Towards an Enhanced Noncoherent Massive MU-MIMO System

Fahad Abdalrahman Alsifiany

Newcastle University

Newcastle upon Tyne, UK

A thesis submitted for the degree of

Doctor of Philosophy

September 2019

I dedicate this thesis to

my loving parents, **Abdalrahman** and **Saadah**,

my beloved wife, **Khairiya**,

my lovely daughter, **Farah**, and my lovely sons, **Faris**, **Faisal** and **Yousef**,

my brothers and sisters, **Abdalraheem**, **Eid**, **Omar**, **Rahmah**, **Fatimah**, **Jamilah**,

and **Monerah**,

and to the memory of my beloved brother **Mohammed**.

Declaration

NEWCASTLE UNIVERSITY

SCHOOL OF ELECTRICAL AND ELECTRONIC ENGINEERING

I, Fahad Abdalrahman Alsifiany, 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: Fahad Abdalrahman Alsifiany

Date:

SUPERVISOR'S CERTIFICATE

This is to certify that the entitled thesis "Towards an Enhanced Noncoherent Massive MU-MIMO System" has been prepared under my supervision in the School of Electrical and Electronic Engineering / Newcastle University for the degree of PhD in Electrical and Electronic Engineering.

Signature:

Supervisor: Prof. Jonathon Chambers FREng

Date:

Acknowledgements

In the name of Allah, the most Beneficent and the most Merciful. In the beginning, I would like to thank Allah for giving me the strength and the courage throughout my PhD study.

First, I would like to thank my supervisors, Prof. Jonathon Chambers and Dr. Aissa Ikhlef for all their limitless support and inspiration. I am grateful to them for providing me with great motivation, enthusiasm, technical insight and encouragement during my PhD study. This work would not have been possible without them. I wish that I will have more opportunities to work with them in the future. I also would like to thank Dr. Charalampos Tsimenidis and Prof. Said Boussakta for their administrative support.

Also, I would like to express my gratitude and appreciation to my sponsor: King Fahad Security Collage at the Ministry of Interior - Saudi Arabia, for offering the financial support for sponsoring my study in Newcastle University.

Finally, I really would like to express my sincere heartfelt thanks, appreciations and gratefulness to all my family members, especially my mother and father for their endless love, prayers, attention, and their support throughout my PhD study and before. I would also like to thank my wife and my children for their constant support and great patience. I extend my profound gratitude to all my friends and all people who have helped me throughout my research years.

Abstract

Many multiple-input multiple-output (MIMO) downlink transmission schemes assume channel state information (CSI) is available at the receiver/transmitter. In practice, knowledge of CSI is often obtained by using pilot symbols transmitted periodically. However, for some systems, due to high mobility and the cost of channel training and estimation, CSI acquisition is not always feasible. The problem becomes even more difficult when many antennas are used in the system and the channel is changing very rapidly before training is completed. Moreover, as the number of transmit/receive antennas grows large, the number of pilot symbols, system overheads, latency, and power consumption will grow proportionately and thereby the system becomes increasingly complex. As an alternative, a noncoherent system may be used wherein the transmitter/receiver does not need any knowledge of the CSI to perform precoding or detection. This thesis focuses on the design of a noncoherent downlink transmission system to jointly improve the performance and achieve a simple low complexity transmission scheme in three MIMO system scenarios: low rate differential spacetime block coding (STBC) in a downlink multiuser (MU-MIMO) system; high rate differential algebraic STBC in a downlink MU-MIMO system; and differential downlink transmission in a massive MU-MIMO system. Three novel design methods for each of these systems are proposed and analysed thoroughly.

For the MIMO system with a low rate noncoherent scheme, a differential STBC MU-MIMO system with a downlink transmission scheme is considered. Specifically, downlink precoding combined with differential modulation (DM) is used to shift the complexity from the receivers to the transmitter. The block diagonalization (BD) precoding scheme is used to cancel co-channel interference (CCI) in addition to exploiting its advantage of enhancing diversity. Since the BD scheme requires channel knowledge at the transmitter, the downlink spreading technique along with DM is also proposed, which does not require channel knowledge neither at the transmitter nor at the receivers. The orthogonal spreading (OS) scheme is

employed to have similar principle as code division multiple access (CDMA) multiplexing scheme in order to eliminate the interference between users. As a STBC scheme, the Alamouti code is used that can be encoded/decoded using DM thereby eliminating the need for channel knowledge at the receiver. The proposed schemes yield low complexity transceivers while providing good performance.

For the MIMO system with a high rate noncoherent scheme, a differential STBC MU-MIMO system that operates at a high data rate is considered. In particular, a full-rate full-diversity downlink algebraic transmission scheme combined with a differential STBC systems is proposed. To achieve this, perfect algebraic space time codes and Cayley differential (CD) transforms are employed. Since CSI is not needed at the differential receiver, differential schemes are ideal for multiuser systems to shift the complexity from the receivers to the transmitter, thus simplifying user equipment. Furthermore, OS matrices are employed at the transmitter to separate the data streams of different users and enable simple single user decoding. In the OS scheme, the transmitter does not require any knowledge of the CSI to separate the data streams of multiple users; this results in a system which does not need CSI at either end. With this system, to limit the number of possible codewords, a sphere decoder (SD) is used to decode the signals at the receiving end. The proposed scheme yields low complexity transceivers while providing full-rate full-diversity system with good performance.

Lastly, a differential downlink transmission scheme is proposed for a massive MIMO system without explicit channel estimation. In particular, a downlink precoding technique combined with a differential encoding scheme is used to simplify the overall system complexity. A novel precoder is designed which, with a large number of transmit antennas, can effectively precancel the multiple access interference (MAI) for each user, thus enhancing the system performance. Maximising the worst case signal-to-interference-plus-noise ratio (SINR) is adopted to optimise the precoder for the users in which full power space profile (PSP) knowledge is available to the base station (BS). Also, two suboptimal solutions based on the matched and the orthogonality approach of PSP are provided to separate the data streams of multiple users. The decision feedback differential detection (DFDD) technique is employed to further improve the performance.

In summary, the proposed methods eliminate MAI, enhance system performance,

and achieve a simple low complexity system. Moreover, transmission overheads are significantly reduced, the proposed methods avoid explicit channel estimation at both ends.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms

5G	Fifth generation
AN	Artificial noise
AP	Access point
AWGN	Additive white Gaussian noise
BD	Block diagonalization
BER	Bit-error rate
BLER	Block error rate
BPSK	Binary phase shift keying
BS	Base station
CCI	Co-channel interference
CD	Cayley differential
CDMA	Code division multiple access
CM	Coherent modulation
CPU	Central processing unit
CSI	Channel state information
dB	Decibel
DFDD	Decision feedback differential detection
DFT	Discrete Fourier transform

DM	Differential modulation
DM-RS	Demodulation reference signals
DPC	Dirty-paper-coding
DPSK	Differential phase shift keying
DSTM	Differential spacetime modulation
EH	Energy harvester
EV	Eavesdropper
FDD	Frequency-domain duplex
FLOP	Floating point operation
G2-STBC	Alamouti spacetime block code
GLRT	Generalised likelihood ratio test
GSP	Gram-Schmidt process
i.i.d.	Independent and identically distributed
IU	Information user
LDC	Linear dispersion block codes
LP	Linear programming
MAC	Media access control
MAI	Multiple access interference
MIMO	Multiple-input multiple-output
MISO	Multiple-input single-output
ML	Maximum-likelihood
MMSE	Minimum mean-square error
MRC	Maximum-ratio combining
MSDD	Multiple symbol differential detection

MSE	Mean square error
MU-MIMO	Multiuser multiple-input multiple-output
OS	Orthogonal spreading
PAPR	Peak-to-average power ratio
PDF	Probability density function
PEP	Pairwise error probability
PSK	Phase-shift keying
PSP	Power space profile
QAM	Quadrature amplitude modulation
QO-STBC	Quasi-orthogonal spacetime block code
QPSK	Quadrature phase shift keying
RF	Radio frequency
RS	Reference signals
Rx	Receiver
RZF	Regularized zero-forcing
SC	Selection combining
SD	Sphere decoder
SER	Symbol error rate
SIMO	Single-input multiple-output
SINR	Signal-to-interference-plus-noise ratio
SISO	Single-input single-output
SLNR	Signal-to-leakage-noise ratio
SNR	Signal-to-noise ratio
STBC	Space-time block coding

SU-MIMO	Single-user multiple-input multiple-output
SVD	Singular-value decomposition
SWIPT	Simultaneous wireless information and power transfer
TAST	Threaded algebraic spacetime
TDD	Time domain duplex
Tx	Transmitter
UWB	Ultra-wideband
V-BLAST	Bell laboratories layered spacetime
VP	Vector perturbation
ZF	Zero-forcing

Notations

$\boldsymbol{A}$	Uppercase boldface characters for matrices
$\boldsymbol{a}$	Lowercase boldface characters for vectors
a	Lowercase characters for scalar
$\boldsymbol{I}_n$	$n \times n$ identity matrix
$\mathbf{0}_{m\times n}$	$m \times n$ zero matrix
$\mathbf{1}_n$	$n \times 1$ all-one column vector
$\text{diag}(\boldsymbol{x})$	Diagonal matrix with the elements of $\boldsymbol{x}$ in the main diagonal
$\Re(\boldsymbol{Q})$	The real part of matrix $\boldsymbol{Q}$
$\Im(\boldsymbol{Q})$	The imaginary part of matrix $\boldsymbol{Q}$
$\boldsymbol{S} \succeq 0$	Indicates that $\boldsymbol{S}$ is a positive semi-definite matrix
$\boldsymbol{S} \succ 0$	Indicates that $\boldsymbol{S}$ is a positive definite matrix
$\rho(\boldsymbol{S})$	The largest eigenvalue of matrix $\boldsymbol{S}$
$\mathcal{R}$	The set of real numbers
$\mathcal{R}_+^n$	The set of nonnegative real numbers
$\mathcal{S}_+^n$	The set of symmetric positive semidefinite $n \times n$ real matrices
$\mathbb{C}^{m\times n}$	The set of all complex $m \times n$ matrices
$\mathbb{R}^{m\times n}$	The set of all real $m \times n$ matrices
$(\cdot)^*$	Conjugate
$(\cdot)^T$	Transpose
$(\cdot)^H$	Conjugate transpose
$(\cdot)^\dagger$	The pseudo-inverse
$\text{trace}(\cdot)$	Trace of a matrix
$\log(\cdot)$	Natural logarithm

$\log_2(\cdot)$	Logarithm to base 2
$\lvert\cdot\rvert$	Absolute value of scalars
$\lVert\cdot\rVert_F$	Frobenius norm of matrices
$\det(\cdot)$	The determinant operation
$\prod(\cdot)$	The product operation
$\boldsymbol{x} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$	A complex Gaussian random vector with zero mean and covariance matrix $\boldsymbol{\Sigma}$.
$\mathbb{E}\{\cdot\}$	The statistical expectation value of the a discrete random variable $\boldsymbol{a}$
$\{\boldsymbol{a}_n\}$	A set of all vectors indexed by n
$\{a_{m,n}\}_m$	A set of all scalars indexed by m
$[\boldsymbol{a}]_i$	The ith entry of vector $\boldsymbol{a}$
$(x)_n$	The x modulo n
$\otimes$	The Kronecker product
$\circ$	The Hadamard product
$[\boldsymbol{A}]_{n,m}$	The nth entry of the mth column of $\boldsymbol{A}$
$\dim(\boldsymbol{a})$	The length of the vector $\boldsymbol{a}$
$\boldsymbol{x} \succeq 0$	Equivalent to $[\boldsymbol{x}]_i \geq 0$, $i = 1, ..., \dim(\boldsymbol{x})$
$\boldsymbol{B} = \text{null}(\boldsymbol{A})$	The null space of $\boldsymbol{A}$ which satisfies $\boldsymbol{AB} = \mathbf{0}$ and $\boldsymbol{BB}^H = \boldsymbol{I}$
$\text{cov}(x, y)$	The covariance between the random variables x and y
$\text{var}(x)$	Variance of x
$f(N) \overset{N\to\infty}{\to} a$	Equivalent to $\lim_{N\to\infty} f(N) = a$
$[x]^+$	Equivalent to $\max(x, 0)$

Chapter 1

Introduction

1.1 Overview

Wireless communication has undergone a remarkable and rapid evolution. In the past decade, the wireless evolution has reached a level where researchers, designers and manufacturers increasingly depart from the conventional concept of communicating on a single point-to-point basis with a central controlling base station (BS). Indeed, it was shown in [1–4] that multiuser multiple-input multiple-output (MU-MIMO) offers numerous benefits over conventional point-to-point MIMO. Firstly, MU-MIMO works with cheap single-antenna terminals. Additionally, MU-MIMO does not require a rich scattering environment. Furthermore, resource allocation in MU-MIMO is simplified because all the active terminals use all the frequency and time bins. However, MU-MIMO is constrained in terms of scalability.

Massive MIMO (also referred to as hyper MIMO or, large-scale MIMO) counters this constraint by using a large excess of service-antennas over time-domain duplex (TDD) operation and active terminals [2]. These extra antennas focus the system energy into smaller regions of space to improve the radiated energy efficiency and throughput. The other noted benefits of large scale MIMO include reduced latency, inexpensive power components, robustness against deliberate jamming and simplification of the media access control (MAC) layer [5]. On the other hand, it has been demonstrated that cooperative communication can improve the capacity and/or the coverage of wireless networks. Thus, massive MIMO and cooperative diversity schemes, can be used together to mitigate fading and further improve the performance of wireless networks. While massive MIMO delivers numerous benefits, it comes with new challenges that need to be addressed.

There exists a large and growing body of literature that supports MU-MIMO transmit precoding techniques [6, 7]. In this kind of scheme, the availability of channel state information (CSI) at the transmitter makes it possible for the precoder to precancel co-channel interference (CCI) at each mobile user. There are two steps that allow the base station (transmitter) to obtain CSI from the receiver in downlink precoding. First, the BS transmits training sequences in TDD and then it uses feedback from the mobile nodes (receiver). However, the most critical factor that affects the throughput gains provided by MU-MIMO and increases complexity at both the transmitter and receiver, is the feedback mechanism, since all the information must be passed by the BS before channel conditions change and the BS periodically receive CSI from mobile nodes, thus incurring overhead. Furthermore, the cost of obtaining CSI at the transmitter requires coordinating the process between the transmitters and receivers in order to decode the transmitted signals effectively.

The impact of the receiver estimation process and/or overhead problem can be potentially solved by using differential modulation (DM). The mobile node in this scheme would not require the capability for coherent detection in order to measure channel conditions using training sequences, i.e. the mobile nodes could not be used either to decode unknown data transmission or to estimate CSI to feed it back to the BS.

This PhD thesis will concentrate on designing, analysing and evaluating new advanced downlink signal transmission techniques for noncoherent scheme. I will consider both conventional and massive MU-MIMO wireless communication systems. The evaluation process will be carried out using MATLAB.

1.2 Contributions

This thesis covers three topics (presented in Chapters 3 to 5) related to the design of noncoherent downlink transmission systems: exploiting low rate differential spacetime block code (STBC) in downlink MU-MIMO systems; exploiting high rate differential algebraic STBC in downlink MU-MIMO systems; and differential downlink transmission in massive MU-MIMO systems. The contributions in each of the three chapters are outlined as follows:

1. **Exploiting low rate differential STBC in downlink MU-MIMO systems**: Chapter 3 considers a STBC code MU-MIMO system with downlink transmission system. The main contributions are as follows:

   (a) The use of DM for downlink transmission in a MU-MIMO system is considered to shift the complexity from the receivers to the transmitter. Specifically, I show how to use DM combined with the block diagonalization (BD) and orthogonal spreading (OS) schemes.

   (b) DM is considered for both BD and OS schemes based on the Alamouti STBC in order to eliminate the need for estimating the composite channels formed by the precoders and the channels at the receivers.

   (c) Since the BD scheme still requires CSI at the transmitter, a downlink OS scheme combined with DM is proposed. In the OS scheme, unlike the BD scheme, the transmitter does not require any knowledge of the CSI to separate the data streams of multiple users. Therefore, implementing the OS scheme with the DM will result in a system that does not need CSI at either ends. The transmission overhead is

significantly reduced using the proposed scheme, since neither feedback nor the estimation of the composite channels are required.

2. **Exploiting high rate differential algebraic STBC in downlink MU-MIMO systems**: Chapter 4 considers a STBC code MU-MIMO system that operates at a high data rate with full diversity. In particular, I propose to use a full rate downlink algebraic transmission scheme combined with a differential space-time scheme for MU-MIMO systems. The main contributions presented in this chapter are as follows:

   (a) In this work, the use of high rate Cayley differential (CD) STBC for downlink transmission in a MU-MIMO system is considered. Specifically, I show how to use differential STBC combined with full rate full diversity perfect algebraic STBC. The use of differential STBC in a multiuser scenario simplifies the complexity of the receivers, since neither feedback nor the estimation of the CSI is required at the receiver.

   (b) Differential STBC is considered based on the orthogonal spreading technique in order to separate the data streams of multiple users. With the use of orthogonal spreading, the transmitter needs no knowledge of the CSI to design the spreading matrices.

   (c) At the receiver of each user, a sphere decoder (SD) is implemented for high rate coherent and differential perfect algebraic STBC to limit the set of candidate symbols to those within a sphere of some radius d.

   (d) The proposed schemes facilitate multiple user data separation, enhancing full rate full diversity, and achieving low complexity receivers and transmitters through the use of differential STBC.

3. **Differential downlink transmission in massive MU-MIMO systems**: Chapter 4 considers a differential downlink transmission scheme for a massive MU-MIMO system without explicit channel estimation. The main contributions of this chapter are:

   (a) I propose a differential MIMO downlink transmission framework, in which a BS is equipped with a massive antenna array that precodes transmit signals to separate the data streams of multiple users.

   (b) A novel downlink precoding design is proposed by employing knowledge of the power space profile (PSP) of users to achieve a low-complexity differential massive

MIMO system.

(c) I provide an optimal solution for the precoder based on a max-min signal-to-interference-plus-noise ratio (SINR) problem formulation. The optimised precoder can effectively precancel the interference between users, thus enhancing overall system performance.

(d) Further, I provide two suboptimal solutions suitable for the low interference system based on the matched and the orthogonality approach of PSP of each user.

(e) The proposed schemes facilitate precancelling MAI, enhance system performance, and provide simple transmitter and receiver schemes. Consequently, since the proposed scheme avoids channel estimation, the system overheads and latency will be reduced significantly.

1.3 List of Publications

This thesis has three contribution chapters. The list of publications is categorised based on these chapters as follows.

1. **Chapter 3: Exploiting low rate differential STBC in downlink MU-MIMO Systems**

   - F. Alsifiany, A. Ikhlef, and J. A. Chambers.: "Exploiting differential modulation within a multiuser-MIMO system with downlink precoding", presented at the 11th International Conference on Mathematics in Signal Processing, Birmingham, UK, December 2016.

   - F. Alsifiany, A. Ikhlef, and J. Chambers, "On differential modulation in downlink multiuser MIMO systems," in 2017 25th European Signal Process. Conf. (EUSIPCO), Kos, Greece, Aug. 2017, pp. 558-562.

2. **Chapter 4: Exploiting high rate differential algebraic STBC in downlink MU-MIMO Systems**

   - F. Alsifiany, A. Ikhlef, and J. Chambers, "Exploiting high rate differential algebraic space-time block code in downlink multiuser MIMO systems," IET Commun., vol. 12, no. 17, pp. 2188-2197, Oct. 2018.

3. **Chapter 5: Differential downlink transmission in massive MU-MIMO Systems**

    - F. Alsifiany, A. Ikhlef, M. Alageli, and J. Chambers, "Differential downlink transmission in massive MU-MIMO systems," *IEEE Access*, vol. 7, pp. 86 906-86 919, 2019.

4. **Other contributions/In preparation**

    - M. Alageli, A. Ikhlef, F. Alsifiany, M. A. M. Abdullah, G. Chen, and J. Chambers, "Optimal downlink transmission for cell-free SWIPT massive MIMO systems with active eavesdropping," IEEE Transactions on Information Forensics and Security, pp. 1-1, Nov. 2019.
    - F. Alsifiany, A. Ikhlef, M. Alageli, and J. Chambers, "Secrecy design for cell-free noncoherent massive MIMO systems with passive eavesdropper". This is under preparation.

1.4 Thesis Outline

This thesis is organized as follows:

Chapter 2 presents general background and a focused literature review including, multiple-antenna systems, massive MIMO and differential space-time modulation. In addition, a survey on previous works related to the design of noncoherent multiuser MIMO systems is discussed.

Chapter 3 is the first contribution chapter that includes the use of low rate differential STBC. The use of DM is combined with BD and OS techniques for a downlink transmission MU-MIMO system to enhance and form the precoder at the transmitter. The DM technique is considered based on the Alamouti STBC to eliminate the need for estimating the composite channels. Since neither feedback nor the estimation of the composite channels are required, the transmission overhead is reduced using the proposed scheme.

Chapter 4 focuses on exploiting high rate differential STBC. Full rate full diversity STBC is accomplished by using algebraic STBC along with a differential MU-MIMO system whereby the receiver does not need any knowledge of the CSI to perform data detection. Differential STBC is considered based on the OS technique to separate the data streams of multiple users and enhance the system performance. In addition, a sphere decoder is implemented in this chapter to minimise the search set of candidate symbols and hence speed up the detection process.

Chapter 5 introduces a novel precoder design suitable for a noncoherent downlink massive MU-MIMO system. An optimal solution for the differential precoder based on max-min SINR is presented in this chapter. The optimised precoder can effectively eliminate the interference between users and thus enhance the system performance. In addition, two suboptimal differential precoders suitable for the low interference system are provided in this chapter.

Finally, the conclusions are drawn in Chapter 6 and ideas for future work are suggested.

Chapter 2

Background and Focused Literature Review

2.1 Multiple-Antenna Systems

In multiple antenna systems, more than one antenna is used at the transmitter and/or at the receiver leading to multiple wireless channels between the transmitter and the receiver [3]. Given fixed transmission power and bandwidth, building a digital wireless communication system that offers high data rate and less error rate is the core part of the trade-off between spectral efficiency and power efficiency. Utilizing multiple antennas at the transmitter and/or at the receiver can exploit the spatial domain and results in better trade-off between spectral efficiency and power efficiency [3]. Many transmission and reception techniques have been developed to effectively achieve better performance provided by multiple antenna systems. One of the important techniques, is the space time block coding (STBC) technique.

2.1.1 Space-Time Block Coding

STBC is a transmission technique in which multiple data streams are transmitted over multiple antennas. In particular, the information is coded in the spatial and time domains to improve the spectral efficiency and/or reliability of data transmissions [1, 3, 4]. The transmitted signals are distorted by channel fading and corrupted by noise at the receiver. The receiver processes the coded signals and decodes the information [3]. STBC allows diversity gain to be achieved and in some cases array (also known as coding) gain.

2.1.1.1 Diversity Gain

In a wireless channel the transmit signal fluctuates and fades on its way to the receiver. Diversity techniques are used in communication systems to combat fading [4]. There are various forms of diversity systems. In receive diversity, a single-input multiple-output (SIMO) system is used and the received signals at the antennas are combined using different combining algorithms such as selection combining (SC) and maximal ratio combining (MRC). In SIMO systems, diversity is measured by the number of independent fading branches which is in this case equal to the number of receive antennas. In transmit diversity, a multiple-input single-output (MISO) system is used and appropriate design techniques are needed at the transmitting antennas to extract diversity. One of the most popular techniques is STBC, which allows diversity to be extracted in the absence of channel knowledge at the transmitter. To enhance the transmit and the receive diversity at the same time, a MIMO system is used.

2.1.1.2 Array Gain

Array gain, also known as power gain, is an increase in the average power induced by using multiple antenna at the receiver and/or at the transmitter [4]. For example, in SIMO channels, the average increase in signal-to-noise ratio (SNR) at the receiver is proportional to the number of receive antennas, which is due to the coherent combination of the received signals.

2.1.2 MIMO Systems

A conventional point-to-point MIMO communication system consists of a single transmitter and a single receiver (user k) or multiple receivers (users K), each equipped with multiple antennas [8]. Fig. 2.1 shows a simple MIMO communication system that consists of multiple antennas at the transmitter (Tx) and multiple antennas at the receiver (Rx). The MIMO system can be either for a single-user MIMO (SU-MIMO) where all data streams are directed to one user, or for MU-MIMO where data streams are broadcast to multiple users.

Typically, communication between the transmitting antennas and the receiving antennas is direct, without cooperation, relaying, and/or any form of interference from any third party. However, in modern communication systems, communication might involve multiple nodes such as relay networks and ad hoc wireless networks, where there is indirect communication between multiple transmitting and receiving antennas [9]. Fig. 2.1 below is an illustration of a space-time transmission scheme in a point-to-point MIMO communication system. In this MIMO configuration, the transmitter has N_t transmit antennas and the receiver has N_r^k receive antennas, where k is the user index. There is a wireless channel between each pair of transmit and receive antennas.

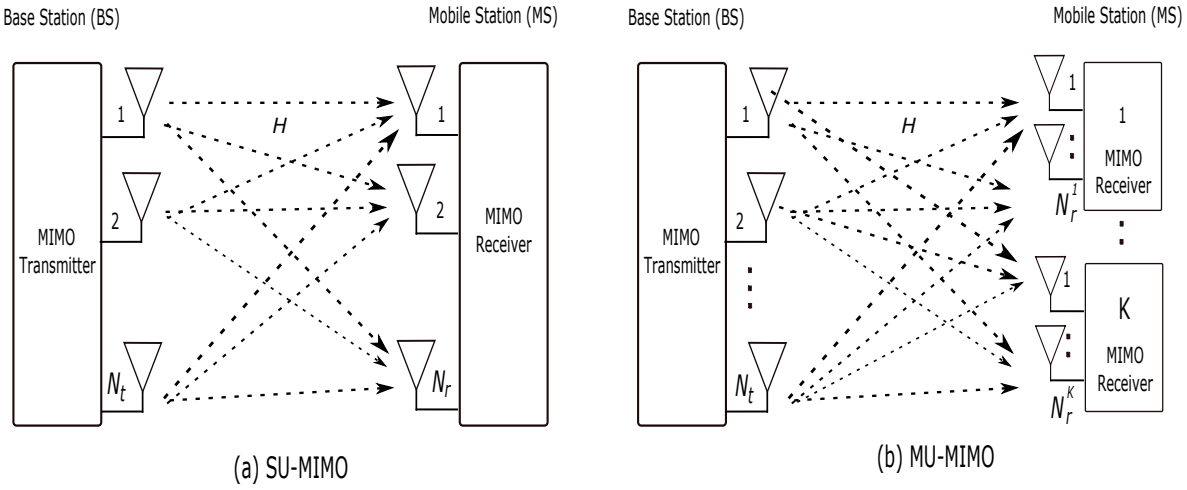

Figure 2.1. MIMO configurations.

Assuming that the channels are constant over T transmission intervals, the system equation of the multiple-antenna system can be represented as

$$\begin{bmatrix} y_{11} & \cdots & y_{1T} \\ \vdots & \ddots & \vdots \\ y_{N_r1} & \cdots & y_{N_rT} \end{bmatrix} = \sqrt{\rho} \begin{bmatrix} h_{11} & \cdots & h_{1N_t} \\ \vdots & \ddots & \vdots \\ h_{N_r1} & \cdots & h_{N_rN_t} \end{bmatrix} \begin{bmatrix} s_{11} & \cdots & s_{1T} \\ \vdots & \ddots & \vdots \\ s_{N_t1} & \cdots & s_{N_tT} \end{bmatrix} + \begin{bmatrix} z_{11} & \cdots & z_{1T} \\ \vdots & \ddots & \vdots \\ z_{N_r1} & \cdots & z_{N_rT} \end{bmatrix} \tag{2.1}$$

which can also be compactly expressed as

$$\boldsymbol{Y} = \sqrt{\rho}\boldsymbol{H}\boldsymbol{S} + \boldsymbol{Z} \tag{2.2}$$

where ρ is the average transmit power, $\boldsymbol{Y}$ is the $N_r \times T$ received signal matrix, $\boldsymbol{S}$ is the $N_t \times T$ transmitted coded signal matrix, $\boldsymbol{H}$ is the $N_r \times N_t$ frequency flat channel matrix, and $\boldsymbol{Z}$ is the $N_r \times T$ additive noise matrix.

The horizontal axis in matrix $\boldsymbol{S}$ represents the spatial domain and its vertical axis represents the temporal domain. This coding method is called space-time coding [2]. Furthermore, increasing the number of transmitting and receiving antennas in the system might increase the data rate and minimise the probability of error. Point-to-point MIMO systems potentially provide much higher spectral efficiency and/or reliability than traditional SISO [10, 11]. In the next subsection one of the most popular STBC techniques is described, namely the Alamouti STBC Scheme.

2.1.3 Alamouti Space Time Code Scheme

One of the elegant code examples associated with conventional MIMO is the Alamouti 2×2 space time coding scheme [12]. The Alamouti scheme is a transmit diversity scheme that is suitable for systems that do not need knowledge of the CSI at the transmitter [1]. In other words, the MIMO Alamouti code is a special case of the general case given in (2.1). The received signal for an Alamouti 2×2 MIMO system can be represented as

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \sqrt{P} \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix} + \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} , \tag{2.3}$$

.

where $(.)^*$ denotes complex conjugate; and can be compactly expressed as

$$\boldsymbol{Y_2} = \sqrt{\rho}\boldsymbol{H_2}\boldsymbol{S_2} + \boldsymbol{Z_2} \tag{2.4}$$

For decoding, the maximum likelihood (ML) decoder is applied:

$$\hat{\mathbf{S}}_2 = \arg \min_{\boldsymbol{S}\in\mathcal{C}} \|\boldsymbol{Y}_2 - \sqrt{\frac{\rho T}{N_t}}\boldsymbol{H}_2\boldsymbol{S}_2\| \,, \tag{2.5}$$

where, $\|.\|$ denotes the Frobenius norm, $\mathcal{C}$ is an available known finite codebook, $\boldsymbol{S}$ is the possible transmitted symbol vector based on the available data. If the cardinality of $\mathcal{C}$ is M, then the transmission rate of this code is given by

$$R = \frac{\log_2 M}{T}. \tag{2.6}$$

Therefore, the most important factor in designing the STBC is to design the set $\mathcal{C}$.

2.2 Massive MIMO

In the last ten years, the number of wirelessly connected devices has grown exponentially. More devices continue to be connected in wireless networks and the growth is expected to skyrocket with the concept of Internet of Things taking shape [13]. All these connected devices require high bandwidth so as to support real-time videos, games and movies. As a result, there is a demand for increase in the throughput of wireless networks. The growth gives rise to a concern regarding efficient consumption of energy for wireless networks and devices. In order to keep up with the exponential trend, industry needs systems that have a high throughput and serve many users at the time while consuming less energy as compared to traditional systems [13]. A solution that offers these requirements is massive MIMO technology. In this technology, several antenna arrays are attached to a BS either in a collocated or distributed arrangement and use the same frequency-time allocation to simultaneously serve multiple users [13]. The propagation channels become favorable for most environments when massive antennas are used in a BS. Linear processing is nearly optimum since the channel matrices are almost orthogonal between the base stations and the users. Using massive MIMO, a high throughput and great energy efficiency can be achieved because of the multiplexing and array gains [13]. Furthermore, massive MIMO technology is able to offer a good service that is uniform to all users when a proper power control scheme is used [13]. Given the capabilities of massive MIMO systems,

the efficiency of wireless devices has increased several folds. This section discusses massive MIMO technology, its working principles, and the challenges the technology faces.

2.2.1 Massive MU-MIMO

Advancement in technology has led to massive MU-MIMO. In multiple antennas technology, a BS is equipped with several antennas to serve several single-antenna users by sharing multiplexing gain among the users [14]. Massive MIMO is a MU-MIMO technology in which hundreds of BS antennas serve hundreds of users at the same time using a shared frequency resource. When the BS is equipped with antennas past a hundred, it is referred to as "Hyper MIMO" or "Full-Dimension MIMO". Massive MIMO is ten times more efficient in practical applications than conventional point-to-point MIMO and its signal processing techniques are far simpler [14]. Correspondingly, in multiple antennas systems, the channel estimation overhead depends on the number of BS antennas and hence a simple linear processing technique that operates at optimal capacity is required to handle the large dimensional vectors that result from multiple users and multiple antennas. In such case, TDD operation is the preferred method of operation as the increase in the number of antennas does not require additional resources for channel estimation [13]. In other words, with TDD, the channel estimation overhead is independent of the number of BS antennas so that the number of antennas can be increased to the desired level without any effect on the estimation overhead [13]. Accordingly, there exist three operations in a coherence interval of TDD, namely: uplink transmission, downlink transmission and channel estimation as depicted in Fig. 2.2 [13].

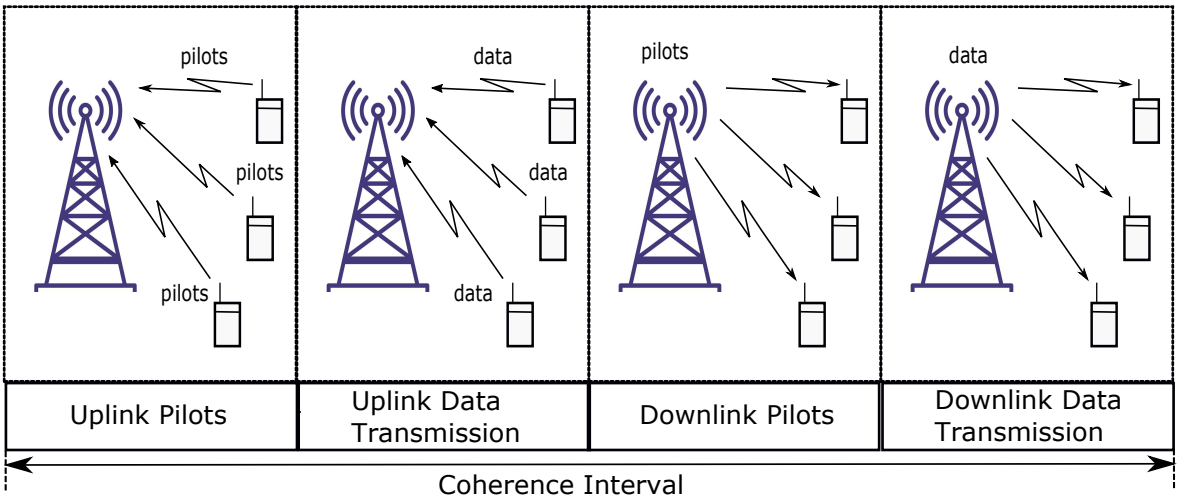

Figure 2.2. TDD Massive MIMO protocol.

2.2.2 Channel Estimation

The BS requires CSI for detection of user transmitted signals in the uplink and for precoding of those signals in the downlink. The BS learns the CSI through uplink training where each user sends a pre-assigned orthogonal pilot sequence to the corresponding BS. The BS uses the received pilot sequences for channel estimation [15]. Every user requires partial knowledge of CSI so as to be able to detect signals from the BS. To obtain this information, downlink training to the users is performed; alternatively blind channel estimation can be implemented. The user only requires the effective channel gain for signal detection since the signals are beamformed to the user via linear precoding methods [15]. Thus the BS requires less time to convey the CSI information to the users.

2.2.2.1 Channel Coherence Time

In wireless communications systems, a communication channel may change with time. Coherence time is the time duration over which the channel impulse response is considered to be constant (invariant). Such channel variation is significant in wireless communications systems, due to Doppler effects. In MU-MIMO system, some theoretic measurements studies about channel coherence time appeared in [16–18]. These studies suggest that high capacity can be achieved with no prior knowledge of the channel if the channel does not change too frequently. How frequently the channel may change is not completely clear.

In this thesis, it is assumed that the channel coherence time is large enough and changes slowly and extends over several matrix transmission periods that is at least twice the number of transmit antennas.

2.2.2.2 Fading Channel Model

Fading is the time variation of received signal power due to changes in paths or transmission medium. The types of fading in wireless communication system is depicted in Fig. 2.3 [4, 19]. In flat fading, the bandwidth of the channel is larger than the bandwidth of the signal. Hence, all frequency components of the signal will experience the same magnitude of fading. In frequency-selective fading, the coherence bandwidth of the channel is smaller than the bandwidth of the signal.

In this thesis, it is assumed that the fading of the channel is flat. Also, it is assumed that the implementation of fading model or fading distribution is Rayleigh flat fading. In Rayleigh

model, only non line of sight (NLOS) components are simulated between transmitter and receiver. It is assumed that there is no LOS path exists between transmitter and receiver [4, 19].

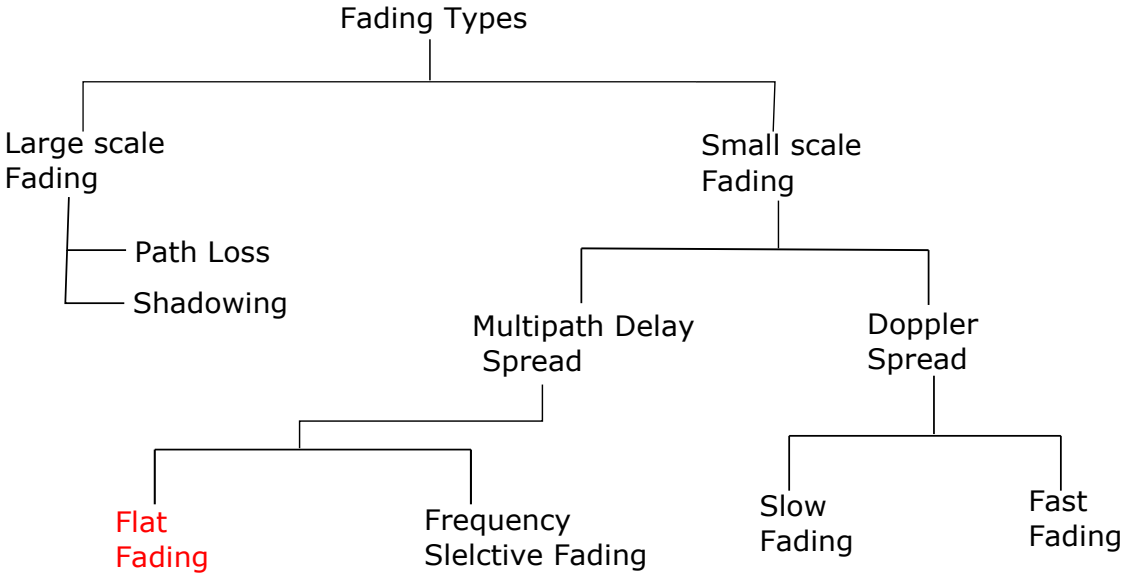

Figure 2.3. Types of fading in wireless communication.

2.2.2.3 Channel Estimation for TDD Operation

For TDD systems, different time slots are used by the uplink and downlink operations whilst the same frequency resource is used [14, 20]. In TDD uplink operation, all the users first synchronously send uplink data signals. Next, the users send orthogonal pilot sequences where these pilot signals are detected by the BS and used to estimate the CSI, detect the uplink data, and generate precoding vectors for downlink data transmission. The channel usage in this case is equivalent to the number of users sending the pilot sequences, i.e., K channel uses.

In TDD downlink operation, the BS can utilize the precoding by using the uplink channel estimation since there is channel reciprocity. On the other hand, for the effective channel gain, the users can use pilot signals beamformed by the BS; a process that needs again a minimum of channel uses equivalent to the number of users. i.e., K channel uses. As a result the training process needs in total, at least $2K$ channel uses, where K is the total number of users.

2.2.2.4 Channel Estimation for FDD Operation

In frequency-domain duplex (FDD) systems, the downlink and uplink operations require different frequency resources and therefore there is no channel reciprocity. In the downlink transmission, the BS requires CSI for pre-coding of the signals to be transmitted to the users. Each user receives pilot sequences from the BS that it uses them to estimate the channels [21]. The

channel estimates are sent back to the BS during the uplink transmission. The uplink transmission involves BS decoding the received signals; a process that requires CSI. Orthogonal pilot sequences are sent by the users to the BS whereupon they are used for channel estimation and the process has K channel usage [21]. An increase in BS antenna number results in a higher number of channel replies, leading to consumption of more resources in the uplink operation. For this reason, TDD is preferred to FDD for massive MIMO systems [15].

2.2.3 Uplink Transmission

It is the part of the coherence interval used by the users to send data to the BS. The same time-frequency resource is shared among the users. To detect the signal from the users, the BS utilizes channel estimation and linear combination mechanisms [13].

2.2.4 Downlink Transmission

In the downlink coherence interval, the BS uses the same time-frequency allocation to communicate with the users. The symbols to be used by the users, coupled with the channel estimates of the BS, are used by the BS to precode the signals that are then transmitted by the BS antennas [13].

2.2.5 Favorable Propagation Conditions in Massive MIMO

The future belongs to the technology that will be able to simultaneously support multiple users at a high throughput while consuming less power. Massive MIMO serves these needs and therefore will continue being deployed in high capacity wireless networks. The most important observations and results for very large random vectors as in massive MIMO are the so-called favourable propagation conditions. Favourable propagation defined as mutual orthogonality among the vector-valued channels to the terminals, where the column-vectors of the propagation vectors are asymptotically orthogonal [22]. Let $\boldsymbol{p} \triangleq [p_1 \cdots p_n]^T$ and $\boldsymbol{q} \triangleq [q_1 \cdots q_n]^T$ be $n \times 1$ vectors with independent and identically distributed (i.i.d.) zero mean random variables with $\mathbb{E}\{|p_i|^2\} = \sigma_p^2$ and $\mathbb{E}\{|q_i|^2\} = \sigma_q^2$, $i = 1, \cdots, n$. Then from the law of large numbers, it follows

$$\frac{1}{n}\boldsymbol{p}^H\boldsymbol{p} \overset{a.s.}{\to} \sigma_p^2, \text{ and } \quad \frac{1}{n}\boldsymbol{p}^H\boldsymbol{q} \overset{a.s.}{\to} 0, \text{ as } n \to \infty \tag{2.7}$$

where $\overset{a.s.}{\to}$ denotes almost sure convergence. Also, from the central limit theorem, we have

$$\frac{1}{\sqrt{n}}\boldsymbol{p}^H\boldsymbol{q} \overset{d}{\to} \mathcal{CN}(0, \sigma_p^2\sigma_q^2), \text{ as } n\to\infty \tag{2.8}$$

where $\overset{d}{\to}$ denotes convergence in distribution. The conditions in (2.7)-(2.8) are satisfied under the most favourable propagation conditions, where the column-vectors of the propagation vectors are asymptotically orthogonal.

