Low-Complexity Antenna Selection Techniques for Massive MIMO Systems

Zaid Abduladheem Abdullah

Newcastle University

Newcastle upon Tyne, UK.

A thesis submitted for the degree of

Doctor of Philosophy

February 2019

To my beloved parents ...

Declaration

NEWCASTLE UNIVERSITY

SCHOOL OF ENGINEERING

I, Zaid Abduladheem Abdullah, declare that this thesis is my own work and it has not been previously submitted, either by me or by anyone else, for a degree or diploma at any educational institute, school or university. To the best of my knowledge, this thesis does not contain any previously published work, except where another person's work used has been cited and included in the list of references.

Signature:

Student: Zaid Abduladheem Abdullah

Date:

SUPERVISOR'S CERTIFICATE

This is to certify that the entitled thesis "Low-Complexity Antenna Selection Techniques for Massive MIMO Systems" has been prepared under my supervision at the school of Engineering / Newcastle University, for the degree of PhD in Electrical and Electronic Engineering.

Signature:

Supervisor: Dr. Charalampos Tsimenidis

Date:

Acknowledgements

When I look back at the years, I remember when I finished my undergraduate degree in my beloved country, Iraq. For a while, I felt as if all the doors were closed in front of me. Six years later, here I am, about to submit my PhD thesis, and I am filled with gratitude for everyone who helped me along the way.

To my Creator, thank you for giving me the strength to overcome all the challenges that I had to face in my life, especially during the last few years. Thank you for blessing me with a loving family, loyal friends, and so many other blessings that I can not even count.

To my supervisor Dr Charalampos Tsimenidis, thank you for your help and support during my time as your student. I feel extremely lucky to have you as my supervisor, and I appreciate all the constructive comments and guidance that I received from you over the years. Also I am truly fortunate to have my second supervisor Dr Martin Johnston. Your help, advice, support, and encouragement helped me enormously throughout my studies, and had a significant influence on my life.

I am also extremely grateful to Newcastle University for sponsoring my PhD, it is an opportunity that I truly appreciate. Among all the universities in the world, Newcastle was my first and only choice, and I could not be happier for choosing this great University to pursue my studies. I would also like to thank the staff members of Newcastle University, espically Prof. Satnam Dlay, who did everything he can so that I would have this scholarship.

To all my dear friends that I was lucky to meet here in Newcastle, and especially the ones in the intelligent sensing and communications group, thank you for the good times that we had together, and all the special memories that will live with me for many, many years to come.

Last but not least, to my lovely mother, Wafaa, and my dear father, Abduladheem, there are no words that can describe my appreciation and gratitude for your efforts, support, and love throughout my life. Thank you from the bottom of my heart for

putting me and my siblings first, and for all the sacrifices that you both made so that we can have the best possible life. Also, to my two brothers, Ahmed and Ali, and my three sisters, Sarah, Noor, and Zahar, thank you for believing in me, and for all the encouragement and support, you are the best siblings that anyone could ever ask for.

Abstract

Massive Multiple-Input Multiple-Output (M-MIMO) is a state of the art technology in wireless communications, where hundreds of antennas are exploited at the base station (BS) to serve a much smaller number of users. Employing large antenna arrays can improve the performance dramatically in terms of the achievable rates and radiated energy, however, it comes at the price of increased cost, complexity, and power consumption.

To reduce the hardware complexity and cost, while maintaining the advantages of M-MIMO, antenna selection (AS) techniques can be applied where only a subset of the available antennas at the BS are selected. Optimal AS can be obtained through exhaustive search, which is suitable for conventional MIMO systems, but is prohibited for systems with hundreds of antennas due to its enormous computational complexity. Therefore, this thesis address the problem of designing low complexity AS algorithms for multi-user (MU) M-MIMO systems.

In chapter 3, different evolutionary algorithms including bio-inspired, quantum-inspired, and heuristic methods are applied for AS in uplink MU M-MIMO systems. It was demonstrated that quantum-inspired and heuristic methods outperform the bio-inspired techniques in terms of both complexity and performance.

In chapter 4, a downlink MU M-MIMO scenario is considered with Matched Filter (MF) precoding. Two novel AS algorithms are proposed where the antennas are selected without any vector multiplications, which resulted in a dramatic complexity reduction. The proposed algorithms outperform the case where all antennas are activated, in terms of both energy and spectral efficiencies.

In chapter 5, three AS algorithms are designed and utilized to enhance the performance of cell-edge users, alongside Max-Min power allocation control. The algorithms aim to either maximize the channel gain, or minimize the interference for the worst-case user only.

The proposed methods in this thesis are compared with other low complexity AS schemes and showed a great performance-complexity trade-off.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms

ABC	Artificial Bee Colony
ADC	Analogue to Digital Converter
AF	Amplify-and-Forward
AS	Antenna Selection
$AWGN$	Additive White Gaussian Noise
BAB	Branch And Bound
BS	Base Station
CSI	Channel State Information
CTS	Classical Tabu Search
DF	Decode-and-Forward
DPC	Dirty Paper Coding
dB	Decibel
EE	Energy Efficiency
EPA	Equal Power Allocation
FDD	Frequency Division Duplex
$FLOP$	Floating-point Operation
GA	Genetic Algorithm
GSM	Generalized Spatial Modulation

$i.i.d.$	independent and identically distributed
LNA	Low Noise Amplifier
LP	Linear Programming
MF	Matched Filter
$M-MIMO$	Massive Multiple-Input Multiple-Output
$MIMO$	Multiple-Input Multiple-Output
ML	Maximum-likelihood
$MMSE$	Minimum Mean Square Error
$MMPA$	Max-Min Power Allocation
MRC	Maximum Ratio Combining
MRT	Maximum Ratio Transmission
MSE	Mean square error
MU	Multi-User
$NOMA$	Non-Orthogonal Multiple Access
$OFDM$	Orthogonal Frequency Division Multiplexing
PA	Power Allocation
$PAPR$	Peak to Average Power Ratio
PPA	Pilot Power Allocation
PSK	Phase Shift Keying
PSO	Particle Swarm Optimization
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
$QPSK$	Quadrature Phase Shift Keying
QTS	Quantum-inspired Tabu Search

$qubit$	quantum bit
RMV	Rectangular Maximum Volume
RF	Radio Frequency
SIC	Successive Interference Cancellation
$SINR$	Signal to Interference plus Noise Ratio
$SIRAS$	Semiblind Interference Rejection Antenna Selection
$SISO$	Single-Input Single-Output
$SLNR$	Signal to Leakage plus Noise Ratio
SM	Spatial Modulation
SNR	Signal to Noise Ratio
TDD	Time Division Duplex
TS	Tabu Search
$UCAS$	User-Centric Antenna Selection
ZF	Zero-forcing

Notations

$\mathbf{A}$	Uppercase boldface characters for matrices
$\mathbf{a}$	Lowercase boldface characters for vectors
$\mathbf{A}^T$	Transpose of the matrix **A**
$\mathbf{A}^*$	Conjugate the matrix **A**
$\mathbf{A}^H$	Hermitian the matrix **A**
$\text{Tr}[\mathbf{A}]$	Trace of the matrix **A**
$\|\mathbf{a}\|$	The Euclidean norm of the vector **a**
$\mathbf{a}_i$	The i^{th} row of matrix **A**
$\mathbf{a}_i^c$	The i^{th} column of matrix **A**
$a_{i,j}$	The i^{th} element of the j^{th} column of matrix **A**
$\boldsymbol{I}_N$	The $N \times N$ identity matrix
$\mathbf{0}_{M\times N}$	The $M \times N$ zero matrix
$\Re[a]$	The real part of the complex number a
$\Im[a]$	The imaginary part of the complex number a
$a!$	The factorial of the integer number a
$\binom{a}{b}$	The binomial coefficient which is evaluated as $\frac{a!}{b!(a-b)!}$

Chapter 1

Introduction

1.1 Why massive MIMO?

Since the start of the new millennium, the demand for mobile devices has increased enormously, and the growth of data traffic over wireless channels has followed Moore's law. For over two decades, researchers have shown that using systems with multiple antennas, known as Multiple-Input Multiple-Output (MIMO), can enhance the performance dramatically, and since then, MIMO systems have been part of many communications standards. However, for most of today's MIMO systems, the Base Station (BS) is equipped with only few antennas (less than 10), and although the use of conventional MIMO systems has shown great performance advantages, it is very unlikely to cope with the fast increasing demand for higher throughputs.

The exponential growth over data traffic during the last decade predicts that the number of connected devices is expected to reach 25 billion by the year 2020 [1], out of which 15 billion are phones, tablets, PCs, and laptops [2]. To meet the demands of such high data traffic, large scale MIMO, also known as Massive MIMO (M-MIMO), systems were first introduced in [3], where a BS with tens to hundreds of antenna elements is assigned to serve a much smaller number of users in the same time-frequency resources. Since then, M-MIMO systems have gathered the attention of researchers from all over the globe, and it was shown that employing very large number of antennas have tremendous advantages. For example, with massive arrays, the radiated power can be reduced dramatically, and the energy transmitted per bit vanishes when the number of antennas grows to infinity. Furthermore, the effect of uncorrelated noise as well as the small-scale fading are eliminated, and near optimal performance can be obtained using simple linear signal processing techniques.

1.2 Aims and motivation

Although there are tremendous theoretical advantages of using M-MIMO, there are many serious practical considerations that limit the applications of such systems. For example, using massive number of antennas at the BS at once means that each antenna must be connected to a separate Radio Frequency (RF) chain, and unlike the antenna elements, RF chains are expensive, since each RF chain consists of Low Noise Amplifier (LNA), mixer, and Analog to Digital Converter (ADC) [4]. More importantly, RF chains are highly power demanding elements, and they consume 50%-80% of the total trasceiving power [5]. Therefore, activating large number of RF chains at the same time will degrade the energy efficiency performance of the system dramatically.

One way to reduce the cost, complexity, and power consumption while preserving the great advantages offered by M-MIMO systems, is through applying Antenna Selection (AS) techniques, where massive number of antenna elements are placed at the BS with limited number of RF chains.

For over a decade, AS has been a widely studied topic in conventional MIMO systems [6–20]. It was first introduced as a tool to reduce the hardware complexity at the transmitter and the receiver of multiple antenna systems. Moreover, dedicating a separate RF chain for each radiating element was found to be the main reason behind causing higher and unstable power consumption. However, the work originated on AS in conventional MIMO systems have proven not to be suitable to apply for M-MIMO, due to their enormous complexity requirement for systems with massive arrays.

In addition, Power Allocation (PA) schemes can be applied at the BS for many reasons, such as: minimizing the total transmission power under a certain Quality of Service (QoS), maximizing the total sum rate, increasing the fairness among the users, or maximizing the Energy Efficiency (EE) of the system.

1.3 Challenges and solutions in M-MIMO systems

As explained above, employing large number of RF chains will impose extremely high hardware complexity and cost, while degrading the energy efficiency performance of the system dramatically. In addition, the existing algorithms for AS in conventional MIMO systems are not suitable to be applied in M-MIMO, due to their enormously high computational complexity requirement for systems with massive number of antenna elements.

It is worth to mention that compared to the uplink case, performing AS in the downlink scenario gives more options for performance improvement, especially when linear methods are applied such as Matched Filter (MF) for signal processing. The reason behind this is that the BS knows the effect of activating each antenna on the Signal to Interference plus Noise Ratio (SINR) for each user. In contrast, and for the uplink case, the received signal at each antenna is a combination of signals transmitted from all users, and therefore it is more difficult to decide which antenna is suffering from the highest interference, as well as to separate the interference from different users.

Furthermore, in practical scenarios the users are uniformly distributed and have different distances from the BS. Therefore, to increase the fairness among the users, and to ensure that each user meet a predefined threshold of QoS, max-min PA techniques can be applied at the BS, where more power is allocated for users who are located far away from the BS.

Accordingly, in this thesis we design low complexity AS methods for MU M-MIMO systems in both uplink and downlink transmission scenarios. Different methodologies have been adopted in our work, and the proposed methods target both high performance and low implementation complexity in terms of floating point operations (FLOPs). Furthermore, max-min PA in a single-cell M-MIMO is also applied to enhance the rate for the worst case user and increase the fairness between users.

1.4 Literature review on AS and PA in M-MIMO systems

Since our work in this thesis involves both PA and AS, we carry an extensive survey on the related work on these subjects in the following subsections.

1.4.1 Related work on AS in M-MIMO systems

There has been a considerable amount of work recently on AS in M-MIMO systems. for example, tshe authors in [21,22] proposed low complexity AS methods for energy efficient M-MIMO systems, however, their algorithms were designed for point to point case, i.e. only one user is scheduled to be served at a given time instant. In contrast, the authors in [23] proposed an energy efficient AS method for Multi-User (MU) M-MIMO downlink transmission. However, maximizing the EE results in poor spectral efficiency, and vice versa, therefore in this thesis we address the problem of maximizing the SINR for any arbitrary number of selected antennas. The authors in [24] proposed a novel iterative AS method for an uplink point to point

M-MIMO with Maximum Ratio Combining (MRC) receiver, under spatially correlated channels. Their work focused on minimizing the Mean Square Error (MSE) of the received signal to improve the error rate performance for a single user scenario. For capacity maximization in an interference-free scenarios, the authors in [25] proposed a Rectangular Maximum-Volume (RMV) theory based AS, for downlink M-MIMO systems, while the authors in [26] proposed a Branch And Bound (BAB) based AS for an uplink M-MIMO transmission. Although their methods demonstrate high performance, they suffer from relatively high complexity. Furthermore, the authors in [27, 28] applied AS in M-MIMO systems on measured channels using both linear and cylindrical arrays equipped with 128 antennas, to maximize the sum rate with Zero Forcing (ZF) precoder. In [29, 30], the authors designed an AS algorithm with quartic complexity in point to point M-MIMO systems. While in [31], the authors designed a joint AS and user scheduling in downlink MU M-MIMO systems with ZF precoder, while in [32] the authors designed a joint antenna and user selection for M-MIMO with non-orthogonal multiple access (NOMA) systems. Moreover, the authors in [33] studied the trade-off between energy and spectral efficiencies under random AS in MU M-MIMO downlink systems. In addition, the authors in [34, 35] proposed a novel AS by exploiting the known constructive interference at the BS for downlink transmission. However, their method works efficiently only on low modulation Phase Shift Keying (PSK) signalling, and it is data dependent, which means that extremely fast RF switching are required. Furthermore, the same authors proposed in [36, 37] a novel joint precoding and AS scheme relying on the constructive interference. In [38], the authors proposed a joint AS and user scheduling scheme under ZF precoder. In [39] the authors proposed a two-steps AS method for point to point M-MIMO systems, where in the first step the algorithm focus on selecting the antennas that has the least spatial correlation between them, followed by the second step which focuses on the performance. Moreover, in [40], the authors proposed sub-optimal AS methods for multi-cell cooperative M-MIMO systems, their methods focused on maximizing the Signal to Noise Ratio SNR while maintaining low computational complexity. The authors in [41] proposed an AS method for power minimization in MU downlink M-MIMO systems, under a predefined QoS requirement. Furthermore, in [42] the authors proposed a bidirectional branch and bound method for AS in M-MIMO, although their method demonstrates high performance, it suffers from large complexity requirement. In [43], the authors proposed an AS under interference alignment, where the interference is forced to zero through transmitter-receiver beamforming. Finally, the authors in [44], proposed joint AS and PA schemes for MU downlink M-MIMO systems with linear precoding techniques.

1.4.2 Literature review on PA in M-MIMO

The authors in [45] considered an uplink MU M-MIMO system, and proposed an optimal power control that jointly selects the training duration, training pilots power, and data power. Their results show that at low SNR, higher power needs to be adopted for training pilots than at high SNR. Furthermore, they showed that the optimal training duration is always equal to the number of users regardless of the SNR regime. In [46], the authors studied joint PA and pilots assignment to maximize the EE of a multi-cell MU M-MIMO system with pilot contamination. Moreover, the authors in [47] investigated PA strategies for multi-cell MU M-MIMO system for sum rate maximization in both uplink and downlink transmission, under ZF equalizer and precoder. Their work demonstrated that optimal PA can highly improve the system performance for practical number of antennas at the BS (e.g., tens to hundreds), however, for a fixed number of users, equal PA becomes optimal as the number of antennas at the BS tends to infinity. Joint PA and user association was proposed in [48] for multi-cell M-MIMO downlink system, under both Maximum Ration Transmission (MRT) and ZF precoding techniques. The authors aimed to minimize the total power transmission by optimizing the subset of BSs to serve each user. Furthermore, the authors in [49] proposed a pilot power allocation (PPA) through user grouping, in a multi-cell M-MIMO systems. An EE PA method was proposed in [50] for downlink M-MIMO with MF precoding. Their work showed an improved EE with reduced transmission power compared to other PA schemes that ignores the interuser interference. While in [51], the authors proposed an optimal PA schemes for MU M-MIMO systems under ZF detector. In their work, an optimal power control was carried on both pilot and data signals based on the large scale fading to maximize the total sum rate. The authors in [52] proposed a pilot design with uplink PA for multi-cell MU M-MIMO systems to mitigate the pilot contamination problem, and max-min power control was applied to ensure fairness between users. In [53], an uplink power control was investigated in multi-cell MU M-MIMO systems, where the BS is equipped with large but finite number of antennas, with Minimum Mean Square Error (MMSE) receivers, under pilot-contaminated channel estimation. In addition, the authors in [54] developed a low complexity PA scheme for EE MU M-MIMO systems with ZF processing, and they proposed a joint PA, number of antenna elements, and user scheduling, assuming a Time Division Duplex (TDD) downlink scenario. In [55], the authors applied optimal PA for multi-pair amplify and forward (AF) M-MIMO relaying under imperfect Channel State Information (CSI). Also with AF M-MIMO relaying, the authors in [56] proposed an EE PA scheme, where the EE performance was theoretically analysed by employing random matrix theory and large system

analysis. Finally, in [57], the authors applied PA for M-MIMO with decode and forward (DF) relaying, assuming MMSE channel estimation and ZF transceivers.

1.5 Thesis organization and contribution

The aim of this thesis is to design and investigate novel AS schemes in MU M-MIMO systems, in both uplink and downlink scenarios, as illustrated in Fig. 1.1. In the uplink case, the motivation was to optimize the sum rate capacity, i.e. the capacity that can be obtained by employing successive interference cancellation (SIC) methods. For the downlink case, the motivation was to maximize the SINR under MF precoding, in both cases when the users are equidistant to the BS, and for uniformly distributed users.

This chapter presents the motivation behind employing M-MIMO systems, and the key role of AS in reducing the cost and complexity of such systems. In addition, an extensive survey is carried out about the previous work on AS as well as PA in M-MIMO systems.

Chapter 2 introduces the required background knowledge on the signal processing techniques for MIMO systems, such as: different methods of linear and non-linear equalization and precoding techniques. Moreover, the idea of AS in both uplink and downlink cases is explained, with some benchmark methods of AS in conventional MIMO systems. In addition, a brief summary of the main challenges in employing M-MIMO systems are also briefly explained in both time and frequency division duplex.

In Chapter 3, an uplink M-MIMO system is considered, and several evolutionary methods for AS are applied to maximize the sum rate capacity. We design a classical Tabu Search (CTS) method, and also apply a Quntum-inspired TS (QTS) for AS and compare our techniques with well known bio-inspired algorithms, such as: Genetic Algrotihm (GA), Artificial Bee Colony (ABC), as well as Particle Swarm Optimization (PSO). The proposed methods outperform the bio-inspired techniques in terms of both performance and complexity requirements.

In Chapter 4, we aim to achieve a further reduction in complexity, in terms of both precoding and AS schemes. Accordingly, a MF precoder is applied in MU M-MIMO downlink transmission, with equidistant users. Two novel AS schemes are proposed to maximize the SINR, and hence maximizing the sum rate. Both algorithms achieve dramatic complexity reduction by avoiding vector multiplications during the iterative selection process. Our results demonstrate that the proposed methods can achieve higher sum rate and energy efficiency with a subset of the available antennas than activating the full antenna subset.

In Chapter 5, a uniformly distributed users were assumed in a downlink M-MIMO scenario

with MF precoding. Three low complexity AS methods were proposed to maximize the SINR for the cell-edge users. The first algorithm aims to maximize the channel gain for the worst-case user, while the second method focuses on minimizing the interference, and the last method aims to minimize the highest interference terms between any two given users. Furthermore, a max-min power control is carried out to further enhance the achievable rate of the worst-case-users, and hence higher fairness is achieved.

Chapter 6 concludes this thesis, and few conclusions are derived from each chapter to summarize the work in this thesis. Furthermore, many interesting directions for future work are also briefly explained.

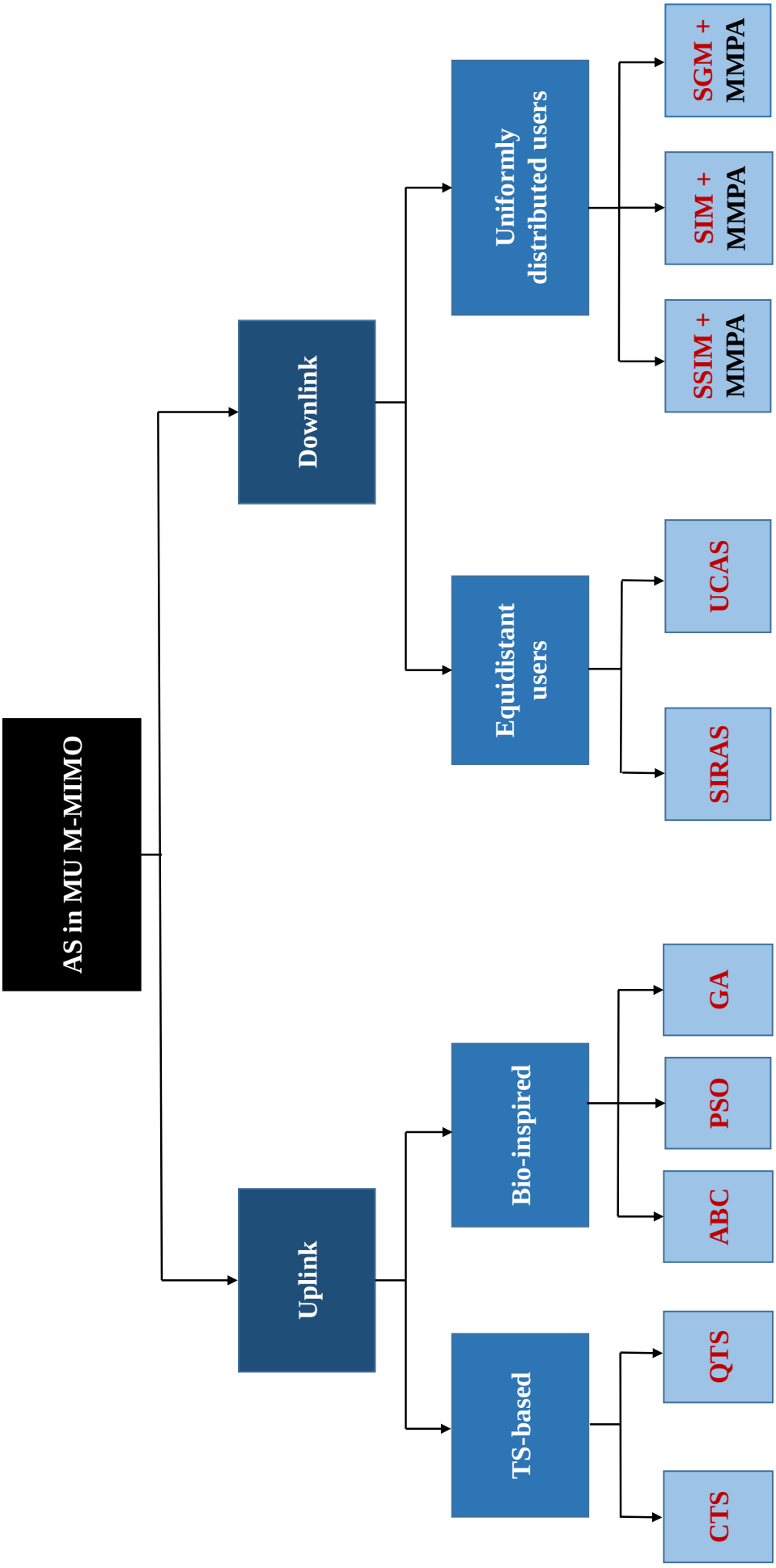


Figure 1.1: Thesis structure

1.6 Publications arising from this research

- **Z. Abdullah**, C. C. Tsimenidis and M. Johnston, "Tabu Search vs. Bio-inspired Algorithms for Antenna Selection in Spatially Correlated Massive MIMO Uplink Channels," in Proc. 24th European Signal Processing Conference (EUSIPCO), Budapest, 2016.

- **Z. Abdullah**, C. C. Tsimenidis and M. Johnston, "Quantum-inspired Tabu Search Algorithm for Antenna Selection in Massive MIMO Systems," in Proc. Wireless Communications and Networking Conference (WCNC), Barcelona, 2018.

- **Z. Abdullah**, C. C. Tsimenidis and M. Johnston, "Efficient Low-Complexity Antenna Selection Algorithms in Multi-User Massive MIMO Systems with Matched Filter Precoding," submitted to IEEE Transactions on Vehicular Technology, July. 2018.

- **Z. Abdullah**, C. C. Tsimenidis, M. Alageli, and M. Johnston, "Cell-Edge-Aware Antenna Selection and Power Allocation in Massive MIMO Systems," submitted to IEEE Global Communications Conference.

Chapter 2

MIMO Communications systems: Detection, Precoding, and Antenna Selection techniques

2.1 MIMO systems

The term MIMO simply refers to a system where multiple antennas are used at the transmitter and receiver ends. Considering a system with N_t transmit antennas and N_r receive antennas, The received signal can be given as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2.1}$$

where $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ is the received signal by the N_r antennas, $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is the independent and identically distributed (i.i.d) wireless channels between the transmitter and receiver, with zero mean and unit variance. $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ is the transmitted information symbols vector, and $\mathbf{n} \in \mathbb{C}^{N_r \times 1}$ is the AWGN noise vector at the receiver with zero mean and variance of σ_n^2, i.e. $\mathbf{n} \sim \mathcal{CN}(\mathbf{0},\, \sigma_n^2 \boldsymbol{I}_{N_r})$. In a vector form, the equation in (2.1) can be expressed as

$$\begin{bmatrix} y_1 \\ \vdots \\ \vdots \\ y_{N_r} \end{bmatrix} = \begin{bmatrix} h_{1,1} & \dots & \dots & h_{1,N_t} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ h_{N_r,1} & \dots & \dots & h_{N_r,N_t} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_{N_t} \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ \vdots \\ n_{N_r} \end{bmatrix}, \tag{2.2}$$

where $h_{i,j}$ represents the channel coefficient between the i^{th} receive and the j^{th} transmit antennas.

2.1.1 Point-to-Point MIMO systems

In point to point MIMO systems, two different approaches can be employed to either enhance the signal reliability, or the throughput. Consider a single-input single-output (SISO) system, i.e. both transmitter and receiver are equipped with only one antenna, if the channel between the transmitter and the receiver is experiencing a deep fade, the received signal will likely contain errors. One way to improve the communication reliability is by sending the same signal through different antennas, as long as these antennas have enough separation to obtain independent channels, and at least one of the channels is strong. This technique is known as *diversity gain* [58]. On the other hand, MIMO systems can be employed to increase the throughput of the system via *spatial multiplexing* [58], where different data streams can be transmitted through different antennas, and separated at the receiver end, using either linear or non-linear detection methods.

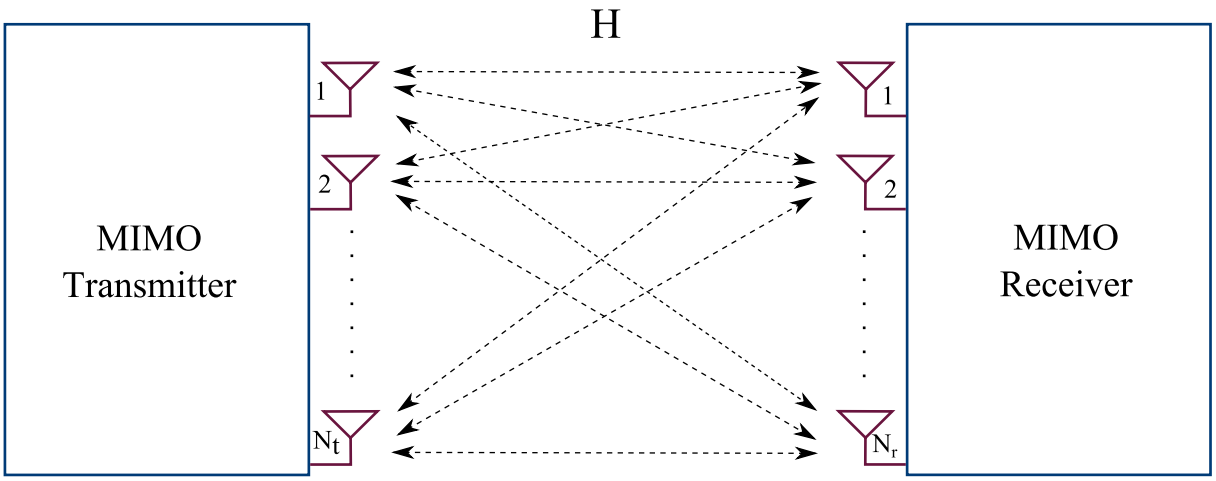


Figure 2.1: Point-to-Point MIMO systems.

2.1.2 Multi-User MIMO systems

In MU MIMO systems, multiple users communicate simultaneously with the BS using the same time-frequency resources as shown in Fig. 2.2. However, since the mobile users do not have the same computational capacity to run the same amount of processing as the BS, it is desirable that the BS will handle most of the processing, in both uplink and downlink scenarios.

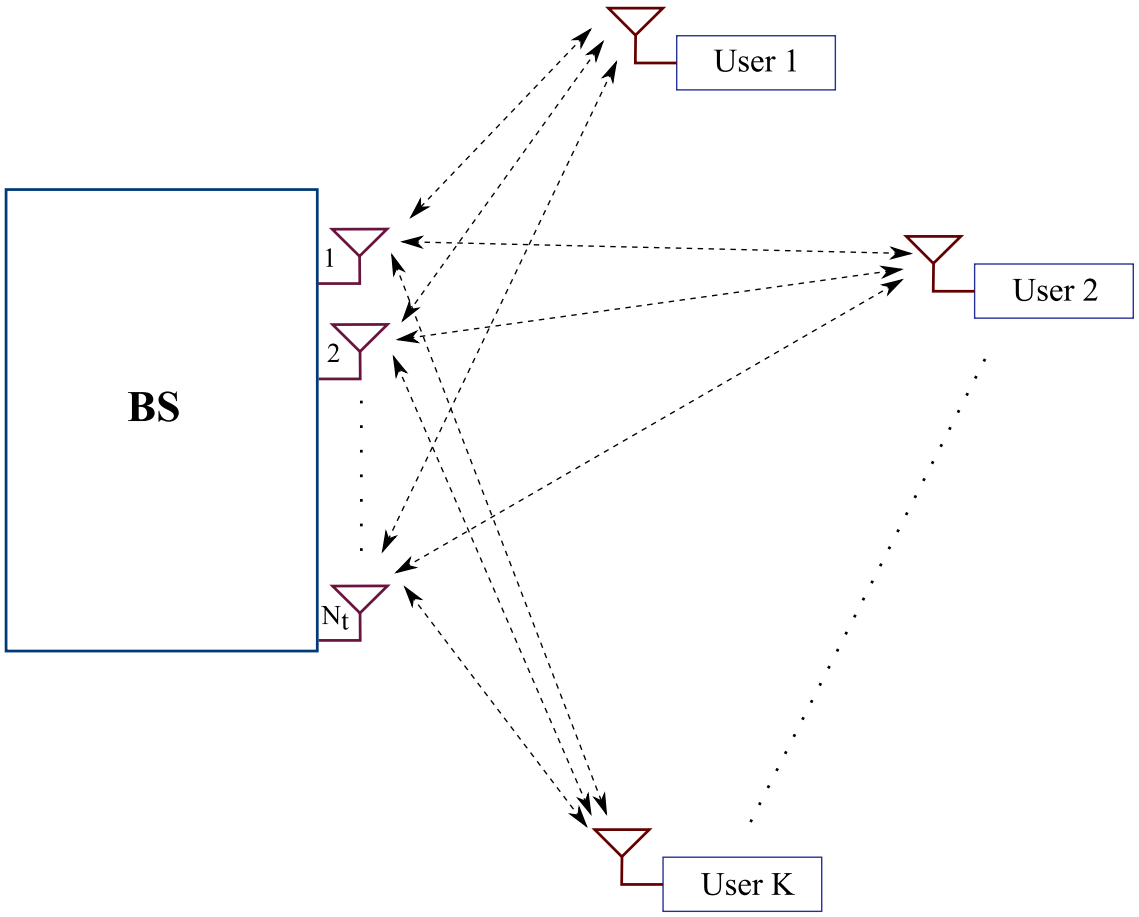


Figure 2.2: Multi-User MIMO systems.

2.2 Detection methods in MIMO systems

Consider a K single-antenna users transmitting their data to a BS equipped with N antennas, i.e. uplink transmission with spatial multiplexing scenario. These data streams will be combined at the receiver side, and therefore need to be separated to obtain the original data sent by the different users. The detection methods are applied at the BS to obtain an estimate of the unknown transmitted vector $\mathbf{x}$, for a given channel matrix $\mathbf{H}$, and received vector $\mathbf{y}$. There are two different types of detection methods: Linear and Non-linear detectors.

2.2.1 Non-linear detection methods

Non-linear detectors are known for their high performance, and they are able to achieve the MIMO capacity [9], which for an uplink case with perfectly known channel at the BS and an SNR of γ, can be given as [59]

$$\mathbf{C}_{\text{ul}} = \log_2 \det \left(\mathbf{I}_{N_r} + \frac{\gamma}{N_t} \mathbf{H}\mathbf{H}^H \right). \tag{2.3}$$

However, that comes at the price of high complexity. In the following section, we will review two well known non-linear detection schemes: Maximum likelihood (ML), and SIC.

2.2.1.1 ML detector

In any communications system, the presence of the AWGN is inescapable, and therefore errors are likely to occur. However, ML detector provides the optimum estimate of the transmitted vector $\mathbf{x}$ [60, 61]. Let $\mathcal{M}$ represents the constellation set that contains the $\mathrm{M-PSK}$ or $\mathrm{M-QAM}$ signals, where QAM refers to Quadrature Amplitude Modulation. For a system with K single-antenna users, the number of different candidates for the transmitted vector $\mathbf{x}$ is M^K, where M is the size of the set $\mathcal{M}$. For example, for system with 2 users transmitting a Quadrature PSK (QPSK) signal, there will be 16 candidate solutions as demonstrated in Table 2.1.

ML detector aims to minimize the Euclidean distance of the noise, by carrying an exhaustive search over all possible vector candidates in the following form

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathcal{M}^K} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 . \tag{2.4}$$

Although ML detector provides optimal solution, it suffers from an enormous complexity that grows exponentially with increasing the number of users or the size of the constellation set $\mathcal{M}$, which makes it only applicable for systems with small number of users, and small constellation sets.

2.2.1.2 SIC detection

SIC detection methods are also non-linear techniques, and they rely on performing a linear equalization techniques to remove the effect of the channel first, and then detect one stream of data at each step. The procedure of SIC can be explained as follow [60]

- At the beginning, the strongest symbol in a given data stream is detected using a linear detector such as Zero Forcing, referred to as ZF-SIC. The received streams from different users are ordered based on their received SNRs.

Table 2.1: ML candidate solutions when $K = 2$, with QPSK constellation.

Candidate index	Candidate Solution	Candidate index	Candidate Solution
1	$\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$	9	$-\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$
2	$\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$	10	$-\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$
3	$\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$, $-\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$	11	$-\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$, $-\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$
4	$\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$, $-\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$	12	$-\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$, $-\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$
5	$\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$	13	$-\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$
6	$\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$	14	$-\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$
7	$\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$, $-\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$	15	$-\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$, $-\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$
8	$\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$, $-\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$	16	$-\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$, $-\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$

- After detecting the strongest data, the BS estimates the interference caused by the detected symbol, and then subtract the interference caused by it from the received signal.
- The output signal after subtracting the interference will then be used to detect the next strongest symbol and subtract its interference and so on. The interference cancellation will run until detecting the weakest received data stream.

It should be noted that although SIC methods have less complexity than ML, they still impose high complexity requirement for system with large number of users or antennas, and therefore they are unsuitable solutions for M-MIMO systems. Instead, simple linear detection methods can achieve near optimal solutions when the BS is equipped with very large number of antenna elements.

2.2.2 Linear detection techniques

Linear detectors in general require low complexity, and therefore they are highly used with systems with high dimensions, such as M-MIMO. The main idea behind these techniques is to equalize the effect of the channel at the receiver side first, then find the minimum euclidean distance between the equalized symbols and the correspondent constellation. Two different linear detectors are presented in this section, which are ZF and MF.

2.2.2.1 Matched filter detector

MF represents the simplest form of linear detection, where the received signal is multiplied at the BS with the Hermitian transpose of the channel matrix $\mathbf{H}$. However, MF does not null the interference between different users, and it treats it merely as a noise. Assuming K single antenna users transmitting their data to a BS with N antennas, the received signal in (2.1) can be re-written as

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\mathbf{x} + \mathbf{n} \\ &= \sum_{i=1}^{K} \mathbf{h}_i x_i + \mathbf{n} \\ &= \mathbf{h}_k x_k + \sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{h}_i x_i + \mathbf{n}, \end{aligned} \tag{2.5}$$

To estimate the signal sent by the k^{th} user, the BS multiplies the received signal with $\mathbf{h}_k^H$ [60], i.e.

$$\begin{aligned} \hat{x}_k &= \mathbf{h}_k^H \mathbf{y} \\ &= \mathbf{h}_k^H \left(\sum_{i=1}^{K} \mathbf{h}_i x_i + \mathbf{n} \right) \\ &= \mathbf{h}_k^H \mathbf{h}_k x_k + \sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{h}_k^H \mathbf{h}_i x_i + \mathbf{h}_k^H \mathbf{n}, \end{aligned} \tag{2.6}$$

where the first term in the right hand side of (2.6) represents the desired signal, the second term represents the interference, while the last term is the noise. A hard decision will then be made to map $\hat{x}_k$ to the nearest symbol in the alphabet of the utilized modulation set. In a general vector form, MF detection can be expressed as follow

$$\hat{\mathbf{x}}_{MF} = \mathbf{H}^H \mathbf{y}. \tag{2.7}$$

The downside of the MF detector is the interuser interference, which becomes the dominant degradation factor at high SNRs. However, with M-MIMO, the increase in the interuser interference term is smaller than the gain for the desired signal, therefore as the number of antennas goes to infinity, the simple MF becomes a near optimal solution.

2.2.2.2 Zero forcing detector

The ZF is another form of linear detection, and it eliminates the interuser interference by multiplying the received signal with the pseudo inverse of the channel matrix, $\mathbf{H}^{\dagger}$, which can be given as [60]

$$\mathbf{H}^{\dagger} = (\mathbf{H}^{H}\mathbf{H})^{-1}\mathbf{H}^{H}, \tag{2.8}$$

let $\boldsymbol{q}_k$ be the k^{th} row of $\mathbf{H}^{\dagger}$, the signal sent by the k^{th} user can be estimated as follow

$$\begin{aligned} \hat{x}_k &= \boldsymbol{q}_k \mathbf{y} \\ &= \boldsymbol{q}_k(\mathbf{H}\mathbf{x} + \boldsymbol{n}) \\ &= x_k + \mathbf{q}_k \mathbf{n}. \end{aligned} \tag{2.9}$$

In a general vector form, the ZF solution can be expressed as follow

$$\hat{\mathbf{x}}_{ZF} = \mathbf{H}^{\dagger}\mathbf{y}. \tag{2.10}$$

Although ZF cancels the interference, it does amplify the noise, i.e. it causes noise enhancement at the receiver side [59, 61, 62].

