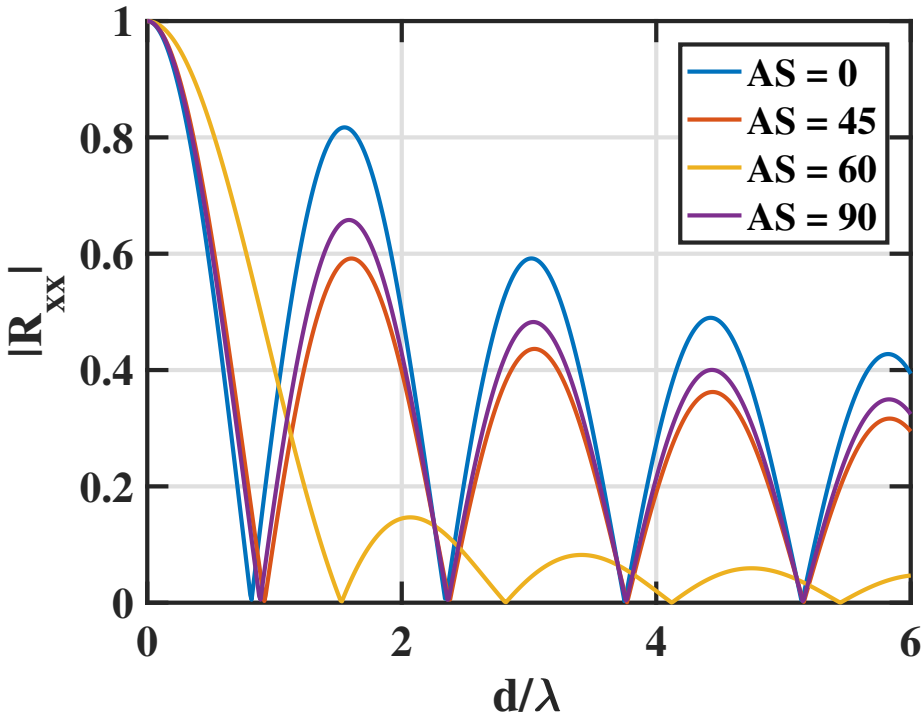


Figure 4.6: Spatial correlation coefficient using Kronecker channel model

The impact of varying N_s is shown in Figure 4.5. In the simulations, for $N_{PCA} = 3$ and $P = 10000$, the achievable mean sum capacity is approximately 8 using LDAS or PCA(N), i.e. $N_s = 8$ RF chains are required. However, PCA(1) uses $N_s = 9$ and the mean-selection method uses $N_s = 10$ RF chains to achieve the same sum capacity. The PAS distribution and the spatial correlation (SC) coefficient for the Uniform PAS model are shown in Figure 4.6. For the same antenna spacing the SC coefficients decrease as AS increases. The SC coefficients become nearly zero at a certain level, for example, it can be seen that they become nearly zero at the integer multiples of 0.5λ, 0.7λ, 0.8λ and 1.7λ, when AS is 0°, 45°, 60° and 90° respectively.

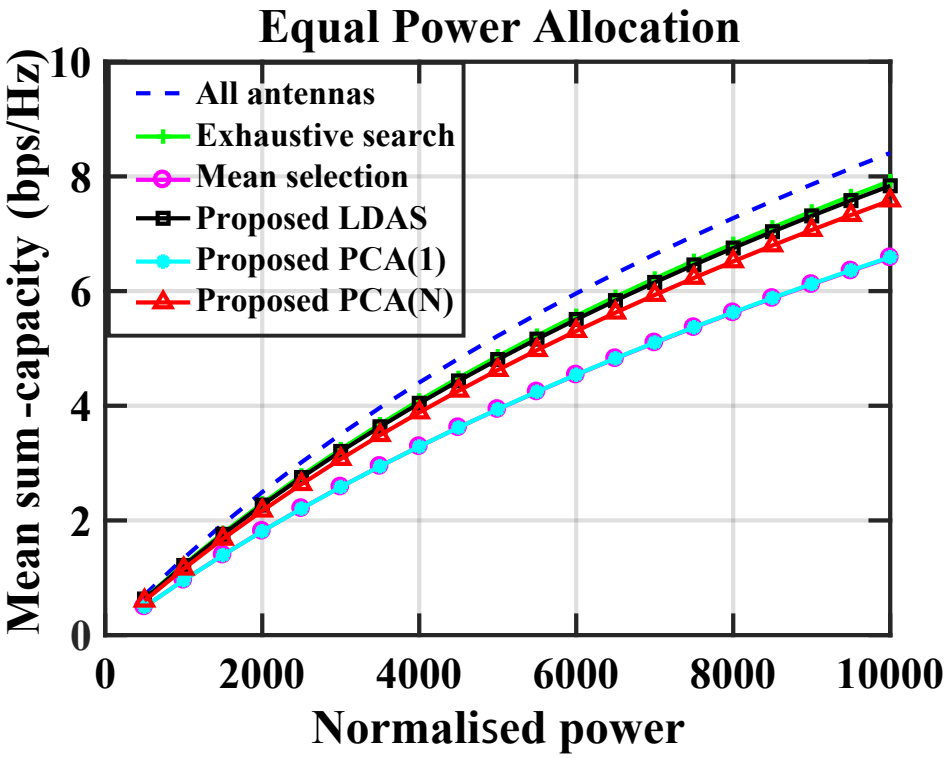


Figure 4.7: Mean sum-capacity versus P ($N_s = 10$) using Weichselberger channel model

In the second part of the simulations, the channel matrix $\mathbf{H}$ was generated by using (4.3). In Figures 4.7and 4.8, $N_t = 12$ and $K = 6$ were used. The mean sum capacity versus the available normalised transmit power (P) with different numbers N_s of selected antennas is plotted using the Monte Carlo method. An equal-power allocation and a water-filling algorithm is used to compute the mean sum capacity for different total power values. These figures show the throughput performance when no antennas were removed and the antenna-selection methods used LDAS, PCA-based selection with $N_{PCA} = 1$ (PCA(1)), PCA-based selection with $N_{PCA} = 3$ (PCA(N)). The LDAS, PCA(1) and PCA(N) selection techniques perform better than mean selection, and the difference between the selection methods increases as N_s is reduced (results for $N_s = 10$ and $N_s = 6$ are shown). PCA(N) and PCA(1) are comparable with PCA(N), being slightly better, however both are better than LDAS.

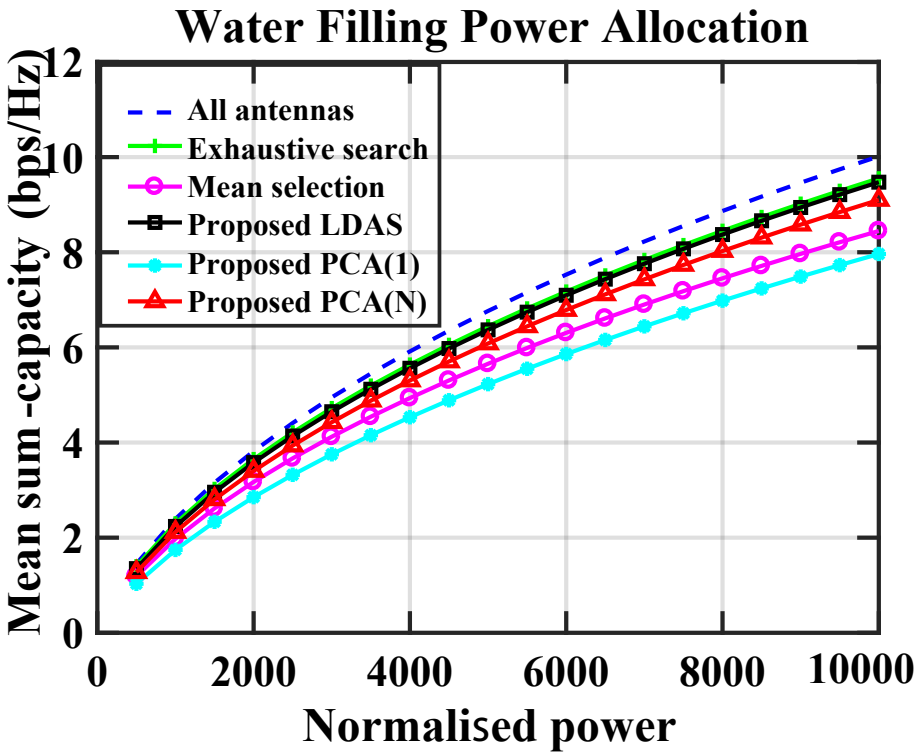


Figure 4.8: Mean sum-capacity versus P ($N_s = 10$) using Weichselberger channel model

Figures 4.9 and 4.10 show the difference made by the utilisation of expensive radio-frequency chains in the equal-power and water-filling allocation power algorithms. For instance, using equal-power allocation the proposed methods of LDAS or PCA(N), $N_s = 8$ RF chains are required to obtain the throughput of approximately 8, however, the other techniques required $N_s = 9$ and $N_s = 10$ to achieve the same sum-capacity. With the water-filling power allocation the trend is the same but the mean sum-rate is more than with equal power allocation.