2.2.6 The Capacity of Massive MIMO

This type of MIMO achives higher data rate through multiplexing and its diversity allows the enhancement of the reliability for the communication link [14]. Point-to-point MIMO's throughput increases linearly with the number of antennas at the receiver or transmitter. The received signal $\boldsymbol{y} \in C^{N_r\times 1}$, can be expressed as

$$\boldsymbol{y} = \sqrt{\rho}\boldsymbol{H}\boldsymbol{s} + \boldsymbol{z}, \tag{2.9}$$

where $\boldsymbol{s} \in C^{N_t\times 1}$ is the transmit signal, $\boldsymbol{z} \in C^{N_r\times 1}$ is the additive noise vector modelled as zero-mean complex circularly symmetric Gaussian random vector with an identity covariance matrix, i.e., $\boldsymbol{z} \sim \mathcal{CN}(0, \boldsymbol{I}_{N_r})$, $E\left\{\|\boldsymbol{s}\|^2\right\} = 1$ is the normalized total power of the transmit signal. With these assumptions, the scalar ρ is the transmit power. The achievable rate for i.i.d. Gaussian transmitted signal with CSI available at the receiver is expressed as

$$C = \log_2 \det\left(\boldsymbol{I} + \frac{\rho}{N_t}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right) \frac{\frac{\text{bits}}{\text{s}}}{\text{Hz}}, \tag{2.10}$$

with normalized propagation coefficients for the matrix $\boldsymbol{H}$, as $\text{Tr}(\boldsymbol{H}\boldsymbol{H}^H) \approx N_rN_t$. The upper and lower bounds for the capacity are expressed as [20]:

$$\log_2(1+\rho N_r) \leq C \leq \min(N_t, N_r)\log_2\left(1+\frac{\rho\max(N_t, N_r)}{N_t}\right). \tag{2.11}$$

The sum rate is dependent on the distribution of the singular values of $\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}$. With the exponential growth of transmit and receive antennas for the massive MIMO system, we have the following two cases:

1- First case, if the number of receiver antennas is held constant while the transmit antenna number tends to infinity, i.e., $N_t \to \infty$, then, for coherent modulation (CM) system, the rate

that can be achieved in this case is approximated as

$$C \approx N_r \log_2(1+\rho) \quad \frac{\frac{\text{bits}}{\text{s}}}{\text{Hz}}. \tag{2.12}$$

2- Second case, holding constant the number of antennas for transmission while that for receiver antenna is let to tend to infinity, i.e., $N_r \to \infty$, the resultant rate is approximated as:

$$C \approx N_t \log_2\left(1+\frac{\rho N_r}{N_t}\right) \quad \frac{\frac{\text{bits}}{\text{s}}}{\text{Hz}}. \tag{2.13}$$

It can be seen from (2.12) and (2.13) that when MIMO systems use many antennas, as in the case with massive MIMO systems, the highest rate is achieved, i.e., the upper bound in (2.11), [15].

This implies it is possible to serve more users at a better throughput at the same time sharing the same frequency resource without any increase in the transmitted power per terminal. Additionally, the quality of service is maintained and the transmit power is slashed by 3 dB if the number of BSs is doubled [14]. Since the antenna array and multiplexing gains above are arrived at using optimal processing and by favourable propagation model, there is a concern as to whether similar gains are attainable using linear processing. There is a temptation to employ point-to-point MIMO that yields low dimensions and employs complex signal processing instead of the massive MIMO which is simple to process. Linear processing is optimal in massive MIMO and leads to similar gains since an increase in BS antennas results in a favourable channel. The favourable channel results from the law of large numbers [15]. Higher wireless throughput is obtainable with an increase in user number and base station antenna arrays whilst simple linear processing delivers the array and multiplexing gains [20]. If antenna arrays are raised in base stations, the sum rate obtained via linear processing nears the sum rate deduced from optimal receivers.

2.2.7 Challenges of Massive MIMO

Even though massive MIMO is aimed at achieving better capacity that allows many users to be served simultaneously while maintaining a high throughput, the system is faced with several challenges as discussed below. Energy efficiency and simpler linear signal processing are some of the strong points of the massive MIMO system; however, the system faces challenges related to these benefits that need to be tackled. These challenges include pilot contamination, unfavourable propagation, channel reciprocity, precoding techniques, signal detection and channel estimation. Additionally the system is faced with low mutual coupling, compact and

low profile, correlation coefficient and channel characterization problems [21]. Some of these challenges are discussed below.

2.2.7.1 Pilot Contamination

Practical wireless network systems use multiple cells that share time-frequency allocation due to the scarcity of these resources. Since the channel coherence interval is limited, it becomes impractical to give all the cells orthogonal pilot sequences in a multicell system. As such the pilot sequences are usually reused. Due to the reuse, contamination of the channel estimate will occur as a result of the pilot sequences that have been released by other users [20]. When contamination occurs, the systems performance is significantly reduced. The problem is persistent even when the antenna arrays of the BS are significantly increased and therefore seriously hampers the massive MIMO technology. Researchers have put considerable effort to eliminate this problem through methods such as channel estimation using the eigenvalue decomposition method where both contamination and decontamination methods are put forward. Through this method, channel pilot contamination can be controlled by employing a pilot allocation scheme that is covariance-aware [13]. However, there is much to do to mitigate this challenge and research in this area is still ongoing.

2.2.7.2 Unfavorable Propagation

Massive MIMO is designed to work in favourable environments of propagation which is rarely the case in practice. In a practical environment where the massive MIMO system is likely to be deployed, the number of users may be higher as compared to the number of scatterers for the signal or even the user channels directed to the BS may be sharing scatterers leading to unfavourable propagation environments.

2.2.7.3 Physical Array Size and Antenna Spacing

Depending on carrier frequency, the area to be covered, and the nature of the use, there are different shapes and sizes of the antenna arrays. In cellular communications, the most popular array models are the linear, the planar, and the cylindrical arrays (see Fig 7.11 in [23]). Relative to the wavelength, the important factors of an antenna array are the antenna spacing and its total size [23]. Massive MIMO systems with a large number of antennas at the BS have been shown to produce high spectral and energy efficiencies under the assumptions of increasing the physical space between the antennas at the BS, i.e., there is no spatial correlation among

antennas at the BS [24, 25]. However, in practical systems it is obvious that as you move to a higher number of antennas at the BS, antennas would have to be placed much closer together thereby increasing the spatial correlation and the array size.

In massive MIMO systems, favourable propagation makes fading channel behave as if it was deterministic channel, and hence the assumptions of uncorrelated channel make the performance analysis and optimisation of massive MIMO tractable. On the other hand, in spatially correlated systems, the channels between different antennas are often correlated and therefore the potential of multi antenna gains and favourable propagation may not always be obtainable. Therefore, it is observed that antenna correlation does have great impact on the bit error rate (BER) performance of a massive MIMO system. A possible solution to this problem is locating the BS antennas in a distributed manner over a wide area to minimise sharing of scatterers among users [14]. In this thesis, it is assumed that there is no correlation between transmit antennas at the BS as they are spaced at a minimum of 0.5λ, where λ is the wavelength.

2.2.7.4 Channel Reciprocity

Reciprocity in massive MIMO systems refers to the channel impulse from the BS to the mobile terminal being the same as that of the user terminal towards the BS [14]. The technique is founded on the TDD principle but most of the equipment used in the mobile terminal does not have hardware that can support reciprocity for the downlink and uplink channels. There have been propositions to solve this issue which includes calibration of the base station's hardware chains. One such method is multiuser beamforming where a base station with more than ten antennas serves several terminal antennas at the same time. The technique employs internal calibration of the hardware to allow for implicit beamforming that enables it to attain baseband processing capabilities. In a massive MIMO system, terminal hardware does not need to be calibrated for uplink and downlink reciprocity. Instead, the BS's hardware is calibrated to give beams that are coherent with the terminal equipment hardware channels [15]. To do this, one antenna is taken as the reference for the others and then the antennas share information amongst them and hence are able to compute compensation factors required for each antenna. An alternative method is to lower the gain slightly and then totally omit calibration and the two ends operate well.

2.2.7.5 Channel Estimation

For the downlink part of the channel estimation process of the coherence interval, a challenge is to achieve maximum usage, orthogonal pilot signals are required due to the huge number of antennas propagating in a mobile environment [20]. The resource requirements, and in particular frequency and time, rise as the quantity of antenna arrays are increased. Another problem is the increase of responses to the base station when the BS antennas increase, leading to a strain on the uplink resources.

2.2.7.6 Precoding Technique

The downlink channel requires the development of pre-coding techniques that act to improve the performance of a massive MIMO system. A possible solution is the use of Hermitian methods which are linear and more practical [14]. According to the review article [20], linear and non-linear precoding technique can be used for massive MIMO in the BS. In comparison between linear and non-linear precoding techniques, the linear precoding techniques such as matched filter (MF) and zero-forcing (ZF) are near-optimal and thus are less complex than non-linear precoding techniques. On the other hand, non-linear precoding techniques such as dirty-paper-coding (DPC), vector perturbation (VP) and lattice-aided methods, have better performance but with higher complexity.

When the CSI is perfectly known to the BS, i.e. $\boldsymbol{H}$ is known, the linear precoding techniques are more practical. When MF is used, the transmit signal can be written as

$$\boldsymbol{A}^{MF} = \frac{1}{\sqrt{\alpha}}(\boldsymbol{H}) \tag{2.14}$$

where $\boldsymbol{A}$ is an $N_r \times N_t$ precoded transmit signal of the data symbols $\boldsymbol{s}$, and α normalizes the average transmit power. When ZF is used, the transmit signal from the BS can be written as

$$\boldsymbol{A}^{RZF} = \frac{1}{\sqrt{\alpha}}\boldsymbol{H}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H})^{-1}. \tag{2.15}$$

The MF precoder has better performance than ZF in the low spectral efficiency region, while the ZF is better in the high spectral efficiency [20]. When regularized ZF (RZF) is used, the transmit signal from the BS can be written as

$$\boldsymbol{A}^{RZF} = \frac{1}{\sqrt{\alpha}}\boldsymbol{H}(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H} + \delta\boldsymbol{I})^{-1}, \tag{2.16}$$

where δ is a parameter subject to optimisation. Setting $\delta = 0$ results in the ZF precoder while $\delta \to \infty$ gives the MF precoder.

2.2.7.7 Signal Detection

It is desirable to have a simple linear processing technique that can handle data in large volumes that normally results from multiple antenna arrays and such mechanisms are yet to be fully developed [21]. Several methods have been put forward to tackle the issue, including minimum mean-square error (MMSE) and MF precoding/detection. Another linear method that has been proposed is ZF. MMSE is able to obtain a performance similar to those of MF and ZF, albeit with a lesser number of antennas [20]. Non-linear alternatives such as lattice reduction, Tabu search and likelihood ascent search give improved performance but have a disadvantage of high complexity which is not desirable. When the BS knows the CSI, linear detection can be considered. By using a linear detector, the received signal in (2.9) is separated into streams by multiplying it with $\boldsymbol{A}^{\mathrm{H}}$ as follows

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{A}^{\mathrm{H}}\boldsymbol{y}. \\ &= \sqrt{\rho}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{H}s + \boldsymbol{A}^{\mathrm{H}}\boldsymbol{z}. \end{aligned} \tag{2.17}$$

Let r_k and s_k be the elements of the kth user of the $N_t \times 1$ vectors $\boldsymbol{r}$ and $\boldsymbol{s}$, respectively. Then the *downlink* received signal for the kth user is given by

$$r_k = \sqrt{\rho_k}\, \boldsymbol{a}_k^{\mathrm{H}}\boldsymbol{h}_k s_k + \sum_{\substack{q=1 \\ q\neq i}}^{K} \sqrt{\rho_q}\boldsymbol{a}_q^{\mathrm{H}}\boldsymbol{h}_k s_q + \boldsymbol{a}_q^{\mathrm{H}} z_k, \tag{2.18}$$

where $\boldsymbol{a}_k$ and $\boldsymbol{h}_k$ are the vectors of the kth user of the matrices $\boldsymbol{A}$ and $\boldsymbol{H}$. While the *uplink* received signal from the kth user is given by

$$r_k = \sqrt{\rho_k}\, \boldsymbol{a}_k^{\mathrm{H}}\boldsymbol{h}_k s_k + \sum_{\substack{q=1 \\ q\neq i}}^{K} \sqrt{\rho_k}\boldsymbol{a}_k^{\mathrm{H}}\boldsymbol{h}_q s_k + \boldsymbol{a}_k^{\mathrm{H}} z. \tag{2.19}$$

2.3 Differential Space-Time Modulation

The MU-MIMO systems that I have reviewed in the previous sections assume coherent detection at the receiver, i.e., the receiver knows the channel $\boldsymbol{H}$. In practical systems, the CSI estimation is carried out by using pilots sequences transmitted with the data that are known

to the receiver. The accuracy of the CSI estimation depends on the number of pilots symbols used in the estimation process. The higher the number of pilots the higher the accuracy of CSI estimation. However, for some systems, due to high mobility and the cost of channel training and estimation, CSI acquisition is impossible [26]. Furthermore, the resource requirements, and in particular frequency and time, rise as the quantity of antenna arrays are increased. Another problem is the increase of responses to the base station when the BS antennas increase, leading to a strain on the uplink resources. One alternative method is to encode the transmitted data differentially, and to decode differentially using differential space-time modulation (DSTM) without any knowledge of the CSI at the receiver. The receiver in this case does not need any knowledge of the CSI to perform detection. Consequently, differential detection does not employ pilot symbol transmission so that it helps keep the receiver structure as simple as possible. Figures 2.4 (a) and (b) show the system configurations of the coherent and noncoherent schemes, respectively. It can be seen that the coherent setup involves a channel estimation and feedback stage to acquire the CSI, while the noncoherent scheme does not require either the transmitter or the receiver to do any channel estimation process to know the channel. In this section, details of differential space-time modulation for SISO and MIMO systems are described, including their encoding, decoding, code construction, and performance analysis.

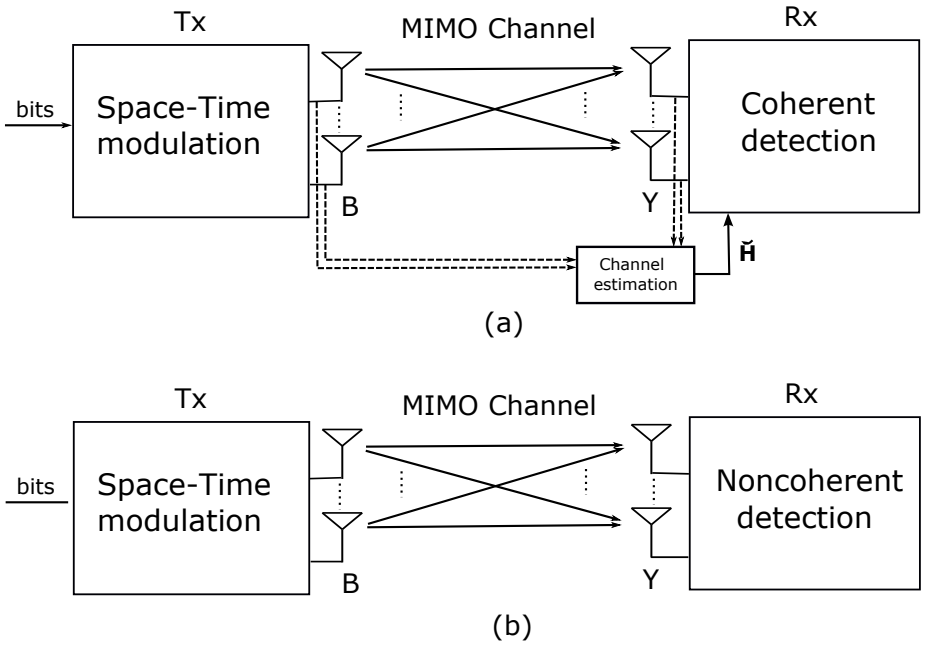

Figure 2.4. MIMO system, (a) with coherent detection, (b) with noncoherent detection

2.3.1 Differential Space-Time Modulation for SISO System

The vector $\boldsymbol{s} = [s_1, s_2, \cdots, s_\tau \cdots, s_N] \in \mathbb{C}^{1\times N}$, is the information vector with elements drawn from an M-ary PSK constellation as:

$$\mathcal{M} = \left\{ e^{j2\pi i/M} \;\middle|\; i = 0, 1, \cdots, M-1 \right\}, \tag{2.20}$$

where N denotes the block length of the coherence time intervals. The PSK modulator group

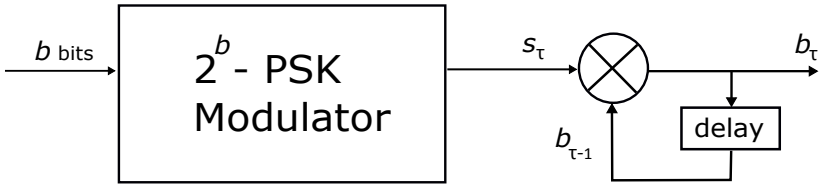

Figure 2.5. DPSK modulator.

every b bits at its input and maps them onto one of the constellation information symbols. It is assumed $\mathbb{E}\{|s_\tau|^2\} = 1$. As shown in Fig. 2.5, a sequence of symbols of the s_τ at time τ, is then differentially encoded to generate the differential PSK (DPSK) modulated signal vector $\boldsymbol{b} \in \mathbb{C}^{1\times(N+1)}$ via the rule

$$b_\tau = b_{\tau-1} s_\tau = b_0 \prod_{i=1}^{\tau} s_i. \tag{2.21}$$

The transmit information signal vector $\boldsymbol{b}$ comprises the initial reference symbol $b_0 = 1$ that is known to the receiver and the following N differentially encoded symbols in the form of $\boldsymbol{b} = [b_0, b_1, \cdots, b_N]$. In flat fading propagation, the received signal at time τ is given by

$$y_\tau = \sqrt{\rho} h b_\tau + z_\tau, \tag{2.22}$$

where h is the fading coefficient between the transmit and receive antennas which is modelled as a complex Gaussian random variable with zero mean and unit variance. z_τ is the noise sample at time τ which is modelled as a complex Gaussian random variable with zero mean and unit variance. It is assumed that the channel h changes slowly (channel coherence time is large enough) and extends over several matrix transmission periods. In such a case, the BS transmission starts with a reference symbol $b_0 = 1$, followed by several information symbols. When encoding using (2.21), the decoding process for s_τ would be according to the last two received signal of y, i.e., y_τ and $y_{\tau-1}$. Thus, to recover s_τ, the receiver computes $y_\tau\ y^*_{\tau-1}$ as

follows

$$y_\tau y_{\tau-1}^* = (\sqrt{\rho}hb_\tau + z_\tau)(\sqrt{\rho}h^* b_{\tau-1}^* + z_{\tau-1}^*) \tag{2.23}$$

$$= |h|^2(\rho b_\tau b_{\tau-1}^*) + \sqrt{\rho}hb_\tau z_{\tau-1}^* + \sqrt{\rho}h^* b_{\tau-1}^* z_\tau + z_\tau z_{\tau-1}^* \tag{2.24}$$

$$= \rho|h|^2|b_{\tau-1}|^2 s_\tau + \sqrt{\rho}hb_\tau z_{\tau-1}^* + \sqrt{\rho}h^* b_{\tau-1}^* z_\tau \tag{2.25}$$

$$= \rho|h|^2 s_\tau + \sqrt{\rho}hb_\tau z_{\tau-1}^* + \sqrt{\rho}h^* b_{\tau-1}^* z_\tau \tag{2.26}$$

In (2.23), the channel h is assumed to be constant over two consecutive symbol intervals. The last two terms in (2.26) can be considered as a noise term with zero mean and variance $2|h|^2\rho$. As the channel h is unknown to the receiver, the differential receivers recovers s_τ according to the decision rule [27]

$$\hat{s}_\tau = \arg \min_{s_\tau \in \mathcal{M}} |y_\tau y_{\tau-1}^* - s_\tau|. \tag{2.27}$$

On the other hand, as a comparison, the coherent detection when h is known to the receiver first computes $h^* y_\tau$ as follows

$$h^* y_\tau = \sqrt{\rho}|h|^2 b_\tau + h^* z_\tau. \tag{2.28}$$

Note that the noise variance in (2.28) is $|h|^2$. Therefore, the optimal receiver would be [27]

$$\hat{s}_\tau = \arg \min_{b_\tau \in \mathcal{M}} |h^* y_\tau - \sqrt{\rho}|h|^2 b_\tau|^2. \tag{2.29}$$

The difference in performance between coherent detection and differential detection can be evaluated using SNR in both cases. From (2.26) and (2.28), the received SNR for differential detection is $|h|^2\rho/2$, while the received SNR for coherent detection is $|h|^2\rho$. By comparing the two cases, it is clear that coherent detection outperforms the differential detection by approximately 3 dB [27].

2.3.2 Differential Space-Time Modulation for MIMO System

In this subsection, the DSTM system design is extended to the case of MIMO systems. I will consider here the DSTM model designed by [28, 29]. The model designed by [28, 29] is of interest because it can be used for any number of transmit and receive antennas and applied to any constellation.

2.3.2.1 System Model and Signal Structure

Consider a SU-MIMO downlink broadcast channel where the BS transmits multiple streams to a single user with multiple receive antennas as shown in Fig. 2.1-(a). The BS has N_t transmit antennas and the receiver has N_r receive antennas.

The channel matrix $\boldsymbol{H} \in \mathbb{C}^{N_r \times N_t}$ is assumed to be a Rayleigh flat fading MIMO channel, where the element $h_{i,j}$ is the channel coefficient between the jth transmit antenna and the ith receive antenna during coherence interval. The elements of $\boldsymbol{H}$ are samples of i.i.d. complex Gaussian process with distribution $\mathcal{CN}(0,1)$.

Let $\boldsymbol{S} \in \mathbb{C}^{T \times T}$ be a transmitted signal, satisfying the following condition:

$$\boldsymbol{S}^H \boldsymbol{S} = \boldsymbol{S}\boldsymbol{S}^H = \boldsymbol{I}_T\,, \quad \forall\, \boldsymbol{S} \in \mathbb{G} \tag{2.30}$$

where, $\mathbb{G}$ is any $T \times T$ unitary group matrices that can be selected from any constellation ($\mathcal{C}$). Let $\boldsymbol{D}$ be $N_t \times T$ matrix such that $\boldsymbol{DS} \in \mathbb{C}^{N_t \times T}$ for all $\boldsymbol{S} \in \mathbb{G}$. The collection of the group matrices can be represented as

$$\mathbb{M} \stackrel{\text{def}}{=} \{\boldsymbol{DS} : \boldsymbol{S} \in \mathbb{G}\}\,, \tag{2.31}$$

where $\mathbb{M}$ represents a multichannel group code of length T over the constellation $\mathcal{C}$. To make $\mathbb{M}$ unitary, $\boldsymbol{D}$ should be unitary as well, i.e, $\boldsymbol{D}^H\boldsymbol{D} = N_t\boldsymbol{I}_T$. Thus, to construct $\mathbb{M}$, $\boldsymbol{D}$ can be chosen to be in $\mathcal{C}$ and $\mathbb{G}$ to be any group of $T \times T$ permutation matrices. Let M denote the cardinality of the set $\mathbb{G}$. Therefore, the rate of the code $R = \frac{\log_2 M}{T}\ \frac{\text{bits}}{\text{s}}$. For an M-ary PSK, $\mathbb{G}$ group code can be constructed such that [28]

$$\mathbb{G} = \left\{1, \Phi_M, \Phi_M^2, \cdots, \Phi_M^{M-1}\right\}, \tag{2.32}$$

where $\Phi_M = e^{2\pi j/M}$. The group codes are either cyclic or dicyclic. The group (M,i) is cyclic if i is odd, thus the group can be generated as follows [28]

$$\mathbb{G} = \left\langle \begin{bmatrix} \Phi_M & 0 \\ 0 & \Phi_M^i \end{bmatrix} \right\rangle\,, \qquad \forall\, 0 < i < M, \quad i \text{ is odd.} \tag{2.33}$$

On the other hand, the group (M,i) is dicyclic if i is even and $M \geq 8$, then it can be generated as follows [28]

$$\mathbb{G} = \left\langle \begin{bmatrix} \Phi_{M/2} & 0 \\ 0 & \Phi_{M/2}^* \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right\rangle \qquad \forall M \geq 8, \;\; i \text{ is even.} \tag{2.34}$$

The information symbols are independent of each other which will permit the use of a wider variety of unitary matrix groups. The following examples can illustrate the idea of the group codes.

Example 1. ***Cyclic group***. *For $N_t = T = 2$, $R = 1$, $M = 4$, $i = 3$, the pair [28]*

$$\mathbb{G} = \left\{ \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \pm \begin{bmatrix} j & 0 \\ 0 & -j \end{bmatrix} \right\}$$
$$\boldsymbol{D} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

is a group code over the QPSK constellation $\mathcal{C} = \{1, j, -1, -j\}$.

Example 2. ***Dicyclic group***. *For $N_t = T = 2$, $R = 1.5$, $M = 8$, $i = 4$, the pair [28]*

$$\mathbb{G} = \left\{ \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \pm \begin{bmatrix} j & 0 \\ 0 & -j \end{bmatrix}, \pm \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \pm \begin{bmatrix} 0 & -j \\ j & 0 \end{bmatrix} \right\}$$
$$\boldsymbol{D} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

is a group code over the QPSK constellation $\mathcal{C} = \{1, j, -1, -j\}$.

Example 3. ***Dicyclic group***. *For $N_t = T = 2$, $R = 2$, $M = 16$, $i = 8$, the pair [28]*

$$\mathbb{G} = \left\{ \pm \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \pm \begin{bmatrix} \frac{1+j}{\sqrt{2}} & 0 \\ 0 & \frac{1-j}{\sqrt{2}} \end{bmatrix}, \pm \begin{bmatrix} j & 0 \\ 0 & -j \end{bmatrix}, \pm \begin{bmatrix} \frac{-1+j}{\sqrt{2}} & 0 \\ 0 & \frac{-1-j}{\sqrt{2}} \end{bmatrix}, \right.$$
$$\left. \pm \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \pm \begin{bmatrix} 0 & \frac{-1+j}{\sqrt{2}} \\ \frac{1+j}{\sqrt{2}} & 0 \end{bmatrix}, \pm \begin{bmatrix} 0 & j \\ 0 & j \end{bmatrix}, \pm \begin{bmatrix} 0 & \frac{1+j}{\sqrt{2}} \\ \frac{-1+j}{\sqrt{2}} & 0 \end{bmatrix} \right\}$$
$$\boldsymbol{D} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

is a group code over the 8PSK constellation $\mathcal{C} = \left\{1, \frac{1+j}{\sqrt{2}}, j, \frac{-1+j}{\sqrt{2}}, -1, \frac{-1-j}{\sqrt{2}}, -j, \frac{1-j}{\sqrt{2}}\right\}$.

2.3.2.2 Differential Encoding for MIMO

Differential encoding can be performed in two ways: overlap and non-overlap. In the overlap case, the channel coherence time changes slowly (e.g. the channel stays constant over several

transmission blocks) [28]. Clearly, as a result, the transmission starts with a reference matrix code, and is then followed by several information blocks.

In the non-overlap case, the channel coherence time is fixed over just two consecutive blocks [28]. Consequently, the transmission starts by transmitting a known reference code matrix at the beginning of each new coherence time, followed by just one information block.

i- Overlap Case

The transmitted signal matrix is encoded differentially as follows [28]:

$$\boldsymbol{B}_\tau = \boldsymbol{B}_{\tau-1}\boldsymbol{S}_\tau = \boldsymbol{B}_0 \prod_{i=1}^{\tau} \boldsymbol{S}_i, \qquad \tau = 1, ..., N\ , \tag{2.35}$$

where $\boldsymbol{B}_\tau$ is a unitary differentially encoded data matrix, N denotes the block length of the coherence time intervals. Note that $\boldsymbol{B}_0 = \boldsymbol{D}$ to initiate the transmission, where $\boldsymbol{D}$ is an $N_t \times T$ matrix such that

$$\boldsymbol{D} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \tag{2.36}$$

ii- Non-Overlap Case

According to [28], the transmitted symbols can also be encoded using the non-overlap scheme as follows

$$\boldsymbol{B}_\tau = \boldsymbol{D}\boldsymbol{S}_\tau\ , \qquad \tau = 1, ..., N. \tag{2.37}$$

This type of encoding is simple, however, its spectral efficiency is almost half that of the overlap case.

2.3.2.3 Differential Detection for MIMO

In DSTM, the transmissions are implemented in blocks, i.e., the transmitted space-time signal is a $T \times T$ matrix. For simplicity, $N_t = T$ is assumed as in [28]. Let $\bar{\boldsymbol{Y}}$ denote the matrix having all $N + 1$ received signal blocks, i.e.,

$$\bar{\boldsymbol{Y}} = [\boldsymbol{Y}_0\ \ \boldsymbol{Y}_1\ \ \cdots, \boldsymbol{Y}_\tau, \cdots\ \ \boldsymbol{Y}_N]\,, \tag{2.38}$$

and, $\boldsymbol{Y}_\tau$ can be expressed as

$$\boldsymbol{Y}_\tau = \sqrt{\rho}\,\boldsymbol{H}\boldsymbol{B}_\tau + \boldsymbol{Z}_\tau, \qquad \tau = 0, 1, ..., N. \tag{2.39}$$

The channel coefficients $\boldsymbol{H}$ change slowly, so that the channel coherence time extends over several matrix transmission. In such a scheme, the BS transmission starts with a reference matrix, followed by several information matrices. Now, the encoding process will be according to (2.35), however for simplicity, the decoding process for $\boldsymbol{S}_\tau$ would not be based on the entire received sequence of $\bar{\boldsymbol{Y}}$, instead it would be based on every two consecutive/adjacent blocks of $\bar{\boldsymbol{Y}}$ as in the following notation [28]

$$\bar{\boldsymbol{Y}} = \left[\underbrace{\boldsymbol{Y}_0\boldsymbol{Y}_1}\ \underbrace{\boldsymbol{Y}_1\boldsymbol{Y}_2}\cdots\underbrace{\boldsymbol{Y}_{\tau-1}\boldsymbol{Y}_\tau}\cdots\underbrace{\boldsymbol{Y}_{N-1}\boldsymbol{Y}_N}\right]. \tag{2.40}$$

In other words, the last two blocks of $\bar{\boldsymbol{Y}}$ can be defined as

$$\bar{\boldsymbol{Y}}_\tau \overset{\text{def}}{=} \begin{bmatrix}\boldsymbol{Y}_{\tau-1} & \boldsymbol{Y}_\tau\end{bmatrix}. \tag{2.41}$$

The code matrices that affected $\boldsymbol{Y}_\tau$ are

$$\bar{\boldsymbol{C}}_\tau = \begin{bmatrix}\boldsymbol{B}_{\tau-1} & \boldsymbol{B}_{\tau-1}\boldsymbol{S}_\tau\end{bmatrix}. \tag{2.42}$$

The differential data words matrices $\bar{\boldsymbol{C}}_\tau\bar{\boldsymbol{C}}_\tau^H$ can be expressed as

$$\bar{\boldsymbol{C}}_\tau^H\bar{\boldsymbol{C}}_\tau = \begin{bmatrix} N_t\boldsymbol{I}_{N_t} & N_t\boldsymbol{S}_\tau^H \\ N_t\boldsymbol{S}_\tau & N_t\boldsymbol{I}_{N_t}\end{bmatrix}, \tag{2.43}$$

where $N_t = T$, $\boldsymbol{S}_\tau\boldsymbol{S}_\tau^H = \boldsymbol{S}_\tau^H\boldsymbol{S}_\tau = \boldsymbol{I}_T$, $\boldsymbol{D}_\tau\boldsymbol{D}_\tau^H = \boldsymbol{D}_\tau^H\boldsymbol{D}_\tau = T\boldsymbol{I}_{N_t} = N_t\boldsymbol{I}_T$, $\boldsymbol{B}_0 = \boldsymbol{D}$, and $\boldsymbol{B}_\tau\boldsymbol{B}_\tau^H = \boldsymbol{B}_\tau^H\boldsymbol{B}_\tau = N_t\boldsymbol{I}_{N_t}$. When $\boldsymbol{B}_{\tau-1}$ is known to the receiver, the optimal differential decoder to decode $\boldsymbol{S}_\tau$ would be as [28, 29]

$$\hat{\mathbf{S}}_\tau = \arg\max_{\boldsymbol{S}_\tau\in\mathbb{G}}\ \text{trace}\left\{\bar{\boldsymbol{Y}}_\tau\bar{\boldsymbol{C}}_\tau^H\bar{\boldsymbol{C}}_\tau\bar{\boldsymbol{Y}}_\tau^H\right\} \tag{2.44}$$

$$= \arg\max_{\boldsymbol{S}_\tau\in\mathbb{G}}\ \Re\left\{\text{trace}\left\{\boldsymbol{Y}_{\tau-1}\boldsymbol{S}_\tau\,\boldsymbol{Y}_\tau^H\right\}\right\}, \tag{2.45}$$

$$= \arg\max_{\boldsymbol{S}_\tau\in\mathbb{G}}\ \Re\left\{\text{trace}\left\{\boldsymbol{S}_\tau\,\boldsymbol{Y}_\tau^H\boldsymbol{Y}_{\tau-1}\right\}\right\}, \tag{2.46}$$

and when $\boldsymbol{H}$ is available at the receiver, coherent detection can be exploited, in this scheme decoder is given by

$$\hat{\mathbf{S}}_\tau = \arg \max_{\boldsymbol{S}_\tau \in \mathbb{G}} \Re \left\{ \text{trace} \left\{ \boldsymbol{H}\ \boldsymbol{B}_{\tau-1}\ \boldsymbol{S}_\tau\ \boldsymbol{Y}_\tau^H \right\} \right\}. \tag{2.47}$$

2.3.2.4 Capacity of Noncoherent MIMO Channels

In MIMO noncoherent downlink transmission in which the receiver has no prior knowledge of the channel. The length of the channel coherence time, T, determines the upper bound on the downlink capacity. Let $\boldsymbol{y}_l$ denote the lth row of $\boldsymbol{Y}$, defined in (2.39) , which can be expressed as

$$\boldsymbol{y}_l = \sqrt{\rho}\boldsymbol{h}_l\boldsymbol{B} + \boldsymbol{z}_l, \tag{2.48}$$

where $\boldsymbol{h}_l$ and $\boldsymbol{z}_l$ denote the lth row of $\boldsymbol{H}$ and lth row of $\boldsymbol{Z}$, respectively. Therefore, the covariance matrix is given as

$$\boldsymbol{\Lambda} = \mathbb{E}\left\{ \boldsymbol{h}_l^H \boldsymbol{h}_l \right\} \tag{2.49}$$

$$= \boldsymbol{I} + \rho \boldsymbol{B}^H \boldsymbol{B}. \tag{2.50}$$

The probability density function of $\boldsymbol{Y}$ conditioned on $\boldsymbol{B}$ is given as [29, 30]

$$P(\boldsymbol{Y}|\boldsymbol{B}) = \frac{\exp\left(-\text{trace}\left\{ \boldsymbol{\Lambda}^{-1} \boldsymbol{Y}^H \boldsymbol{Y} \right\} \right)}{\pi^{TN_r} \left(\det(\boldsymbol{\Lambda}) \right)^{N_r}} \tag{2.51}$$

where $\det(\cdot)$ denotes the determinant of a matrix. The mutual information of the channel entropy is given as

$$I(\boldsymbol{Y}; \boldsymbol{B}) = H(\boldsymbol{Y}) - H(\boldsymbol{Y}|\boldsymbol{B}), \tag{2.52}$$

where $H(\boldsymbol{Y})$ is the entropy of $\boldsymbol{Y}$ and $H(\boldsymbol{Y}|\boldsymbol{B})$. In order to obtain the capacity of noncoherent channel, the general entropy form in (2.52) is maximised such that [17]

$$C = \frac{1}{T} \sup_{P(\boldsymbol{B})} I(\boldsymbol{Y}; \boldsymbol{B}) \tag{2.53}$$

where $P(\boldsymbol{B})$ is the joint p.d.f of $P(\boldsymbol{B})$. To solve this optimisation problem in (2.53), the authors in [16] have found that the capacity of the differential MIMO system can be geometrically interpreted as sphere packing in the Grassmann manifold. The Grassmann manifold is defined as follows: given a Stiefel Manifold $\mathbb{S}(M_1, M_2)$ which represents the space of all M_1-by-M_2 complex unitary matrices, then, the Grassmann manifold $\mathbb{G}(M_1, M_2)$ is the space of the ratio of

the Stiefel Manifold $\mathbb{S}(M_1, M_2)$ to the Stiefel Manifold $\mathbb{S}(M_2, M_2)$[1] [16].

Define $\varsigma \stackrel{\text{def}}{=} \min(N_t, N_r, \lfloor T/2 \rfloor)$, where $\lfloor x \rfloor$ denotes the integer part of x. For the coherence time T, the capacity of the noncoherent channel at high average SNR regime, which I denoted by C_{noncoh} is [16]

$$C_{noncoh} = \varsigma \left(1 - \frac{\varsigma}{T}\right) \log_2 \text{SNR} + c, \qquad \frac{\frac{\text{bits}}{\text{s}}}{\text{Hz}} \tag{2.54}$$

where

$$\begin{aligned} c =& \frac{1}{T} \log_2 \left|\mathbb{G}\left(T, \varsigma\right)\right| + \varsigma \left(1 - \frac{\varsigma}{T}\right) \log_2 \frac{T}{\varsigma \pi e} \\ &+ \left(1 - \frac{\varsigma}{T}\right) \mathbb{E}\left[\log_2 \det(\boldsymbol{H}\boldsymbol{H}^H)\right], \end{aligned} \tag{2.55}$$

and $|\mathbb{G}(T, N)|$ is a natural measure on the Grassmann manifold defined as [16]

$$|\mathbb{G}(T, \varsigma)| = \frac{\prod_{i=T-\varsigma+1}^{T} \frac{2\pi^i}{(i-1)!}}{\prod_{i=1}^{\varsigma} \frac{2\pi^i}{(i-1)!}}. \tag{2.56}$$

Note that when $T \to \infty$ in (2.54), the capacity for the noncoherent MIMO channels approaches that of coherent MIMO channels case. Fig. 2.6 serves as an example of the capacity achieved when the channel $\boldsymbol{H}$ is known for coherent system and when it is unknown for noncoherent system. The SNR is chosen to be high for an $N_t = N_r = 8$ MIMO system. Three values of T are considered, $T = 25$, $T = 50$, and $T = 100$.

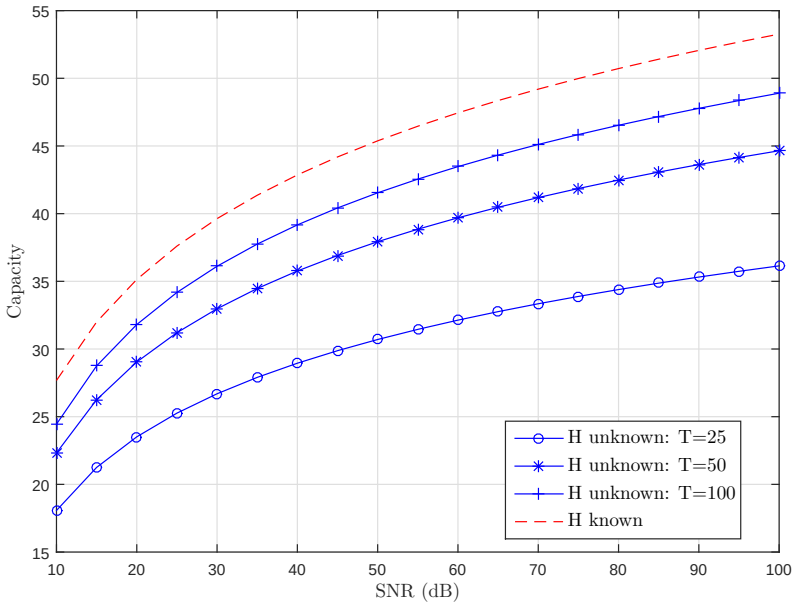

Figure 2.6. Capacity comparisons of an $N_t = N_r = 8$ MIMO channel for the coherent and noncoherent cases with $T = 25$, $T = 50$, and $T = 100$.

[1]For more details regarding the Stiefel and Grassmann manifolds, the reader is recommended to see [16].

2.4 Focused Literature Review

This section reviews the previous literature work related to the design of noncoherent downlink transmission systems. The review covers the design in three different categories of MIMO systems (which specify the topics of the next three chapters): low rate differential STBC in a downlink MU-MIMO system; high rate differential algebraic STBC in a downlink MU-MIMO system; and differential downlink transmission in a massive MU-MIMO system.

2.4.1 Low Rate Differential STBC in a Downlink MU-MIMO System

For the MU-MIMO downlink, the availability of channel state information at the transmitter makes it possible for the precoder to precancel the CCI at each user. The authors in [6] proposed a framework that uses BD to cancel the CCI and assumed full CSI knowledge at the transmitter. The CSI between the transmitter and the receivers is estimated at the receivers then fed back to the transmitter. This leads to increased complexity of the receivers. In [7], the authors proposed a method that combines the precoding technique in [6] and the Alamouti STBC. The proposed method provides a substantial gain in terms of spatial diversity with a low decoding complexity. However, for the decoding process, each receiver still needs to know the composite channel formed by the precoder and the channel in order to coherently decode the Alamouti STBC. In practice, each receiver acquires the composite channel by direct estimation. The prior focus of STBC MU-MIMO downlink transmission techniques has been on cases where CSI is available at the receivers and transmitter. However, for some systems, due to high mobility and the cost of channel training and estimation, CSI acquisition is impossible [26].

Therefore, designing and implementing a framework that requires neither feedback nor the estimation of the composite channels has been an unresolved topic.

2.4.2 High Rate Differential Algebraic STBC in a Downlink MU-MIMO System

The transmission of an orthogonal STBC over a MIMO channel in [1, 31] was proposed to achieve full diversity with a low complexity receiver. However, orthogonal STBC suffers from an inability to work with a greater number of antennas at full transmission rates. When decoding complexity is not an issue, non-orthogonal full-rate full-diversity algebraic STBC may be used [32, 33]. For MIMO systems, there are many previous employed space-time codes that provide a higher rate with full diversity in a trade-off with complexity, such as threaded algebraic space-

time (TAST) block codes, the classic Bell Laboratories layered space-time (V-BLAST) and linear dispersion block codes (LDC) [34, 35]. However, the minimum determinants of these codes are generally non-zero, but vanish as the spectral efficiency of the signal constellation is increased. The authors in [36] have constructed full rate and full diversity perfect algebraic STBC with a non-vanishing determinant when the spectral efficiency increases.

In the MU-MIMO downlink, transmit diversity can be applied using downlink transmission techniques, such as the orthogonal spreading multiplexing code. The authors in [7, 37] used this technique to decompose the MU-MIMO channels into parallel single user non-interfering channels, and hence CCI was eliminated. Implementing the orthogonal spreading technique at the transmitter (e.g. a base station) helps maintain simplicity in the receiver, so that simple linear decoding approaches are applicable at the receiving end (e.g. end users). In a coherent scenario, this approach was later considered in [38] as a multiplexing scheme for a MU-MIMO system, and was combined with full rate full diversity algebraic STBC. The proposed method cancels the CCI and provides a substantial gain in terms of full rate and spatial diversity. However, for the decoding process, each receiver still needs to know the CSI to coherently decode the algebraic STBC. In practice, each receiver acquires the composite channel by direct estimation, which leads to increased complexity of the receivers.

The prior focus of noncoherent MU-MIMO downlink transmission techniques assumed CSI is available at the receivers and transmitter with low data rate. However, it is not always feasible or advantageous to adopt channel estimation based schemes, essentially for high mobility system or when many antennas are used [26]. One way to get high rates noncoherent scheme is to encode the transmitted data differentially using a CD transform, and to decode differentially without any knowledge of the CSI at the receiver [18]. Our previous work in [39] has dealt with implementing the MU-MIMO downlink transmission of an Alamouti STBC combined with differential modulation, which does not require channel knowledge for decoding. The scheme provides low complexity transceivers while providing good performance. However, our work in [39] cannot provide a comprehensive high rate differential scheme in downlink scenario.

Hence, there is much interest in designing a high data rate noncoherent system that does not need CSI at either the transmitter or receiver for downlink MU-MIMO system.

2.4.3 Differential Downlink Transmission in a Massive MU-MIMO System

Much of the research on MIMO downlink transmission designs assumes perfect CSI at the transmitter. The availability of CSI at both ends makes it possible for the system to eliminate the multiple access interference (MAI) between users. However, due to various reasons, such as pilot contamination from training sequence reuse in massive MIMO, perfect CSI estimation is unattainable [26]. In [7], the authors proposed a framework that uses the block diagonalisation method to cancel MAI between users. The proposed method provides a substantial gain in terms of spatial diversity with a low decoding complexity. However, for the decoding process, each user still needs to know the channel in order to decode the information signal coherently.

Our work in [39] proposed a downlink spreading scheme combined with DM detection to eliminate the need of estimating the CSI at the BS and users. The scheme provides both low complexity transceivers and good performance. However, for large number of users, the proposed scheme in [39] does not provide a comprehensive high rate differential scheme in a downlink scenario due to the long length of the spreading code. Our work in [40] proposed a full rate downlink algebraic transmission scheme combined with a differential space-time scheme. The proposed scheme provides a full-rate full-diversity system and does not require any knowledge of the CSI to separate the data streams of multiple users. In this approach, however, the BS typically employs only a few antennas, and thus the corresponding improvement in spectral efficiency and system simplicity is still relatively modest.