2.3 Precoding methods in MIMO systems

In the downlink scenario, the BS transmit its data to the users simultaneously, and precoding schemes need to be applied to convert the K symbols message vector into an N data vector, where N and K are the number of antennas at the BS and the number of users, respectively. The maximum achievable capacity in the downlink scenario, with perfect knowledge of the channel matrix and equal power allocation among the antennas, can be given as [7]

$$\mathbf{C}_{\text{dl}} = \log_2 \det\Big(\mathbf{I}_K + \gamma\, \mathbf{H}^{H}\mathbf{H}\Big). \tag{2.11}$$

Precoding can be applied via either linear or non-linear methods, where the latter achieve higher throughput at the cost of increased complexity.

2.3.1 Non-linear precoding schemes

Although the purpose behind applying precoding techniques is mainly to reduce the complexity at the receiver end, the complexity of Non-linear schemes might be infeasible to apply even at the BS due to its high complexity. However, in this section we explain the optimum precoding scheme in terms of the achievable sum rate capacity, i.e. *"dirty paper coding"*, in the following subsection.

2.3.1.1 Dirty paper coding

In order to achieve the sum rate capacity of MU MIMO systems giving in 2.11, Dirty Paper Coding (DPC) needs to be applied [63–66], where the BS selects codewords for each user to eliminate the effect of interuser interference.

The name *"dirty paper coding"* refers to the fact that transmitting signal through a channel with interference is similar to writing on a paper with dirt spots. Assume that we have the message *"Hello World"* that we aim to transmit to the receiver through the interfering channel, as shown in Figs. 2.3a, and 2.3b, respectively. One way to transmit the signal is by avoiding the dirt spots as much as possible, however some dirt will remain in the message and the receiver might recover the original message with errors as shown in Fig. 2.3c. The optimum way to transmit the message without being affected by the dirt spots, is through DPC, where the transmitter and receiver agree upon a certain codeword that can adopt perfectly to the dirt spots, as if they did not exist, as shown in Fig. 2.3d.

Although DPC proves to be optimum in terms of performance, its implementation impose a significant complexity which makes it a prohibited approach for practical scenarios. Instead, linear precoding techniques can be applied which have sub-optimal performance with affordable complexity.

2.3.2 Linear precoding schemes

In practical scenarios, the mobile user does not have the knowledge of the channels between the BS and other users, therefore, precoding techniques are applied at the BS to maximize the SNR at the receiver end. In other words, in both uplink and downlink, the aim is to perform the complex signal processing at the BS rather than at the mobile user. Two main linear precoding techniques are presented in this section, MF and ZF.

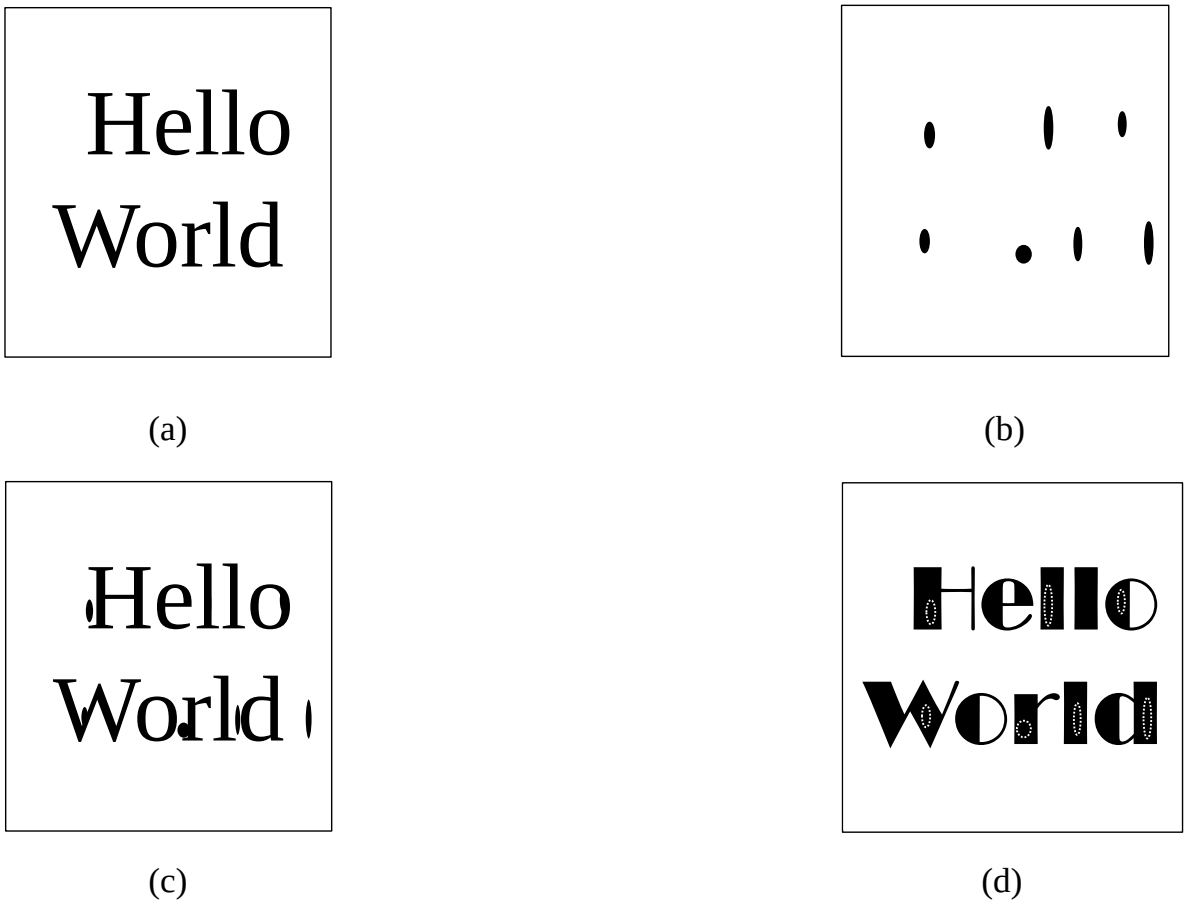


Figure 2.3: Writing on dirty paper: (a) "Hello World" is the original message that we aim to transmit, (b) The paper with dirt spots that we need to write our message on, (c) The transmitter aims to avoid the dirt spots, however, some dirt still exist, (d) The transmitter and receiver agree on a codeword that will adopt to the dirt spots and null the interference completely.

2.3.2.1 Matched filter precoding

MF precoding is the simplest form of linear precoding, and for a user k, it can be expressed as

$$\mathbf{w}_k^{MF} = \frac{\mathbf{h}_k^*}{\|\mathbf{h}_k\|}, \tag{2.12}$$

where the term in the denominator of (2.12) is the scaling factor, and used to prevent the BS from exceeding the power limit. The received signal at the k^{th} user can be given as

$$y_k^{MF} = \sqrt{p_k}\mathbf{h}_k^T\mathbf{w}_k^{MF}x_k + \sum_{\substack{i=1 \\ i\neq k}}^{K}\sqrt{p_i}\mathbf{h}_k^T\mathbf{w}_i^{MF}x_i + n_k, \tag{2.13}$$

where $\sqrt{p_k}$ is the power allocated for the k^{th} user, and the second term in the right hand side of (2.13) represents the interuser interference.

2.3.2.2 Zero forcing precoding

The precoding matrix utilizing the ZF precoder can be expressed as

$$\mathbf{W}^{ZF} = \gamma_{ZF}\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}, \tag{2.14}$$

where γ_{ZF} is the scaling factor that is needed to satisfy the total power constraint, and it can be given as [67]

$$\gamma_{ZF} = \frac{1}{\sqrt{\text{Tr}[(\mathbf{H}\mathbf{H}^H)^{-1}]}}. \tag{2.15}$$

The data symbols are multiplied by the precoding matrix before being transmitted through the channel, and the received signal by the K users can be given as

$$\mathbf{y}^{ZF} = \mathbf{H}\mathbf{W}^{ZF}\mathbf{x} + \mathbf{n}. \tag{2.16}$$

The main advantage of ZF over MF precoding, is that it can null the interuser interference at the receiver end. However, there are many disadvantages for this type of precoding. One of the main limitations of ZF is that when the number of users grows large for a fixed number of antennas at the BS, ZF precoder suffers from a sum rate loss. For example, in [27], the authors showed that when the BS was equipped with 16 antennas, the sum rate achieved with 4 users was higher than that with 16 users. The reason behind this is that ZF wastes large amount of the available power just to null the interference, which results in low signal power. Moreover, ZF suffers from degraded performance when there are users located on the cell-edge, also known as the *"near-far problem"*. In addition, ZF precoder is very sensitive to different types of distortions such as unmodeled interference [68, 69]. In contrast, MF precoding has lower complexity and can accommodate more users than ZF, however, its performance is limited by the interuser interference especially at high SNR regime.

2.4 MIMO systems with limited number of RF chains

There is no doubt that employing multiple antennas can increase the capacity and reliability of any communication system. However, employing large number of antennas, at the transmitter or receiver, comes at the price of increased hardware complexity, cost, and power consumption, which is caused by the demand for increased number of RF chains. Each RF chain consists of power amplifier, ADC, LNA, and mixers [70], [71]. For those reasons, and unlike the antenna elements, RF chains are not only expensive, but also highly power demanding elements. Accordingly, reducing the number of RF chains can dramatically improve the EE of the communication systems.

Two main methods have been proposed to employ MIMO systems with reduced number of RF chains. The first approach is AS, where only the subset of antennas that maximize a given performance metric is selected. The second approach is called Spatial Modulation (SM), where the index of the antenna that is transmitting the signal is also carrying information. In the following, we present the two aforementioned approaches.

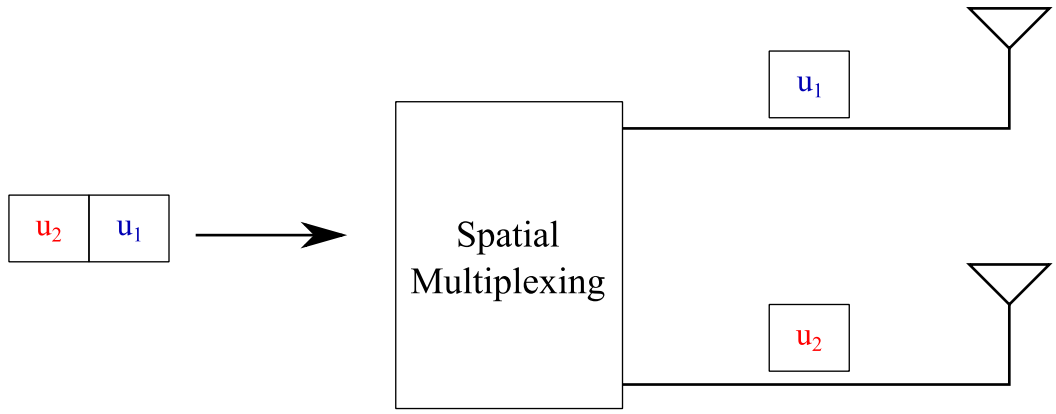
2.4.1 Spatial modulation

In SM, the transmitter is equipped with multiple antennas, however, only one of them is activated in a given time instance. The main difference between the spatial modulation and spatial multiplexing is depicted in Fig. 2.4, for a system equipped with two transmit antennas and two BPSK symbols to be transmitted in a single channel use.

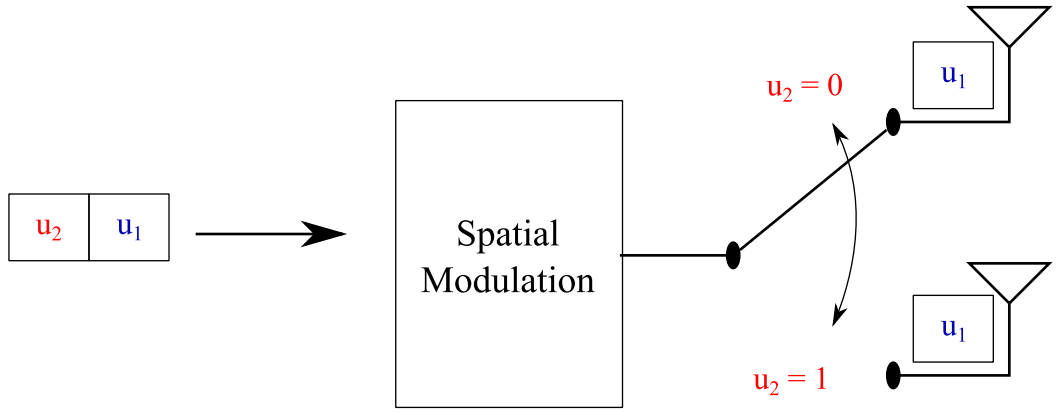
On one hand, transmitting the two symbols utilizing spatial multiplexing concept will lead to each antenna being allocated one information symbol to transmit. Consequently, both RF chains need to be activated at the transmitter, which can result in poor EE performance for the system. On the other hand, SM can be utilized, where the first symbol can be assigned to one antenna, while the second symbol determines the transmit antenna index. In other words, the first and second information symbols are explicitly and implicitly transmitted, respectively, in a single channel use utilizing only one RF chain.

In general, assuming that the transmitter is equipped with N antenna elements, and the cardinality of the PSK constellation set equals to M, the achievable rate of SM system can be given as follow [72–75]

$$R_{SM} = \log_2 N + \log_2 M. \tag{2.17}$$



(a)



(b)

Figure 2.4: The concept of spatial multiplexing and spatial modulation

2.4.2 Generalized spatial modulation

The generalized spatial modulation (GSM) is an extension of the SM system, however, here the transmitter can be equipped with multiple RF chains, that can be activated simultaneously instead of only one RF in the SM. Although GSM suffers from higher detection complexity than SM, it offers higher achievable rates for two reasons: First, the number of activated RF chains is higher than that in SM. Second, the number of implicitly encoded bits is higher due to the increased number of activated antenna subsets. The achievable rate for system employing GSM can be given as follow

$$R_{GSM} = \log_2 \left\lfloor \binom{N}{N_{RF}} \right\rfloor + N_{RF} \log_2 M. \tag{2.18}$$

However, one of the main challenges for both SM and GSM is the fast RF switching required

due to their encoding mechanism. In other words, SM and GSM both require RF switching that operates at the symbol rate, with low insertion losses. In addition, and for the GSM techniques, low complexity detection methods are essential due to the prohibited complexity of ML detector for higher number of transmit antennas.

2.4.3 Antenna selection

Another way to maintain the advantages offered by MIMO systems while reduce the cost, complexity, and power consumption is through applying AS techniques, where only a subset of the antennas are activated. AS techniques can be applied at both transmitter and receiver, whether its an uplink or a downlink scenario.

Although SM can achieve higher rates than AS, there are many advantages for applying AS over SM. For example, the detection of the activated antennas when SM is applied impose higher complexity at the receiver side. However, no such a problem exist in AS. Moreover, SM requires extremely fast RF switching since it is data dependent technique, while AS can be performed over several channel realizations, especially when dealing with slow time varying channels.

Optimal AS can be obtained via exhaustive search, where all the different combinations of antenna subset are tested, and the subset that gives the best performance metric will be selected. However, with large number of antennas, this method becomes prohibited, due to its enormous complexity requirement. Therefore, suboptimal solutions with lower complexity should be applied.

AS methods can be applied in both MU MIMO systems, or point to point MIMO systems, as demonstrated in Figs 2.5 and 2.6, respectively. In point to point MIMO systems, the AS can be applied at the transmitter, receiver, or both at the same time. While in MU MIMO systems, AS is usually applied at the BS since, in most scenarios, the mobile users are equipped with only one antennas. AS methods can be applied to optimize different performance metrics. Most of the work on AS has focused to improve one of the following metrics: maximizing the capacity, minimizing the error rate, or to maximize the EE of a communication system. However, in this section, we focus on AS methods to maximize the capacity of MIMO systems.

The idea behind the capacity maximizing AS is to select the subset of antennas that achieve the maximum capacity with reduced dimensional systems. Different algorithms have been proposed, and some benchmark techniques are given in the next sections.

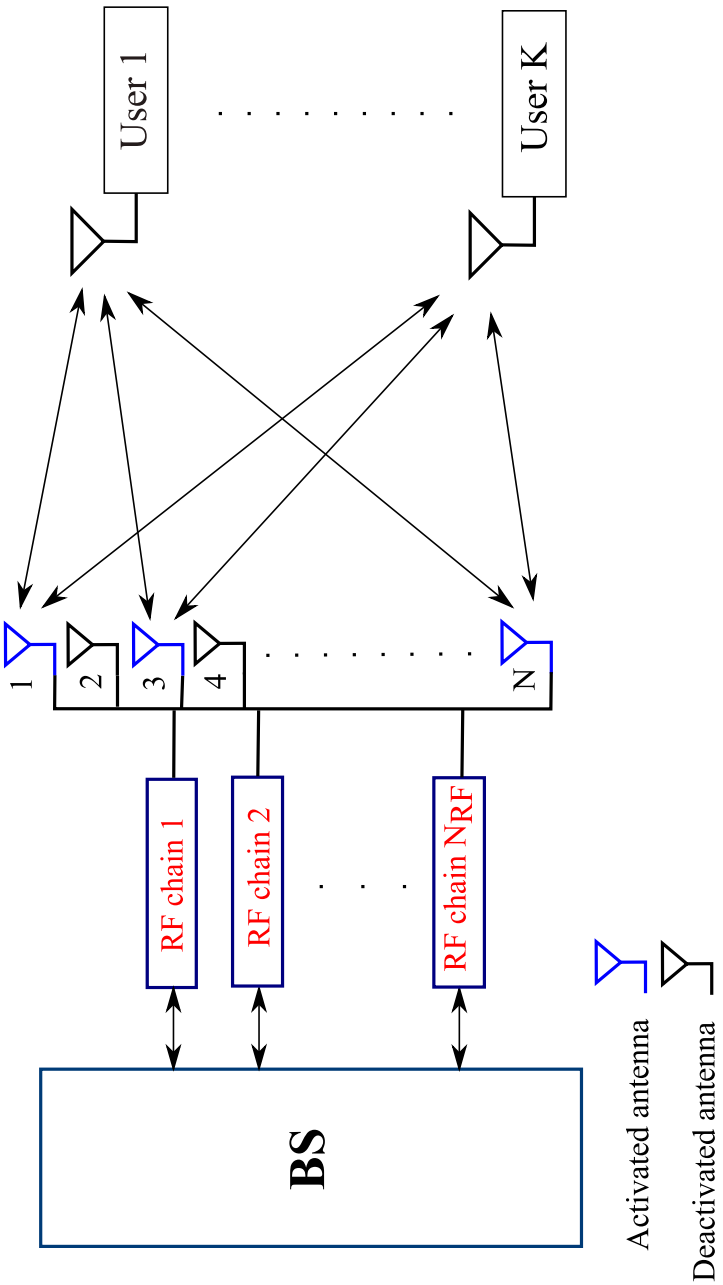


Figure 2.5: Antenna selection in Multi-user MIMO systems.

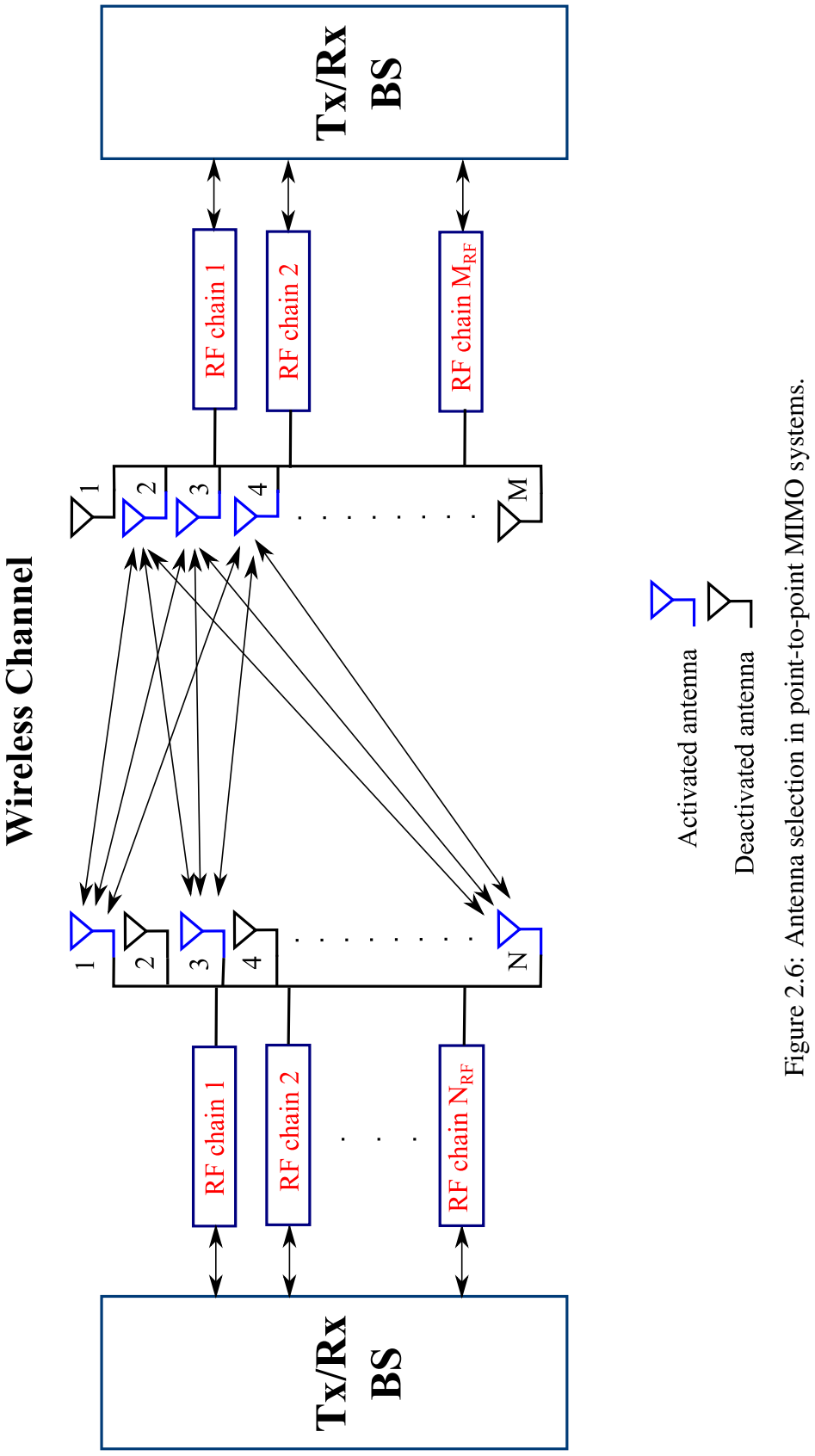


Figure 2.6: Antenna selection in point-to-point MIMO systems.

2.4.3.1 Optimal AS for capacity maximization

As mentioned earlier, optimal AS can be applied by evaluating the capacity for all different antenna subset combinations, which for an uplink system with K single antenna users and N antennas at the Bs, can be expressed as the following optimization problem

$$\begin{aligned} &\underset{\mathbf{\Delta}}{\text{maximize}} \quad \log_2 \det\left(\mathbf{I}_N + \frac{\gamma}{K}\,\mathbf{\Delta}\mathbf{H}\mathbf{H}^H\right) \\ &\text{subject to} \\ &\Delta_{n,n} \in \{0,1\}, \qquad (2.19a) \\ &\sum_{n=1}^{N} \Delta_{n,n} = N_s, \qquad (2.19b) \end{aligned}$$

where N_s is the number of selected antennas at the BS. Although this method provides the optimal solution, it has a binomial coefficient of $\binom{N}{N_s}$ different combinations, therefore it becomes prohibited for large values of N and N_s.

Two well known methods that have near optimal performance in terms of capacity, known as: Incremental AS and Decremental AS, were proposed by [12] for spatial multiplexing MIMO systems, and are explained in the following subsections.

2.4.3.2 Incremental AS for capacity maximization

The incremental AS method starts with an empty set of antennas, and then at each iteration, the antenna that will maximize the capacity will be selected. Assume that after n iterations, the antennas indexed with $\{s_1, s_2, ..., s_n\}$ have been selected, and $\mathbf{H}_{\mathcal{S}}$ represents the $n \times K$ submatrix of $\mathbf{H}$, where $\mathcal{S}$ is a set containing the indices of selected antennas at a given iteration. Appending the $(n^*)^{th}$ antenna yields to

$$\begin{aligned} C(\mathbf{H}_{\mathcal{S}}, \mathbf{h}_{n^*}) &= \log_2 \det\left(\mathbf{I}_K + \gamma(\mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}} + \mathbf{h}_{n^*}^H\mathbf{h}_{n^*})\right) \\ &= \log_2 \det\left(\mathbf{I}_K + \gamma\,\mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}}\right) \\ &+ \log_2\left(1 + \gamma\mathbf{h}_{n^*}(\mathbf{I}_N + \gamma\mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}})^{-1}\mathbf{h}_{n^*}^H\right). \qquad (2.20) \end{aligned}$$

Therefore, selecting the antenna n^* that will lead to maximizing the capacity can be expressed as

$$s_{n+1} = \arg\max_{n^* \notin \mathcal{S}} \mathbf{h}_{n^*}(\gamma^{-1}\mathbf{I}_K + \mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}})^{-1}\,\mathbf{h}_{n^*}^H, \qquad (2.21)$$

Algorithm 1 Incremental AS algorithm

Input N, N_s, γ, and $\mathbf{H}$,
Initialize
  $\mathbf{A} = \gamma \mathbf{I}_K$, $\mathcal{S} = \mathbf{0}_{1\times N}$,
$s_1 = \arg\max_{1\leq n^* \leq N} \|\mathbf{h}_{n^*}\|^2$,
$\mathcal{S}_{s_1} = 1$,
for $n = 1 \rightarrow N_s - 1$
  update $\mathbf{A} \leftarrow \mathbf{A} - \mathbf{A}\mathbf{h}_{s_n}^H (1 + \mathbf{h}_{s_n}\mathbf{A}\mathbf{h}_{s_n}^H)^{-1}\mathbf{h}_{s_n}\mathbf{A}$,
  $s_{n+1} = \arg\max_{n^* \notin \mathcal{S}=1} \mathbf{h}_{n^*}\mathbf{A}\mathbf{h}_{n^*}^H$,
  $\mathcal{S}_{s_{n+1}} = 1$,
end for
Output $[\mathbf{H}_j]_{j\in\mathcal{S}=1}$,

The selection process in (2.21) requires large number of matrix inversions, which leads to an extremely high complexity. Therefore, a recursive update approach based on the matrix inversion lemma can be applied to reduce the required complexity [76]. Let $\mathbf{A}$ denotes an $n \times n$ positive definite matrix, and $\mathbf{a}$ be an $n \times 1$ vector, then the inverse of $(\mathbf{A}+\mathbf{a}\mathbf{a}^H)$ can be computed as follow [12]

$$(\mathbf{A} + \mathbf{a}\mathbf{a}^H)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{a}(1 + \mathbf{a}^H\mathbf{A}^{-1}\mathbf{a})^{-1}\mathbf{a}^H\mathbf{A}^{-1}. \tag{2.22}$$

By employing the selection rule in (2.21) and the update in (2.22), the incremental AS method can be applied as shown in Algorithm 1. This algorithm is an attractive solution when the number of selected antennas is relatively small compared to the number of antenna at the BS. Furthermore, this algorithm is designed to maximize the sum rate capacity, which can only be obtained through SIC in the uplink, or DPC in the downlink scenarios. Therefore, for systems with MF detection/precoding, applying this algorithm will result in poor performance, since it ignores the interference which becomes the dominant degradation factor, especially at moderate to high SNRs.

As demonstrated in [44], when MF precoding scheme is utilized, there are antennas that will cause more interference than desired signal gain. Therefore, applying an algorithm that does not take the interference into account, such as the incremental AS described above, when MF is applied will not lead to the selection of the desired antennas, i.e. the antennas that lead to high SINRs.

2.4.3.3 Decremental AS for capacity maximization

When the number of selected antenna is close to the number of receive antennas at the BS, the decremental method becomes more attractive approach than the incremental selection. The

Algorithm 2 Decremental AS algorithm

Input N, N_s, γ, and $\mathbf{H}$,
Initialize
  $\mathbf{A} = (\gamma^{-1}\mathbf{I}_K + \mathbf{H}^H\mathbf{H})^{-1}$, $\mathcal{S} = \mathbf{1}_{1\times N}$,
$s = \arg\min_{1\leq n^*\leq N} \mathbf{h}_{n^*}\mathbf{A}\mathbf{h}_{n^*}^H$,
$\mathcal{S} = \mathcal{S} \setminus s$,
for $n = 1 \rightarrow N - N_s - 1$
  update $\mathbf{A} \leftarrow \mathbf{A} + \mathbf{A}\mathbf{h}_s^H(1 - \mathbf{h}_s\mathbf{A}\mathbf{h}_s^H)^{-1}\mathbf{h}_s\mathbf{A}$,
  $s = \arg\min_{n^*\in\mathcal{S}} \mathbf{h}_{n^*}\mathbf{A}\mathbf{h}_{n^*}^H$,
  $\mathcal{S} = \mathcal{S} \setminus s$,
end for
Output $[\mathbf{H}_j]_{j\in\mathcal{S}}$,

decremental AS methods starts with a full set of activated antennas. At each iteration, the decremental method deactivates one antenna that will lead to the minimum reduction in the system capacity. The implementation of the decremental AS is similar to the incremental AS, and it is shown in Algorithm 2. As for the incremental selection method, the decremental technique was also designed to maximize the sum rate capacity.

It was shown that the decremental AS methods outperforms the incremental AS technique. The reason behind that is with the decremental selection, the selection of the deactivated antenna takes into account the contribution of all remaining (activated) antennas, while in the incremental method, the selection in done taking into account only the contribution of the appended antenna [12].

In terms of complexity requirement, both methods, incremental and decremental AS, require relatively high complexity for systems with large dimensions. The incremental selection method requires an upper bound of NN_sK^3 complex multiplications/additions, while the decremental method suffers from higher complexity since it includes a matrix inversion at the start of the algorithm.

2.5 MIMO systems with massive antenna arrays

In conventional MIMO systems, the transmitter/receiver are equipped with few antennas only, and although it shows a decent performance gain compared to SISO systems, there is still space for dramatic improvements. To fulfil the demand for higher achievable rates, and simplifying the signal processing at the same time, MIMO systems with massive antenna arrays, also known as Massive MIMO, have been introduced as a potential technology for next generation of wireless communications systems.

M-MIMO is a scalable version of conventional MU MIMO systems, where the BS is equipped with very large number of antennas to serve multiple users utilizing the same time/frequency

resources. Employing massive antenna arrays can have great benefits, such as: 10 times or more increase in capacity, dramatic reduction in the radiated energy, and a near optimal performance can be obtained using linear signal processing techniques instead of more complex non-linear techniques [77, 78].

Fig. 2.7 illustrates the concept of M-MIMO systems setup. BSs are usually equipped with very large number of antennas N, serving K user terminals in the same time-frequency resources, $N \gg K$. Increasing the number of antennas while keeping the same number of users lead to a logarithmic increase in throughput, with linear increment in training time.

In the uplink scenario, the user terminals send their data to the BS simultaneously, and the BS needs to separate the received signal to recover the K data streams sent by the users. In contrast, the BS needs to apply channel precoding techniques before sending the information to the users, to ensure that each user only receives the data intended for it. In both cases, only the BS needs to have knowledge of CSI, to reduce the amount of processing at the user end.

2.5.1 Channel model in MU M-MIMO systems

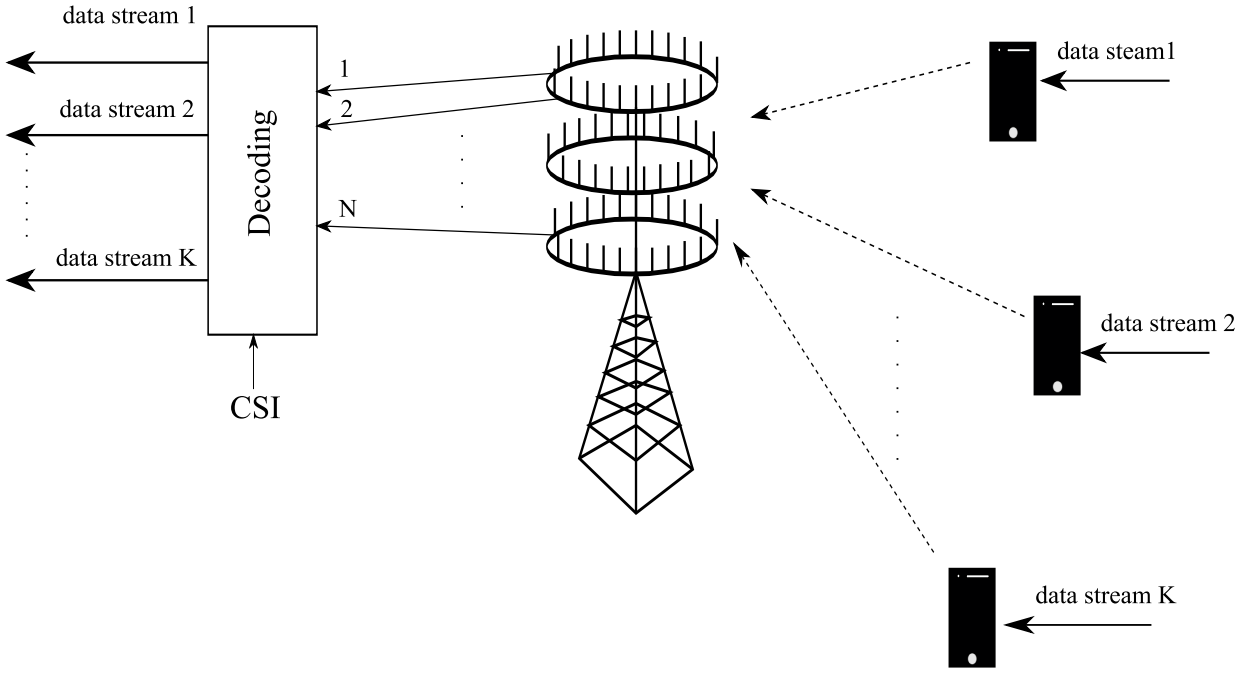
In a single-cell MU M-MIMO, where one BS is equipped with N antennas, serving K single-antenna users, the path gain between the k^{th} user and the n^{th} antenna at the BS consists of small and large scale components, and can be expressed as

$$h_{n,k} = g_{n,k}\sqrt{\beta_k}, \tag{2.23}$$

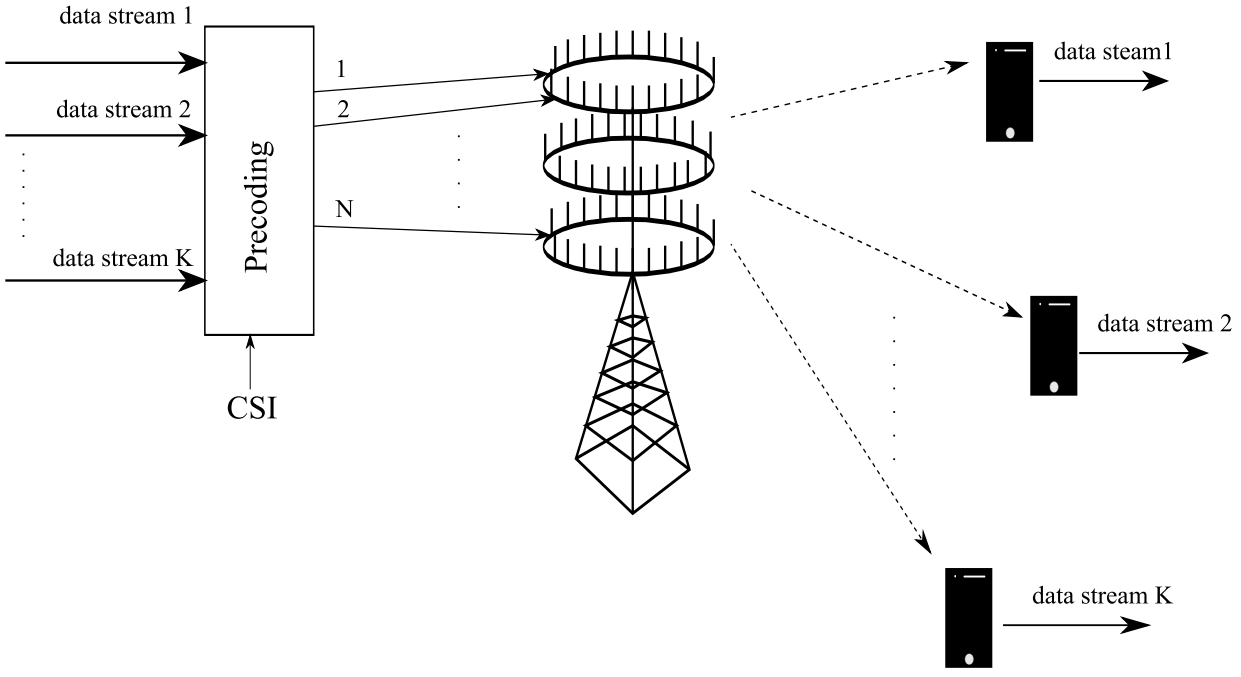
where $g_{k,n}$ is the complex small scale fading, and it is different between each antenna and the k^{th} user, while β_k represents the large scale fading component, and it is irrelevant of the antenna index at the BS for a given user, since the distance between a certain user and the BS is much greater than the distance between different antennas at the same BS. In a vectors form, the channel matrix can be represented as

$$\mathbf{H} = \begin{pmatrix} h_{1,1} & \dots & h_{1,K} \\ \vdots & \ddots & \vdots \\ h_{N,1} & \dots & h_{N,K} \end{pmatrix} = \mathbf{G}\mathbf{D}^{1/2}, \tag{2.24}$$

where



(a) Uplink MU M-MIMO



(b) Downlink MU M-MIMO

Figure 2.7: Multi-User Massive MIMO systems: (a) uplink scenario, (b) downlink scenario.

$$\mathbf{G} = \begin{pmatrix} g_{1,1} & \cdots & g_{1,K} \\ \vdots & \ddots & \vdots \\ g_{N,1} & \cdots & g_{N,K} \end{pmatrix}, \tag{2.25}$$

$$\mathbf{D} = \begin{pmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_K \end{pmatrix}. \tag{2.26}$$

Obtaining the knowledge of CSI is a crucial aspect in any communication system, and M-MIMO system is no exception. Two different schemes can be adopted for data transmission in M-MIMO. The first scheme is called *time division duplex*, where the uplink/downlink transmissions occur in different time instances, however, the bandwidth is completely utilized during the entire transmission interval. The second scheme, called *frequency division duplex*, utilizes different frequency bands for the uplink and downlink transmissions.

2.5.1.1 CSI in time division duplex scenarios

In order to perform a detection in the uplink, or a precoding in the downlink, the CSI needs to be known at the BS. To obtain the CSI, the BS needs to estimate the values of the channel gains between the users and its antennas. Most of the work on M-MIMO considers a time division duplex (TDD) scenario, where the channels are reciprocal in the uplink and downlink cases.

Fig. 2.8 illustrates the transmission protocol in TDD scenario for a single coherent interval, where the users send their data followed by training pilots used by the BS to estimate their channels. Unlike the data where it can be sent simultaneously by all users, the pilot signals needs to be orthogonal, so that the BS can have an accurate estimate of the channels for each user. Moreover, the length of the training depends on the number of users in the cell, and is irrelevant to the number of antennas at the BS. Once the BS have an estimate of the CSI, it can use the same CSI for downlink transmission due to the channel reciprocity property in TDD [79].