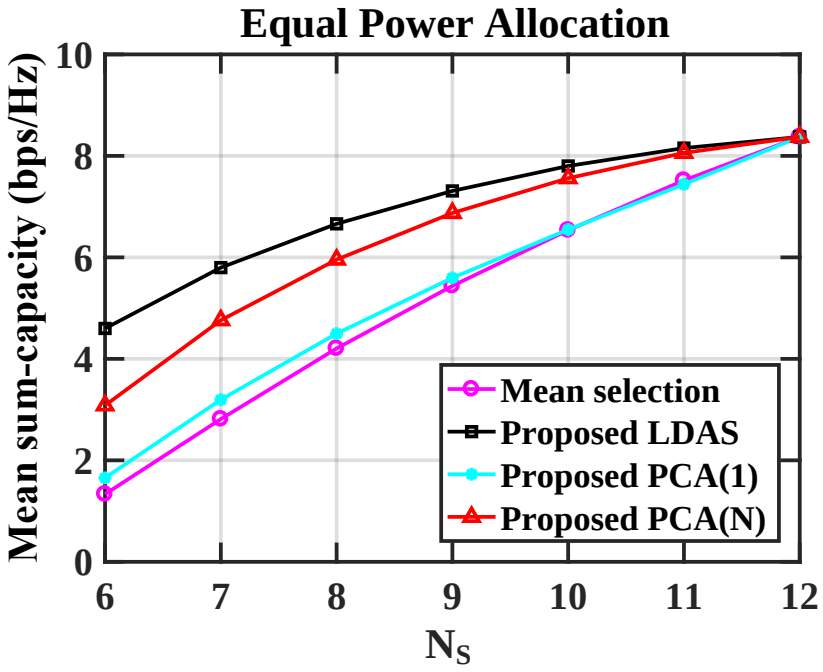


Figure 4.9: Number of RF chains using Weichselberger channel model

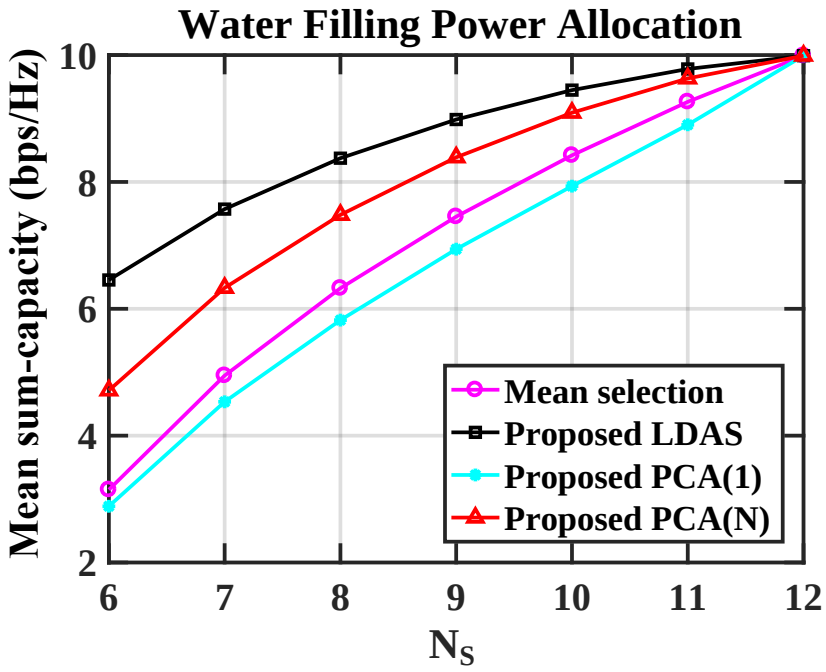


Figure 4.10: Number of RF chains using Weichselberger channel model

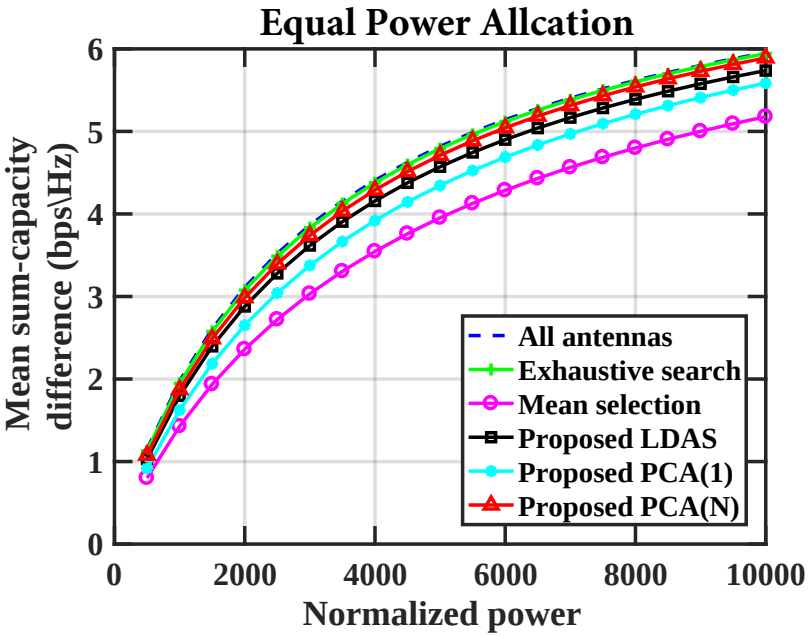


Figure 4.11: Mean sum-capacity difference (bps/Hz) Versys Power $P = 10000$ (Weichselberger channel model Vs Kronecker channel model)

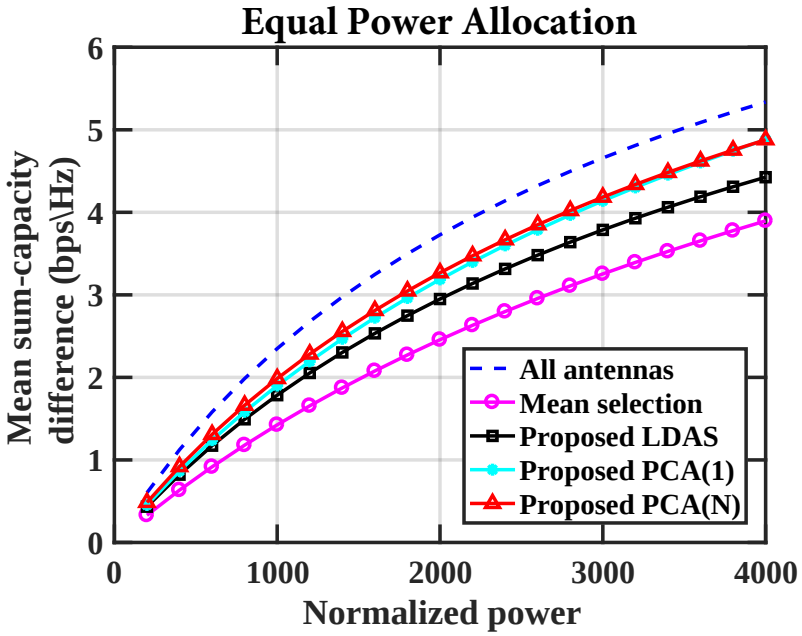


Figure 4.12: Mean sum-capacity difference (bps/Hz) Vs Power $P = 4000$ (Weichselberger channel model Vs Kronecker channel model)

In Figure 4.11 and 4.12, the comparison of sum-capacity using Weichselberger channel model and Kronecker channel model with normalized power $P = 10000$ and $P = 4000$ is plotted. An ensemble of 5000 matrices were used and a difference of sum-capacity is shown in mean selection, LDAS, PCA(1) and PCA(N). An exhaustive search algorithm is shown in Figure 4.11. It can be predicted that Weichselberger model perform better than Kronecker channel model in this case.

4.5 Summary of the Chapter

Antenna selection in MU-MIMO broadcast at BSs is an effective technique to use RF units efficiently. In this chapter, we present two methods of antenna selection to reduce the required number of RF chains. The proposed techniques are based on non-central PCA (LDAS and PCA-based antenna selection) which uses Kronecker and Weichselberger channel modelling. It is also shown how the mean sum capacity varies with different channel conditions and which antenna-selection schemes are preferable for those conditions.

Technical skill is mastery of complexity, while creativity is mastery of simplicity.

Erik Christopher Zeeman (1925 - 2016)

CHAPTER 5

DIMENSIONALITY REDUCTION OF LARGE MATRICES

In this chapter Section 5.1 briefly explains the importance of reducing matrix size and Section 5.2 gives a brief introduction to linear precoding techniques.The analysis of channel complexity reduction using PCA is explored in Section 5.3. The computational complexity in terms of FLOPs is presented in Section 5.4. Simulation results are discussed in Section 5.5 and Section 5.6 concludes this chapter.

5.1 Introduction

In Chapters 3 and 4 we extensively discussed the selection of the best transmitting antennas at the base station to maximise the sum-capacity. The selection process reduces the hardware system complexity in uncorrelated and spatially correlated channel models.

In this chapter, we present the technique that reduces the dimensions of large matrices using the Floating Point Operations (FLOPS) method, and the sum-capacity is evaluated with the down-sized matrix. The conventional methods of ZF and MMSE precoding schemes are used to examine the throughput with a huge matrix size. However, with a negligible loss of sum rate, the proposed principal component analysis technique can reduce the extensive-computations of large matrices compared with the computations used in conventional techniques.

In massive MIMO, when the number of transmit antennas at the base station is significantly larger than the number of users served, the channel of each user to/from the BS is nearly orthogonal to that of any other user, and by using linear precoding at the transmitter we can possibly approach the best link capacity [122]. The authors in [130] derive an approximation of the achievable sum rates with different linear precoding techniques in massive MIMO; when the number of antennas grows without bound that is a system with an unlimited number of BS antennas. Therefore, linear precoding of the transmitted signal of an antenna array in massive MIMO is often said to direct a signal from the antenna array towards one or more receivers.

The problem here is the computation of large precoding matrices in massive MIMO, which becomes a bottleneck in the transmission of data. A possible solution is in the reduction of the complexity of a channel, which enables the system to have

a reduced computational demand [131]. There are numerous complexity-reduction techniques as discussed in Chapter 2.

In this chapter, we study a complexity-reduction scheme in a massive MISO system using an algorithm of PCA. The main motivation of using PCA in this work is its simplicity and its potential to achieve a balance between reduced complexity and loss of information. The reduction in the complexity of a channel can be interpreted in terms of lossless and lossy complexity reduction. In lossless complexity reduction, redundant bits and information found to be less relevant are discarded once they have gone through the reduction process. Part of the information is lost permanently in the lossy complexity reduction method, but, it has the potential to achieve a better reduction ratio (RR) to benefit a wide range of applications, with a trade-off between reducing complexity and losing information [131].

The critical parameter that needs to be set in PCA analysis for reducing channel complexity with a negligible loss of sum rate is the number of singular values, i.e. Principal Components (PCs). Typically these PCs have an exponentially decaying trend, meaning that the ordered PCs decrease rapidly and only a few PCs are required to closely approach the sum rate with much reduced complexity.

The key feature of this novel technique is to find a new subspace using channel statistics and project it into a subspace which is uncorrelated. Furthermore, the reduced channel then uses the ZF precoding scheme, under the assumption of equal transmit power for individual data streams, to observe system performance.

Consider a downlink MU-MIMO wireless communication system employing a single BS that supports K users each equipped with one antenna (Figure 5.1). The BS has N_t transmit antennas that are used to transmit data to the K users such that

$N_t \geq K$. The downlink received signal y_k at user k is determined by

$$y_k = \mathbf{h}_k \mathbf{x} + n_k, \qquad k = 1, \ldots, K, \qquad \textit{Eq. (5.1)}$$

where $\mathbf{h}_k \in \mathbb{C}^{1 \times N_t}$ represents the MISO channel, $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ is the transmitted signal and $\{n_k\}$ are i.i.d. complex Gaussian noise terms with unit variance. The power constraint for the input signal is $\mathbb{E}[\mathbf{x}^H \mathbf{x}] \leq P$, which implies that the total transmit power is not dependent on the number of transmit antennas. We assume that the transmit antennas and users are spaced sufficiently far apart such that the entries of $\mathbf{h}_k$, for $k = 1, \ldots, K$ can be modelled as a set of i.i.d. zero mean circularly symmetric complex Gaussian random variables. These entries have unit variance, that is $\mathbf{h}_k$, $k = 1, \ldots, K$ is distributed as $\mathcal{CN}(0, \mathbf{I})$, where $\mathbf{I}$ is the identity matrix.

5.2 Linear Precoding

Let us consider the overall channel matrix $\mathbf{H} \in \mathbb{C}^{K \times N_t} \mathcal{N}(0,1)$. Assuming a subset of k users $\mathcal{J} = \{j_1, \ldots, j_k\} \subset \{1, \ldots, K\}$ stacking on top of each other then the channel matrix $\mathbf{H}_\mathcal{J}$ can be written as $\mathbf{H}_\mathcal{J} = [\mathbf{h}_1, \ldots, \mathbf{h}_K]$, where $\mathbf{h}_i \in \mathbb{C}^{1 \times N_t} (i = 1, \ldots, K)$ are the rows of matrix $\mathbf{H}_\mathcal{J}$ and represent the i^{th} user channel. In order to keep the interference from other users to a minimum and to serve multiple users simultaneously, the transmitted signal is precoded by the BS under the assumption of complete knowledge of channel state information at the transmitter side (CSIT). Let a vector $\mathbf{u} = (u_1, \ldots, u_k)^T$ of size $K \times 1$ be the user data to be transmitted, and it is assumed to be precoded by the matrix $\mathbf{G} \in \mathbb{C}^{N_t \times K}$ at the BS defined as $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_K]$, where $\mathbf{g}_i \in \mathbb{C}^{N_T \times 1}$ $(i = 1, \ldots, K)$ are the column vectors of the precoding matrix $\mathbf{G}$, so that