In order to improve the spectral efficiency and to simplify the required signal processing, a massive MIMO downlink system is employed [41, 42], where the BS is equipped with a very large number of transmit antennas. In practice, the demodulation reference signals (DM-RS) are used to support channel estimation and data demodulation. In DM-RS, the estimation of the channel for coherent detection is often obtained by training and tracking, e.g. using reference signals (RS), or pilots. However, it is not always feasible to use training-based schemes, with systems that have a large number of antennas. As the number of transmit antennas grows large such as in the case of massive MIMO, the channel estimation process, system overheads, and latency will grow proportionately [43]. Discussion of DM-RS improvements are ongoing in 3GPP release 15 standardization [44], hence, it is natural to adopt differential modulation with massive MIMO to reduce the overhead and latency of DM-RS.

A well-established method to enhance DM is multiple symbol differential detection (MSDD). The authors in [45] point out a 3dB performance improvement simply by demodulating the

received symbols jointly as a block, instead of one at a time using the MSDD detection technique. The authors of [46, 47] developed MSDD detection for the uplink MIMO system in ultra-wideband (UWB) systems. Essentially, the authors in [47] adopted decision feedback differential detection (DFDD) for a massive MIMO system, as this approach improves the performance of MSDD. However, the multiuser transmission scheme in [47] suffers from severe MAI without a proper precoding design scheme. Furthermore, prior research on MSDD and MAI cancellation has mainly been focused on uplink transmission, where cancellation was implemented at the BS receiver, and therefore complexity was not a significant concern [46, 47]. For downlink transmission, however, interference cancellation at end users increases receiver complexity, and for this reason, it is better to account for interference cancellation at the BS instead of receivers.

Therefore, it is of interest to design massive MIMO systems that precode the transmitted signals at the BS, enhance system performance, provide simple transceiver schemes, and avoid channel estimation at both ends.

Chapter 3

Exploiting Low Rate Differential STBC in Downlink MU-MIMO Systems

3.1 Introduction

Future wireless systems require effective transmission techniques to support high data rate and reliable communications. As such, a potential technique to utilize as part of multiple antenna systems to enhance system diversity is STBC [1]. In the multiuser multiple-input multiple-output MU-MIMO downlink, transmit diversity gain can be maximised by using downlink transmission techniques such as transmit precoding, e.g., the block diagonalization (BD), and transmit spreading, such as the orthogonal spreading (OS) scheme. These techniques allow the MU-MIMO channels to be decomposed into parallel single user non-interfering channels, and hence eliminate CCI [6, 7].

Implementing a precoding technique at the transmitter (e.g., a base station) helps maintain simplicity in the receiver, so that simple linear decoding approaches are applicable at the receivers (e.g., end users). For the MU-MIMO downlink, the knowledge of CSI at the transmitter helps the precoder to cancel the CCI at each user. The authors in [6] assumed full CSI knowledge at the transmitter and then they built up the BD precoder to cancel the CCI. In fact, the CSI is estimated first at the receivers and then fed back to the transmitter. This signalling process between the receivers and the transmitter will increase the complexity of the receivers, e.g., mobile devices. In [7], the authors proposed a method that combines the precoding method in [6] and the Alamouti STBC. The proposed method provides a fundamental gain in terms of low complexity decoding and spatial diversity. However, in order to coherently decode the Alamouti STBC, each receiver still needs to know the composite channel formed by the precoder and the channel. In practice, each receiver obtains the composite channel by direct estimation.

The prior focus of STBC MU-MIMO downlink transmission techniques has been on cases where CSI is available at the receivers and transmitter. However, for some systems, due to high mobility and the cost of channel training and estimation, CSI acquisition is impossible [26]. One alternative method for such systems is DM. In this work, the use of DM for downlink transmission in a MU-MIMO system is considered. Specifically, the use of DM combined with the BD and OS schemes is shown. Furthermore, DM is considered for both schemes based on the Alamouti STBC in order to eliminate the need for estimating the composite channels formed by the precoders and the channels at the receivers. In the BD scheme, the use of DM is to simplify the complexity of the receivers by eliminating the need for CSI as well as to cancel CCI. In particular, in order to have low complexity receivers, it is assumed that the channels are estimated at the transmitter, since it can tolerate more complexity compared to the receivers. Once the channels are estimated at the BS, the transmitter computes the precoder as in [6, 7].

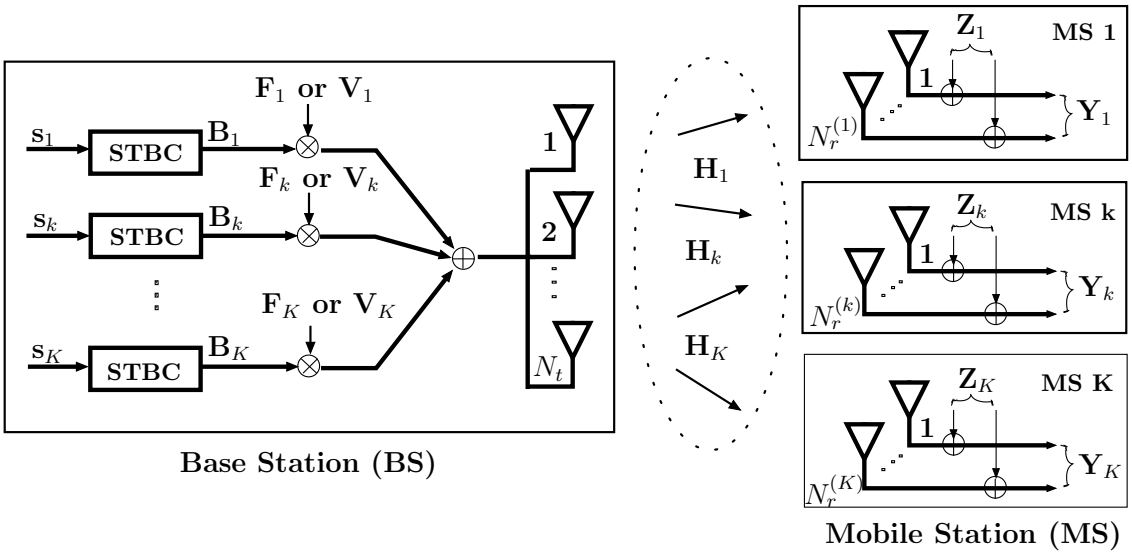

Figure 3.1. STBC MU-MIMO downlink transmission system.

However, since the BD scheme still requires CSI at the transmitter, a downlink OS scheme combined with DM is proposed. In the OS scheme, unlike the BD scheme, the transmitter does not require any knowledge of the CSI to separate the data streams of multiple users [35, 38]. Therefore, implementing the OS scheme with the DM will result in a system that does not need CSI at either ends. The proposed schemes facilitate the pre-cancelling of CCI, enhance diversity, as well as achieve a low complexity transmitter and receivers. Moreover, transmission overhead is significantly reduced using the proposed scheme, since neither feedback nor the estimation of the composite channels are required. Note that the BD scheme uses the spatial dimension to cancel CCI, whereas the OS scheme uses the time dimension.

The rest of this chapter is organized as follows. Section 3.2 introduces the system model of STBC MU-MIMO. Section 3.3 describes downlink transmission for interference cancellation. Section 3.5 presents differential STBC for MU-MIMO system with downlink transmission. In Section 3.6, the simulation results are shown. Finally, a chapter summary is given in Section 3.7.

3.2 System Model

Consider a MU-MIMO downlink broadcast channel where the BS transmits multiple streams to K users (e.g., mobile stations), as shown in Fig. 3.1. The BS has N_t transmit antennas and each user has $N_r^{(k)}$, $k = 1, \cdots, K$, receive antennas. The total number of receive antennas for all users is N_r, i.e., $N_r = \sum_{k=1}^{K} N_r^{(k)}$.

3.2.1 Channel Model

The channel matrix $\boldsymbol{H}_k \in \mathbb{C}^{N_r^{(k)} \times N_t}$ for each user k is a Rayleigh flat fading matrix given by

$$\boldsymbol{H}_k = \begin{bmatrix} h_{1,1} & \cdots & h_{1,N_t} \\ \vdots & \ddots & \vdots \\ h_{N_r^{(k)},1} & \cdots & h_{N_r^{(k)},N_t} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_1 \\ \vdots \\ \boldsymbol{h}_{N_r^{(k)}} \end{bmatrix}, \tag{3.1}$$

where the element $h_{i,j}^{(k)}$ is the channel coefficient between the jth transmit antenna and the ith receive antenna of user k, and $\mathbb{C}$ denotes the set of complex numbers. It is assumed that the channel coefficients are quasi-static over T transmission slots. The elements of $\boldsymbol{H}_k$ are independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance, i.e., $\mathcal{CN}(0,1)$.

3.2.2 Space-Time Block Coding - Alamouti Code

The multiple data streams $\boldsymbol{s}_k$ for each user are encoded by the Alamouti encoder to generate the STBC codeword. Let $\boldsymbol{B}_k \in \mathbb{C}^{2\times 2}$, $k = 1, \cdots, K$, be the transmitted Alamouti STBC signal, satisfying the following condition [7, 29]:

$$\boldsymbol{B}_k^H \boldsymbol{B}_k = \boldsymbol{B}_k \boldsymbol{B}_k^H = \boldsymbol{I}_2. \tag{3.2}$$

The generator matrix for the Alamouti code is given as

$$\boldsymbol{B}_k = \frac{1}{\sqrt{2}} \begin{bmatrix} s_{1,k} & -s_{2,k}^* \\ s_{2,k} & s_{1,k}^* \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{b}_{1,k} & \boldsymbol{b}_{2,k} \end{bmatrix} \tag{3.3}$$

where $s_{1,k}$ and $s_{2,k} \in \mathcal{C}$ are the two input symbols to the Alamouti STBC encoder for user k. $\mathcal{C}$ and $(.)^H$ denote the constellation set and the Hermitian operator, respectively.

3.3 Downlink Transmission for Interference Cancellation

In this section, two different methods are used to cancel CCI in downlink transmission. The first scheme, referred to as the BD scheme, is suitable for the case where the CSI is available at the transmitter and the second scheme, referred to as the OS scheme, is suitable for the case where the CSI is not available at the transmitter.

3.3.1 BD Scheme

The received signal $\boldsymbol{Y}_k^{(\mathrm{BD})} \in \mathbb{C}^{N_r^{(k)} \times 2}$ at the kth user can be expressed as

$$\begin{aligned}\boldsymbol{Y}_k^{(\mathrm{BD})} &= \boldsymbol{H}_k \boldsymbol{F}_k \boldsymbol{B}_k + \boldsymbol{H}_k \sum_{j=1, j\neq k}^{K} \boldsymbol{F}_j \boldsymbol{B}_j + \boldsymbol{Z}_k \\ &= \boldsymbol{H}_k \boldsymbol{F}_k \boldsymbol{B}_k + \boldsymbol{P}_k + \boldsymbol{Z}_k \ , \end{aligned} \tag{3.4}$$

where $\boldsymbol{F}_k \in \mathbb{C}^{N_t \times 2}$ is the precoding matrix, $\boldsymbol{Z}_k \in \mathbb{C}^{N_r^{(k)} \times 2}$ is an additive white Gaussian noise (AWGN) noise matrix. $\boldsymbol{P}_k \in \mathbb{C}^{N_r^{(k)} \times 2}$ is the CCI component at the kth user. Note that, at the BS, the precoding matrix $\boldsymbol{F}_k$ for the kth user is multiplied by the symbol matrix $\boldsymbol{B}_k$ and added to the precoded signals from the other users to produce the composite transmitted matrix, i.e., $\sum_{k=1}^{K} \boldsymbol{F}_k \boldsymbol{B}_k$.

The BD method employs precoding matrices $\boldsymbol{F}_k, k = 1, \cdots, K$, to completely suppress the CCI at the receivers. To cancel the CCI, the following constraint should be satisfied [6, 7]

$$\boldsymbol{H}_j \boldsymbol{F}_k = 0 \ , \quad j, k = 1, ..., K, \ j \neq k. \tag{3.5}$$

Let $\bar{\boldsymbol{H}}_k \in \mathbb{C}^{\bar{N}_r^{(k)} \times N_t}$, where $\bar{N}_r^{(k)} = N_r - N_r^{(k)}$, denote the channel matrix for all K users excluding the kth user's channel, which is defined as

$$\bar{\boldsymbol{H}}_k = \left[\boldsymbol{H}_1^H \ \cdots \ \boldsymbol{H}_{k-1}^H \ \boldsymbol{H}_{k+1}^H \ \cdots \ \boldsymbol{H}_K^H \right]^H . \tag{3.6}$$

Therefore, the zero-interference constraint in (3.5) is re-expressed as

$$\bar{\boldsymbol{H}}_k \boldsymbol{F}_k = 0 \ , \qquad k = 1, ..., K. \tag{3.7}$$

According to [7], to satisfy (3.7), one solution is to construct $\boldsymbol{F}_k$ as

$$\boldsymbol{F}_k = (\boldsymbol{I} - \bar{\boldsymbol{H}}_k^{\dagger} \bar{\boldsymbol{H}}_k) \boldsymbol{\Phi}_k \ , \tag{3.8}$$

where $\boldsymbol{\Phi}_k \in \mathbb{C}^{N_t \times 2}$ is an eigenmode selection matrix, and $(.)^{\dagger}$ denotes the pseudo-inverse. The magnitude, i.e, the vector norm of the precoding matrix $\boldsymbol{F}_k$ has to be unity to ensure a constant transmission power for the kth user, i.e.,

$$\boldsymbol{F}_k^H \boldsymbol{F}_k = \boldsymbol{I}_2 \ , \qquad k = 1, \cdots, K. \tag{3.9}$$

Therefore, to satisfy (3.9), the unitary $\boldsymbol{F}_k$ matrix can be constructed as a linear combination of the column space spanning vectors of $(\boldsymbol{I} - \bar{\boldsymbol{H}}_k^{\dagger}\bar{\boldsymbol{H}}_k)$, which can be obtained by the Gram-Schmidt orthogonalization (GSO), or the standard QR decomposition. In this chapter, QR decomposition is used for its simplicity.

To compute $\boldsymbol{\Phi}_k$, a singular value decomposition (SVD) of $\boldsymbol{H}_k(\boldsymbol{I} - \bar{\boldsymbol{H}}_k^{\dagger}\bar{\boldsymbol{H}}_k)$ is performed. This is done by selecting the two singular vectors corresponding to the two largest singular values of $\boldsymbol{H}_k(\boldsymbol{I} - \bar{\boldsymbol{H}}_k^{\dagger}\bar{\boldsymbol{H}}_k)$. The resulting received signal for the kth user after cancelling out the CCI is given by

$$\boldsymbol{Y}_k^{(\mathrm{BD})} = \boldsymbol{H}_k\boldsymbol{F}_k\boldsymbol{B}_k + \boldsymbol{Z}_k = \breve{\boldsymbol{H}}_k\boldsymbol{B}_k + \boldsymbol{Z}_k, \tag{3.10}$$

where $\breve{\boldsymbol{H}}_k \in \mathbb{C}^{N_r^{(k)}\times 2}$ is the effective channel for user k. Note that the CCI elimination applies as long as the transmit precoding can be implemented, which requires that the number of transmit antennas is sufficient to achieve full diversity with the given number of receive antennas. According to [7] the minimum number of transmit antennas should satisfy

$$N_{t,min} = \max_i \left[\sum_{k=1,k\neq i}^{K} N_r^{(k)} \right] + 2. \tag{3.11}$$

For the case where all users have the same number of antennas, i.e, $N_r^{(1)} = N_r^{(2)} = \cdots = N_r^{(K)} = N$, (3.11) becomes

$$N_{t,min} = [(K-1)N] + 2. \tag{3.12}$$

3.3.1.1 Channel Estimation Error Model

The high mobility and multipath propagation may result in MAI for the BD scheme, which destroys the orthogonality of the precoder matrix [48]. The estimated channel error matrix can be modelled as

$$\ddot{\boldsymbol{H}}_k = \boldsymbol{H}_k + \boldsymbol{E}_k, \tag{3.13}$$

where $\boldsymbol{H}_k$ is the perfect channel value and $\boldsymbol{E}_k$ is the error matrix. Entries of $\boldsymbol{E}_k$ are i.i.d. Gaussian variables with distribution zero mean and covariance of σ^2. Section 3.6 shows the impact of the error on channel estimation.

3.3.2 OS Scheme

In the OS case, the received signal matrix $\boldsymbol{Y}_k^{(\mathrm{OS})} \in \mathbb{C}^{N_r^{(k)} \times KN_t}$ for the kth user is given by [38]

$$\boldsymbol{Y}_k^{(\mathrm{OS})} = \boldsymbol{H}_k \boldsymbol{B}_k \boldsymbol{V}_k + \boldsymbol{H}_k \sum_{j=1, j\neq k}^{K} \boldsymbol{B}_j \boldsymbol{V}_j + \boldsymbol{Z}_k, \tag{3.14}$$

where $\boldsymbol{V}_k \in \mathbb{C}^{N_t \times KN_t}$ is the orthogonal spreading matrix for user k, $\boldsymbol{Z}_k \in \mathbb{C}^{N_r^{(k)} \times KN_t}$ is an AWGN noise matrix. The composite transmitted matrix is $\sum_{k=1}^{K} \boldsymbol{B}_k \boldsymbol{V}_k$. Note that, in order to apply Alamouti STBC along with the orthogonal spreading code, the number of transmit antennas at the BS has to be limited to two, i.e., $N_t = T = 2$.

In the OS scheme, each user is assigned a unique orthogonal spreading code to separate the data of the users at the receivers. The STBC codeword for each user is multiplexed by its own specific spreading code and then transmitted. As in the BD method case, to eliminate CCI, the spreading code matrix has to obey the following conditions

$$\boldsymbol{V}_k \boldsymbol{V}_k^H = \boldsymbol{I}_{N_t}\,, \qquad k = 1, ..., K. \tag{3.15}$$

$$\boldsymbol{V}_j \boldsymbol{V}_k^H = 0\,, \qquad k, j = 1, ..., K, \text{ and } j \neq k. \tag{3.16}$$

The OS code for each user can be constructed as a submatrix of the Hadamard matrix, or from a discrete Fourier transform (DFT) matrix. Hadamard matrices are of interest because of their simplicity where it can be constructed easily. Hadamard transform is an example of a generalized form of Fourier transforms and it can be built out of size-2 DFTs, and is equivalent to a multidimensional DFT. Hadamard codes are a set of orthogonal codes which are built repeatedly from the basic building block

$$\boldsymbol{A}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix} \tag{3.17}$$

according to

$$\boldsymbol{A}_{2^{n+1}} = \frac{1}{\sqrt{2^{n+1}}} \begin{bmatrix} \boldsymbol{A}_{2^n} & \boldsymbol{A}_{2^n} \\ \boldsymbol{A}_{2^n} & -\boldsymbol{A}_{2^n} \end{bmatrix}, \tag{3.18}$$

where the dimension of the Hadamard matrix in (3.18) is $2^{n+1} \times 2^{n+1}$. Note that in this work $2^{n+1} = KN_t$.

Due to the orthogonality of the spreading matrices used at the transmitter, at each receiver,

the original information signal is retrieved by despreading the received signal with the synchronized duplicate of the spreading code. Therefore, the received signal matrix $\boldsymbol{Y}_k^{(\mathrm{OS})}$ in (3.14) for the kth user is despread by multiplying it with $\boldsymbol{V}_k^H$, which yields

$$\hat{\boldsymbol{Y}}_k^{(\mathrm{OS})} = \boldsymbol{Y}_k^{(\mathrm{OS})}\boldsymbol{V}_k^H = \boldsymbol{H}_k\boldsymbol{B}_k + \hat{\boldsymbol{Z}}_k, \tag{3.19}$$

where $\hat{\boldsymbol{Y}}_k^{(\mathrm{OS})} \in \mathbb{C}^{N_r^{(k)} \times N_t}$ is the despread received signal, and $\hat{\boldsymbol{Z}}_k \in \mathbb{C}^{N_r^{(k)} \times N_t}$ is the despread AWGN noise.

3.3.2.1 Spreading Code Error Model

It is not always feasible to have perfect orthogonality for the OS precoder, especially when the user is moving so fast that the OS code is changing very rapidly. For the OS scheme, the error spreading matrix can be expressed as

$$\bar{\boldsymbol{V}}_k = \boldsymbol{V}_k + \alpha\boldsymbol{V}_j, \qquad j \neq k. \tag{3.20}$$

where α is the error coefficient. For example, the conditions for the orthogonality of spreading code matrix for User 1 against User 2 are as follows

$$\bar{\boldsymbol{V}}_1\bar{\boldsymbol{V}}_1^H = \boldsymbol{I}_{N_t} + \alpha^2\boldsymbol{I}_{N_t}, \tag{3.21}$$

$$\boldsymbol{V}_2\bar{\boldsymbol{V}}_1^H = \alpha\boldsymbol{I}_{N_t}. \tag{3.22}$$

Section 3.6 illustrates the impact of error on the spreading code.

3.3.3 Complexity Analysis for the Precoders

In this section, the computational complexity with the notion of flops is introduced here, where flops denotes the floating point operations. At the transmitter, the BD scheme uses the spatial dimension to cancel CCI, whereas the OS scheme uses the time dimension. In the BD scheme, in order to cancel CCI completely, the system must satisfy [6, 7]

$$N_t \geq \left(\sum_{j=1, j\neq k}^{K} N_r^{(j)} + 2 \right). \tag{3.23}$$

The complexity of the BD scheme is mainly based on the pseudo-inverse $\bar{\boldsymbol{H}}_k^{\dagger} = \bar{\boldsymbol{H}}_k^H \left(\bar{\boldsymbol{H}}_k \bar{\boldsymbol{H}}_k^H\right)^{-1}$, and the QR decomposition of $(\boldsymbol{I} - \bar{\boldsymbol{H}}_k^{\dagger}\bar{\boldsymbol{H}}_k)$. The complexity of both the pseudo-inverse operation and the QR decomposition follows [49, 50]

$$\mathcal{O}\left(KN_t\left(\sum_{j=1,j\neq k}^{K} N_r^{(j)}\right)^2\right). \tag{3.24}$$

In the OS scheme, the precoder is independent from the number of receive antennas. Thus, the complexity of the OS scheme is only based on Hadamard matrix construction which is already given. Hence, it does not incur any computational complexity. Obviously, the OS scheme has lower computational complexity than the BD scheme, but in terms of throughput, the OS scheme throughput is K times smaller than that of the BD scheme. Note that, the computational complexity at the receiver side for both schemes is the same, and the following section gives more details about the DM decoder.

3.4 Coherent STBC for MU-MIMO with Downlink Transmission

3.4.1 Coherent Detection

When the channel state information is available at the receiver, the received Alamouti STBC signal can be detected coherently. The decoder processes for both precoding schemes are provided as follows:

3.4.1.1 Coherent Detection for BD

To decode the Alamouti space-time code coherently for precoding scheme BD, the mobile nodes (the receivers) need the effective channel matrix $\breve{\boldsymbol{H}}_k$ and the received signal $\boldsymbol{Y}_k^{(\mathrm{BD})}$. The combining signals are computed based on the standard Alamouti decoder, it follows the following general case

$$\begin{bmatrix}\hat{s}_{k,1}\\ \hat{s}_{k,2}\end{bmatrix} = \begin{bmatrix}\breve{h}_{k,1,1}^{*} & \breve{h}_{k,1,2}\\ \breve{h}_{k,1,2}^{*} & -\breve{h}_{k,1,1}\end{bmatrix}\begin{bmatrix}y_{k,1,1}\\ y_{k,1,2}^{*}\end{bmatrix} + \begin{bmatrix}\breve{h}_{k,2,1}^{*} & \breve{h}_{k,2,2}\\ \breve{h}_{k,2,2}^{*} & -\breve{h}_{k,2,1}\end{bmatrix}\begin{bmatrix}y_{k,2,1}\\ y_{k,2,2}^{*}\end{bmatrix} + \cdots + \cdots \begin{bmatrix}\breve{h}_{k,N_r^{(k)},1}^{*} & \breve{h}_{k,N_r^{(k)},2}\\ \breve{h}_{k,N_r^{(k)},2}^{*} & -\breve{h}_{k,N_r^{(k)},1}\end{bmatrix}\begin{bmatrix}y_{k,N_r^{(k)},1}\\ y_{k,N_r^{(k)},2}^{*}\end{bmatrix}, \tag{3.25}$$

where $\hat{s}_{k,j}$ is the jth estimated transmitted symbol of user k, $\breve{h}_{k,i,j}$ is the effective channel coefficient intended for the jth transmit symbol (always j is either 1 or 2) and the ith receive antenna of user k during coherence interval time, $y_{k,j,i}$ is the received signal for the kth user at the receiver i intended for the jth transmit symbol.

Then, after computing the combined signals, the ML decoder which is the optimum receiver, can be used. The transmitted symbols according to the ML detection are computed as follows

$$\hat{\boldsymbol{b}}_k = \arg \min_{\boldsymbol{b} \in \mathcal{C}} \left\| \boldsymbol{y}_k - \breve{\boldsymbol{H}}_k \boldsymbol{b}_k \right\| \tag{3.26}$$

3.4.1.2 Coherent Detection for OS

For the OS schemes, to decode the Alamouti space-time code coherently, the mobile nodes need only the channel matrix $\boldsymbol{H}_k$ and the despread received signal $\hat{\boldsymbol{Y}}_k^{(\mathrm{OS})}$. Hence, the combining signals are computed based on the Alamouti decoder as follows

$$\begin{aligned} \begin{bmatrix} \hat{s}_{k,1} \\ \hat{s}_{k,2} \end{bmatrix} = & \begin{bmatrix} h^*_{k,1,1} & h_{k,1,2} \\ h^*_{k,1,2} & -h_{k,1,1} \end{bmatrix} \begin{bmatrix} \hat{y}_{k,1,1} \\ \hat{y}^*_{k,1,2} \end{bmatrix} + \begin{bmatrix} h^*_{k,2,1} & h_{k,2,2} \\ h^*_{k,2,2} & -h_{k,2,1} \end{bmatrix} \begin{bmatrix} \hat{y}_{k,2,1} \\ \hat{y}^*_{k,2,2} \end{bmatrix} + \quad \cdots \\ & + \quad \cdots \quad \begin{bmatrix} h^*_{k,N_r^{(k)},1} & h_{k,N_r^{(k)},2} \\ h^*_{k,N_r^{(k)},2} & -h_{k,N_r^{(k)},1} \end{bmatrix} \begin{bmatrix} \hat{y}_{k,N_r^{(k)},1} \\ \hat{y}^*_{k,N_r^{(k)},2} \end{bmatrix}, \end{aligned} \tag{3.27}$$

where $h_{k,i,j}$ is the channel coefficient intended for the jth transmit symbol the ith receive antenna of user k during coherence interval time and $\hat{y}_{k,j,i}$ is the despread received signal for the kth user at the receiver i intended for the jth transmit symbol. Then, the transmitted symbols can be recovered using the ML decoder as follows

$$\hat{\boldsymbol{b}}_k = \arg \min_{\boldsymbol{b} \in \mathcal{C}} \| \hat{\boldsymbol{y}}_k - \boldsymbol{H}_k \boldsymbol{b}_k \| \tag{3.28}$$

3.5 Differential STBC for MU-MIMO with Downlink Transmission

In this section, the differential encoding and decoding process for downlink transmission in a MU-MIMO system is discussed. In particular, this section demonstrates how to use the BD and OS schemes in differential STBC MU-MIMO systems.

3.5.1 Differential Encoding

The particular encoding algorithm utilized for DM builds upon the works in [28, 29]. The algorithm requires that unitary STBCs such as the Alamouti code are used. In the encoding process, the $\boldsymbol{B}_0$ matrix is used as a reference code, in which the transmitted matrix for the initial block of each user k is set to be identity as

$$\boldsymbol{B}_{0,k} = \boldsymbol{I}_T, \qquad k = 1, \cdots, K. \tag{3.29}$$

Then, for the BD scheme, the unitary Alamouti STBC matrices are encoded differentially for the subsequent blocks as follows

$$\boldsymbol{B}_n^{(\mathrm{BD})} = \sum_{k=1}^{K} \boldsymbol{F}_k \left(\prod_{i=0}^{n} \boldsymbol{B}_{i,k} \right) \ , \qquad n = 0, ..., N. \tag{3.30}$$

For the OS scheme, the encoding process is as follows

$$\boldsymbol{B}_n^{(\mathrm{OS})} = \sum_{k=1}^{K} \left(\prod_{i=0}^{n} \boldsymbol{B}_{i,k} \right) \boldsymbol{V}_k \ , \qquad n = 0, ..., N, \tag{3.31}$$

where $\boldsymbol{B}_n^{(q)}, q \in \{\mathrm{BD}, \mathrm{OS}\}$, is the nth encoded block, $N+1$ is the total number of encoded signal blocks, and $\boldsymbol{F}_k$ and $\boldsymbol{V}_k$ represent the precoding matrix and spreading matrix for user k, respectively.

The performance of the differential modulation system depends on the length of time over which the channel coefficients remain constant. Ordinarily, the reference (known) symbol $\boldsymbol{B}_{0,k}$ must be sent periodically, based on the channel coherence time. Accordingly, generating the downlink precoding matrix $\boldsymbol{F}_k$ or the downlink spreading matrix $\boldsymbol{V}_k$ for the new channel coefficient matrix only needs to be done when there are new channel coefficients.

3.5.2 Differential Decoding

For the MU-MIMO downlink system, the differential transmissions are implemented in blocks, in which each user k receives the sum of all the transmit waveforms of other users; then the received signal blocks for each user must be detected independently. Thus, if $\boldsymbol{G}_k$ denotes the matrix having all $N+1$ received signal blocks for the kth user, i.e.,

$$\boldsymbol{G}_k = \left[\boldsymbol{Y}_{0,k} \ \ \boldsymbol{Y}_{1,k} \ \ \cdots \ \ \boldsymbol{Y}_{N,k} \right], \tag{3.32}$$

then the received signal block at the kth user during the nth iteration block, i.e., $\boldsymbol{Y}_{n,k}$ can be expressed as

$$\boldsymbol{Y}_{n,k} = \boldsymbol{H}_k \boldsymbol{B}_n^{(q)} + \boldsymbol{Z}_{n,k}, \qquad n = 0, ..., N, \tag{3.33}$$

where $q \in \{\text{BD}, \text{OS}\}$, and $\boldsymbol{Z}_{n,k}$ is the kth user AWGN noise during the nth block. For DM encoding, it is assumed that the channel matrix $\boldsymbol{H}_k$ changes slowly (channel coherence time is large enough) and extends over several matrix transmission periods. In such a case, the BS transmission starts with a reference matrix, followed by several information matrices. When encoding using (3.30) or (3.31), the decoding process for $\boldsymbol{B}_{n,k}$ would be according to the last two blocks of $\boldsymbol{G}_k$ as in the following notation [28, 29]

$$\boldsymbol{G}_k = \left[\underbrace{\boldsymbol{Y}_{0,k}\boldsymbol{Y}_{1,k}} \cdots \underbrace{\boldsymbol{Y}_{n-1,k}\boldsymbol{Y}_{n,k}} \cdots \underbrace{\boldsymbol{Y}_{N-1,k}\boldsymbol{Y}_{N,k}} \right]. \tag{3.34}$$

For the BD method, to make this more explicit, define

$$\boldsymbol{Y}_{n,k} \triangleq \begin{bmatrix} \boldsymbol{Y}_{n-1,k} \\ \boldsymbol{Y}_{n,k} \end{bmatrix} \triangleq \begin{bmatrix} \boldsymbol{H}_k \boldsymbol{B}_{n-1}^{(q)} + \boldsymbol{Z}_{n-1,k} \\ \boldsymbol{H}_k \boldsymbol{B}_n^{(q)} + \boldsymbol{Z}_{n,k} \end{bmatrix}, \tag{3.35}$$

and recall from (3.5) that the interference of other users is suppressed, thus the two blocks in (3.35) become a single user block matrix as

$$\boldsymbol{Y}_{n,k} \triangleq \begin{bmatrix} \boldsymbol{H}_k \boldsymbol{F}_k \boldsymbol{B}_{n-1,k} + \boldsymbol{Z}_{n-1,k} \\ \boldsymbol{H}_k \boldsymbol{F}_k \boldsymbol{B}_{n-1,k} \boldsymbol{B}_{n,k} + \boldsymbol{Z}_{n,k} \end{bmatrix}. \tag{3.36}$$

The code matrices that affect $\boldsymbol{Y}_{n,k}$ are

$$\boldsymbol{D}_{B_{n,k}} = \begin{bmatrix} \boldsymbol{B}_{n-1,k} \\ \boldsymbol{B}_{n-1,k}\boldsymbol{B}_{n,k} \end{bmatrix}. \tag{3.37}$$

Assuming that $N_t = T$, and using these results, as well as (3.2) and (3.9), the matrices in (3.37) can be expressed as

$$\boldsymbol{D}_{B_{n,k}}^H \boldsymbol{D}_{B_{n,k}} = 2\boldsymbol{I}_{N_t}, \tag{3.38}$$

therefore, these matrices represent unitary block codes. When $\boldsymbol{B}_{n-1,k}$ is known to the receiver, the optimal decoder for this block is the quadratic receiver as [28]

$$\hat{\boldsymbol{B}}_{n,k} = \arg \max_{\boldsymbol{B}_{n,k}} \text{trace} \left\{ \boldsymbol{Y}_{n,k} \boldsymbol{D}_{B_{n,k}} \boldsymbol{D}_{B_{n,k}}^H \boldsymbol{Y}_{n,k}^H \right\}. \tag{3.39}$$

Since it follows

$$\boldsymbol{D}_{B_{n,k}}\boldsymbol{D}_{B_{n,k}}^{H} = \begin{bmatrix} \boldsymbol{I}_T & \boldsymbol{B}_{n,k}^{H} \\ \boldsymbol{B}_{n,k} & \boldsymbol{I}_T \end{bmatrix}, \tag{3.40}$$

the decoder in (3.39) can be re-written as follows [28], [29]

$$\hat{\boldsymbol{B}}_{n,k} = \arg \max_{\boldsymbol{B}_{n,k}} \text{trace} \left\{ \begin{bmatrix} \boldsymbol{Y}_{n-1,k} \\ \boldsymbol{Y}_{n,k} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_T & \boldsymbol{B}_{n,k}^{H} \\ \boldsymbol{B}_{n,k} & \boldsymbol{I}_T \end{bmatrix} \begin{bmatrix} \boldsymbol{Y}_{n-1,k} \\ \boldsymbol{Y}_{n,k} \end{bmatrix}^{H} \right\}$$

$$= \arg \max_{\boldsymbol{B}_{n,k}} \Re \left\{ \text{trace} \left\{ \boldsymbol{B}_{n,k}\, \boldsymbol{Y}_{n,k}^{H} \boldsymbol{Y}_{(n-1),k} \right\} \right\}, \tag{3.41}$$

where $\Re(.)$ denotes the real part, and trace$(.)$ denotes the trace of a matrix. Similarly, the equivalent differential decoder for the OS scheme can be constructed. Note that when the CSI is available at the receiver, the standard Alamouti decoder is used before the maximum likelihood (ML) detection is implemented upon the combined signals.

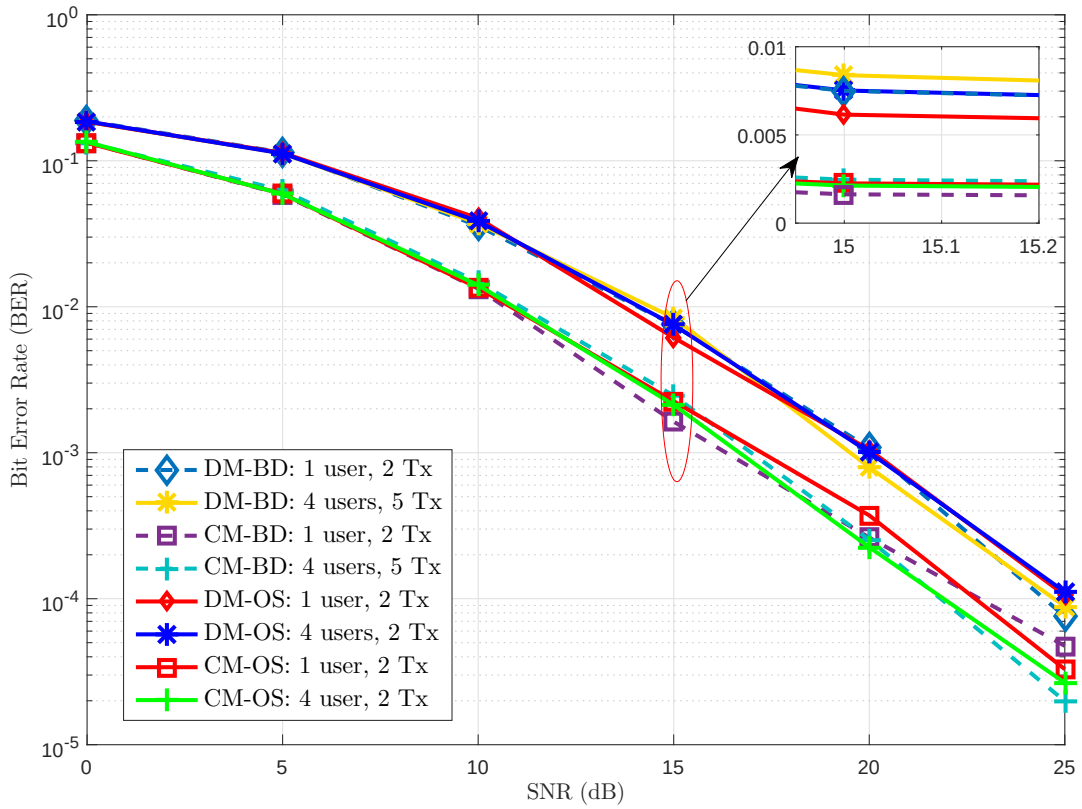

Figure 3.2. BER performance of MU-MIMO STBC downlink precoding with coherent and differential detection using BD and OS schemes for $N_r^{(k)} = 1$.

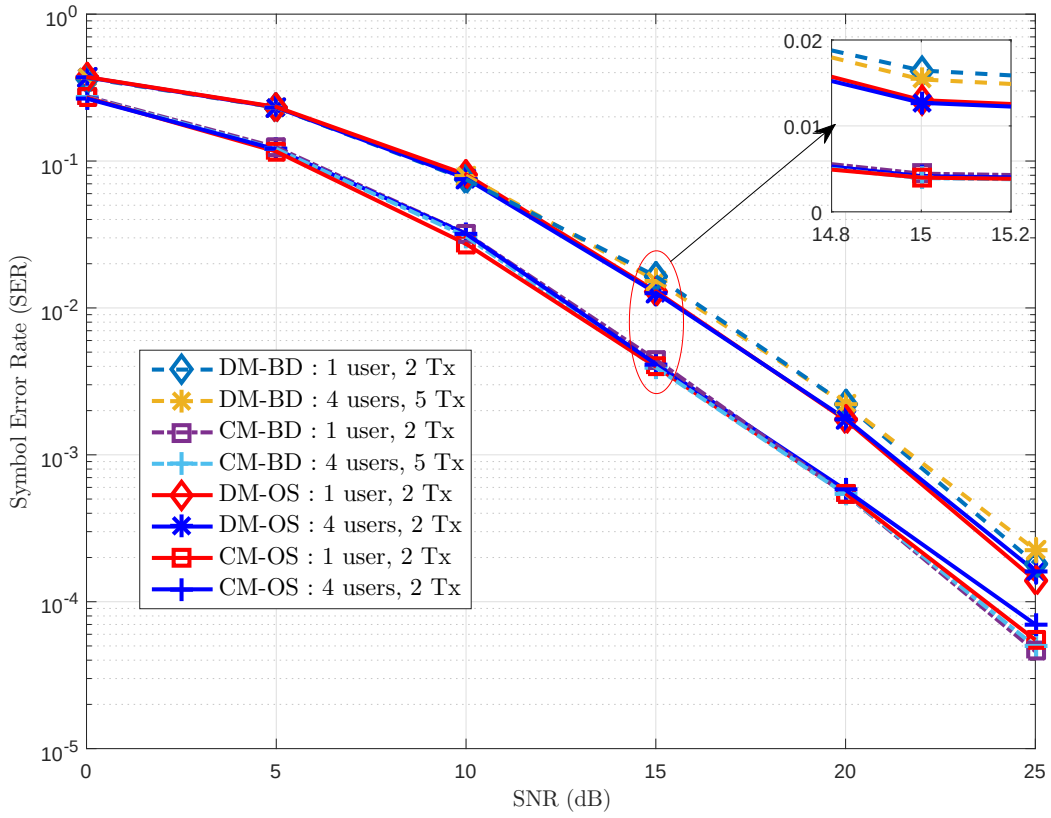

Figure 3.3. SER performance of MU-MIMO STBC downlink precoding with coherent and differential detection using BD and OS schemes for $N_r^{(k)} = 1$.

3.6 Simulations Results and Discussion

In this section, the performances of the differential and coherent Alamouti STBC for MU-MIMO downlink transmission are examined. Alamouti codes with QPSK are used throughout the simulation.

Fig. 3.2 and Fig. 3.3 plot the bit error rate (BER) and the symbol error rate (SER) for coherent modulation (CM) and DM with one receive antenna per user. For the BD scheme, the performance curve is plotted for a single user system with 2 transmit antennas at the BS and a four-user system with 5 transmit antennas at the BS. For the OS scheme, the number of transmission antenna has been set to be always two against 1 and 4 users. It is observed that CM and DM for both BD and OS schemes achieve the same performance as a single-user STBC-MISO link; that is, CCI is completely eliminated and full diversity is achieved with the Alamouti code. Ordinarily, the differential detection underperforms the coherent detection by about 3 dB.

Fig. 3.4 and Fig. 3.5 illustrate the results of repeating the experiment with two receive an-

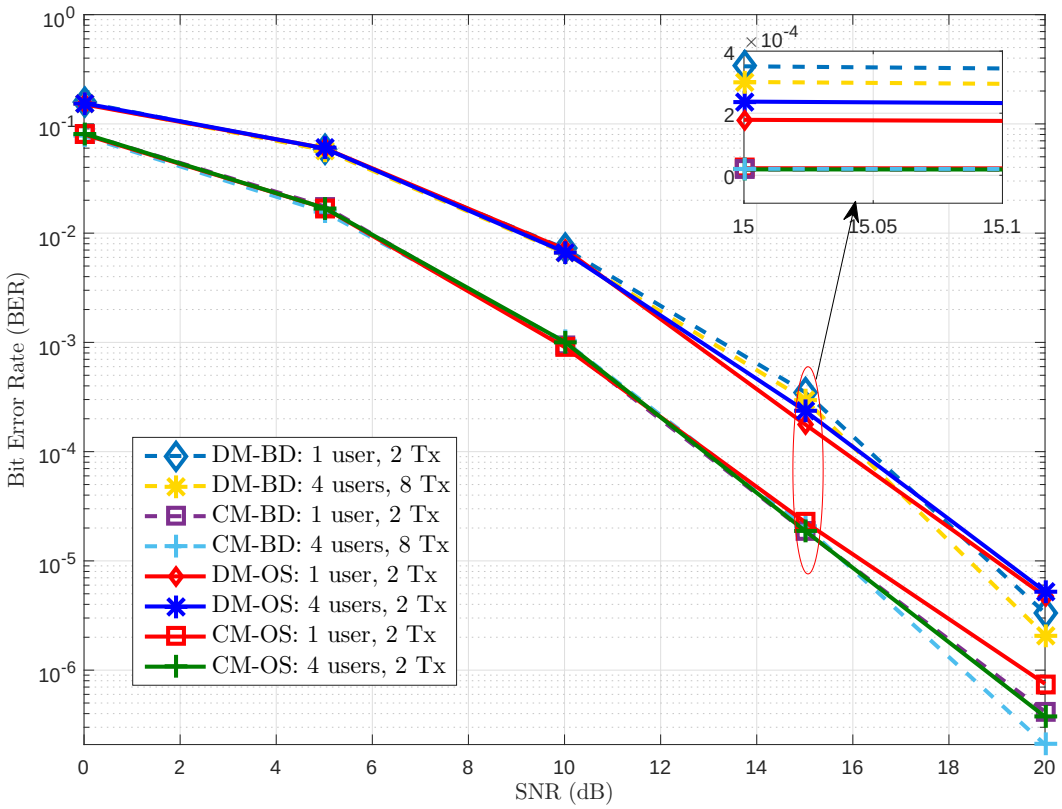

Figure 3.4. BER performance of MU-MIMO STBC downlink precoding with coherent and differential detection using BD and OS schemes for $N_r^{(k)} = 2$.

tennas per user. Similarly, the MU-MIMO system of CM and DM for both schemes behave as a single user STBC-MIMO link, but with better performance than the one receive antenna per user system. For the BD scheme, CCI elimination requires that the number of transmit antennas is sufficient to achieve full diversity with the given number of receive antennas, so $N_t = 8$ is chosen. For the OS scheme, the same performance has been observed with fixed number of transmit antennas, e.g., $N_t = 2$. Consequently, unlike the BD scheme, the number of receive antenna per user is independent of the number of transmission antenna.