2.5.1.2 CSI in frequency division duplex scenarios

In frequency division duplex (FDD), the uplink and downlink transmissions occur on different frequency bands. Therefore, and unlike the TDD, the assumption of channel reciprocity does not hold in this case. The estimate of the uplink channels can be obtained in a similar manner

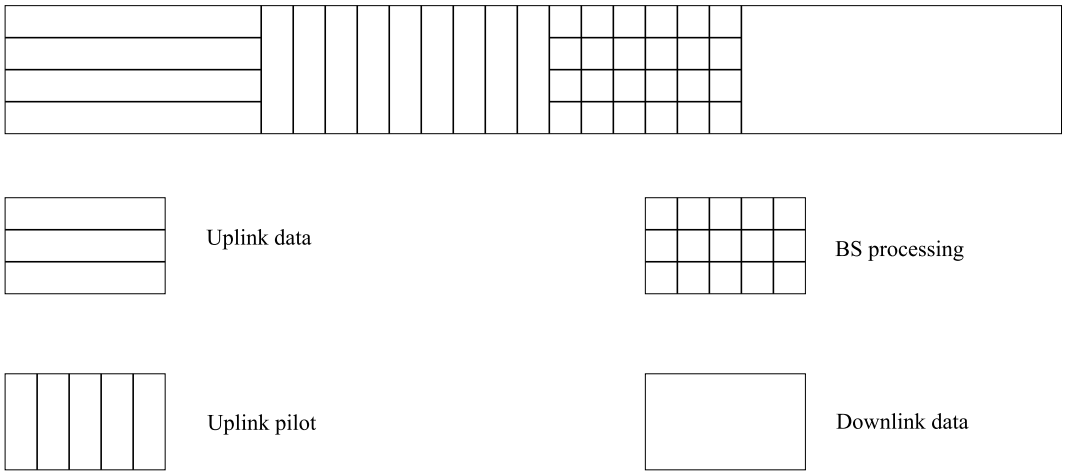


Figure 2.8: TDD protocol during one coherent interval.

to that in TDD scenario, however, the real challenge is to obtain CSI for the downlink channels. To achieve that, the BS needs to send pilot signals to the users, and then each user will estimate the channels and feed them back to the BS over a control channel.

2.5.1.3 TDD vs FDD in M-MIMO systems

Although TDD is more likely to be adopted in M-MIMO systems than FDD, the latter achieves higher rates. For example, assume that B is the system bandwidth, and P is the received power at the user terminal, while the noise spectral density is N_o. Utilizing TDD will lead to the following achievable rate, in bits/sec, for the downlink case [78]

$$R_{TDD} = \frac{B}{2}\log_2\left(1+\frac{P}{BN_o}\right), \tag{2.27}$$

where the division of the bandwidth by 2 is due to the fact that the transmission is occupying half of the time only. In contrast, for the same specifications, the achievable rate for FDD can be expressed as [78]

$$R_{FDD} = \frac{B}{2}\log_2\left(1+\frac{P}{(B/2)N_o}\right), \tag{2.28}$$

where the division of the bandwidth by 2 outside the logarithm and inside it is due to the fact that only half the frequency is utilized for downlink transmission.

2.6 Challenges in M-MIMO systems

Although M-MIMO has significant advantages, there are many factors that limits either the performance or the application of such systems and need to be considered. In the following, we discuss the main challenges in M-MIMO systems.

2.6.1 Complexity requirement with massive arrays

As explained earlier in this chapter, the use of large number of RF chains impose a real challenge in M-MIMO. Although the radiated power can be decreased dramatically in M-MIMO, the energy efficiency suffers from large degradation due to the power consumed by the RF chains. In addition, the use of large RF chains results in extremely high signal processing at the BS. Therefore, using limited number of RF chains can improve the EE, cost, and reduce the amount of data processing. The use of limited RF chains means that the asymptotic advantages of M-MIMO can no longer be obtained. However, applying AS or SM with M-MIMO are far superior to any conventional MIMO systems.

2.6.2 Equality for all users

In realistic scenarios, the number of antennas at the BS, even with M-MIMO, are limited. Therefore, users who are located at the cell-edge might suffer from poor service due to the high interference and weak signal gain. Although PA techniques can be utilized to enhance the service of the worst-case users, the performance can still be improved for all users.

Recent work in M-MIMO have proposed the concept of cell-free systems to tackle this issue [80–84]. In cell-free, as illustrated in Fig. 2.9, the antennas are distributed in the cell as access points, where each access point will have one or few antenna elements only. In this case, there are no boundaries for the cell, and everyone can enjoy the high quality service. However, cell-free MIMO suffers from other challenges that do not exist in centralized M-MIMO, such as the synchronization of the access points as well as higher complexity.

2.6.3 Channel estimation in FDD systems

As mentioned earlier in this chapter, the uplink channels in FDD systems are estimated via orthogonal pilots by the users. However, in the downlink scenario, the BS sends pilot signals to the users, and the estimation will be fed back over a control channel to the BS. The main

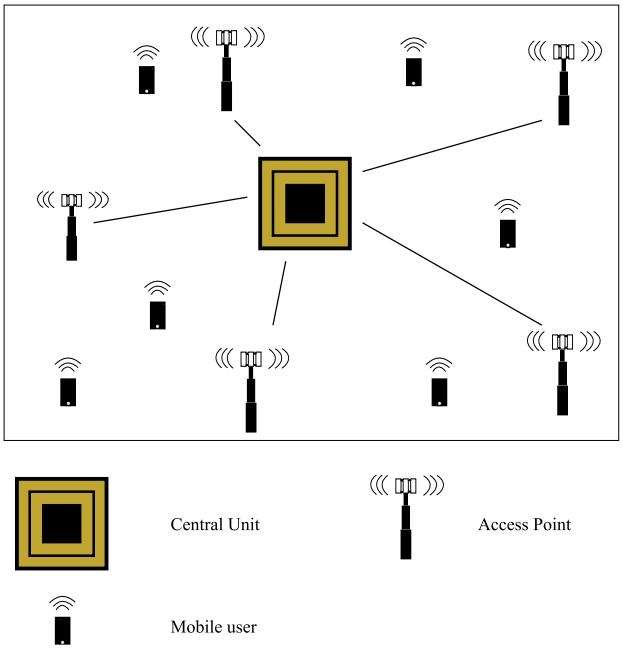


Figure 2.9: Cell Free M-MIMO systems.

disadvantage of this method is that the pilots sent by the BS needs to be orthogonal, and therefore as the number of antennas grows, most of the coherent interval (the interval over which the channel is considered to be flat in both time and frequency), will be occupied only for training purposes, which leads to an extremely short interval for data transmission.

To discuss the CSI in FDD with more details, assume that the BS is equipped with N antennas, and K single-antenna users are being served under FDD scenario. In the uplink, the minimum number of orthogonal pilots that needs to be sent is K, in addition to N coefficients which are fed back by the users to inform the BS about the downlink channels. In the downlink case, the BS needs to send N orthogonal pilots to the users to perform the CSI estimation in the downlink case. In contrast, TDD only requires K orthogonal pilots, which is not only much smaller than the number of pilots in FDD, but also is irrelevant of the number of antennas at the BS. However, the applications of FDD in M-MIMO might be feasible in certain circumstances. For example, the authors in [85] justified the use of FDD in M-MIMO and demonstrated its feasibility under certain conditions such as the knowledge of channel covariance matrix.

2.6.4 Pilot contamination in TDD systems

One of the main challenges in M-MIMO is the pilot contamination. As discussed earlier in this chapter, and for TDD scenarios, the users need to send orthogonal pilots to the BS in order to obtain CSI. For multi-cell scenario, ideally, the users in neighbouring cells need to send orthogonal pilots as well to eliminate any interference between the cells. For example, Let $\boldsymbol{\psi}_{k,l} = \{\psi_{k,l}^1, \ldots, \psi_{k,l}^\tau\}$ be a pilot sequence of length τ from the k^{th} user to the BS in cell l, assuming orthogonal pilots between users in different cells leads to

$$\boldsymbol{\psi}_{k,l}^H \boldsymbol{\psi}_{m,j} = \delta[k-m]\ \delta[l-j], \tag{2.29}$$

where $\delta[i]$ is equal to 1 if $i = 0$, and 0 otherwise. However, assuming orthogonal pilots within the same cell as well as neighbouring cells will greatly limits the number of users that can be served in M-MIMO systems [77]. Therefore, the pilots are re-used in neighbouring cells to serve larger number of users, which leads to degradation in the channel estimation accuracy. This phenomenon is known as *pilot contamination*, which is illustrated in Fig. 2.10, and its effect can not be eliminated by increasing the number of antenna elements at the BS.

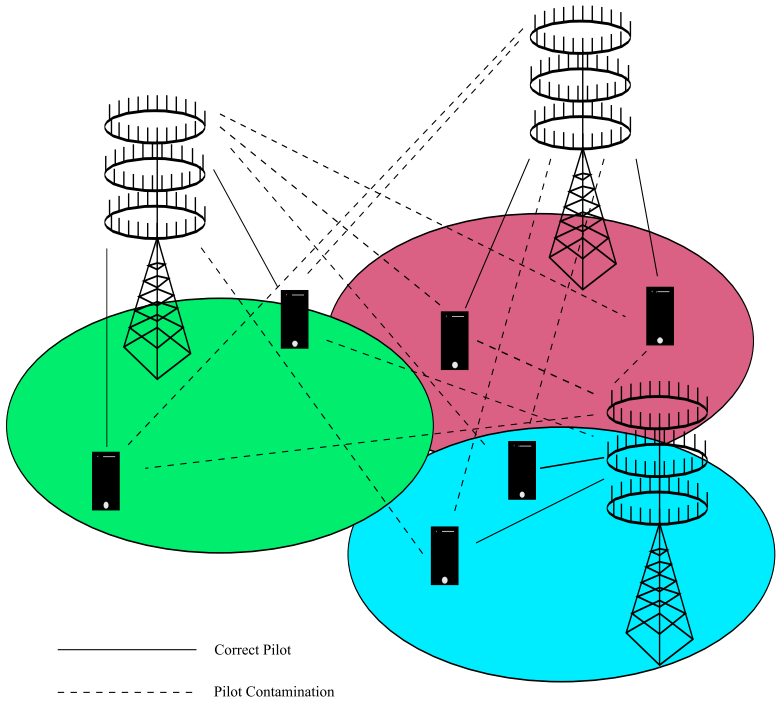


Figure 2.10: Illustration of pilot contamination in M-MIMO.

2.7 Chapter Summary

In this chapter, we presented the different types of conventional MIMO systems, and the common linear and non-linear techniques used for signal processing, in terms of both detection and precoding. Moreover, the advantages of employing systems with limited number of RF chains are justified. Accordingly, SM, GSM, and AS were explained in details alongside with their advantages and challenges, where it was shown that SM and GSM achieve higher throughputs at the cost of increased detection complexity. In contrast, AS can maintain low complexity requirements at the receiver end, and two well-known near-optimal AS algorithms were presented for conventional MIMO systems. In addition, MU M-MIMO systems were explained as a tool to meet the demand of achieving much higher throughputs compared to conventional MIMO systems. Finally, the main advantages and challenges of M-MIMO systems were explained in details for both FDD and TDD scenarios.

Chapter 3

Evolutionary optimization algorithms for Antenna Selection in M-MIMO Systems

3.1 Introduction

Since mid-twentieth century, several scientists have suggested that evolution could be utilized as an optimization tool for solving different engineering problems [86]. In this chapter we apply different evolutionary algorithms for AS in M-MIMO system, these algorithms can be divided into three different categories, which are: bio-inspired, heuristic, and quantum-inspired techniques.

3.1.1 Bio-inspired optimization

Bio-inspired techniques have been used extensively to solve several engineering problems. The main idea behind any bio-inspired algorithm is to evolve a population of candidates towards a better solution for a given problem, by using operators inspired by natural selection. Three bio-inspired algorithms are applied in this chapter for AS in M-MIMO systems, and they are: Genetic Algorithm, Particle Swarm Optimization, and Artificial Bee Colony.

3.1.1.1 Genetic algorithm

Genetic Algorithms (GAs) have been used since the 1950s. One of the first people who worked on these algorithms and also had the most influence on this field than any other was John Holland of the University of Michigan [87]. Holland represented GA as a method for moving from a certain population of "chromosomes" to another population by using genetics-inspired operations such as $Crossover$, $Reproduction$, and $Mutation$ [86].

GA have been used as an optimization tool for different engineering-related problems. For example, in [88–93] the authors have applied GA for AS in MIMO systems, while in [94], GA was used for joint precoding and AS. Moreover, the authors in [95] have studied the convergence performance of MIMO detection utilizing GA against ML approach.

3.1.1.2 Particle swarm optimization

Particle Swarm Optimization (PSO) algorithm was first developed by [96], and is a class of evolutionary algorithms based on the intelligent behavior of biological organisms. The term "swarm" refers to a collection of interacting agents. For example, a flock of birds can be thought of as a swarm whose individual agents are birds, or a crowd is a swarm whose agents are people, and so forth [87].

Similar to GA, PSO have been applied by researchers to tackle engineering optimization problems. In [97] a modified PSO algorithm was designed for adaptive filtering, while in [98] PSO was used for user scheduling in downlink MIMO systems. Moreover, in [99] the authors designed a two-stage PSO algorithm for transmit beamforming. Finally, in [100–103] the authors applied PSO algorithms for AS in MIMO systems.

3.1.1.3 Artificial bee colony

Artificial Bee Colony (ABC) is among the recent bio-inspired optimization algorithms, and was developed by [104] to tackle optimization problems based on the intelligent behavior of honey bees on finding food sources. In ABC algorithm, the colony of artificial bees consists of three different types of bees, and they are used to search for the best solution. These bees are: Employed, Onlooker, and Scout bees [105].

Since the ABC algorithm has been recently proposed, it has not been widely explored yet, however, the authors in [106] and [107] employed the ABC algorithm for Peak to Average Power Ratio (PAPR) reduction and interference cancellation, respectively, in MIMO-OFDM systems, while in [108], the authors applied ABC algorithm for finding a suitable number of active users as well as optimizing the size of antenna elements for energy-efficient M-MIMO systems.

3.1.2 Quantum and classical tabu search algorithms

The term Tabu Search (TS) refers to a list of moves or actions that are tabooed for a certain number of iterations. However, different actions are tabooed in classical and quantum-inspired

methods as explained in the following sections.

3.1.2.1 Classical tabu search

Classical Tabu Search (CTS) is a metaheuristic search method, and it was first proposed by Glover in [109]. CTS is a local search algorithm, that uses the memory to avoid revisiting the previous moves (or solutions) to ensure an efficient search of the neighborhood. A special matrix, called Tabu matrix, is used to save the previous visited solutions in the neighborhood, and forbid using them for a certain number of upcoming iterations.

CTS has been widely applied for different optimization metrics in communications systems, for example in [110] the authors proposed a CTS for beamforming in Millimeter-Wave M-MIMO systems. While in [111] and [112] CTS method was proposed for detection in M-MIMO, furthermore, the authors in [113] utilized CTS approach for detection in conventional MIMO systems.

3.1.2.2 Quantum-inspired tabu search

Quantum-inspired Tabu Search (QTS) is an evolutionary algorithm proposed by [114] to solve the 0/1 Knapsack problem. It takes the concept of tabu search and combines it with the characteristics of quantum computing, such as superposition. The QTS algorithm does not use memory to prevent certain moves or solutions. Instead, at each iteration, it prevents applying the unitary operator on certain quantum bits, referred to as *qubits*, when a predefined condition is met.

different Quantum-inspired algorithms have been proposed recently, for example in [115] the authors applied a Quantum-inspired evolutionary algorithm for detection in M-MIMO systems, while the authors in [116] applied a Quantum-inspired algorithm for user selection and PA in cognitive MIMO systems.

3.2 Motivation and contributions

As demonstrated in Chapter 2, exhaustive search method can not be applied for AS in M-MIMO systems due to its enormous complexity requirement. In contrast, evolutionary algorithms have been widely applied as low complexity optimization tools and can achieve near optimal performance. Therefore, the motivation in this chapter is to apply and design different evolutionary algorithms for AS in MU M-MIMO systems operating in the uplink scenario.

Our novel contributions in this chapter can be summarized in the following list

1. We develop an ABC, a CTS, as well as a QTS algorithms for AS in M-MIMO systems. We assume an uplink transmission scenario, and the cost function under optimization is the sum rate capacity.

2. The optimization parameter for the QTS method, which is the rotation angle (θ) that controls the evolution process of the QTS algorithm, is also found by exhaustive search from 0^o to 180^o for the AS problem. The values of θ from 180^o to 360^o will give identical results, since only the phase will be changing.

3. The complexity of the different algorithms was measured in terms of the CPU time required under different number of antennas and iterations.

4. The developed algorithms were compared with other well-known evolutionary algorithms found in the literature, namely the GA and the PSO algorithms, and it was demonstrated that classical and quantum-inspired tabu search algorithms outperform bio-inspired algorithms in terms of both performance and complexity.

5. For the QTS algorithm, the effect of the rotation angle θ on the system capacity is demonstrated for different initializations of α^2 and β^2, and conclusions were drawn based on the observations. Moreover, the effect of changing the number of transmit, receive, and selected antennas on the optimum rotation angle was also tested.

6. Finally, the convergence behaviour of the applied AS algorithms was also studied, and it was demonstrated that QTS algorithm is more flexible than the classical tabu search as well as the bio-inspired algorithms, since it only requires finding the optimum rotation angle to evolve the system towards better solutions.

3.3 System Model

Consider an uplink M-MIMO system working in the TDD mode, where the users send an orthogonal pilots to the BS to obtain CSI. It should be noted that here we assume perfect CSI at the BS. Moreover, the BS is equipped with N_r receive antennas[1], while N_t represents the number of transmit antennas ($N_r \gg N_t$). This system can be represented by the following equation

[1]Here we assume that the N_r antennas at the BS experience spatial correlation due to the limited spacing between them.

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{w}, \tag{3.1}$$

where $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ is the transmitted signal vector, $\mathbf{G} \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix, $\mathbf{w} \in \mathbb{C}^{N_r \times 1}$ is the additive Gaussian noise with zero mean and variance of σ_n^2, and $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ is the received signal vector. Throughout this work, we consider spatial correlation between the antennas at the BS. The adopted correlation model is introduced and explained in the following Subsection.

3.3.1 Spatial correlation channel model

The correlated channel matrix $\mathbf{G}$ in (3.1) can be described using the Kronecker model as follows [117]

$$\mathbf{G} = \mathbf{R}_R^{1/2}\mathbf{H}\,\mathbf{R}_T^{1/2}, \tag{3.2}$$

where $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is a Gaussian matrix, with coefficients assumed to be independent and identically distributed (i.i.d.), with zero mean and unit variance. $\mathbf{R}_R$ and $\mathbf{R}_T$ are the receive and transmit correlation matrices, respectively. It should be clarified that the operator $(.)^{1/2}$ in (3.2) represents the Hermitian square root of a matrix. In this chapter, we consider a multi-user uplink scenario, where there is no correlation between the users, while the antennas at the BS are spatially correlated due to the limited spacing between them. The spatially correlated channel matrix can then be given as

$$\mathbf{G} = \mathbf{R}_R^{1/2}\mathbf{H}. \tag{3.3}$$

The model of the $N_r \times N_r$ correlation matrix was assumed to have exponential correlation structure, which is a common model and can effectively measure the level of spatial correlation [117]. In this model, the correlation matrix can be implemented using only one coefficient $\rho \in \mathbb{C}$ with $|\rho| \leq 1$ as follows

$$R_{mn} = \begin{cases} \rho^{|m-n|} & , m \geq n \\ (\rho^{|m-n|})^* & , m < n, \end{cases} \tag{3.4}$$

where R_{mn} is the correlation between the m^{th} and n^{th} receive antennas, and $|\ .\ |$ is the absolute

value operator.

3.4 Antenna selection problem formulation and algorithms

We consider a BS with massive number of antenna elements ($N_r \geq 100$), and choose the best subset of these antennas to maximize the system capacity. For a MIMO system, the capacity can be given using the following equation [79]

$$\mathbf{C} = \log_2 \left(\det \left(\mathbf{I}_{N_r} + \frac{\gamma}{N_t} \mathbf{G}\mathbf{G}^H \right) \right), \tag{3.5}$$

where $\mathbf{I}_{N_r}$ is the $N_r \times N_r$ identity matrix, γ is the SNR, and $\mathbf{G}^H$ is the Hermitian (conjugate transpose) of the channel matrix.
Out of the available N_r antennas at the BS, we employ the optimization algorithms to choose the best N_s antennas that can maximize the capacity.

For simplicity, we will define the antenna selection operator as

$$\mathbf{s} = [s_1, s_2, ...s_{N_r}], \tag{3.6}$$

where

$$s_i = \begin{cases} 1 & \text{if the antenna is selected} \\ 0 & \text{Otherwise .} \end{cases} \tag{3.7}$$

At first, $\mathbf{s}$ is initialized with zeros, and once the optimization algorithm choose the best antenna subset, the location of these antennas will become 1s, while the rest of the elements will remain 0s. The optimized capacity can be then calculated as

$$\mathbf{C} = \log_2 \left(\det \left(\mathbf{I}_{N_r} + \frac{\gamma}{N_t} \mathrm{diag}(\mathbf{s}) \mathbf{G}\mathbf{G}^H \right) \right), \tag{3.8}$$

where $\mathrm{diag}(\mathbf{s})$ is an $N_r \times N_r$ diagonal matrix with $\mathbf{s}$ is its diagonal entry. Before we introduce the different AS algorithms, the basics of Quantum computing, which the QTS algorithm relies on, are introduced in the next section.

3.5 The basics of quantum computing

3.5.1 Quantum bit or "qubit"

In classical computers, the smallest unit of information is the bit, and it occurs in one of two possible states, i.e. "0" or "1". The analogous concept of a classical bit in a quantum computer is the quantum bit, or "qubit" [118]. The qubit can be in the $|0\rangle$ state, $|1\rangle$ state, or in a linear combination, i.e. superposition of the two states

$$|\psi\rangle = \alpha\,|\mathbf{0}\rangle + \beta\,|\mathbf{1}\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \tag{3.9}$$

where α and β are real or complex numbers, and $|\alpha|^2$ and $|\beta|^2$ represent the probabilities of the qubit to be found in the "0" and "1" states, respectively. Moreover, following the rule of probability, it should be noted that $|\alpha|^2 + |\beta|^2 = 1$. The Dirac notation $|.\rangle$, pronounced as "ket", is the standard notation in quantum mechanics, and it is used to represent the state of qubits [119]. When measuring the qubit, it can only give a "0" or "1", which means that the measurement process will destroy any superposition states and the system will collapse into one state only [118].

3.5.2 Multiple qubits

For a system with a string of n qubits, known as Q-bit individual (Q) [120], it can be represented in the following form

$$Q = \left[\begin{array}{c|c|c|c} \alpha_1 & \alpha_2 & ... & \alpha_n \\ \beta_1 & \beta_2 & ... & \beta_n \end{array}\right], \tag{3.10}$$

where $|\alpha_i|^2 + |\beta_i|^2 = 1$, for $i = 1, 2, ..., n$. In quantum computing, a system with n qubits can represent 2^n states at the same time before performing the measurement process. While in a classical system, a string of length n bits can only represent one of the 2^n possible solution. This is the key point that highlights the powerfulness of quantum systems.

3.5.3 Evolution of quantum systems

In Quantum systems, all the candidate (possible) solutions have the same probability of observation at the beginning. The main goal is to iteratively evolve the system using certain operations such that when performing the measurement process, the desired solution will be observed with high probability, since the measurement process will destroy any superposition state.

Unitary operators, also known as quantum gates, are applied to evolve quantum systems. A unitary operator can be defined as a matrix that obeys the condition $\mathbf{U}^{\dagger}\mathbf{U} = \mathbf{I}$. Where $\mathbf{U}^{\dagger}$ is the adjoint (conjugate transpose) of $\mathbf{U}$, and $\mathbf{I}$ is the identity matrix. There are many unitary operators that can be used to evolve a quantum system [121], such as the Hadamard gate, bit-flip gate, phase-flip gate, and the rotation operator. However, in our work, we apply the rotation operator $\mathbf{R}_{\theta}$ [118]:

$$\mathbf{R}_{\theta} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \tag{3.11}$$

$$|\psi\rangle' = \mathbf{U}\,|\psi\rangle, \tag{3.12}$$

where $|\psi\rangle'$ is the new quantum state after applying the unitary operator. It is worth mentioning that applying the unitary operator will not affect the sum of the probabilities of any qubit in the Q-bit individual. For a single qubit, the effect of applying the rotation operator can be seen in Fig. 3.1, where the new qubit state is a rotated version of the original qubit state by an angle of θ in the anticlockwise direction if θ was positive, and in the clockwise direction if θ was negative. However, this geometrical representation is only valid when α and β are both real numbers.

In the following sections, we will introduce and explain the different evolutionary algorithms adopted in this chapter to tackle the optimization problem of AS in MU M-MIMO system.

3.6 Bio-inspired algorithms for antenna selection in M-MIMO systems

As mentioned earlier, three bio-inspired methods were used in this chapter for AS in M-MIMO systems, which are: PSO, GA, and ABC, and they are explained in details in the following subsections.

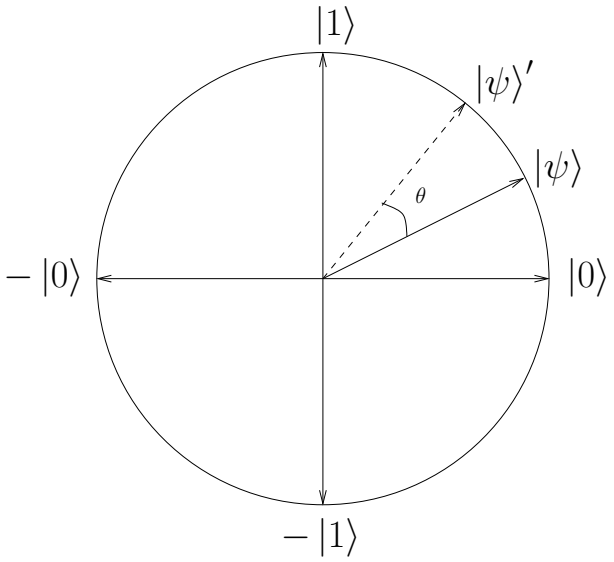


Figure 3.1: The effect of $\mathbf{R}_\theta$ on a single qubit

3.6.1 PSO algorithm for AS

At first, a certain number of particles (i.e. possible solutions for the AS problem) are generated randomly. Each possible solution can be represented as a vector of length N_r with N_s number of 1s located randomly along the vector. The capacity represented by the fitness value is measured for each possible solution. Additionally, the velocity is calculated, which directs the particle to fly towards the best solution. In PSO, every particle is influenced by its neighbors (called $local\ best$) as well as by the best particle among the group (called $global\ best$). The velocity of the particles is given as

$$\begin{aligned} v_i(t) = v_i(t-1) + \ & rand_1 \times k_1(p_{li} - f_i(t-1)) \\ + \ & rand_2 \times k_2(p_{gi} - f_i(t-1)), \end{aligned} \tag{3.13}$$

where $v_i(t)$ represents the velocity of the current iteration for the i^{th} antenna, and $v(t-1)$ is the velocity of the previous iteration. $rand_1$ and $rand_2$ are random numbers drawn from a uniform distribution between 0 and 1, and manipulating these values can have an effect on the convergence behaviour of the algorithm. k_1 and k_2 are weighting factors with arbitrary values, they are assumed to have a value of 2 in our simulations. p_{li} represents the local best solution for the i^{th} antenna depending on the two neighbors of the current particle. In our simulations,

the first and last particles were assumed to be connected, i.e. the neighbors of the first particle are the second and last particle of the population. Moreover, the neighbors of the last particle are the first and the second last particle in the population. p_{gi} is the global best solution among the whole population for the current antenna, and finally, $\mathbf{f}$ is a vector initialized with zeros, and updated after each iteration as follow

$$f_i(t) = f_i(t-1) + v_i(t). \tag{3.14}$$

After each iteration, the global best and the local best solutions are updated before the next iteration starts. Furthermore, at the end of the iterations, the N_s maximum values of $\mathbf{f}$ will be chosen as the surviving antennas, while the rest of antennas will be ignored.

Algorithm 3 PSO AS Algorithm

Input N, N_r, $max\ iterations$,
Initialize $\mathbf{f}_n = \mathbf{0}_{N_r \times max\ iterations}$, $n = 1 : N$,
Generate N random particles,
Measure the capacity for the N particles,
$iter \leftarrow 1$,
while $iter \leq max\ iterations$
    Find $local\ best$ and $global\ best$ solutions,
    Measure the velocity for each antenna in each particle,
    Update $\mathbf{f}$,
    Select the highest N_s values of $\mathbf{f}$ for each particle,
    Measure the modified capacity for all the particles,
    $iter = iter + 1$,
end while

3.6.2 GA algorithm for antenna selection

At the beginning, a certain number of chromosomes (i.e. possible solutions for the AS problem) are generated randomly. Every possible solution can be represented as a vector of "genes" (in this case bits, 0s or 1s), and the number of 1s in each chromosome equal to N_s. The fitness value for each chromosome (i.e. possible solution) will then be calculated and the best K chromosomes will be chosen to a mating pool for the reproduction process.

3.6.2.1 Reproduction process

In the reproduction process, the best K chromosomes will be paired off randomly into pairs of chromosomes, these chromosomes will then go through certain operations to produce a new

population of chromosomes. In our simulations, K was equal to half the number of chromosomes.

3.6.2.2 Crossover process

In this process, a mask of length N_r is generated randomly with values of 0s and 1s, where the probability of each bit being 0 is equal to the probability of being 1 (50% each). In the $crossover$ process, and for each gene of the chromosome, if the values for the two chromosomes in each pair were not equal, and the value of the mask was 1 at the location of the current gene, then the two chromosomes will exchange their genes to produce a new chromosome.

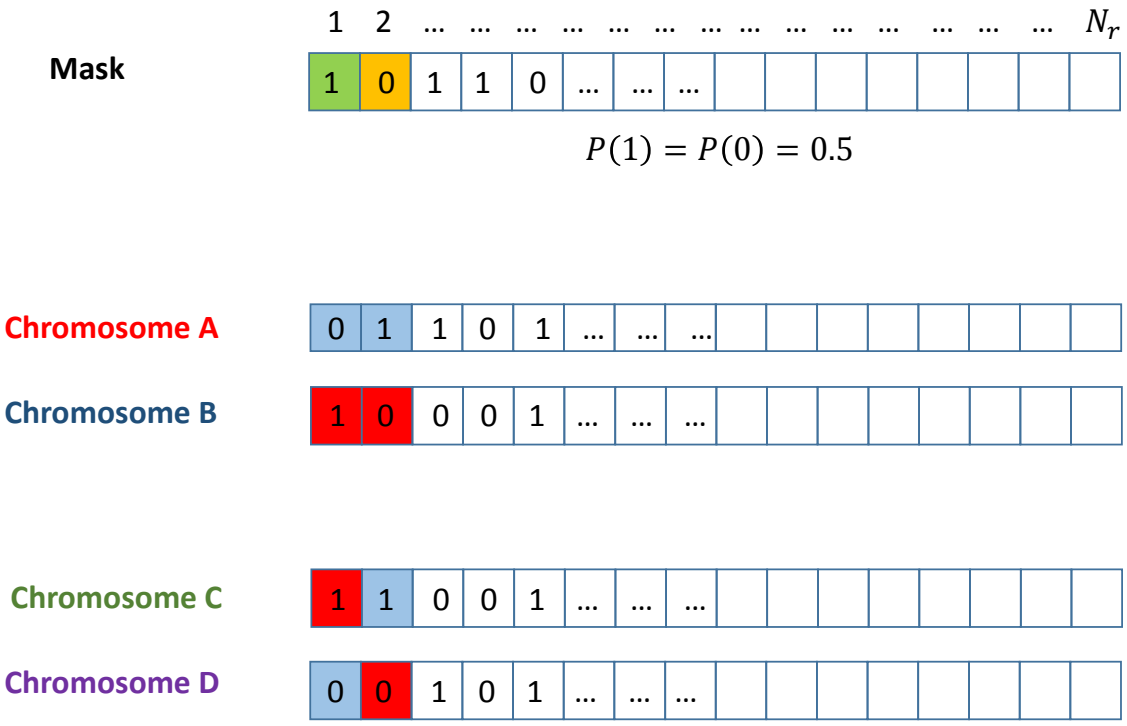


Figure 3.2: Crossover process.

However, this might cause a problem, since the total number of 1s in the new chromosome might be less or more than N_s. To overcome this issue, after generating each chromosome, the number of 1s within this chromosome will be checked. If it is less than N_s, then random locations of the chromosome will change their values from 0 to 1, until the total number of 1s is

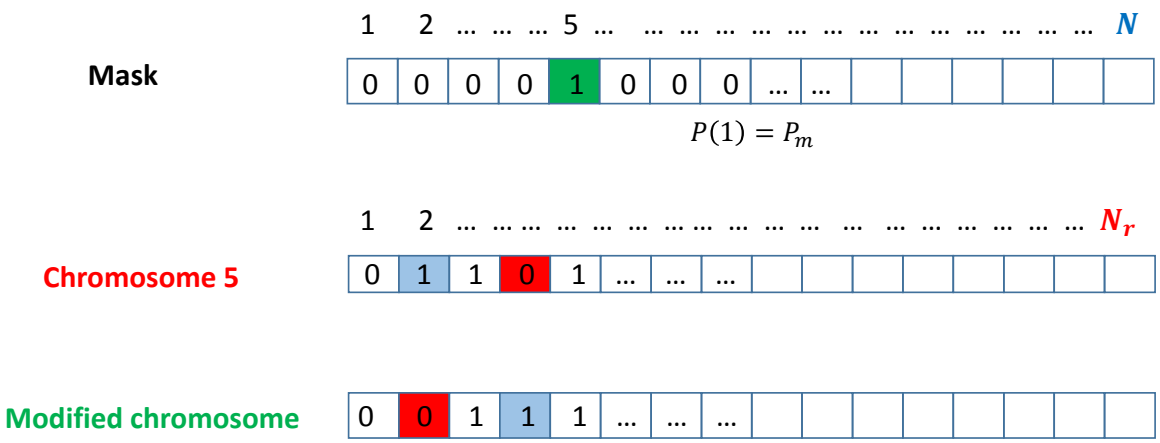


Figure 3.3: Mutation process.

equal to N_s. In contrast, if the total number of 1s within any generated chromosome is greater than N_s, then random genes will be ignored so that the total number of 1s in any chromosome will be equal to N_s.

3.6.2.3 Mutation process

The last process of the GA algorithm is the $Mutation$ process, where a mutation mask will be generated that consists of 0s and 1s according to the mutation probability P_m. In our simulation, P_m was set equal to 0.09, if the element of the mask was equal to 1. Subsequently, two random genes in the corresponding chromosome will exchange their information. If the two genes have the same information, implying that both of them were zeros or ones, then the chromosome will remain the same after the $mutation$ process.

After finishing all the steps, the fitness value will be calculated for the new population and the best K chromosomes will go through the same process in the next iteration until the maximum number of iterations has been reached. In the final step, the chromosome with the highest fitness value will be chosen.

Algorithm 4 GA AS Algorithm

1: **Input** : N, K, and $Max\ iter$,
2: Generate N chromosomes randomly,
3: **for** $iter$ = 1 : $Max\ iter$
4: Evaluate the capacity for the N chromosomes,
5: Select the best K chromosomes for the Reproduction process,
6: Apply Crossover process for the selected chromosomes,
7: Repair the new chromosomes,
8: Apply Mutation process,
9: **end for**

3.6.3 ABC algorithm for antenna selection

In the ABC algorithm, every bee represents a possible solution for the AS optimization problem. There are three different types of bees used in this algorithm, they are: Employed bees (EB), Onlooker bees (OB), and Scouts. At first, a certain number of EB (initial solutions) are generated randomly and their nectar amount, which is the capacity in this case, is measured. The total number of solution, or food sources, is equal to the number of EB. Every solution can be represented as a vector of length N_r, which is the number of parameters ($0s$ and $1s$) in the solution, and the number of $1s$ in any solution is equal to N_s. These bees share their nectar amount with the bees waiting on the dance area in the hive. Every EB will return to the same food position visited by itself after sharing its nectar amount, and modifies its solution by changing the parameters randomly, i.e. changing the location of the $0s$ and $1s$, then measures the modified fitness value. If the value of the modified solution for every bee is better than the previous one, then the bee will forget its old solution and memorizes the position of the new food source, otherwise the bee will return to the initial position.

The OB will then choose a food source (solution) depending on the nectar amount measured by the employed bee by using the following equation

$$\text{Source}_i = \frac{f_i}{\sum_{n=1}^{N} f_n}, \tag{3.15}$$

where f_i represents the fitness value of the source i, and N is the total number of possible solutions (EB). Once the OB has chosen the food source, then it will try to improve its solution using other food sources by the following equation

$$v_{ij} = |x_{ij} - x_{kj}|\,, \tag{3.16}$$

where $k \in \{1, 2, ...N\}$ is a randomly selected index, and $i \neq k$. $j \in \{1, 2, ...N_r\}$ represents the parameter index of the food source i. The total number of OB is equal to the number of solutions N.

This modification on the food sources might cause the problem of having more or less than N_s number of $1s$ in any solution. To tackle this issue, random parameters will be chosen to change their values so that the total number of $1s$ is equal to N_s in all the food sources. After that the fitness value of the modified solution will be calculated, and if it shows an improvement compared to the old solution, then it will memorize the modified solution, otherwise, the old solution will be used.

Finally, in order to search the area for the best food source and not getting stuck in limited number of solutions, one scout bee will be sent at each iteration to perform random search and calculate the fitness value and compare it with the worst solution in the population, if it was better then that food source will be replaced with the new food source found by the scout bee, otherwise the population will remain the same without any further changes for the next iteration.

For the next iteration, the OB will modify the solutions provided by the EB and a scout bee will be sent to perform a random selection until a certain number of iterations has been reached, and the best food source will be chosen for the antenna selection operation.

Algorithm 5 The ABC Algorithm

Input: N, $maximum\ iterations$,
Generate N possible food sources randomly for the N Employed bees (EB),
Measure the nectar amount, and share it with the Onlooker bees (OB) at the hive,
Each EB modifies its solution and measure the modified nectar amount, every EB keeps the best food source only,
$iter \leftarrow 1$,
while $iter \leq maximum\ iterations$
    Every OB chooses a food source based on Eq. (3.15),
    Every OB modifies its solution according to Eq. (3.16),
    Repair the modified solutions, and measure their fitness values,
    Compare the old solutions with the modified ones, keep the one with highest capacity,
    Send one Scout bee (SB) and compare its nectar amount with the worst solution, and keep it if it shows improvement,
    $iter = iter + 1$,
end while

3.7 Metaheuristic methods for AS in M-MIMO systems

The second category of the evolutionary methods is the heuristic techniques used in this chapter, which are: CTS and QTS. In the following, the two methods are introduced and explained in details for the AS in M-MIMO systems. It should be noted that, to the best of our knowledge, both of these methods have not been applied for AS before.

3.7.1 CTS algorithm for antenna selection

In the CTS algorithm, an initial solution will be generated randomly. This solution can be represented as a vector of length N_r with N_s number of $1s$ and the rest are $0s$. At each iteration, a certain number of neighbors will be generated and their fitness values will be calculated. The best among these neighbors will be chosen as the next move for the next iteration, even if its fitness value is less than the fitness value of the current solution. The reason behind this is to ensure exploring the area as wide as possible without getting stuck in certain locations.

We define the neighbor in this algorithm as a solution differs with the current solution by a very few number of antenna locations, we call them the tabu antennas. For example: choosing two out of the N_s antennas and change their locations while keeping the locations for the rest of the antennas fixed.