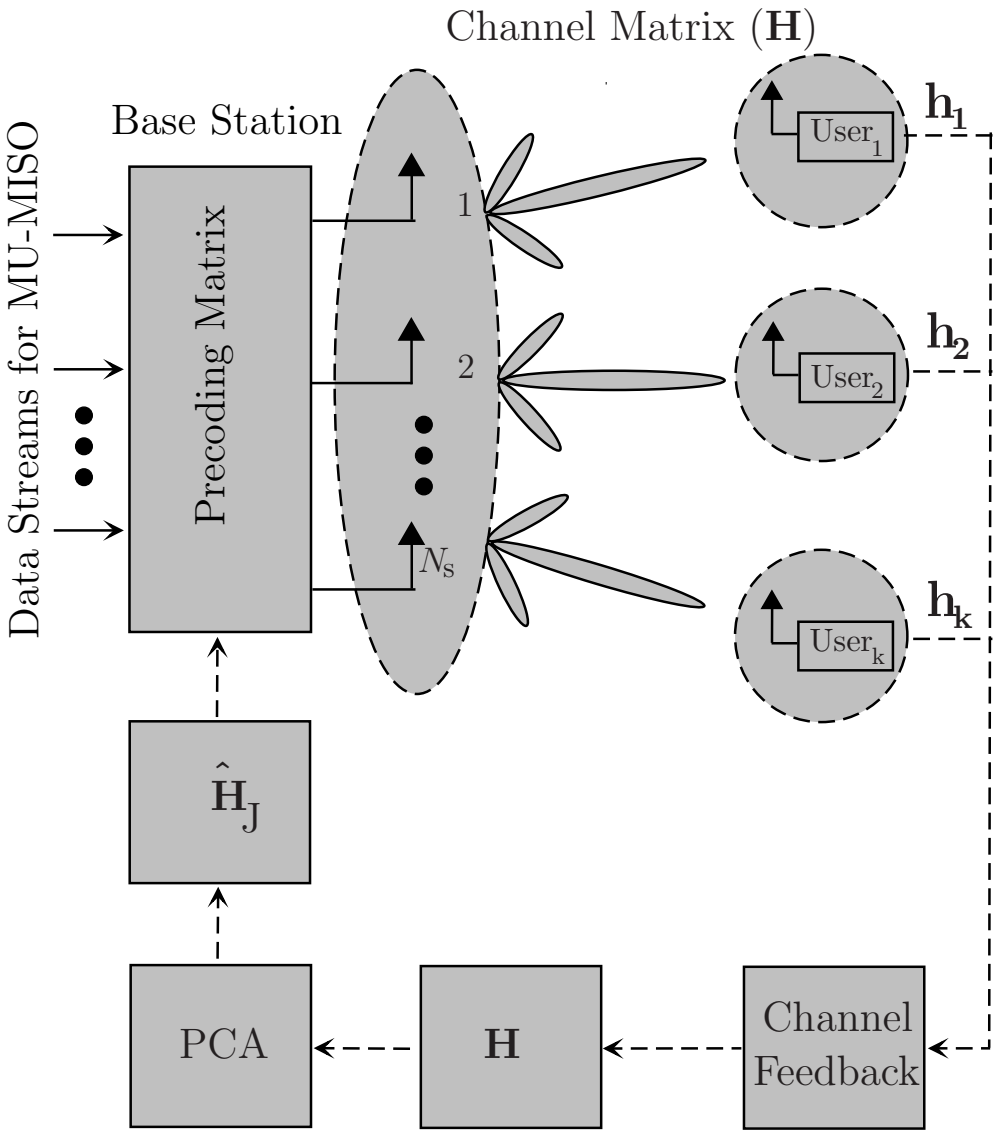


Figure 5.1: The MU-MISO System Model

$\mathbf{x} = \mathbf{Gu}$. The received signal y_k at user k is given by

$$y_k = \mathbf{h}_k \mathbf{g}_k u_k + \sum_{j=1, j\neq k}^{K} \mathbf{h}_k \mathbf{g}_j u_j + n_k. \qquad \textit{Eq. (5.2)}$$

5.2.1 Zero Forcing Precoding

In this technique precoding vectors are designed to mitigate co-channel interference (CCI) arising from other users [132]. Consider a scenario of a downlink MU-MIMO system when the BS is equipped with multiple antennas and each user terminal is equipped with a single antenna for a simple receiver. In this case the results in [18] show that if multiple users are served simultaneously then, by using the spatial-division multiple-access (SDMA) technique at the BS, the sum capacity will be enhanced. However the SDMA technique causes a problem of multiuser interference which can be mitigated by using the precoding scheme. Consider a case when the number of transmit antennas is greater than the number of users, i.e. $N_t \geq K$. Let the Moore-Penrose inverse of $\mathbf{H}_{ZF} = \mathbf{H}_{\mathcal{J}}^H(\mathbf{H}_{\mathcal{J}}\mathbf{H}_{\mathcal{J}}^H)^{-1}$. Denote $\mathbf{h}_{zfk},\ k = 1,\ldots,K$ as the k^{th} column of the Moore-Penrose inverse of $\mathbf{H}_{ZF}$. Each column of the precoding matrix is obtained by $\mathbf{g}_{zfi} = \mathbf{h}_{zfi}/||\mathbf{h}_{zfi}||$ and the complete precoding matrix is given as $\mathbf{G}_{ZF} = [\mathbf{g}_{zf1},\ldots,\mathbf{g}_{zfK}]$, where $\mathbf{g}_{zfi} \in \mathbb{C}^{N_t \times 1}(i = 1,\ldots,K)$ are the normalized values of the corresponding $\mathbf{h}_{zfi}$ column vector. The ergodic sum capacity under the assumption of equal power allocation is written as

$$\mathbf{R}_{ZF} = E_h\left\{\sum_{k=1}^{K}\log_2\left(1+\frac{SNR}{N_t}\|\mathbf{h}_k\mathbf{g}_{zfk}\|^2\right)\right\}. \qquad \textit{Eq. (5.3)}$$

5.2.2 Minimum Mean Square Error Precoding

In a point-to-point MISO system, ZF precoding enables interference-free communication for all individual users. However, the channel gains $||\mathbf{h}_k||$ are small for those channels which are in bad condition. To compensate for this bad channel condition a large power is needed. Hence the MMSE precoder introduces a regularisation fac-

tor $\beta\mathbf{I}$ in channel inversion, where β is a design parameter. This design parameter needs to be optimised to achieve a maximum sum-capacity. The optimal choice of β is $\frac{N_t}{SNR}$, which depends on the design problem [18]. Similarly to a well-known MMSE receiver, the MMSE procedure is applied at the transmitter side [133] by assuming the channel state information to be known at the BS. The MMSE precoding matrix $\mathbf{G}_{MMSE}$ can be defined as $\mathbf{G}_{MMSE} = [\mathbf{g}_{mmse1}, \ldots, \mathbf{g}_{mmseK}]$. The precoding vectors of the $\mathbf{G}_{MMSE}$ matrix can be calculated as, $\mathbf{g}_{mmsei} = \mathbf{h}_{mmsei}/\|\mathbf{h}_{mmsei}\|$, where $\mathbf{w}_{mmsei} \in \mathbb{C}^{Nt\times 1}$ $(i = 1, \ldots, K)$ represents the MMSE precoding vector of the i^{th} user. The precoding vectors of the $\mathbf{G}_{MMSE}$ matrix can be calculated as $\mathbf{H}_{MMSE} = \mathbf{H}_{\mathcal{J}}^H((\mathbf{H}_{\mathcal{J}}\mathbf{H}_{\mathcal{J}}^H) + \beta\mathbf{I})^{-1} = [\mathbf{h}_{mmse1}, \ldots, \mathbf{h}_{mmseK}]$. The ergodic sum-rate for MMSE is written as

$$\mathbf{R}_{MMSE} = E_h\left\{\sum_{k=1}^{K}\log_2\left(1+\frac{\|\mathbf{h}_k\mathbf{g}_{mmsek}\|^2}{\sum\limits_{j=1,j\neq k}^{K}\|\mathbf{h}_k\mathbf{g}_{mmsej}\|^2+\frac{N_t}{SNR}}\right)\right\}. \qquad \textit{Eq. (5.4)}$$

5.3 Channel Complexity Reduction via PCA

Initially, assume that $\mathcal{J} = \{N_1, N_2, \ldots, N_t\}$ are the indices of each column vector in $\mathbf{H}_{\mathcal{J}_{K\times N_s}}$. The objective is to reduce the complexity of a channel in such a way that it results in the least reduction in sum capacity. Suppose that we have n observations on p variables, then we have a total of $(n \times p)$ measurements. The matrix whose elements in the N_r^{th} row and N_t^{th} column are transmitted through the communications

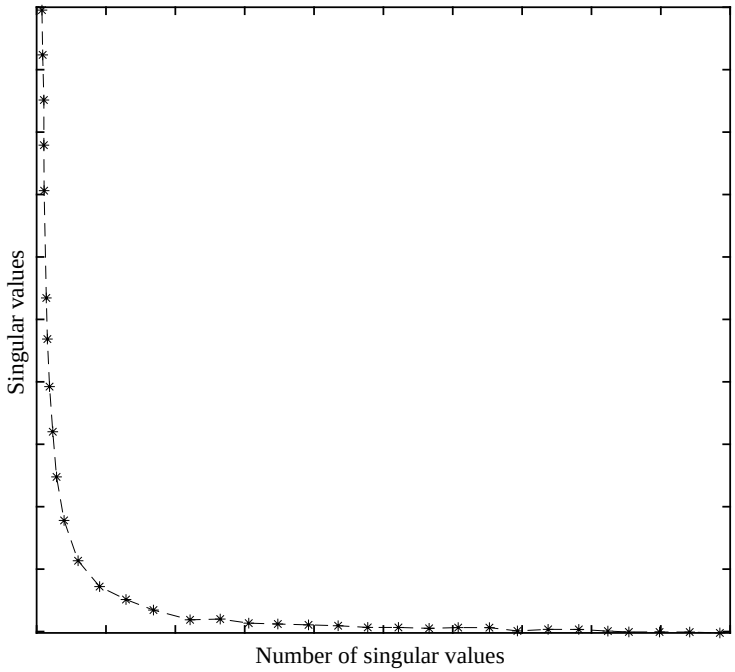


Figure 5.2: Typical decay of the ordered PCs

system to serve as inputs for a given application can be written as

$$\hat{\mathbf{H}}_{\mathcal{I}} = \overbrace{\left\{\begin{pmatrix} \hat{h}_{\mathcal{I}_{11}} & \hat{h}_{\mathcal{I}_{12}} & \dots & \hat{h}_{\mathcal{I}_{N_t}} \\ \hat{h}_{\mathcal{I}_{21}} & \hat{h}_{\mathcal{I}_{21}} & \dots & \hat{h}_{\mathcal{I}_{2N_t}} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{h}_{\mathcal{I}_{N_r1}} & \hat{h}_{\mathcal{I}_{N_r2}} \dots & \dots & \hat{h}_{\mathcal{I}_{N_rN_t}} \end{pmatrix}\right.}^{\text{Random Variables}} = \begin{pmatrix} \hat{\mathbf{w}}_1 \\ \hat{\mathbf{w}}_2 \\ \vdots \\ \hat{\mathbf{w}}_k \end{pmatrix}.$$

Users (Samples)

In order to centralise the data we take the sample mean of each column vector in $\mathcal{I}$, resulting in a row vector. This sample mean row vector is subtracted from each row in the matrix. The zero mean sample covariance matrix is obtained by using $\mathbf{S} = \frac{1}{n-1}\hat{\mathbf{w}}\hat{\mathbf{w}}^T$,. Using Eigen Value Decomposition (EVD), this problem can be

expressed in term of eigenvalues. To begin, suppose that the EVD of $\mathbf{H}_{\mathcal{J}}$ is given by

$$\mathbf{H}_{\mathcal{J}} = \mathbf{U}_{\mathcal{J}} \Lambda_{\mathcal{J}} \mathbf{V}_{\mathcal{J}}, \qquad \textit{Eq. (5.5)}$$

where $\Lambda_{\mathcal{J}} = [y_{ij}]$ is a $K \times N_t$ matrix with singular values λ_i on the diagonal (i.e. $y_{ii} = \lambda_i \; i = 1, \ldots, K$) and y_{ij} are zero elsewhere, $\mathbf{U}_{\mathcal{J}} = u_{ij}$ is a $K \times K$ unitary matrix and $\mathbf{V}_{\mathcal{J}}$ is an $N_t \times N_t$ unitary matrix. The eigenvalues are assumed to be sorted in decreasing order : $\lambda_1 \geq \ldots \geq \lambda_k$. The subset of $\Lambda_{\mathcal{J}}$ is selected which consists of those PCs that make the largest contributions to the eigenvalues, that is $\mathcal{J}_s = \{N_1, N_2, \ldots, N_s\}$ are the selected indices in $\hat{\mathbf{H}}_{\mathcal{J}_s K \times N_s}$, where $N_s << N_t$.