Fig. 3.6 shows the performance of exploiting DM combined with BD and OS schemes with three receive antennas per user. The high mobility and multipath propagation may result in multiple access interference (MAI) in the OS scheme and imperfect channel estimation in the BD scheme, which destroy the orthogonality of the precoders. Hence, Fig. 3.6 also shows the impact of possible errors for both schemes, namely BD and OS, with two users in the system. For the OS scheme, the error spreading matrix is designed according to (3.13) where the values of α are chosen to be 0.1 and 0.2. For the BD scheme, an imperfect channel matrix is designed according (3.20). Entries of $\boldsymbol{E}_k$ are i.i.d. Gaussian variables with distribution having zero mean

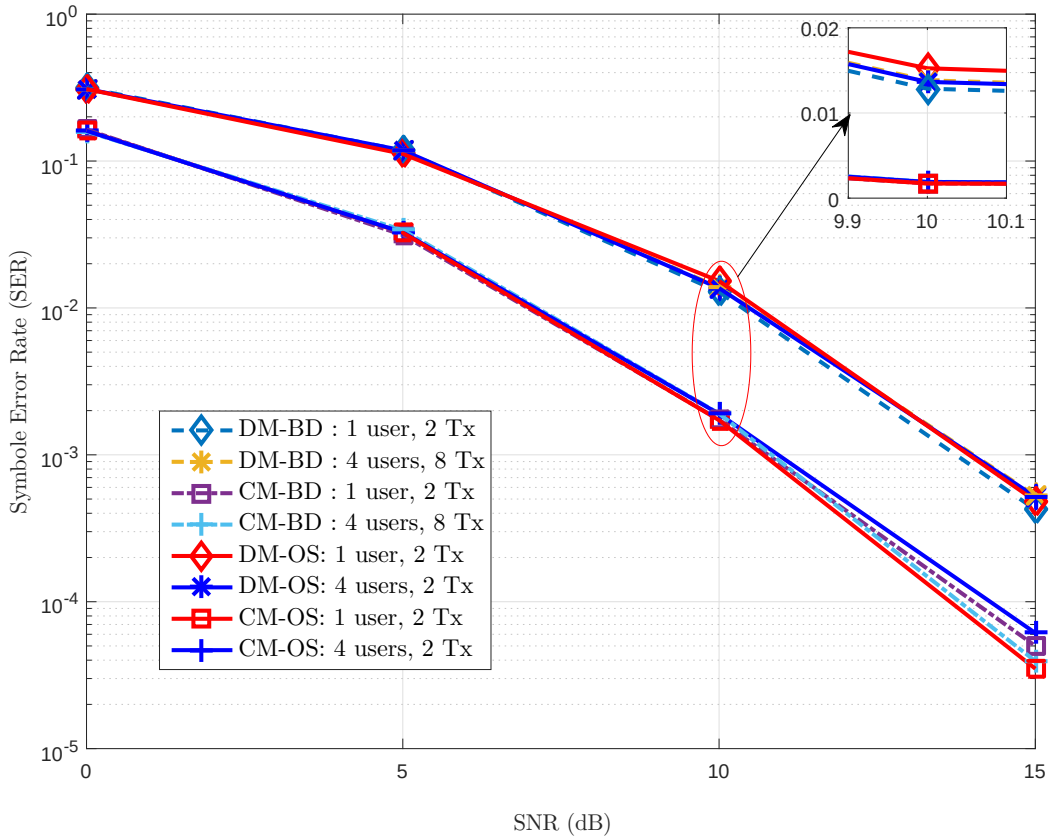

Figure 3.5. SER performance of MU-MIMO STBC downlink precoding with coherent and differential detection using BD and OS schemes for $N_r^{(k)} = 2$.

and covariance of σ^2. The values of σ are chosen to be 0.1 and 0.2. From Fig. 3.6, it is clear that the OS is more robust against errors compared to the BD scheme.

3.7 Summary

In this chapter, a low complexity differential STBC scheme for MU-MIMO with downlink transmission has been proposed. In particular, DM combined with either the BD scheme or the OS scheme overcame the need for CSI at the receivers as well as cancelled CCI. On the other hand the use of STBC can achieve full diversity without needing CSI at the transmitter. It has been demonstrated that implementing the BD scheme with DM will establish a system that does not need CSI at the receivers to decode the signals, while combining the OS scheme with DM will establish a system that requires CSI at neither the transmitter nor at the receivers. The differential modulation for both systems loses typically 3dB in performance relative to the coherent detection method, but this is offset by the reduction in complexity of the receivers and the transmitter. The BD scheme is more complex than the OS scheme; however, the BD scheme

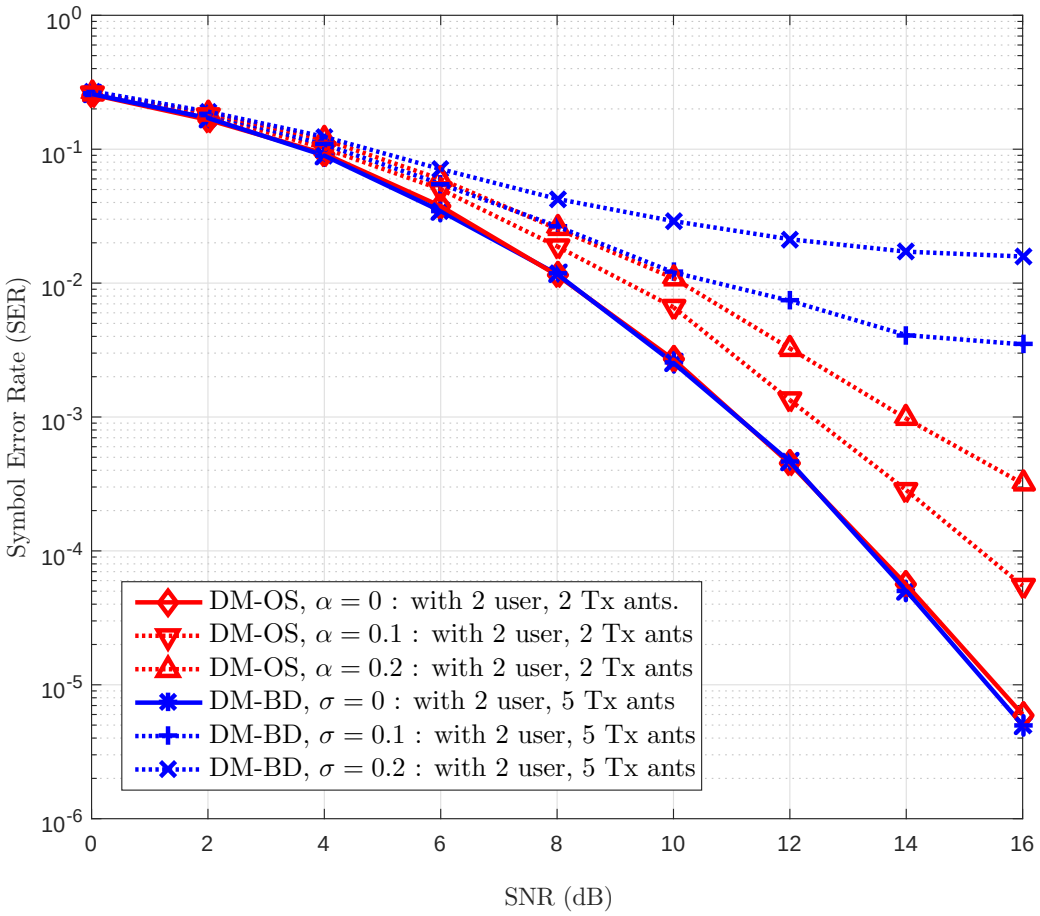

Figure 3.6. SER performance of differential detection system using BD and OS schemes for $N_r^{(k)} = 3$ with the impact of precoding errors on user 1.

has a higher throughput. Moreover, it was shown that the OS is more robust against precoding errors compared to the BD scheme.

Chapter 4

Exploiting High Rate Differential Algebraic STBC in Downlink MU-MIMO Systems

4.1 Introduction

MIMO technology is one of the most important milestones in the development of wireless communications and can be used to increase the spectral efficiency through spatial multiplexing and improve the link reliability through transmit diversity [51]. The MIMO design tradeoffs such as multiplexing, diversity, performance, and complexity in both uncoded and coded MIMO systems play a fundamental role in efficient system planning and deployment [52]. Further, wireless systems require effective transmission techniques to support high data rate and reliable communications. As such, a space-time block code (STBC) is a potential transmission technique which can be utilized, as part of multiple antenna systems, to enhance the spatial diversity of the system [53], and it is already used in standard systems such as the UMTS standard for mobile wireless, the IEEE 802.16 standard for fixed and nomadic wireless, and the IEEE 802.11 standard for wireless LANs [4].

The transmission of an orthogonal STBC over a MIMO channel is shown to have full diversity with a low complexity receiver [1, 31]. However, due to the limitation in the number of antennas, this scheme can not be used to have a very high transmission rate. In other words, the number of receive antennas is limited by the number of transmit antennas, thus this scheme does not operate for very high rate. When decoding complexity is not an issue, one may use non-orthogonal full rate full diversity algebraic STBC [32, 33]. The prior focus of the high rate MU-MIMO downlink transmission techniques has been on cases where CSI is available at the receivers and transmitter. However, for some systems, due to high mobility and the cost of channel training and estimation, CSI acquisition is impossible [26]. One alternative method is to encode the transmitted data differentially using a CD transform, and to decode differentially without any knowledge of the CSI at the receiver [18]. Our previous work in [39] has dealt with implementing the MU-MIMO downlink transmission of an Alamouti STBC combined with differential modulation, which does not require channel knowledge for decoding. The scheme provides low complexity transceivers while providing good performance. However, this work in [39] cannot provide a comprehensive high rate differential scheme in the downlink scenario.

In this chapter, the use of high rate Cayley differential STBC for downlink transmission in a MU-MIMO system is considered. Specifically, the use of differential STBC combined with full rate full diversity perfect algebraic STBC is resolved. The use of differential STBC in a multiuser scenario simplifies the complexity of the receivers, since neither feedback nor the estimation of the CSI are required at the receiver. Furthermore, differential STBC is considered based on the orthogonal spreading technique in order to separate the data streams of multiple

users. With the use of orthogonal spreading, the transmitter needs no knowledge of the CSI to design the spreading matrices. Therefore, implementing the orthogonal spreading scheme with the differential STBC will result in a system in which neither the transmitter nor the receiver needs knowledge of the CSI. At the receiver of each user, a sphere decoder (SD) is implemented for high rate coherent and differential perfect algebraic STBC to limit the set of candidate symbols to those within a sphere of some radius d. The proposed schemes facilitate the multiple user data separation, enhancing full rate full diversity, and achieving low complexity receivers and transmitters through the use of differential STBC. However, the system in this chapter has higher computational complexity thanks to its higher rate.

The rest of the chapter is organized as follows. Section 4.2 introduces the system model of STBC MU-MIMO. Section 4.3 reviews the coherent perfect algebraic STBC for MU-MIMO. Section 4.4 presents the differential perfect algebraic STBC for MU-MIMO. In Section 4.5, the computational complexity and rate analysis of the system are derived. In Section 4.6, the simulation results are shown and, finally, conclusions are drawn in Section 4.7.

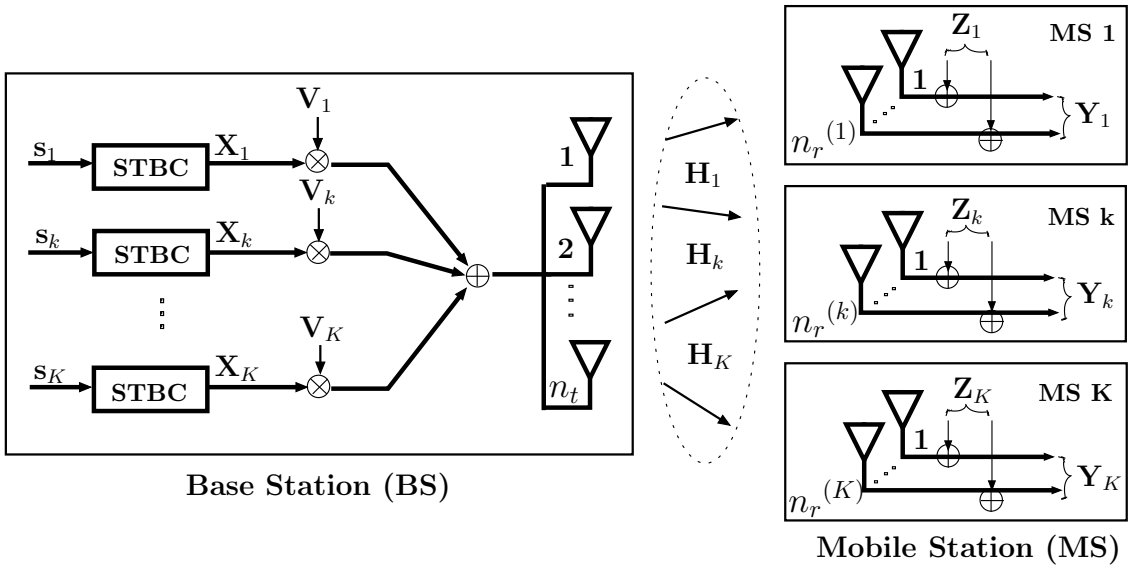

Figure 4.1. An STBC MU-MIMO downlink transmission system.

4.2 System Model

Consider a MU-MIMO downlink broadcast channel where the BS transmits multiple streams to K users (e.g., mobile stations), as shown in Fig. 4.1. The BS has N_t transmit antennas and each user k has N_r^k receive antennas. It is assumed that all users have the same number of receive antennas unless otherwise stated. Further, the superscript k is omitted for simplicity.

The channel matrix $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ for each user k is a Rayleigh flat fading matrix given by

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & \cdots & h_{1,N_t} \\ \vdots & \ddots & \vdots \\ h_{N_r,1} & \cdots & h_{N_r,N_t} \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_{N_r} \end{bmatrix}, \tag{4.1}$$

where the element $h_{i,j}$ is the channel coefficient between the jth transmit antenna and the ith receive antenna of user k, and $\mathbb{C}^{M \times N}$ denotes the set of $M \times N$ complex matrices. The elements of $\mathbf{H}$ are independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance, i.e., $\mathcal{CN}(0, 1)$.

For any kth user, the $N_t \times N_r$ information symbol matrix can be defined as

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_{N_r} \end{bmatrix} \qquad = \begin{bmatrix} s_{1,1} & \cdots & s_{1,N_r} \\ \vdots & \ddots & \vdots \\ s_{N_t,1} & \cdots & s_{N_t,N_r} \end{bmatrix}, \tag{4.2}$$

where $s_{i,j}$, $i = 1, \cdots, N_t$, $j = 1, \cdots, N_r$, are the information symbols taken from the constellation set $\mathcal{C} \in \{\text{QAM, PAM}\}$. In this chapter, a class of linear non-orthogonal STBCs that have full rate and full diversity is considered, such as perfect algebraic STBC [32, 36]. A perfect algebraic STBC codeword is an $N_t \times N_t$ matrix $\mathbf{X}$ whose entries are a linear combination of the input information signals. The spatial and temporal diversity of the codeword $\mathbf{X}$ is integrated into the space-time code design, as will be shown in the next sections.

The received signal matrix $\mathbf{Y} \in \mathbb{C}^{N_r \times KN_t}$ for the kth user is given by

$$\mathbf{Y} = \mathbf{HXV} + \mathbf{H} \sum_{j=1, j \neq k}^{K} \mathbf{X}_j \mathbf{V}_j + \mathbf{Z}, \tag{4.3}$$

where $\mathbf{V} \in \mathbb{C}^{N_t \times KN_t}$ is the orthogonal spreading matrix for user k, $\mathbf{Z} \in \mathbb{C}^{N_r \times KN_t}$ is an AWGN noise matrix. Note that the composite transmitted matrix is $\sum_{k=1}^{K} \mathbf{X}_k \mathbf{V}_k$.

In the orthogonal spreading code matrix, each user is assigned a unique orthogonal spreading code to separate the data of the users at the receivers. To eliminate CCI, the spreading code matrix has to obey the following conditions

$$\mathbf{V}_k \mathbf{V}_k^H = \mathbf{I}_{N_t}, \qquad k = 1, ..., K, \tag{4.4}$$

$$\mathbf{V}_j \mathbf{V}_k^H = 0, \quad k, j = 1, ..., K, \text{ and } j \neq k, \tag{4.5}$$

where $(\cdot)^H$ denotes the Hermitian operator. The orthogonal spreading code for each user can be constructed as a submatrix of the Hadamard matrix, or from a discrete Fourier transform (DFT) matrix. Hadamard matrices are of interest because of their simplicity [39]. The received signal matrix $\mathbf{Y}$ in (4.3) for the kth user is despread by multiplying it with $\mathbf{V}^H$, which yields

$$\hat{\mathbf{Y}} = \mathbf{Y}\mathbf{V}^H = \mathbf{H}\mathbf{X} + \hat{\mathbf{Z}}, \tag{4.6}$$

where

$$\hat{\mathbf{Y}} = \begin{bmatrix} \hat{y}_{1,1} & \cdots & \hat{y}_{1,N_t} \\ \vdots & \ddots & \vdots \\ \hat{y}_{N_r,1} & \cdots & \hat{y}_{N_r,N_t} \end{bmatrix}, \tag{4.7}$$

and

$$\hat{\mathbf{Z}} = \mathbf{Z}\mathbf{V}^H = \begin{bmatrix} \hat{z}_{1,1} & \cdots & \hat{z}_{1,N_t} \\ \vdots & \ddots & \vdots \\ \hat{z}_{N_r,1} & \cdots & \hat{z}_{N_r,N_t} \end{bmatrix}. \tag{4.8}$$

A brief review of the MU-MIMO high rate perfect algebraic STBC system is now presented, where the CSI is available only at the receiver.

4.3 Review of the Coherent Perfect STBC for MU-MIMO with Downlink Transmission

In this section, a coherent scheme is considered where the receiver knows the CSI. The scheme transmits data in linear combination over space and time. The design criterion of perfect STBC is to minimise the maximum pairwise error probability (PEP), where the ML detection might receive the distorted version $\hat{\mathbf{X}}$ of the original transmitted signal $\mathbf{X}$, and the PEP is given as [36, 53]

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \leq \frac{4^{rN_r}}{\left(\prod_{i=1}^{r} \lambda_i\right)^{N_r} \rho^{rN_r}}, \tag{4.9}$$

where r is the rank of the codeword difference matrix $(\mathbf{X} - \hat{\mathbf{X}})$, ρ is the signal-to-noise ratio (SNR) per receive antenna, $\lambda_i, i = 1, \cdots, r$, are the eigenvalues of $(\mathbf{X} - \hat{\mathbf{X}})(\mathbf{X} - \hat{\mathbf{X}})^H$, the minimum value of rN_r is the diversity gain, and the minimum value of $\left(\prod_{i=1}^{r} \lambda_i\right)^{1/r}$ is the coding gain.

4.3.1 Encoding of Coherent Perfect Algebraic STBC for MU-MIMO with Downlink Transmission

For coherent perfect STBC, the input symbol vectors, $\mathbf{s}_1, \cdots, \mathbf{s}_{N_r}$, are first rotated by the real or complex rotation matrix $\mathbf{M} \in \mathbb{C}^{N_t \times N_t}$ and then threaded into different layers l, where $l = 1, \cdots, N_t$. In other words, let $x_{1l}, x_{2l}, \cdots, x_{N_t l}$ be the symbols transmitted in the lth layer, i.e., [36]

$$\mathbf{x}_l = \mathbf{M}\mathbf{s}_l. \tag{4.10}$$

Thus, a layer can be viewed as an array of size $N_t \times N_t$. Any element of this array can be specified by two indices, (a, t), where a denotes the spatial domain and t denotes the temporal domain. Let l_i, $1 \leq i \leq N_t$ denote the ith layer. Hence, a layer can be formed such that

$$l_i = \left\{ \left(\left(t + i - 1 \right)_{N_t}, t \right) : \ 0 \leq t < N_t \right\}, \tag{4.11}$$

where $(x)_{N_t}$ denotes x modulo N_t operation. Accordingly, consecutive symbols from the same codeword are transmitted from different transmit antennas in different time slots. This method of transmission maximises the spatial and temporal diversity of the system.

The rotation matrix $\mathbf{M}$ (real or complex) is designed to maximise the distance between the symbol vectors to minimise the error rate and is constructed from an algebraic number field $\mathbb{Q}(\theta)$ of degree N_t generated by an algebraic number θ as in [54, 55].

The perfect algebraic STBC, as proposed in [36], is constructed based on cyclic division algebra theory for the special cases of $N_t = 2, 3, 4, 6$. To thread the symbols into the perfect algebraic STBC, the rotated symbol vectors are applied to the code block by

$$\mathbf{X} = \sum_{l=1}^{N_t} \text{diag}(\mathbf{M}\mathbf{s}_l) \cdot \mathbf{e}^{l-1}, \tag{4.12}$$

where $\text{diag}(\cdot)$ denotes the diagonal of a matrix, the threading matrix $\mathbf{e}$ is given as follows

$$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & \cdots & \gamma \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}, \tag{4.13}$$

and $\gamma = \sqrt{-1}$ is chosen by using Class Field Theory that ensures the transmitted code block has a non-vanishing determinant [36]. For multiple users equipped with different numbers of receive antennas, one important design parameter to consider is the number of threads l. Therefore, in this chapter, it is assumed the total number of layers is limited by the number of receive antennas N_r per user, i.e. $l = N_r$. Hence, the perfect STBC codeword can be rewritten as

$$\mathbf{X} = \sum_{l=1}^{N_r} \text{diag}(\mathbf{M}\mathbf{s}_l) \cdot \mathbf{e}^{l-1}. \tag{4.14}$$

Example 4. *For $N_t = 4$, number of users $K = 2$, User 1 equipped with $N_r = 3$ and User 2 equipped with $N_r = 1$. Then, the perfect algebraic STBC codeword for User 1 with $l = 3$ layers is in the form of*

$$\mathbf{X}_{4\times 4} = \begin{bmatrix} x_{11} & 0 & \gamma x_{13} & \gamma x_{12} \\ x_{22} & x_{21} & 0 & \gamma x_{23} \\ x_{33} & x_{32} & x_{31} & 0 \\ 0 & x_{43} & x_{42} & x_{41} \end{bmatrix},$$

of course, a higher rate code can be implemented by increasing the number of threads per user.

4.3.2 Decoding Coherent Perfect Algebraic STBC for MU-MIMO with Downlink Transmission

The sphere decoding approach is one of the most important decoding schemes for high data rate transmission systems over MIMO channels. The sphere decoder is basically a distance-based decoder that limits the number of possible codewords by considering only those codewords within a sphere centered at the received signal vector [56]. The kth user received spread signal can be expressed in terms of its vectorization as [35, 38]

$$\begin{aligned} vec\big(\hat{\mathbf{Y}}^T\big) &= vec\Big((\mathbf{H}\mathbf{X})^T\Big) + vec\big(\hat{\mathbf{Z}}^T\big) \\ &= \mathcal{B}_c\, vec(\mathbf{S}) + vec\big(\hat{\mathbf{Z}}^T\big), \end{aligned} \tag{4.15}$$

where

$$vec\big(\hat{\mathbf{Y}}^T\big) = [\hat{y}_{1,1}, \cdots, \hat{y}_{1,N_t}, \cdots, \hat{y}_{N_r,1}, \cdots, \hat{y}_{N_r,N_t}]^T,$$

$$vec\big(\hat{\mathbf{Z}}^T\big) = [\hat{z}_{1,1}, \cdots, \hat{z}_{1,N_t}, \cdots, \hat{z}_{N_r,1}, \cdots, \hat{z}_{N_r,N_t}]^T,$$

$$vec(\mathbf{S}) = [s_{1,1}, \cdots, s_{N_t,1}, \cdots, s_{1,N_r}, \cdots, s_{N_t,N_r}]^T,$$

and $\mathcal{B}_c$ is the new $N_tN_r \times N_tN_r$ effective channel matrix of the coherent perfect STBC which is given by

$$\mathcal{B}_c = \tilde{\mathbf{H}} \cdot \big(\mathbf{I}_{N_r} \otimes \mathbf{M}\big), \tag{4.16}$$

where the $N_tN_r \times N_tN_r$ matrix $\tilde{\mathbf{H}}$ is given by

$$\tilde{\mathbf{H}} = \begin{bmatrix} \text{diag}(\mathbf{h}_1) & \cdots & \big(\text{diag}(\mathbf{h}_1)\mathbf{e}^{N_r-1}\big)^T \\ \vdots & \ddots & \vdots \\ \text{diag}(\mathbf{h}_{N_r}) & \cdots & \big(\text{diag}(\mathbf{h}_{N_r})\mathbf{e}^{N_r-1}\big)^T \end{bmatrix}, \tag{4.17}$$

and $\otimes$ denotes the Kronecker matrix product. The underlying complex system in (4.15) can be converted into an equivalent real system by separating the real and imaginary parts of the received vector to define the following $2N_tN_r \times 1$ signal

$$\mathcal{Y} = \mathcal{H}_c\mathcal{S} + \mathcal{Z}, \tag{4.18}$$

where

$$\mathcal{Y} = \Big[\Re\big(vec(\hat{\mathbf{Y}}^T)\big) \quad \Im\big(vec(\hat{\mathbf{Y}}^T)\big)\Big]^T,$$

$$\mathcal{S} = \Big[\Re\big(vec(\mathbf{S})\big) \quad \Im\big(vec(\mathbf{S})\big)\Big]^T,$$

$$\mathcal{Z} = \Big[\Re\big(vec(\hat{\mathbf{Z}}^T)\big) \quad \Im\big(vec(\hat{\mathbf{Z}}^T)\big)\Big]^T,$$

and

$$\mathcal{H}_c = \begin{bmatrix} \Re\big(\mathcal{B}_c\big) & -\Im\big(\mathcal{B}_c\big) \\ \Im\big(\mathcal{B}_c\big) & \Re\big(\mathcal{B}_c\big) \end{bmatrix}.$$

In (4.18), a simple linear system of equation is constructed and can be decoded using the sphere decoder technique, which can be implemented to decode the kth user symbols $\hat{\mathcal{S}}$ such that

$$\hat{\mathcal{S}} = \arg \min_{\mathbf{S} \in \mathcal{C}^{\mathbf{n}}} \|\mathcal{Y} - \mathcal{H}_c\mathcal{S}\|^2, \tag{4.19}$$

where $n = 2 \times N_t \times N_r$. The differential STBC is now presented and then it is shown how to combine it with full-rate full-diversity perfect algebraic STBC through the use of the Cayley transform.

4.4 Differential Perfect STBC for MU-MIMO with Downlink Transmission

In this section, the differential encoding and decoding process for downlink transmission in a MU-MIMO system is discussed. In particular, this section demonstrates how to use a high rate space-time coding such as the perfect STBC with differential STBC for MU-MIMO systems. Here, it is assumed neither the transmitter nor the receiver has prior knowledge of the CSI. One method of implementing differential STBC with multiple antennas and a high data rate is to encode the transmitted data differentially using a CD transform, and to decode differentially without any knowledge of the CSI. The proposed work combines the perfect algebraic STBC with the CD transform that constructs full rate and full diversity differential STBC. In the following, a review of the differential STBC and the CD transform is presented and then the utilization of the CD transform with perfect algebraic STBC in a MU-MIMO framework is derived.

4.4.1 Differential STBC for MU-MIMO System

In differential STBC, the communications are done in blocks of N_t transmissions, which implies that the transmitted signal for any user k is an $N_t \times N_t$ matrix. The received despread signal block in (4.6) for the kth user at the τ-th block, $\tau = 0, ..., N$, can be re-expressed as

$$\hat{\mathbf{Y}}_\tau = \mathbf{H}\mathbf{X}_\tau + \hat{\mathbf{Z}}_\tau, \tag{4.20}$$

where $\hat{\mathbf{Y}}_\tau$, $\mathbf{X}_\tau$, and $\hat{\mathbf{Z}}_\tau$ are the despread received signal matrix, the transmitted perfect algebraic STBC matrix, and the despread noise matrix for the kth user at the τ-th block, respectively. The transmitted perfect algebraic STBC codeword matrix is encoded differentially as follows [28, 29]

$$\mathbf{X}_\tau = \mathbf{X}_{\tau-1}\mathbf{U}_{z_\tau}, \tag{4.21}$$

where $\mathbf{U}_{z_\tau}$ is a unitary data matrix utilized by the Cayley transform (it is specified $\mathbf{U}_{z_\tau}$ in the next subsection), $z_\tau \in \{0, \cdots, L-1\}$ is the transmitted data, and $\mathbf{X}_{\tau-1}$ is the transmitted matrix of the previous block. The transmitted matrix for the initial block of each user k is set to be identity, i.e. $\mathbf{X}_0 = \mathbf{I}_{N_t}$.

If it is assumed that the channels stay constant for two consecutive blocks, i.e. $\mathbf{H}_\tau =$

$\mathbf{H}_{\tau-1} = \mathbf{H}$, then (4.20) can be written as

$$\begin{aligned}\hat{\mathbf{Y}}_\tau &= \mathbf{H}\mathbf{X}_{\tau-1}\mathbf{U}_{z_\tau} + \hat{\mathbf{Z}}_\tau \\ &= \hat{\mathbf{Y}}_{\tau-1}\mathbf{U}_{z_\tau} + \hat{\mathbf{Z}}_\tau - \hat{\mathbf{Z}}_{\tau-1}\mathbf{U}_{z_\tau}.\end{aligned}$$

Therefore, the fundamental differential system equation for the kth user is given as

$$\hat{\mathbf{Y}}_\tau = \hat{\mathbf{Y}}_{\tau-1}\mathbf{U}_{z_\tau} + \hat{\mathbf{Z}}'_\tau, \tag{4.22}$$

where

$$\hat{\mathbf{Z}}'_\tau = \hat{\mathbf{Z}}_\tau - \hat{\mathbf{Z}}_{\tau-1}\mathbf{U}_{z_\tau}. \tag{4.23}$$

Since $\mathbf{U}_{z_\tau}$ is a unitary matrix, the entries of the additive noise term $\hat{\mathbf{Z}}'_\tau$ are i.i.d. $\mathcal{CN}(0,2)$. Thus, $\hat{\mathbf{Z}}'_\tau$ is statically independent of $\mathbf{U}_{z_\tau}$ and has twice the power. Therefore, for the kth user, the maximum-likelihood (ML) decoder of the differential STBC is

$$\hat{z}_\tau = \arg \max_{n=0,\cdots,L-1} \left\| \hat{\mathbf{Y}}_\tau - \hat{\mathbf{Y}}_{\tau-1}\mathbf{U}_n \right\|_F^2, \tag{4.24}$$

thus, the receiver does not need CSI to perform the decoding process. The PEP of transmitting $\mathbf{U}_n$ and mistakenly decoding $\mathbf{U}_{n'}$ has the following upper bound [28, 29]

$$P_e(\mathbf{U}_n \to \mathbf{U}_{n'}) \leq \frac{1}{2}\prod_{i=1}^{r}\left[1 + \frac{\rho^2}{4\big(1+2\rho\big)}\sigma_i^2\left(\mathbf{U}_n - \mathbf{U}_{n'}\right)\right]^{-N_r}, \tag{4.25}$$

where $\sigma_i(\cdot)$ denotes the ith singular value of the codeword difference matrix. Hence, good constellations $\mathbf{U}_1 \cdots \mathbf{U}_N$ have singular values

$$\sigma_i(\mathbf{U}_n - \mathbf{U}_{n'}), \qquad i = 1, \cdots, r, \tag{4.26}$$

that are as large as possible for $n \neq n'$. For large SNR, the probability of error depends dominantly on the product [17]

$$\prod_{i=1}^{r}\sigma_i(\mathbf{U}_n - \mathbf{U}_{n'}) = \left|\det\left(\mathbf{U}_n - \mathbf{U}_{n'}\right)\right| \tag{4.27}$$

Therefore, at high SNRs, the one inside brackets in (4.25) is neglected which implies the fol-

lowing upper bound based on the nonzero singular values, see eq.(9) in [18]

$$P_e(\mathbf{U}_n \to \mathbf{U}_{n'}) \lesssim 8^{rN_r} \cdot \frac{\rho^{-rN_r}}{|\det(\mathbf{U}_n - \mathbf{U}_{n'})|^{2N_r}}. \tag{4.28}$$

The diversity gain is defined to be G_d and the coding gain to be G_c. The diversity gain is given by

$$G_d = rN_r. \tag{4.29}$$

The differential STBC also achieves full diversity order of N_tN_r if the unitary matrix is fully diverse. Using (4.29) and the right-hand side of (4.28), it follows

$$\begin{aligned} & P_e(\mathbf{U}_n \to \mathbf{U}_{n'}) \\ & \lesssim 8^{G_d} |\det(\mathbf{U}_n - \mathbf{U}_{n'})|^{-2N_r} \rho^{-G_d} \\ & \lesssim \left[\left(8^{G_d} |\det(\mathbf{U}_n - \mathbf{U}_{n'})|^{-2N_r} \right)^{-G_d^{-1}} \right]^{-G_d} \cdot \rho^{-G_d} \\ & \lesssim \left(G_c \cdot \rho \right)^{-G_d}, \end{aligned} \tag{4.30}$$

where

$$G_c = \left(8^{G_d} |\det(\mathbf{U}_n - \mathbf{U}_{n'})|^{-2N_r} \right)^{-G_d^{-1}}. \tag{4.31}$$

By using (4.29) in (4.31), it follows

$$G_c \approx |\det(\mathbf{U}_n - \mathbf{U}_{n'})|^{\frac{2}{r}}. \tag{4.32}$$

The PEP will be lower in the case that it receives multiple replicas of the signal using diversity. In other words, diversity is the slope of the error probability curve in terms of the received SNR in a log-log scale. In this case, taking the $\log$ for both sides of (4.30) implies that

$$\log(P_e) = -G_d \big[\log(G_c) + \log(\rho) \big], \tag{4.33}$$

or more explicitly

$$\log(G_c) = \frac{\log(P_e)}{-G_d} - \log(\rho). \tag{4.34}$$

This coding gain ratio is a measure of the worst case separation between encoded symbols. It therefore determines the worst case for PEP, and hence the block error rate. The differential STBC also achieves full diversity order of N_tN_r if the unitary matrix is fully diverse, i.e.,

$r = N_t$. Therefore, to minimise the PEP, the following conditions should be satisfied [36, 53]:

- In order to maximise the diversity gain, the rank criterion r of $(\mathbf{U}_n - \mathbf{U}_{n'})$ should be maximised.
- In order to maximise the coding gain G_c, the minimum determinant of $(\mathbf{U}_n - \mathbf{U}_{n'})$ should be maximised.
- Non-vanishing minimum determinant on the coding gain.

4.4.2 Differential Perfect Algebraic STBC for a MU-MIMO System

For MU-MIMO differential transmission schemes, the information must first be encoded in a unitary matrix to ensure the same transmit power for different blocks. This can be achieved by applying the Cayley transform as [18]

$$\mathbf{U}_{z_\tau} = (\mathbf{I}_{N_t} - j\mathbf{A}_\tau)(\mathbf{I}_{N_t} + j\mathbf{A}_\tau)^{-1}, \tag{4.35}$$

where $j = \sqrt{-1}$, and $\mathbf{A}_\tau$ is an $N_t \times N_t$ Hermitian matrix at block time τ, (the subscript on $\mathbf{A}$ from now on is dropped for simplicity). As proposed in [18], the output of the Caylay transform is unitary if, and only if, $\mathbf{A}$ is a Hermitian matrix. Therefore, the Hermitian property must be ensured for the transmitted perfect STBC signals. Furthermore, according to [18, 35], the Hermitian constraints require real constellations and real rotation matrices to maintain the Hermitian property for matrix $\mathbf{A}$.

For differential perfect algebraic STBC, the input symbol vectors, $\mathbf{s}_1, \cdots, \mathbf{s}_{N_r}$, are first rotated by the real rotation matrix $\mathbf{M} \in \mathbb{R}^{N_t \times N_t}$ and then threaded differentially into different layers l, where $l = 1, \cdots, N_t$. Let l_i, $1 \leq i \leq N_t$ denote the ith layer. Hence, a layer can be formed such that [35]

$$l_i = \left\{ \left((N_t - i - t)_{N_t}, t \right) : 0 \leq t < N_t \right\}. \tag{4.36}$$

Therefore, the placement of the real rotated symbols into the code block in differential threading is very similar to the case for coherent encoding, but reversed. Then, the differential perfect algebraic STBC codeword can be expressed as

$$\mathbf{X} = \sum_{l=1}^{N_r} \text{diag}(\mathbf{M}\mathbf{s}_l) \cdot \mathbf{a} \cdot \left(\mathbf{e}^{l-1}\right)^T, \tag{4.37}$$

where $\mathbf{e}$ is defined in (4.13) but with differential case, here $\gamma = 1$ is used,

$$\mathbf{a} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 1 & 0 & 0 & 0 \\ 1 & \cdots & 0 & 0 & 0 \end{bmatrix}, \tag{4.38}$$

and $\mathbf{M} \in \mathbb{R}^{N_t \times N_t}$ is a real rotation matrix as in [35, 54, 55]. The code block generated in the differential perfect algebraic STBC case must be Hermitian. Then the Hermitian conversion for the matrix $\mathbf{A}$ in differential perfect STBC can be written as [35]

$$\mathbf{A} \triangleq \frac{1}{\sqrt{2}} \Bigg(\Big(j \cdot triu(\mathbf{X}) + (tril(\mathbf{X}))^H \Big) + \Big((-j \cdot triu(\mathbf{X}))^H + (tril(\mathbf{X})) \Big) \Bigg) + \text{diag}\big(\text{diag}(\mathbf{X})\big), \tag{4.39}$$

where $triu(\cdot)$ denotes the $N_t \times N_t$ matrix that contains only the above diagonal elements of $\mathbf{X}$, and $tril(\cdot)$ denotes the $N_t \times N_t$ matrix that contains only the below diagonal elements of $\mathbf{X}$. Therefore, if the input matrix $\mathbf{A}$ is Hermitian, the Cayley transformed matrix in (4.35) will be unitary. Furthermore, since the differential perfect algebraic STBC requires multiplying the new code block by the previous block, the resulting new transmitted output in (4.21) remains unitary.