After choosing the best among the neighbors, the old and new locations of the tabu antennas will be stored in the tabu matrix, and they can not be used for the next L iterations, where L is the length of the tabu matrix. The first set of tabu antennas to enter the tabu matrix will be the first one to leave it.

In the next iteration, the new solution will be used and new neighbors will be generated, and the best one will be considered as the next move, and the tabu antennas will be stored in the tabu matrix and so on until we reach the maximum number of iterations. At the end, the solution with the best fitness value in all the iterations will be declared as the final solution for the antenna selection problem.

Algorithm 6 The CTS Algorithm

Input: N, $maximum\ iterations$,
Generate random solution, then Evaluate its Capacity and store it in R_{max},
$iter \leftarrow 1$,
while $iter \leq maximum\ iterations$
    Generate N neighbors and Evaluate their capacities,
    Select the neighbor with highest capacity as the $next\ move$,
    Update the $tabu\ matrix$,
    Update R_{max},
    $iter = iter + 1$,
end while

3.7.2 QTS algorithm for antenna selection

The QTS algorithm starts by initializing every qubit in Q with equal probabilities of α and β, i.e. $|\alpha|^2 = |\beta|^2 = 0.5$. A matrix $\mathbf{V} \in \mathbb{R}^{N \times N_r}$ will be generated with random values between 0 and 1, where N is the number of possible solutions. This matrix will be used to generate the possible solutions or neighbors. Each solution in the neighborhood will be generated by comparing each value of $V_{i,j}$ with $|\beta_j|^2$ from Q, where $i = 1, 2, ..., N$ and $j = 1, 2, ..., N_r$. If the value of $V_{i,j}$ was less than $|\beta_j|^2$, then $X_{i,j}$ will be 1, otherwise it will be 0. Furthermore, $\mathbf{X} \in \mathbb{R}^{N \times N_r}$ is a matrix containing the set of possible solutions. Once the possible solutions are set, they need to be repaired, since the number of $1s$ in any solution might be less or more than N_s. To tackle this problem, the number of 1s within any possible solution will be checked. If it was less than N_s, then random locations of the solution will change their values from 0 to 1, until the total number of 1s is equal to N_s. In contrast, if the total number of 1s within any generated solution was greater than N_s, then random antenna locations will be ignored so that the total number of 1s in any solution will be equal to N_s. After repairing all the solutions, their capacities will be measured. The best solution $\mathbf{s}^b$ and the worst solution $\mathbf{s}^w$ will be used to update Q for the next iteration. If s_j^b and s_j^w were equal, i.e. they are both either $1s$ or $0s$, then the j^{th} qubit in Q will be tabooed. Otherwise, the rotation operator in (3.11) will be applied with either a positive or a negative value of θ, depending on the location of the qubit, to evolve the qubit towards a better solution as shown in Table 3.1.

In the next iteration, a new matrix of random numbers $\mathbf{V}$ will be generated and compared with the updated values of $|\beta|^2$ from Q to generate a new set of solutions. The solution with the highest capacity throughout the maximum number of iterations will be chosen as the final solution for the antenna selection problem.

Table 3.1: Rotation operator lookup table

Qubit location	s_j^b	s_j^w	Action on qubit
First or Third quadrant	0	0	Taboo
	0	1	$\mathbf{R}_{-\theta}$
	1	0	$\mathbf{R}_{\theta}$
	1	1	Taboo
Second or Forth quadrant	0	0	Taboo
	0	1	$\mathbf{R}_{\theta}$
	1	0	$\mathbf{R}_{-\theta}$
	1	1	Taboo

Algorithm 7 The QTS Algorithm

1: $iter \leftarrow 1$,
2: Initialize Q with $|\alpha|^2 = |\beta|^2 = 0.5$,
3: Initialize highest capacity R_{max} = 0,
4: **while** $iter \leq maximum\ iterations$
5: Generate random matrix $\mathbf{V}$,
6: Compare $\mathbf{V}$ with $|\beta|^2$ to $obtain\ X$,
7: Repair $\mathbf{X}$ and $evaluate$ the Capacity,
8: Update R_{max},
9: Find $\mathbf{s^b}$ and $\mathbf{s^w}$ to $update\ Q$ using the $Lookup\ table$,
10: $iter = iter + 1$,
11: **end while**

3.8 Processing time evaluation

To address the underlying complexity, the CPU time was measured for the different algorithms on a 3.4 GHz intel Core i5 PC, with 8 GB of RAM using the MATLAB R2014a software program. In Tables 3.2, 3.3, and 3.4 the SNR value was fixed at 0 dB, and the simulations were carried out 50 times for each algorithm. In Tables 3.2 and 3.4, the number of iterations were 50, while in Table 3.3, 75 iterations were used. On the other hand, the number of initial solutions were 20 for PSO, GA and ABC respectively, and 10 neighbors for CTS and QTS in Tables 3.2 and 3.3, while 40 initial solutions and 20 neighbors were considered in Table 3.3. The reason behind this is that the complexity level of these algorithms depends on both the number of iterations and the population size. In each case, the capacity as well as the CPU time were captured and compared for the different algorithms.

It should be noted that the processing time required for the employed evolutionary algorithms is normalized in the aforementioned tables, since it depends on the specifications of the computer machine or the DSP board that runs these algorithms. In other words, performing the same simulation codes on different machines will result in different processing time.

As shown from Tabules 3.2, 3.3, and 3.4, CTS and QTS require 50% shorter CPU time than any bio-inspired algorithm and achieve at the same time higher capacities. Moreover, QTS

Table 3.2: CPU time required for the different algorithms at SNR = 0 dB, N_t = 10, N_r = 400, N_s = 100, and 50 iterations

	Algorithm Specifications	Capacity (bps/Hz)	CPU time (normalized)
PSO	20 particles	34.4627	0.8352
GA	20 chromosomes	35.5471	0.7376
ABC	20 food sources	34.5867	1
CTS	10 neighbours	35.9624	0.3913
QTS	10 neighbours	36.4729	0.3377

Table 3.3: CPU time required for the different algorithms at SNR = 0 dB, N_t = 10, N_r = 400, N_s = 100, and 75 iterations

	Algorithm Specifications	Capacity (bps/Hz)	CPU time (normalized)
PSO	20 particles	34.4775	0.8954
GA	20 chromosomes	36.1473	0.7380
ABC	20 food sources	34.6487	1
CTS	10 neighbours	36.0814	0.3967
QTS	10 neighbours	36.6816	0.3489

outperform CTS in terms of both complexity and performance.

3.9 Simulation Results and Discussion

In this section, different simulation results will be presented for different evolutionary AS algorithms. It should be noticed that unless stated otherwise, the number of possible solutions N for the bio-inspired algorithms was set to 20, while 10 possible solutions was considered for the metaheuristic methods. Moreover, the correlation factor $|\rho|$ was set to 0.8, while the number of iterations was 25 for all algorithms. Furthermore, the number of tabu antennas was set to 2, while the length of tabu matrix L was chosen to be 20. Finally, the number of transmit antennas N_t was 10.

Table 3.4: CPU time required for the different algorithms at SNR = 0 dB, N_t = 10, N_r = 400, N_s = 100, and 50 iterations

	Algorithm Specifications	Capacity (bps/Hz)	CPU time (normalized)
PSO	40 particles	34.3620	0.8481
GA	40 chromosomes	36.2724	0.7505
ABC	40 food sources	34.7633	1
CTS	20 neighbours	36.5691	0.3811
QTS	20 neighbours	36.9215	0.3397

At the beginning, the best value of θ was determined by running extensive simulations, and it was found to be around 41^o as Fig 3.4 shows. The value of θ in [114] was set to 0.01π (1.8^o), however, this value is clearly not optimum for the antenna selection problem. The simulations were carried out for 1000 different channel observations for each value of θ from 0 to π with a step of 1^o. It is worth noting that the values of θ from π to 2π will give identical results to the values from 0 to π, since only the phase of $\mathbf{R}_\theta$ will be different.

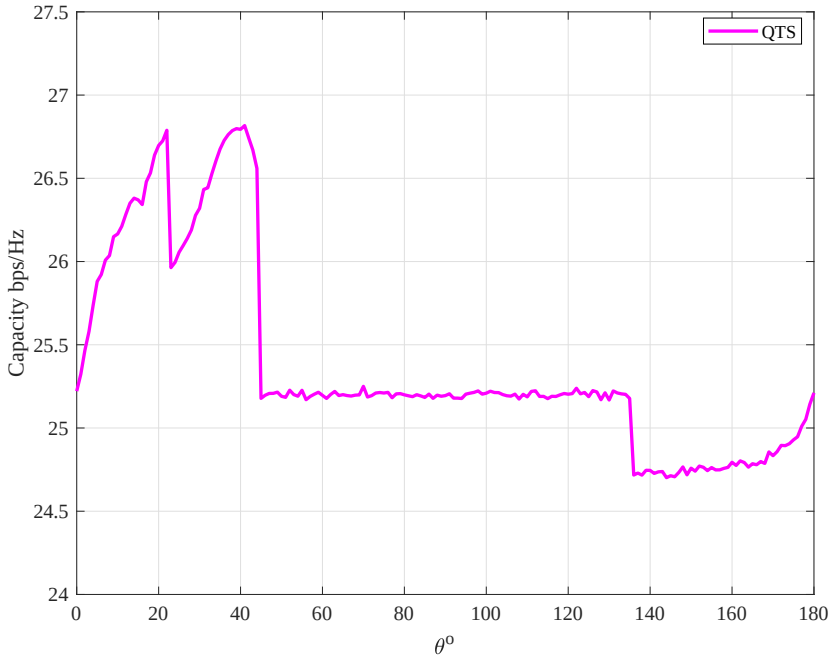


Figure 3.4: Capacity vs θ^o for $N_r = 200$, and $N_s = 50$ at 0 dB SNR.

3.9.1 Effect of changing the number of antennas and β on the rotation angle

The validity of the optimum rotation angle θ was tested for different number of transmit, and receive antennas. As shown in Fig. 3.5, the same behaviour is observed for the different cases, and changing the number of antennas has no effect on the rotation angle. This is a great advantage for the QTS algorithm over CTS and all other evolutionary algorithms. For example, in CTS algorithm, the parameters of this algorithm, i.e. the length of tabu matrix and the number of tabu antennas, are dynamic, and they are directly related to the number of iterations, number of selected antennas, as well as number of receive antennas. Therefore, changing any of the settings in the system means that these parameters need to be re-optimized. Moreover, the mutation probability P_m, number of selected chromosomes for reproduction in GA, the weights of the local global best antennas in PSO, ... etc, all need to be optimized in order to enhance the performance of these algorithms. In contrast, in QTS, and as shown in Fig. 3.5, the only parameter that needs optimizing is the angle of rotation, and it does not depend on the specifications of the system.

The performance of the QTS algorithm was further tested for different scenarios. For example, in Fig. 3.6, different initial values of β, which were used to initialized each qubit in Q, were applied, and the performance of the system was observed. The values of β that gave the best performance were found to be $\sqrt{0.3}$ and $\sqrt{0.5}$, while the worst performance was observed when β was set to $\sqrt{0.1}$ and $\sqrt{0.9}$. Furthermore, each value of β has different optimum angle θ. It is worth to mention that sum of α^2 and β^2 should always be 1.

3.9.2 Capacity vs SNR for different AS algorithms

The system capacity was adopted as the performance metric for the different AS schemes applied in this chapter. Figs. 3.7 and 3.8 show the achievable capacity for the different evolutionary algorithms as well as Random Selection (RS) criteria, and for different number of receive and selected antennas. QTS outperforms all other AS schemes for both cases, followed by CTS, then GA. In contrast, PSO algorithm show the worst performance between the evolutionary algorithms followed by the ABC scheme.

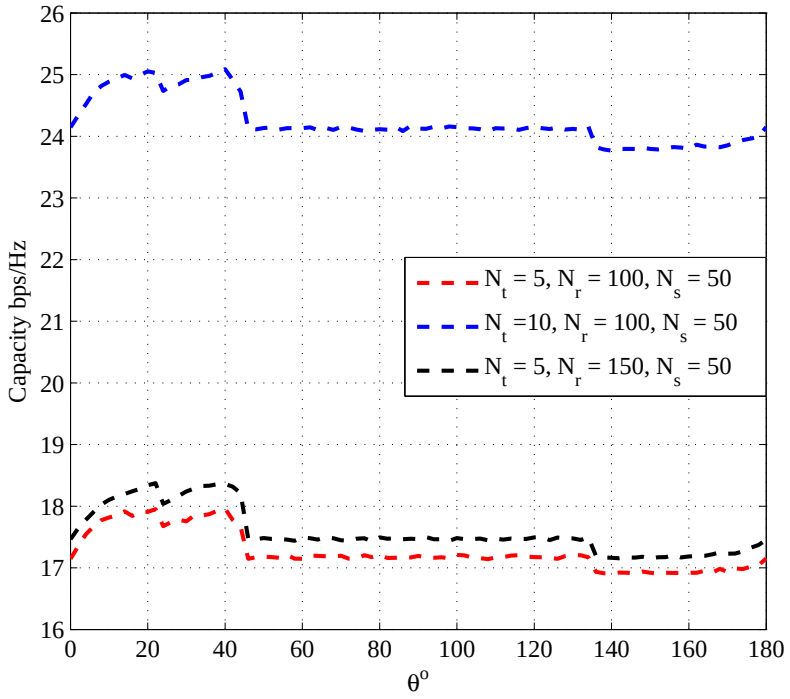


Figure 3.5: Capacity vs θ^o for different number of transmit, receive, and selected antennas when β =0.5, Number of iterations = 15, and SNR = 0 dB.

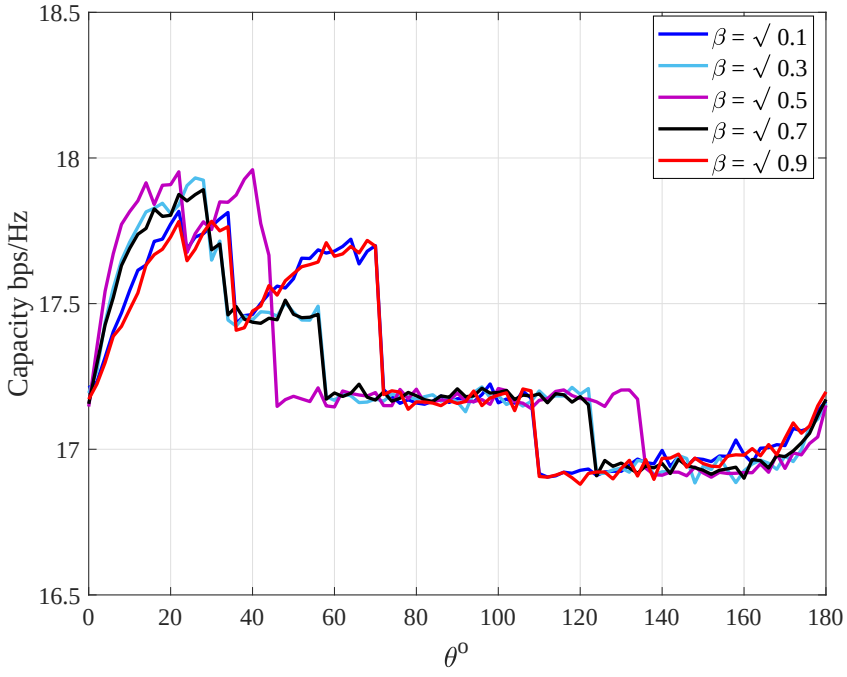


Figure 3.6: Capacity vs θ^o for different values of β with N_t = 5, N_r = 100, N_s = 50, N=10, Number of iterations = 15, and SNR = 0 dB.

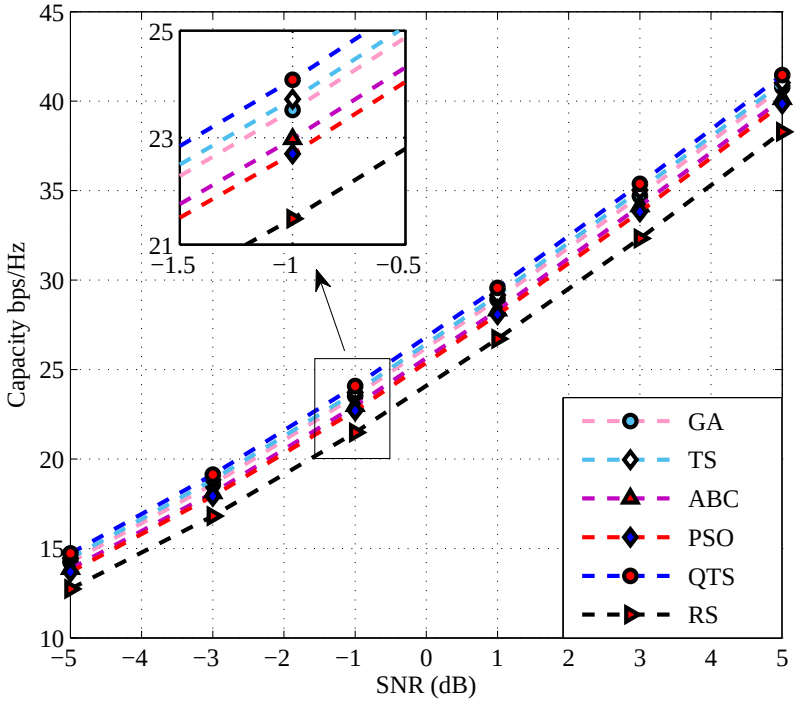


Figure 3.7: Capacity vs SNR different AS algorithms with N_r = 200, N_s = 50.

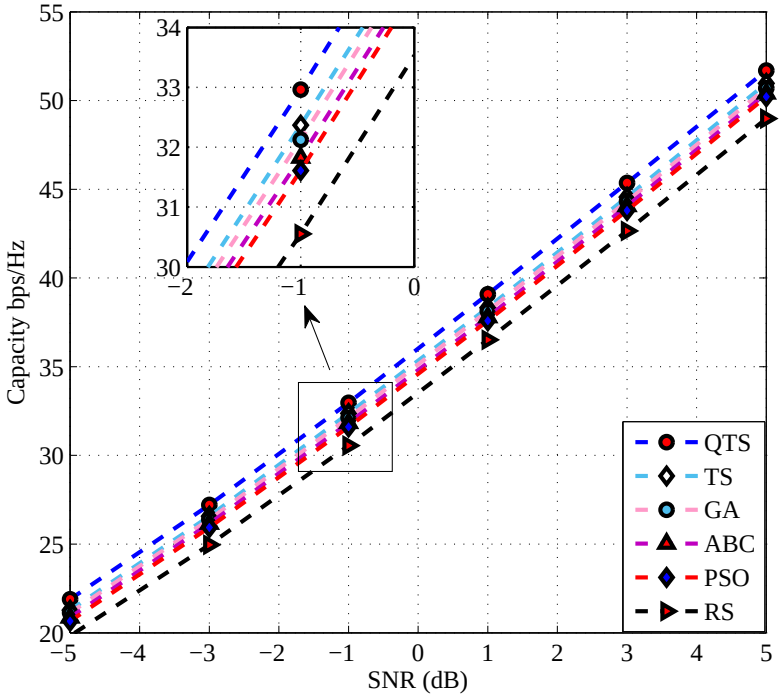


Figure 3.8: Capacity vs SNR for different AS algorithms with N_r = 400, N_s = 100.

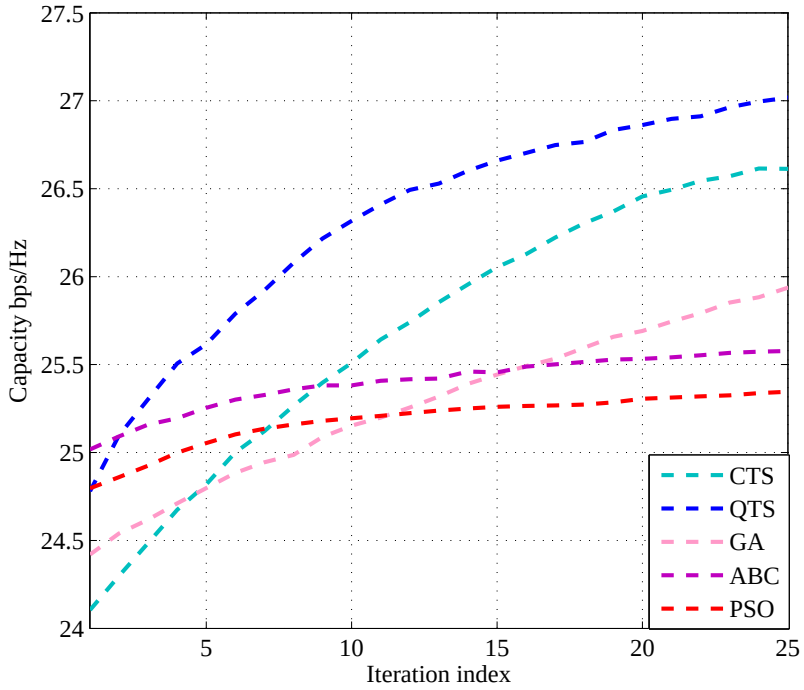


Figure 3.9: Convergence speed of different AS algorithms when N_r = 200, N_s = 50, SNR = 0 dB, N = 16.

3.9.3 Convergence speed for different AS algorithms

The convergence speed is one of the most important measures of any evolutionary algorithm, since it gives a better understanding on the efficiency of the evolving process. The convergence behaviour for the different AS algorithms was tested for different number of receive antennas, and size of populations.

Fig. 3.9 shows the convergence behaviour when $N_r = 200$, and the population size $N = 16$ for all algorithms. The QTS shows the best convergence speed among the applied algorithms followed by CTS then GA. Interestingly, the PSO and ABC algorithms considerably outperform both CTS and GA when the number of iterations is small. However, both of ABC and PSO show very small increase in capacity per iteration. Furthermore, when the size of population is small, for example $N = 8$, PSO and ABC both show the same performance as shown in Fig. 3.10.

3.9.4 Effect of spatial correlation on the performance

In this section, we observe the effect of the spatial correlation on the capacity of the system for different AS schemes. It should be noted that in our work we assume that the users do not experience any correlation between them, and the spatial correlation only exists between the

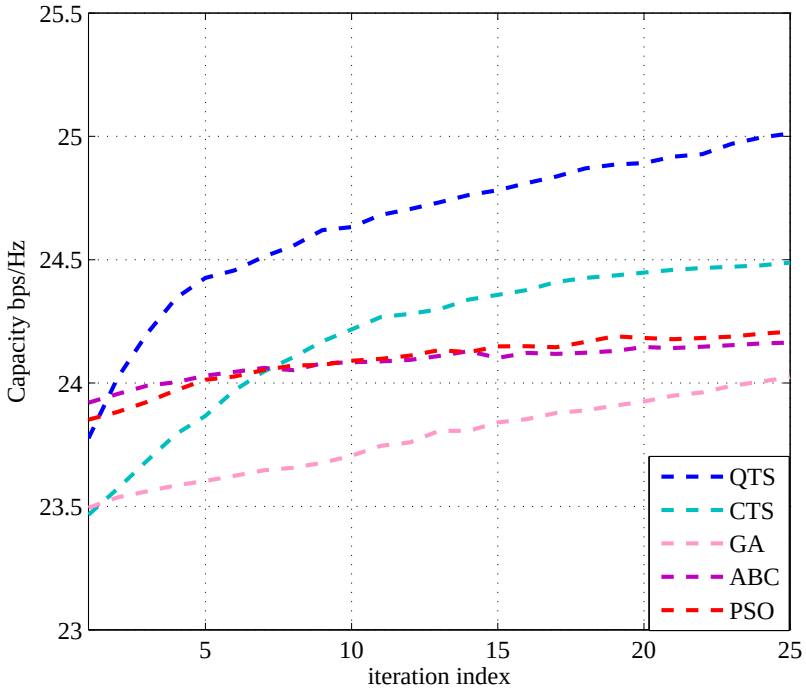


Figure 3.10: Convergence speed of different AS algorithms when N_r = 100, N_s = 50, SNR = 0 dB, N = 8.

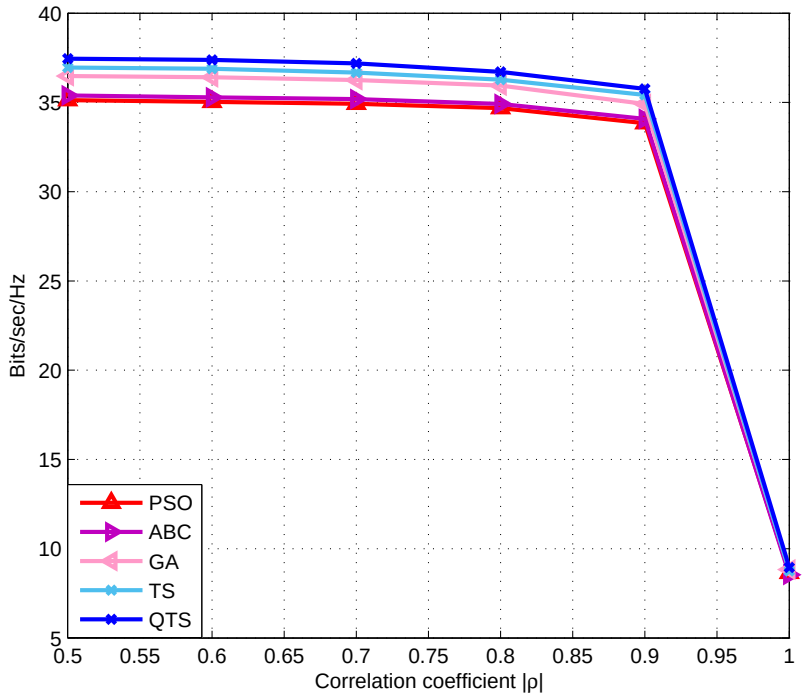


Figure 3.11: Capacity vs $|\rho|$ for different AS algorithms with SNR = 0 dB, N_r = 400, and N_s = 100.

antennas at the BS. As Fig. 3.11 demonstrates, all the different AS algorithms show similar behaviour, where their performance is highly degraded when $|\rho|$ is ≥ 0.9.

3.10 Critical evaluation of different evolutionary algorithms

In this section, we evaluate the adopted optimization methods for AS in MU M-MIMO system. For bio-inspired methods, GA requires lower complexity compared to PSO and ABC methods. Moreover, GA has better convergence behaviour when the number of iterations is large. In contrast, PSO and ABC demonstrated higher rates when small number of iterations is applied, i.e. less than 10 iterations as demonstrated in Fig. 3.9. However, they show small improvement in terms of sum rate capacity per iteration. Consequently, GA outperforms PSO and ABC as the number of iterations becomes large.

For Tabu-based methods, both CTS and QTS show good convergence behaviour with low complexity. However, QTS is preferable to tackle the AS optimization problem since it has only one optimization parameter, which is the rotation angle. In contrast, CTS has many optimization parameters such as: The number of tabu-antennas and the length of tabu-matrix, which are directly affected by the number of iterations and the number of antennas.

It should be noted that although QTS showed superior performance compared to the other adopted schemes in this chapter, this only holds for the AS case. For example, for user scheduling or detection, QTS might not show the best performance compared to GA, CTS, ABS, or PSO methods, since these algorithms evolve in a controlled random way which makes it difficult to predict their behaviour in solving different optimization problems.

3.11 Chapter Summary

In this chapter, bio-inspired and metaheuristic algorithms were applied for AS in M-MIMO systems. The bio-inspired algorithms included GA, PSO, and ABC, while the metaheuristic methods included CTS and QTS algorithms. Different simulation results were presented to compare the different algorithms, including: system capacity, complexity requirement, as well as convergence speed. The QTS algorithm outperformed all other evolutionary algorithms in terms of both complexity and system capacity, followed by CTS then GA. In contrast, PSO algorithm showed the worst performance followed by the ABC algorithm which had the highest complexity among the different algorithms. Moreover, in order to evolve the QTS algorithm towards better solution, the optimum rotation angle θ needs to be optimized, and was found in our work through exhaustive search. Our simulations demonstrated that the optimum angle depends on the initialization of each qubit in the Q-bit individual Q, but it is irrelevant of the number of transmit, receive, or selected antennas. In contrast, ABC, PSO, GA, and CTS algorithms have multiple dynamic parameters, and they need to be optimized in order to enhance their performance.

Chapter 4

Reduced Search Space Antenna Selection Methods for M-MIMO with MF precoding

4.1 Introduction

In the previous chapter, the cost function considered for optimization was the sum rate capacity, which can be obtained via SIC for the uplink, or DPC in the downlink scenario. Such methods require high complexity signal processing techniques as shown in Chapter 2. Furthermore, although evolutionary algorithms have low complexity compared to the exhaustive search method, or other high complexity AS schemes, they still require large number of cost function evaluations, in our case was the system capacity, which involves large number of complex vector multiplications.

In this chapter, and similar to our work in chapter 3, we focus on maximizing the spectral efficiency of MU M-MIMO system, by designing novel algorithms for AS. However, we aim to achieve further complexity reduction in terms of both precoding and algorithm design. In addition, we evaluate the EE of the proposed system for different number of selected antennas and over a wide range of SNRs.

Considering a system with N antennas at the BS serving K single-antenna users in the same time-frequency resources, our contributions can be summarized as follow

1. We design a User-Centric AS (UCAS) algorithm with reduced search space, where the available antennas are divided into K groups, and each group corresponds to one user only, where the k^{th} group contains the antennas that have the maximum channel norms for the k^{th} user. Therefore, to maximize the SINR for the k^{th} user at a given iteration, the proposed algorithm selects an antenna from the k^{th} group only. This reduces the search

space by a factor of K.

2. The second proposed algorithm, called a Semiblind Interference Rejection Antenna Selection (SIRAS), was designed to reject the highest number of interuser interference terms at any given iteration, based on the signs of the interference terms only.

3. The proposed designing criteria for both algorithms was intended to avoid any vector multiplications during the iterative selection process, by storing the multiplications of each two entries in the channel matrix before the iterative algorithms start. This resulted in dramatic complexity reduction in terms of number of floating-point operations (FLOPs) required for their implementations.

4. The performance of the proposed algorithms is evaluated in terms of the achievable sum rate and EE, under perfect and imperfect CSI, and showed a significant gain compared to other low complexity AS algorithms.

5. The complexity in terms of number of FLOPs is evaluated for the proposed algorithms, and compared with the complexity of other AS schemes found in the literature. The proposed schemes show a great performance-complexity trade-off.

4.2 System Model

In this section, we introduce in details the system model adopted in this chapter, and also the SINR, sum rate achieved with MF, as well as the EE for such scenario.

4.2.1 Channel, signal, and noise models

We consider in this work a single cell MU M-MIMO system operating in the downlink scenario as depicted in Fig. 4.1. A TDD transmission is assumed, where the users send orthogonal pilots to the BS to obtain CSI. Let $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_K]^T \in \mathbb{C}^{K \times N}$ be the channel matrix, where $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$ represents the channel coefficients between the BS and the k^{th} user, with $h_{k,n} \sim \mathcal{CN}(0, \sigma_h^2)$, and $\mathbf{w}_k \in \mathbb{C}^{N \times 1}$ denotes the unit norm precoding vector for the k^{th} user, and it can be given as

$$\mathbf{w}_k = \frac{\mathbf{h}_k^*}{\|\mathbf{h}_k\|}. \tag{4.1}$$

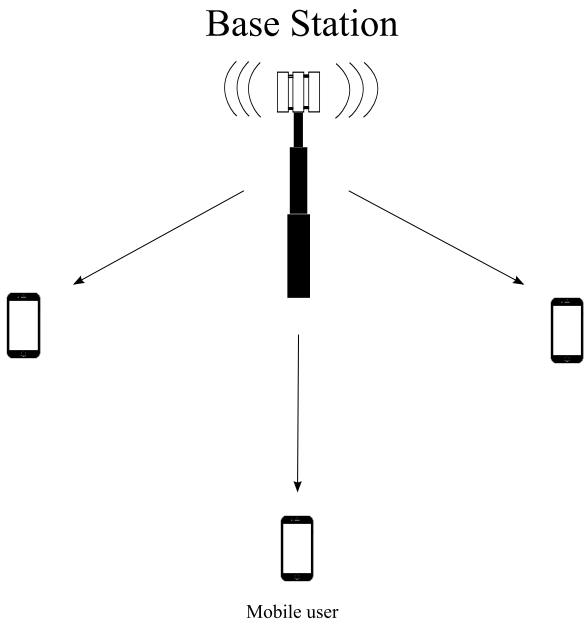


Figure 4.1: Multi-user M-MIMO downlink system.

Assuming that the BS has a perfect knowledge of $\mathbf{H}$, and MF precoding is applied at the transmitter, the received signal at the k^{th} user can be given as

$$y_k = \sqrt{p_k}\mathbf{h}_k^T\mathbf{w}_k x_k + \sum_{\substack{i=1 \\ i\neq k}}^{K} \sqrt{p_i}\mathbf{h}_k^T\mathbf{w}_i x_i + n_k, \tag{4.2}$$

where, $n_k \sim \mathcal{CN}(0,\, \sigma_n^2)$ is the AWGN, $\boldsymbol{x} = [x_1, ..., x_K]^T$ is the information symbols vector intended for the K users, with $E\{\boldsymbol{x}\boldsymbol{x}^H\} = \boldsymbol{I}_K$, $\sqrt{p_k}$ is the power allocated for the k^{th} user, and it follows the constraint

$$\sum_{i=1}^{K} p_i = P_T, \tag{4.3}$$

where P_T is the total transmission power in Watts available at the BS for all users.

4.2.2 SINR, achievable rates, and energy efficiency

In this work, we will use the sum rate and EE metrics to demonstrate the efficiency of the proposed AS algorithms. For systems with MF precoding, the SINR for the k^{th} user can be expressed as [122]

$$\gamma_k = \frac{p_k \left|\mathbf{h}_k^T\mathbf{w}_k\right|^2}{\sum_{i=1,\ i\neq k}^{K} p_i \left|\mathbf{h}_k^T\mathbf{w}_i\right|^2 + \sigma_n^2}, \tag{4.4}$$

and the achievable rate for the same user is a function of the SINR, and it can be given as [122]

$$R_k = \log_2(1 + \gamma_k), \tag{4.5}$$

while the total sum rate is the sum of achievable rates for all users, i.e.

$$R = \sum_{k=1}^{K} R_k. \tag{4.6}$$

Furthermore, the EE in bits/Joule can be defined as the total bandwidth, multiplied by the total sum rate over the total power consumed [123], i.e.

$$\text{EE} = \frac{B \,.\, R}{P_{total}}, \tag{4.7}$$

where B is the bandwidth, and P_{total} is the total power consumed at the transmitter and the receiver, and can be given as [124]

$$P_{total} = \frac{P_T}{\eta} + N_{RF} P_{tx} + K P_{rx}, \tag{4.8}$$

where η is the power amplifier efficiency, N_{RF} is the number of activated RF chains at the transmitter, while P_{tx} and P_{rx} are the circuit power consumption per RF chain at the transmitter and the receiver, respectively [124].

The aim is to design efficient low complexity AS algorithms to select N_s out of the available N antennas at the BS to maximize the SINR. It should be noted that finding the optimal number of antennas is not considered in this work, since optimizing the number of antennas to maximize the total sum rate will result in poor performance in terms of EE, and vice versa. Moreover, throughout this work, we assume frequency-flat fading channels. Similar assumptions in M-MIMO systems with AS were made in [125], [43]. However, for system with frequency selective channels, Orthogonal Frequency Division Multiplexing (OFDM) can be utilized, and the proposed algorithms can be applied after averaging over all subcarriers as in [27].

4.3 Low complexity AS algorithms in M-MIMO

It is well known that optimal AS can be achieved through exhaustive search, however, this method is prohibited for systems with large number of antennas due to its enormous complexity. Therefore, in this section we present two low complexity AS algorithms found in the literature

and shall be used for comparison reasons with our proposed methods.

4.3.1 Maximum SNR AS algorithm

The Maximum SNR (MS) algorithm represents the simplest selection method. After calculating the channel norms for each antenna, the N_s antennas with highest channel norms will be selected as the final solution, this can analytically be given as

$$\Delta = \arg\max_{N_s} \left\{ \left\| \mathbf{h}_1^c \right\|, \left\| \mathbf{h}_2^c \right\|, ..., \left\| \mathbf{h}_N^c \right\| \right\}, \tag{4.9}$$

where Δ is a set containing the indices of the N_s antennas that have the highest channel norms.

4.3.2 Greedy AS algorithm

The authors in [44] showed that for systems with MF precoding, removing certain antennas can increase the total sum rate. Two algorithms were proposed by the authors, the first is a high complexity iterative algorithm, where at each iteration all the available antennas are tested, and the antenna that causes the largest performance degradation is removed. Although this algorithm guarantees near optimal results, its high complexity makes it an impractical solution for the AS problem. The second algorithm that was proposed by the same authors is a low complexity greedy selection algorithm, where the antennas are sorted based on their channel norms values in a descending order, and at each iteration, the antenna with highest channel norm will be tested, if it increases the total sum rate then it will be activated, otherwise it will be deactivated. Since the authors consider a very similar scenario to that we consider in our work, we shall compare our results with their algorithm in terms of sum rate performance, EE, as well as implementation complexity. The greedy selection method is described in Algorithm 8.

Algorithm 8 Greedy AS Algorithm

Input:
  N, N_s, and $\mathbf{H}$,
Initialize:
  $\Lambda \leftarrow 1:N$, (Λ is a set of available antennas),
  $\mathcal{N} = \mathbf{0}_{1\times N}$, ($\mathcal{N}$ is the set of selected antennas),
  $R_{max} = 0$,
while $\Lambda \neq \emptyset$ and $\sum_{n=1}^{N} \mathcal{N}_n < N_s$ **do**
  $n^{\star}$ = $\arg\max_{n\in\Lambda} \|\mathbf{h}_n^c\|$,
  $\mathcal{N}_{n^{\star}} \leftarrow 1$,
  Evaluate the Sum rate R in (4.6) using the activated antennas in $\mathcal{N}$,
  if $R > R_{max}$
    $R_{max} = R$,
  else
    $\mathcal{N}_{n^{\star}} \leftarrow 0$,
  end if
  $\Lambda = \Lambda \setminus n^{\star}$,
end while
Output: $[\mathbf{H}_j^c]_{j\notin\Lambda}$

4.4 Proposed AS algorithms

Although MF is one of the most attractive forms of linear precoding, it suffers from inter-user interference, which causes a dramatic performance degradation. Both of the proposed algorithms depend on the channel correlation matrix $\mathbf{U} = \mathbf{H}\mathbf{H}^{\mathbf{H}}$, which can be expressed as

$$\mathbf{U} = \begin{pmatrix} \mathbf{h}_1^T\mathbf{h}_1^* & \mathbf{h}_1^T\mathbf{h}_2^* & \cdots & \cdots & \mathbf{h}_1^T\mathbf{h}_K^* \\ \mathbf{h}_2^T\mathbf{h}_1^* & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \mathbf{h}_{K-1}^T\mathbf{h}_K^* \\ \mathbf{h}_K^T\mathbf{h}_1^* & \cdots & \cdots & \cdots & \mathbf{h}_K^T\mathbf{h}_K^* \end{pmatrix}, \tag{4.10}$$

where the elements at the diagonal of (4.10) are related to the desired signal gain for each user, while the elements in the upper and lower triangle parts are directly related to the inter-user interference, since the interference from user j to user i is $\frac{\left|\mathbf{h}_i^T\mathbf{h}_j^*\right|^2}{\|\mathbf{h}_j\|^2}$, for $i, j = 1, 2, \ldots K$, and $i \neq j$. Each element in (4.10) is a summation of N multiplications between two complex numbers, i.e.

$$\mathbf{h}_i^T \mathbf{h}_j^* \;=\; h_{i,1}h_{j,1}^* + h_{i,2}h_{j,2}^* + ... + h_{i,N}h_{j,N}^*. \tag{4.11}$$

Both of the proposed algorithms aim to increase the SINR, and hence increase the sum rate and EE, while maintaining low implementation complexity requirement.