In PCA analysis the significance of the PCs decreases rapidly as shown in Figure 5.2. Small PCs mainly represent noise and data dependencies, therefore only a few PCs are needed for a good approximation, meaning that only a few singular values are considered, reducing the channel complexity. The reduction in the complexity achieved by using PCA can be measured in the reduction ratio (RR). It is defined as the ratio of the size of the full-dimensional channel to the size of the lower-dimensional channel. The channel complexity reduction algorithm performed at the transmitting end can be summarised by Algorithm 1.

Algorithm 1: Complexity reduction using PCA

Input : $\mathbf{H}_{K\times N_t}, N_t, K, N_s$;
Output: $\hat{\mathbf{H}}_{\mathcal{I}_s K\times N_s}$, where $N_s << N_t$;
Initialisation: Let $\mathcal{I} = \{N_1, N_2, \ldots, N_t\}$, be the indices of each column vectors in the matrix $\mathbf{H}_{K\times N_t}$;

1 **if** $(N_t > K)$ **then**
2 **for each of the column indices of** $\mathcal{I}$ **do**

a) Compute the mean of each column vector in the matrix $\mathcal{I}, \hat{\bar{h}}_k = \frac{1}{n}\sum_{j=1}^{n} \hat{h}_{jk}$, yields $\hat{\bar{\mathbf{h}}} = [\hat{\bar{h}}_1, \ldots, \hat{\bar{h}}_{N_t}]^T$;

b) Find a new origin by $\hat{h}_{jk} \leftarrow (\hat{h}_{jk} - \hat{\bar{h}}_k)$;

c) Write row vector as $\hat{\mathbf{h}} = [\hat{h}_1, \ldots, \hat{h}_{N_t}]^T$;

d) Compute eigenvalue matrix by $\mathbf{S} = \frac{1}{n-1}\hat{\mathbf{h}}\hat{\mathbf{h}}^T$;

e) Compute Eigenvalue Decomposition of $\mathbf{S} = \mathbf{U}\Lambda\mathbf{V}^{-1}$, yields eigenvalues and corresponding eigenvectors of $\mathbf{S}$

f) Sort eigenvalues of $\mathbf{S}$ such that $\Lambda = \{\lambda_k > \lambda_{k-1} >, \ldots, \lambda_1\}$

g) Sort eigenvectors in order of eigenvalues of $\mathbf{S}$ in descending order.

h) Select the largest contribution to the eigenvalues, that is $\mathcal{I}_s = \{N_1, N_2, \ldots, N_s\}$.

3 **end**
4 **Result** $\mathcal{I}_s = \{N_1, N_2, \ldots, N_s\}$ are the selected indices of $\hat{\mathbf{H}}_{\mathcal{I}_{K\times N_s}}$, where $N_s << N_t$. Now calculate the sum rate with ZF by using $\hat{\mathbf{H}}_{\mathcal{I}_s}$.
5 **end**

5.4 Computational Complexity

To design an efficient, low-complexity algorithm in signal-processing tasks, a detailed analysis of the number of required floating-point operations (FLOPs) is often inevitable. In a real arithmetic case, multiplication followed by an addition needs 2 FLOPs. However, with complex-valued quantities, 8 FLOPs are needed when a multiplication is followed by an addition. Therefore, the complexity in complex matrix

multiplication is 4 times that of its real counterpart. The computational complexity of conventional ZF, MMSE and the proposed PCA algorithms are analysed in terms of FLOPs. The required numbers of FLOPs of these techniques are simulated under different system dimensions, and numerical values are shown later in Table 5.1. The required number of FLOPs for ZF and MMSE is equal to $16n^3 + 3n^2 - 2n$ and $16n^3 + 3n^2$ respectively [134, 135], where n represents the system dimension, which is defined as the number of rows and columns in a matrix. The computational complexity of PCA – the complexity reduction technique– is determined by the number of datapoints/PCs D, and the system dimension n is given by $(D + n^2)n$ [136]. In addition, in case of ZF and MMSE, full dimensions of n is used and in a proposed technique of PCA half number of dimensions are utilized to produce the results.

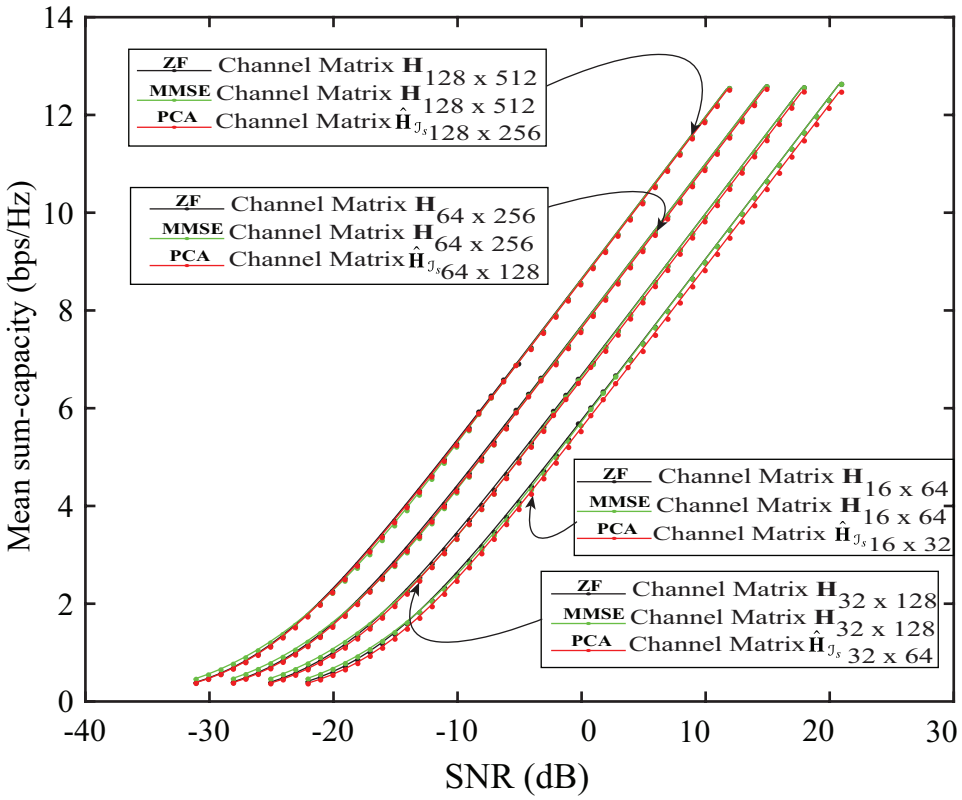


Figure 5.3: Sum rate simulations using PCA

5.5 Simulation Results and Discussion

For the simulation results, we use a Rayleigh-fading channel whose characteristics may be found in [137]. The simulations are implemented in MATLAB by using Monte-Carlo method, where 5000 realisation of channel matrix are taken to demonstrate the throughput performance of the proposed algorithm with multiple settings of the system dimensions. Four different scenarios are discussed, as shown in Figure 5.3. A system with a large number of datasets/variables is considered and the channel complexity is reduced using PCA. A reduction ratio of 0.5 in each case is achieved by using $RR = \frac{N_s}{N_t}$. For instance, consider a case when $K = 16$, $N_t = 64$ and $N_s = 32$. The conventional precoding techniques (ZF and MMSE) use all the available data sets in channel $\mathfrak{I} = \{N_1, N_2, \ldots, N_t\}$ to compute the sum rate. But, in our proposed technique, out of all the available data sets we only use $\mathfrak{I}_s = \{N_1, N_2, \ldots, N_s\}$ inputs, meaning that instead of using the 64 available variables only 32 are used to calculate the throughput of the system in this particular case. Therefore, with a negligible loss of sum capacity (Figure 5.3) the channel complexity is reduced significantly as shown in Figure 5.4. This figure shows that the proposed PCA algorithm involves a lower channel complexity than the conventional ZF and MMSE algorithms. It can be noticed that, with the increase of the system dimensions (n), the channel complexity reduced significantly. Similarly, in the other three cases of system dimensions illustrated in Figure 5.3, half of the available random variables are used to calculate the sum rate for multiple users, and the analysis of the required number of FLOPs is shown in Figure 5.4. In a nutshell, by using half of the available variables in a data set, PCA has approximately the same throughput but with much less computational complexity than the conventional techniques.

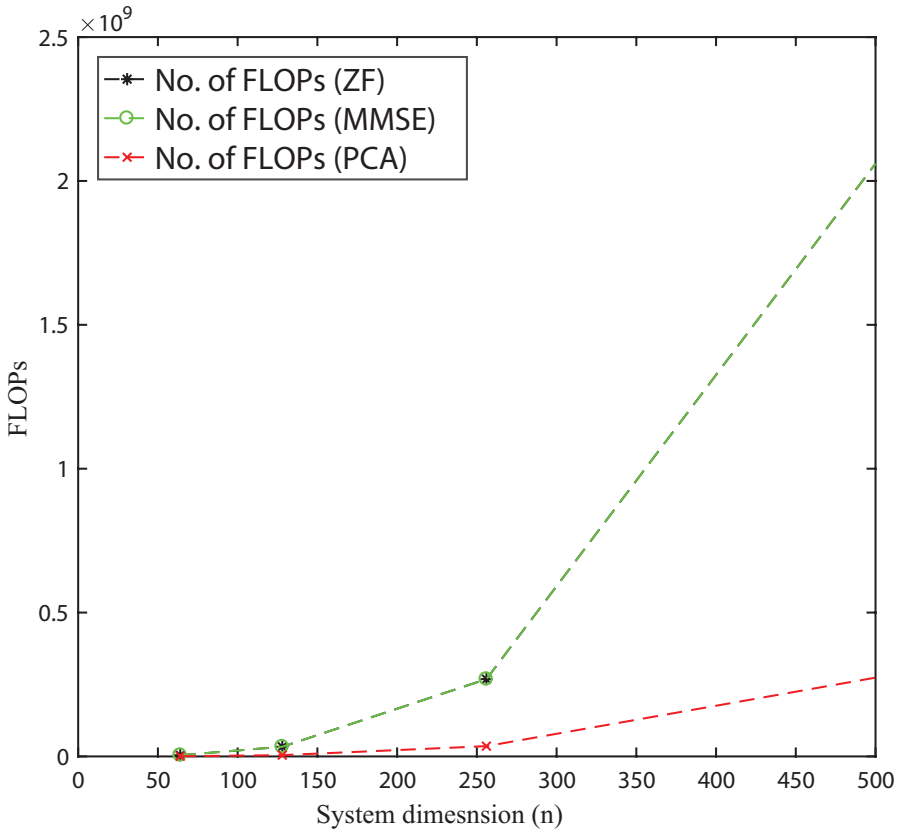


Figure 5.4: Computational complexity in FLOPs for massive MU-MISO system

The numerical values of computational complexity in terms of FLOPs for different system settings are shown in Table 5.1. It is clear that ZF and MMSE show high complexity in different system dimensions, and the computational complexity of PCA is much less by (86%) than with the conventional techniques. The numerical values of the system capacity at an SNR of 10 dB show a negligible loss in throughput as compared to the traditional techniques. The computational time of the channel using conventional techniques and PCA are also analysed and shown in Table 5.2. Conventional techniques utilise all the available variables in a data set ($\mathcal{I} = \{1,2,\ldots,N_t\}$) to compute the sum rate, with a computational time depicted in Table 5.2. However, PCA uses $\mathcal{I}_s = \{1,2,\ldots,N_s\}$ as the data set of a channel, to