Example 5. *For $N_t = 4$, $N_r = 4$, and $K = 1$, thus $l = 4$ layers. Then, the Hermitian matrix $\mathbf{A}$ is in the form of*

$$\mathbf{A} = \begin{bmatrix} x_{14} & \frac{x_{23}+jx_{13}}{\sqrt{2}} & \frac{x_{32}+jx_{12}}{\sqrt{2}} & \frac{x_{41}+jx_{11}}{\sqrt{2}} \\ \frac{x_{23}-jx_{13}}{\sqrt{2}} & x_{22} & \frac{x_{31}+jx_{21}}{\sqrt{2}} & \frac{x_{44}+jx_{24}}{\sqrt{2}} \\ \frac{x_{32}-jx_{12}}{\sqrt{2}} & \frac{x_{31}-jx_{21}}{\sqrt{2}} & x_{34} & \frac{x_{43}+jx_{33}}{\sqrt{2}} \\ \frac{x_{41}-jx_{11}}{\sqrt{2}} & \frac{x_{44}-jx_{24}}{\sqrt{2}} & \frac{x_{43}-jx_{33}}{\sqrt{2}} & x_{42} \end{bmatrix}.$$

Therefore, with this formulation, given the invertible equivalent channel matrix and the transmitted codeword block $\mathbf{X}$, it is easy to determine the input symbols, by using the Hermitian matrix $\mathbf{A}$ as a roadmap. The matrix $\mathbf{A}$ points out the elements of the $\mathbf{X}$ matrix that include each symbol, and they can be scaled and summed to form the best estimate of the input symbol. For

example, from position (4,1) and (1,4) in matrix $\mathbf{A}$, thus

$$a_{41} = \frac{x_{41} - jx_{11}}{\sqrt{2}}, \qquad a_{14} = \frac{x_{41} + jx_{11}}{\sqrt{2}},$$

and these are the only symbols involved in these positions from matrix $\mathbf{X}$. Accordingly, the original input symbols from the transmitted codeword block $\mathbf{X}$ are as follows

$$x_{41} = \frac{1}{\sqrt{2}} \left\{a_{41} + a_{14}\right\}, \qquad x_{11} = \frac{1}{\sqrt{2}} \Im \left\{a_{14} - a_{41}\right\}.$$

4.4.3 Decoding the Differential Perfect Algebraic STBC for a MU-MIMO System

For the MU-MIMO downlink system, the differential transmissions are implemented in blocks, in which each user k receives the sum of all the transmit waveforms of other users; then the received signal blocks for each user must be detected independently. Thus, if $\mathbf{G}$ denotes the matrix having all $N+1$ received signal blocks for the kth user, i.e.,

$$\mathbf{G} = \left[\hat{\mathbf{Y}}_0 \; \hat{\mathbf{Y}}_{\tau-1} \; \hat{\mathbf{Y}}_\tau \; \cdots \; \hat{\mathbf{Y}}_N\right]. \tag{4.40}$$

When encoding using (4.21), the decoding process for $\mathbf{X}_\tau$ for the kth user would be according to the last two blocks of $\mathbf{G}$ as in the following notation

$$\mathbf{G} = \left[\underbrace{\hat{\mathbf{Y}}_0\hat{\mathbf{Y}}_1} \cdots \underbrace{\hat{\mathbf{Y}}_{\tau-1}\hat{\mathbf{Y}}_\tau} \cdots \underbrace{\hat{\mathbf{Y}}_{N-1}\hat{\mathbf{Y}}_N}\right]. \tag{4.41}$$

Then the combined information between the unitary matrix $\mathbf{U}_{z_\tau}$ and the received signal blocks $(\hat{\mathbf{Y}}_{\tau-1}, \hat{\mathbf{Y}}_\tau)$ in the differential scheme at the kth user can be expressed as

$$\begin{bmatrix} \hat{\mathbf{Y}}_{\tau-1} \\ \hat{\mathbf{Y}}_\tau \end{bmatrix} = \mathbf{H} \begin{bmatrix} \mathbf{X}_{\tau-1} \\ \mathbf{X}_{\tau-1}\mathbf{U}_{z_\tau} \end{bmatrix} + \begin{bmatrix} \hat{\mathbf{Z}}_{\tau-1} \\ \hat{\mathbf{Z}}_\tau \end{bmatrix}. \tag{4.42}$$

For differential perfect algebraic STBC encoding, it is assumed that for any user k the channel matrix $\mathbf{H}$ changes slowly (channel coherence time is large enough) and extends over several matrix transmission periods. In such a case, the base station transmission starts with a reference matrix $\mathbf{X}_0$, followed by several information matrices. The Hermitian matrix $\mathbf{A}$ is used to form an equivalent channel model for differential decoding. An easier way to represent this model is

to rewrite the differential receiver equation using the Cayley transform [18]

$$\begin{aligned}\hat{\mathbf{Y}}_\tau &= \hat{\mathbf{Y}}_{\tau-1}\mathbf{U}_{z_\tau} + \hat{\mathbf{Z}}_\tau - \hat{\mathbf{Z}}_{\tau-1}\mathbf{U}_{z_\tau} \\ &= \hat{\mathbf{Y}}_{\tau-1}(\mathbf{I}_{N_t} - j\mathbf{A})(\mathbf{I}_{N_t} + j\mathbf{A})^{-1} + \hat{\mathbf{Z}}_\tau \\ &\quad - \hat{\mathbf{Z}}_{\tau-1}(\mathbf{I}_{N_t} + j\mathbf{A})^{-1}(\mathbf{I}_{N_t} - j\mathbf{A}).\end{aligned}$$

By multiplying both sides by $(\mathbf{I}_{N_t} + j\mathbf{A})$, I have

$$\begin{aligned}\hat{\mathbf{Y}}_\tau(\mathbf{I}_{N_t} + j\mathbf{A}) = \hat{\mathbf{Y}}_{\tau-1}(\mathbf{I}_{N_t} - j\mathbf{A}) + \hat{\mathbf{Z}}_\tau(\mathbf{I}_{N_t} + j\mathbf{A}) \\ - \hat{\mathbf{Z}}_{\tau-1}(\mathbf{I}_{N_t} - j\mathbf{A}),\end{aligned}$$

which can be simplified as

$$\begin{aligned}\hat{\mathbf{Y}}_\tau - \hat{\mathbf{Y}}_{\tau-1} = -j\,(\hat{\mathbf{Y}}_\tau + \hat{\mathbf{Y}}_{\tau-1})\mathbf{A} + \hat{\mathbf{Z}}_\tau(\mathbf{I}_{N_t} + j\mathbf{A}) \\ - \hat{\mathbf{Z}}_{\tau-1}(\mathbf{I}_{N_t} - j\mathbf{A}).\end{aligned} \tag{4.43}$$

Note that due the differential detection with matrix $\mathbf{A}$ as a unitary Cayley transform, the additive noise in (4.43) has the covariance

$$2(\mathbf{I}_{N_t} + j\mathbf{A})(\mathbf{I}_{N_t} - j\mathbf{A}) = 2(\mathbf{I}_{N_t} + \mathbf{A}^2), \tag{4.44}$$

which results in some performance degradation. Then, the ML decoder can be given as [18]

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}_1 \cdots \mathbf{s}_{N_T}} \left\| \left(\hat{\mathbf{Y}}_\tau - \hat{\mathbf{Y}}_{\tau-1}\right) - \left(\frac{1}{j}(\hat{\mathbf{Y}}_\tau + \hat{\mathbf{Y}}_{\tau-1})\mathbf{A}\right) \right\|^2. \tag{4.45}$$

To find the ML solution vectors without an exhaustive search, the sphere decoding method is used, as it considers only a small set of vectors rather than all possible transmitted signal vectors. The sphere decoder representation is obtained by constructing an equivalent channel model for the differential system equation in (4.43). Let $\mathbf{C} = \hat{\mathbf{Y}}_\tau - \hat{\mathbf{Y}}_{\tau-1}$, and $\mathbf{B} = -j\,(\hat{\mathbf{Y}}_\tau + \hat{\mathbf{Y}}_{\tau-1})$, then the differential equivalent channel is

$$\mathbf{C} = \mathbf{B}\mathbf{A} + \hat{\mathbf{Z}}_d, \tag{4.46}$$

where $\hat{\mathbf{Z}}_d = \hat{\mathbf{Z}}_\tau(\mathbf{I}_{N_t} + j\mathbf{A}) - \hat{\mathbf{Z}}_{\tau-1}(\mathbf{I}_{N_t} - j\mathbf{A})$ is the additive Gausssian noise with zero mean and covariance $2(\mathbf{I}_{N_t} + \mathbf{A}^2)$. Now, the received spread signal for the kth user is vectorized as

$$\begin{aligned} vec\big(\mathbf{C}^T\big) &= vec\Big((\mathbf{BA})^T\Big) + vec\big(\hat{\mathbf{Z}}_d^T\big) \\ &= \mathcal{B}_d\, vec(\mathbf{S}) + vec\big(\hat{\mathbf{Z}}_d^T\big), \end{aligned} \tag{4.47}$$

where $\mathcal{B}_d$ is the new $N_tN_r \times N_tN_r$ effective channel matrix of the differential perfect STBC which is given by

$$\mathcal{B}_d = \tilde{\mathbf{B}}\tilde{\mathbf{A}} \cdot \big(\mathbf{I}_{N_r} \otimes \mathbf{M}\big), \tag{4.48}$$

where the $N_tN_r \times N_tN_r$ matrix $\tilde{\mathbf{B}}$ is given by

$$\tilde{\mathbf{B}} = \begin{bmatrix} \big(\text{diag}(\mathbf{b}_1)\mathbf{a}\big)^T & \cdots & \big(\text{diag}(\mathbf{b}_1)\mathbf{a} \cdot \mathbf{e}^{N_r-1}\big)^T \\ \vdots & \ddots & \vdots \\ & & \\ \big(\text{diag}(\mathbf{b}_{N_r})\mathbf{a}\big)^T & \cdots & \big(\text{diag}(\mathbf{b}_{N_r})\mathbf{a} \cdot \mathbf{e}^{N_r-1}\big)^T \end{bmatrix}. \tag{4.49}$$

The $N_tN_r \times N_tN_r$ block diagonal matrix $\tilde{\mathbf{A}}$ is in the form of

$$\tilde{\mathbf{A}} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_{N_r} \end{bmatrix}, \tag{4.50}$$

where $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{N_r}$ are the scaled submatrices of the original Hermitian matrix $\mathbf{A}$, and each is of size $N_t \times N_t$. To define the block diagonal matrix $\tilde{\mathbf{A}}$ completely, an example is given here.

Example 6. *By using the same entities as in Example 5. Then, the submatrices of the* 16×16 *block diagonal matrix* $\tilde{\mathbf{A}}$ *are in the form*

$$\mathbf{A}_1 = \begin{bmatrix} j & 0 & 0 & 1 \\ 0 & j & 1 & 0 \\ 0 & -j & 1 & 0 \\ -j & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} j & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -j & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{A}_3 = \begin{bmatrix} j & 1 & 0 & 0 \\ -j & 1 & 0 & 0 \\ 0 & 0 & j & 1 \\ 0 & 0 & -j & 1 \end{bmatrix}, \quad \mathbf{A}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & j & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -j & 0 & 1 \end{bmatrix}.$$

Now, the complex received vector in (4.47) is converted to its equivalent real and imaginary parts, i.e,

$$\mathcal{R} = \mathcal{H}_d \mathcal{S} + \mathcal{Z}_d, \tag{4.51}$$

where

$$\mathcal{R} = \Big[\Re\big(vec(\mathbf{C}^T)\big) \quad \Im\big(vec(\mathbf{C}^T)\big)\Big]^T,$$

$$\mathcal{S} = \Big[\Re\big(vec(\mathbf{S})\big) \quad \Im\big(vec(\mathbf{S})\big)\Big]^T,$$

$$\mathcal{Z}_d = \Big[\Re\big(vec(\hat{\mathbf{Z}}_d^T)\big) \quad \Im\big(vec(\hat{\mathbf{Z}}_d^T)\big)\Big]^T,$$

and

$$\mathcal{H}_d = \begin{bmatrix} \Re\big(\mathcal{B}_d\big) & -\Im\big(\mathcal{B}_d\big) \\ \Im\big(\mathcal{B}_d\big) & \Re\big(\mathcal{B}_d\big) \end{bmatrix}.$$

The sphere decoder can be implemented to decode the kth user symbols $\hat{\mathcal{S}}$ such that

$$\hat{\mathcal{S}} = \arg \min_{\mathbf{S} \in \mathcal{C}^{\mathbf{n}}} \|\mathcal{R} - \mathcal{H}_d \mathcal{S}\|^2. \tag{4.52}$$

4.5 Computational Complexity and Rate Analysis

The matrix $\mathcal{H}_d$ in (4.51) has the size of $2N_tN_r \times 2N_tN_r$, thus the system has $2N_tN_r$ equations and $2N_tN_r$ unknowns. The sphere decoder usually benefits from having more equations and less unknowns because the computational complexity is polynomial, yet goes exponential when the difference between the number of equations and unknowns grows large. To allow for a low-complexity decoder and to have at least as many equations as unknowns when $N_t \geq N_r$, the number of threads l per block per user is constrained by [34]

$$l \leq \min\big(N_t, N_r\big), \tag{4.53}$$

and since the number of symbols per block per user is $q = N_t l$. Hence, in this chapter, the following constraint is imposed

$$q \leq \min\left(N_t^2, N_t N_r\right). \tag{4.54}$$

In this case, the maximum rate of the code essentially depends on the number of threads l per block per user, the total number of q symbols per block per user sent in that thread, the cardinality of constellation L, and the orthogonal spreading code period per user. Since the channel is used N_t times, the system transmission rate per channel per user is

$$R = \frac{q}{K \cdot N_t} \cdot \log_2(L) \qquad bits/sec/Hz.. \tag{4.55}$$

There are K users in the system, each transmitting q symbols per block. Therefore, the total bit rate per system is

$$R = \frac{q}{N_t} \cdot \log_2(L) \qquad bits/sec/Hz. \tag{4.56}$$

Note that the rate is independent of the number of users. Through a wise choice of the number of threads per block $l \leq \min\left(N_t, N_r\right)$, systems that achieve this transmission rate will have full rate and full diversity [32].

4.5.1 Rate Analysis

As discussed earlier, the differential perfect algebraic scheme achieves full diversity full rate over a MU-MIMO channel where at different time slots and different antennas, different symbols are transmitted. Table 4.1 briefly summarises and compares the rate parameters of differential perfect algebraic STBC with other practical STBC schemes that offer reasonable data rates and diversity such as differential Alamouti code (G2-STBC) [39], and differential quasi-orthogonal code (QO-STBC) [4]. In terms of the MIMO's diversity feature shown in Table 4.1, the three MIMO schemes of Algebraic, Alamouti, and quasi-orthogonal STBCs are capable of attaining the full diversity order of $N_t N_r$, which minimises the PEP of (4.25) according to its rank criterion. With regards to the transmission rate as seen in Table 4.1, differential perfect algebraic STBC introduced in this chapter is capable of achieving the full MIMO transmission rate, provided that the parameters satisfy $q = N_t N_r$, which results in a maximised rate gain of $R = \frac{q}{N_t} \cdot \log_2(L)$. In the other STBCs shown in Table 4.1, every element of a codeword matrix is a linear combination of the input symbols and limited by a fixed number of transmit

Table 4.1. The rate and diversity parameters of classic STBCs representatives.

STBC Scheme	Parameters				
Algebraic perfect	$N_t > 1$	$N_r \geq 1$	$G_d = N_t N_r$	$q = N_t N_r$	$R = \frac{q}{N_t} \cdot \log_2(L)$
G2-STBC [39]	$N_t = 2$	$N_r \geq 1$	$G_d = N_t N_r$	$q = 2$	$R = \log_2(L)$
QO-STBC [4]	$N_t = 4$	$N_r \geq 1$	$G_d = N_t N_r$	$q = 4$	$R = \log_2(L)$

antennas, i.e., $N_t = 2$ or $N_t = 4$. The number of symbols is selected such that an orthogonal STBC is feasible. Such a limit on the number of symbols is not necessary if the orthogonality condition of the STBC is relaxed as in differential perfect algebraic STBC. For example, with $N_t = 2$, $N_r = 2$, and 8-PAM; rates for differential perfect algebraic STBC and G2-STBC, are 6, 3, respectively. Similarly, with $N_t = 4$, $N_r = 4$, and 8-PAM; rates for differential perfect algebraic STBC and QO-STBC, are 12, 3, respectively. That shows the difference in the rate.

4.5.2 Complexity Analysis

To bring more insight on the computational complexity, the notion of flops is introduced in this section, where flops denote the floating point operations (FLOPs). The total number of FLOPs is used to measure the computational complexity of different schemes. The total FLOPs needed for the matrix operations is summarised below [49, 57]:

- Multiplication of $m \times n$ and $n \times p$ complex matrices: $8mnp - 2mp$;
- QR decomposition of an $m \times n$ $(m \leq n)$ complex matrix: $16(n^2 m - nm^2 + \frac{1}{3}m^3)$;
- SVD of an $m \times n$ $(m \leq n)$ complex matrix where only $\mathbf{\Sigma}$ and $\mathbf{V}$ are obtained: $32(nm^2 + 2m^3)$;
- SVD of an $m \times n$ $(m \leq n)$ complex matrix where $\mathbf{U}$, $\mathbf{\Sigma}$ and $\mathbf{V}$ are obtained: $8(4n^2 m + 8nm^2 + 9m^3)$;
- Inversion of an $m \times m$ real matrix using Gauss-Jordan elimination: $4m^3/3$.

Table 4.2. Computational complexity of coherent perfect algebraic STBC

Steps	Operation	Flops	Case $(2,2,2)\times 6$
1	$\sum_{l=1}^{N_r} \text{diag}(\mathbf{M}\mathbf{s}_l)\cdot \mathbf{e}^{l-1}$	$\mathcal{O}\big(KN_r(16N_t^3-2N_t^2)\big)$	20304
2	$\mathbf{HXV}$	$\mathcal{O}\big(K(16KN_t^2N_r - 2KN_tN_r)\big)$	10152
3	$\mathbf{YV}^H$	$\mathcal{O}\big(K(8KN_t^2N_r - 2N_tN_r)\big)$	5112
4	$\mathbf{I}_{N_r}\otimes\mathbf{M}$	$\mathcal{O}\big(K(N_t^2N_r^2)\big)$	432
5	$\tilde{\mathbf{H}}\cdot\big(\mathbf{I}_{N_r}\otimes\mathbf{M}\big)$	$\mathcal{O}K(8N_t^3N_r^3-2N_t^2N_r^2)$	40608
			Total=76608

Table 4.3. Computational complexity of the proposed differential perfect STBC

Steps	Operation	Flops	Case $(2,2,2)\times 6$
1	$\sum_{l=1}^{N_r} \text{diag}(\mathbf{M}\mathbf{s}_l)\cdot \mathbf{a}\cdot \left(\mathbf{e}^{l-1}\right)^T$	$\mathcal{O}\big(2KN_r(24N_t^3-2N_t^2)\big)$	61344
2	$\mathbf{HXUV}$	$\mathcal{O}\big(K(24KN_t^3-2KN_t^2)\big)$	46008
3	$\mathbf{YV}^H$	$\mathcal{O}\big(K(8KN_t^2N_r-2N_tN_r)\big)$	5112
4	$\mathbf{I}_{N_r}\otimes\mathbf{M}$	$\mathcal{O}\big(K(N_t^2N_r^2)\big)$	432
5	$\tilde{\mathbf{B}}\tilde{\mathbf{A}}\cdot\big(\mathbf{I}_{N_r}\otimes\mathbf{M}\big)$	$\mathcal{O}K(16N_t^3N_r^3-2N_t^2N_r^2)$	82080
			Total=194976

Table 4.4. Computational complexity of differential Alamouti STBC.

Steps	Operation	Flops	Case $(2,2,2)\times 6$
1	$\bar{\mathbf{H}}^\dagger$	$\mathcal{O}\big(K(\frac{4}{3}\bar{n}_r^3+16\bar{n}_r^2n_t-2N_t\bar{n}_r)\big)$	4720
2	$(\mathbf{I}-\mathbf{H}^\dagger\bar{\mathbf{H}})\mathbf{\Phi}$	$\mathcal{O}\big(K(8N_t^2\bar{n}_r+14N_t^2-4N_t)\big)$	4896
3	$\mathbf{QR}\big((\mathbf{I}-\bar{\mathbf{H}}^\dagger\bar{\mathbf{H}})\big)$	$\mathcal{O}\big(K(\frac{16}{3}N_t^3)\big)$	3456
4	$\mathbf{SVD}\big(\mathbf{H}(\mathbf{I}-\bar{\mathbf{H}}^\dagger\bar{\mathbf{H}})\big)$	$\mathcal{O}\big(K(64N_r^3+8N_t^2N_r+32N_tN_r^2-2N_tN_r)\big)$	5496
5	$\mathbf{HFX}$	$\mathcal{O}\big(K(16N_tN_r+24N_r)\big)$	720
			Total=19288

Table 4.2, Table 4.3, and Table 4.4 show the operations and the required FLOPs for the algorithms of the coherent perfect algebraic STBC, the differential perfect algebraic STBC, and the differential G2-STBC in [39], respectively. For illustration, it is assumed that the system has $K = 3$ users, each user with $N_r = 2$ receive antennas, and $N_t = 6$ transmit antennas; this scenario is denoted as $(2, 2, 2) \times 6$. For simplicity, and without loss of generality, it is also assume that all users have the same number of receive antennas. Note that $\bar{n}_r = \sum_{j=1, j\neq k}^{K} N_r^j$. Clearly, the proposed differential perfect algebraic STBC scheme requires the highest complexity.

Furthermore, the cases of the computational complexity of the system dimensions are shown in Fig. 4.2 and Fig. 4.3. First, in Fig. 4.2, the number of receive antennas for each user is set at $N_r = 2$ and the number of users K is increased. Similarly, in Fig. 4.3, the number of users is fixed to be $K = 4$ while the number of receive antennas for each user is increased gradually. From both figures, the computational complexity of the proposed system is higher and increases exponentially. The reason is that, the differential perfect algebraic STBC scheme requires higher rate and as a result the number of antennas increases exponentially and thus the size of unknown variables for the equivalent channel matrix equation also increases exponentially. It is also observed that varying the number of receive antennas has much higher impact on the complexity than varying the number of users. Therefore, for a high rate system that supports different type of terminals, it is better to keep the number of receive antennas for each terminal as low as possible.

As shown above, it is worth noting that the perfect STBC combined with differential STBC scheme proposed in this chapter relaxes the orthogonality conditions of the standards orthogonal STBC codes such as G2-STBC and QO-STBC code. Therefore, in this chapter, the number of transmit symbols per block in the downlink is much higher and that will result in an increase of the overall rate of the system. Thus the system can transmit and receive in high rate without needing the CSI.

4.6 Simulation Results and Discussion

In this section, the performance of the differential perfect algebraic space-time modulation scheme for the MU-MIMO downlink transmission is examined. In this section, the channel is modeled as quasi-static, where the fading block matrix between the transmitter and receiver is constant (but unknown) between two successive channel uses. The SNR per user is defined as SNR=$N_r\rho$. The Monte Carlo simulation is used to evaluate the performance in terms of the block error rate (BLER) and BER.

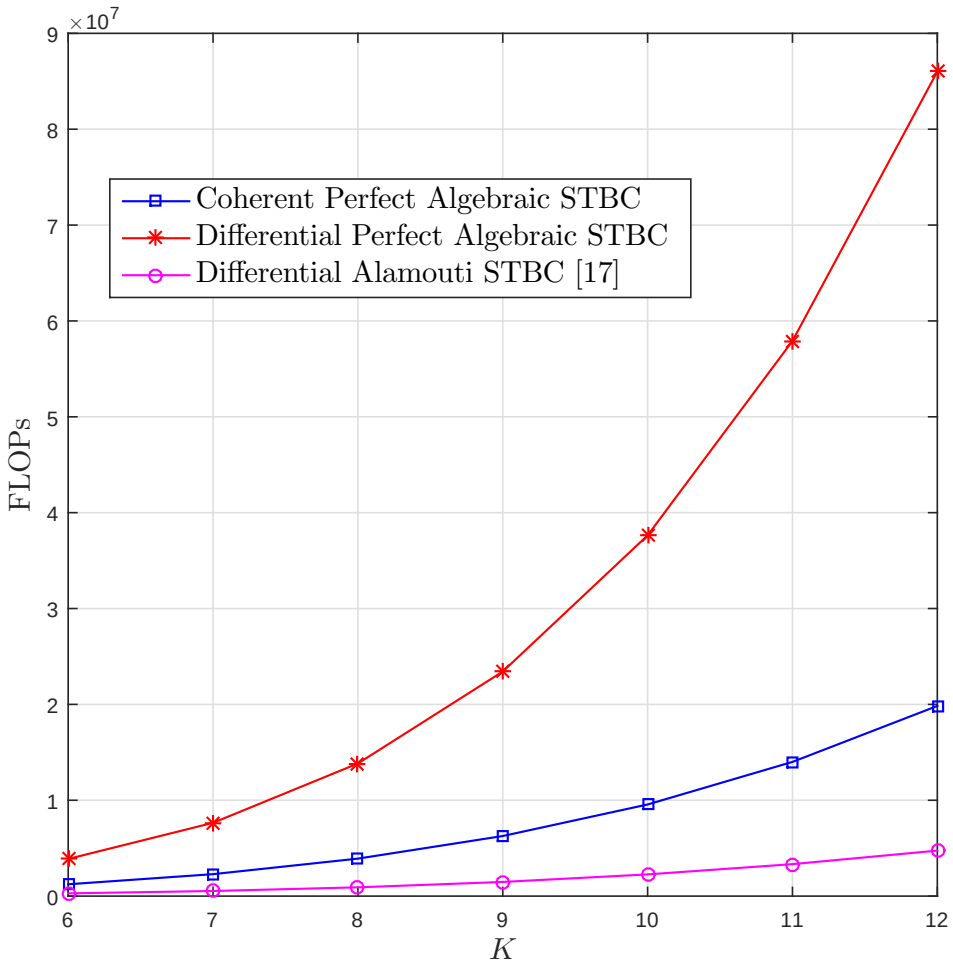

Figure 4.2. Comparison of the computational complexity for differential perfect algebraic, coherent perfect algebraic, and differential Alamouti with $N_r = 2$, and $N_t = K \times N_r$.

The Differential Perfect Algebraic STBC with Multipe Users: In Fig. 4.4 the BLER performance curve is first simulated and plotted for one, two, four, and eight users system. Each user has two receive antennas and 4-PAM symbols are used. The base station has four transmit antennas. In this case, each STBC block has $q = 8$ symbols per user with $R = 4$. The transmitted codeword X for each user consists of two layers 4×4 perfect algebraic STBC , i.e, $l = N_r = 2$. The 4×4 rotation matrix M is given in [54, 55]. Hadamard matrices are used as the orthogonal spreading matrices to cancel the CCI. It is shown that the MU-MIMO system for all cases (e.g., in the case of one, two, four, and eight users) achieves the same performance as a single-user MIMO system; that is, on multiple users, the orthogonal spreading codes are allowed to eliminate CCI. This results in every user being processed as if it was a single-user case, so that the results for every user are identical and the CCI is completely eliminated and full rate full diversity is achieved with the differential perfect algebraic STBC.

Diversity Gain, Coding Gain, and Rates : In Fig. 4.4, the slope of the BLER curves for

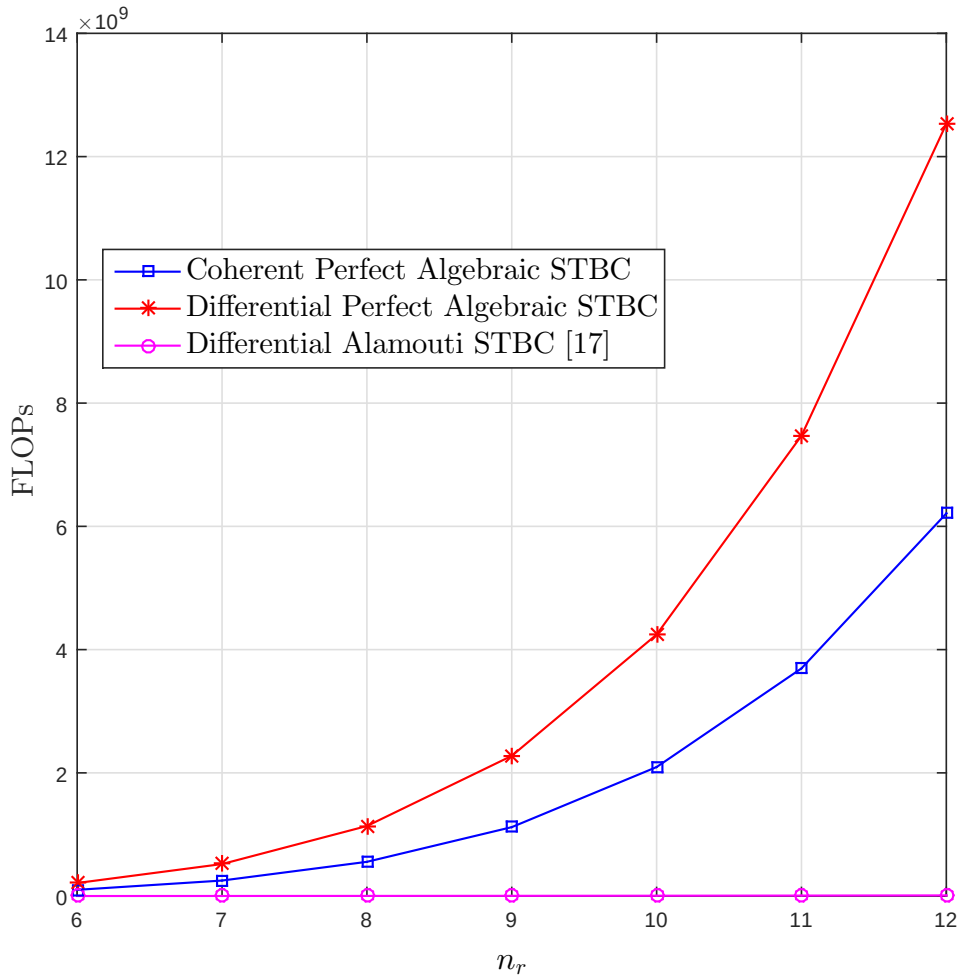

Figure 4.3. Comparison of the computational complexity for differential perfect algebraic, coherent perfect algebraic, and differential Alamouti with $K = 4$, and $N_t = K \times N_r$.

high SNR converges to the slope determined by the diversity. As shown in (4.29), this slope is $-rN_r$ (in a log-log scale) where r is the rank of the codeword difference matrix. For differential algebraic STBC codes, the number of threads l determines the rank r, and this is limited by the number of receive antennas, N_r, if $N_r \leq N_t$. Hence the diversity slope is $-N_r^2$. Further, $-rN_r$ is also related to the number of symbols encoded in each codeword block over symbol time periods. Thus the number of symbols per block is N_tN_r. For the case of Fig. 4.4, the system has $l = 2$ threads and 4 symbol time periods in the 4×4 algebraic codeword. Each thread has 4 symbols and only 2 symbols are encoded and transmitted in any one symbol time period. Clearly, when 2 threads are populated, 8 entries of the 4×4 code block are populated and the other 8 are filled with zeros. Therefore, the rank of the codeword difference matrix $r = 2$ for this approach of coding. Thus $G_d = rN_r = 4$. In Fig. 4.4, the diversity line based on the actual BLER curves is plotted using the estimated lower bounds formula in (4.34) of the coding gain G_c. The BLER curves appear to approach a slope of -4 asymptotically and the diversity gain

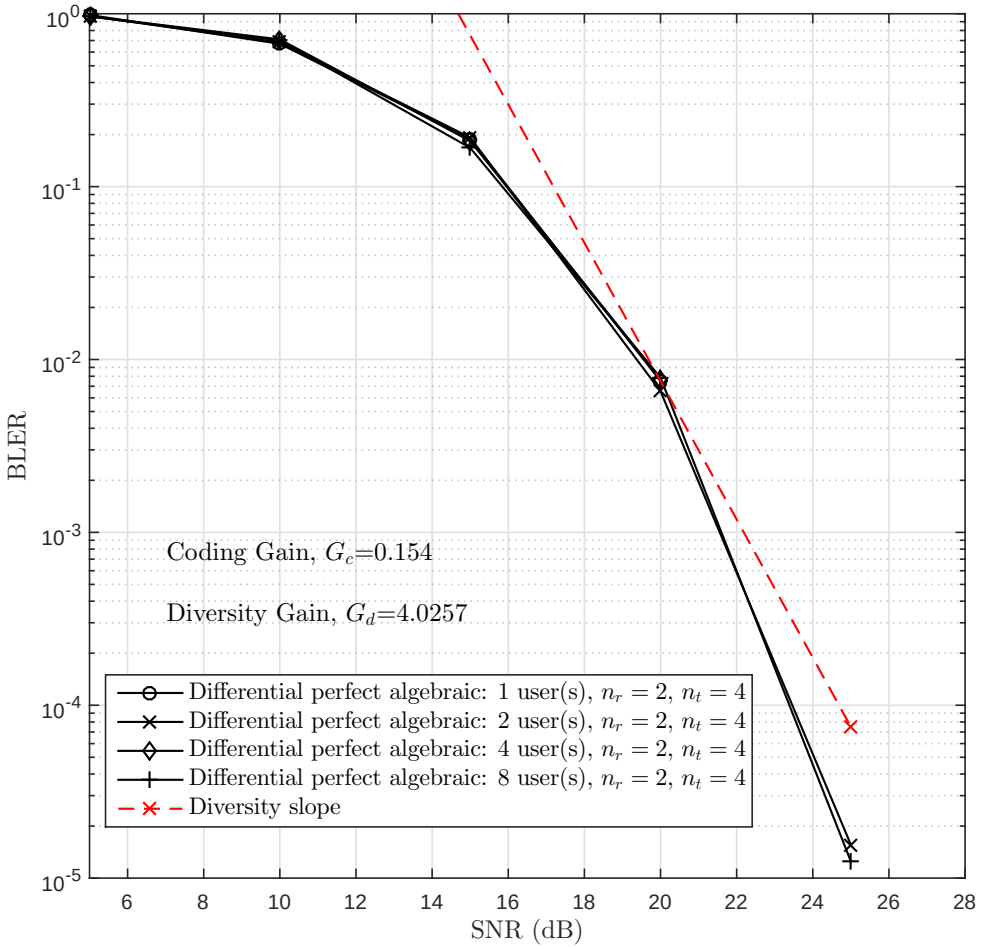

Figure 4.4. BLER performance of the proposed MU-MIMO STBC downlink transmission with differential algebraic STBC with a one, two, four, and eight users system model, $R = 4$, $N_t = 4$, $N_r = 2$, $l = 2$, and 4-PAM.

is $G_d = 4.0257$, i.e., the full diversity for this case is achieved with the rate of $R = 4$. Using the method above, the estimated value for a lower bound for coding gain is $G_c = 0.154$ (as a linear ratio).

Differential Algebraic Versus Orthogonal Differential Alamouti: In Fig. 4.5, the BLER performance is plotted and compared between the differential algebraic STBC and the differential Alamouti code in [39]. The differential Alamouti code is used as a benchmark scheme. First, the performance for both schemes is examined at the same rate $R = 2$ with $N_t = 2$ and $N_r = 2$. To get $R = 2$ for both schemes, 4-PAM constellation size for the differential Alamouti and 2-PAM for differential algebraic are used by using (4.56). The figure shows that the performance of the proposed scheme outperforms the differential Alamouti for the same rate. Second, the rate of the proposed scheme is increased from $R = 2$ to $R = 4$ and to $R = 8$, then it is compared to the differential Alamouti. The differential Alamouti with $R = 2$ is initially better at low SNR.

However, the differential algebraic curves converge and outperform at high SNR, even if their rates are two or four times the rate of the differential Alamouti scheme thanks to their steeper diversity slope, i.e., $G_d = 4$ and $G_d = 16$.

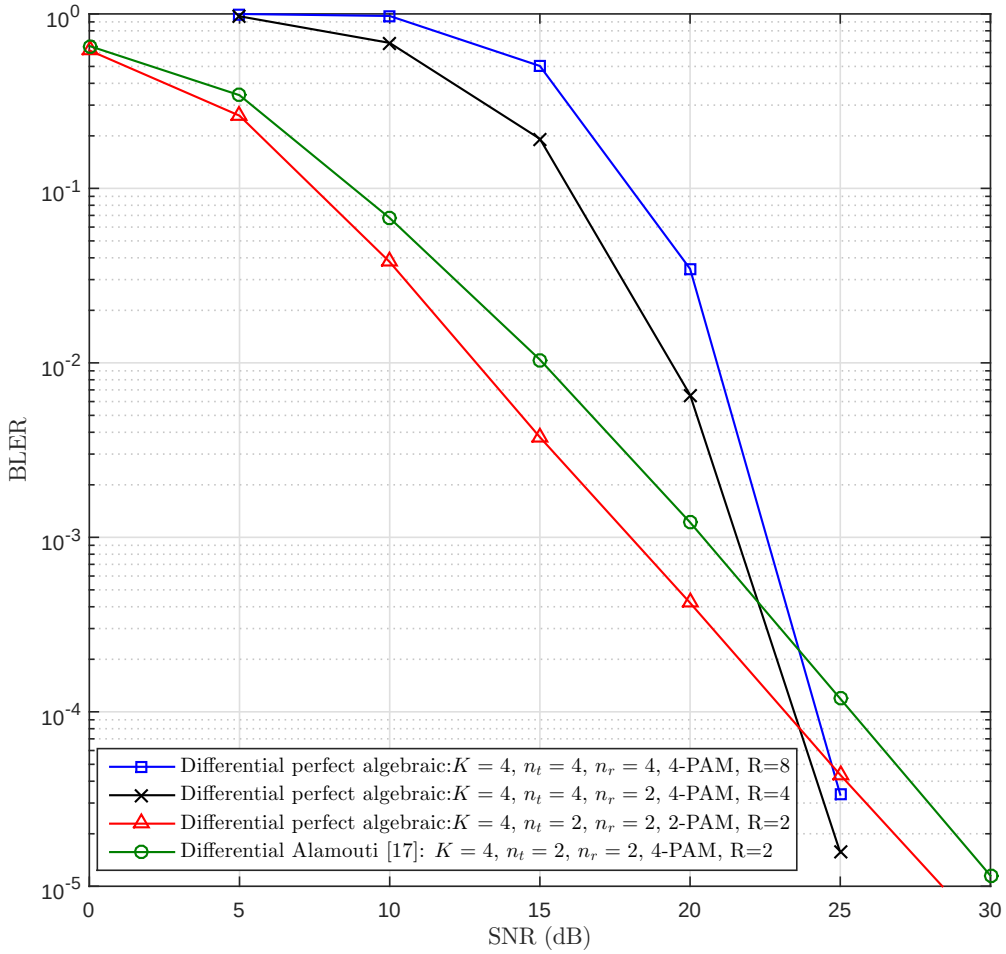

Figure 4.5. BLER performance of the proposed MU-MIMO STBC downlink transmission for differential algebraic and the orthogonal Alamouti Code for different rates.

The Impact of Multiple Receive Antenna Diversity: In Fig. 4.6, a two-user system is assumed; User 1 is equipped with one receive antenna, and User 2 has three receive antennas, and a 4-PAM constellation is used. The base station has four transmit antennas. The rate for User 1 is 2 bits/sec/Hz with one layer and for User 2 is 6 bits/sec/Hz with three layers. The BER performance of the system is shown in Fig. 4.6. The performance of User 2 is seen to outperform that of User 1 at high SNR, even though its rate is three times the rate of User 1, because of its receive antenna diversity.

Coherent Algebraic Versus Differential Algebraic STBC: In Fig. 4.7, the BLER performance is plotted and compared between the coherent algebraic and the differential algebraic STBC. For

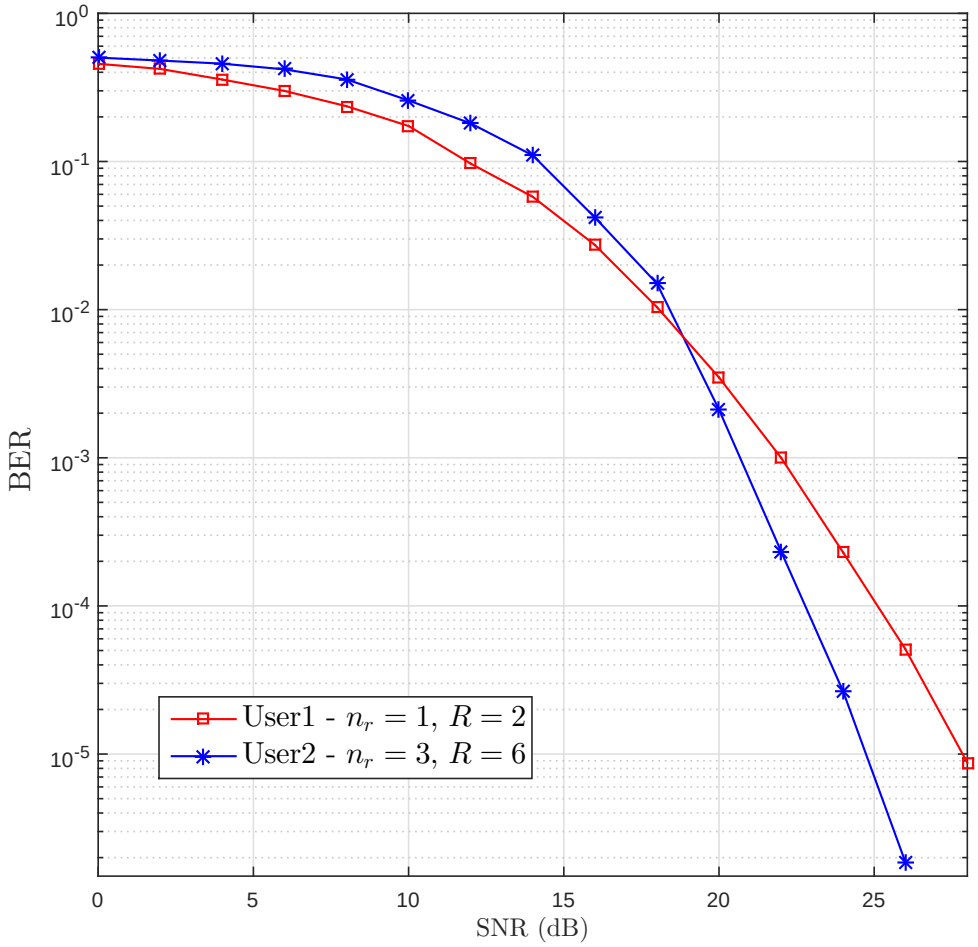

Figure 4.6. BER performance of the proposed MU-MIMO STBC downlink transmission with differential algebraic STBC with a two-user system model with two different rates and layers, $N_t = 4$, and 4-PAM.

a fair comparison of the two schemes, identical setup is assumed, namely 4 users, $N_t = 4$, each user has $N_r = 2$, $l = 2$ threads, $R = 4$, 4-PAM, sphere decoder, and with unitary Cayley matrix. Further, to quantify this performance loss in both schemes, the receiver's SNR is calculated for the same unit transmit power. For the differential algebraic detection scheme, the power of noise at the receiver is approximately two times the power of the noise for the coherent detection as shown in (4.44). Therefore, the received SNR of the differential detection scheme is approximately half of that of the coherent detection scheme for the same transmission power. This results in about a 4 dB difference in the performance at high SNRs as expected. The coding gain for both schemes is calculated, and it is $G_c = 0.421$ for coherent scheme and $G_c = 0.154$ for differential. The diversity for both schemes is $G_c \approx -4$.

The Impact of Multiple Access Interference (MAI): Fig. 4.8 illustrates the results of repeating the same experiment for a two user system where each user has two receive antennas but with

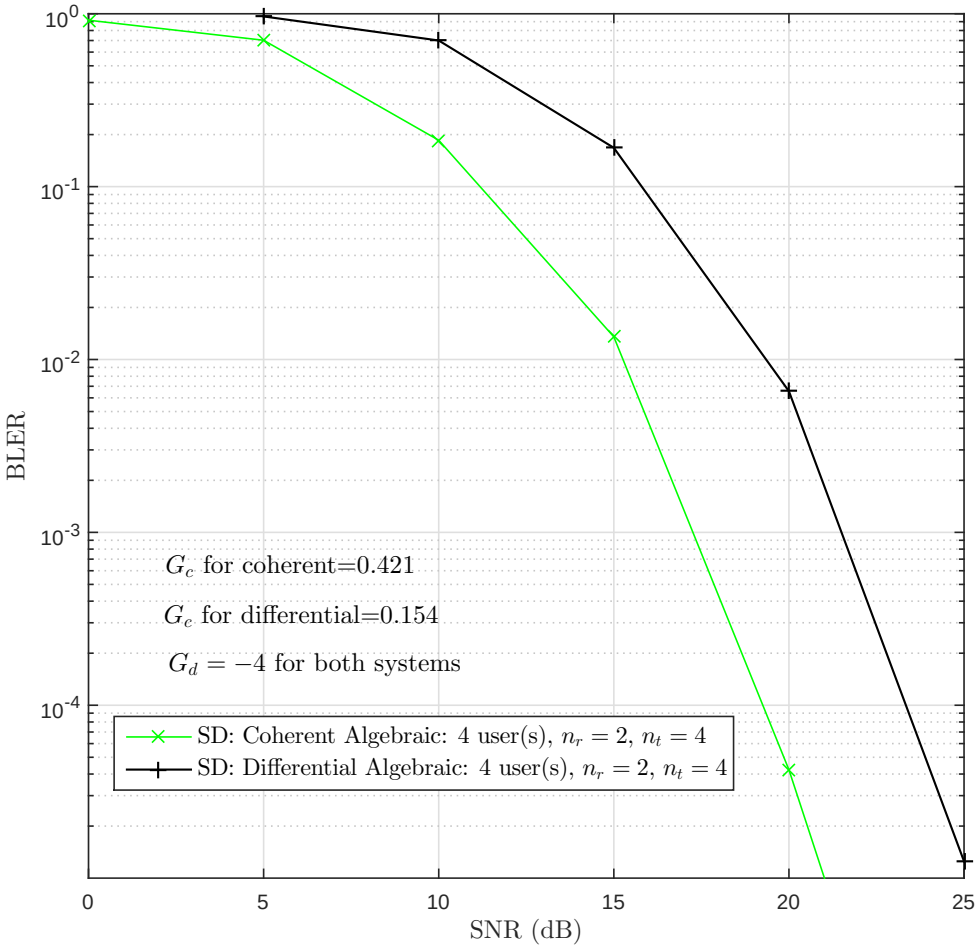

Figure 4.7. BLER performance of the proposed MU-MIMO STBC downlink transmission for differential algebraic and coherent algebraic STBC with four users system, $R = 4$, $N_t = 4$, $N_r = 2$, $l = 2$, and 4-PAM.

higher rate $R = 6$ and 8-PAM. The performance of the system in Fig. 4.8 underperforms that of the system in Fig. 4.4 because of its higher rate. Further, in this figure, the effect of error in the spreading matrices $\mathbf{V}$ is examined. Increasing the number of users in the system, the high mobility, and multipath propagation may result in MAI in orthogonal spreading matrices, which destroy the orthogonality of the transmitted signals for multiple users. For the two user system, let the error spreading matrix for User 1 be $\bar{\mathbf{V}}_1 = \mathbf{V}_1 + \alpha\mathbf{V}_2$, where α is the error coefficient. Therefore, the conditions for the orthogonality of the spreading matrix for User 1 and User 2 are as follows

$$\bar{\mathbf{V}}_1\bar{\mathbf{V}}_1^H = \mathbf{I}_{N_t} + \alpha^2\mathbf{I}_{N_t}. \tag{4.57}$$

$$\mathbf{V}_2\bar{\mathbf{V}}_1^H = \alpha\mathbf{I}_{N_t}. \tag{4.58}$$

The values of α are chosen to be 0.03, 0.05, and 0.08. It is shown that the error in the orthogonality of the spreading matrix $\mathbf{V}$ occurs among users when $\alpha > 0$.