4.4.1 Proposed User-Centric Antenna Selection Algorithm

In this method, the antennas at the BS are divided into K groups based on their channel norms, where each group corresponds to one user. Moreover, each group should be allocated exactly N/K antennas, therefore, any group that reaches its maximum limit will not be considered in the allocation process for the remaining set of available antennas. The n^{th} antenna at the BS will be allocated to the k^{th} group, denoted as $\mathcal{G}_k$, only if it satisfies the following condition

$$h_{k,n} \in \mathcal{G}_k \iff |h_{k,n}| > |h_{i,n}|\,, \forall i \in \mathcal{K} \setminus k. \tag{4.12}$$

where $\mathcal{K}$ is a set containing the indices of groups with less than N/K antennas, the proposed grouping technique is described in Algorithm 9. To reduce the computational complexity, we store the values of the N complex multiplications in (4.11), and for each element in (4.10), in a matrix $\mathbf{\Xi} \in \mathbb{C}^{K^2 \times N}$, this will result in avoiding any vector multiplication later on during the iterative selection process. The matrix $\mathbf{\Xi}$ can be represented as

$$\mathbf{\Xi} = \begin{pmatrix} \xi_{1,1} & \cdots & \cdots & \xi_{1,N} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \xi_{K,1} & \cdots & \cdots & \xi_{K,N} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \xi_{K^2,1} & \cdots & \cdots & \xi_{K^2,N} \end{pmatrix}, \tag{4.13}$$

where $\xi_{1,1} = h_{1,1}h_{1,1}^*$, $\xi_{1,N} = h_{1,N}h_{1,N}^*$, $\xi_{K,1} = h_{1,1}h_{K,1}^*$, $\xi_{K,N} = h_{1,N}h_{K,N}^*$, $\xi_{K^2,1} = h_{K,1}h_{K,1}^*$ and $\xi_{K^2,N} = h_{K,N}h_{K,N}^*$. Therefore, the SINR for the k^{th} user depends mainly on rows $\{K(k-1)+1, K(k-1)+2, \ldots, Kk\}$ in $\mathbf{\Xi}$. At a given iteration, the algorithm aims to maximize the SINR for the k^{th} user only, by selecting a single antenna from $\mathcal{G}_k$. Let $\mathcal{R}_k$ denotes the set of row indices in $\mathbf{\Xi}$ that corresponds to user k, with $\mathcal{R}_{k_i}$ being the i^{th} element of $\mathcal{R}_k$, for example,

Algorithm 9 Proposed grouping strategy for the UCAS method

Input K, N, and $\mathbf{H}$,
Initialize
  $\mathcal{G}_k = \mathbf{0}_{\frac{N}{K}\times 1}, \forall k \in \{1:K\}$,
  $t_k = 0, \forall k \in \{1:K\}$, (Number of antennas in each group),
  $\mathcal{K} = 1:K$, (Available groups with less than $\frac{N}{K}$ antennas),
for $n = 1 \rightarrow N$
  $k^* = \arg\max_{k\in\mathcal{K}} |h_{k,n}|$,
  $t_{k^*} = t_{k^*} + 1$,
  $\mathcal{G}_{k^*}(t_{k^*}) = n$, ($\mathcal{G}_k(i)$ is the i^{th} element of $\mathcal{G}_k$),
  if $t_{k^*} = \frac{N}{K}$
    $\mathcal{K} = \mathcal{K} \setminus k^*$,
  end if
end for
Output $\mathcal{G}_k, \forall k \in \{1:K\}$

$\mathcal{R}_1 = \{1, 2, \ldots, K\}$, while $\mathcal{R}_{1_1} = 1$, and let $\boldsymbol{\omega} = [\omega_1, \omega_2, \ldots, \omega_{K^2}]^T$ be a vector initialized with zeros, and used to update the values of the signal and interference terms after each selected antenna, since selecting any antenna from any group will result in adding a column from $\boldsymbol{\Xi}$ to $\boldsymbol{\omega}$. To clarify this, assume at iteration t, the antenna $\zeta^{[t]}$ was selected, the vector $\boldsymbol{\omega}$ will then be updated as follow

$$\boldsymbol{\omega}^{[t]} = \boldsymbol{\omega}^{[t-1]} + \boldsymbol{\xi}^c_{\zeta^{[t]}}, \tag{4.14}$$

each antenna is selected to maximize the SINR for a certain user k, this can analytically expressed as

$$\zeta^{[t]} = \arg\max_{ns\in\mathcal{G}_k} \frac{\left|\omega^{[t-1]}_{\mathcal{R}_{k_k}} + \xi_{\mathcal{R}_{k_k},ns}\right|^2}{\sum_{\substack{i=1\\ i\neq k}}^{K} \left|\omega^{[t-1]}_{\mathcal{R}_{k_i}} + \xi_{\mathcal{R}_{k_i},ns}\right|^2 + \sigma_n^2}, \tag{4.15}$$

note that the selection process was not only achieved without any vector multiplications, but also the size of the search space was reduced by K times. For example, in selecting the 1^{st} antenna, the search space was reduced from N to N/K. Moreover, it should be noted that $\mathcal{G}_k$ is updated after the selection, by removing the index that corresponds to the selected antenna before the next iteration starts. In the next iteration, the selection will be carried out in a similar manner to maximize the SINR for the next user, i.e. the $(k+1)^{th}$ user. It is worth to mention that exactly N_s/K antennas are selected from each group to ensure fairness between all users. The UCAS scheme is fully described in Algorithm 10.

Algorithm 10 Proposed UCAS algorithm

Input K, N, N_s, $\boldsymbol{\Xi}$, $\mathcal{G}_k(\forall k)$, $\mathcal{R}_k(\forall k)$, and $\mathbf{H}$,
Initialize
  $\boldsymbol{\omega}^{[0]} = \mathbf{0}_{K^2\times 1}$, $t \leftarrow 0$,
for $l = 1 \rightarrow N_s/K$
  for $k = 1 \rightarrow K$
    $t \leftarrow t + 1$,
    $\zeta^{[t]} = \arg\max_{n_s \in \mathcal{G}_k} \frac{\left|\omega^{[t-1]}_{\mathcal{R}_{k_k}} + \xi_{\mathcal{R}_{k_k},n_s}\right|^2}{\sum_{\substack{i=1\\ i\neq k}}^{K}\left|\omega^{[t-1]}_{\mathcal{R}_{k_i}} + \xi_{\mathcal{R}_{k_i},n_s}\right|^2 + \sigma_n^2}$,
    $\boldsymbol{\omega}^{[t]} = \boldsymbol{\omega}^{[t-1]} + \boldsymbol{\xi}^c_{\zeta^{[t]}}$,
    $\mathcal{G}_k \leftarrow \mathcal{G}_k \setminus \zeta^{[t]}$,
  end for
end for
Output $[\mathbf{H}^c_j]_{j\notin\mathcal{G}_{k,\forall k}}$

4.4.2 Proposed Semiblind Interference Rejection Antenna Selection Algorithm

The SIRAS algorithm aims to minimize the interuser interference by minimizing the terms in the upper and lower triangle parts of (4.10) relying on the signs of the interference terms only, hence the name semiblind. Furthermore, since $\mathbf{h}_i^T\mathbf{h}_j^* = (\mathbf{h}_j^T\mathbf{h}_i^*)^*$, it is sufficient to minimize the elements in the upper triangle part only, and that will lead to the exact same minimization for the terms in the lower triangle part, and vice versa. The algorithm starts by storing the $M \times N$ complex multiplications in $\boldsymbol{\Phi}$, where $M = (K^2 - K)/2$ is the number of interference terms in the upper triangle part of (4.10), and it can be expressed as follow

$$\boldsymbol{\Phi} = \begin{pmatrix} \phi_{1,1} & \cdots & \cdots & \phi_{1,N} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \phi_{M,1} & \cdots & \cdots & \phi_{M,N} \end{pmatrix}, \tag{4.16}$$

since the upper triangle part of (4.10) is considered in this work, the N elements in the first row of (4.16) are the values of the N complex number multiplications in the first interference term in (4.10), i.e. $\phi_{1,n} = h_{1,n}h^*_{2,n}$, while $\phi_{M,n} = h_{K-1,n}h^*_{K,n}$, for $n = 1, 2, ..., N$. The algorithm then selects its first antenna based on the maximum total channel norms

$$\zeta^{[0]} = \arg\max_{n\in 1:N} \|\mathbf{h}^c_n\|, \tag{4.17}$$

where $\zeta^{[0]}$ is the first selected antenna. After selecting the first antenna, a vector $\boldsymbol{\psi} \in \mathbb{C}^{M\times 1}$ will be initialized with the column in $\boldsymbol{\Phi}$ that corresponds to the selected antenna, i.e. $\boldsymbol{\psi}^{[0]} = \boldsymbol{\phi}^{c}_{\zeta^{[0]}}$. Note that the M values in $\boldsymbol{\psi}$ are complex, and they are directly related to the interference between the users at any given iteration, hence minimizing these values will result in minimizing the total interference and therefore higher SINR. Moreover, selecting any antenna will result in adding a column from $\boldsymbol{\Phi}$ to $\boldsymbol{\psi}$. Assume that at iteration t, the antenna $\zeta^{[t]}$ was selected, then $\boldsymbol{\psi}$ will be updated as follow

$$\boldsymbol{\psi}^{[t]} = \boldsymbol{\psi}^{[t-1]} + \boldsymbol{\phi}^{c}_{\zeta^{[t]}}, \tag{4.18}$$

therefore, the goal is to select the antenna that corresponds to the column in $\boldsymbol{\Phi}$ which will minimize the M complex values in $\boldsymbol{\psi}$. In other words, at iteration t, the algorithm aims to select the antenna $\zeta^{[t]}$ from the set $\mathcal{S}$, where the n^{th} antenna (denoted as λ_n) belongs to $\mathcal{S}$ if it satisfies the following condition

$$\begin{aligned}\lambda_n \in \mathcal{S} \iff & \left\{\text{sign}(\Re[\phi_{m,n}]) \neq \text{sign}(\Re[\psi_m^{[t-1]}])\right\} \cap \\ & \left\{\text{sign}(\Im[\phi_{m,n}]) \neq \text{sign}(\Im[\psi_m^{[t-1]}])\right\}, \forall m, n \in \mathcal{A},\end{aligned} \tag{4.19}$$

where $\mathcal{A}$ is a set containing the indices of available antennas at a given iteration. However, it is not guaranteed to find an antenna that satisfies the condition in (4.19), therefore, we relax this condition and select the antenna that has the highest number of opposite signs between $\boldsymbol{\Phi}$ and $\boldsymbol{\psi}$. In the next iteration, the vector $\boldsymbol{\psi}$ will be updated according to (4.18), and the same procedure will be repeated again until maximum number of selected antennas is reached. The proposed method is described in details in Algorithm 11.

4.5 Complexity Analysis

In order to show the efficiency of the proposed methods, the complexity of the proposed algorithms, MS method, as well as the greedy selection proposed in [44] will be evaluated in terms of number of FLOPs required for their implementation. We follow the analysis in [126], where the addition between two real numbers is equivalent to 1 FLOP, while their multiplication is equivalent to 4 additions. Moreover, we assume that comparing the signs of two real numbers is equivalent to 1 addition, while finding the square root of a real number and division between two real numbers are equivalent to one real multiplication.

It should be noted that the analysis carried out in this section is for the software complex-

Algorithm 11 Proposed SIRAS algorithm

Input K, N, N_s, M, $\mathbf{\Phi}$, and $\mathbf{H}$,
Initialize:
    $\boldsymbol{\psi} = \mathbf{0}_{M\times 1}$,
    $\mathcal{M} = \mathbf{0}_{1\times N}$, (set containing the indices of selected antennas)
    $\mathcal{A} = 1 \rightarrow N$, (set containing the indices of available antennas)
$\zeta^{[0]} = \arg\max_{n\in 1:N} \|\mathbf{h}_n^c\|$,
$\boldsymbol{\psi}^{[0]} = \boldsymbol{\phi}^c_{\zeta^{[0]}}$,
for $t = 1 \rightarrow N_s$-1
    $\mathcal{M}_{\zeta^{[t-1]}} = 1$,
    $\mathcal{A} = \mathcal{A} \setminus \zeta^{[t-1]}$,
    $\boldsymbol{\lambda} = \mathbf{0}_{1\times N-t}$, (vector containing the number of opposite signs between $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ for each available antenna),
    for $n = 1 \rightarrow N - t$
        for $m = 1 \rightarrow M$
            if sign($\Re[\phi_{m,\eta_n}]$) $\neq$ sign($\Re[\psi_m^{[t-1]}]$)
                $\lambda_n = \lambda_n + 1$,
            end if
            if sign($\Im[\phi_{m,\eta_n}]$) $\neq$ sign($\Im[\psi_m^{[t-1]}]$)
                $\lambda_n = \lambda_n + 1$,
            end if
        end for
    end for
    $\zeta^{[t]} = \arg\max \boldsymbol{\lambda}$,
    $\boldsymbol{\psi}^{[t]} = \boldsymbol{\psi}^{[t-1]} + \boldsymbol{\phi}^c_{\zeta^{[t]}}$,
end for
Output: $[\mathbf{H}_j^c]_{j\notin\mathcal{A}}$

ity of the system only in terms of the number of operations required, which in return will be reflected on the time processing required to apply these algorithms in practical scenarios.

4.5.1 Complexity of MS scheme

The MS method has the lowest complexity, however that comes at the price of performance degradation. Finding the channel norms for the N antennas requires $N(10K+3)$ FLOPs, while sorting the antennas requires $N\log_{10} N$ FLOPs, therefore, the total complexity is

$$\mathcal{C}_{MS} = N(10K + \log_{10} N + 3). \tag{4.20}$$

4.5.2 Complexity of greedy AS algorithm

The greedy algorithm starts by finding the channel norms for the N antennas and sorting them, which requires $N(10K+\log_{10} N+3)$ FLOPs. The algorithm will then starts its iterations, where at each iteration the sum rate in (4.6) will be evaluated. Finding the MF precoding weights for all the users requires $K\sum_{l=1}^{Ns}(18l+3)$ FLOPs, while evaluating the SINRs for all users require $\sum_{l=1}^{N_s}(20K^2l + 20K^2 + 3K)$ Flops. Finally, calculating the sum rate after obtaining the SINR

values takes $(2K-1)N_s$ FLOPs, therefore, the total complexity of the greedy algorithm can be given as

$$\begin{aligned}\mathcal{C}_{greedy} &= N(10K+\log_{10}N+3)+N_s(20K^2+8K-1)\\ &\quad+\sum_{l=1}^{Ns} l(20K^2+18K).\end{aligned} \tag{4.21}$$

However, here we only consider an ideal case for the greedy algorithm by assuming that each selected antenna will boost the total sum rate, therefore, the actual complexity of the greedy algorithm is higher than that shown in this section, otherwise it would have shown the exact same sum rate performance as the MS scheme.

4.5.3 Complexity of UCAS algorithm

The UCAS algorithm starts by evaluating the matrix $\boldsymbol{\Xi}$ that has $K^2\times N$ complex multiplications, which results in $18K^2N$ FLOPs. Then, the absolute channel values between the N antennas and K users are evaluated before being allocated to $\mathcal{G}_k, \forall k$, which results in $14KN$ FLOPs. The N_s/K iterations will then start to select K antennas at each iteration, which results in $\sum_{l=1}^{N_s/K}(20K+4)(N-Kl)$ FLOPs. Finally, the vector $\boldsymbol{\omega}$ is updated after selecting each antenna, which requires $2K^2N_s$ FLOPs. Therefore, the total number of FLOPs required by the UCAS algorithm can be given as

$$\begin{aligned}\mathcal{C}_{UCAS} &= KN(18K+14)+2K^2N_s\\ &\quad+\sum_{l=1}^{N_s/K}(20K+4)(N-Kl).\end{aligned} \tag{4.22}$$

4.5.4 Complexity of SIRAS algorithm

The SIRAS algorithm requires $9N(K^2-K)$ FLOPs to obtain $\boldsymbol{\Phi}$, after that, $N(10K+4)$ FLOPs are required to select the first antenna. The N_s-1 iterations will then start, and at each iteration the vector $\boldsymbol{\lambda}$ is used to store the number of opposite signs, which takes $\sum_{l=1}^{N_s-1}2(N-l)(K^2-K)$ FLOPs. Finding the N_s-1 antennas from $\boldsymbol{\lambda}$ requires $\sum_{l=1}^{N_s-1}(N-l)$ FLOPs, and finally, updating the vector $\boldsymbol{\psi}$ results in $(N_s-1)(K^2-K)$ FLOPs. Therefore, the total complexity for the SIRAS algorithm can be given as

Table 4.1: Number of additions and multiplications for different AS schemes

Algorithm	Operator	Additions	Multiplications
MS	$\|\mathbf{h}_n^c\|, \forall n$	$N(2K-1)$	$N(2K+1)$
	Sorting	$N \log N$	$--$
	Total	$N(2K+\log N-1)$	$N(2K+1)$
Greedy	$\|\mathbf{h}_n^c\|, \forall n$	$N(2K-1)$	$N(2K+1)$
	Sorting	$N \log N$	$--$
	$\mathbf{w}_k, \forall k$	$\sum_{l=1}^{N_s} K(2l-1)$	$\sum_{l=1}^{N_s} K(4l+1)$
	$\gamma_k, \forall k$	$K\sum_{l=1}^{N_s}(4Kl-1)$	$K\sum_{l=1}^{N_s}(4Kl+5K+1)$
	R	$(2K-1)N_s$	$--$
	Total	$N(2K+\log N-1)-N_s+\sum_{l=1}^{N_s} l(4K^2+2K)$	$N(2K+1)+N_s(5K^2+2K)+\sum_{l=1}^{N_s} l(4K^2+4K)$
UCAS	$\mathbf{\Xi}$	$2K^2N$	$4K^2N$
	$\mathcal{G}_{k,\forall k}$	$2KN$	$3KN$
	$\zeta^{[t]}, \forall t \in \{1,...,N_s\}$	$\sum_{l=1}^{N_s/K}(\frac{N}{K}-l)4K^2$	$\sum_{l=1}^{N_s/K}(\frac{N}{K}-l)(4K^2+K)$
	$\boldsymbol{\omega}^{[t]}, \forall t \in \{1,...,N_s\}$	$2K^2N_s$	$--$
	Total	$K(2KN+2N+2KN_s)+\sum_{l=1}^{N_s/K}(\frac{N}{K}-l)4K^2$	$N(4K^2+3K)+\sum_{l=1}^{N_s/K}(\frac{N}{K}-l)(4K^2+K)$
SIRAS	$\mathbf{\Phi}$	$N(K^2-K)$	$2N(K^2-K)$
	$\|\mathbf{h}_n^c\|, \forall n$	$N(2K-1)$	$N(2K+1)$
	$\zeta^{[0]}$	N	$--$
	$\boldsymbol{\lambda}$	$\sum_{l=1}^{N_s-1} 2(N-l)(K^2-K)$	$--$
	$\zeta^{[t]}, \forall t \in \{1,...,N_s-1\}$	$\sum_{l=1}^{N_s-1}(N-l)$	$--$
	$\Psi^{[t]}, \forall t \in \{1,...,N_s-1\}$	$(N_s-1)(K^2-K)$	$--$
	Total	$N(K^2+K)+(N_s-1)(K^2-K)+\sum_{l=1}^{N_s-1}(N-l)(2K^2-2K+1)$	$N(2K^2+1)$

$$\mathcal{C}_{SIRAS} = N(9K^2+K+4)+(K^2-K)(N_s-1) + \sum_{l=1}^{N_s-1}(N-l)(2K^2-2K+1). \tag{4.23}$$

The number of additions and multiplications for the different AS schemes are shown for each operator in Table 4.1. It should be noted that in the aforementioned table, the number of additions also includes the comparison between the values or signs of two real numbers, while the number of multiplications includes the division as well as finding the square root of a real number. In addition, a critical example for the required number of additions and multiplications for the different AS schemes is illustrated in Table 4.2, for the case when $N = 256$, $N_s = 64$, and $K = 8$.

Fig 4.2 shows the complexity in number of FLOPs for different number of selected antennas. The MS algorithm has the lowest complexity and it does not depend on the number of selected antennas, since all the antennas will be sorted in descending order and the first N_s antenna indices will be selected. In contrast, the complexity of the Greedy, SIRAS and UCAS

Table 4.2: Number of additions and multiplications for different AS schemes when N = 256, Ns = 64, and K = 8

Algorithm	Operator	Additions	Multiplications
SM	$\|\mathbf{h}_n^c\|, \forall n$	3840	4352
	Sorting	768	$--$
	Total	4608	4352
Greedy	$\|\mathbf{h}_n^c\|, \forall n$	3840	4352
	Sorting	768	$--$
	$\mathbf{w}_k, \forall k$	32768	67072
	$\gamma_k, \forall k$	531968	553472
	R	960	$--$
	Total	570304	624896
UCAS	$\boldsymbol{\Xi}$	32768	65536
	$\mathcal{G}_{k,\forall k}$	4096	6144
	$\zeta^{[t]}$	56320	58080
	$\boldsymbol{\omega}^{[t]}$	8192	$--$
	Total	101376	129760
SIRAS	$\boldsymbol{\Phi}$	14336	28672
	$\|\mathbf{h}_n^c\|, \forall n$	3840	4352
	$\zeta^{[0]}$	256	$--$
	$\boldsymbol{\lambda}$	1580544	$--$
	$\zeta^{[t]}$	14112	$--$
	$\Psi^{[t]}$	3528	$--$
	Total	1616616	33024

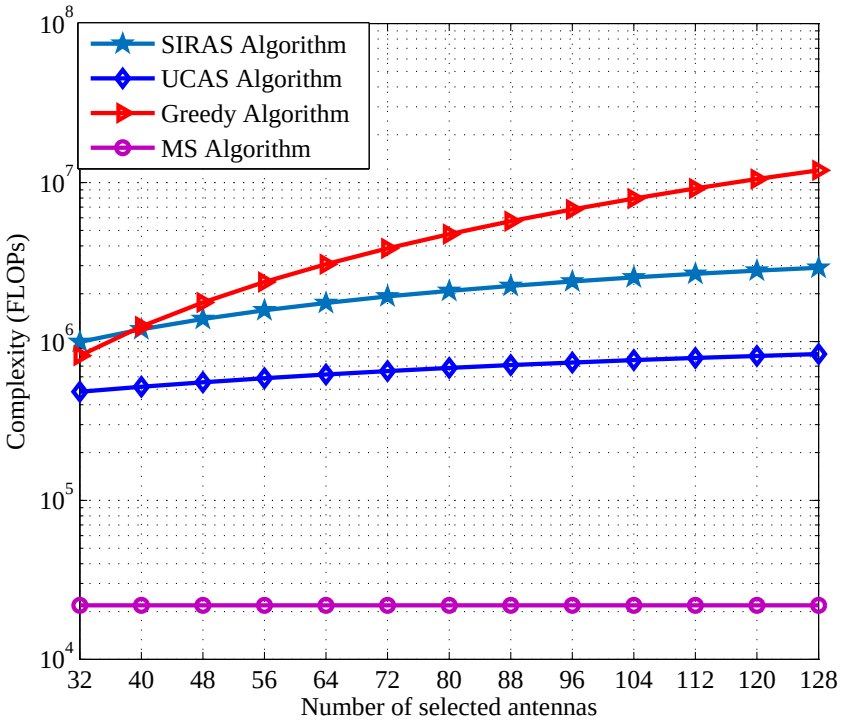


Figure 4.2: Complexity in number of FLOPs vs number of selected antennas for different AS schemes when $N = 256$, and $K = 8$.

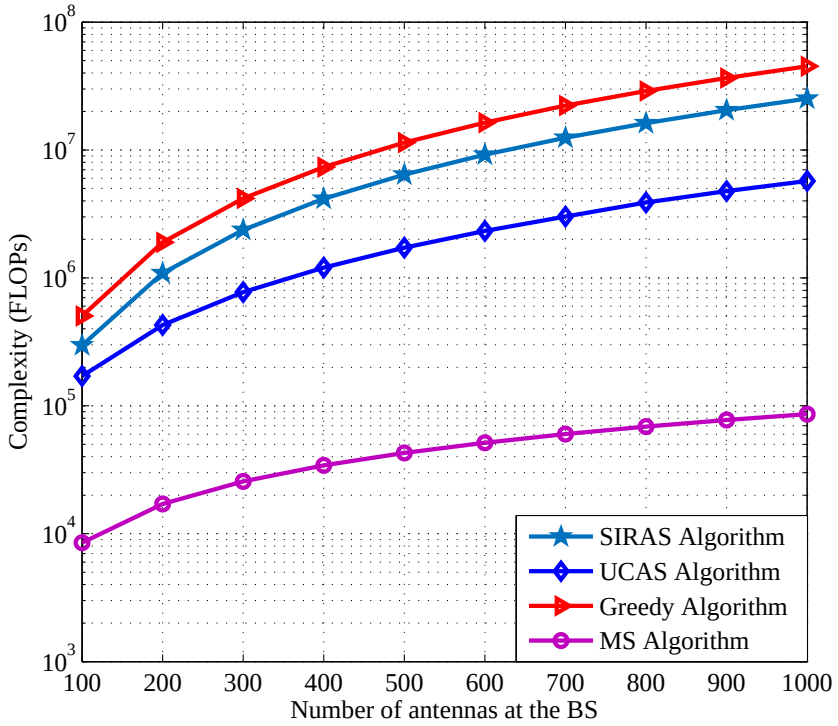


Figure 4.3: Complexity in number of FLOPs vs number of antennas at the BS for different AS schemes when $N_s = N/4$, and $K = 8$.

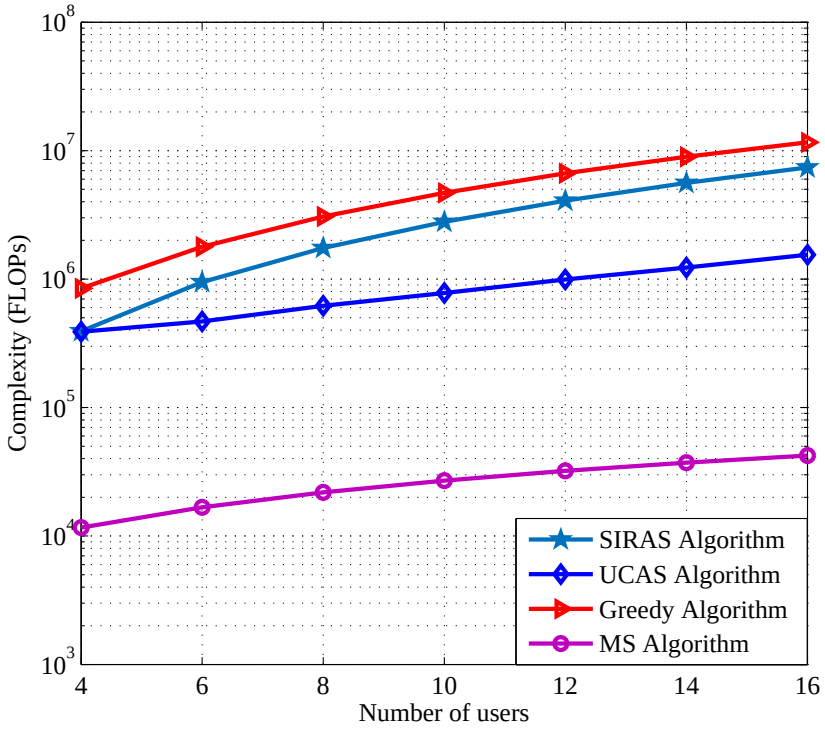


Figure 4.4: Complexity in number of FLOPs vs number of users for different AS schemes when $N = 256$, $N_s = N/4$.

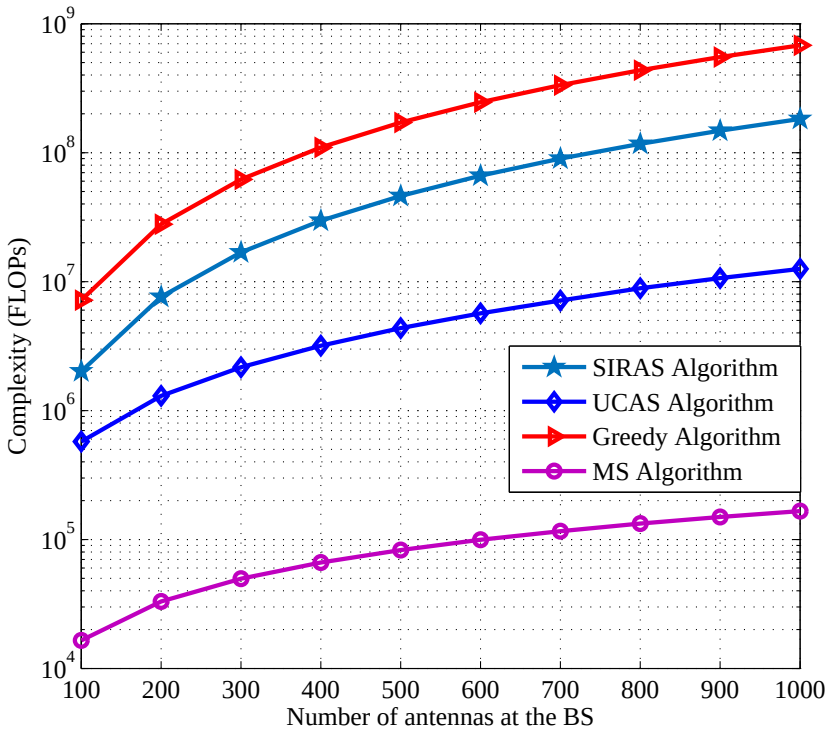


Figure 4.5: Complexity in number of FLOPs vs number of antennas at the BS for different AS schemes when $N_s = N/2$, and $K = 16$.

algorithms all increase when the number of selected antennas increase. The greedy algorithm requires the highest complexity among the three algorithm when the number of selected antenna is equal to or larger than 40.

Figs. 4.3 and 4.4 show the complexity for different AS schemes for different number of antennas N at the BS and different number of users K, respectively, with $N_s = N/4$. In both results, the MS algorithm has the lowest complexity, followed by the UCAS algorithm, then the SIRAS, while the greedy algorithm requires the highest complexity among the adopted schemes. Fig. 4.5 demonstrates the complexity for different number of antennas at the BS when $N_s = N/2$, and 16 users are being served, the UCAS algorithm show a significant reduction in complexity compared to the greedy selection algorithm. Furthermore, the actual complexity for the greedy algorithm is higher than the one shown in this section, and here we only consider an ideal case.

4.6 Numerical Results and Discussion

Before discussing the results, we need to introduce the different parameters used in this work. First, we define the SNR per user as

$$SNR = \frac{\sigma_h^2}{\sigma_n^2}\frac{P_T}{K}, \tag{4.24}$$

where the channel variance σ_h^2 can be found by using the path loss formula for a general urban channel model, which can be given as [127]

$$\mathrm{PL}\ (\mathrm{dB}) = 10\log_{10} d^{\alpha} + \beta, \tag{4.25}$$

where d is the distance between the users and the BS, α is the path loss component, and β is the fixed-loss component. In our simulations, d was set to 100 meters, α was assumed to be 2 [128], and β = 10 dB, therefore, σ_h^2 = 10^{-5}, while the noise variance σ_n^2 was assumed to have a value of 10^{-9}. For systems where users are uniformly distributed in the cell, PA techniques can be applied to ensure that each user will meet a minimum pre-defined threshold of SINR, however, this is out of the scope of this work. In addition, the bandwidth B was 20 MHz, the efficiency of the power amplifier η was 0.35, while the receive and transmit circuit power consumption per RF chain P_{rx} and P_{tx} were set to 62.5 mW and 48.2 mW, respectively [124]. Finally, the simulations were averaged over 10^4 different channels realizations for each SNR value.

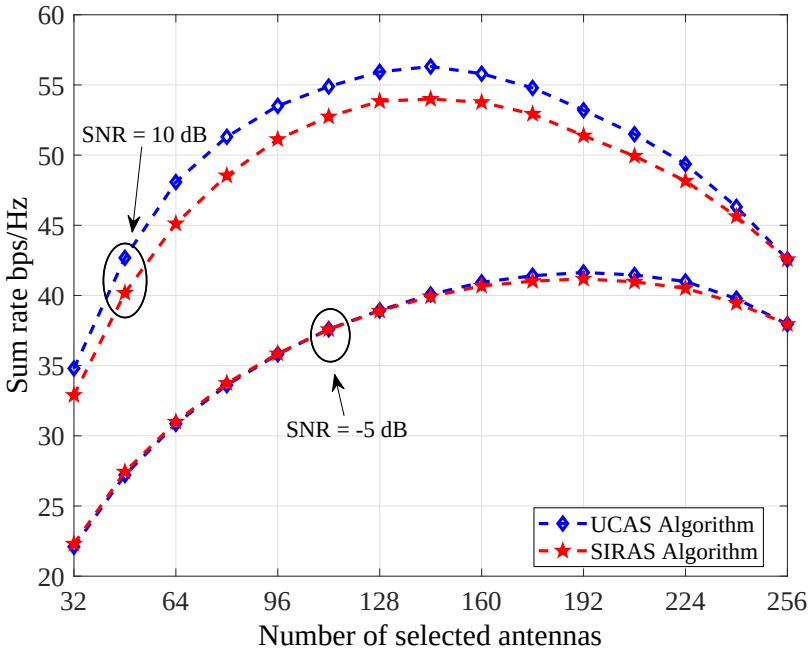


Figure 4.6: Sum rate vs number of selected antennas for the proposed AS schemes when $N = 256$ and $K = 8$.

4.6.1 Impact of AS on spectral and energy efficiencies

The advantage of designing efficient AS schemes when MF precoding is applied, in terms of spectral and energy efficiencies, can be seen in Figs. 4.6 and 4.7. Both of the proposed algorithms achieve higher rates with better EE than the full system case, where all the antennas are employed, i.e. when $N_s = N$. Furthermore, in terms of sum rate performance, increasing the SNR results in decreasing the number of antennas required to match the achievable rate where all antennas are activated. The reason behind this is that as the SNR increases, the effect of the noise becomes negligible, and the main factor that degrades the performance is the inter-user interference, and our proposed algorithms selects the antennas that highly minimizes the inter-user interference. In other words, although selecting a subset of the available antennas will reduce the signal gain, but the resultant SINR will be higher, hence higher rates are achieved. Moreover, in terms of EE, applying AS techniques can dramatically improve the system performance as can be seen in Fig. 4.7. For example, when the number of selected antennas is 64, the UCAS and SIRAS algorithms outperform the full system case by 190.8 and 184 $\mathrm{Mbits/Joule}$, respectively.

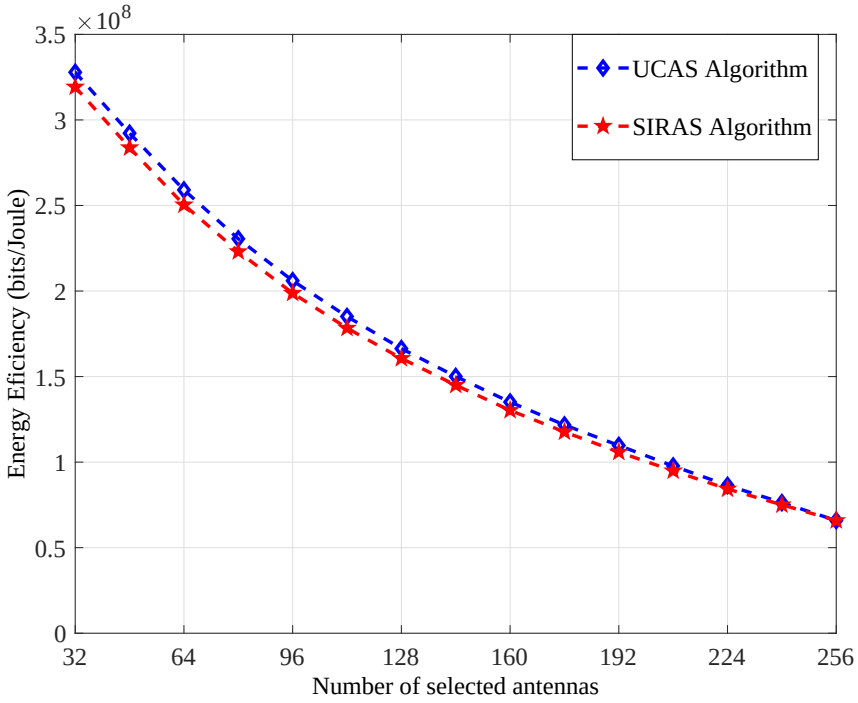


Figure 4.7: Energy Efficiency vs number of selected antennas for the proposed AS schemes when $N = 256$, $K = 8$, and SNR = 10 dB.

4.6.2 Achievable rates with fixed number of selected antennas

In this subsection, we show the performance of the proposed algorithms in terms of the achievable rates, and compare our results with MS and greedy selection techniques for a wide range of SNR values.

Fig. 4.8 shows the total sum rate when the BS is equipped with 128 antennas, out of which 32 are selected, for different AS schemes. The proposed algorithms outperform significantly both greedy and MS methods. Moreover, the UCAS slightly outperforms the SIRAS algorithm when the SNR is less than 10 dB, while they both show the same performance for higher SNR values. For this scenario, employing all the available antennas at the BS achieves higher rates than our proposed algorithms when the number of selected antennas is 32.

Fig. 4.9 shows the total sum rate when 256 antennas are placed at the BS, and 64 antennas are selected by the proposed algorithms. The proposed methods not only outperform the greedy and MS techniques, but also the case where all the antennas are activated for SNR values higher than 0.5 dB and 3.5 dB for the UCAS and SIRAS algorithms, respectively.

From Fig. 4.9, the UCAS algorithm outperforms the full system, greedy, and MS selection methods at 10 dB SNR by 5.44, 10.12, and 19.64 bps/Hz, respectively, which correspond to a significant 108.8, 202.4, and 392.8 Mbps for a system with 20 MHz of bandwidth, while the

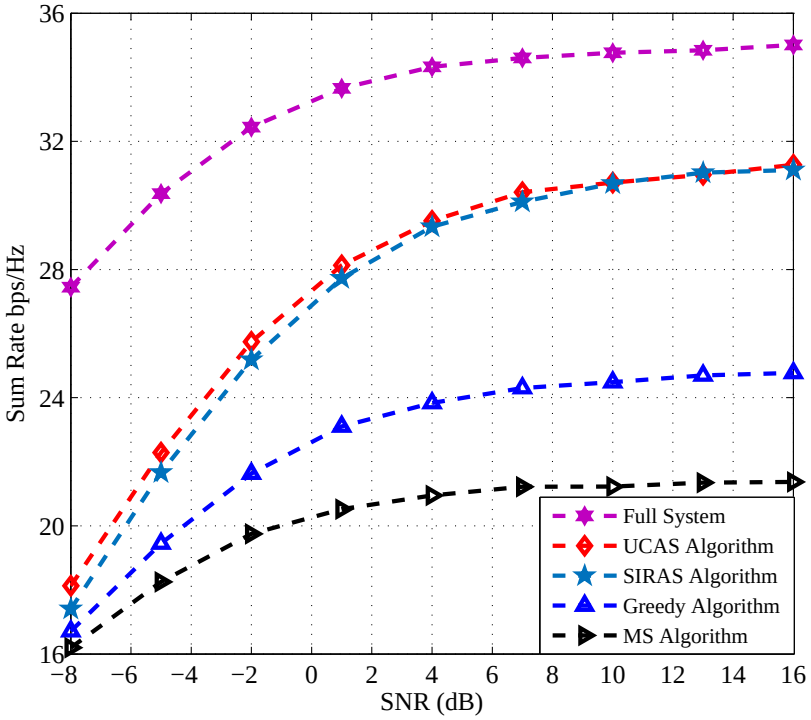


Figure 4.8: Sum rate vs SNR (dB) for different AS schemes when $N = 128$, $N_s = 32$, and $K = 8$.