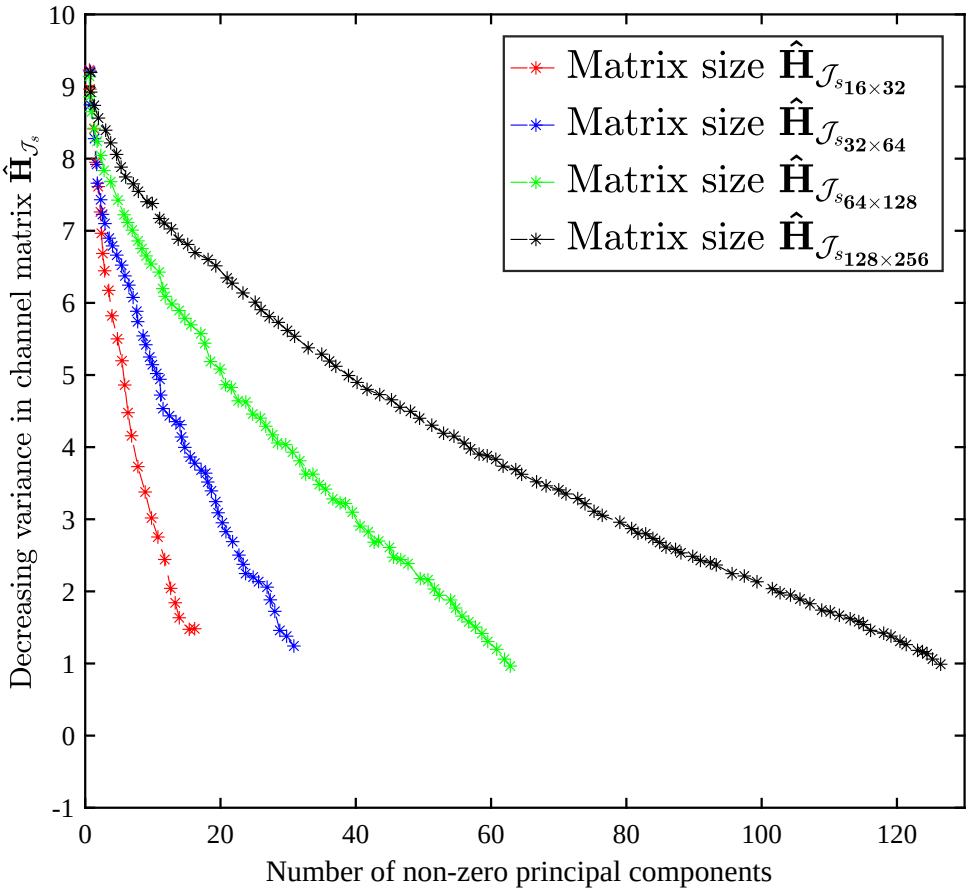


Figure 5.5: Decreasing-variance trend of PCA-based channel matrix $\hat{\mathbf{H}}_{\mathcal{J}_s}$

calculate the throughput of the system with a computational time which is close to that of the conventional techniques.

The decreasing variance, that is the typical decay of the ordered PCs of a reduced PCA-based channel matrix, is shown in Figure 5.5. It indicates the use of the number of non-zero PCs which are required to reduce the channel complexity efficiently. In Figure 5.6 the exponential decay of the largest eigenvalues is plotted. It follows a conventional decay of PCs as shown in Figure 5.2. The exponential decay can be interpreted as determining which linear combination of the original random variables explains most of the variability in the data set.

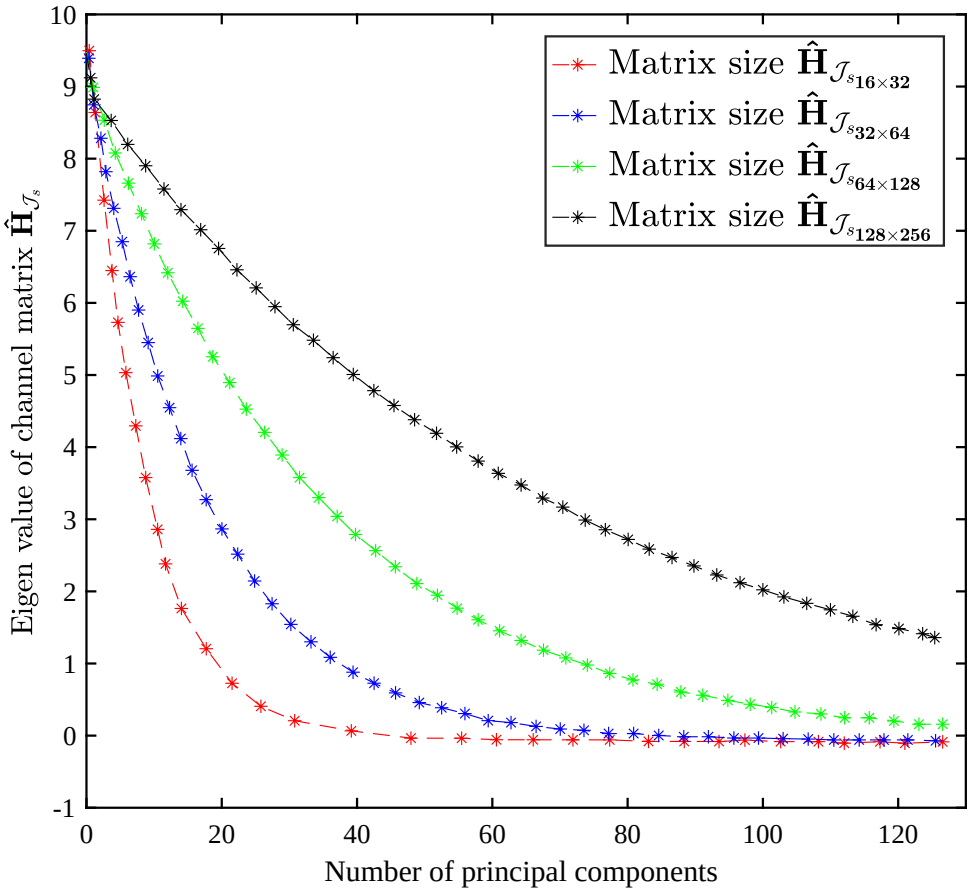


Figure 5.6: Exponential decay of the largest PCs

The plots in Figures 5.7 and 5.8 were simulated using the Kronecker-based channel model as shown in (4.2). Once the channel matrices were simulated, the linear precoding techniques such as ZF and MMSE were implemented to examine the system throughput. These techniques uses the full matrix dimensions, i.e. $\mathcal{J} = \{1,2,\ldots,N_t\}$ are utilised to compute the sum-rate. However, by using the analysis of PCA, out of all the available data sets we only use $\mathcal{J}_s = \{1,2,\ldots,N_s\}$. Therefore, it reduces the computational complexity of the system in a correlated environment.

5.6 Summary of the Chapter

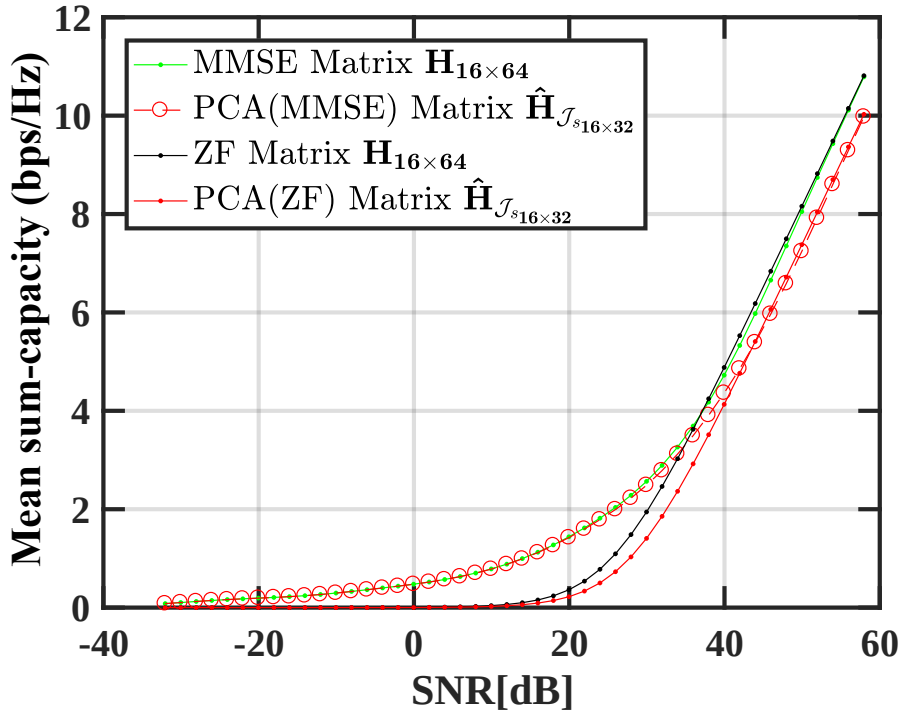
In this chapter, a novel PCA-based channel complexity reduction scheme for downlink massive MISO systems is presented. The results show that a channel using PCA has approximately the same throughput as one using the ZF and MMSE algorithms, but the proposed method has much reduced complexity both in time and by (86%) in FLOPS. A reduction ratio of 0.5 is achieved in each case with a negligible loss in sum capacity.

Table 5.1: Computational Complexity in FLOPs for multiple system dimensions

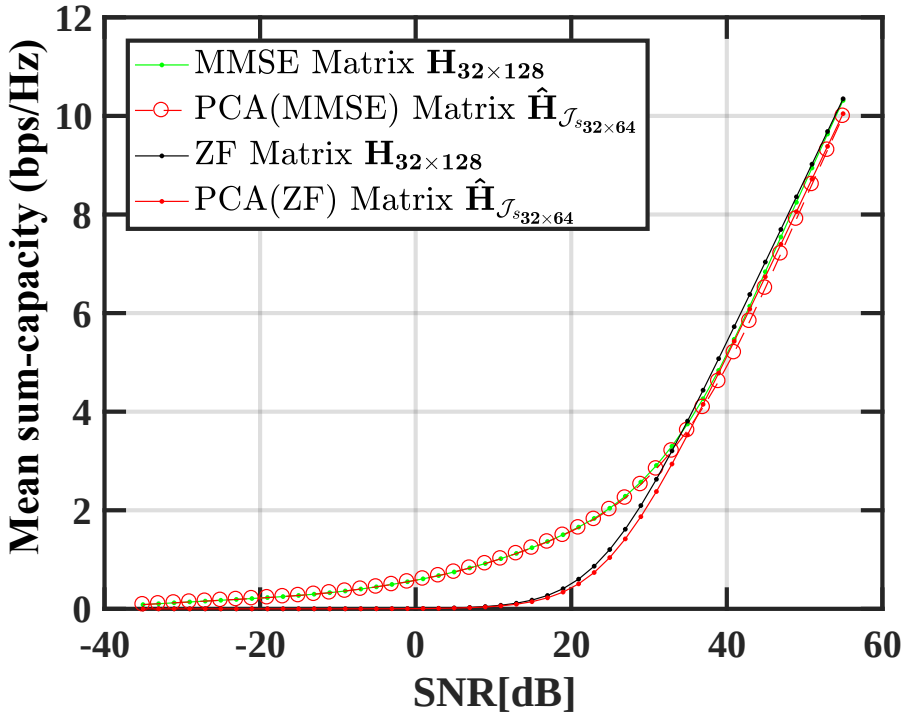
Matrix Size	FLOPs (ZF)	Capacity (bps/Hz) (SNR = 10 dB)	FLOPs (MMSE)	Capacity (bps/Hz) (SNR = 10 dB)	Matrix Size	FLOPs (PCA)	Capacity (bps/Hz) (SNR = 10 dB)	%age reduction in FLOPs
$\mathbf{H}_{16\times64}$	4.21×10^6	9.95	4.21×10^6	9.17	$\hat{\mathbf{H}}_{\mathcal{I}_{s16\times32}}$	5.61×10^5	9.00	86.66
$\mathbf{H}_{32\times128}$	3.36×10^7	10.17	3.36×10^7	10.16	$\hat{\mathbf{H}}_{\mathcal{I}_{s32\times64}}$	4.47×10^6	9.71	86.69
$\mathbf{H}_{64\times256}$	2.69×10^8	10.87	2.69×10^8	10.87	$\hat{\mathbf{H}}_{\mathcal{I}_{s64\times128}}$	3.57×10^7	10.81	86.70
$\mathbf{H}_{128\times512}$	2.15×10^9	11.88	2.15×10^9	11.88	$\hat{\mathbf{H}}_{\mathcal{I}_{s128\times256}}$	2.85×10^8	11.86	86.71
ZF $\{16n^3+3n^2-2n\}$ [134]			MMSE $\{16n^3+3n^2\}$ [135]			PCA $\{(D+n^2)n\}$ [136]		