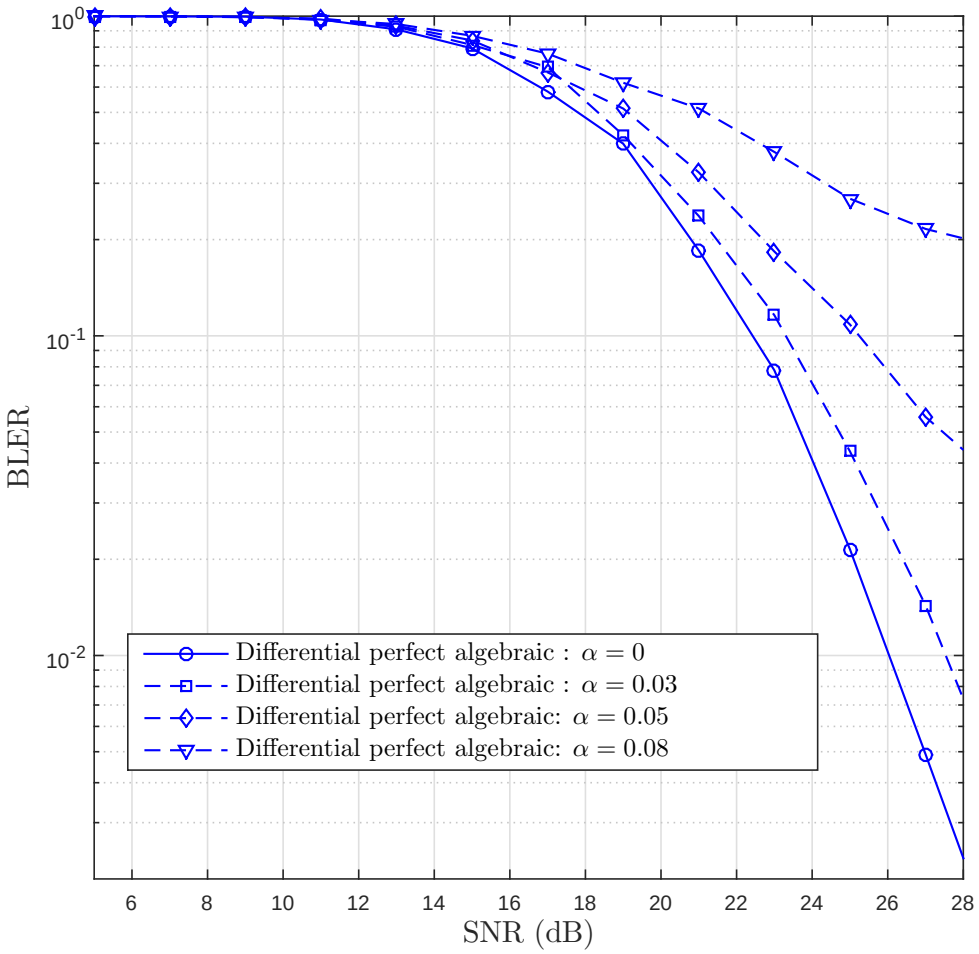

Figure 4.8. BLER performance of the proposed MU-MIMO STBC downlink transmission with differential algebraic STBC using a two user system model, $R = 6$, $N_t = 4$, $N_r = 2$, and 8-PAM.

4.7 Summary

In this chapter, a differential perfect algebraic STBC scheme for MU-MIMO with downlink transmission has been proposed. The Cayley differential STBC that has been introduced does not require channel knowledge, either at the transmitter or receiver. To simplify the receivers' equipment in the MU-MIMO system, the impact of the receiver channel estimation process and/or overhead problem can potentially be solved and avoided by using the Caylay differential STBC. Furthermore, to achieve a full-rate full-diversity noncoherent system the differential STBC combined with perfect algebraic STBC is derived. Due to the multiple users, there is a

need for the separation of the data streams and this is achieved by use of orthogonal spreading matrices. For this system, to limit the number of possible codewords, a near-optimal sphere decoder is performed to decode the signals at the receiver. The proposed schemes yield low complexity transceivers while also providing high rate with good performance. However, the system in this chapter has higher computational complexity because of its higher rate. Monte Carlo simulation results demonstrate the effectiveness of the proposed schemes.

Chapter 5

Differential Downlink Transmission in Massive MU-MIMO Systems

5.1 Introduction

MIMO technology helps in improving wireless multiple access and can be used to increase the spectral efficiency and improve the link reliability at low power operation [52, 58]. With multiple transmit antennas at the BS, the system can spatially multiplex multiple data streams for multiple users at the same frequency and time. The spatial multiplexing property becomes more effective as the number of antennas becomes large where the system is referred to as massive MIMO [59]. Such properties allow massive MIMO architecture to be an important part of many wireless communications standards, such as LTE and 5G networks.

In massive MIMO, precoding design requires channel knowledge at the transmitter to cancel the interference between users. However, due to the large number of transmit antennas, the estimation of all channel coefficients becomes impossible. Instead, in this chapter, differential transmit precoding schemes is considered which avoid the channel estimation. Therefore, a differential MIMO downlink transmission framework is proposed, in which a BS is equipped with a massive antenna array that precodes transmit signals to separate the data streams of multiple users. In particular, to achieve a low-complexity differential massive MIMO system, a novel downlink precoding design is proposed by employing knowledge of the power space profile (PSP) of users. It is assumed that the PSP for each user is estimated at the BS, since it can tolerate more complexity compared to receivers. Once the PSPs are estimated at the BS, the transmitter computes the precoder. More precisely, an optimal solution for the precoder based on a max-min signal-to-interference-plus-noise ratio (SINR) problem formulation is provided. The optimised precoder can effectively precancel the interference between users, thus enhancing overall system performance. In addition, two suboptimal solutions suitable for the low interference system based on the matched and the orthogonality approach of PSP of each user are also provided. The proposed schemes facilitate precancelling MAI, enhance system performance, and provide simple transmitter and receiver schemes. Consequently, since the proposed schemes avoid channel estimation, the system overheads and latency will be reduced significantly.

The remainder of this chapter is organised as follows. Section 5.2 introduces the system model of the differential massive MIMO system. Section 5.3 describes the downlink transmit precoding approach. Section 5.4 presents differential detection for a massive MIMO system. In Section 5.5, simulation results are shown. Finally, conclusions are drawn in Section 5.6.

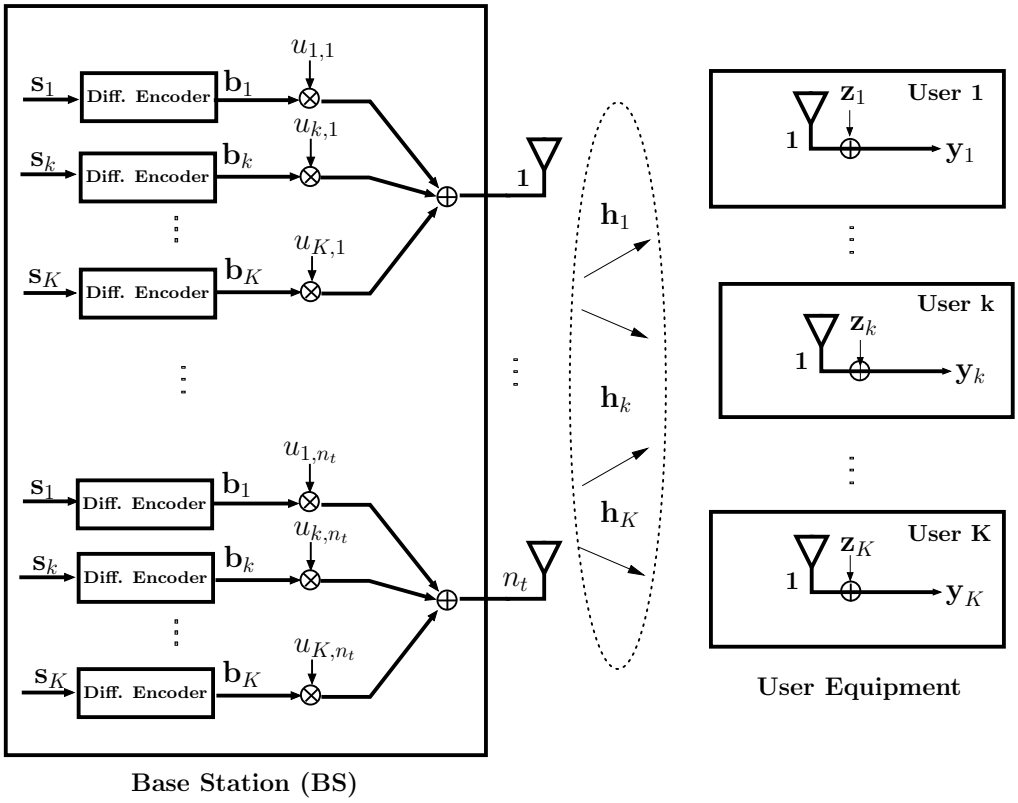

Figure 5.1. Differential massive MIMO downlink system model.

5.2 System Model

Consider a single-cell massive MIMO downlink broadcast channel. The BS has n_t transmit antennas, which simultaneously transmit multiple streams to K single-antenna users, as shown in Fig.1. The number of transmit antennas is assumed to be very large[1] ($n_t \gg 1$). It is assumed that all users are equipped with a single-antenna for the decoding process which is a realistic assumption for the massive MIMO system, where the large number of transmit antennas at the BS provides a mutual orthogonality among the vector-valued channels to the users (so-called favourable propagation) [61]. For downlink massive MIMO transmission, multiple-antennas at each user increases receiver complexity and overhead. Instead, it is more practical to have a simple, inexpensive, and power efficient single-antenna receiver. Further, equivalent capacity can be achieved by serving K single-antenna users instead of one user having K-multiple-antennas users, thereby serving more users in the cell [61].

[1]The assumption of $n_t \to \infty$ is valid and commonly used in the massive MIMO literature. However, the system's performance can be tested for the practical case of large but limited number of antennas [60].

5.2.1 Differential Massive MIMO System Model

For any kth user, $\boldsymbol{s}_k = [s_{k,1}, s_{k,2}, \cdots, s_{k,N}] \in \mathbb{C}^{1\times N}$, $1 \leq k \leq K$, is the information vector with elements drawn from an M-ary PSK constellation as:

$$\mathcal{M} = \left\{ e^{j2\pi i/M} \;\middle|\; i = 0, 1, \cdots, M-1 \right\}, \tag{5.1}$$

where N denotes the block length of the coherence time intervals. In the context of differential massive MIMO system, a sequence of symbols of the kth user $s_{k,\tau}$, $1 \leq \tau \leq N$, is differentially encoded into the transmit symbol vector $\boldsymbol{b}_k \in \mathbb{C}^{1\times(N+1)}$ via the rule

$$b_{k,\tau} = s_{k,\tau} b_{k,\tau-1} = b_{k,0} \prod_{i=1}^{\tau} s_{k,i}. \tag{5.2}$$

The transmit information signal vector $\boldsymbol{b}_k$ comprises the initial reference symbol $b_{k,0} = 1$ and the following N differentially encoded symbols in the form of $\boldsymbol{b}_k = [b_{k,0}, b_{k,1}, \cdots, b_{k,N}]$.

No prior information about the channel is considered to be available at the BS. The channel vector between the BS and user k, $\boldsymbol{h}_k = [h_{k,1}, h_{k,2}, \cdots, h_{k,n_t}]^T \in \mathbb{C}^{n_t\times 1}$, models independent fast fading and slow fading PSP attenuation, where the PSP is denoted as $g_{k,m}$, $1 \leq m \leq n_t$, PSP is specified later in Section 5.2.2. It is assumed that the channel coefficients remain constant over the block length and vary independently from one block to another. The coefficient $h_{k,m}$ can be written as

$$h_{k,m} = \sqrt{g_{k,m}}\, \tilde{h}_{k,m}, \qquad m = 1, 2, \cdots, n_t, \tag{5.3}$$

where $\tilde{h}_{k,m}$ is the fast-fading coefficient from the kth user to the mth transmit antenna of the BS, which is modelled as an independent over m and identically distributed (i.i.d.) complex Gaussian random variable with zero-mean and unit-variance, i.e., $\tilde{h}_{k,m} \sim \mathcal{CN}(0, 1)$. $g_{k,m}$ models the PSP attenuation between the mth antenna at the BS and user k, which is assumed to be independent over m and to be constant over many coherence time intervals N and known *a priori* to the BS. It is considered that the value of $\tilde{h}_{k,m}$ remains stationary for a sufficiently long transmission time. Then, it follows that

$$\boldsymbol{h}_k = \boldsymbol{G}_k^{1/2}\, \tilde{\boldsymbol{h}}_k, \qquad k = 1, \cdots, K, \tag{5.4}$$

where $\tilde{\boldsymbol{h}}_k = [\tilde{h}_{k,1}, \tilde{h}_{k,2}, \cdots, \tilde{h}_{k,n_t}]^T \in \mathbb{C}^{n_t\times 1}$, and $\boldsymbol{G}_k = \mathrm{diag}(\boldsymbol{g}_k) = \mathrm{diag}(g_{k,1}, g_{k,2}, \cdots, g_{k,n_t}) \in \mathbb{R}^{n_t\times n_t}$. Therefore, the variance of $\{\boldsymbol{h}_k\}$ is determined by the user PSP, where the channel vari-

ance is equal to the power profile, i.e., $h_{k,m} \sim \mathcal{CN}(0, g_{k,m})$.

It is assumed that the multiuser system adopts a linear transmission and reception strategy. The BS performs transmit beamforming and communicates simultaneously with all users. The instantaneous transmitted signal matrix $\boldsymbol{B} \in \mathbb{C}^{n_t \times (N+1)}$ for the kth user can then be expressed as

$$\boldsymbol{B} = \sum_{k=1}^{K} \sqrt{p_k} \boldsymbol{u}_k \boldsymbol{b}_k, \tag{5.5}$$

where $\boldsymbol{u}_k = \left[u_{k,1}, u_{k,2}, \cdots, u_{k,n_t}\right]^T \in \mathbb{R}^{n_t \times 1}$ is the normalised differential transmit precoder (beamformer) of the kth user, where $\|\boldsymbol{u}_k\|^2 = 1$. p_k is the downlink average transmit power of the kth user. A total power constraint at the BS is considered, namely

$$\mathbb{E}\left\{\mathrm{trace}(\boldsymbol{B}^H \boldsymbol{B})\right\} = \bar{P}. \tag{5.6}$$

The received signal vector $\boldsymbol{y}_k \in \mathbb{C}^{1 \times (N+1)}$ for the kth user is given by

$$\boldsymbol{y}_k = \sqrt{p_k} \boldsymbol{h}_k^H \boldsymbol{u}_k \boldsymbol{b}_k + \boldsymbol{h}_k^H \sum_{\substack{q=1 \\ q \neq k}}^{K} \sqrt{p_q} \boldsymbol{u}_q \boldsymbol{b}_q + \boldsymbol{z}_k \tag{5.7}$$

$$= \sqrt{p_k} \boldsymbol{h}_k^H \boldsymbol{u}_k \boldsymbol{b}_k + \underbrace{\boldsymbol{w}_k + \boldsymbol{z}_k}_{\stackrel{\text{def}}{=} \bar{\boldsymbol{z}}_k}, \tag{5.8}$$

where the term $\sqrt{p_k} \boldsymbol{h}_k^H \boldsymbol{u}_k \boldsymbol{b}_k$ represents the desired signal at the kth user, $\boldsymbol{w}_k = \boldsymbol{h}_k^H \sum_{\substack{q=1 \\ q \neq k}}^{K} \sqrt{p_q} \boldsymbol{u}_q \boldsymbol{b}_q \in \mathbb{C}^{1 \times (N+1)}$ is the MAI component against the kth user, and $\boldsymbol{z}_k \in \mathbb{C}^{1 \times (N+1)}$ is the noise vector modelled as zero-mean complex circularly symmetric Gaussian random variables, i.e., $\boldsymbol{z}_k \sim \mathcal{CN}(0, \sigma_{z_k}^2 \boldsymbol{I}_{n_t})$, and $\bar{\boldsymbol{z}}_k = \boldsymbol{w}_k + \boldsymbol{z}_k$ denotes the sum of the noise and MAI component from all other interference users.

Assuming that the information transmitted symbols $\boldsymbol{b}_k$ are uncorrelated, the average SINR_k at the kth user can be expressed as follows

$$\mathrm{SINR}_k = \mathbb{E}\left[\frac{p_k \left|\boldsymbol{h}_k^H \boldsymbol{u}_k\right|^2}{\sum_{\substack{q=1 \\ q \neq k}}^{K} p_q \left|\boldsymbol{h}_k^H \boldsymbol{u}_q\right|^2 + \sigma_{z_k}^2}\right]. \tag{5.9}$$

5.2.2 Colocated Antenna System with a Uniform Linear Array Model

Now, the PSP model is derived by following the approach in [47]. As shown in Fig. 5.2, the users are randomly distributed in front of a large uniform antenna array at the BS. It is assumed

that the BS has full knowledge of any user's location information. The location of the users is determined by the following parameters: $r_{k,m}$ is the distance between the antenna index m and the kth user; l_k is the direct orthogonal distance between the kth user and the array; and, l_a is the antenna spacing[1]. Let m_k denote the antenna element closest to the kth user according to the Euclidean distance. From algebraic geometry, it follows

$$r_{k,m} = \sqrt{l_k^2 + l_a^2 |m - m_k|^2} \tag{5.10}$$

$$= l_k \sqrt{1 + |m - m_k|^2 / l_{k,r}^2}, \quad m = 1, \cdots, n_t, \tag{5.11}$$

where $l_{k,r} \stackrel{\text{def}}{=} l_k / l_a$ denotes the normalised relative distance of the kth user to the array. It is assumed that the average transmit power obeys the path loss model with *path loss exponent* γ. Hence, the path loss for the kth user at antenna m is given by

$$g_{k,m} = r_{k,m}^{-\gamma}. \tag{5.12}$$

Using exponential and logarithmic properties, yields

$$\begin{aligned} g_{k,m} &= \exp\left\{ -\gamma \log_e \left(l_k \sqrt{1 + |m - m_k|^2 / l_{k,r}^2} \right) \right\} \\ &= \exp\left\{ -\gamma \log_e(l_k) + \frac{-\gamma \log_e \left(1 + |m - m_k|^2 / l_{k,r}^2\right)}{2} \right\}. \end{aligned}$$

Since $\log_e(1 + x) \approx x$ for small x, yields

$$\begin{aligned} g_{k,m} &= \exp\left\{ -\gamma \log_e(l_k) \right\} \cdot \exp\left\{ \frac{-\gamma |m - m_k|^2}{2 l_{k,r}^2} \right\} \\ &= l_k^{-\gamma} \cdot \exp\left\{ \frac{-|m - m_k|^2}{2\zeta_k^2} \right\} \qquad (5.13) \\ &= \beta_k \cdot \exp\left\{ \frac{-|m - m_k|^2}{2\zeta_k^2} \right\}, \qquad (5.14) \end{aligned}$$

where $\beta_k = l_k^{-\gamma}$ and $\zeta_k^2 \stackrel{\text{def}}{=} l_{k,r}^2 / \gamma$. Therefore, the PSP is well approximated by a Gaussian function with mean m_k and channel variance ζ_k^2.

Remark 1. *In practical systems, the power space profile (PSP), $g_{k,m}$, of each user (which includes the path loss exponent) varies very slowly with time compared to the fast fading coefficients $\tilde{h}_{k,m}$. In this context, for massive MIMO systems, it is reasonable to assume that the*

[1] it is assumed that there is no correlation between transmit antennas at the BS as they are spaced at a minimum of 0.5λ.

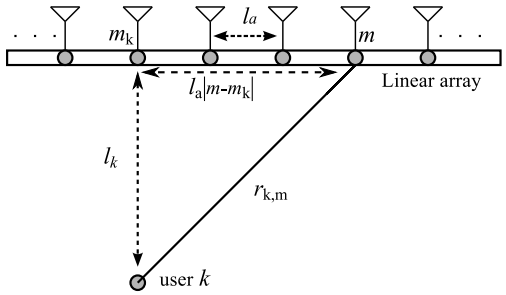

Figure 5.2. Geometric system model of user's location.

PSPs of the users of the system are known at the BS [62–64]. In the uplink, it is assumed that each user transmits N (i.i.d.) symbols as $\boldsymbol{s}_k = [s_{k,1}, s_{k,2}, \cdots, s_{k,\tau}, \cdots, s_{k,N}]$. For each symbol the BS can calculate the PSP for each user accurately by ***averaging*** *the received uplink signal over different data slots indexed by τ as [62]*

$$\hat{\boldsymbol{G}}_k = diag(\hat{\boldsymbol{g}}_k) = \mathbb{E}\left[\boldsymbol{r}_{k,\tau}\boldsymbol{r}_{k,\tau}^H\right], \tag{5.15}$$

where $\boldsymbol{r}_{k,\tau}$ is the uplink received signal vector for the kth user at the antennas of the BS during the τth time slot, which is given as

$$\boldsymbol{r}_{k,\tau} = \boldsymbol{h}_k s_{k,\tau} + \boldsymbol{z}_{k,\tau}. \tag{5.16}$$

Hence, with the assumption of channel reciprocity, the PSP is calculated for each user during the uplink as in (5.15), which is assumed to be equivalent to the PSP in the downlink.

Note that PSP profile estimation is less challenging than estimating the actual channel state information, in which the PSP can remain constant over many coherence time intervals. However, the actual estimation process of $g_{k,m}$ is beyond the scope of this work, thus it is assumed that the PSP profiles is (perfectly) known in our system.

5.3 Downlink Transmit Precoding

In massive MIMO, transmit precoding is used to cancel inter-user interference. Conventional transmit precoding design requires channel knowledge at the transmitter. However, in massive MIMO, the number of transmit antennas is very large, i.e., $n_t \gg 1$. Hence, the estimation of all channel coefficients $h_{k,m}$ quickly becomes unfeasible. Instead, differential transmit precoding schemes could be considered which avoid the need for explicit channel estimation. After estimating the PSP profile $g_{k,m}$ at the BS, it is used for designing the transmit precoder for each user

k to separate different users. Now, an asymptotic analysis of SINR and the proposed precoder design strategies for the differential massive MIMO framework is presented.

5.3.1 Asymptotic Analysis of SINR

As a consequence of employing a large number of antennas at the BS $n_t \to \infty$ (as in this case of massive MIMO), the downlink channel vectors of independent users have a large degree of orthogonality, i.e.,

$$\frac{1}{n_t}\,\boldsymbol{h}_k^H \boldsymbol{h}_k \overset{N\to\infty}{\to} \frac{\boldsymbol{g}_k^T \boldsymbol{g}_k}{n_t}, \qquad \frac{1}{n_t}\,\boldsymbol{h}_k^H \boldsymbol{h}_{q\neq k} \overset{N\to\infty}{\to} \frac{\boldsymbol{g}_k^T \boldsymbol{g}_{q\neq k}}{n_t}.$$

The orthogonality between different user's channels is determined by the orthogonality between the small fading vectors $\{\tilde{\boldsymbol{h}}_k\}$, and the orthogonality between the PSPs $\{\boldsymbol{g}_k\}$.

Theorem 1. *From the law of large random numbers and under the most favourable propagation conditions, where the column-vectors of the propagation vectors are asymptotically orthogonal, the expected value of* SINR_k *can be calculated when* $n_t \to \infty$. *Since* $\{\boldsymbol{h}_k^H\}$ *has Gaussian distribution with zero-mean and covariance of* $\boldsymbol{G}_k = \mathrm{diag}\left(\boldsymbol{g}_k\right)$, *hence the desired signal* $\sqrt{p_k}\boldsymbol{h}_k^H \boldsymbol{u}_k$ *is also Gaussian distributed with zero mean and variance* $p_k \sum_{m=1}^{n_t} g_{k,m} u_{k,m}^2$, *where the sum of multiple Gaussian variables is also a Gaussian variable. Similarly, the interference component of* SINR_k *is also a Gaussian signal with variance* $\sum_{q=1,q\neq k}^{K} p_q \sum_{m=1}^{n_t} g_{k,m} u_{q,m}^2$. *The variance for the AWGN noise is* $\sigma_{z_k}^2$. *Therefore, the expected value of* SINR_k *as* $n_t \to \infty$ *is*

$$\mathrm{SINR}_k \overset{N\to\infty}{\to} \frac{\mathbb{E}\left[p_k \left|\boldsymbol{h}_k^H \boldsymbol{u}_k\right|^2\right]}{\mathbb{E}\left[\sum_{\substack{q=1\\ q\neq k}}^{K} p_q \left|\boldsymbol{h}_k^H \boldsymbol{u}_q\right|^2 + \sigma_{z_k}^2\right]} \tag{5.17}$$

$$= \frac{p_k \sum_{m=1}^{n_t} g_{k,m} u_{k,m}^2}{\sum_{\substack{q=1\\ q\neq k}}^{K} p_q \sum_{m=1}^{n_t} g_{k,m} u_{q,m}^2 + \sigma_{z_k}^2}. \tag{5.18}$$

Proof. See Appendix A □

5.3.2 Suboptimal Precoders

5.3.2.1 Matched PSP Precoder

The first precoder design strategy is to match the beamformer vector to the PSP profile of the transmit antennas to separate different users, i.e., $u_{k,m}^2 \sim g_{k,m}$, which can be written as

$$u_{k,m} = \sqrt{\beta_k \cdot \exp\left\{\frac{-|m - m_k|^2}{2\zeta_k^2}\right\}}, \tag{5.19}$$

where it is assumed here the BS has full knowledge of the channel parameter ζ_k, and the antenna index m_k which is the closest to the user k with maximum average power. For the power allocation in matched PSP scheme, the downlink transmit power is allocated equally between users, i.e., $p_k = \bar{P}/K$.

5.3.2.2 Orthogonal PSP Precoder

In the orthogonal PSP precoder, the beamformer for each user has to be distinguished and identified from other users. In the orthogonal precoder scheme, each user is assigned a unique orthogonal PSP to enhance data separation between users. The orthogonal PSP for each user is then multiplexed by its own power profile.

The orthogonal precoder for each user can be constructed using the Gram-Schmidt process (GSP). Let the vector $\boldsymbol{v}_k = \left[v_{k,1}, v_{k,2}, \cdots, v_{k,n_t}\right]^T \in \mathbb{R}^{n_t \times 1}$ represent the user's PSP vector. The elements of vector $\boldsymbol{v}_k$ are computed by matching their value to the power profile of the transmit antennas, i.e., $v_{k,m}^2 \sim g_{k,m}$. The Gram-Schmidt process takes a finite, linearly independent set $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k, \cdots, \boldsymbol{v}_K\}$ for $K \leq n_t$ and generates an orthogonal set $\bar{S} = \{\bar{\boldsymbol{v}}_1, \bar{\boldsymbol{v}}_2, \cdots, \bar{\boldsymbol{v}}_k, \cdots, \bar{\boldsymbol{v}}_K\}$ which spans the same K-dimensional subspace of $\mathbb{R}^{n_t \times 1}$ as S. The projection operator is defined as [19]

$$\text{proj}_{\bar{\boldsymbol{v}}}(\boldsymbol{v}) = \frac{\langle \bar{\boldsymbol{v}}, \boldsymbol{v} \rangle}{\langle \bar{\boldsymbol{v}}, \bar{\boldsymbol{v}} \rangle} \bar{\boldsymbol{v}}, \tag{5.20}$$

where $\langle \bar{\boldsymbol{v}}, \boldsymbol{v} \rangle$ denotes the inner product of the vectors $\bar{\boldsymbol{v}}$ and $\boldsymbol{v}$, i.e., $\langle \bar{\boldsymbol{v}}, \boldsymbol{v} \rangle = \bar{\boldsymbol{v}}^T \boldsymbol{v}$ for vectors in $\mathbb{R}^{n_t \times 1}$. The Gram-Schmidt process then works as follows:

$$\begin{aligned}
\bar{\boldsymbol{v}}_1 &= \boldsymbol{v}_1, \\
\bar{\boldsymbol{v}}_2 &= \boldsymbol{v}_2 - \text{proj}_{\bar{\boldsymbol{v}}_1}(\boldsymbol{v}_2), \\
\bar{\boldsymbol{v}}_3 &= \boldsymbol{v}_3 - \text{proj}_{\bar{\boldsymbol{v}}_1}(\boldsymbol{v}_3) - \text{proj}_{\bar{\boldsymbol{v}}_2}(\boldsymbol{v}_3),
\end{aligned}$$

$$\vdots$$

$$\bar{\boldsymbol{v}}_k = \boldsymbol{v}_k - \sum_{i=1}^{k-1} \text{proj}_{\bar{\boldsymbol{v}}_i}(\boldsymbol{v}_k). \tag{5.21}$$

Note that, the Gram-Schmidt precoder for the first user is equal to the original PSP for the first user, i.e., $\bar{\boldsymbol{v}}_1 = \boldsymbol{v}_1$, hence the user's separation works only for the received signal of the first user. To enhance the separation for the received signal of other users, each element of the orthonormal vector $\bar{\boldsymbol{v}}_k$ is multiplied by its own specific original power profile elements of $\boldsymbol{v}_k$ and then normalise them, which yields

$$\boldsymbol{u}_k = \frac{\boldsymbol{v}_k \circ \bar{\boldsymbol{v}}_k}{\|\boldsymbol{v}_k \circ \bar{\boldsymbol{v}}_k\|}, \tag{5.22}$$

where $\circ$ denotes the Hadamard product. For power allocation in orthogonal PSP precoder, the downlink transmit power is allocated equally between users, i.e., $p_k = \bar{P}/K$.

5.3.3 Optimal PSP Precoders

In this precoder, the joint optimisation of power and downlink precoder is considered for the PSP among all users simultaneously using the max-min formulation problem. A max-min formulation guarantees a fair quality of service among all users.

5.3.3.1 SINR Optimal PSP Precoder

In optimal PSP precoder, I maximise the worst case SINR jointly among all user. Starting from (5.18), the corresponding optimisation problem can be written as

$$\underset{\substack{p_k, u_{k,m} \\ k\in[1,K], m\in[1,n_t]}}{\text{maximise}} \quad \min_{k\in[1,K]} \left(\frac{p_k \sum_{m=1}^{n_t} g_{k,m} u_{k,m}^2}{\sum_{\substack{q=1 \\ q\neq k}}^{K} p_q \sum_{m=1}^{n_t} g_{k,m} u_{q,m}^2 + \sigma_{z_k}^2} \right), \tag{5.23a}$$

$$\text{subject to} \quad \sum_{k=1}^{K} \sum_{m=1}^{n_t} p_k u_{k,m}^2 \leq \bar{P}, \tag{5.23b}$$

$$p_k u_{k,m}^2 \geq 0, \ \ \forall k, m. \tag{5.23c}$$

Problem (5.23) can be recast as

$$\underset{\substack{p_k, \boldsymbol{c}_k \\ k\in[1,K]}}{\text{maximise}} \quad \min_{k\in[1,K]} \left(\frac{p_k \boldsymbol{f}_k^T \boldsymbol{c}_k}{\boldsymbol{f}_k^T \sum_{\substack{q=1 \\ q\neq k}}^{K} p_q \boldsymbol{c}_q + 1} \right), \tag{5.24a}$$

$$\text{subject to} \quad \sum_{k=1}^{K} p_k \mathbf{1}_{n_t}^T \boldsymbol{c}_k \leq \bar{P}, \tag{5.24b}$$

$$p_k \boldsymbol{c}_k \geq 0, \ \ \forall k, \tag{5.24c}$$

where $\boldsymbol{f}_k$ and $\boldsymbol{c}_k$ are defined as

$$\boldsymbol{f}_k = \frac{1}{\sigma_{z_k}^2} \left[g_{k,1}, g_{k,2}, \cdots, g_{k,m}, \cdots, g_{k,n_t} \right]^T, \tag{5.25}$$

and

$$\boldsymbol{c}_k = \left[u_{k,1}^2, u_{k,2}^2, \cdots, u_{k,m}^2 \cdots, u_{k,n_t}^2 \right]^T. \tag{5.26}$$

It can be seen that the cost-function in (5.24a) is non-linear and non-convex over the optimisation variables p_k, and $\boldsymbol{c}_k$ for $k \in [1, K]$. In the following, an optimal solutions for the design problems is provided. The feasibility of problem (5.24) can be examined by solving it with the objective function replaced by constant values, i.e., finding a common domain which satisfies all problem constraints. Without loss of generality, it is assumed that the problem is feasible. Next, the optimisation problem is solved optimally through recasting the non-convex constraints. Now, define a $Kn_t \times 1$ vector $\boldsymbol{v}$ as

$$\boldsymbol{v} = \left[p_1(\boldsymbol{c}_1)^T, p_2(\boldsymbol{c}_2)^T, \cdots, p_K(\boldsymbol{c}_K)^T \right]^T. \tag{5.27}$$

In addition, define other variables $\boldsymbol{w}_k$ and $\bar{\boldsymbol{w}}_k$ of size $Kn_t \times 1$ as

$$\boldsymbol{w}_k = \left[\mathbf{0}_{(k-1)n_t \times 1}^T, \boldsymbol{f}_k^T, \mathbf{0}_{(K-k)n_t \times 1}^T \right]^T, \tag{5.28}$$

and

$$\bar{\boldsymbol{w}}_k = \left[\boldsymbol{f}_1^T, \cdots, \boldsymbol{f}_{k-1}^T, \mathbf{0}_{n_t \times 1}^T, \boldsymbol{f}_{k+1}^T, \cdots, \boldsymbol{f}_K^T \right]^T, \tag{5.29}$$

where $\mathbf{0}_{m \times 1}$ denotes an $m \times 1$ vector whose elements are zero. Next, the SINR_k optimisation problem in (5.24) may be written in a more convenient form by using (5.27), (5.28), and (5.29), which yields

$$\underset{\boldsymbol{v}}{\text{maximise}} \ \min_{k \in [1,K]} \left(\frac{\boldsymbol{w}_k^T \boldsymbol{v}}{\bar{\boldsymbol{w}}_k^T \boldsymbol{v} + 1} \right), \tag{5.30a}$$

$$\text{subject to} \quad \mathbf{1}_{n_t}^T \boldsymbol{v} \leq \bar{P}, \tag{5.30b}$$

$$\boldsymbol{v} \geq 0. \tag{5.30c}$$

To convexify the cost-function (5.30a), which comprises a product of fractional terms, the numerators and denominators of the fractions are substituted by exponential variables as follows [65]

$$e^{\alpha_k} = \boldsymbol{w}_k^T \boldsymbol{v}, \qquad \forall k, \tag{5.31a}$$

$$e^{\tilde{\alpha}_k} = \bar{\boldsymbol{w}}_k^T \boldsymbol{v} + 1, \quad \forall k. \tag{5.32a}$$

Then, by using the properties of the exponential and according to (5.31a) and (5.32a), the problem in (5.30) can be formalised as

$$\underset{\substack{\boldsymbol{v},\alpha_k,\tilde{\alpha}_k \\ k\in[1,K]}}{\text{maximise}} \ \min_{k\in[1,K]} \big(e^{(\alpha_k-\tilde{\alpha}_k)}\big), \tag{5.33a}$$

$$\text{subject to} \qquad \mathbf{1}_{n_t}^T \boldsymbol{v} \leq \bar{P}, \tag{5.33b}$$

$$\boldsymbol{v} \geq 0, \tag{5.33c}$$

$$e^{\alpha_k} \leq \boldsymbol{w}_k^T \boldsymbol{v}, \qquad \forall k, \tag{5.33d}$$

$$e^{\tilde{\alpha}_k} \geq \bar{\boldsymbol{w}}_k^T \boldsymbol{v} + 1, \quad \forall k. \tag{5.33e}$$

It can be seen that the exponential parameters e^{α_k} and $e^{\tilde{\alpha}_k}$ in (5.33d) and (5.33e) are constrained by the expressions on the right hand sides of (5.31a) and (5.32a), respectively. The objective function in (5.33a) consists of an exponential function which is non-convex, and thus it can be linearised by using the monotonicity property of the exponential function. Hence, the objective function in (5.33a) can be defined as follows

$$e^{(\alpha_k-\tilde{\alpha}_k)} \overset{\text{def}}{=} \alpha_k - \tilde{\alpha}_k, \qquad \forall k. \tag{5.34}$$

Next, to deal with the non-convex constraint (5.33e), the exponential term $e^{\tilde{\alpha}_k}$ is linearised by using the first order Taylor approximation as follows [66]

$$e^{\tilde{\alpha}_k} = e^{\ddot{\alpha}_k}\big(1 + \tilde{\alpha}_k - \ddot{\alpha}_k\big), \qquad \forall k, \tag{5.35}$$

where $\ddot{\alpha}_k$ is the point where the linear approximation is made. Therefore, from (5.34) and (5.35), problem (5.33) can be reformulated as

$$\underset{\substack{\boldsymbol{v},\alpha_k,\tilde{\alpha}_k \\ k\in[1,K]}}{\text{maximise}} \ \min_{k\in[1,K]} \alpha_k - \tilde{\alpha}_k, \tag{5.36a}$$

$$\text{subject to} \quad \mathbf{1}_{n_t}^T \boldsymbol{v} \leq \bar{P}, \tag{5.36b}$$

$$\boldsymbol{v} \geq 0, \tag{5.36c}$$

$$e^{\alpha_k} \leq \boldsymbol{w}_k^T \boldsymbol{v}, \qquad \forall k, \tag{5.36d}$$

$$e^{\ddot{\alpha}_k}\left(1 + \tilde{\alpha}_k - \ddot{\alpha}_k\right) \geq \bar{\boldsymbol{w}}_k^T \boldsymbol{v} + 1, \quad \forall k. \tag{5.36e}$$

Now the above problem (5.36) is convex and can be solved iteratively using CVX optimisation software [67]. The initial value of $\ddot{\alpha}_k$ is updated by the optimised value of $\tilde{\alpha}_k$, $\forall k$, obtained in the previous iteration. The iterations continue until the error, $\sum_{k=1}^{K} |\ddot{\alpha}_k - \tilde{\alpha}_k|$, converges to a certain threshold. Algorithm 1 is provided to solve the above optimisation function. Here $\boldsymbol{\alpha} = [\alpha_1 \cdots \alpha_K]^T$, $\tilde{\boldsymbol{\alpha}} = [\tilde{\alpha}_1 \cdots \tilde{\alpha}_K]^T$, and $\ddot{\boldsymbol{\alpha}} = [\ddot{\alpha}_1 \cdots \ddot{\alpha}_K]^T$.

Algorithm 1 Algorithm for solving problem (5.36)

1: Set threshold = ϵ
2: Initialize $\ddot{\boldsymbol{\alpha}}^{[i]}$, $\tilde{\boldsymbol{\alpha}}^{[i]}$, $i = 0$
3: **while** $\mathbf{1}_{n_t}^T \left(|\ddot{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}}|\right) > \epsilon$ or $i = 0$ **do**
4: increment $i = i + 1$.
5: update the initial values $\ddot{\boldsymbol{\alpha}}^{[i]} = \tilde{\boldsymbol{\alpha}}^{[i-1]}$.
6: solve problem (5.36) using CVX and calculate $\boldsymbol{v}^{[i]}$, $\boldsymbol{\alpha}^{[i]}$, $\tilde{\boldsymbol{\alpha}}^{[i]}$.
7: **until** Convergence.
8: **end while**
9: Find $\boldsymbol{u}_k$ and p_k of each user from $\boldsymbol{v}$ as in (5.26) and (5.27).

Remark 2. *There is an alternative approach for designing the transmit precoder based on maximising the worst case of signal-to-leakage-noise ratio (SLNR). The SLNR is defined as the ratio of received signal power at the desired user to received signal power at the other users (the leakage) [68]. The average* SLNR_k *at the* k*th user can be expressed as*

$$\mathrm{SLNR}_k \overset{N\to\infty}{\to} \frac{\mathbb{E}\left[p_k \left|\boldsymbol{h}_k^H \boldsymbol{u}_k\right|^2\right]}{\mathbb{E}\left[p_k \sum_{\substack{q=1 \\ q\neq k}}^{K} \left|\boldsymbol{h}_q^H \boldsymbol{u}_k\right|^2 + \sigma_{z_k}^2\right]} \tag{5.37}$$

$$= \frac{p_k \sum_{m=1}^{n_t} g_{k,m} u_{k,m}^2}{p_k \sum_{\substack{q=1 \\ q \neq k}}^{K} \sum_{m=1}^{n_t} g_{q,m} u_{k,m}^2 + \sigma_{z_k}^2}. \tag{5.38}$$

The proof of (5.38) is similar to the proof of (5.18) in Theorem 1. The optimisation solution for maximising the worst case SLNR of (5.38), (max-min SLNR), jointly among all users provides the same performance as in the proposed optimal PSP SINR precoder, (max-min SINR).

5.3.4 Computational Complexity Analysis for the PSP Precoders

In this section, the computational complexity is quantified for the proposed PSP precoders for the optimal and the suboptimal solutions. The computational process is done based on the size of input data, the floating point operations (FLOPs), the type of the optimisation problems, the number of the required iterations, and the methods used in finding the solution.

5.3.4.1 Complexity of Suboptimal Solutions

The notion of FLOPs is introduced. The total number of FLOPs is used to measure the computational complexity of matrix operations. The total FLOPs needed for some matrix operations is listed below [49]:

- Multiplication of $m \times n$ and $n \times p$ complex matrices: $\mathcal{O}(8mnp - 2mp)$;
- Inversion of an $m \times m$ real matrix using Gauss-Jordan elimination: $\mathcal{O}(4m^3/3)$.
- GSP to an $m \times n$ $(m \geq n)$ complex matrix: $\mathcal{O}(8n^2(m - n/3))$.
- Hadamard product for two $m \times m$ matrices: $\mathcal{O}(m)$.
- L^2-norm of an $m \times 1$ real vector: $\mathcal{O}(3m)$. According to the aforementioned summary of FLOPs operations, the computational complexity of the suboptimal PSP precoder is

$$\mathcal{O}\bigg(K\Big(13n_t - \frac{8}{3}\Big)\bigg). \tag{5.39}$$

5.3.4.2 Complexity of Optimal Solution

Now, the complexity of optimising the downlink PSP precoder is calculated in which it is formulated as a linear programming (LP) problem in (5.36). The computational complexity of such LP problems has been studied in Chapter 6 in [69] where the complexity is calculated in terms of the number of optimisation variables n, number of constraints m and the size of

input data $\dim(\boldsymbol{p})$, where $\boldsymbol{p}$ is the vector of input data. To apply the complexity evaluation steps given in chapter 6 in [69], problem (5.36) is recast into its standard LP form. This can be achieved by replacing the min operator in the objective function by new slack variable π and K scalar constraints (see (5.40d)). Therefore, (5.40) is an equivalent and standard LP recast of the original problem (5.36). Note that the constraint (5.36d) is linearised in a similar way as used for (5.36e) since the used CVX's solvers such as SDPT3 and SeDuMi do not support the exponential function.

$$\underset{\substack{\boldsymbol{v},\ \alpha_k,\ \tilde{\alpha}_k,\ \pi \\ k\in[1,K]}}{\text{maximise}} \ \pi \tag{5.40a}$$

$$\text{subject to} \ \ \mathbf{1}_{n_t}^T \boldsymbol{v} \leq \bar{P}, \tag{5.40b}$$

$$\boldsymbol{v} \geq 0, \tag{5.40c}$$

$$\alpha_k - \tilde{\alpha}_k \geq \pi, \ \ \forall k, \tag{5.40d}$$

$$e^{\hat{\alpha}_k}\big(1 + \alpha_k - \hat{\alpha}_k\big) \leq \boldsymbol{w}_k^T \boldsymbol{v}, \quad \forall k, \tag{5.40e}$$

$$e^{\ddot{\alpha}_k}\big(1 + \tilde{\alpha}_k - \ddot{\alpha}_k\big) \geq \bar{\boldsymbol{w}}_k^T \boldsymbol{v} + 1, \ \ \forall k. \tag{5.40f}$$

Problem (5.40) contains $n = (n_t + 2)K + 1$ scalar variables, $m = (n_t + 3)K$ scalar constraints, and require the input data vector $\boldsymbol{p} = [n,\ m,\ \boldsymbol{w}_1^T,\ \ldots,\ \boldsymbol{w}_K^T,\ \bar{\boldsymbol{w}}_1^T,\ \ldots,\ \bar{\boldsymbol{w}}_K^T,\ \hat{\alpha}_1,\ \ldots,\ \hat{\alpha}_K,\ \ddot{\alpha}_1,\ \ldots,\ \ddot{\alpha}_K]$. According to these problem parameters, the complexity of achieving a per-iteration solution within the an accuracy ϵ is [69]

$$\mathcal{O}(1)\sqrt{m+n}\ \ln\left(\frac{\dim(\boldsymbol{p}) + \|\boldsymbol{p}\|_1 + \epsilon^2}{\epsilon}\right), \tag{5.41}$$

where $\mathcal{O}(1)$ is the complexity of a real operation. According to (5.41) and the aforementioned problem parameters, the per-iteration complexity asymptotically (as $n_t,\ K \to \infty$ and $n_t \gg K$) converges to

$$\mathcal{O}\bigg(Kn_t\big[\ln(2K^2 n_t) + \ln\big(\frac{1}{\epsilon}\big)\big]\bigg). \tag{5.42}$$

Obviously, from (5.39) and (5.42), the optimal PSP precoder has lower computational complexity than the suboptimal PSP precoders, where the main parameters are the total number of users K and the total number of transmit antennas n_t. More on the comparison between (5.39) and (5.42) will be explored in Section 5.5.