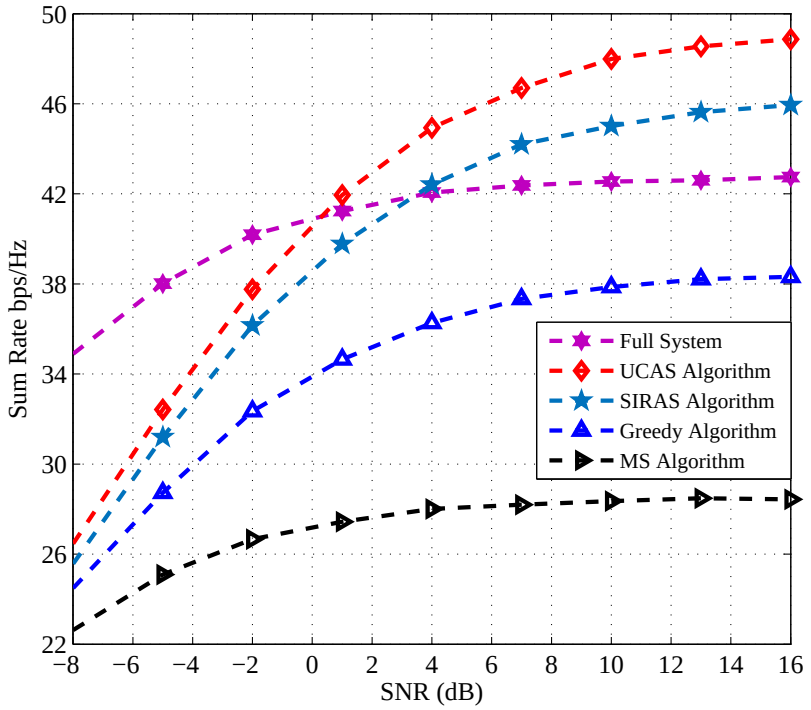


Figure 4.9: Sum rate vs SNR (dB) for different AS schemes when $N = 256$, $N_s = 64$, and $K = 8$.

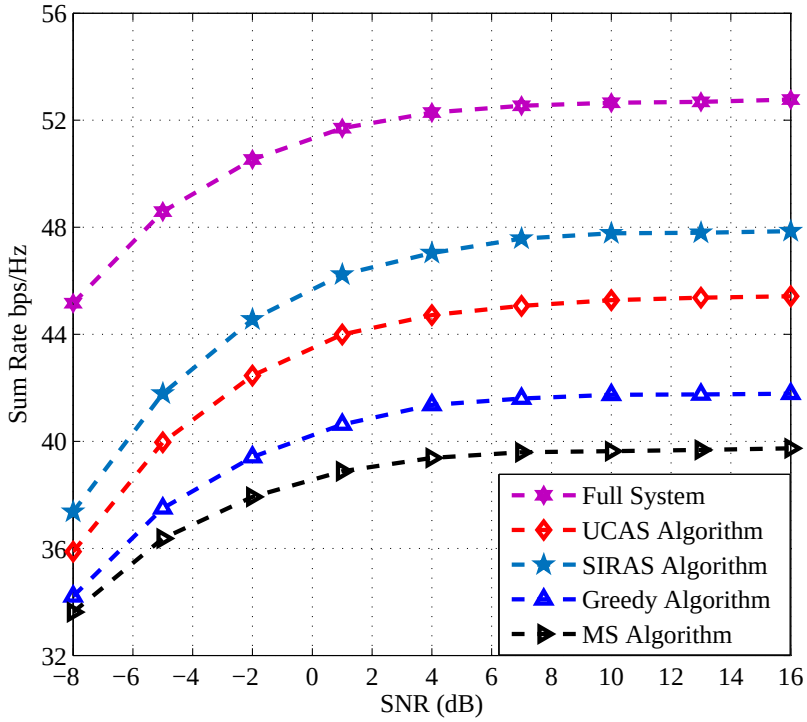


Figure 4.10: Sum rate vs SNR (dB) for different AS schemes when $N = 128$, $N_s = 64$, and $K = 16$.

SIRAS method shows an improvement of 2.46, 7.14, and 16.66 bps/Hz compared to the full system, greedy, and MS selection methods, respectively, at the same SNR value.

Figs. 4.10 and 4.11 show the achievable rates when the number of users is 16, and the BS is equipped with 128 and 256 antennas, respectively, and $N_s = N/2$. In both cases, the SIRAS algorithm show better performance than UCAS, and the proposed methods significantly outperform the greedy and MS selection methods. For example, when the BS is equipped with 256 antennas, the SIRAS algorithm achieves the same rate as the full system case, and outperform the UCAS, greedy, and MS methods by 2.57, 7.15, and 13.97 bps/Hz, respectively, at SNR of 10 dB.

The reason behind the SIRAS outperforming the UCAS method is that increasing the number of users will increase the number of interference terms. Accordingly, the SIRAS method selects the antennas the will reject the interference for all the users at the same time. In contrast, the UCAS selects at each iteration the antenna that will maximize the SINR for one user only, and as the number of users becomes large, this method becomes less efficient that the SIRAS as demonstrated in Figs. 4.10 and 4.11.

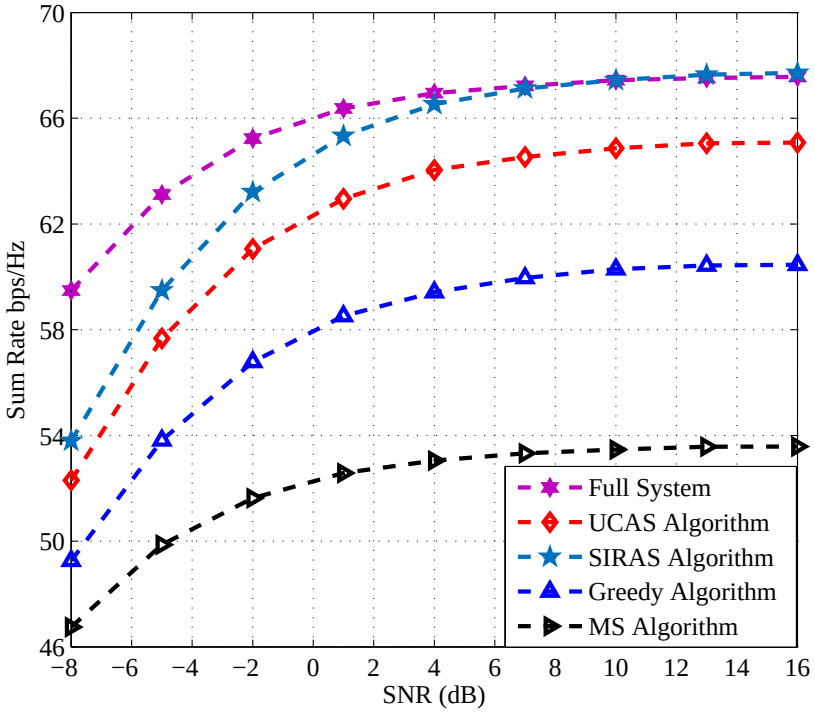


Figure 4.11: Sum rate vs SNR (dB) for different AS schemes when $N = 256$, $N_s = 128$, and $K = 16$.

4.6.3 Energy efficiency performance

The same scenarios are considered for the EE, where in the first scenario $N = 128$ antennas are available at the BS, while in the second scenario $N = 256$, and the number of selected antennas $N_s = N/4$ in both cases.

As Figs. 4.12 and 4.13 show, at moderate SNR values, the proposed methods outperform significantly all other schemes for both scenarios. Furthermore, activating all the antennas results in extremely poor EE performance, which validates the importance of AS for energy efficient systems. Finally, the UCAS shows better EE performance than SIRAS when the number of selected antennas was 64, especially at SNR values of less than 25 dB, while both algorithms show the same performance when 32 antennas were selected for all the SNR values between 10 to 40 dB.

4.6.4 Achievable rates with imperfect channel state information

In practical scenarios, the complex channel matrix $\mathbf{H}$ is estimated at the BS through pilot signals sent by the users, and channel estimation errors arise in any practical system. In this subsection, we evaluate the performance of the proposed algorithms under imperfect CSI. The estimated

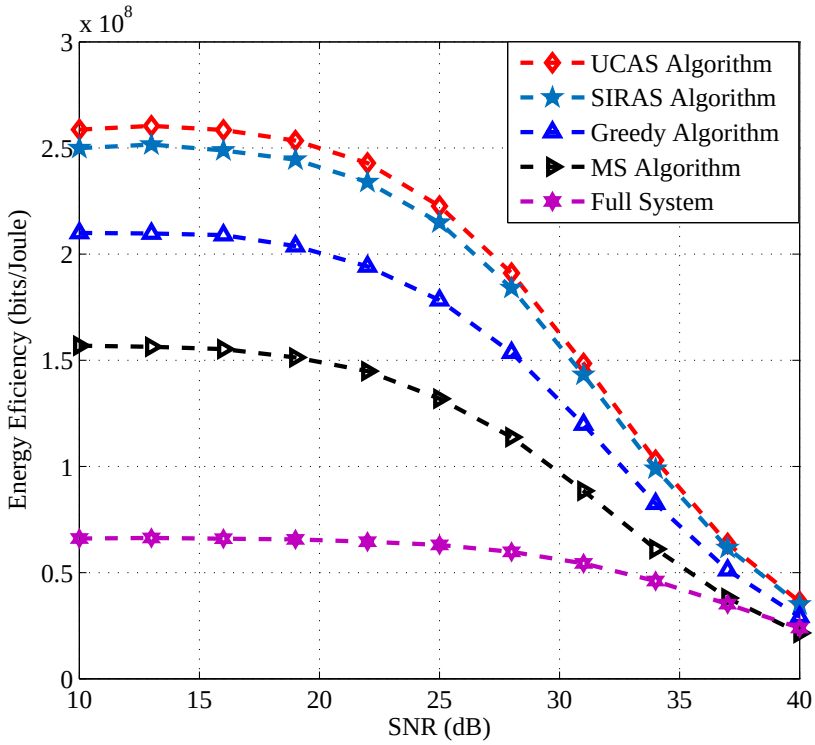


Figure 4.12: EE vs SNR (dB) for different AS schemes when $N = 256$, $N_s = 64$, and $K = 8$.

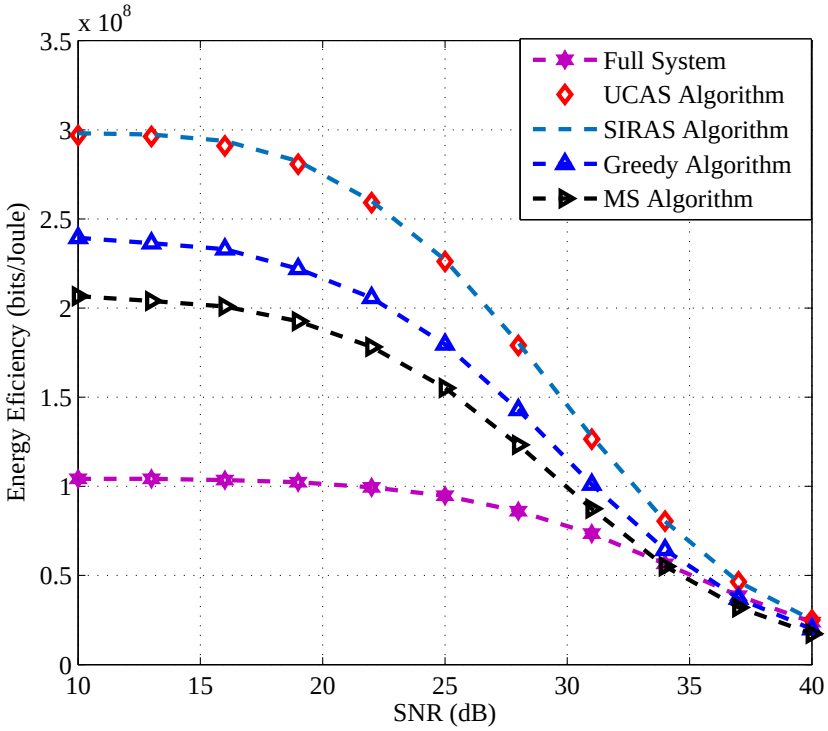


Figure 4.13: EE vs SNR (dB) for different AS schemes when $N = 128$, $N_s = 32$, and $K = 8$.

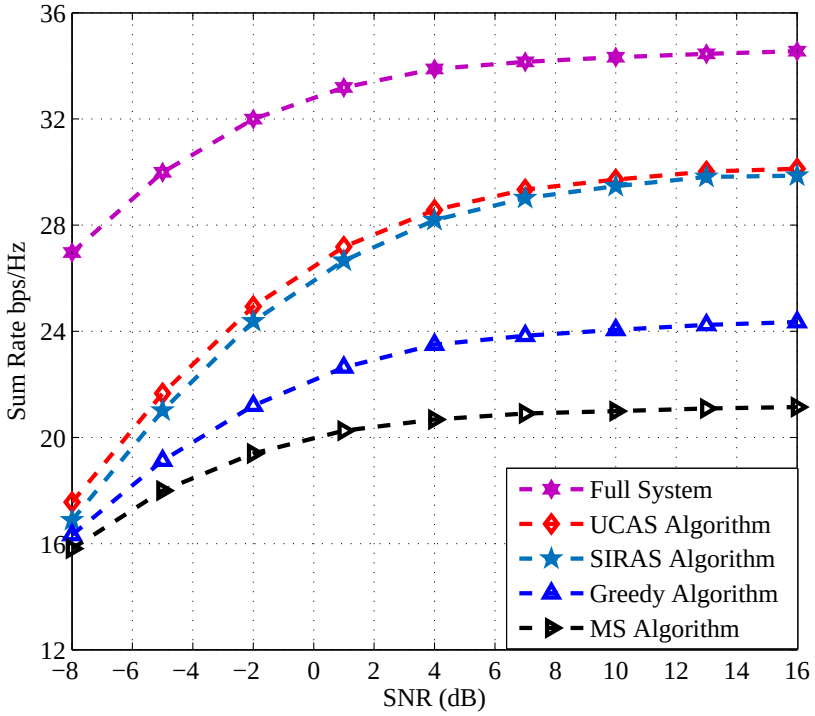


Figure 4.14: Sum rate vs SNR (dB) for different AS schemes when $N = 128$, $N_s = 32$, $K = 8$ and $\epsilon = 0.2$.

channel matrix $\widehat{\mathbf{H}}$ can be given as [129], [130]

$$\widehat{\mathbf{H}} = \mathbf{H} + \epsilon \boldsymbol{E}, \tag{4.26}$$

where $\epsilon \boldsymbol{E}$ represents the channel estimation error term, and is uncorrelated with $\mathbf{H}$. The entries of $\boldsymbol{E}$ are independent and identically distributed random variables with zero mean and variance of σ_h^2. Furthermore, ϵ controls the estimation accuracy, and $\epsilon = 0$ means the BS has a perfect CSI.

Figs. 4.14 and 4.15 demonstrate the performance of different AS schemes under imperfect CSI. In the aforementioned Figs., ϵ was set to 0.2, and the proposed method show small degradation compared to the perfect CSI case. Furthermore, from Fig. 4.15, the UCAS and SIRAS algorithms outperform the full system case at 2 dB and 7 dB, respectively, compared to 0 dB and 4 dB for perfect CSI.

Fig. 4.16 show the achievable rates of the proposed methods under different amount of channel estimation error. As expected, the achievable rates are effected by the estimation accuracy for both methods. In addition, at higher SNR, both methods becomes more sensitive under imperfect CSI. The reason behind this is that as the SNR increases, the noise effect becomes

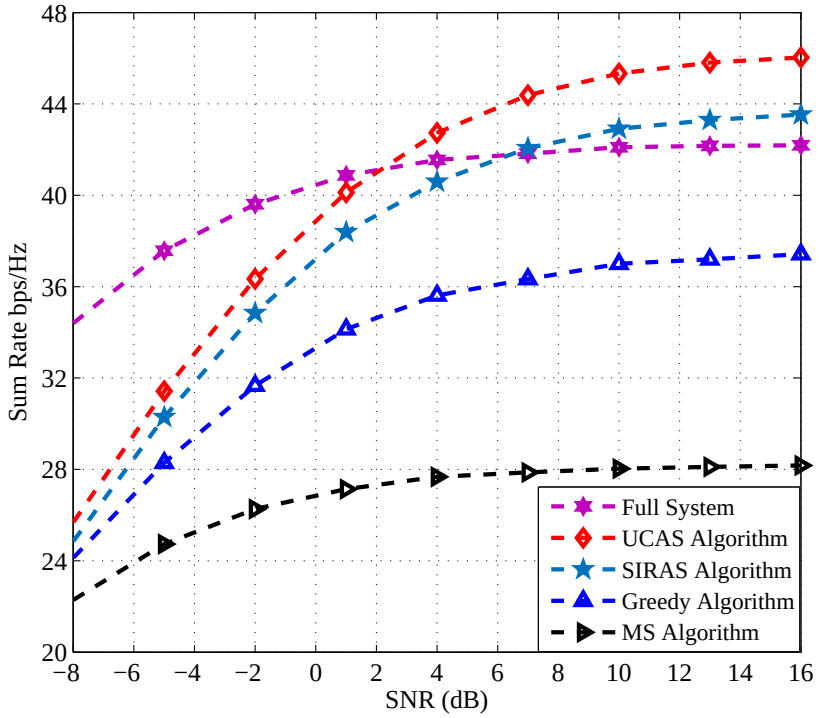


Figure 4.15: Sum rate vs SNR (dB) for different AS schemes when $N = 256$, $N_s = 64$, $K = 8$ and $\epsilon = 0.2$.

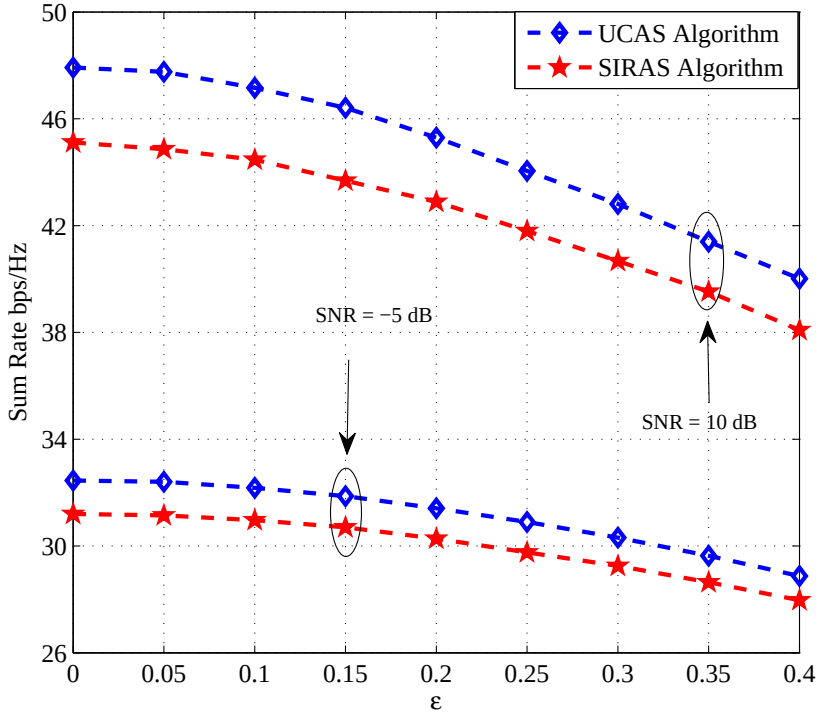


Figure 4.16: Sum rate vs ϵ for the proposed AS schemes when $N = 256$, $N_s = 64$, and $K = 8$

negligible, and the presence of CSI errors lead to the selection of sub-optimal antennas. For example, when the SNR $= -5\ dB$, the UCAS and SIRAS algorithms suffer a 3.57 and 3.23 bps/Hz rate loss when $\epsilon = 0.4$ compared to the perfect CSI case, respectively. While the achievable rates for the same algorithms degrade by 7.92 and 7.11 bps/Hz when $\epsilon = 0.4$, respectively, at SNR of 10 dB compared to perfect CSI case.

4.7 AS design considerations for different types of precoding schemes

In this chapter, the AS methods proposed were specifically designed for systems where MF precoding is applied. However, in certain circumstances, other types of linear precoding are preferred, and it is important to discuss the designing criteria of AS with different precoding schemes. In M-MIMO systems, one of the main challenges in designing any algorithm, especially iterative algorithms, is the associated computational complexity. Therefore, avoiding vector multiplications, as shown in this chapter, is extremely important to significantly reduce the number of FLOPs required.

For example, assuming the ZF precoder is applied, the designing criteria can target the maximization of the weights of the ZF precoder matrix, which can be given as follows [67]

$$\mathbf{W}^{ZF} = \gamma_{ZF}\mathbf{H}^{H}(\mathbf{H}\mathbf{H}^{H})^{-1}, \tag{4.27}$$

where γ_{ZF} is the normalization factor for ZF precoder, and it can be given as follows [67]

$$\gamma_{ZF} = \frac{1}{\sqrt{\text{Tr}[(\mathbf{H}\mathbf{H}^{H})^{-1}]}}. \tag{4.28}$$

Therefore, one way to design AS algorithm when ZF precoding is utilized, is through selecting the antennas that will maximize the normalization factor, which will lead to higher weights for the precoding matrix. A special matrix can also be used to avoid vector multiplications and reduce the computational complexity. In addition, efficient grouping of the antennas can also result in dramatic complexity reduction, due to the reduced search space, as demonstrated with the UCAS method. Therefore, antenna grouping can be considered for AS design with different types of linear precoding.

4.8 Chapter Summery

In this chapter, a single cell scenario was considered where a BS equipped with N antennas communicating with K single antenna users in the same time-frequency resources, with MF precoding. Two novel and highly efficient AS algorithms were proposed for the aforementioned scenario. Both algorithms were designed in a way such that no vector multiplications were involved during the iterative selection process, which in return resulted in a dramatic complexity reduction. A user-centric approach was adopted for the first algorithm, where the antennas at the BS are divided into K groups, with each group containing the N/K antennas that has the maximum channel gain for the correspondent user. This resulted in a search space reduction by a factor of K. The second algorithm focused on rejecting the interuser interference terms based on the signs of the interference terms only. For the user-centric method, the antenna that maximize the SINR for a certain user is selected at each iteration, while for the interference rejection method, the antenna that rejects the highest number of interference terms at a given iteration is selected. Several scenarios were considered, and our results demonstrated that when the number of users is 8, the UCAS algorithm outperforms the SIRAS method in terms of the achievable rates, while the latter showed better performance when 16 users were served. The performance of the proposed methods was compared with other low complexity AS schemes in terms of spectral and energy efficiencies, under perfect and imperfect CSI at the BS. Moreover, the implementation complexity was evaluated in terms of number of FLOPs, and the proposed methods showed a great performance-complexity trade-off.

Chapter 5

Cell-Edge-Aware Antenna Selection and Power Allocation for Multi-User M-MIMO Systems

5.1 Introduction

In realistic scenarios, the users are randomly distributed within the cell, and have different distances from the BS, and hence experiencing different large scale fading effects. PA techniques are employed to tackle this issue and make sure that each user will meet a predefined minimum QoS threshold. So far, the work on AS in M-MIMO systems have only focused on improving the overall performance of the system. In this chapter we design and utilize AS methods, in addition to PA, to enhance the rates of users who have low SINRs, such as cell-edge users.

A MU downlink scenario is considered in this chapter, where a BS employing a MF precoder is communicating with K single-antenna users in the same time-frequency resources. Our contribution in this chapter can be summarized as follow

1. We propose a low complexity Successive Gain Maximization (SGM) algorithm, where at each iteration, one antenna is chosen to maximize the gain of the user with minimum SNR. Compared to the conventional Maximum SNR based AS, we show that maximizing the gain successively not only increase the worst-case rate, but also the total sum rate for all users.

2. Relying on the channel correlation matrix, we propose two iterative, and low complexity antenna selection algorithms to minimize the inter-user interference. The first algorithm

is the Successive Interference Minimization (SIM), where at each iteration, one antenna is selected to minimize the interference for the user who is experiencing the highest interference. While the second algorithm is the Simplified Successive Interference Minimization (SSIM), where at a given iteration, one antenna is selected to minimize only the highest interference term in the channel correlation matrix.

3. The complexity of the three proposed algorithms is evaluated in terms of number of FLOPs and compared with other low complexity AS algorithms found in the literature. Our proposed methods showed promising trade-offs in terms of complexity and performance.

4. Our results demonstrate that minimizing the interference achieves higher rates than maximizing the channel gain. However, the latter achieves a considerably higher rate for the worst-case user at low SNRs.

5. Finally, we build upon the change of variables idea proposed in [131], to perform fair resource allocation between the users, by maximizing the worst-case rate. The complexity of the PA problem is also evaluated in terms of number of FLOPs.

5.2 System Model

A similar system model to that in Chapter 4 is considered in this work, where a BS equipped with N antennas serving K single-antenna users in the same time-frequency resources, and MF precoding is applied before transmitting the information symbols. A Time Division Duplex (TDD) scenario is considered, and the channel is assumed to be perfectly known at the BS through uplink pilot signals sent by the K users. However, in this chapter we consider a realistic scenario, where users are distributed randomly in the cell with minimum and maximum distances from the BS equal to $\mathrm{d}_{\min}$ and $\mathrm{d}_{\max}$, respectively, as illustrated in Fig. 5.1. Moreover, we develop low complexity AS algorithms to select N_s out of the available N antennas at the BS, to either minimize the interference, or to maximize the channel gain for the worst-case user. It is worth to mention that in this work we consider flat fading channels, which for example, has been considered previously with M-MIMO AS in [125], [43].

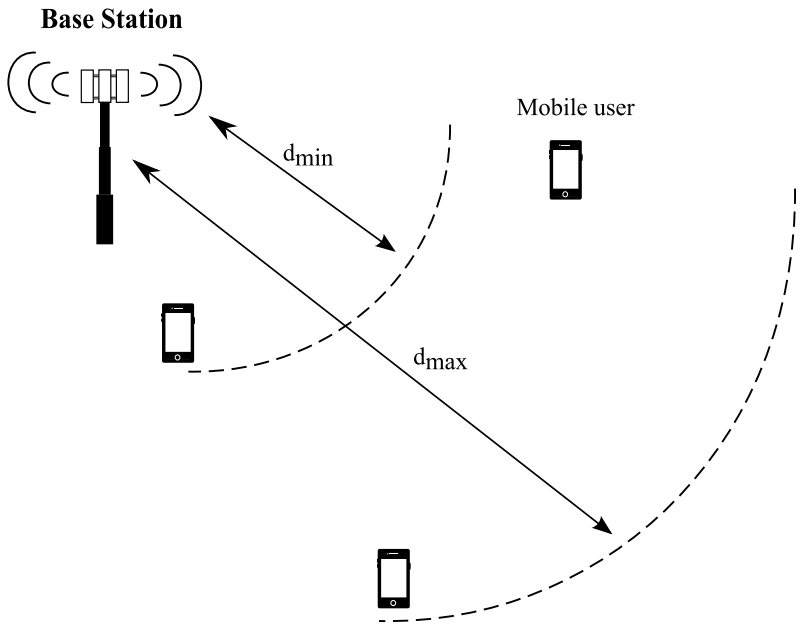


Figure 5.1: Uniformly distributed MU M-MIMO system.

5.3 Proposed antenna selection algorithms

In practical scenarios, users are distributed in the cell, and do not have equal distances from the BS. Consequently, users who are close to the BS will have higher SINRs, and hence better rates, while those who are on the edge of the cell will experience weak service. Therefore, selecting the antennas for all the users at once will neither be optimal, nor fair, since the selection will be highly influenced by users with high SINRs. In our work, only one antenna is selected at each iteration, and the selection is done either to maximize the channel gain, or to minimize the interference, of the worst-case user in that iteration. The three algorithms rely on the channel correlation matrix $\mathbf{U} = \mathbf{H}\mathbf{H}^{\mathbf{H}}$, which can be expressed as

$$\mathbf{U} = \begin{pmatrix} \mathbf{h}_1^T\mathbf{h}_1^* & \mathbf{h}_1^T\mathbf{h}_2^* & \cdots & \cdots & \mathbf{h}_1^T\mathbf{h}_K^* \\ \mathbf{h}_2^T\mathbf{h}_1^* & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \mathbf{h}_{K-1}^T\mathbf{h}_K^* \\ \mathbf{h}_K^T\mathbf{h}_1^* & \cdots & \cdots & \cdots & \mathbf{h}_K^T\mathbf{h}_K^* \end{pmatrix}, \tag{5.1}$$

where the k^{th} element in the diagonal of $\mathbf{U}$ represents the channel gain for k^{th} user, while the upper and lower triangle parts of $\mathbf{U}$ are directly related to the interference between the users. The proposed algorithms aim to achieve high performance, fairness, as well as low

computational complexity by avoiding vector multiplications to reduce the number of FLOPs.

5.3.1 Successive gain maximization AS algorithm

The SGM algorithm aims to maximize the diagonal of $\mathbf{U}$, where each element of the diagonal represents the channel gain for a given user k before normalization, and it can be expressed as

$$\begin{aligned}\mathbf{h}_k^T\mathbf{h}_k^* &= h_{k,1}h_{k,1}^* + h_{k,2}h_{k,2}^* + ... + h_{k,N}h_{k,N}^* \\ &= \sum_{n=1}^{N}\left(\Re\{h_{k,n}\}^2 + \Im\{h_{k,n}\}^2\right). \end{aligned} \tag{5.2}$$

To reduce the number of FLOPs, the values of the N multiplications in (5.2) are stored in one row of the matrix $\boldsymbol{V} \in \mathbb{R}^{K\times N}$, which can be expressed as

$$\boldsymbol{V} = \begin{pmatrix} v_{1,1} & \dots & \dots & v_{1,N} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ v_{K,1} & \dots & \dots & v_{K,N} \end{pmatrix}, \tag{5.3}$$

where $v_{1,1} = h_{1,1}h_{1,1}^*$, $v_{1,N} = h_{1,N}h_{1,N}^*$, $v_{K,1} = h_{K,1}h_{K,1}^*$, and $v_{K,N} = h_{K,N}h_{K,N}^*$. The first antenna is then selected based on the maximum channel gain as follow

$$\xi^{[0]} = \arg\max_{n\in 1:N} \|\mathbf{h}_n^c\|, \tag{5.4}$$

after selecting the first antenna, a vector $\boldsymbol{\phi} \in \mathbb{R}^{K\times 1}$ will be initialized with the K values in $\boldsymbol{V}$ associated with the selected antenna, and will be updated after selecting each of subsequent antennas. The $N_s - 1$ iterations will then start, and at each iteration, the antenna that will maximize the gain for the worst-case user, i.e. the user with lowest SNR, will be chosen as follow

$$\xi^{[t]} = \arg\max_{n\in\eta} \boldsymbol{v}_{u^*,n}, \tag{5.5}$$

where $\xi^{[t]}$ is the antenna selected at iteration t, η is a set of available antennas, and u^* is the user with lowest SNR at a given iteration t, and can be expressed as $u^* = \arg\min \boldsymbol{\phi}^{[t-1]}$, since the values in $\boldsymbol{\phi}^{[t-1]}$ are the channel gains for each user before normalization. Afterwards, the vector $\boldsymbol{\phi}$ will be updated as follow

$$\boldsymbol{\phi}^{[t]} = \boldsymbol{\phi}^{[t-1]} + \boldsymbol{v}^c_{\xi^{[t]}}. \tag{5.6}$$

Algorithm 12 Proposed SGM AS algorithm

Input K, N, N_s, $\mathbf{H}$,
Initialize:
  $\boldsymbol{V} = \mathbf{0}_{K\times N}$, $\boldsymbol{\phi} = \mathbf{0}_{K\times 1}$,
  $\chi = \mathbf{0}_{N\times 1}$, ($\chi$ sets the locations of selected antennas to "1"),
  $\eta = 1:N$, (set of available antennas),
for $k = 1 \leftarrow K$
  for $n = 1 \leftarrow N$
    $v_{k,n} = \Re\{h_{k,n}\}^2 + \Im\{h_{k,n}\}^2$,
  end for
end for
$\xi^{[0]} = \arg\max_{n\in 1:N} \|\mathbf{h}_n^c\|$,
$\boldsymbol{\phi}^{[0]} = \boldsymbol{v}_{\xi^{[0]}}^c$,
$\chi_{\xi^{[0]}} = 1$,
$\eta = \eta \setminus \xi^{[0]}$,
for $t = 1 \leftarrow N_s$-1
  $u^* = \arg\min \boldsymbol{\phi}^{[t-1]}$,
  $\xi^{[t]} = \arg\max_{n\in\eta} v_{u^*,n}$,
  $\boldsymbol{\phi}^{[t]} = \boldsymbol{\phi}^{[t-1]} + \boldsymbol{v}_{\xi^{[t]}}^c$,
  $\chi_{\xi^{[t]}} = 1$,
  $\eta = \eta \setminus \xi^{[t]}$,
end for
Output: $[\mathbf{H}_j^c]_{j\notin\eta}$

The same procedure will be repeated for subsequent iterations until maximum number of selected antennas is reached. The SGM method is described in detail in Algorithm (12).

5.3.2 Successive interference minimization antenna selection algorithm

The main drawback of MF precoding is the inter-user interference, and it causes dramatic performance degradation, especially for a BS serving large number of users. The SIM algorithm aims to minimize the inter-user interference by minimizing the terms in the upper and lower triangle parts of (5.1). If we neglect the normalization at this point, the interference caused by user j to user k can be expressed as

$$\begin{aligned} \left|\mathbf{h}_k^T \mathbf{h}_j^*\right|^2 &= \left|h_{k,1}h_{j,1}^* + h_{k,2}h_{j,2}^* + ... + h_{k,N}h_{j,N}^*\right|^2 \\ &= \left|\sum_{n=1}^{N} h_{k,n}h_{j,n}^*\right|^2, \end{aligned} \tag{5.7}$$

while the total interference for user k can be given as

$$\Upsilon_k = \sum_{\substack{j=1 \\ j\neq k}}^{K} \left|\mathbf{h}_k^T \mathbf{h}_j^*\right|^2. \tag{5.8}$$

For a system serving K users, there are L interference terms, where $L = K^2 - K$. Each of these L terms is an addition of N complex number multiplications. In our work, we store the $L \times N$ complex multiplication values in $\mathbf{\Theta} \in \mathbb{C}^{L \times N}$ and use them at later stages when selecting the antennas. The matrix $\mathbf{\Theta}$ can be expressed as

$$\mathbf{\Theta} = \begin{pmatrix} \theta_{1,1} & \dots & \dots & \theta_{1,N} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \theta_{K-1,1} & \dots & \dots & \theta_{K-1,N} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \theta_{K^2-K,1} & \dots & \dots & \theta_{K^2-K,N} \end{pmatrix}, \tag{5.9}$$

where $\theta_{1,1} = h_{1,1}h^*_{2,1}$, $\theta_{1,N} = h_{1,N}h^*_{2,N}$, $\theta_{K-1,1} = h_{1,1}h^*_{K,1}$, $\theta_{K-1,N} = h_{1,1}h^*_{K,N}$, $\theta_{K^2-K,1} = h_{K,1}h^*_{K-1,1}$, and $\theta_{K^2-K,N} = h_{K,N}h^*_{K-1,N}$. The algorithm then starts by selecting the antenna with highest channel norms. After that, the L interference terms in $\mathbf{\Theta}$ associated with the selected antenna are used to initialize the vector $\boldsymbol{\lambda} \in \mathbb{C}^{L \times 1}$. It is worth to mention that, as shown in (5.9), the $(K-1)$ interference terms for the k^{th} user are indexed between rows $(k-1)(K-1)+1$ and $(k-1)(K-1)+K-1$, in both $\mathbf{\Theta}$ and $\boldsymbol{\lambda}$. The $N_s - 1$ iterations will then start, where at each iteration, the algorithm finds the user with highest interference, and then selects one antenna that will reduce the interference for that user to its minimum value as follow

$$k^* = \arg \max_{1 \le k \le K} \sum_{i=1}^{K-1} \left| \lambda_{\varepsilon_k + i} \right|^2, \tag{5.10}$$

where $\varepsilon_k = (k-1)(K-1)$, and k^* is the user with the highest interference at a given iteration t. The selected antenna can be expressed as

$$\xi^{[t]} = \arg \min_{ns \in \eta} \sum_{i=1}^{K-1} \left| \lambda_{\varepsilon^* + i} + \theta_{\varepsilon^* + i, ns} \right|^2, \tag{5.11}$$

where $\varepsilon^* = (k^* - 1)(K-1)$, $\xi^{[t]}$ is the selected antenna at iteration t. The interference vector $\boldsymbol{\lambda}$ is updated after selecting each antenna as follow

$$\boldsymbol{\lambda}^{[t]} = \boldsymbol{\lambda}^{[t-1]} + \boldsymbol{\theta}^c_{\xi^{[t]}}, \tag{5.12}$$

the same procedure will be repeated for $N_s - 1$ iterations. The steps of the SIM method are

shown in Algorithm (13).