Table 5.2: Computational time in seconds for multiple system dimensions

Matrix Dimension	ZF	MMSE	Matrix Dimension	PCA
$\mathbf{H}_{16\times64}$	3.89×10^{-4}	4.12×10^{-4}	$\hat{\mathbf{H}}_{\mathcal{I}_{s16\times32}}$	3.63×10^{-4}
$\mathbf{H}_{32\times128}$	1.5×10^{-4}	1.5×10^{-4}	$\hat{\mathbf{H}}_{\mathcal{I}_{s32\times64}}$	1.4×10^{-4}
$\mathbf{H}_{64\times256}$	4.90×10^{-2}	5.5×10^{-2}	$\hat{\mathbf{H}}_{\mathcal{I}_{s64\times128}}$	4.9×10^{-2}
$\mathbf{H}_{128\times512}$	2.16×10^{-2}	2.16×10^{-2}	$\hat{\mathbf{H}}_{\mathcal{I}_{s128\times256}}$	2.13×10^{-2}

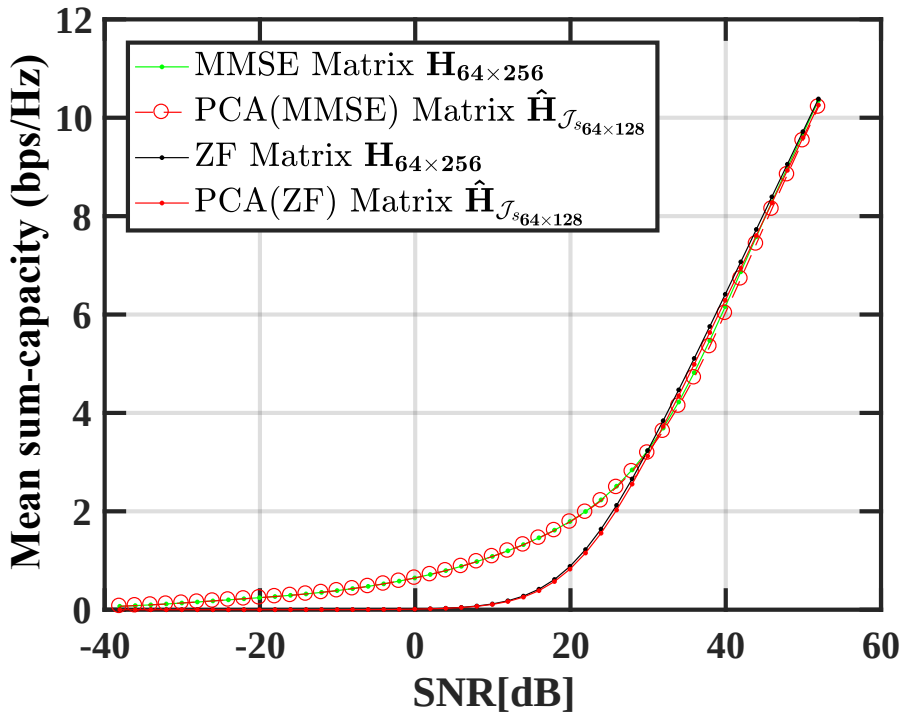


(a)

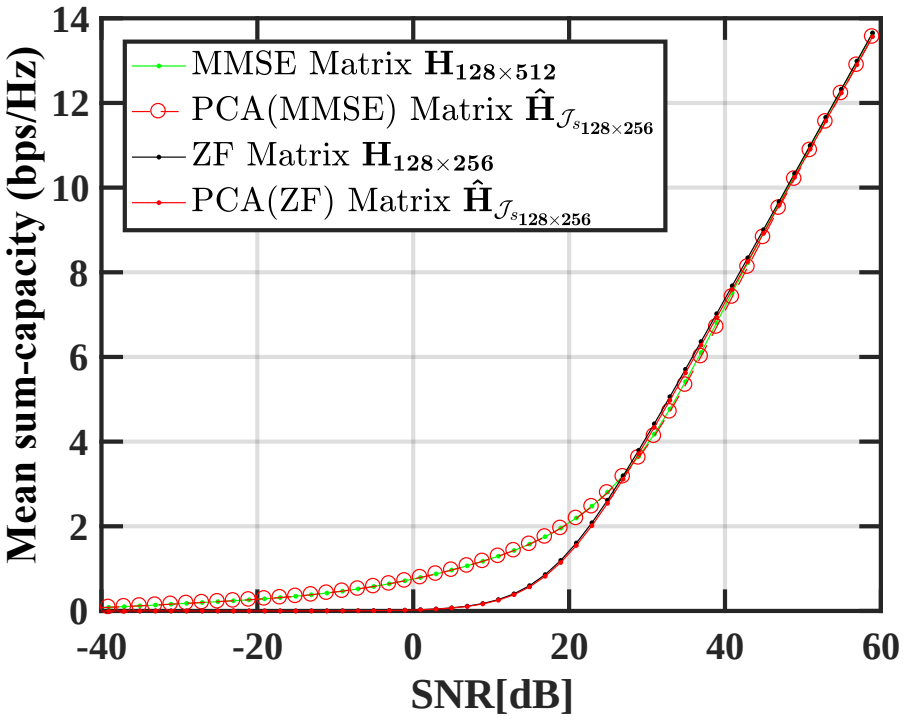


(b)

Figure 5.7: (a) Sum-capacity simulations using PCA with 16×32 matrix dimension (b) Sum-capacity simulations using PCA with 32×64 matrix dimension



(a)



(b)

Figure 5.8: (a) Sum-capacity simulations using PCA with 64×128 matrix dimension (b) Sum-capacity simulations using PCA with 128×256 matrix dimension

Do not judge me by my successes, judge me by how many times I fell down and got back up again.

Nelson Mandela (1918-2013)

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

This chapter gives a brief summary of the thesis in Section 6.1 and then discusses future research work based on the results of this dissertation in Section 6.2.

This thesis explored the problem of antenna selection, signal modelling in the presence of correlation and dimensionality reduction of large matrices in massive MIMO broadcast wireless channels. The results in this work were presented mainly in four parts: a) novel antenna-selection algorithms, b) application of novel antenna selection algorithms to spatially correlated channels, c) large matrix computations were reduced using a dimensionality-reduction technique, d) analysis of matrix dimension reduction for spatially correlated channel models.

Antenna selection is a promising low-complexity solution that solves the pressing problem of the increased hardware and signal-processing complexity of MIMO systems. Optimal antenna selection in massive MIMO is required to minimise hardware complexity, i.e. an efficient use of radio-frequency units. Current antenna-selection algorithms, were investigated and, using PCA, two semi-heuristic antenna-selection algorithms namely non-central PCA and LDAS, were proposed. Using analytic methods PCA eigenvalues were decomposed into two components: the mean-channel-gain component and the channel correlation component. Using simulation we show that the proposed antenna-selection methods perform much better using mean-channel-gain selection and show how antenna selection depends on the channel matrix structure. It was shown that these algorithms perform reasonably well as compared to other antenna-selection algorithms, while having less hardware complexity. The simulation results show how the mean sum-capacity varied with different channel conditions, and which antenna selection schemes were preferred for those conditions. The performance of these algorithms was analysed with linear precoding at the base station, and simulated results showed that the proposed algorithms reduced the required number of radio-frequency units without reducing their effectiveness.

Signal modelling in a MIMO system architecture using different channel models was employed. The most common analytical channel models used were the Kronecker and the Weichselberger models. The Kronecker channel model approximates the correlation matrix at both ends of the link using Power Azimuth Spectrum (PAS) models. The correlation matrices in this model were assumed to be separable at both the transmitter and receiver. In contrast, a slightly more complex model, known as the Weichselberger channel model, was also discussed. This model uses a spatial structure to determine the Direction of Departure (DOD) and Direction of Arrival (DOA) of the receive and transmit correlation matrices. These correlation matrices were approximated at receiver and transmitter by using PAS models, and these matrices overcome the assumptions of the Kronecker model. However, this model requires the additional knowledge of the power coupling matrix. Once the MIMO channel matrix was modelled in different realistic environments, the proposed antenna-selection algorithms were implemented and the system throughput was analysed. It was clear that the system capacity was significantly higher in ideal scenarios i.e. using a Rayleigh-fading channel model. However, the performance of the proposed algorithms in approximated real channel matrices was optimal and less complex than the exhaustive-search antenna-selection algorithms.

To reduce the matrix size, the analysis techniques of PCA were used. The throughput of the system with full dimensions and of a system with optimal dimensions were evaluated in an uncorrelated channel model. The FLOP technique was then implemented to calculate the numerical value of the number of operations involved in computations. It was shown by using the FLOP method that the number of computations while sacrificing negligible sum capacity. Linear precoding was used at the transmitter side and the performance of PCA was compared with conventional

full-dimensional systems. The numerical results show that a PCA dimensionality-reduction technique at 10 dB SNR has approximately the same throughput as one using the ZF and MMSE algorithms. However, ZF and MMSE algorithms used a matrix dimension as $\mathbf{H}_{16\times 64}$ and the proposed algorithm used a matrix dimension as $\hat{\mathbf{H}}_{\mathcal{J}_{s16\times 32}}$, to obtain the throughput at 10 dB. Therefore, the proposed method has much reduced complexity in both time and by 86% in FLOPS. A reduction ratio of 0.5 was achieved in each case with a negligible loss. In addition, the signal correlations were modelled using spatially correlated channel models, and the correlation matrices were approximated using PAS models. The sum rates were evaluated in different realistic environments in a full dimensional system and a system with the optimal dimension. The FLOPs technique was then implemented to compute the system dimensions in both scenarios. It was shown that there was much less computational complexity of the system with optimal dimensions. However, this comes with a negligible throughput loss in the sum rate. For instance, the numerical value of the throughput of a system at an SNR of 10 dB was shown to approach the sum capacity of a full-dimensional system.

6.2 Future Directions

The results presented in this thesis lead to many interesting open questions related to MIMO broadcast wireless channels. Some suggested directions for future research are as follows.

Firstly, a possible extension of the work presented in this dissertation is on time-varying wireless channels. Most of the current antenna-selection algorithms assume a quasi-stationary wireless channel. The interesting question to ask is how antenna-

selection algorithms behave when the channel characteristics vary with time. In this work, antenna-selection algorithms use perfect channel-state information available at the transmitter for the selection process. However, in time-varying channels this channel-state information becomes quickly outdated. Also there are delays involved in sending this channel information back to the base station. Therefore, it is interesting to investigate the deficiencies in sum-capacity due to imperfect and outdated channel-state information at the base station.

Secondly, the analytical channel models for MIMO systems are developed. For instance, the Rayleigh model (i.i.d.), spatially correlated channel models (the Kronecker model, the Weichselberger model) are explored and the system performance is analysed. However, the temporal correlation of channel vectors is not examined. The question to ask is how antenna-selection algorithms behave in the presence of temporal-correlation. Thus, to evaluate the shortcomings in sum-capacity to analyse the antenna-selection algorithm in time-varying channels with imperfect channel-state information at the base station should be addressed.