5.4 Differential Detection for Massive MIMO with Downlink Transmission

In this section, the differential encoding and decoding process for the downlink transmission in a massive MIMO system is discussed. Here, it is assumed that neither the transmitter nor the receiver has prior knowledge of the CSI.

5.4.1 Multiple Symbols Differential Detection

The simpler suboptimal method of implementing DM detection with massive MIMO is to encode the transmitted data differentially and to decode only the last two consecutive received symbols, e.g. $N = 2$, without any knowledge of the CSI. In contrast, the optimal method is to decode a block of N consecutive information symbols jointly without any knowledge of the CSI by performing MSDD, e.g. $N \gg 2$, which results in a 3dB performance improvement compared to DM detection [1] [45, 46]. In MSDD for the downlink system, the differential transmissions are implemented in blocks, in which each user k receives the sum of all the transmit waveforms of other users; then, the received signal blocks for each user must be detected independently. The measurements at the receiver are collected by spatial autocorrelation, thus the generalised likelihood ratio test (GLRT) optimisation criterion is adopted whereby the maximisation of the likelihood function is performed not only over the unknown symbols but also over unknown channels [70].

Now, the differential decoder for the downlink transmission is constructed. From the received signal $\boldsymbol{y}_k$ in (5.8) and using the differential encoding rule in (5.2), it follows

$$\boldsymbol{y}_k = \sqrt{p_k}\boldsymbol{h}_k^H \boldsymbol{u}_k \Big[b_{k,0}\ b_{k,1}\ b_{k,2} \cdots b_{k,N} \Big] + \bar{\boldsymbol{z}}_k \tag{5.43}$$

$$= \sqrt{p_k}\boldsymbol{h}_k^H \boldsymbol{u}_k b_{k,0} \Big[1\ s_{k,1}\ s_{k,1}s_{k,2} \cdots \prod_{i=1}^{N} s_{k,i} \Big] + \bar{\boldsymbol{z}}_k \tag{5.44}$$

$$= \quad \phi_k \boldsymbol{\alpha}_k + \bar{\boldsymbol{z}}_k \tag{5.45}$$

where $\boldsymbol{\alpha}_k = \big[1\ s_{k,1}\ s_{k,1}s_{k,2} \cdots \prod_{i=1}^{N} s_{k,i}\big]$ denotes the unknown information symbols, and $\phi_k = \sqrt{p_k}\boldsymbol{h}_k^H \boldsymbol{u}_k b_{k,0}$ contains the unknown channel coefficients $\boldsymbol{h}_k$ scaled by predefined initial symbol $b_{k,0}$, the transmit power p_k, and the transmit precoder $\boldsymbol{u}_k$. Since ϕ_k is unknown in the absence

[1] In MSDD, there is a 3dB gain when using large values of N, yet the cardinality of the search set grows exponentially with N, i.e., $|\mathcal{M}| = \mathcal{M}^{N+1}$. However, to achieve low complexity design, one can resort to using an edge computing platform or the conventional DM detection.

of channel knowledge $\boldsymbol{h}_k$, a possible alternative to the ML decoder to detect the information symbols $\boldsymbol{s}_k$ is the GLRT approach. When the $\bar{\boldsymbol{z}}_k$ is AWGN noise, the conditional probability density function (PDF) for the multivariate Gaussian distribution $(\boldsymbol{y}_k - \phi_k \boldsymbol{\alpha}_k)$ is [46]

$$\frac{1}{\sqrt{2\pi\sigma_{\bar{\boldsymbol{z}}_k}^2}} \exp\left(-\frac{\left(\boldsymbol{y}_k - \phi_k \boldsymbol{\alpha}_k\right)^H \left(\boldsymbol{y}_k - \phi_k \boldsymbol{\alpha}_k\right)}{2\sigma_{\bar{\boldsymbol{z}}_k}^2} \right). \tag{5.46}$$

The maximisation of the log-likelihood metric is performed not only over the candidate information symbols $\tilde{\boldsymbol{s}}_k$ of $\boldsymbol{s}_k$, but also over the candidate value $\tilde{\phi}_k$ of ϕ_k, and then it follows that [46]

$$\Lambda\left[\boldsymbol{y}_k | \tilde{\boldsymbol{s}}_k, \tilde{\phi}_k\right] = \boldsymbol{y}_k \tilde{\phi}_k^* \tilde{\boldsymbol{\alpha}}_k^H + \tilde{\phi}_k \tilde{\boldsymbol{\alpha}}_k \boldsymbol{y}_k^H - \tilde{\phi}_k \tilde{\boldsymbol{\alpha}}_k \tilde{\phi}_k^* \tilde{\boldsymbol{\alpha}}_k^H, \tag{5.47}$$

where $\tilde{\boldsymbol{\alpha}}_k = \left[1 \ \tilde{s}_{k,1} \ \tilde{s}_{k,1}\tilde{s}_{k,2} \cdots \prod_{i=1}^{N} \tilde{s}_{k,i}\right]$ is the candidate value for $\boldsymbol{\alpha}_k$. The log-likelihood metric in (5.47) can be reformulated to yield the equivalent metric

$$\begin{aligned} \Lambda\left[\boldsymbol{y}_k | \tilde{\boldsymbol{s}}_k, \tilde{\phi}_k\right] &= \boldsymbol{y}_k \tilde{\phi}_k^* \tilde{\boldsymbol{\alpha}}_k^H + \tilde{\phi}_k \tilde{\boldsymbol{\alpha}}_k \boldsymbol{y}_k^H \\ &\quad - \left(1 + \sum_{\tau=1}^{N} \prod_{i=1}^{\tau} |\tilde{s}_{k,i}|^2\right) |\tilde{\phi}_k|^2. \end{aligned} \tag{5.48}$$

Therefore, the GLRT-based decision metric for MSDD is given by

$$\hat{\boldsymbol{s}}_k = \arg\max_{\tilde{\boldsymbol{s}}_k} \left\{ \max_{\tilde{\phi}_k} \left\{ \Lambda\left[\boldsymbol{y}_k | \tilde{\boldsymbol{s}}_k, \tilde{\phi}_k\right] \right\} \right\}. \tag{5.49}$$

Then, solving (5.49) by first keeping $\tilde{\boldsymbol{s}}_k$ fixed and compute

$$\Lambda\left[\boldsymbol{y}_k | \tilde{\boldsymbol{s}}_k\right] = \max_{\tilde{\phi}_k} \left\{ \Lambda\left[\boldsymbol{y}_k | \tilde{\boldsymbol{s}}_k, \tilde{\phi}_k\right] \right\}. \tag{5.50}$$

To calculate (5.50), Taylor's theorem can be used for variational techniques by imposing

$$\tilde{\phi}_k = \phi_k^\star + \lambda_k \epsilon_k, \tag{5.51}$$

where $\phi_k^\star \in \mathbb{C}$ denotes the optimal solution to ϕ_k, $\epsilon_k \in \mathbb{C}$ measures the error of $\tilde{\phi}_k$ from the optimal solution, and $\lambda_k \in \mathbb{R}$ is the error real coefficient. Substituting (5.51) into (5.48) and then taking the first-order derivative of $\Lambda\left[\boldsymbol{y}_k | \tilde{\boldsymbol{s}}_k, \tilde{\phi}_k\right]$ with respect to λ_k at zero, which yields

$$\frac{\partial}{\partial \lambda_k} \Lambda\left[\boldsymbol{y}_k | \tilde{\boldsymbol{s}}_k, \tilde{\phi}_k\right] \Bigg|_{\lambda_k = 0} = \epsilon_k^* \boldsymbol{y}_k \tilde{\boldsymbol{\alpha}}_k^H + \epsilon_k \tilde{\boldsymbol{\alpha}}_k \boldsymbol{y}_k^H$$

$$- \left(1 + \sum_{\tau=1}^{N} \prod_{i=1}^{\tau} |\tilde{s}_{k,i}|^2\right) \left(\epsilon_k^* \phi_k^\star + \epsilon_k (\phi_k^\star)^*\right). \quad (5.52)$$

By setting the derivative in (5.52) equal to zero, the optimal value of ϕ_k is obtained as

$$\phi_k^\star = \frac{\boldsymbol{y}_k \tilde{\boldsymbol{\alpha}}_k^H}{1 + \sum_{\tau}^{N} \prod_{i=1}^{\tau} |\tilde{s}_{k,i}|^2} \quad (5.53)$$

Then, substituting the optimal value $\phi_k^\star$ in (5.53) by $\tilde{\phi}_k$ in (5.48) yields

$$\begin{aligned} \Lambda\left[\boldsymbol{y}_k | \tilde{\boldsymbol{s}}_k\right] &= \frac{\boldsymbol{y}_k \tilde{\boldsymbol{\alpha}}_k^H \tilde{\boldsymbol{\alpha}}_k \boldsymbol{y}_k^H}{1 + \sum_{\tau=1}^{N} \prod_{i=1}^{\tau} |\tilde{s}_{k,i}|^2} \\ &= \frac{\tilde{\boldsymbol{\alpha}}_k \boldsymbol{y}_k^H \boldsymbol{y}_k \tilde{\boldsymbol{\alpha}}_k^H}{1 + N} \end{aligned} \quad (5.54)$$

Note that the channel estimate is not explicitly used during the stages of MSDD. When using the M-ary PSK constellation, the MSDD detection problem in (5.54) can be simplified as [46]

$$\hat{\boldsymbol{b}}_k = \underset{\tilde{\boldsymbol{b}}_k \in \mathcal{M}^{N+1}, \tilde{b}_o = 1}{\arg\max} \; \tilde{\boldsymbol{b}}_k \boldsymbol{Y}_k \tilde{\boldsymbol{b}}_k^H \quad (5.55)$$

$$= \underset{\tilde{\boldsymbol{b}}_k \in \mathcal{M}^{N+1}, \tilde{b}_o = 1}{\arg\max} \; \Re\left\{ \sum_{\tau=1}^{N} \tilde{b}_{k,\tau}^* \sum_{l=0}^{\tau-1} \tilde{b}_{k,l} \cdot y_{k,l,\tau} \right\}, \quad (5.56)$$

where $\boldsymbol{Y}_k = \boldsymbol{y}_k^H \boldsymbol{y}_k \in \mathbb{C}^{(N+1)\times(N+1)}$ is the autocorrelation matrix of the received signal comprised of the correlation coefficients $y_{k,l,\tau}$, $\tau = 1, \ldots, N$, $l = 0, \ldots, \tau - 1$, between the lth and the τth received differential signals. To get the information symbols $\boldsymbol{s}_k$, it can be seen that $\boldsymbol{s}_k$ is directly obtained as

$$s_{k,\tau} = b_{k,\tau} \cdot b_{k,\tau-1}^*. \quad (5.57)$$

In (5.55), the differential decoder uses one side of the complex-conjugate symmetry of the correlation coefficients, thus $y_{k,l,\tau} = y_{k,\tau,l}^*$. Further, the diagonal elements of $\boldsymbol{Y}_k$ can be neglected as they do not influence the decision metrics, i.e., $y_{k,l,l} = y_{k,\tau,\tau} = 0$.

Example 7. *For $N = 3$, it follows*

$$\boldsymbol{Y}_k = \begin{bmatrix} 0 & y_{k,0,1} & y_{k,0,2} & y_{k,0,3} \\ y_{k,0,1} & 0 & y_{k,1,2} & y_{k,1,3} \\ y_{k,0,2} & y_{k,1,2} & 0 & y_{k,2,3} \\ y_{k,03,l} & y_{k,1,3} & y_{k,2,3} & 0 \end{bmatrix}. \quad (5.58)$$

Now assuming that at time instant l and τ, the BS transmits the differential signals $b_{k,l}$ and $b_{k,\tau}$. Therefore, from (5.8) and (5.56), the correlation coefficients ($y_{k,l,\tau} = y_{k,\tau,l}$) constructed from the autocorrelation received signal matrix can be represented as

$$\begin{aligned} y_{k,l,\tau} &= y_{k,l}^* y_{k,\tau} \\ &= \left(\sqrt{p_k}\boldsymbol{h}_k^T \boldsymbol{u}_k^* b_{k,l}^* + \bar{z}_{k,l}^*\right)\left(\sqrt{p_k}\boldsymbol{h}_k^H \boldsymbol{u}_k b_{k,\tau} + \bar{z}_{k,\tau}\right) \\ &= p_k |\boldsymbol{h}_k^H \boldsymbol{u}_k|^2 \cdot b_{k,l}^* b_{k,\tau} + \eta_{k,l,\tau}, \end{aligned} \tag{5.59}$$

where $p_k|\boldsymbol{h}_k^H \boldsymbol{u}_k|^2$ represents the captured energy of the signal of the kth user and

$$\begin{aligned} \eta_{k,l,\tau} = &\sqrt{p_k}\boldsymbol{h}_k^T \boldsymbol{u}_k^* b_{k,l}^* \bar{z}_{k,\tau} \\ &+ \sqrt{p_k}\bar{\boldsymbol{z}}_{k,l}^* \boldsymbol{h}_k^H \boldsymbol{u}_k b_{k,\tau} + \bar{z}_{k,l}^* \bar{z}_{k,\tau}, \end{aligned} \tag{5.60}$$

are composed of all terms corrupted by noise and MAI.

5.4.2 Decision Feedback Differential Detection

In order to improve the performance further, DFDD is adopted in this chapter. This approach leads to better performance compared to MSDD. Different from [47], the DFDD is derived for the downlink transmission instead of the uplink. In DFDD, the decisions are made successively, adding all previous decisions in the decision of the current symbol. In this decoding algorithm, the decoder detects symbols one by one. After finding the best candidate for the first symbol, the effects of this symbol in all of the receiver equations are added and considered. Then, the second symbol is detected from the new sets of equations. The effects of the first and second detected symbols are added and then considered to derive a new set of equations. The process continues until all symbols are detected. Of course, the order in which the symbols are detected will impact the end solution. The algorithm includes three steps, i.e. decision, process, and ordering.

5.4.2.1 Decision Process

From the description given above and starting with $b_{k,0} = 1$, the decision process means that the information symbols in (5.56) are detected one by one as

$$\hat{b}_{k,\tau} = \underset{\tilde{b}_{k,\tau} \in \mathcal{M}^{N+1}}{\arg\max}\, \Re\Big\{ \tilde{b}_{k,\tau}^* \sum_{l=0}^{\tau-1} \tilde{b}_{k,l} \cdot y_{k,l,\tau} \Big\} \tag{5.61}$$

$$= \exp\left(j \cdot \theta_{\text{PSK}}\Big\{ \sum_{l=0}^{\tau-1} \tilde{b}_{k,l} \cdot y_{k,l,\tau} \Big\}\right), \tag{5.62}$$

where

$$\theta_{\text{PSK}}\{x\} \overset{\text{def}}{=} \frac{2\pi}{M} \cdot \left\lfloor \frac{M}{2\pi} \cdot \arg(x) \right\rceil, \tag{5.63}$$

and

$$\Delta\theta_{\text{PSK}}\{x\} \overset{\text{def}}{=} \left| \arg\left\{ \exp\Big(j \cdot \big(\arg(x) - \theta_{\text{PSK}}(x)\big)\Big)\right\}\right| \tag{5.64}$$

quantises the phase of a complex number $x \in \mathbb{C}$ to the M phase values of M-ary PSK, and computes the quantisation error, respectively. The operation $\lfloor x \rceil$ in (5.63) takes as input a real number x and gives as output a reduction into the interval $[\pi, 2\pi]$. The purpose of this step is to decide which transmitted symbol to detect at each stage of the decoding.

5.4.2.2 Optimum Decision Ordering

It is well known from decision feedback equalisation in MIMO systems, also known as BLAST [71], that sorting the decisions in an optimised order improves performance. The symbol with lowest quantisation error in (5.64) is the best in this step. The decision order can be achieved by reordering the columns and rows of the $\boldsymbol{Y}_k$ matrix. That is, the best transmitted symbols in the $\boldsymbol{Y}_k$ matrix are denoted by $(\hat{\tau}_0, \hat{\tau}_1, \cdots, \hat{\tau}_N)$, where $\hat{\tau}_i, \hat{\tau}_l \in \{0, \cdots, N\}$, $\hat{\tau}_i \neq \hat{\tau}_l$ for $i \neq l$. Then, the symbols transmitted in the $\hat{\tau}_i$th index are defined by $(b_{k,\hat{\tau}_0}, b_{k,\hat{\tau}_1}, \cdots, b_{k,\hat{\tau}_i}, \cdots, b_{k,\hat{\tau}_N})$.

Now, the initial transmitted symbol is setted to identity, i.e., $b_{k,\hat{\tau}_0} = 1$. Then, the first decided symbol should be the $\hat{\tau}_1$th symbol, where

$$[\hat{\tau}_0, \hat{\tau}_1] = \underset{\substack{l \in \{0,\cdots,N\}, \tau \in \{1,\cdots,N\} \\ l \leq \tau}}{\arg\min} \left| \Delta\theta_{\text{PSK}}\{y_{k,l,\tau}\} \right|, \tag{5.65}$$

and the estimate for the $b_{k,\hat{\tau}_1}$ symbol is obtained from

$$b_{k,\hat{\tau}_1} = \exp\left(j \cdot \theta_{\text{PSK}}\Big\{ y_{k,\hat{\tau}_0,\hat{\tau}_1} \Big\}\right). \tag{5.66}$$

Taking the previous decision into account, the symbol that is decided next can be obtained successively from

$$\hat{\tau}_i = \underset{\substack{\tau \in \{1,\cdots,N\} \\ /\{\hat{\tau}_0,\cdots,\hat{\tau}_{i-1}\}}}{\arg\min} \left| \Delta\theta_{\text{PSK}}\Big\{ \sum_{l=0}^{i-1} b_{k,\hat{\tau}_l} \cdot y_{k,\hat{\tau}_l,\tau} \Big\} \right|, \tag{5.67}$$

and its value can be obtained from

$$b_{k,\hat{\tau}_i} = \exp\left(j \cdot \theta_{\text{PSK}}\Big\{\sum_{l=0}^{i-1} b_{k,\hat{\tau}_l} \cdot y_{k,\hat{\tau}_l,\hat{\tau}_i}\Big\}\right). \tag{5.68}$$

This ordering scheme has attempted to provide reliable decisions for the first decided symbols, which will impact the decision for subsequent symbols, and thus improve performance. Further, it must be noted that the actual realisations of the channel vectors $\{\boldsymbol{h}_k\}$ are not needed to decode the information signals.

5.5 Simulation Results and Discussion

In this section, the performance of the differential massive MIMO downlink transmission is examined. The channel is assumed to be modeled as quasi-static, where the block fading channel between the transmitter and receiver is constant (but unknown) during N successive channel uses, i.e., the block length of the coherence time intervals. The fast fading coefficients for each user $\tilde{\boldsymbol{h}}_k = [\tilde{h}_{k,1}, \cdots, \tilde{h}_{k,n_t}]^T$ are mutually independent and modeled as independent and identically distributed (i.i.d.) complex Gaussian random variable with zero-mean and unit-variance, i.e., $\tilde{h}_{k,m} \sim \mathcal{CN}(0,1)$.

Throughout this section, it is assumed the following; urban area cellular radio model for γ, one receive antenna per user, the noise power $\sigma^2_{z_k} = 0$ dB, the constellation size is 4-PSK, the length of the transmission block is set to $N = 200$, and the DFDD detection technique is used for differential detection. Table 5.1 shows the values of PSP parameters to be used in (5.14) whenever needed throughout the simulation section. Note that using ζ without the superscript k means that the values of ζ are equal for all users, i.e., $\zeta_1 = \cdots = \zeta_K = \zeta$. The array geometry of BS antennas is co-located uniform linear array as in Fig. 5.2. The Monte Carlo simulation is used to evaluate the performance in terms of BER.

5.5.1 Single-User Scenario

The BER performance curve is first simulated and plotted for only one user. It is assumed that the user's location is in front of the centre of the antenna array, i.e., $m_1 = 50$. The BS has $n_t = 100$ transmit antennas. This case is examined using the three proposed precoders, e.g., matched PSP precoder, orthogonal PSP precoder, and optimal PSP precoder. In addition, these precoders are compared against the unity precoder (equal power allocation), where the precoder vector elements are all set to one, i.e., $\{\boldsymbol{u}_k\} = \mathbf{1}_{n_t}$ and then normalised. The channel parameter

Table 5.1. The values of PSP parameters to be used in (5.14).

PSP Parameters List					
Path loss exponent γ	Antenna spacing l_a	Orthogonal distance l_k	Relative distance $l_{k,r}$	Peak amplitude β_k	Channel variance ζ_k
3.6	0.3m	3m	10m	0.01	5
3.6	0.3m	5m	16m	0.003	8
3.6	0.3m	10m	33m	0.0002	16
3.6	0.3m	15m	50m	0.00005	26

is set to $\zeta = 8$. When there is no interference, Fig. 5.3 shows that the performance of the proposed precoders schemes, e.g., matched, orthogonal, and optimal precoders, outperforms the one that does not perform any kind of optimisation for the precoding vector, e.g., the unity precoder. Clearly, in the interference-free system, the performance of the optimal PSP precoder is slightly better than the other two precoders but the difference is very small. It should be noted that in a coherent system, it is well known that the matched (to the channel) filter maximises the SNR for the single user case. This is valid for both conventional and massive MIMO systems. However, in a noncoherent system, the matched PSP precoder is matched only to the PSP and not to the channel itself. Therefore, the matched PSP precoder does not necessarily maximise the SNR. In the optimal PSP precoder design, the optimiser tends to allocate the power to the channels that have significant gains. In other words, as the PSP coefficients are positive, the optimised precoder (that maximises the SNR and improves the BER) will have only coefficients corresponding to the largest coefficients of the PSP greater than zero and the rest are equal to zero.

5.5.2 Multiple-User Scenario

Fig. 5.4 shows the coefficients of the proposed precoders, i.e., matched, orthogonal, and optimal PSP in the case of $K = 3$ users and $n_t = 100$. The users are placed in front of the uniform array at equal distance l_k from the BS but with different positions (angles) $m_1 = 20$, $m_2 = 50$, and $m_3 = 80$. Since it is assumed l_k is equal for all users, then $\zeta_1 = \zeta_2 = \zeta_3 = \zeta = 16$. In

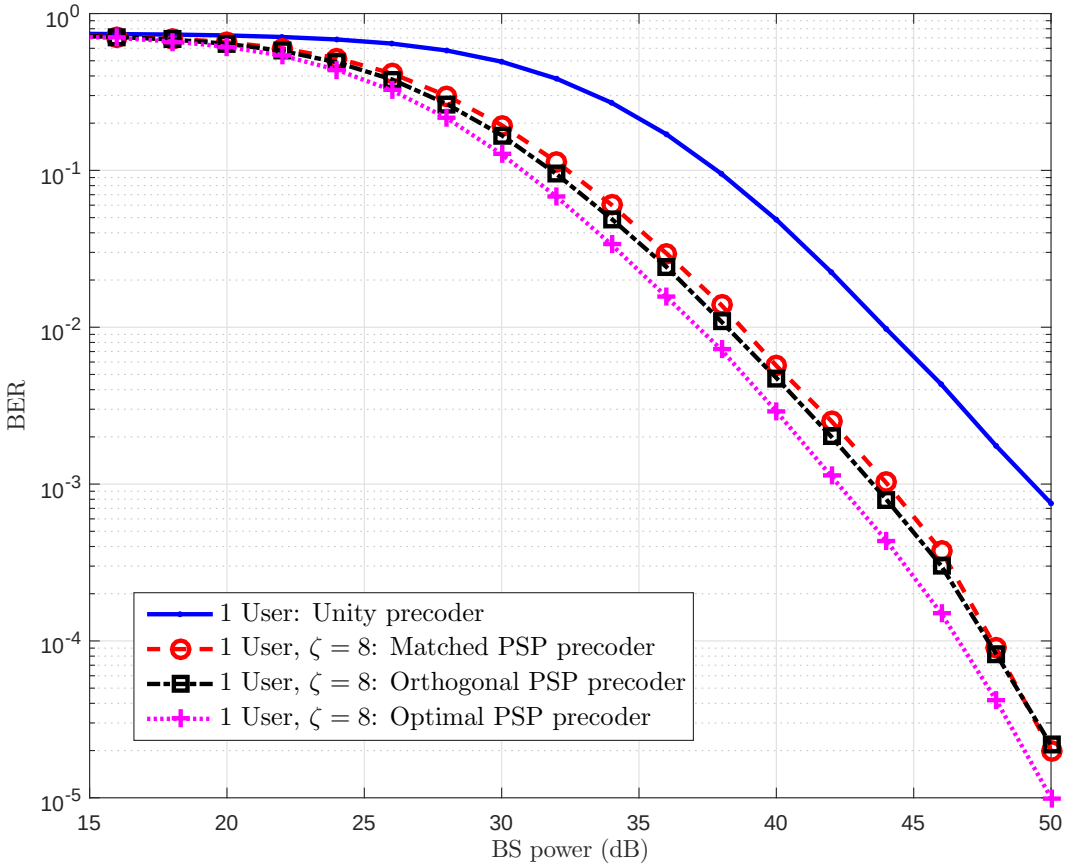

Figure 5.3. BER performance of the proposed differential MIMO downlink transmission with single user. $n_t = 100$, $\zeta = 8$.

Fig 4-(a), the BS first uses (5.14) to generate the PSP, $\{\boldsymbol{g}_k\}$, for each user, in which the BS uses them as an input to the three designed precoders. Fig 4-(a) shows the generated PSP for the three users (blue: user 1; red: user 2; black: user 3). It is evident that in the matched PSP precoder in Fig 4-(b), the precoder coefficients for the three users overlap significantly. For the orthogonal precoder in Fig 4-(c), the overlap between the precoder coefficients is reduced by using the Gram-Schmidt process. In the optimal PSP precoder in Fig 4-(d), the overlap between the precoder coefficients is minimised and the user is mostly separated. It is worth mentioning that if the following three conditions are satisfied, namely n_t is very large, ζ_k is small, and l_k is small, then $\boldsymbol{g}_k^T \boldsymbol{u}_q \approx 0$ for $k \neq q$. The value of ζ_k is affected by the user's distance l_k from the BS, the shorter the user's distance to the BS the smaller the value of ζ_k, which minimises the interference between users.

In Fig. 5.5, the performance of the proposed PSP precoders is compared in terms of BER. It is assumed that $K = 3$, $m_1 = 20$, $m_2 = 50$, $m_3 = 80$, $n_t = 100$. In Fig. 5.5, for any value of ζ, the performance of the optimal PSP precoder outperforms the other precoders. The

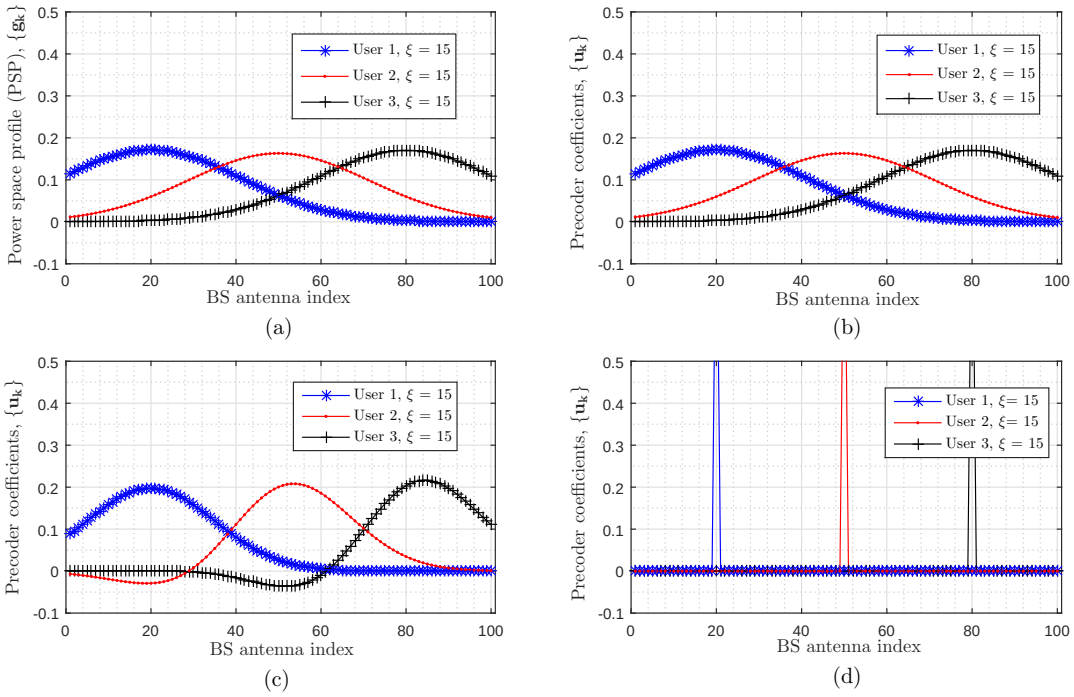

Figure 5.4. The proposed coefficients for the precoders of different users for $K = 3$, $n_t = 100$, and $\zeta = 16$. (a) Normalized PSP of different users, (b) Normalized matched PSP precoder's coefficients, (c) Normalized orthogonal PSP precoder's coefficients, (d) Normalized optimal PSP precoder's coefficients.

matched PSP precoder is not robust against interference at high BS power and thus has the worst performance. In the case of $\zeta = 5$ for all users, the performance of the precoders is almost the same and this is because of using small value of ζ in which the users do not overlap and hence are separated very well. In Fig. 5.5 also, in the presence of interference between users, the value of the power profile parameters such as ζ can impact the precoders' performance. Fig. 5.5 shows the effect of adjusting ζ on the performance of the matched, orthogonal, and optimal precoders. Note that when the value of the channel variance is increased for all users from $\zeta = 5$ to $\zeta = 8$ and then $\zeta = 16$, the power profile significantly overlaps between users hence causing a degradation in the system performance. Hence, for large orthogonality between users' channels (small value of channel parameter ζ), the performance of matched precoder design is close to the optimal design performance. The larger the orthogonality the closer the performance.

Fig. 5.6 investigates the impact of increasing the number of users on the system performance in terms of BER using the three proposed precoders. It is assumed that: $n_t = 100$, $\zeta = 5$, and $K = 2$, $K = 4$, and $K = 6$. For $K = 2$, the positions are set to [25 75], for $K = 4$, the positions are set to [20 40 60 80], and for $K = 6$, the positions are set to [15 30 45 60 75 90]. It is shown that differential massive MIMO systems with fewer users outperform those with a large number of users. However, using an optimal PSP precoder with the most appropriate

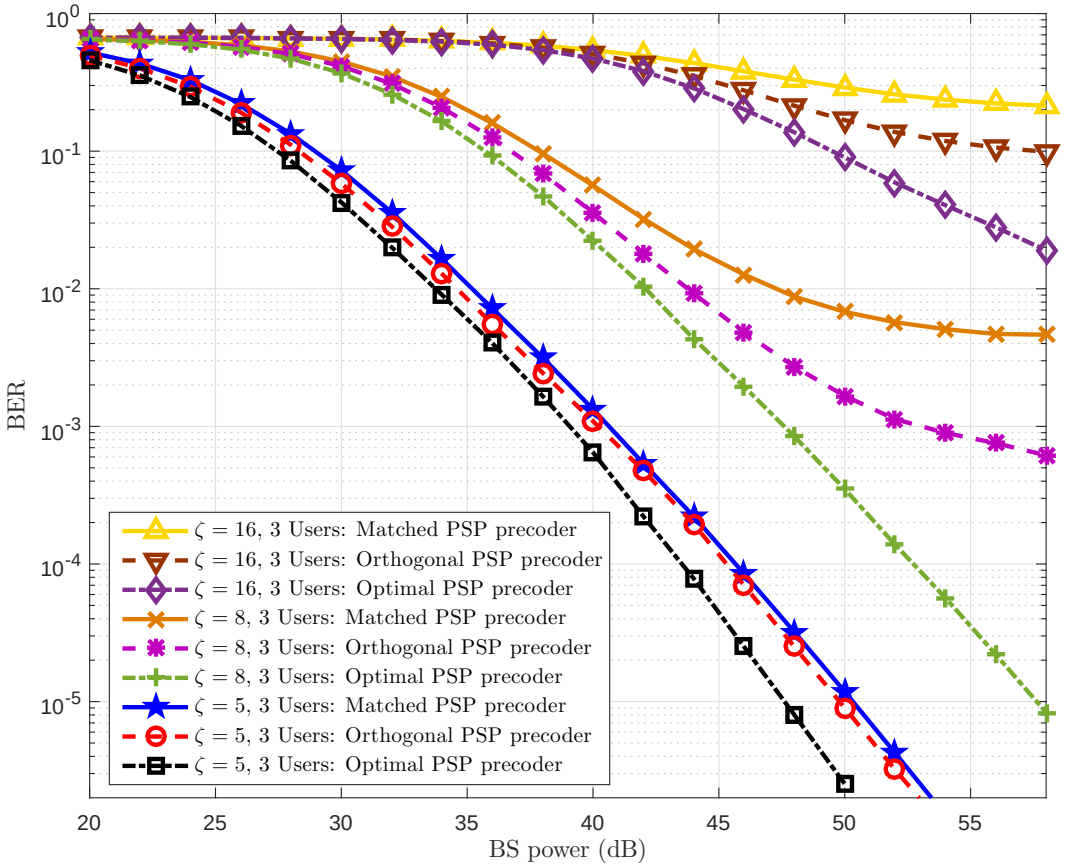

Figure 5.5. BER performance of the proposed differential MIMO downlink transmission with $K = 3$, $n_t = 100$. The values of PSP parameter are $\zeta = 5$, $\zeta = 8$ and $\zeta = 16$.

number of n_t and/or value of ζ can minimise the overlap between users and thereby reduce loss of performance.

Fig. 5.7 examines the influence of increasing the number of transmit antennas, e.g., $n_t = 100$ to $n_t = 200$, on the system performance using the optimal PSP precoder. Three users, $K = 3$, are placed in front of the uniform array at different positions [30 60 90] and different distance l_k from the BS, which yields $\zeta_1 = 5$, $\zeta_2 = 8$, and $\zeta_3 = 16$. From Fig. 5.7, it can be seen that differential massive MIMO systems with higher number of transmit antennas outperform those with lower number of antennas. Therefore, as $n_t \to \infty$ the degree of orthogonality between users becomes large which can minimise the interference between users and improve the overall performance of the system. The larger the number of transmit antennas the better the performance.

Fig. 5.8 and Fig. 5.9 show the computational complexity of the system. In Fig. 5.8, first the number of users is set to $K = 6$ and then the number of transmit antenna n_t is increased. Similarly, in Fig. 5.9, the number of transmit antennas is fixed to be $n_t = 100$ while the number

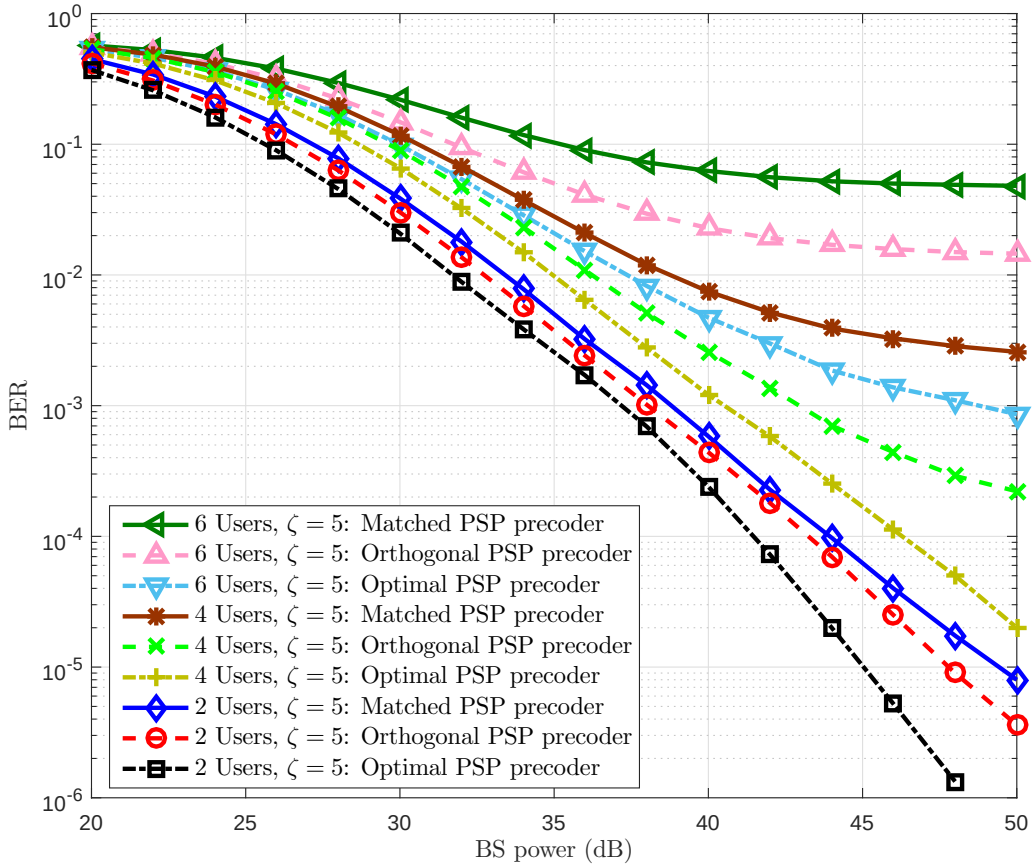

Figure 5.6. BER performance of the proposed differential MIMO downlink transmission with $n_t = 100$, $\zeta = 5$. Users cases are $K = 2$, $K = 4$, and $K = 6$.

of users in the system increases gradually. From both figures, the computational complexity of the suboptimal PSP precoders are higher than the optimal PSP precoder. It is also evident that varying the number of transmit antennas at the BS has higher impact on the complexity than varying the number of users. Therefore, the optimal PSP precoder yields a low complexity scheme while providing good performance.

5.6 Summary

This chapter proposed three precoding schemes, namely the matched, orthogonal and optimal PSP precoders, for downlink transmission in massive MIMO systems with differential encoding and detection. With a large number of transmit antennas at the BS and full knowledge of the PSP, the proposed low-complexity downlink precoding techniques allow MAI between users to be eliminated. In a multiuser scenario, the optimal PSP precoder can effectively separate the data streams of different users, thus enhancing the system performance. In the detection scenario, the DFDD technique is used to detect the differential information signals. Simulations

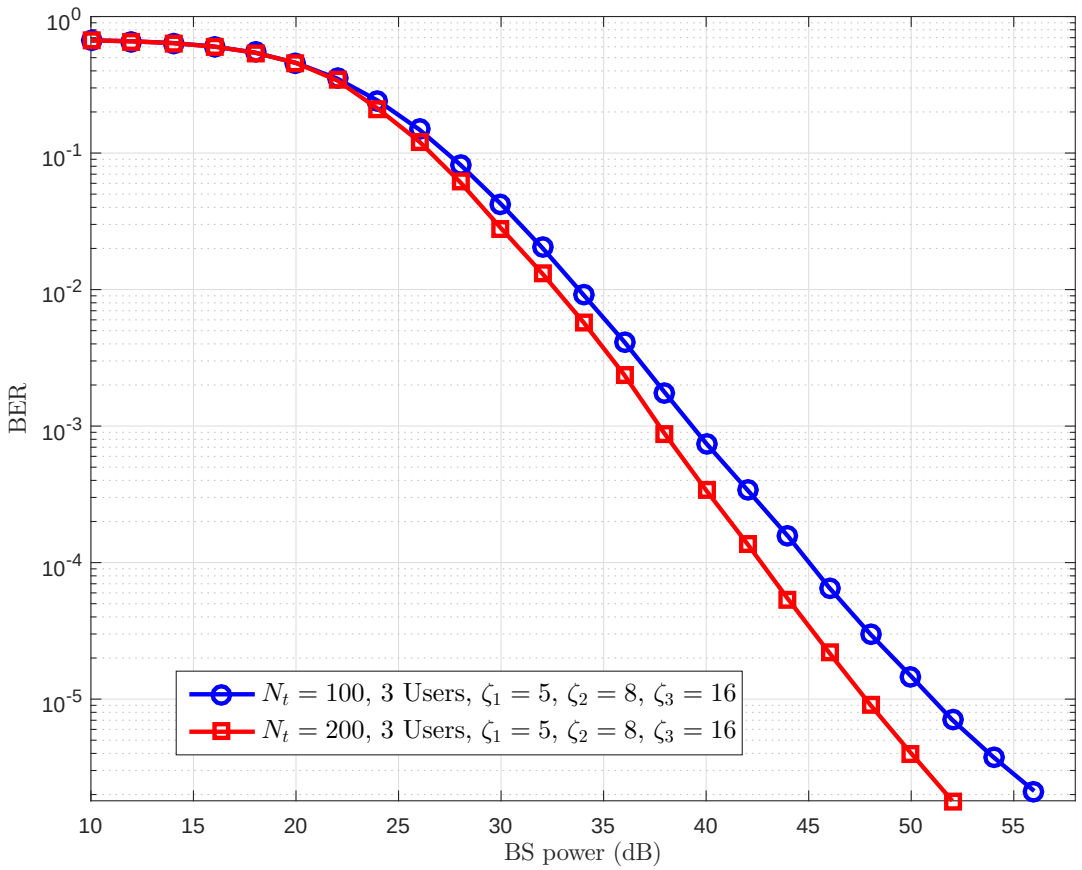

Figure 5.7. BER performance of the proposed differential MIMO downlink transmission with $K = 3$, $n_t = 100$, and $n_t = 200$, using different values of channel variance between users; $\zeta_1 = 5$, $\zeta_2 = 8$, and $\zeta_3 = 16$.

show that the proposed schemes are effective precoding techniques for a massive MIMO system in a scenario where the channel is unknown at both the transmitter and receiver.