Algorithm 13 Proposed SIM AS algorithm

Input K, N, N_s, $\mathbf{H}$, $L = K^2 - K$,
Initialize:
  $\boldsymbol{\Theta} = \mathbf{0}_{L\times N}$, $\boldsymbol{\lambda} = \mathbf{0}_{L\times 1}$,
  $\chi = \mathbf{0}_{N\times 1}$, $\eta = 1:N$, and $l = 0$,
for $i = 1 \rightarrow K$
  for $j = 1 \rightarrow K$ & $j \neq i$
    $l \leftarrow l + 1$,
    for $n = 1 \rightarrow N$
      $\theta_{l,n} = h_{i,n} h^*_{j,n}$,
    end for
  end for
end for
$\xi^{[0]} = \arg\max_{n\in 1:N} \|\mathbf{h}^c_n\|$,
$\chi_{\xi^{[0]}} = 1$,
$\eta = \eta \setminus \xi^{[0]}$,
$\boldsymbol{\lambda}^{[0]} = \boldsymbol{\theta}^c_{\xi^{[0]}}$,
for $t = 1 \rightarrow N_s - 1$
  for $k = 1 \rightarrow K$
    $\varepsilon_k = (k-1)(K-1)$,
    $\Upsilon_k = \sum_{i=1}^{K-1} |\lambda_{\varepsilon_k + i}|^2$,
  end for
  $k^* = \arg\max \Upsilon$,
  $\xi^{[t]} = \arg\min_{ns\in\eta} \sum_{i=1}^{K-1} |\lambda_{\varepsilon^*+i} + \theta_{\varepsilon^*+i,ns}|^2$,
  $\boldsymbol{\lambda}^{[t]} = \boldsymbol{\lambda}^{[t-1]} + \boldsymbol{\theta}^c_{\xi^{[t]}}$,
  $\chi_{\xi^{[t]}} = 1$,
  $\eta = \eta \setminus \xi^{[t]}$,
end for
Output: $[\mathbf{H}^c_j]_{j\notin\eta}$

5.3.3 Simplified successive interference minimization antenna selection algorithm

The SSIM method is similar to the SIM algorithm, with less complexity. Instead of minimizing the $(K-1)$ interference terms for a certain user each iteration, the SSIM minimizes only the maximum interference term in $\mathbf{U}$. Moreover, the elements in the upper triangle part are the complex conjugate of the elements in lower triangle part of $\mathbf{U}$, i.e. $\mathbf{h}_i^T \mathbf{h}_j^* = (\mathbf{h}_j^T \mathbf{h}_i^*)^*$, hence

$\left|\mathbf{h}_i^T \mathbf{h}_j^*\right|^2 = \left|\mathbf{h}_j^T \mathbf{h}_i^*\right|^2$. Therefore, reducing the interference in the upper triangle part leads to the exact same interference reduction in the lower triangle part and vice versa. Thus, only M interference terms need to be optimized, where $M = L/2$. In our work, we optimize the upper triangle part of $\mathbf{U}$. The $M \times N$ complex multiplication values are stored in $\mathbf{\Psi} \in \mathbb{C}^{M \times N}$, which can be represented as

$$\mathbf{\Psi} = \begin{pmatrix} \psi_{1,1} & \cdots & \cdots & \psi_{1,N} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \psi_{M,1} & \cdots & \cdots & \psi_{M,N} \end{pmatrix}, \tag{5.13}$$

where $\psi_{1,1} = h_{1,1} h_{2,1}^*$, $\psi_{1,N} = h_{1,N} h_{2,N}^*$, $\psi_{M,1} = h_{K-1,1} h_{K,1}^*$, and $\psi_{M,N} = h_{K-1,N} h_{K,N}^*$. The antenna that has maximum channel norms will be selected as the first antenna, and a vector $\boldsymbol{\omega} \in \mathbb{C}^{M \times 1}$ will then be initialized with the M inter-user interference terms in $\mathbf{\Psi}$ associated with the first selected antenna. Each term in $\boldsymbol{\omega}$ is a function of the interference between two users i and j, where $i, j = 1, ..., K$, and $i \neq j$. The algorithm then starts its $N_s - 1$ iterations. At each iteration, the algorithm finds the maximum term in $\boldsymbol{\omega}$, and then selects the antenna that will reduce the interference between the two given users to its minimum value. This can analytically be expressed as

$$\xi^{[t]} = \arg\min_{ns \in \eta} \left|\omega_{m^\star} + \psi_{m^\star, ns}\right|^2, \tag{5.14}$$

where $m^\star$ is the maximum term in $\boldsymbol{\omega}$, $\xi^{[t]}$ is the selected antenna, η is the set of available antennas, and t is the current iteration. After each iteration, the M interference terms in $\boldsymbol{\omega}$ are updated with the interference values caused by adding the new selected antenna as follow

$$\boldsymbol{\omega}^{[t]} = \boldsymbol{\omega}^{[t-1]} + \boldsymbol{\psi}^c_{\xi[t]}, \tag{5.15}$$

the same procedure will be repeated until the N_s antennas are selected. The steps of SSIM method are described in detail in Algorithm (14).

5.4 Sum rate optimization through power allocation

After selecting the antennas, we shift our focus to optimize the allocation of total available power at the BS. Recalling the formula in (4.6), the sum rate for a downlink multi-user system

Algorithm 14 Proposed SSIM AS algorithm

Input K, N, N_s, $\mathbf{H}$, $M = (K^2 - K)/2$,
Initialize
  $\boldsymbol{\Psi} = \mathbf{0}_{M\times N}$, $\boldsymbol{\omega} = \mathbf{0}_{M\times 1}$,
  $\chi = \mathbf{0}_{N\times 1}$, $\eta = 1 : N$, $m = 0$,
for $i = 1 \rightarrow K - 1$
  for $j = i + 1 \rightarrow K$
    $m = m + 1$,
    for $n = 1 \rightarrow N$
      $\psi_{m,n} = h_{i,n} h^*_{j,n}$,
    end for
  end for
end for
$\xi^{[0]} = \arg\max_{n\in 1:N} \|\mathbf{h}^c_n\|$,
$\chi_{\xi^{[0]}} = 1$,
$\eta = \eta \setminus \xi^{[0]}$,
$\boldsymbol{\omega}^{[0]} = \boldsymbol{\psi}^c_{\xi^{[0]}}$,
for $t = 1 \rightarrow N_s - 1$
  $m^\star = \arg\max \boldsymbol{\omega}^{[t-1]}$,
  $\xi^{[t]} = \arg\min_{ns\in\eta} |\omega_{m^\star} + \psi_{m^\star,ns}|^2$,
  $\boldsymbol{\omega}^{[t]} = \boldsymbol{\omega}^{[t-1]} + \boldsymbol{\psi}^c_{\xi^{[0]}}$,
  $\chi_{\xi^{[t]}} = 1$,
  $\eta = \eta \setminus \xi^{[t]}$,
end for
Output: $[\mathbf{H}^c_j]_{j\notin\eta}$

can be expressed as

$$\begin{aligned} R &= \sum_{k=1}^{K} R_k = \sum_{k=1}^{K} \log_2(1+\gamma_k) \\ &= \log_2\left(\prod_{k=1}^{K}(1+\gamma_k)\right) \\ &= \log_2\left(\prod_{k=1}^{K}\left[\frac{\sum_{i=1}^{K} p_i \left|\mathbf{h}_k^T\mathbf{w}_i\right|^2 + \sigma_n^2}{\sum_{\substack{i=1\\ i\neq k}}^{K} p_i \left|\mathbf{h}_k^T\mathbf{w}_i\right|^2 + \sigma_n^2}\right]\right). \end{aligned} \tag{5.16}$$

Maximizing the total sum rate will result in allocating high power to users with high SINRs. In contrast, users located far away from the BS will be allocated low or no power at all. Therefore, maximizing the total sum rate is neither realistic, nor fair. To achieve fairness, more power should be allocated to users experiencing poor channel conditions, this approach is known as max-min optimization. However, complete fairness between the users might be costly, and the total sum rate will experience dramatic degradation. Therefore, we add a per user power constraint, such that no user should be allocated less than a predefined threshold of the total available power. This way, we aim to maximize the rate for the worst-case user without causing large degradation in the total sum rate. The problem can be formulated as

$$\underset{\{p_1,...,p_K\}}{\text{maximize}} \quad \min_k \quad R_k$$

subject to

$$\sum_{k=1}^{K} p_k \leq P_T, \tag{5.17a}$$

$$p_k \geq \frac{P_T}{K}\,\zeta,\ \forall k, \tag{5.17b}$$

where p_k is the power allocated for the k^{th} user, P_T is the total transmission power in Watts available at the BS, ζ is a parameter that controls the per-user power constraint threshold, and $0 < \zeta < 1$. Due to the inter-user interference, the objective function in (5.17) is a non-convex problem. Therefore, we use the idea proposed by [131], which relies on change of variables, as follow

$$e^{u_k} = \sum_{i=1}^{K} p_i \left|\mathbf{h}_k^T \mathbf{w}_i\right|^2 + \sigma_n^2,\ \forall k, \tag{5.18a}$$

$$e^{s_k} = \sum_{\substack{i=1 \\ i\neq k}}^{K} p_i \left|\mathbf{h}_k^T \mathbf{w}_i\right|^2 + \sigma_n^2,\ \forall k. \tag{5.18b}$$

After applying simple logarithmic and exponential functions, the objective function in (5.17) can be expressed as maximize$\{\min_k\ u_k - s_k\}$. Where $\boldsymbol{u} = [u_1, ..., u_K]^T$ and $\boldsymbol{s} = [s_1, ..., s_K]^T$ are slack variables. Moreover, the constraints of u_k and s_k are the expressions in the right hand side of (5.18a) and (5.18b), respectively. Therefore, the optimization problem can be re-written as follow

$$\underset{\{p_1,...,p_K\},\, \boldsymbol{u},\, \boldsymbol{s}}{\text{maximize}} \quad \min_{k} \;\; u_k - s_k$$

subject to

$$\sum_{i=1}^{K} p_i \left|\mathbf{h}_k^T \mathbf{w}_i\right|^2 + \sigma_n^2 \geq e^{u_k},\ \forall k, \tag{5.19a}$$

$$\sum_{\substack{i=1 \\ i\neq k}}^{K} p_i \left|\mathbf{h}_k^T \mathbf{w}_i\right|^2 + \sigma_n^2 \leq e^{s_k},\ \forall k, \tag{5.19b}$$

$$\sum_{k=1}^{K} p_k \leq P_T, \tag{5.19c}$$

$$p_k \geq \frac{P_T}{K}\,\zeta,\ \forall k. \tag{5.19d}$$

The only non-convex constraint in (5.19) is (5.19b). Therefore, the exponential term e^{s_k} needs to be linearised. To tackle this problem, a first order Taylor approximation is applied, such that $e^{s_k} = e^{\bar{s}_k}(s_k - \bar{s}_k + 1), \forall k$. Where the linearization is made around the points $\bar{\boldsymbol{s}} = [\bar{s}_1, ..., \bar{s}_K]^T$. Therefore, problem (5.19) can be reformulated to

$$\underset{\{p_1,...,p_K\},\, \boldsymbol{u},\, \boldsymbol{s}}{\text{maximize}} \quad \min_{k} \;\; u_k - s_k$$

subject to

$$\sum_{\substack{i=1 \\ i\neq k}}^{K} p_i \left|\mathbf{h}_k^T \mathbf{w}_i\right|^2 + \sigma_n^2 \leq e^{\bar{s}_k}(s_k - \bar{s}_k + 1),\ \forall k, \tag{5.20a}$$

$$\text{In addition to constraints } (5.19a),\ (5.19c),\ (5.19d). \tag{5.20b}$$

At this point, all the constraints are convex, and the optimization problem in (5.20) can be solved iteratively, as shown in Algorithm (15), using the CVX software [132]. It is worth mentioning that the initial values of $\{\bar{s}_1, ..., \bar{s}_K\}$ for the first iteration, were obtained based on Equal Power Allocation (EPA) criteria, and can be expressed as follow

$$\bar{s}_k = \log\left(\frac{P_T}{K}\sum_{\substack{i=1\\ i\neq k}}^{K}\left|\mathbf{h}_k^T\mathbf{w}_i\right|^2 + \sigma_n^2\right),\ \forall k. \tag{5.21}$$

Algorithm 15 Iterative algorithm to solve problem (5.20)

1: **Initialize**:
2: $\bar{\boldsymbol{s}}^{[0]}$, $t = 1$, $\tau = 0.001$, (τ is the error tolerance)
3: **define** err $\triangleq \sum_{k=1}^{K} |s_k - \bar{s}_k|$,
4: **while** err $> \tau$ **do**
5: Solve (5.20) and Calculate $\boldsymbol{s}^{[t]}$,
6: Update $\bar{\boldsymbol{s}}^{[t]} = \boldsymbol{s}^{[t]}$,
7: Increment $t = t + 1$,
8: **end while**

5.5 Complexity Analysis of PA and AS Techniques

In this section, we evaluate the computational complexity of the max-min PA (MMPA) and AS algorithms used in this work. Since we perform power optimization after choosing the antennas, their complexity will be evaluated separately.

5.5.1 Computational complexity of the proposed AS algorithms

In this subsection, the complexity of the proposed AS algorithms used in this chapter is evaluated in terms of number of FLOPs. The same complexity analysis used in chapter 4 are followed in this section. Furthermore, the number of additions and multiplications for the proposed methods are given for each operator in Table 5.1.

5.5.1.1 Complexity analysis of the SGM algorithm

The SGM algorithm starts by evaluating the matrix $\boldsymbol{V}$ which requires $9KN$ FLOPs, the first antenna is then selected which results in $N(10K + 4)$ FLOPs. The $N_s - 1$ iterations will then start, and finding the user with minimum channel gain requires $K(N_s - 1)$ FLOPs, and selecting the $N_s - 1$ antennas takes $\sum_{l=1}^{N_s-1}(N-l)$ FLOPs. Finally, updating the vector $\boldsymbol{\phi}$ results in $K(N_s - 1)$ FLOPs. Therefore, the total complexity for the SGM algorithm is

$$\mathcal{C}_{SGM} = N(19K + 4) + 2K(N_s - 1) + \sum_{l=1}^{N_s-1}(N - l). \tag{5.22}$$

5.5.1.2 Complexity analysis of the SIM algorithm

The SIM algorithm requires $18N(K^2-K)$ FLOPs to obtain the matrix $\mathbf{\Theta}$, followed by $N(10K+4)$ FLOPs for selecting the first antenna $\xi^{[0]}$. During the N_s-1 iterations, finding the user with maximum interference requires $(N_s-1)(18K^2-12K)$ FLOPs, while selecting the (N_s-1) antennas result in $\sum_{l=1}^{N_s-1}(N-l)(20K-20)$ FLOPs. Finally, updating the vector $\boldsymbol{\lambda}$ requires $2(N_s-1)(K^2-K)$ FLOPs. Therefore, the total number of FLOPs required for the SIM algorithm can be given as

$$\begin{aligned} \mathcal{C}_{SIM} &= N(18K^2-8K+4)+(N_s-1)(20K^2-14K) \\ &+ \sum_{l=1}^{N_s-1}(N-l)(20K-20). \end{aligned} \tag{5.23}$$

5.5.1.3 Complexity analysis of the SSIM algorithm

Evaluating the matrix $\mathbf{\Psi}$ requires $9N(K^2-K)$ FLOPs, which followed by $N(10K+4)$ FLOPs for selecting the first antenna. During the N_s-1 iterations, finding the highest interference term requires $7(K^2-K)(N_s-1)$ FLOPs, while selecting the antennas results in $\sum_{l=1}^{N_s-1}20(N-l)$ FLOPs. Finally, updating $\boldsymbol{\omega}$ results in $(K^2-K)(N_s-1)$ FLOPs. The total complexity of the SSIM method can therefore given as

$$\begin{aligned} \mathcal{C}_{SSIM} &= N(9K^2+K+4)+8(N_s-1)(K^2-K) \\ &+ \sum_{l=1}^{N_s-1}20(N-l). \end{aligned} \tag{5.24}$$

Figs. 5.2 shows that when the number of antennas at the BS is 256, and out of which 64 are selected by the different AS schemes, the MS algorithm requires the lowest complexity for any number of users served in the cell. In contrast, the greedy algorithm has the highest complexity among the AS schemes used in this work for the given scenario. For the proposed AS method, the SGM has the lowest complexity since it only deals with real numbers during the iterative selection process. In contrast, the SSIM has the highest complexity between the proposed algorithms, since at a given iteration, the available antennas are tested to minimize the total interference for the worst case user. Finally, the SSIM has a lower complexity than the

Table 5.1: Number of Additions and Multiplications for the proposed AS schemes

Algorithm	Operator	Additions	Multiplications
SGM	$\boldsymbol{V}$	KN	$2KN$
	$\xi^{[0]}$	$2KN$	$N(2K+1)$
	u^*	$K(N_s-1)$	$--$
	$\xi^{[t]}, \forall t \in \{1,...,N_s-1\}$	$\sum_{l=1}^{N_s-1}(N-l)$	$--$
	$\boldsymbol{\phi}^{[t]}, \forall t \in \{1,...,N_s-1\}$	$K(N_s-1)$	$--$
	Total	$3KN+2K(N_s-1)+\sum_{l=1}^{N_s-1}(N-l)$	$N(4K+1)$
SIM	$\boldsymbol{\Theta}$	$2N(K^2-K)$	$4N(K^2-K)$
	$\xi^{[0]}$	$2KN$	$N(2K+1)$
	k^*	$2K^2(N_s-1)$	$(N_s-1)(4K^2-3K)$
	$\xi^{[t]}, \forall t \in \{1,...,N_s-1\}$	$\sum_{l=1}^{N_s-1}(N-l)(4K-4)$	$\sum_{l=1}^{N_s-1}(N-l)(4K-4)$
	$\boldsymbol{\lambda}^{[t]}, \forall t \in \{1,...,N_s-1\}$	$2(N_s-1)(K^2-K)$	$--$
	Total	$2NK^2+(N_s-1)(4K^2-2K)+\sum_{l=1}^{N_s-1}(N-l)(4K-4)$	$N(4K^2-2K+1)+(N_s-1)(4K^2-3K)+\sum_{l=1}^{N_s-1}(N-l)(4K-4)$
SSIM	$\boldsymbol{\Psi}$	$N(K^2-K)$	$2N(K^2-K)$
	$\xi^{[0]}$	$2KN$	$N(2K+1)$
	m^*	$(K^2-K)(N_s-1)$	$\frac{3}{2}(K^2-K)(N_s-1)$
	$\xi^{[t]}, \forall t \in \{1,...,N_s-1\}$	$\sum_{l=1}^{N_s-1}4(N-l)$	$\sum_{l=1}^{N_s-1}4(N-l)$
	$\boldsymbol{\omega}^{[t]}, \forall t \in \{1,...,N_s-1\}$	$(K^2-K)(N_s-1)$	$--$
	Total	$N(K^2+K)+2(N_s-1)(K^2-K)+\sum_{l=1}^{N_s-1}4(N-l)$	$N(2K^2+1)+\frac{3}{2}(K^2-K)(N_s-1)+\sum_{l=1}^{N_s-1}4(N-l)$

SIM algorithm, since the SSIM tries to minimize only the highest interference term in $\mathbf{U}$. Both SIM and SSIM algorithms deal with complex numbers, which requires more FLOPs than the SGM algorithm, which only deals with real number operations.

Fig. 5.3 demonstrates the complexity when increasing the number of antennas at the BS. The greedy algorithm requires the highest complexity followed by the SIM scheme when the number of selected antennas is $N/2$. However, reducing the number of selected antennas to $N/8$ results in higher complexity requirement for the SIM method between the applied AS schemes when the number of users is 8, as demonstrated in Fig.5.4. Furthermore, regardless of the number of selected antennas, the MS has the lowest complexity, followed by SGM, and then SSIM method.

Fig. 5.5 shows the complexity in number of FLOPs for the different AS schemes for a different number of selected antennas when the BS is equipped with 128 to serve 8 users. Unlike the proposed methods, the complexity of the greedy algorithm is highly affected by increasing the number of selected antennas, and hence requires the highest complexity when the number of selected antennas is higher than 30. In contrast, the MS shows the lowest complexity and it does not depend on the number of selected antennas.

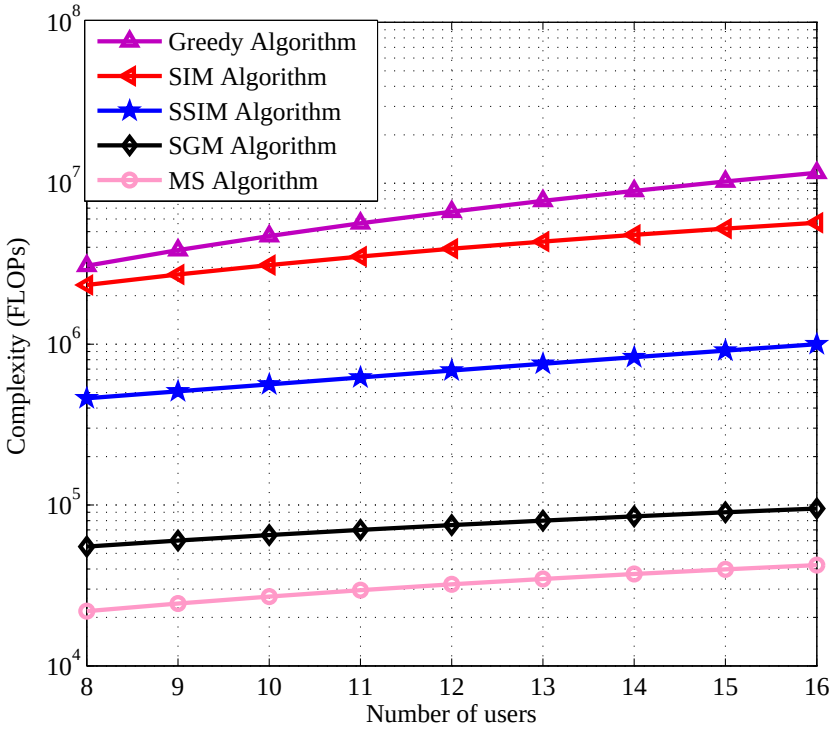


Figure 5.2: Complexity of different algorithms in number of FLOPs, for $N = 256$ and $N_s = 64$.

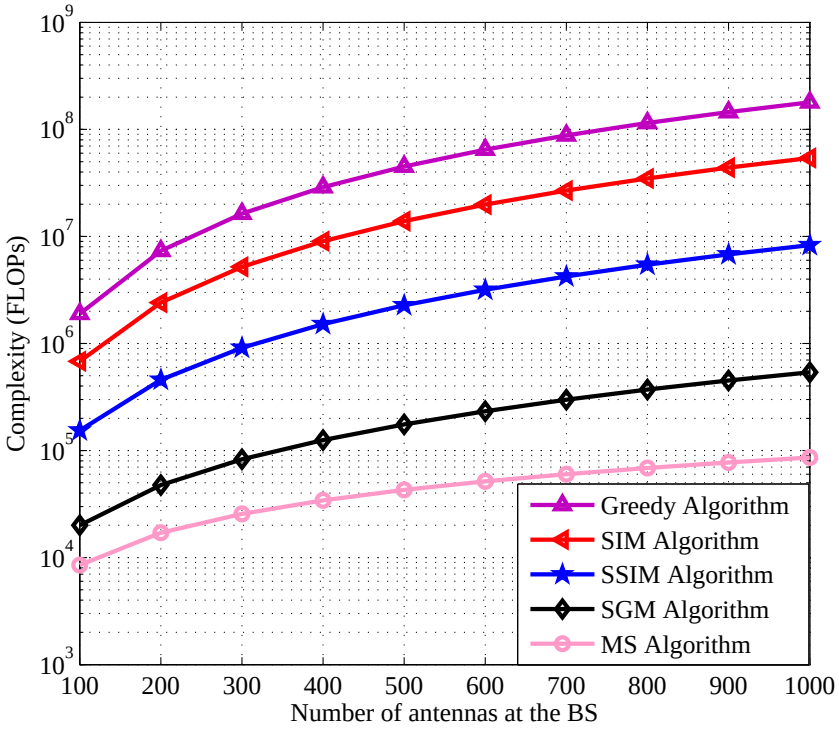


Figure 5.3: Complexity of different algorithms in number of FLOPs, for $N_s = N/2$ and $K = 8$.

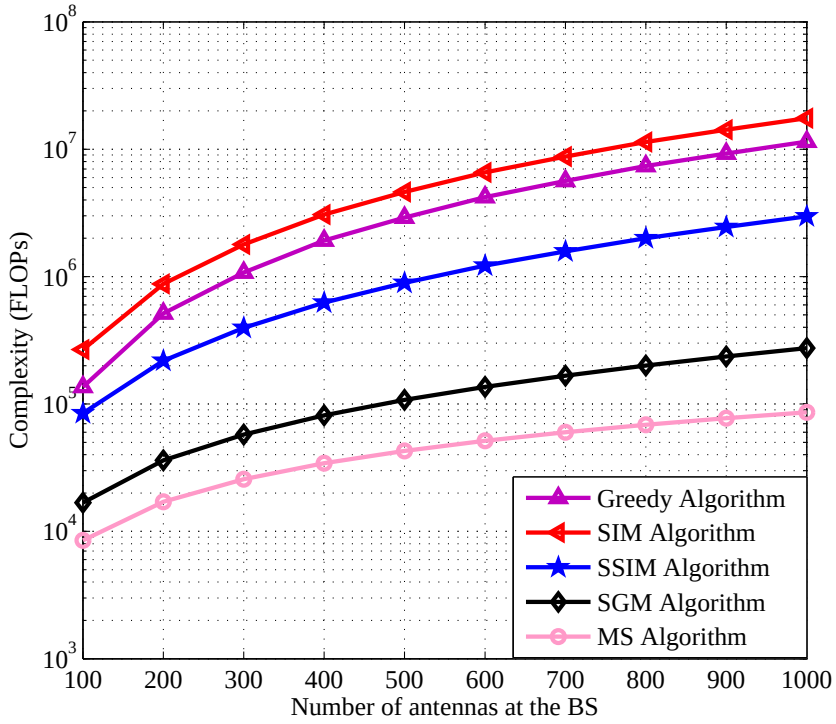


Figure 5.4: Complexity of different algorithms in number of FLOPs, for $N_s = N/8$ and $K = 8$.

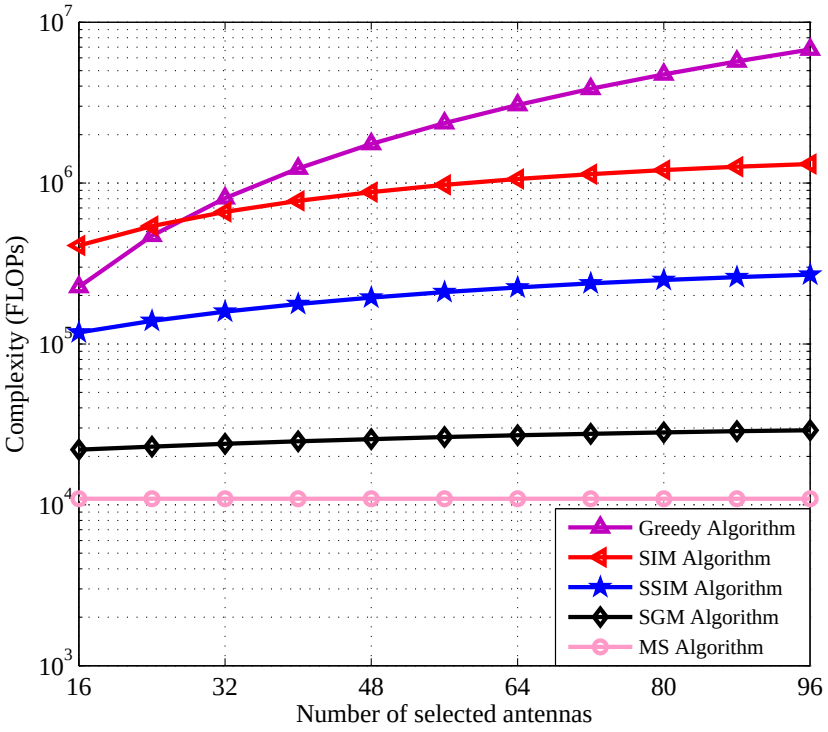


Figure 5.5: Complexity of different algorithms in number of FLOPs, for $N = 128$ and $K = 8$.

5.5.2 Complexity analysis of MMPA

The PA problem in (5.20) belongs to the class of linear programming (LP), in which the objective function and the constraints are linear in the optimization variables. The complexity of achieving a solution of LP within an accuracy ϵ has been studied in [133]. The complexity of (5.20) depends on the number of optimization variables, number of per scalar value constraints, and the size of input data. Note that the solvers which are based on interior-point algorithm — such those employed by CVX — do not support entropy functions such as logarithmic and exponential functions. Therefore, the constraints in (5.19a), which include exponential functions, are linearised as $\sum_{i=1}^{K} p_i|\mathbf{h}_k^T\mathbf{w}_i|^2 + \sigma_n^2 \geq e^{\bar{u}_k}(u_k - \bar{u}_k + 1),\ \forall\, k$, where $\bar{u}_k$ is the point around which the linearization is made. To apply the complexity analysis steps provided in [133], we have to recast (5.20) into its standard LP form. First, the $\min$ operator in the objective function $\min_k u_k - s_k$ is replaced by a new slack variable ν along with adding the constraints $(u_k - s_k) \geq \nu,\ \forall\, k$. The per-iteration standard LP of (5.20) will become

$$\underset{\{p_1,\ldots,p_K\},\, \boldsymbol{u},\, \boldsymbol{s},\, \nu}{\text{maximize}} \quad \nu$$

subject to

$$u_k - s_k \geq \nu,\ \forall\, k, \tag{5.25a}$$

$$\sum_{i=1}^{K} p_i|\mathbf{h}_k\mathbf{w}_i|^2 + \sigma_n^2 \geq e^{\bar{u}_k}(u_k - \bar{u}_k + 1),\ \forall\, k, \tag{5.25b}$$

$$\text{In addition to constraints } (5.20a),\ (5.19c),\ (5.19d). \tag{5.25c}$$

The problem in (5.25) has the following parameters: number of optimization variables, $n_v = 3K + 1$, number of per scalar value constraints, $n_c = 4K + 1$, number of input data $\dim(\boldsymbol{d}) = 2K + 5$, where $\boldsymbol{d} = [n_v,\ n_c,\ [\bar{u}_1, \bar{s}_1;\ \ldots;\ \bar{u}_K, \bar{s}_K],\ \sigma^2,\ P_t,\ \zeta]$ is the vector of input data. Given these parameters, the per-iteration computational complexity of achieving a solution of (25) within an accuracy ϵ is equal to [133]

$$\mathcal{C}_{(\boldsymbol{d},\epsilon)} = (n_v + n_c)^{\frac{3}{2}}\, n_v^2 \ln\left(\frac{\dim(\boldsymbol{d}) + \|\boldsymbol{d}\|_1 + \epsilon^2}{\epsilon}\right), \tag{5.26}$$

where $\|\boldsymbol{a}\|_1$ is the 1-norm of $\boldsymbol{a}$ and can be defined as $\|[a_1, ..., a_L]\|_1 = \sum_{l=1}^{L} |a_l|$. The result in (5.26) has an asymptotic complexity of $\mathcal{O}(K^{\frac{7}{2}}[\ln(K) + \ln(\frac{1}{\epsilon})])$.

5.6 Results and Discussions

In this section, we show the simulation results for different AS algorithms under EPA and MMPA schemes. Before discussing the results, we introduce the different parameters used throughout our work. We start by defining the SNR for user k as follow

$$SNR_k = \frac{\sigma_{h_k}^2}{\sigma_n^2}\frac{P_T}{K}. \tag{5.27}$$

To find the average SNR, we need to find the maximum and minimum variance of the channel, $\sigma_{h_{max}}^2$ and $\sigma_{h_{min}}^2$. We apply the same values for the path loss component and the fix-loss component in Chapter 4, and assume that the minimum and maximum distances between any user and the BS, $\mathrm{d_{min}}$ and $\mathrm{d_{max}}$, were equal to 25 and 250 meters, respectively. Therefore, the maximum channel variance $\sigma_{h_{max}}^2$ is 1.6×10^{-4}, which corresponds to a user located at $\mathrm{d_{min}}$ meters away from the BS, hence the maximum SNR can be given as

$$SNR_{max} = \frac{\sigma_{h_{max}}^2}{\sigma_n^2}\frac{P_T}{K}, \tag{5.28}$$

while the worst-case user, who is located at $\mathrm{d_{max}}$ meters away from the BS, has a channel variance of $\sigma_{h_{min}}^2 = 1.6 \times 10^{-6}$, hence the minimum SNR can be expressed as

$$SNR_{min} = \frac{\sigma_{h_{min}}^2}{\sigma_n^2}\frac{P_T}{K}. \tag{5.29}$$

Therefore, the average SNR can be expressed as

$$SNR_{average} = 0.5\left(SNR_{max} + SNR_{min}\right), \tag{5.30}$$

while the noise variance σ_n^2 was assumed to be 1.6×10^{-5}. Furthermore, we assume throughout our simulations that the number of antenna elements $N = 128$, number of selected antennas $N_s = 16$, and the number of users $K = 5$. Finally, the value of ζ was set to 0.75, which indicates that each user should be allocated at least 75% of the average power.

5.6.1 Achievable rates with EPA

Fig. 5.6 shows the total sum rate of the EPA strategy for different AS schemes. The SIM achieves a superior performance compared to all other alogrithms, especially at high SNRs,

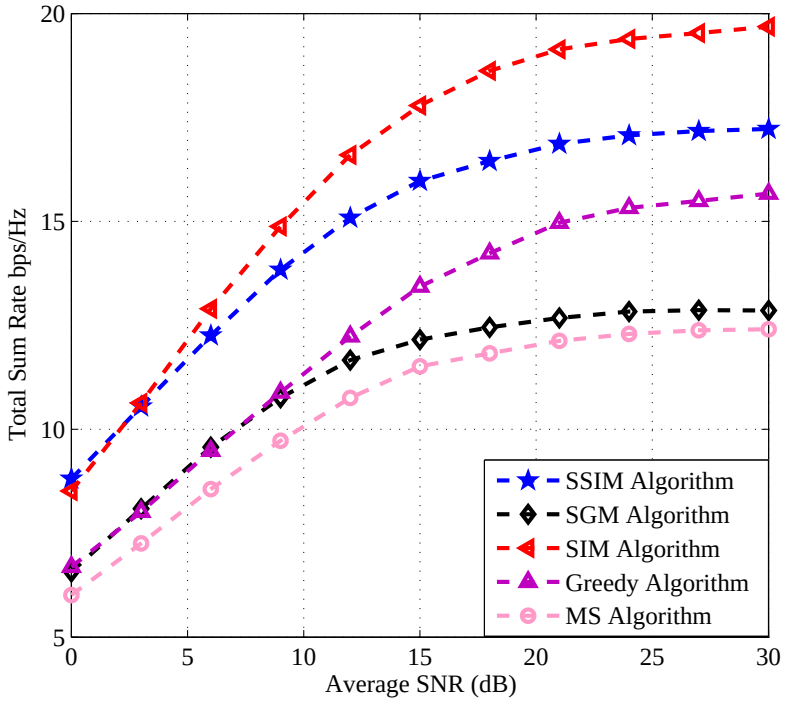


Figure 5.6: Achievable sum rate vs. Average SNR under EPA of different algorithms, for $N = 128$, $N_s = 16$, and $K = 5$.

followed by the SSIM, and then the greedy algorithm. In contrast, the MS algorithm has the worst performance among the applied AS techniques for all SNR values investigated.

Since the users do not have equal distances from the BS, it is important to measure the achievable rate of the worst-case user. Interestingly, Fig. 5.7 demonstrates the importance of SGM algorithm, since it achieves the highest worst-case rate when the average SNR is less than 12 dB. Furthermore, the SIM algorithm shows the best performance at high SNRs followed by both SSIM and greedy selection algorithms. However, the SSIM exhibits a higher total sum rate than the greedy algorithm as shown in Fig. 5.6. In terms of total achievable sum rate, the SIM and SSIM algorithms significantly outperform the greedy algorithm by 4.36 and 2.54 bps/Hz at an average SNR of 15 dB, respectively. While the same algorithms outperform the MS selection by 6.28 and 4.46 bps/Hz for the same SNR value. On the other hand, considering the worst-case user only, the SGM algorithm outperforms all the other selection techniques applied in this work by approximately 0.33 bps/Hz, which corresponds to 6.6 Mbps for a system with 20 MHz bandwidth at 0 dB of average SNR.

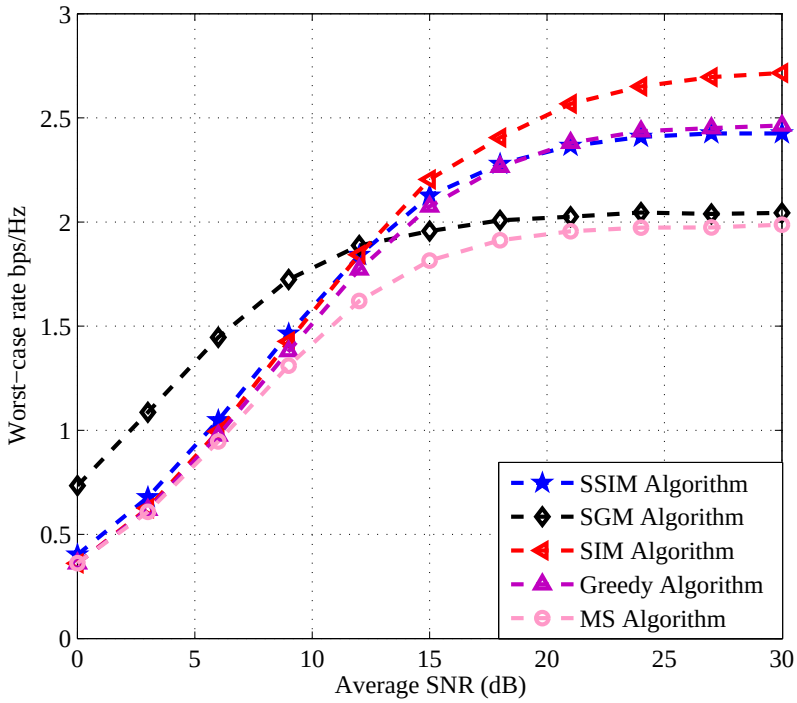


Figure 5.7: Worst-case rate vs. Average SNR under EPA of different algorithms, for $N = 128$, $N_s = 16$, and $K = 5$.

5.6.2 Achievable rates with MMPA

In this section, we show the achievable rates for the proposed algorithms after carrying MMPA. The idea of this approach is to enhance the rates of users who are located far away from the BS, by allocating them more power than users who are closer to the BS. However, this will result in a small degradation in the total sum rate. Fig. 5.8 shows the achievable rate for the worst-case user with EPA and MMPA. At low SNRs, the SGM algorithm achieves the highest worst-case rate before and after applying MMPA, while at high SNRs, SIM algorithm achieves the highest rate with EPA and MMPA, followed by SSIM algorithm. Fig. 5.9 shows the total sum rate before and after applying MMPA. The SIM outperforms both SSIM and SGM, especially at high SNRs. While the SGM has the lowest total sum rate between the proposed algorithms for all SNR values.

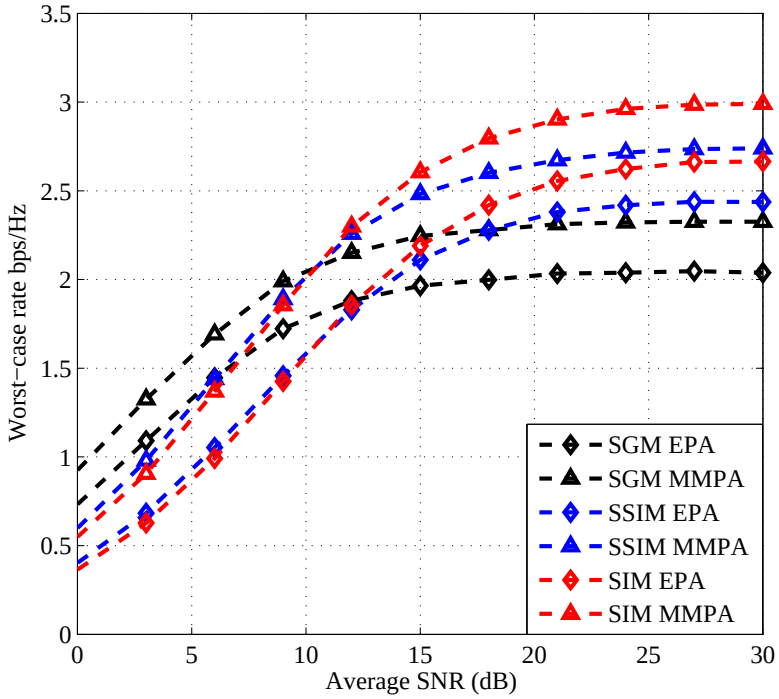


Figure 5.8: Worst-case rate vs. Average SNR under EPA and MMPA of different algorithms, for $N = 128$, $N_s = 16$, and $K = 5$.

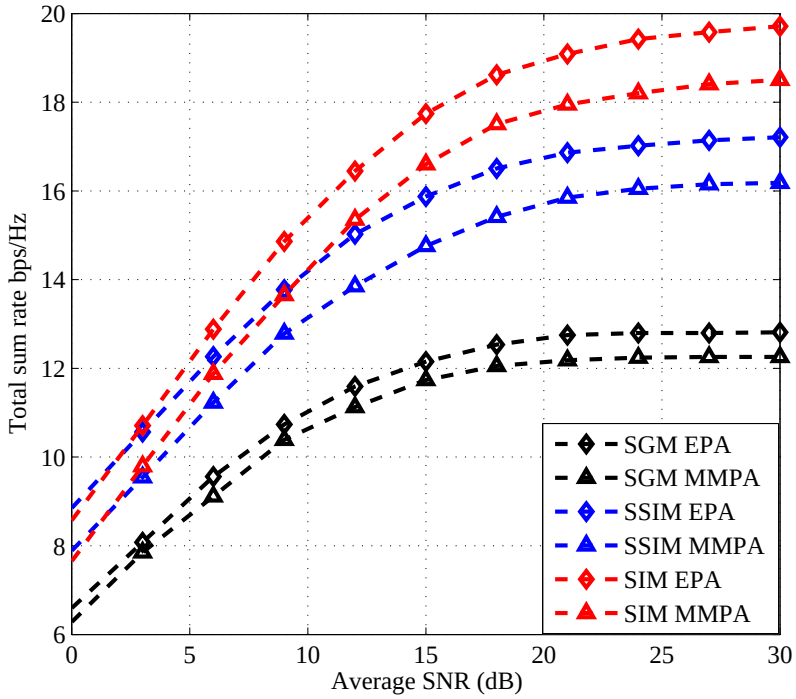


Figure 5.9: Achievable sum rate vs. Average SNR under EPA and MMPA of different algorithms, for $N = 128$, $N_s = 16$, and $K = 5$.