REFERENCES

[1] Qualcom, “The 100 x data challenge,,” January 2017. [Online]. Available: https://www.qualcomm.com/solutions/automotive/drive-data-platform

[2] Ericson, “5g radio access - research and vision,” June 2013. [Online]. Available: https://www.ericsson.com/assets/local/publications/white-papers/wp-5g.pdf

[3] Cisco, “Cisco visual networking index: Global mobile data traffic forcast update, 2016-2021,” June 2017. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf

[4] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[5] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[6] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Transactions on ecommunications*, vol. 10, pp. 585–595, 1999.

[7] J. H. Winters, “On the capacity of radio communications systems with diversity in rayleigh fading environments,” *IEEE Journal of Selected Areas in Communications*, vol. 5, no. 5, pp. 871–878, June 1987.

[8] J. Lee, Y. Kim, Y. Kwak, J. Zhang, A. Papasakellariou, T. Novlan, C. Sun, and Y. Li, "LTE-advanced in 3gpp Rel -13/14: an evolution toward 5g," *IEEE Communications Magazine*, vol. 54, no. 3, pp. 36–42, Mar. 2016.

[9] L. Liu, R. Chen, S. Geirhofer, K. Sayana, Z. Shi, and Y. Zhou, "Downlink MIMO in LTE-advanced: SU-MIMO vs. MU-MIMO," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 140–147, February 2012.

[10] P. H. Kuo, "New physical layer features of 3gpp LTE release-13 [Industry Perspectives]," *IEEE Wireless Communications*, vol. 22, no. 4, pp. 4–5, Aug. 2015.

[11] G.Caire and S. S. Shlomo, "On achievable rates in a multi-antenna broadcast downlink," 2000.

[12] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, May 2003.

[13] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, June 2003.

[14] G. Caire and S. Shamai, "On the achievable trhoughput of a multiantenna gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, June 2003.

[15] M. Costa, "Writing on dirty paper (Corresp.)," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[16] Q. Spencer, C. Peel, A. Swindlehurst, and M. Haardt, "An introduction to multiuser mimo downlink," *IEEE Communications Magazine*, vol. 42, pp. 60–67, October 2004.

[17] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2013.

[18] F. Kaltenberger, M. Kountouris, L. Cardoso, R. Knopp, and D. Gesbert, "Capacity of linear multi-user MIMO precoding schemes with measured channel data," in *2008 IEEE 9th Workshop on Signal Processing Advances in Wireless Communications*, July 2008, pp. 580–584.

[19] Y. Gao, H. Vinck, and T. Kaiser, "Massive mimo antenna selection: switching architectures, capacity bounds, and optimal antenna selection algorithms," *IEEE Transactions On Signal Processing*, vol. 66, no. 5, pp. 1346–1360, March 2018.

[20] T. L. Marzeeta, "Noncooperative cellular wireless with unlimited number of base station antennas," *IEEE Transactions on Wireless Communications,*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[21] F. Rusek, D. Persson, L. B. K, E. G. Larsson, T. L. Marzeeta, O. Edfors, and F. Tufvesson, "Scalling up mimo: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, January 2013.

[22] V. Venkateswaran and A. J. van der Veen, "Analog Beamforming in MIMO Communications With Phase Shift Networks and Online Channel Estimation,"

IEEE Transactions on Signal Processing, vol. 58, no. 8, pp. 4131–4143, Aug. 2010.

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

[24] A. F. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Microwave Magazine*, vol. 5, no. 1, pp. 46–56, March 2004.

[25] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 68–73, October 2004.

[26] M. L. B, J. Choi, J. Bang, and B.-C. Kang, "An energy efficient antenna selection for large scale green mimo systems," in *IEEE international symposium Circuits and System (ISCAS)*. IEEE, May 2013, pp. 950–953.

[27] X. Gao, O. Edfors, J. liu, and F. Tufvesson, "Antenna selection in measured massive mimo channels using convex optimisation," *IEEE GLOBECOM Workshop on Emerging Technologies for LTE Advanced and Beyond-4G*, pp. 129–134, 2013.

[28] X. Cheng and Y. He, "Geometrical model for massive mimo systems," *IEEE Vehicular Technology Conference*, pp. 1–6, 2017.

[29] K. Liu, V. Raghavan, and A. M. Sayeed, "Capacity scaling and spectral efficiency in wide-band correlated MIMO channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2504–2526, October 2003.

[30] Q. U. A. Nadeem, A. Kammoun, M. Debbah, and M. S. Alouini, "A Generalized Spatial Correlation Model for 3d MIMO Channels Based on the Fourier

Coefficients of Power Spectrums," *IEEE Transactions on Signal Processing*, vol. 63, no. 14, pp. 3671–3686, July 2015.

[31] M. T. A. Rana, R. Vesilo, and I. B. Collings, "Antenna selection in massive MIMO using non-central Principal Component Analysis," in *2016 26th International Telecommunication Networks and Applications Conference (ITNAC)*, December 2016, pp. 283–288.

[32] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*, 1st ed. Singapore ; Hoboken, NJ: Wiley-IEEE Press, November 2010.

[33] I. T. Jolliffe, *Principal component analysis*. Springer-Verlang, 1986.

[34] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. D. Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5g: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, June 2017.

[35] F.-L. Luo and C. Zhang, *Signal Processing For 5G: Algorithms And Implementations*. IEEE Press, 2016.

[36] T. L. Marzetta, "Massive MIMO: An Introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015.

[37] S. Wu, C. X. Wang, Y. Yang, W. Wang, and X. Gao, "Performance comparison of massive MIMO channel models," in *2016 IEEE/CIC International Conference on Communications in China (ICCC)*, July 2016, pp. 1–6.

[38] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the Number of RF Chains and Phase Shifters, and Scheduling Design With Hybrid

AnalogDigital Beamforming," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3311–3326, May 2016.

[39] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Massive MIMO in Real Propagation Environments: Do All Antennas Contribute Equally?" *IEEE Transactions on Communications*, vol. 63, no. 11, pp. 3917–3928, November 2015.

[40] D. A. Gore and A. J. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2580–2588, October 2002.

[41] R. W. Heath, S. Sandhu, and A. J. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Communications Letters*, vol. 5, no. 4, pp. 142–144, April 2001.

[42] S. Sanayei and A. Nosratinia, "Asymptotic capacity analysis of transmit antenna selection," in *International Symposium on Information Theory, 2004. ISIT 2004. Proceedings.*, June 2004, pp. 241–.

[43] M. A. Jensen and M. L. Morris, "Efficient capacity-based antenna selection for MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 1, pp. 110–116, January 2005.

[44] R. Narasimhan, "Spatial multiplexing with transmit antenna and constellation selection for correlated MIMO fading channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2829–2838, November 2003.

[45] ——, "Transmit antenna selection based on outage probability for correlated MIMO multiple access channels," *IEEE Transactions on Wireless Communications*, vol. 5, no. 10, pp. 2945–2955, October 2006.

[46] M. A. Khan, R. Vesilo, L. M. Davis, and I. B. Collings, "User and Transmit Antenna Selection for MIMO Broadcast Wireless Channels with Linear Receivers," in *2008 Australasian Telecommunication Networks and Applications Conference*, December 2008, pp. 276–281.

[47] Y. Ni, W. Zhang, and M. Chen, "Antenna subset selection in MU large-scale MIMO systems," in *2013 International Conference on Wireless Communications and Signal Processing*, October 2013, pp. 1–5.

[48] S. Mahboob, R. Ruby, and V. C. M. Leung, "Transmit Antenna Selection for Downlink Transmission in a Massively Distributed Antenna System Using Convex Optimization," in *2012 Seventh International Conference on Broadband, Wireless Computing, Communication and Applications*, November 2012, pp. 228–233.

[49] Z. Shi, X. Zhu, Y. Zhao, and L. Huang, "Interference alignment based on antenna selection for massive MIMO system," in *2015 10th International Conference on Computer Science Education (ICCSE)*, July 2015, pp. 606–610.

[50] M. Hanif, H. C. Yang, G. Boudreau, E. Sich, and H. Seyedmehdi, "Low Complexity Antenna Subset Selection for Massive MIMO Systems with Multi-Cell Cooperation," in *2015 IEEE Globecom Workshops (GC Wkshps)*, December 2015, pp. 1–5.

[51] J. Wang, A. I. Prez-Neira, and M. Gao, "A concise joint transmit/receive antenna selection algorithm," *China Communications*, vol. 10, no. 3, pp. 91–99, March 2013.

[52] Y. Jiang, M. K. Varanasi, and J. Li, "Performance Analysis of ZF and MMSE Equalizers for MIMO Systems: An In-Depth Study of the High SNR Regime," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2008–2026, April 2011.

[53] A. F. Molisch, "Effect of far scatterer clusters in mimo outdoor channel models," 57^{th} *IEEE Vehicular Technology Conference*, vol. 1, pp. 532–538, 2003.

[54] M. Ozcelik, N. Czink, and E. Bonek, "What makes a good MIMO channel model?" in *2005 IEEE 61st Vehicular Technology Conference*, vol. 1, May 2005, pp. 156–160.

[55] W. Weichselberger, "Spatial structure of multiple antenna radio channels: a signal processing viewpoint," Ph.D. dissertation, Technische Universitt Wien, Austria, December 2003.

[56] H. Ozcelik, "Indoor mimo channel models," Ph.D. dissertation, Institutfr Nachrichtentechnik, Technische Universitt Wien, Vienna, Austria, December 2004.

[57] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti, H. Hofstetter, P. Kysti, D. Laurenson, G. Mat, A. F. Molisch, C. Oestges, and H. zcelik, "Survey of Channel and Radio Propagation Models for Wireless MIMO Systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 1, December 2007.

[58] A. M. Rao and D. L. Jones, "Efficient detection with arrays in the presence of angular spreading," *IEEE Transactions on Signal Processing*, vol. 51, pp. 301–302, February 2003.

[59] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell Systems Technical Journal*, vol. 47, pp. 957–1000, July-August 1968.

[60] W. C. Jakes, *Microwave Mobile Communication*. New York, USA: John Wiley & Sons, 1974.

[61] W. C. Y. Lee, "Effect on correlation between two mobile radio base-station antennas," *IEEE Transactions on Communications*, vol. COM-21, no. 11, pp. 1214–1224, November 1973.

[62] L. M. Correia, *Wireless flexible personalised communications : COST 259, European co-operation in mobile radio research*. Chichester ; New York : Wiley, 2001.

[63] J. Ling, U. Tureli, D. Chizhik, and C. Papadias, "Rician Modeling and Prediction for Wireless Packet Data Systems," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4692–4699, November 2008.

[64] V. Erceg, L. Schumacher, P. Kyritsi, A. Molisch, D. S. Baum, A. Y. Gorokhov, C. Oestges, Q. Li, K. Yu, N. Tal, B. Dijkstra, A. Jagannatham, C. Lanzl, V. J. Rhodes, J. Medbo, D. Michelson, M. Webster, E. Jacobsen, D. Cheung, C. Prettie, M. Ho, S. Howard, B. Bjerke, L. Jengx, H. Sampath, S. Catreux, S. Valle, A. Poloni, A. Forenza, and T. R. R. W. Heath, "Tgn channel models, technical report," IEEE P802.11 Wireless LANs, Tech. Rep., May 2004.

[65] J. Medbo and P. Schramm, *Channel models for HIPERLAN/2*, ETSI/BRAN Document NO. 3ERI085B, 1998.

[66] C. Oestges, V. Erceg, and A. J. Paulraj, "A physical scattering model for MIMO macrocellular broadband wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 721–729, June 2003.

[67] V. Erceg, K. V. S. Hari, M. S. Smith, D. Baum, P. Soma, L. J. Greenstein, D. G. Michelson, S. Ghassemzadeh, A. J. Rustako, R. S. Roman, K. P. Sheikh, C. Tappenden, J. M. Costa, C. Bushue, A. Sarajedini, R. Schwartz, D. Branlund, T. Kaitz, and D. Trinkwon, "Channel Models for Fixed Wireless Application," July 2001.

[68] 3GPP, "3gpp tr 25.996: Spatial channel model for multiple input multiple output (mimo) simulations (release 6)," 3rd Generation Partnership Project (3GPP), Technical Report, 2003.