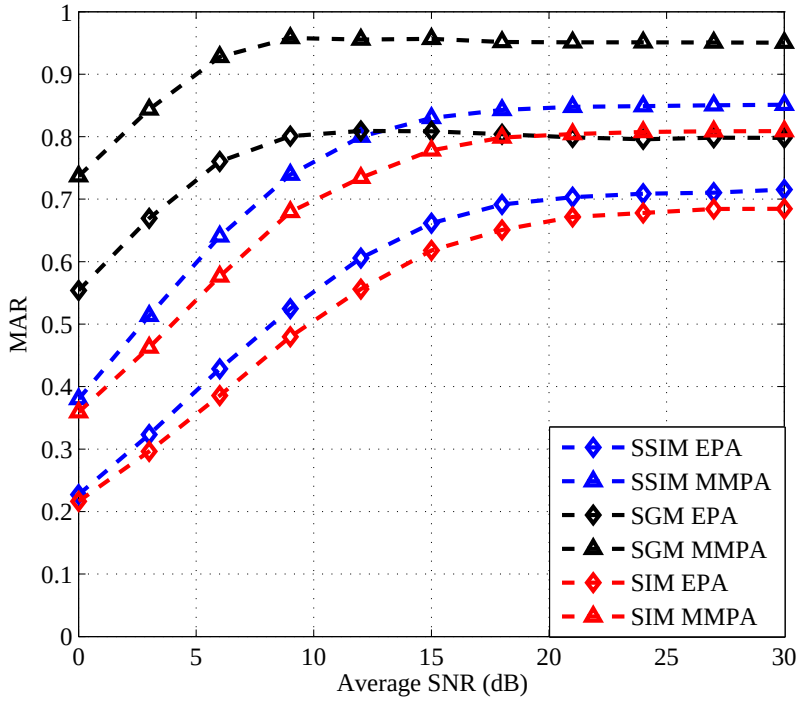


Figure 5.10: MAR ratio vs Average SNR for the three proposed algorithms with EPA and MMPA.

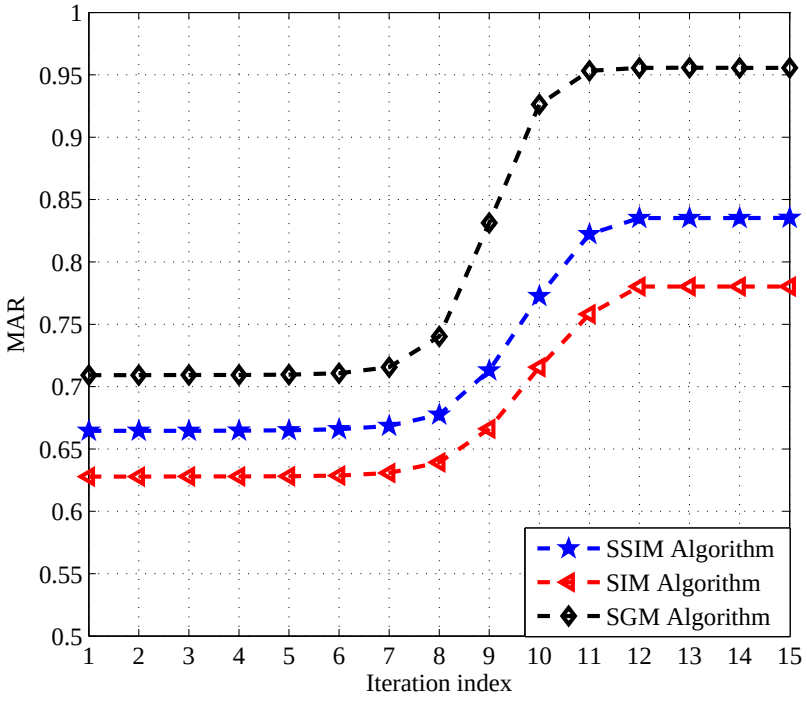


Figure 5.11: Convergence speed vs MAR ratio for $K = 5$, and average SNR of 15 dB.

As shown from Figs. 5.8 and 5.9, applying MMPA increased the worst-case user rate by 0.41, 0.37, and 0.22 bps/Hz for the SIM, SSIM, and SGM algorithms at an average SNR of 15 dB, respectively. However, that comes at the cost of decreasing the total achievable rate by 1.14, 1.11, and 0.41 bps/Hz for the SIM, SSIM, and SGM algorithms at the same SNR value, respectively.

Moreover, Fig. 5.10 shows the Minimum to Average Rate (MAR) ratio, which indicates the level of fairness achieved among the users. It is clear that applying MMPA guarantees higher fairness between the users for all algorithms. Furthermore, SGM algorithm achieves a considerably higher equality among users compared to SSIM and SIM algorithms, especially at low SNRs, which confirms its advantage over the other two schemes for systems with low power budgets. However, at high SNRs, the three methods show high fairness between the users. Furthermore, Fig. 5.11 shows the number of iterations required for the iterative MMPA algorithm to achieve convergence in terms of MAR ratio, for the three proposed algorithms when the average SNR is 15 dB. The three algorithms require 12 iterations to converge with $\tau = 10^{-3}$ bps/Hz of error tolerance.

Figs. 5.12 and 5.13 demonstrate the impact of ζ on the fairness level among the users and the total sum rate, respectively. As shown in Fig. 5.12, SGM method achieves higher fairness regardless of the value of ζ, which explains the small degradation in the total sum rate for the SGM method in Fig. 5.13. In contrast, the SIM method suffers the highest performance degradation in terms of total sum rate, since it has the lowest MAR among the proposed algorithms.

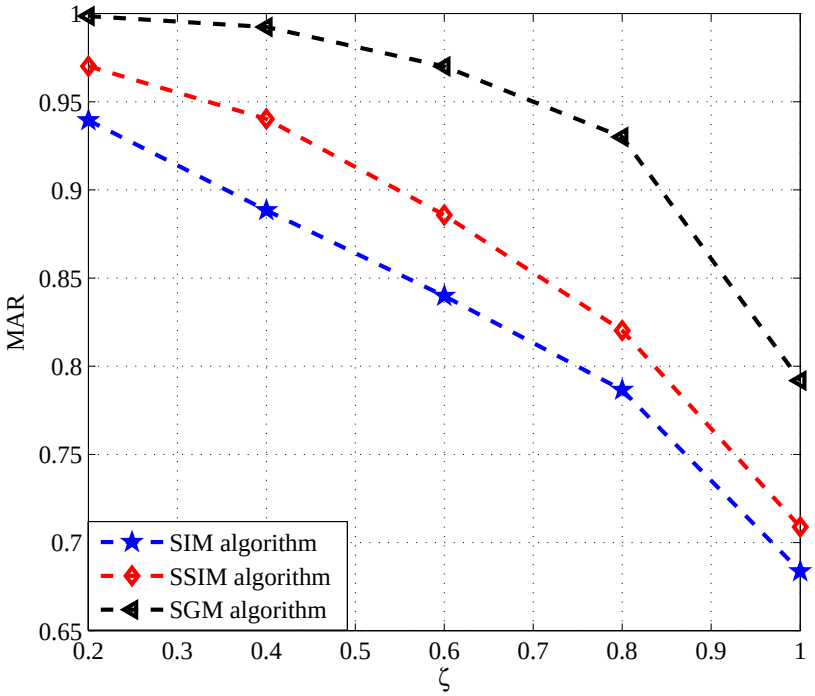


Figure 5.12: ζ vs MAR for $K = 5$, $N = 128$, $N_s = 16$, and SNR = 30 dB.

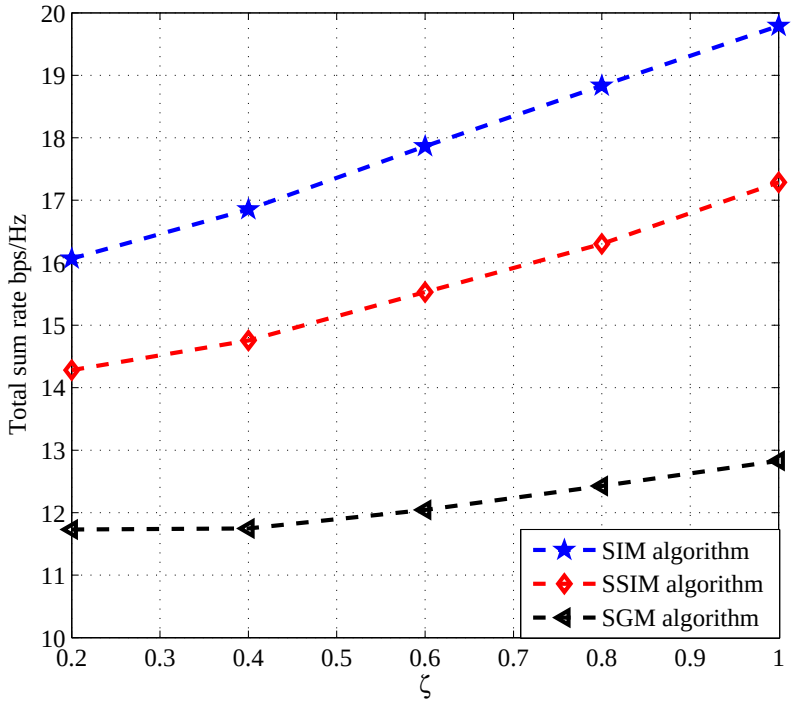


Figure 5.13: ζ vs total sum rate for $K = 5$, $N = 128$, $N_s = 16$, and SNR = 30 dB.

5.7 Chapter Summary

In this chapter, three low complexity AS algorithms were proposed for distributed MU M-MIMO downlink systems when MF precoding is applied. The first algorithm aimed to maximize the channel gain for the worst-case user, and it showed to have the least complexity and the highest worst-case rate at low SNRs among the proposed methods. While the second algorithm aimed to minimize the interference for worst-case user, and it showed to achieve the highest sum rates compared to the other two algorithms, however, that came at the cost of increased complexity. The final algorithm aimed to reduce the highest interference term between any two users, and it showed good trade-off between complexity and performance compared to the other two proposed methods. Furthermore, MMPA was applied to achieve higher fairness among the users, however, that came at the cost of small degradation in the total sum rates. Finally, the complexity of the proposed AS algorithms as well as the PA problem was evaluated in terms of number of FLOPs required for their implementations.

Chapter 6

Conclusions and Future Work

In order to satisfy the ever increasing demand for higher spectral efficiency in wireless communications systems, employing large antenna arrays at the BS seems to be an inevitable approach. However, using large antenna arrays comes at huge price, in terms of cost, complexity, and power consumption. To increase the achievable rates of a conventional MIMO system while maintaining an affordable cost and complexity requirements, AS schemes can be utilized. This thesis addressed the AS schemes in MU M-MIMO systems, in both uplink and downlink scenarios. Moreover, several low complexity, yet highly efficient, AS schemes were proposed for SINR maximization in MU M-MIMO systems.

6.1 Summary and conclusions of the thesis

In this thesis, an overview of MIMO systems was presented in Chapter 2, were linear detection and precoding techniques were first introduced. Moreover, two well-known benchmark AS schemes for conventional MIMO systems were presented. Then M-MIMO systems were introduced and the key rule of AS in systems with massive arrays was explained. After highlighting the main limitations of the existing AS schemes for M-MIMO systems, we present our main research contribution in this thesis.

- In Chapter 3, an uplink MU M-MIMO system was considered, and several evolutionary algorithms were applied to maximize the capacity for interference-free scenarios. The key contributions of this chapter can be summarized as follow

  C3.1 In terms of the achievable rates, the QTS algorithm demonstrated the best performance among the employed evolutionary algorithms for AS in M-MIMO systems. Followed by both CTS and GA. In contrast, the PSO method showed the worst

capacity performance followed by the ABC algorithm. Furthermore, our results demonstrated that the performance of both ABC and PSO methods becomes very undesirable as the number of iterations increases, since both methods have the worst convergence behaviour compared to the other employed evolutionary methods.

C3.2 In terms of complexity requirement, QTS and CTS require much lower complexity than any of the bio-inspired algorithms. Moreover, the ABC method suffers from the highest complexity, followed by PSO, and then GA.

C3.3 In addition to its low complexity requirement and high performance, QTS method had an advantage over all other evolutionary methods in terms of optimizing its parameters, since it only required finding the optimal rotation angle, which did not depend on the dimensions of the system. In contrast, CTS, ABC, GA, and PSO each have several parameters that needs to be optimized in order to obtain the best performance.

- In Chapter 4, a single-cell downlink MU M-MIMO system was consider with MF precoding. Two novel AS algorithms were proposed for SINR maximization with reduced complexity. Our main contributions in this chapter were:

  C4.1 A reduced search space UCAS algorithm was proposed, where each user was assigned a subset of antennas with high channel norms. Therefore, maximizing the SINR for a certain user was achieved by selecting an antenna from the group which was assigned to that user only, this resulted in dramatic complexity reduction while maintaining high performance.

  C4.2 A SIRAS algorithm was proposed, where at each iteration one antenna was selected to reject the highest number of interference terms, without taking into account the path gain of the antennas.

  C4.3 Our results demonstrated that both methods significantly outperformed other low complexity AS methods from the literature. In addition, the proposed methods demonstrated higher sum rates and EE than the case where all the antenna elements were activated. Furthermore, the SIRAS method outperformed the UCAS algorithm as the number of users increased.

  C4.4 A complexity analysis was carried out for the proposed methods in terms of number of floating-point operations. Our analysis demonstrated that the UCAS method requires less complexity than SIRAS algorithm. However, both methods showed a

significant performance-complexity trade-off compared to other low complexity AS methods.

- In Chapter 5, different AS methods were proposed and utilized, alongside max-min PA control, to enhance the performance of cell-edge users. The main contribution work in this chapter can be summarized in the following

  C5.1 Three low complexity AS methods were proposed to enhance the performance of the worst-case user at each iteration. The first algorithm aimed to maximize the channel gain, while the second method aimed to minimize the interuser interference. The last method focused on minimizing the highest interference term only.

  C5.2 Our results demonstrated that maximizing the channel gain is better at low SNRs, since it achieves a considerably higher worst-case rate compared to the other proposed methods. In contrast, at high SNRs, minimizing the interference outperform significantly the SGM method.

  C5.3 MMPA was carried out to maximize the worst-case user rate under total power constraint, and increase the fairness between the users.

  C5.4 The complexity was analysed in terms of number of FLOPs for the different AS schemes as well as the MMPA control. For the AS schemes, the SGM method had the lowest complexity requirement among the proposed methods, while the SIM suffered from the highest complexity. In addition, and for the PA scheme, it was found that the complexity is a function of the number of users, and it does not depend on the number of antennas at the BS.

6.2 Future Work

This thesis addressed the problem of designing low complexity AS schemes for MU M-MIMO systems. The work in this thesis can be further investigated in the following directions:

1. In Chapter 3, both of CTS and GA showed close performance to the QTS algorithm with good convergence behaviour. However, different parameters were used in both methods such as the mutation probability, number of paired chromosomes, random repairing procedure, length of the tabu matrix, and number of tabu antennas, ... etc. Optimizing these parameters can further improve the achievable rates of these methods. Furthermore, the

evolutionary methods employed can also be investigated as a low complexity detection in M-MIMO systems.

2. The AS methods proposed in Chapter 4, can be investigated for a Cell-Free M-MIMO systems where the access points are distributed over large area, for both antenna and access point selection.

3. In Chapter 5, the proposed AS methods can be improved by taking into account the precoding normalization factors for each user during the iterative selection process, since they depend on the large scale fading of each user. Therefore, ignoring these factors can lead to the selection of suboptimal antennas. However, this will result in an increased complexity. Furthermore, for systems with users located at the cell-edge, user scheduling can be applied to achieve higher rates.

4. The same methodology used for PA in Chapter 5, can be utilized for different types of precoding design, such as to maximize the SINR or the Signal to Leakage plus Noise Ratio (SLNR). Furthermore, the joint AS and precoding design can be an interesting approach for future work.

5. The work in Chapter 5 can be further extended for multi-cell M-MIMO scenario, where both AS and PA methods can take into account inter-cell interference in addition to the co-channel interference.

6. The work on AS can be further extended to consider a time-correlated channels, where for different channel realizations, only a subset of the available antennas will go through the selection process, which can result in dramatic complexity reduction.

7. The spectral efficiency of M-MIMO systems can further be improved by employing the generalized spatial modulation (GSM), where the indices of the activated antennas bear information. Therefore, combining AS with GSM can target both high energy and improved spectral efficiencies at the same time. However, detecting the antenna indices at the receiver side will impose higher complexity. Accordingly, reduced complexity detection method for joint AS and GSM is a very attractive direction for future research.

References

[1] S. Sesia, M. Baker, and I. Toufik, *LTE-the UMTS long term evolution: from theory to practice*. John Wiley & Sons, 2011.

[2] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7112–7139, 2014.

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

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

[5] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor, G. Auer, and L. Van Der Perre, "Challenges and enabling technologies for energy aware mobile radio networks," *IEEE Communications Magazine*, vol. 48, no. 11, 2010.

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

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

[8] M. Gharavi-Alkhansari and A. B. Gershman, "Fast antenna subset selection in MIMO systems," *IEEE transactions on signal processing*, vol. 52, no. 2, pp. 339–347, 2004.

[9] A. F. Molisch, M. Z. Win, Y.-S. Choi, and J. H. Winters, "Capacity of MIMO systems with antenna selection," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1759–1772, 2005.

[10] S. Sanayei and A. Nosratinia, "Capacity of MIMO channels with antenna selection," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4356–4362, 2007.

[11] A. Gorokhov, D. Gore, and A. Paulraj, "Receive antenna selection for MIMO flat-fading channels: theory and algorithms," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2687–2696, 2003.

[12] A. Gorokhov, D. A. Gore, and A. J. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: theory and algorithms," *IEEE Transactions on signal processing*, vol. 51, no. 11, pp. 2796–2807, 2003.

[13] B. H. Wang, H. T. Hui, and M. S. Leong, "Global and fast receiver antenna selection for MIMO systems," *IEEE Transactions on Communications*, vol. 58, no. 9, pp. 2505–2510, 2010.

[14] K. T. Phan and C. Tellambura, "Receive antenna selection based on union-bound minimization using convex optimization," *IEEE Signal Processing Letters*, vol. 14, no. 9, pp. 609–612, 2007.

[15] M. Ju, H.-K. Song, and I.-M. Kim, "Joint relay-and-antenna selection in multi-antenna relay networks," *IEEE Transactions on Communications*, vol. 58, no. 12, pp. 3417–3422, 2010.

[16] S. Kim, M. Shin, and C. Lee, "Transmit antenna selection scheme for iterative receivers in MIMO systems," *IEEE Signal Processing Letters*, vol. 14, no. 12, pp. 916–919, 2007.

[17] S. Padmanabhan, R. G. Stephen, C. R. Murthy, and M. Coupechoux, "Training-based antenna selection for PER minimization: A POMDP approach," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3247–3260, 2015.

[18] P. Sinchu, R. G. Stephen, C. R. Murthy, and M. Coupechoux, "A POMDP solution to antenna selection for PER minimization," in *Global Communications Conference (GLOBECOM), 2014 IEEE*. IEEE, 2014, pp. 3850–3855.

[19] A. Dua, K. Medepalli, and A. J. Paulraj, "Receive antenna selection in MIMO systems using convex optimization," *IEEE Transactions on Wireless Communications*, vol. 5, no. 9, pp. 2353–2357, 2006.

[20] M. Sadek, A. Tarighat, and A. H. Sayed, "Active antenna selection in multiuser MIMO communications," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1498–1510, 2007.

[21] H. Li, L. Song, D. Zhu, and M. Lei, "Energy efficiency of large scale MIMO systems with transmit antenna selection," in *Communications (ICC), 2013 IEEE International Conference on*. IEEE, 2013, pp. 4641–4645.

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

[23] A. Bereyhi, S. Asaad, and R. R. Mueller, "Stepwise transmit antenna selection in downlink massive multiuser MIMO," in *WSA 2018; 22nd International ITG Workshop on Smart Antennas*. VDE, 2018, pp. 1–8.

[24] D. Mi, M. Dianati, S. Muhaidat, and Y. Chen, "A novel antenna selection scheme for spatially correlated massive MIMO uplinks with imperfect channel estimation," in *Vehicular Technology Conference (VTC Spring), 2015 IEEE 81st*. IEEE, 2015, pp. 1–6.

[25] H. Tang and Z. Nie, "Rmv antenna selection algorithm for massive MIMO," *IEEE Signal Processing Letters*, vol. 25, no. 2, pp. 239–242, 2018.

[26] 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, 2018.

[27] 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.

[28] X. Gao, O. Edfors, J. Liu, and F. Tufvesson, "Antenna selection in measured massive MIMO channels using convex optimization," in *Globecom Workshops (GC Wkshps), 2013 IEEE*. IEEE, 2013, pp. 129–134.

[29] M. Gkizeli and G. N. Karystinos, "Maximum-SNR antenna selection among a large number of transmit antennas," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 891–901, 2014.

[30] ——, "Maximum-SNR transmit antenna selection with two receive antennas is polynomially solvable," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 4749–4753.

[31] M. Benmimoune, E. Driouch, W. Ajib, and D. Massicotte, "Joint transmit antenna selection and user scheduling for massive MIMO systems," in *Wireless Communications and Networking Conference (WCNC), 2015 IEEE*. IEEE, 2015, pp. 381–386.

[32] X. Liu and X. Wang, "Efficient antenna selection and user scheduling in 5G massive MIMO-NOMA system," in *Vehicular Technology Conference (VTC Spring), 2016 IEEE 83rd*. IEEE, 2016, pp. 1–5.

[33] Z. Liu, W. Du, and D. Sun, "Energy and spectral efficiency tradeoff for massive MIMO systems with transmit antenna selection," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 4453–4457, 2017.

[34] P. V. Amadori and C. Masouros, "Interference-driven antenna selection for massive multiuser MIMO," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 5944–5958, 2016.

[35] ——, "Power efficient massive MU-MIMO via antenna selection for constructive interference optimization," in *Communications (ICC), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1607–1612.

[36] ——, "Large scale antenna selection and precoding for interference exploitation," *IEEE Transactions on Communications*, vol. 65, no. 10, pp. 4529–4542, 2017.

[37] ——, "A mixed-integer programming approach to interference exploitation for massive-MIMO," in *Wireless Communications and Networking Conference Workshops (WCNCW), 2018 IEEE*. IEEE, 2018, pp. 107–112.

[38] M. Olyaee, M. Eslami, and J. Haghighat, "An energy-efficient joint antenna and user selection algorithm for multi-user massive MIMO downlink," *IET Communications*, vol. 12, no. 3, pp. 255–260, 2017.

[39] J. Joung and S. Sun, "Two-step transmit antenna selection algorithms for massive MIMO," in *Communications (ICC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–6.

[40] 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 *Globecom Workshops (GC Wkshps), 2015 IEEE*. IEEE, 2015, pp. 1–5.

[41] S. Qin, G. Li, G. Lv, G. Zhang, and H. Hui, "L1/2-regularization based antenna selection for RF-chain limited massive MIMO systems," in *Vehicular Technology Conference (VTC-Fall), 2016 IEEE 84th*. IEEE, 2016, pp. 1–5.

[42] Y. Gao, W. Jiang, and T. Kaiser, "Bidirectional branch and bound based antenna selection in massive MIMO systems," in *Personal, Indoor, and Mobile Radio Communications (PIMRC), 2015 IEEE 26th Annual International Symposium on*. IEEE, 2015, pp. 563–568.

[43] M. Xie and T.-M. Lok, "Antenna selection in RF-chain-limited MIMO interference networks under interference alignment," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 3856–3870, 2017.

[44] R. Hamdi, E. Driouch, and W. Ajib, "Resource allocation in downlink large-scale MIMO systems," *IEEE Access*, vol. 4, pp. 8303–8316, 2016.

[45] H. Q. Ngo, M. Matthaiou, and E. G. Larsson, "Massive MIMO with optimal power and training duration allocation," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 605–608, 2014.

[46] T. M. Nguyen, V. N. Ha, and L. B. Le, "Resource allocation optimization in multi-user multi-cell massive MIMO networks considering pilot contamination," *IEEE Access*, vol. 3, pp. 1272–1287, 2015.

[47] Q. Zhang, S. Jin, M. McKay, D. Morales-Jimenez, and H. Zhu, "Power allocation schemes for multicell massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 5941–5955, 2015.

[48] T. Van Chien, E. Björnson, and E. G. Larsson, "Joint power allocation and user association optimization for massive MIMO systems." *IEEE Trans. Wireless Communications*, vol. 15, no. 9, pp. 6384–6399, 2016.

[49] P. Liu, S. Jin, T. Jiang, Q. Zhang, and M. Matthaiou, "Pilot power allocation through user grouping in multi-cell massive MIMO systems," *IEEE Transactions on Communications*, vol. 65, no. 4, pp. 1561–1574, 2017.

[50] L. Zhao, H. Zhao, F. Hu, K. Zheng, and J. Zhang, "Energy efficient power allocation algorithm for downlink massive MIMO with MRT precoding," in *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th*. IEEE, 2013, pp. 1–5.

[51] O. Saatlou, M. O. Ahmad, and M. Swamy, "Spectral efficiency maximization of single cell massive multiuser MIMO systems via optimal power control with ZF receiver," in *Personal, Indoor, and Mobile Radio Communications (PIMRC), 2017 IEEE 28th Annual International Symposium on*. IEEE, 2017, pp. 1–5.

[52] T. Van Chien, E. Björnson, and E. G. Larsson, "Joint pilot design and uplink power allocation in multi-cell massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2000–2015, 2018.

[53] K. Guo, Y. Guo, G. Ascheid *et al.*, "Uplink power control with MMSE receiver in multi-cell MU-massive-MIMO systems," in *Communications (ICC), 2014 IEEE International Conference on*. IEEE, 2014, pp. 5184–5190.

[54] Y. Zhang, H. Gao, F. Tan, and T. Lv, "Resource allocation of energy-efficient multi-user massive MIMO systems," in *Globecom Workshops (GC Wkshps), 2016 IEEE*. IEEE, 2016, pp. 1–6.

[55] Y. Liu, Q. Chen *et al.*, "On the energy efficient multi-pair two-way massive MIMO AF relaying with imperfect CSI and optimal power allocation," *IEEE Access*, vol. 6, pp. 2589–2603, 2018.

[56] H. Gao, T. Lv, X. Su, H. Yang, and J. M. Cioffi, "Energy-efficient resource allocation for massive MIMO amplify-and-forward relay systems." *IEEE Access*, vol. 4, pp. 2771–2787, 2016.

[57] F. Tan, T. Lv, and S. Yang, "Power allocation optimization for energy-efficient massive MIMO aided multi-pair decode-and-forward relay systems," *IEEE Transactions on Communications*, vol. 65, no. 6, pp. 2368–2381, 2017.

[58] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[59] J. G. Proakis and M. Salehi, *Digital communications*. McGraw-Hill, 2008.

[60] A. Chockalingam and B. S. Rajan, *Large MIMO systems*. Cambridge University Press, 2014.

[61] L. Bai and J. Choi, *Low complexity MIMO detection*. Springer Science & Business Media, 2012.

[62] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM wireless communications with MATLAB*. John Wiley & Sons, 2010.

[63] N. Jindal and A. Goldsmith, "Dirty-paper coding versus TDMA for MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 51, no. 5, pp. 1783–1794, 2005.

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

[65] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian MIMO broadcast channel," in *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*. IEEE, 2004, p. 174.

[66] M. Costa, "Writing on dirty paper (corresp.)," *IEEE transactions on information theory*, vol. 29, no. 3, pp. 439–441, 1983.

[67] 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, 2014.

[68] T. X. Tran and K. C. Teh, "Spectral and energy efficiency analysis for SLNR precoding in massive MIMO systems with imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 4017–4027, 2018.

[69] M. Sadek, A. Tarighat, and A. H. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE transactions on Wireless Communications*, vol. 6, no. 5, 2007.

[70] T. M. Duman and A. Ghrayeb, *Coding for MIMO communication systems*. John Wiley & Sons, 2008.

[71] A. Gershman and N. Sidiropoulos, *Space-time processing for MIMO communications*. John Wiley & Sons, 2005.

[72] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proceedings of the IEEE*, vol. 102, no. 1, pp. 56–103, 2014.

[73] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: A survey," *IEEE Communications Magazine*, vol. 49, no. 12, 2011.

[74] R. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, p. 2228, 2008.

[75] I. A. Hemadeh, M. El-Hajjar, and L. Hanzo, "Hierarchical multi-functional layered spatial modulation," *IEEE Access*, vol. 6, pp. 9492–9533, 2018.

[76] T. Kailath, *Linear systems*. Prentice-Hall Englewood Cliffs, NJ, 1980, vol. 156.

[77] 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, 2014.

[78] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of massive MIMO*. Cambridge University Press, 2016.

[79] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," *IEEE J. of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, 2014.

[80] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and H. Yang, "Cell-free massive MIMO systems," in *Signals, Systems and Computers, 2015 49th Asilomar Conference on*. IEEE, 2015, pp. 695–699.

[81] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.

[82] ——, "Cell-free massive MIMO: Uniformly great service for everyone," in *Signal Processing Advances in Wireless Communications (SPAWC), 2015 IEEE 16th International Workshop on*. IEEE, 2015, pp. 201–205.

[83] L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki, "Energy efficiency in cell-free massive MIMO with zero-forcing precoding design," *IEEE Communications Letters*, vol. 21, no. 8, pp. 1871–1874, 2017.

[84] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 1, pp. 25–39, 2018.

[85] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—the large-scale array regime," *IEEE transactions on information theory*, vol. 59, no. 10, pp. 6441–6463, 2013.

[86] M. Mitchell, *An introduction to genetic algorithms*. MIT press, 1998.

[87] J. Kennedy, J. F. Kennedy, R. C. Eberhart, and Y. Shi, *Swarm intelligence*. Morgan Kaufmann, 2001.

[88] P. D. Karamalis, N. D. Skentos, and A. G. Kanatas, "Selecting array configurations for MIMO systems: An evolutionary computation approach," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1994–1998, 2004.

[89] B. Makki, A. Ide, T. Svensson, T. Eriksson, and M.-S. Alouini, "A genetic algorithm-based antenna selection approach for large-but-finite MIMO networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 6591–6595, 2017.

[90] J.-K. Lain, "Joint transmit/receive antenna selection for MIMO systems: a real-valued genetic approach," *IEEE Communications Letters*, vol. 15, no. 1, pp. 58–60, 2011.

[91] J. Lu, X. Li, and D. Liu, "Asynchronous transmission of wireless multicast system with genetic joint antennas selection," in *Wireless Communications and Networking Conference Workshops (WCNCW), 2013 IEEE*. IEEE, 2013, pp. 66–70.

[92] Y. Wang and Y. Dong, "A genetic antenna selection algorithm for massive MIMO systems with channel estimation error," in *Advances in Wireless and Optical Communications (RTUWO), 2015*. IEEE, 2015, pp. 1–4.

[93] L. Du, L. Li, and Y. Xu, "A genetic antenna selection algorithm with heuristic beamforming for massive MIMO systems," in *Wireless Personal Multimedia Communications (WPMC), 2016 19th International Symposium on*. IEEE, 2016, pp. 49–52.

[94] W.-H. Fang, S.-C. Huang, and Y.-T. Chen, "Genetic algorithm-assisted joint quantised precoding and transmit antenna selection in multi-user multi-input multi-output systems," *IET communications*, vol. 5, no. 9, pp. 1220–1229, 2011.

[95] K. Obaidullah, C. Siriteanu, S. Yoshizawa, and Y. Miyanaga, "Evaluation of genetic algorithm-based detection for correlated MIMO fading channels," in *Communications and Information Technologies (ISCIT), 2011 11th International Symposium on*. IEEE, 2011, pp. 507–511.

[96] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc of IEEE International Conference on Neural Network. Piscataway: IEEE Press*. sn, 1995, pp. 1942–1948.

[97] D. J. Krusienski and W. K. Jenkins, "A modified particle swarm optimization algorithm for adaptive filtering," in *Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on*. IEEE, 2006, pp. 4–pp.

[98] Y.-Q. Hei, X.-H. Li, K.-C. Yi, and X. Li, "Multiuser scheduling in downlink MIMO systems using particle swarm optimization," in *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE*. IEEE, 2009, pp. 1–5.

[99] H.-Y. Lu and C.-H. Huang, "A two-stage particle swarm optimization based cooperative transmit beamforming for multiuser MIMO systems," in *Communication Technology (ICCT), 2010 12th IEEE International Conference on*. IEEE, 2010, pp. 555–558.

[100] J. Dong, Y. Xie, Y. Jiang, F. Liu, R. Shi, and D. Xiong, "Particle swarm optimization for joint transmit and receive antenna selection in MIMO systems," in *Communication Problem-Solving (ICCP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 237–240.

[101] M. Sinaie, M. S. Tohidi, and P. Azmi, "Particle swarm optimization for energy efficient antenna selection in MIMO broadcasting channel," in *Electrical Engineering (ICEE), 2013 21st Iranian Conference on*. IEEE, 2013, pp. 1–4.

[102] J. Dai and M. Chen, "Improved particle swarm optimization-based antenna selection scheme for correlated MIMO channels," in *Communication Technology (ICCT), 2010 12th IEEE International Conference on*. IEEE, 2010, pp. 512–515.

[103] H. Yongqiang, L. Wentao, and L. Xiaohui, "Particle swarm optimization for antenna selection in MIMO system," *Wireless personal communications*, vol. 68, no. 3, pp. 1013–1029, 2013.

[104] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Technical report-tr06, Erciyes university, engineering faculty, computer engineering department, Tech. Rep., 2005.

[105] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm," *Journal of global optimization*, vol. 39, no. 3, pp. 459–471, 2007.

[106] N. Taşpınar and M. Yıldırım, "A novel parallel artificial bee colony algorithm and its PAPR reduction performance using SLM scheme in OFDM and MIMO-OFDM systems," *IEEE Communications Letters*, vol. 19, no. 10, pp. 1830–1833, 2015.

[107] N. VenkateswaraRao and C. Venkateswarlu, "Hybrid ABC optimization based interference cancellation in MIMO-OFDM," in *Communication and Electronics Systems (ICCES), 2017 2nd International Conference on*. IEEE, 2017, pp. 21–25.

[108] F. Bouchibane and M. Bensebti, "Parameter tuning of artificial bee colony algorithm for energy efficiency optimization in massive MIMO systems," in *Detection Systems Architectures and Technologies (DAT), Seminar on*. IEEE, 2017, pp. 1–5.

[109] F. Glover, "Tabu search-part i," *ORSA Journal on computing*, vol. 1, no. 3, pp. 190–206, 1989.

[110] X. Gao, L. Dai, C. Yuen, and Z. Wang, "Turbo-like beamforming based on tabu search algorithm for millimeter-wave massive MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 5731–5737, 2016.

[111] N. Srinidhi, T. Datta, A. Chockalingam, and B. S. Rajan, "Layered tabu search algorithm for large-MIMO detection and a lower bound on ML performance," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 2955–2963, 2011.

[112] T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Random-restart reactive tabu search algorithm for detection in large-MIMO systems," *IEEE Communications Letters*, vol. 14, no. 12, pp. 1107–1109, 2010.

[113] H. Zhao, H. Long, and W. Wang, "Tabu search detection for MIMO systems," in *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*. IEEE, 2007, pp. 1–5.

[114] Y.-H. Chou, Y.-J. Yang, and C.-H. Chiu, "Classical and quantum-inspired tabu search for solving 0/1 knapsack problem," in *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1364–1369.

[115] M. Teeti, R. Wang, H. Chen, Y. Liu, and Q. Ni, "Quantum-inspired evolutionary algorithm for large-scale MIMO detection," in *Personal, Indoor, and Mobile Radio Communications (PIMRC), 2017 IEEE 28th Annual International Symposium on*. IEEE, 2017, pp. 1–6.

[116] U. Pareek, M. Naeem, and D. C. Lee, "Quantum inspired evolutionary algorithm for joint user selection and power allocation for uplink cognitive MIMO systems," in *Computational Intelligence in Scheduling (SCIS), 2011 IEEE Symposium on*. IEEE, 2011, pp. 33–38.

[117] S. Chatzinotas, M. A. Imran, and R. Hoshyar, "On the multicell processing capacity of the cellular MIMO uplink channel in correlated rayleigh fading environment," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3704–3715, 2009.

[118] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*. Cambridge university press, 2010.

[119] S. Imre and F. Balazs, *Quantum Computing and Communications: an engineering approach*. John Wiley & Sons, 2005.

[120] K.-H. Han and J.-H. Kim, "Quantum-inspired evolutionary algorithm for a class of combinatorial optimization," *IEEE Transactions on evolutionary computation*, vol. 6, no. 6, pp. 580–593, 2002.

[121] S. Imre and F. Balazs, *Quantum Computing and Communications: an engineering approach*. John Wiley & Sons, 2013.

[122] C. Zhang, Y. Huang, Y. Jing, S. Jin, and L. Yang, "Sum-rate analysis for massive MIMO downlink with joint statistical beamforming and user scheduling," *IEEE Transactions on Wireless Communications*, vol. 16, no. 4, pp. 2181–2194, 2017.

[123] C. Jiang and L. J. Cimini, "Antenna selection for energy-efficient MIMO transmission," *IEEE Wireless Communications Letters*, vol. 1, no. 6, pp. 577–580, 2012.

[124] Y. Pei, T.-H. Pham, and Y.-C. Liang, "How many rf chains are optimal for large-scale mimo systems when circuit power is considered?" in *Global Communications Conference, 2012 IEEE*. IEEE, 2012, pp. 3868–3873.

[125] Y. Gao, A. H. Vinck, and T. Kaiser, "Massive MIMO antenna selection: Switching architectures, capacity bounds and optimal antenna selection algorithms," *IEEE Transactions Signal Processing*, 2017.

[126] R. J. Baxley and G. T. Zhou, "Comparing selected mapping and partial transmit sequence for PAR reduction," *IEEE Transactions on Broadcasting*, vol. 53, no. 4, pp. 797–803, 2007.

[127] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 3, pp. 317–325, 1980.

[128] A. Yang, Z. He, C. Xing, Z. Fei, and J. Kuang, "The role of large-scale fading in uplink massive MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 1, pp. 477–483, 2016.

[129] C. Wang, E. K. Au, R. D. Murch, W. H. Mow, R. S. Cheng, and V. Lau, "On the performance of the MIMO zero-forcing receiver in the presence of channel estimation error," *IEEE Transactions on Wireless Communications*, vol. 6, no. 3, 2007.

[130] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, 2003.

[131] P. Zhao, M. Zhang, H. Yu, H. Luo, and W. Chen, "Robust beamforming design for sum secrecy rate optimization in MU-MISO networks," *IEEE Transaction on Information Forensics Security*, vol. 10, no. 9, pp. 1812–1823, 2015.

[132] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21, Apr. 2011," *Available: cvxr. com/cvx*.

[133] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. Siam, 2001, vol. 2.