[69] R. Janaswamy, "Angle and time of arrival statistics for the Gaussian scatter density model," *IEEE Transactions on Wireless Communications*, vol. 1, no. 3, pp. 488–497, July 2002.

[70] R. B. Ertel and J. H. Reed, "Angle and time of arrival statistics for circular and elliptical scattering models," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 11, pp. 1829–1840, November 1999.

[71] J. C. Liberti and T. S. Rappaport, "A geometrically based model for line-of-sight multipath radio channels," in *Proceedings of Vehicular Technology Conference - VTC*, vol. 2, April 1996, pp. 844–848.

[72] J. Fuhl, A. F. Molisch, and E. Bonek, "Unified channel model for mobile radio systems with smart antennas," *Sonar and Navigation IEE Proceedings - Radar*, vol. 145, no. 1, pp. 32–41, February 1998.

[73] P. Petrus, J. H. Reed, and T. S. Rappaport, "Geometrical-based statistical macrocell channel model for mobile environments," *IEEE Transactions on Communications*, vol. 50, no. 3, pp. 495–502, March 2002.

[74] M. P. Lotter and R. V. Rooyen, "Modeling spatial aspects of cellular CDMA/SDMA systems," *IEEE Communications Letters*, vol. 3, no. 5, pp. 128–131, May 1999.

[75] R. J. Piechocki, G. V. Tsoulos, and J. P. McGeehan, "Simple general formula for PDF of angle of arrival in large cell operational environments," *Electronics Letters*, vol. 34, no. 18, pp. 1784–1785, September 1998.

[76] D. D. N. Bevan, V. T. Ermolayev, A. G. Flaksman, and I. M. Averin, "Gaussian channel model for mobile multipath environment," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, no. 9, 2004. [Online]. Available: http://link.springer.com/article/10.1155/S1110865704404028

[77] K. Yu and B. Ottersten, "Models for mimo propagation channels: A review," *Wiley Journal of Wireless Communications and Mobile Computing*, vol. 2, no. 7, pp. 653–666, 2002.

[78] D. Tse and P. P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, May 2005.

[79] K. I. Pedersen, P. E. Mogensen, and B. H. Fleury, "Spatial channel characteristics in outdoor environments and their impact on BS antenna system perfor-

mance," in *48th IEEE Vehicular Technology Conference, 1998. VTC 98*, vol. 2, May 1998, pp. 719–723 vol.2.

[80] S. Venkatesan, S. H. Simon, and R. A. Valenzuela, "Capacity of a Gaussian MIMO channel with non-zero mean," in *2003 IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No.03CH37484)*, vol. 3, October 2003, pp. 1767–1771.

[81] F. Bashar and T. D. Abhayapala, "Performance analysis of spatially distributed MIMO systems," *IET Communications*, vol. 11, no. 4, pp. 566–575, 2017.

[82] E. Zchmann, M. Lerch, S. Caban, R. Langwieser, C. F. Mecklenbrauker, and M. Rupp, "Directional evaluation of receive power, Rician K-factor and RMS delay spread obtained from power measurements of 60 GHz indoor channels," in *2016 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, September 2016, pp. 246–249.

[83] L. Wood and W. S. Hodgkiss, "Impact of Channel Models on Adaptive M-QAM Modulation for MIMO Systems," in *2008 IEEE Wireless Communications and Networking Conference*, March 2008, pp. 1316–1321.

[84] G. D. Durgin and T. S. Rappaport, "Effects of multipath angular spread on the spatial cross-correlation of received voltage envelopes," in *1999 IEEE 49th Vehicular Technology Conference (Cat. No.99CH36363)*, vol. 2, July 1999, pp. 996–1000.

[85] D. Chizhik, F. Rashid-Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels," *IEEE Communications Letters*, vol. 4, no. 11, pp. 337–339, November 2000.

[86] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, March 2000.

[87] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, August 2002.

[88] A. A. Abouda, H. El-Sallabi, L. Vuokko, and S. G. Haggman, "Performance of Stochastic Kronecker MIMO Radio Channel Model in Urban Microcells," in *2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, September 2006, pp. 1–5.

[89] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 90–100, January 2006.

[90] A. M. Tulino, A. Lozano, and S. Verdu, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Transactions on Wireless Communications*, vol. 5, no. 3, pp. 662–671, March 2006.

[91] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632–2639, September 2001.

[92] M. Vu and A. Paulraj, "MIMO Wireless Linear Precoding," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 86–105, September 2007.

[93] T. A. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley-Interscience, 1991.

[94] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, October 1998.

[95] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1165–1175, July 2004.

[96] A. Garcia-Rodriguez and C. Masouros, "Power-Efficient Tomlinson-Harashima Precoding for the Downlink of Multi-User MISO Systems," *IEEE Transactions on Communications*, vol. 62, no. 6, pp. 1884–1896, June 2014.

[97] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1305–1316, July 2004.

[98] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, no. 1, pp. 195–202, January 2005.

[99] C. Masouros, "Correlation Rotation Linear Precoding for MIMO Broadcast Communications," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 252–262, January 2011.

[100] C. Masouros and E. Alsusa, "Dynamic linear precoding for the exploitation of known interference in MIMO broadcast systems," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1396–1404, March 2009.

[101] ——, "Soft Linear Precoding for the Downlink of DS/CDMA Communication Systems," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 203–215, January 2010.

[102] S. M. Razavi, T. Ratnarajah, and C. Masouros, "Transmit-Power Efficient Linear Precoding Utilizing Known Interference for the Multiantenna Downlink," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 9, pp. 4383–4394, November 2014.

[103] A. Goldsmith, *Wireless Communications*. Cambridge University Press, August 2005.

[104] G. H. Golub and C. F. V. Loan, *Matrix Computations*. JHU Press, 2013.

[105] A. M. Ahmadian, W. Zirwas, and R. S. Ganesan, "Low complexity moore-penrose inverse for large como areas with sparse massive mimo channel matrices," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2016, pp. 1–7.

[106] H. V. Nguyen, V. D. Nguyen, and O. S. Shin, "Low-Complexity Precoding for Sum Rate Maximization in Downlink Massive MIMO Systems," *IEEE Wireless Communications Letters*, vol. 6, no. 2, pp. 186–189, April 2017.

[107] H. Sifaou, A. Kammoun, L. Sanguinetti, M. Debbah, and M. S. Alouini, "Power efficient low complexity precoding for massive MIMO systems," in

2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP), December 2014, pp. 647–651.

[108] H. Prabhu, J. Rodrigues, O. Edfors, and F. Rusek, "Approximative matrix inverse computations for very-large MIMO and applications to linear pre-coding systems," in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2013, pp. 2710–2715.

[109] S. Zarei, W. Gerstacker, R. R. Mller, and R. Schober, "Low-complexity linear precoding for downlink large-scale MIMO systems," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, September 2013, pp. 1119–1124.

[110] J. A. Lee and M. Verleysen, *Nonlinear Dimensionality Reduction*, M. Jordan, J. Kleinberg, and B. Scholkopf, Eds. Springer, 2007.

[111] *Spatial modulation for generalized MIMO: Challenges opportunities and implementations*, vol. 102, no. 1, January 2014.

[112] V. S. Krishna and M. R. Bhatnagar, "A joint antenna and path selection technique in single-relay-based df cooperative mimo networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1340–1353, March 2016.

[113] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Massive mimo in real propagation environments: do all antennas contribute equally?" *IEEE Transactions on Communications*, vol. 63, no. 11, pp. 3917–3928, 2015.

[114] H. Li, L. Song, D. Zhu, and M. Lei, "Energy efficiency of large scale mimo systems with transmit antenna selection," in *IEEE International Conference on Communication*, 2013 2013, pp. 4641–4645.

[115] H. Li, L. Song, and M. Debbah, "Energy efficiency of large scale multiple antenna system with transmit antenna selection," *IEEE Transactions on Communication*, vol. 62, no. 2, pp. 638–647, February 2014.

[116] Z. Zhou, S. Zhou, J. Gong, and Z. Niu, "Energy-efficient antenna selection and power allocation for large-scale multiple antenna systems with hybrid energy supply," in *IEEE Global Communications Conference (GLOBECOM)*, December 2014.

[117] B. Lee, L. Ngo, and B. Shim, "Antenna group selection based user scheduling for massive mimo systems," in *IEEE Global Communication Conference*, December 2014.

[118] M. Benmimoune, E. Driouch, W. Ajib, and D. Massicotte, "Joint transmit antenna selection and user scheduling for massive mimo systems," in *IEEE Wireless Communications and Networking Conference (WCNC)*, March 2015.

[119] T.-W. Ban and B. C. Jung, "A practical antenna selection technique in multiuser massive mimo networks," *IEICE Transactions on Communications*, vol. E96-B, no. 11, pp. 2901–2905, 2013.

[120] H. Huang, C. B. Papadias, and S. Venkatesan, *MIMO Communication for Cellular Networks*. Springer, New York, 2012.

[121] K. M. Abadir and J. R. Magnus, *Matrix Algebra*. Cambridge University Press, August 2005.

[122] F. Hu, *Opportunities in 5G networks: A research and development perspective*. CRC press, 2016.

[123] G. N. Kamga, M. Xia, and S. Assa, "Spectral-Efficiency Analysis of Regular- and Large-Scale (Massive) MIMO With a Comprehensive Channel Model," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 4984–4996, June 2017.

[124] G. N. Kamga, M. Xia, and S. Aissa, "Channel modeling and capacity analysis of large MIMO in real propagation environments," in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 1447–1452.

[125] R. H. Gohary and H. Yanikomeroglu, "The ergodic high SNR capacity of the spatially-correlated non-coherent MIMO channel within an SNR-independent gap," in *2015 IEEE Information Theory Workshop - Fall (ITW)*, October 2015, pp. 234–238.

[126] M. Matthaiou, N. D. Chatzidiamantis, and G. K. Karagiannidis, "A New Lower Bound on the Ergodic Capacity of Distributed MIMO Systems," *IEEE Signal Processing Letters*, vol. 18, no. 4, pp. 227–230, April 2011.

[127] M. Matthaiou, N. D. Chatzidiamantis, G. K. Karagiannidis, and J. A. Nossek, "On the Capacity of Generalized- k Fading MIMO Channels," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5939–5944, November 2010.

[128] N. Costa and S. Haykin, *Multiple-Input Multiple-Output Channel Models: Theory and Practice*. John Wiley & Sons, June 2010.

[129] L. Schumacher, K. I. Pedersen, and P. E. Mogensen, "From antenna spacings to theoretical capacities - guidelines for simulating MIMO systems," in *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 2, September 2002, pp. 587–592 vol.2.

[130] J. Hoydis, S. T. Brink, and M. Debbah, "Comparison of linear precoding schemes for downlink massive MIMO," in *2012 IEEE International Conference on Communications (ICC)*, June 2012, pp. 2135–2139.

[131] J. C. S. de Souza, T. M. L. Assis, and B. C. Pal, "Data Compression in Smart Distribution Systems via Singular Value Decomposition," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 275–284, January 2017.

[132] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, March 2006.

[133] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2700–2712, August 2005.

[134] K. Zu, R. C. de Lamare, and M. Haardt, "Lattice reduction-aided regularized block diagonalization for multiuser MIMO systems," in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2012, pp. 131–135.

[135] K. Zu, "Novel Efficient Precoding Techniques for Multiuser MIMO Systems," P.hD., University of York, May 2013. [Online]. Available: http://etheses.whiterose.ac.uk/4458/

[136] L. J. P. van der Maaten, E. O. Postma, and H. J. van den Herik, *Dimensionality Reduction: A Comparative Review*, 2008.

[137] S. Kumar, P. K. Gupta, G. Singh, and D. S. Chauhan, "Performance Analysis of Rayleigh and Rician Fading Channel Models using Matlab Simulation,"

International Journal of Intelligent Systems and Applications; Hong Kong, vol. 5, no. 9, pp. 94–102, August 2013.