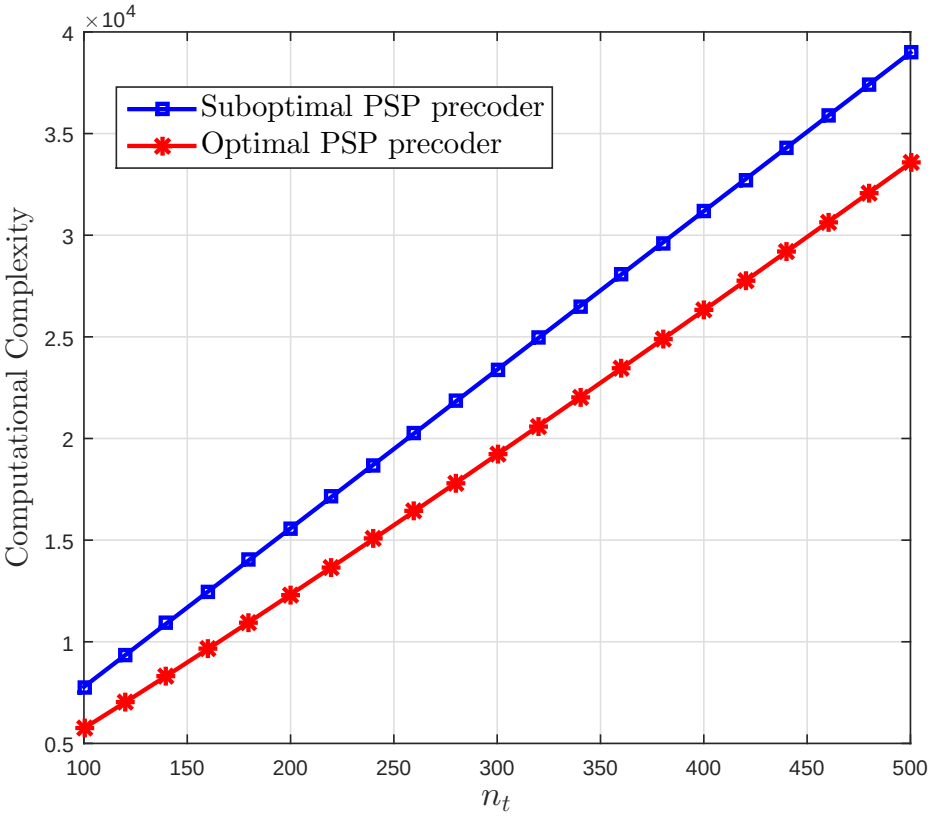

Figure 5.8. Comparison of the computational complexity for suboptimal PSP precoders and optimal PSP precoder with $K = 6$ and $\epsilon = 0.5$.

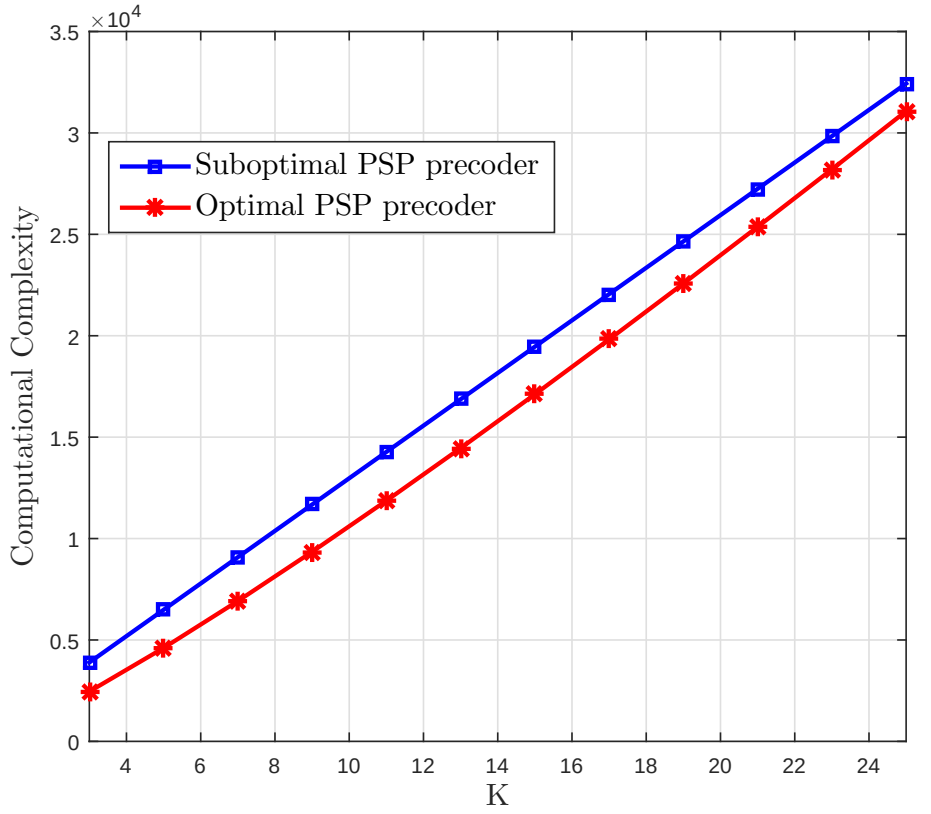

Figure 5.9. Comparison of the computational complexity for suboptimal PSP precoders and optimal PSP precoder with $n_t = 100$ and $\epsilon = 0.5$.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This thesis focused on noncoherent design in downlink transmission MU-MIMO systems and aimed to achieve three objectives: improve the overall system performance, obtain a simple low complexity transceiver at the users' devices, and accomplish downlink transmission schemes that do not require either the transmitter or receiver to know the channel. To this end, three noncoherent MU-MIMO systems have been considered: low rate differential STBC, high rate differential algebraic STBC, and differential downlink transmission for massive MIMO system. In the following, a summary and the main results of each chapter are given:

In Chapter 1, a general introduction to the thesis along with the contributions of the study has been presented.

In Chapter 2, a general background to MIMO systems has been presented. In addition, the noncoherent MIMO systems have been introduced along with the best practice model for a differential modulation system. At the end of Chapter 2, previous focused related work has been discussed.

In Chapter 3, a simple low complexity differential STBC scheme for MU-MIMO with downlink transmission has been proposed. Essentially, the impact of the receiver channel estimation process, overhead problem, and achieving a full diversity system can be potentially solved by using differential STBC. The precoding technique of the transmit signal such as BD and OS, designed to eliminate interference between users, provides in effect a single-user link to the receiver. This technique extends to the DM case in which the receiver does not need any knowledge of the CSI to perform information detection when using the BD scheme with DM. Furthermore, combining the OS scheme with DM will result in a system that does not require any CSI knowledge at either the transmitter or the receivers. The improvement in BER performance from the MIMO technique utilized here occurs when there are sufficient transmit antennas in the system to achieve full diversity with the set of receive antennas. The impact of differential modulation on link performance is 3 dB lower in SNR relative to the coherent modulation case. The reduced complexity of the receivers' equipment achieved by avoiding coherent signal processing can in many cases be more important than compensating for this SNR impact. The BD scheme is more complex than the OS scheme; however, the BD scheme has a higher throughput. Moreover, it was shown that the OS is more robust against precoding errors compared to the BD scheme.

Chapter 4 considers the design of a high data rate differential perfect algebraic STBC scheme for MU-MIMO with downlink transmission. The Caylay differential STBC that has

been used in this chapter does not require either the transmitter or receiver to know the channel. The Caylay differential STBC simplifies the receivers' equipment, provides an effective high rate data transmission scheme, and works with any number of transmit and receive antennas. Moreover, it has been shown that to achieve a full-rate full-diversity differential MIMO system, a perfect algebraic STBC must be combined with the Caylay differential STBC. At the BS, the OS matrices are used to separate the data streams of multiple users and thus enable simple user detection. Using OS matrices along with the Caylay differential STBC will form a system model that does not require any knowledge of the CSI at both ends, i.e., BS and receivers. For this system, after despreading the received signal, near-optimal sphere decoding is performed to decode the signals at the receiver of each user. Since the system model with high data rate involves a high complexity decoding process, the SD technique has been proposed to solve this problem but with less complexity. The proposed schemes yield low complexity transceivers while also providing high rate with good performance. However, the system in this chapter has higher computational complexity because of its higher rate.

Chapter 5 proposes a noncoherent downlink transmission model for massive MIMO system. Noncoherent signalling mitigates one of the key problems that arises in massive MIMO systems. When training-based schemes are used in the uplink of these systems, multiple adjacent cells will reuse the same training pilots symbols, resulting in pilot contamination. Using massive MIMO with a precoding technique creates sharp beams toward end users which will greatly increase the network capacity and improve the BER. Accordingly, three precoding schemes, namely the matched, orthogonal, and optimal PSP precoders have been implemented in this chapter for downlink transmission with differential encoding and detection. Conventional transmit precoding design requires channel knowledge at the transmitter. Instead in this work, differential transmit precoding design is considered which avoids the need for explicit channel estimation at the BS. For this, PSP knowledge is used to design the transmit precoder for each user. With a large number of transmit antennas at the BS and full knowledge of the PSP, the proposed low-complexity downlink precoding techniques allow MAI between users to be eliminated. In a multiuser scenario, the optimal PSP precoder can effectively separate the data streams of different users, thus enhancing the system performance. In the detection scenario, the DFDD technique is used to detect the differential information signals. Simulations show that the proposed schemes are effective precoding techniques for a massive MIMO system in a scenario where the channel is unknown at both the transmitter and receiver.

Appendix B presents the design of secure noncoherent downlink transmission for cell-free differential massive MIMO systems with passive eavesdropper (EV) to jointly improve the

power and data transfer. This system consists of a large number of randomly located access points (APs) that cooperate via a central processing unit and serve multiple users and a single information-untrusted multiple-antenna passive energy-harvester (EH). The EH is interested in energy harvesting, however, it could act as an information eavesdropper by overhearing the signal intended for a certain user. The secrecy rate per AP for a noncoherent system is non-linear in terms of the transmit power elements and that imposes new challenges in formulating a convex power control problem. To overcome with this problem, a new method of balancing the transmit power for APs by reformulating the non-convex problem into a convex problem is derived. To this end, optimal and suboptimal solutions for the constrained noncoherent secrecy problem is also provided.

6.2 Future Work

This thesis focused on the design of a noncoherent downlink transmission system in three different MIMO systems scenarios. However, there are still open gaps regarding the design of efficient and effective noncoherent massive MU-MIMO systems. It is of interest to investigate the following research points:

- MIMO block fading channels arise in many communication scenarios [30]. However, finding transmission techniques that enable the highest rate to be noncoherently communicated over these MIMO channels remains an open gap for research. For example, the authors in [16] showed that the optimal capacity of a noncoherent MIMO system can be geometrically interpreted as sphere packing in the Grassmann manifold at high SNRs. However, the characteristic behaviour of the optimal noncoherent MIMO capacity is still unknown at low and medium SNRs.

- In coherent MIMO systems, information symbols are generated and constructed as a linear combination from QAM and PSK constellations. In contrast, in a noncoherent MIMO system, information symbols are constructed independently from an isotropic distribution and mapped onto one of the complex unitary code matrices [17, 29], which implies that the information symbols are generated in a different way than far a coherent system, so the results of linearity do not apply directly. This transition task of linearity results in implementation challenges such as coding design, labeling, and receiver detection.

- The proposed system in Chapter 4 provides a high data rate noncoherent downlink transmission MU-MIMO system. However, the system in this chapter has higher computa-

tional complexity because of its higher rate. Hence, considering the low computational complexity model for noncoherent downlink transmission MU-MIMO case is a promising avenue of research.

- One of the challenges for a noncoherent downlink transmission MIMO system is the peak-to-average power ratio (PAPR). In differential encoding, the input transmit signals are transformed to unitary code matrices which directly affects the dynamic range and the power amplifier used at the BS. The impact of this problem has yet to be explored.

- In linear uniform collocated massive MIMO antenna geometrical model, all users in each cell are served by an array of colocated antennas mounted at the BS. This design may experience antenna correlations, low coverage, and long dimensions array problems essentially for massive MIMO. Distributed array design such as random cell free array design, planar array design, and cylindrical array design can offer a scalable implementation with much higher probability of coverage than the conventional linear colocated array design [23]. Hence, designing a scalable array model for noncoherent massive MIMO system is a very interesting topic for future work.

- Recently, cell-free massive MIMO systems, where the service antennas are distributed over a large area, have attracted interest from many researchers due to their ability to offer much higher probability of coverage than the collocated massive MIMO [72]. In addition, they have the ability to alleviate the problems of antenna correlations and deep shadowing fading [73]. Therefore, considering and exploring noncoherent MIMO systems in a cell-free context could motivate researchers. Furthermore, the capacity of noncoherent cell-free massive MIMO systems and comparing it with that of coherent cell-free massive MIMO systems needs to be explored.

  The array dimensions and user positions chosen are strange. A 100 element array with 0.5m spacing is 50m long but the user ranges are much smaller than this

- Lately, physical layer security has proven to be a promising security technique to enhance the secrecy of the system and prevent any traffic interception attack from eavesdroppers. Hence, designing a secrecy noncoherent massive MIMO system has yet to be explored.

- The PSP profile estimation for noncoherent massive MIMO, in Chapter 5, is less challenging than estimating the actual channel state information, in which the PSP can remain constant over many coherence time intervals. However, the actual estimation process of PSP in real wireless communication systems has yet to be explored.

- Finally, a Rician fading channel is more common and suitable for small cell architectures. The model of a noncoherent massive MIMO system for Ricean fading is not always the same as how it is designed for a Rayleigh fading channel. Hence, it is of interest to test and investigate a Rician fading channel with a noncoherent massive MIMO system.

Appendix A

Proof of Theorem 1

The expected value of SINR_k at the kth user can be expressed as follows

$$\mathrm{SINR}_k = \mathbb{E}\left\{\frac{p_k|\boldsymbol{h}_k^H\boldsymbol{u}_k|^2}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\} \tag{A.1}$$

$$= \left(p_k\,\mathbb{E}\left\{|\boldsymbol{h}_k^H\boldsymbol{u}_k|^2\right\}\right)\mathbb{E}\left\{\frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\}. \tag{A.2}$$

To calculate the expected value of the norm $|\boldsymbol{h}_k^H\boldsymbol{u}_k|^2$, we first expand it as follows

$$|\boldsymbol{h}_k^H\boldsymbol{u}_k|^2 = \sum_{m=1}^{n_t}|\tilde{h}_{k,m}|^2 g_{k,m}u_{k,m}^2 + \sum_{\mathbb{I}}\tilde{h}_{k,i}^*\tilde{h}_{k,j}\sqrt{g_{k,i}g_{k,j}}u_{k,i}u_{k,j}, \tag{A.3}$$

where $\mathbb{I} = \big\{\{k,i\}_i \times \{k,j\}_j \mid \{k,i\} \neq \{k,j\}\big\}$. Since $\mathbb{E}\{|\tilde{h}_{k,m}|^2\} = 1$ and $\mathbb{E}\{\tilde{h}_{k,i}^*\tilde{h}_{k,j}\} = 0$ are always true[1], then we have

$$\mathbb{E}\left\{|\boldsymbol{h}_k^H\boldsymbol{u}_k|^2\right\} = \sum_{m=1}^{n_t} g_{k,m}u_{k,m}^2. \tag{A.4}$$

By using the result of (A.4) in (A.2), we have

$$\mathrm{SINR}_k = \left(p_k\sum_{m=1}^{n_t} g_{k,m}u_{k,m}^2\right)\mathbb{E}\left\{\frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\}. \tag{A.5}$$

[1] Please note that the expectation is over the fast-fading randomness.

Now, we consider the expectation $\mathbb{E}\left\{\frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\}$. By using the Taylor series expansion, we can write this expectation as [74]

$$\mathbb{E}\left\{\frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\} = \mathbb{E}\left\{\frac{X_k}{Y_k}\right\} = \frac{\mathbb{E}\{X_k\}}{\mathbb{E}\{Y_k\}} - \frac{\text{cov}\left(X_k, Y_k\right)}{\left(\mathbb{E}\{Y_k\}\right)^2} + \frac{\text{var}\left(Y_k\right)\mathbb{E}\{X_k\}}{\left(\mathbb{E}\{Y_k\}\right)^2\mathbb{E}\{Y_k\}}, \tag{A.6}$$

where $X_k = 1$ and $Y_k = \sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2 + \sigma_{z_k}^2$. Now, we calculate the values of $\mathbb{E}\{Y_k\}$ and var(Y_k). Following similar calculation used to obtain (A.3) and (A.4), we have $\mathbb{E}\{|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2\} = \sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2$. Therefore

$$\mathbb{E}\{Y_k\} = \sum_{\substack{q=1\\q\neq k}}^{K} p_q \sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2 + \sigma_{z_k}^2. \tag{A.7}$$

On the other hand, we have

$$\begin{aligned}\text{var}\{Y_k\} &= \mathbb{E}\left\{\left|Y_k - \mathbb{E}\{Y_k\}\right|^2\right\} \\ &= \sum_{\substack{q=1\\q\neq k}}^{K} p_q^2\,\mathbb{E}\left\{\left|\sum_{\mathbb{I}} \tilde{h}_{k,i}^{*}\tilde{h}_{k,j}\sqrt{g_{k,i}g_{k,j}}u_{q,i}u_{q,j}\right|^2\right\} \\ &= \sum_{\substack{q=1\\q\neq k}}^{K} p_q^2 \sum_{\mathbb{I}} g_{k,i}g_{k,j}u_{q,i}^2u_{q,j}^2.\end{aligned} \tag{A.8}$$

Based on (A.7), (A.8) and the orthogonality between $\boldsymbol{g}_k$ and $\boldsymbol{u}_q$ as $n_t \to \infty$ for $k \neq q$, the following inequality $(\mathbb{E}\{Y_k\})^2 \gg \text{var}(Y_k)$ is true. Further, $\text{cov}(X_k, Y_k) = 0$. Applying these results to the series expansion in (A.6) we get

$$\mathbb{E}\left\{\frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\} \approx \frac{\mathbb{E}\{X_k\}}{\mathbb{E}\{Y_k\}} = \frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q \sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2 + \sigma_{z_k}^2}. \tag{A.9}$$

Substituting (A.9) into (A.5), we obtain

$$\text{SINR}_k = \frac{p_k \sum_{m=1}^{n_t} g_{k,m}u_{k,m}^2}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q \sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2 + \sigma_{z_k}^2}. \tag{A.10}$$

This concludes the proof.

Appendix B

Secrecy Design for Noncoherent Cell-Free Massive MIMO Systems with Passive Eavesdropper

B.1 Introduction

Massive MIMO technology, where a base station (BS) equipped with a large number of antennas that serves many information users (IUs) simultaneously, helps in improving wireless communications link and can be used to increase the capacity of the system and improve the link reliability at low power operation [72, 73]. As a consequence of employing large number of antennas in massive MIMO, the downlink channel vectors of independent users have a large degree of orthogonality which simplifies the system design and performance analysis [75]. Massive antenna arrays at the BS can be formed either in colocated or distributed manner. In collocated massive MIMO, all users in each cell are served by an array of colocated antennas mounted at the BS. In contrast, in cell-free massive MIMO, users over a large area are served by a large number of distributed antennas, namely, access points (APs).

Recently, cell-free massive MIMO system has attracted interest from many researchers due to its ability of coping with the problems of antenna correlations and deep shadowing fading [73]. In addition, cell-free massive MIMO offers a scalable implementation with much higher probability of coverage than the conventional colocated massive MIMO, however, with the cost of increased backhaul requirements [72,76–78]. The most important feature of cell-free massive MIMO is that many single-antenna APs serve a much smaller number of users simultaneously along with the use of the asymptotic orthogonality between independent users' channels, i.e.,

enabling channel hardening phenomena [79]. In contrast, in coherent cell-free massive MIMO system, the estimation of the CSI is often obtained by training and tracking between the single-antenna APs and the end users, then the so-obtained CSI and related signalling are shared between the central processing unit (CPU) and the APs. However, it is not always practical to use training-based schemes, especially with systems that have a large number of distributed APs. As the number of distributed single-antenna APs grows large such as in the case of cell-free massive MIMO, the channel estimation process, payload data, backhaul requirements, and the information exchange between the APs and the CPU will grow proportionately [80]. Hence, there is much interest to adopt a low backhaul requirements noncoherent scheme with cell-free massive MIMO that do not require either the transmitter or receiver to know the CSI.

In most practical scenarios, due to the open nature of the wireless communication medium, legitimate transmission information can be easily intercepted and altered by malicious eavesdropper, i.e., eavesdropping attacks. In contrast, physical layer designs of cell-free massive MIMO need to cope with many security challenges such as simultaneous wireless information and power transfer (SWIPT) systems that contain untrusted information from passive energy harvester (EH). The secrecy issues in SWIPT massive MIMO systems, in particular for noncoherent cell-free system, have not been previously studied in the literature. Recently, a large number of existing studies in the broader literature have examined the secrecy problems in SWIPT systems for the *coherent colocated* massive MIMO architecture, where the BS and information users have full knowledge of the channel [74, 81–85]. However, the proposed schemes suffer from backhaul requirements such as CSIs signalling, consequent *noncoherent cell-free* massive MIMO system is needed to address these issues.

Contributions: This chapter proposes a secure system design and the performance evaluation of noncoherent cell-free massive MIMO with SWIPT when the CSI is not known at the APs. In particular, the transmitted information intended for the legitimate IU is to be kept secure and confidential from the eavesdropper, where the wireless communication links between the APs and the IUs are vulnerable to be overheard by an information untrusted passive EH. Hence, the system consists of legitimate multiple antennas IUs and passive multiple antennas eavesdropper which pretends to be an EH receiver in order to overhear the information signals intended to one or more of the legitimate IU. Since the availability of the CSI is not relevant for the noncoherent downlink transmission system, then both the APs and EH have only the statistical value of the CSI available. Further, in downlink noncoherent transmission stage, the artificial noise (AN) can be used to degrade the quality of the information signals directed to EH.

The contributions are summarised as follows

- An achievable noncoherent rate at high average SINR is provided by following the approach of Grassmannian manifold model in [16].
- Deterministic closed-form expressions is derived for ergodic secrecy rate (SR) and average harvested energy rate for cell-free massive MIMO which can be optimised jointly to overcome eavesdropping attacks and improve the system performance.
- An optimised downlink transmission is proposed based on the statistical CSI. The optimisation is done for the information, noise, and energy signals beamformed towards the legitimate IUs and illegitimate EH, respectively. To this end, the downlink power allocation is optimised for information and AN signals at all APs and for all users to maximise the worst-case secrecy rate.
- The optimisation problem is nonlinear and not convex over the optimisation constraints. Thus, by using large vector asymptotic analysis, Taylor series expansion, and the monotonicity property of the exponential function a convex approximation for the ergodic secrecy rate is obtained.
- Finally, the suboptimal optimisation solution is provided with lower complexity.

B.2 System Model

As illustrated in Fig. B.1, the downlink of a cell-free massive MIMO system comprising a very large number of randomly located APs, $\{\text{AP}_1,\ \ldots,\ \text{AP}_N\}$; M information users (IUs) each equipped with K antennas, $\{\text{IU}_1,\ \ldots,\ \text{IU}_M\}$, interested in information decoding; and a K_e-antenna passive information-untrusted energy harvester (EH) interested in energy harvesting, however, it could act as an information eavesdropper by overhearing the signal intended for a certain IU, certain IU, is considered.. The APs are randomly located on a circular area A based on a homogeneous Poisson point process Φ with a density λ, while the users (IUs and the EV) are independently and randomly located on A. All APs are linked to a CPU via an infinite capacity backhaul. The CPU controls a differential downlink transmission via all APs.
Let $\{\text{AP}_1,\ \ldots, \text{AP}_N\}$ be the set of the adopted realisation of APs. $\boldsymbol{H}_i \in \mathbb{C}^{K\times N} = [\boldsymbol{h}_{i,1},\ \ldots, \boldsymbol{h}_{i,K}]^H$ $= [\bar{\boldsymbol{h}}_{i,1},\ \ldots, \bar{\boldsymbol{h}}_{i,K}]^H \boldsymbol{B}_i^{\frac{1}{2}}$ denotes the downlink channel matrix between IU_i and the set of APs, where $\boldsymbol{h}_{i,j} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{B}_i)$ is the channel vector between the jth antenna of IU_i and the APs. $\bar{\boldsymbol{h}}_{i,j} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_N)$ is the small-scale fading vector and $\boldsymbol{B}_i = \text{diag}(\boldsymbol{b}_i) = \text{diag}(b_{i,1},\ \ldots,\ b_{i,N})$

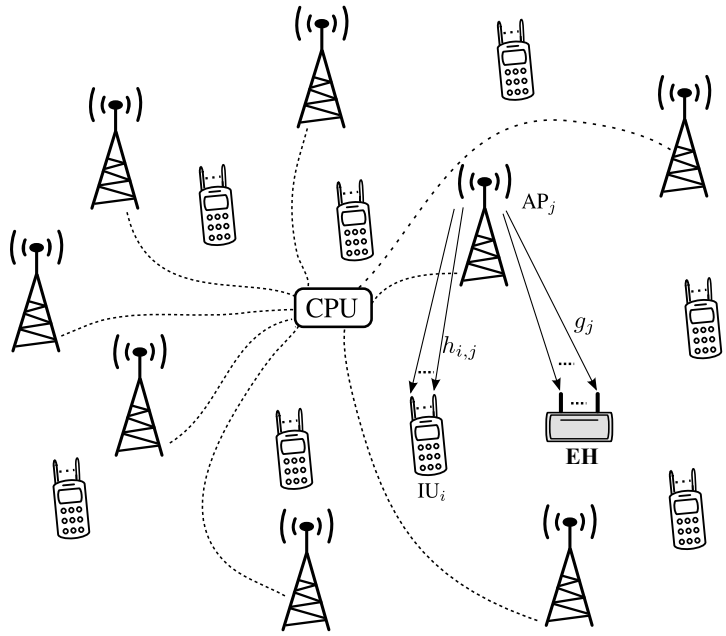

Figure B.1. An illustration of the proposed differential cell-free massive MIMO system.

is a diagonal matrix representing the large-scale fading which is called in this chapter as power space profile (PSP)[1]

Similarly, $\boldsymbol{G} \in \mathbb{C}^{K_e \times N} = [\boldsymbol{g}_1, \ \ldots, \boldsymbol{g}_{K_e}]^H = [\bar{\boldsymbol{g}}_1, \ \ldots, \bar{\boldsymbol{g}}_{K_e}]^H \boldsymbol{B_e}^{\frac{1}{2}}$ denotes the downlink channel matrix between EH and the set of APs, where $\boldsymbol{g}_j \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{B}_e)$ is the channel vector between the jth antenna of EH and the APs. $\bar{\boldsymbol{g}}_j \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_N)$ is the small-scale fading vector and $\boldsymbol{B}_e = \text{diag}(\boldsymbol{b}_e) = \text{diag}(b_{e,1}, \ \ldots, \ b_{e,N})$ is a diagonal matrix containing the large-scale fading coefficients.

The large-scale fading coefficients in the $\{\boldsymbol{B}_i\}$ and $\boldsymbol{B}_e$ change very slowly compared to the small-scale fading coefficients, therefore, it is assumed that $\{\boldsymbol{B}_i\}$ and $\boldsymbol{B}_e$ are perfectly known at the APs [86].

B.2.1 Signal Model

To improve the information security, the APs employ the transmission of the information signal towards the IUs, and the jamming signal (artificial noise (AN)) towards the EV. The base band model of the received signal, $\boldsymbol{Y}_i \in \mathbb{C}^{K \times T}$, at the IU$_i$ over T coherence time slots can be represented as

[1]It is assumed that the spacing between the IU's antennas is larger than the half of the signal wave length. Therefore, there is no correlation between the receive antennas.

$$\boldsymbol{Y}_i = \sum_{j=1}^{M} \boldsymbol{H}_i \boldsymbol{X}_j + \boldsymbol{H}_i \bar{\boldsymbol{X}} + \boldsymbol{W}_i, \tag{B.1}$$

where $\boldsymbol{X}_j \in \mathbb{C}^{N\times T} = [\boldsymbol{x}_{j,1},\ \ldots,\ \boldsymbol{x}_{j,T}]$, $\boldsymbol{x}_{j,t} = \boldsymbol{P}_j^{\frac{1}{2}} \bar{\boldsymbol{x}}_{j,t}$ is the signal intended for IU$_j$ during the tth time slot. $\bar{\boldsymbol{x}}_{j,t} \sim \mathcal{CN}(\boldsymbol{0},\ \boldsymbol{I}_N)$ is the vector bearing information, and $\boldsymbol{P}_j = \text{diag}(\boldsymbol{p}_j) = \text{diag}(p_{j,1},\ \ldots,\ p_{j,N})$ is a matrix containing the average transmit powers for the signal vector $\boldsymbol{x}_{j,t}$[1]. $\boldsymbol{W}_i \sim \mathcal{CN}(\boldsymbol{0},\ \sigma^2 \boldsymbol{I}_N)$ is the matrix of additive Gaussian noise at IU$_i$. $\bar{\boldsymbol{X}} \in \mathbb{C}^{N\times T} = [\bar{\boldsymbol{x}}_1,\ \ldots,\ \bar{\boldsymbol{x}}_T]$, $\bar{\boldsymbol{x}}_t = \bar{\boldsymbol{P}}^{\frac{1}{2}} \bar{\bar{\boldsymbol{x}}}_t$ is the AN signal intended to jam the EV during the tth time slot. $\bar{\bar{\boldsymbol{x}}}_t \sim \mathcal{CN}(\boldsymbol{0},\ \boldsymbol{I}_N)$ is a vector containing the AN symbols, and $\bar{\boldsymbol{P}} = \text{diag}(\bar{\boldsymbol{p}}) = \text{diag}(\bar{p}_1,\ \ldots,\ \bar{p}_N)$ is a matrix containing the average transmit powers for the AN signal vector $\bar{\boldsymbol{x}}_t$.

Similarly, the received signal at the EV is given as

$$\boldsymbol{Y}_e \in \mathbb{C}^{K_e\times T} = \sum_{j=1}^{M} \boldsymbol{G}\boldsymbol{X}_j + \boldsymbol{G}\bar{\boldsymbol{X}} + \boldsymbol{W}_e, \tag{B.2}$$

where $\boldsymbol{W}_e \sim \mathcal{CN}(\boldsymbol{0},\ \sigma^2 \boldsymbol{I}_N)$ is the matrix of additive Gaussian noise at the EV.

The average signal-to-interference-plus-noise ratio SINR (per time slot) at each receive antenna of IU$_i$, SINR$_i$, is[2]

$$\text{SINR}_i = \mathbb{E}\left[\frac{\left|\boldsymbol{h}_{i,j}^H \boldsymbol{x}_{i,t}\right|^2}{\sum_{l\neq i}\left|\boldsymbol{h}_{i,j}^H \boldsymbol{x}_{l,t}\right|^2 + \left|\boldsymbol{h}_{i,j}^H \bar{\boldsymbol{x}}_t\right|^2 + \sigma^2}\right]. \tag{B.3}$$

Given that IU$_k$ is the IU being attacked by the EV, the average SINR (per time slot) at each receive antenna of the EV, SINR$_{e_k}$, is

$$\text{SINR}_{e_k} = \mathbb{E}\left[\frac{\left|\boldsymbol{g}_j^H \boldsymbol{x}_{k,t}\right|^2}{\sum_{l\neq k}\left|\boldsymbol{g}_j^H \boldsymbol{x}_{l,t}\right|^2 + \left|\boldsymbol{g}_j^H \bar{\boldsymbol{x}}_t\right|^2 + \sigma^2}\right]. \tag{B.4}$$

B.2.2 Secrecy Performance of Noncoherent System

In massive MIMO noncoherent downlink transmission in which the receiver has no prior knowledge of the channel. The length of the channel coherence time, T, determines the upper bound on the downlink capacity.

It has been found that the capacity of the differential MIMO system can be geometrically

[1]Since the large-scale fading stays constant over all time slots, $t = 1,\ \ldots,\ T$, then the optimised average transmit power $\boldsymbol{P}_j$ for a certain time slot t is also optimal for any other time slot $m \neq n$.

[2]The average SINR per antenna per time slot is calculated instead of the total SINR, since such an SINR is required to calculate the downlink rate of the differential MIMO system as will be discussed in Subsection B.2.3.

interpreted as sphere packing in the Grassmann manifold [16]. The Grassmann manifold is defined as follows: given a Stiefel Manifold $\mathbb{S}(M_1, M_2)$ which represents the space of all M_1-by-M_2 complex unitary matrices, then, the Grassmann manifold $\mathbb{G}(M_1, M_2)$ is the space of the ratio of the Stiefel Manifold $\mathbb{S}(M_1, M_2)$ to the Stiefel Manifold $\mathbb{S}(M_2, M_2)$[1] [16].

For $N > K$ and the coherence time $T \geq 2K$, the achievable rate at high average SINR regime by a single IU$_i$, C_i; and by the EV when attacking the kth IU, IU$_k$, is C_{E_k}, are [16]

$$C_k = K\left(1 - \frac{K}{T}\right)\log_2 \text{SINR}_k + c_k, \tag{B.5}$$

$$C_{E_k} = K_e\left(1 - \frac{K_e}{T}\right)\log_2 \text{SINR}_{e_k} + c_e \tag{B.6}$$

where

$$\begin{aligned} c_k =& \frac{1}{T}\log_2 \left|\mathbb{G}\left(T, K\right)\right| + K\left(1 - \frac{K}{T}\right)\log_2 \frac{T}{K\pi e} \\ &+ \left(1 - \frac{K}{T}\right)\mathbb{E}\left[\log_2 \det \boldsymbol{H}_k \boldsymbol{H}_k^H\right], \end{aligned} \tag{B.7}$$

$$\begin{aligned} c_e =& \frac{1}{T}\log_2 \left|\mathbb{G}\left(T, K_e\right)\right| + K_e\left(1 - \frac{K_e}{T}\right)\log_2 \frac{T}{K_e\pi e} \\ &+ \left(1 - \frac{K_e}{T}\right)\mathbb{E}\left[\log_2 \det \boldsymbol{G}\boldsymbol{G}^H\right], \end{aligned} \tag{B.8}$$

and $|\mathbb{G}(T, N)|$ is a natural measure on the Grassmann manifold defined as [16]

$$|\mathbb{G}(T, K)| = \frac{\prod_{i=T-K+1}^{T} \frac{2\pi^i}{(i-1)!}}{\prod_{i=1}^{K} \frac{2\pi^i}{(i-1)!}} \tag{B.9}$$

Theorem 2. *For $N \to \infty$, the values of $\mathbb{E}[\log_2 det\, \boldsymbol{H}_k \boldsymbol{H}_k^H]$ in (B.7) and $\mathbb{E}[\log_2 det\, \boldsymbol{G}\boldsymbol{G}^H]$ in (B.8) converge to deterministic values as follows [87]*

$$\mathbb{E}\left[\log_2 det\, \boldsymbol{H}_k \boldsymbol{H}_k^H\right] \overset{N\to\infty}{\to} \log_2 trace\left(\boldsymbol{B}_k\right) \tag{B.10}$$

$$\mathbb{E}\left[\log_2 det\, \boldsymbol{G}\boldsymbol{G}^H\right] \overset{N\to\infty}{\to} \log_2 trace\left(\boldsymbol{B}_e\right). \tag{B.11}$$

Proof. The proof is provided in Appendix B.6.1. □

The high SINR regime assumption in (B.5) and (B.6) indicate that both $\{\text{SINR}_k\}$ and $\{\text{SINR}_{e_k}\}$ are large, but not necessarily equal. In the following, a theorem is provided in which the values of $\{\text{SINR}_k\}$ and $\{\text{SINR}_{e_k}\}$ are calculated for $N \to \infty$.

[1] For more details regarding the Stiefel and Grassmann manifolds, the reader is recommended to see [16].

Theorem 3. *As $N \to \infty$, SINR$_k$ and SINR$_{e_k}$ (given in (B.3) and (B.4)) are accurately approximated as*

$$\begin{aligned} \textit{SINR}_k &\overset{N\to\infty}{\to} \frac{\mathbb{E}\left[\left|\boldsymbol{h}_{k,j}^H \boldsymbol{x}_{k,t}\right|^2\right]}{\mathbb{E}\left[\sum_{k\neq l}\left|\boldsymbol{h}_{k,j}^H \boldsymbol{x}_{l,t}\right|^2 + \left|\boldsymbol{h}_{k,j}^H \bar{\boldsymbol{x}}_t\right|^2 + \sigma^2\right]} \\ &= \frac{\boldsymbol{b}_k^T \boldsymbol{p}_k}{\sum_{k\neq l} \boldsymbol{b}_k^T \boldsymbol{p}_l + \boldsymbol{b}_k^T \bar{\boldsymbol{p}} + \sigma^2} \end{aligned} \tag{B.12}$$

$$\begin{aligned} \textit{SINR}_{e_k} &\overset{N\to\infty}{\to} \frac{\mathbb{E}\left[\left|\boldsymbol{g}_j^H \boldsymbol{x}_{k,t}\right|^2\right]}{\mathbb{E}\left[\sum_{k\neq l}\left|\boldsymbol{g}_j^H \boldsymbol{x}_{l,t}\right|^2 + \left|\boldsymbol{g}_j^H \bar{\boldsymbol{x}}_t\right|^2 + \sigma^2\right]} \\ &= \frac{\boldsymbol{b}_e^T \boldsymbol{p}_k}{\sum_{k\neq l} \boldsymbol{b}_e^T \boldsymbol{p}_l + \boldsymbol{b}_e^T \bar{\boldsymbol{p}} + \sigma^2} \end{aligned} \tag{B.13}$$

Proof. See Appendix B.6.2. □

Assuming that IU$_k$ is the user being attacked, then the secrecy rate of IU$_k$ is defined as

$$R_k \overset{N\to\infty}{\to} \left[C_k - C_{E_k}\right]^+. \tag{B.14}$$

B.2.3 Average Harvested Energy

Assuming that the EH devotes the whole received signal for energy harvesting during a unit time slot duration, and the energy harvesting efficiency of the EH is $0 \leq \zeta \leq 1$. Given IU$_k$, $k \in \{1, \ldots, M\}$ is the user being attacked, the total energy harvested by the EH is expressed as

$$\begin{aligned} E_k &= \zeta \mathbb{E}\left[\left|\boldsymbol{g}_j^H \boldsymbol{x}_{k,t}\right|^2 + \sum_{l\neq k}\left|\boldsymbol{g}_j^H \boldsymbol{x}_{l,t}\right|^2 + \left|\boldsymbol{g}_j^H \bar{\boldsymbol{x}}_t\right|^2\right] \\ &= \zeta\left(\sum_{k=1}^{M} \boldsymbol{b}_e^T \boldsymbol{p}_k + \zeta \boldsymbol{b}_e^T \bar{\boldsymbol{p}}\right). \end{aligned} \tag{B.15}$$

For an optimised downlink transmission, the power allocation will be optimised based on the attacked IU. Therefore, the value of the harvested energy is subjected to the attacked IU.

B.3 Secrecy Problem Formulation

In this section, the focus is on optimising the downlink power allocation for information and AN signals at all APs and for all users, $\{p_{j,1}, \ldots, p_{j,N}\}$ and $\{p_1, \ldots, p_N\}$, to maximise the

worst-case secrecy rate. Let us define

$$\Delta = K\left(1-\frac{K}{T}\right) - K_e\left(1-\frac{K_e}{T}\right). \tag{B.16}$$

Given that IU_k is the IU under attack, then the worst-case secrecy rate, $\min_k R_k$, is

$$\begin{aligned}\min_k R_k \overset{N\to\infty}{\to} \min\Bigg\{&\Delta_1 \log_2\left(\frac{\text{SINR}_k}{\text{SINR}_{e_k}}\right)\\ &+ |\Delta| \log_2\left(\text{SINR}_d\right) + c_k - c_e\Bigg\},\end{aligned} \tag{B.17}$$

where $\text{SINR}_d = \text{SINR}_k$ and $\Delta_1 = K_e\left(1-\frac{K_e}{T}\right)$ if $\text{sign}(\Delta) = +$; $\text{SINR}_d = \text{SINR}_{e_k}$ and $\Delta_1 = K\left(1-\frac{K}{T}\right)$ if $\text{sign}(\Delta) = -$. With the result in (B.17), a worst-case secrecy rate maximisation problem is formulated under constraints on total transmit power at each AP as follows

$$\underset{\{\boldsymbol{p}_j\},\bar{\boldsymbol{p}}}{\text{maximise}} \quad \min_k R_k \tag{B.18a}$$

subject to

$$\sum_{j=1}^{M} p_{j,i} + \bar{p}_i \le P_t, \qquad \forall i, \tag{B.18b}$$

$$E_k \ge \bar{E}, \forall k \tag{B.18c}$$

where P_t is the available power budget at each AP, $p_{j,i}$ is the information signal power between the jth IU and the ith AP, and $\bar{p}_i$ is the AN signal power between the ith AP and the EV, and $\bar{E}$ is a lower bound constraint on the average harvested energy by the EH. The constraint in (B.18b) ensures the average power used at each AP is within the limit of P_t. It is clear that the objective function (B.18a) is non-convex since R_k comprises a logarithm of a multiplicative fractional function. In the following two subsections, an optimal and suboptimal solutions are provided for the constrained secrecy problem (B.18).

B.3.1 Optimal Solution

The per user per AP power allocation can be optimally balanced by reformulating the non-convex problem (B.18) into a convex problem by using the following exponential variable substitution method. Let us consider the case in which $\text{sign}(\Delta) = -$, the expression in (B.17) can

be recast as

$$\min_k R_k \overset{N\to\infty}{\to} \min\left\{\frac{K}{\log_e(2)}\left(1-\frac{K}{T}\right)\log_e\left(\frac{e^{u_k}e^{v_k}}{e^{s_k}e^{t_k}}\right) + \frac{\Delta}{\log_e(2)}\log_e\left(\frac{e^{t_k}}{e^{v_k}}\right)+c_k-c_e\right\}, \tag{B.19}$$

where

$$e^{u_k} = \boldsymbol{b}_k^T\boldsymbol{p}_k, \tag{B.20}$$

$$e^{s_k} = \sum_{i\neq l}\boldsymbol{b}_k^T\boldsymbol{p}_l + \boldsymbol{b}_k^T\bar{\boldsymbol{p}} + \sigma^2, \tag{B.21}$$

$$e^{v_k} = \sum_{l\neq k}\boldsymbol{b}_e^T\boldsymbol{p}_l + \boldsymbol{b}_e^T\bar{\boldsymbol{p}} + \sigma^2, \tag{B.22}$$

$$e^{t_k} = \boldsymbol{b}_e^T\boldsymbol{p}_k. \tag{B.23}$$

Furthermore, by using the properties of logarithmic and exponential functions, it follows

$$\min_k R_k \overset{N\to\infty}{\to} \min\left\{\frac{K}{\log_e(2)}\left(1-\frac{K}{T}\right)\log_e\left(e^{u_k+v_k-s_k-t_k}\right) + \frac{\Delta}{\log_e(2)}\log_e\left(e^{t_k-v_k}\right)+c_k-c_e\right\}, \tag{B.24}$$

which implies

$$\min_k R_k \overset{N\to\infty}{\to} \min\left\{\frac{K}{\log_e(2)}\left(1-\frac{K}{T}\right)\left(u_k+v_k-s_k-t_k\right) + \frac{\Delta}{\log_e(2)}\left(t_k-v_k\right)+c_k-c_e\right\}, \tag{B.25}$$

where $\log_e(e^x) = x$. We assume that our problem is feasible. Next, the optimisation problem is solved optimally through recasting the non-convex constraints. Using (B.25), the corresponding optimisation problem in (B.18) can be written as

$$\underset{\substack{\{\boldsymbol{p}_j\},\bar{\boldsymbol{p}}\\ \{u_k,v_e,s_k,t_e\}}}{\text{maximise}} \quad \min_k \left\{\frac{K}{\log_e(2)}\left(1-\frac{K}{T}\right)\left(u_k+v_k-s_k-t_k\right) + \frac{\Delta}{\log_e(2)}\left(t_k-v_k\right)+c_k-c_e\right\} \tag{B.26a}$$

subject to

$$e^{u_k} \leq \boldsymbol{b}_k^T\boldsymbol{p}_k, \qquad \forall k, \tag{B.26b}$$

$$e^{\bar{s}_k}(s_k-\bar{s_k}+1) \geq \sum_{k\neq l}\boldsymbol{b}_k^T\boldsymbol{p}_l + \boldsymbol{b}_k^T\bar{\boldsymbol{p}} + \sigma^2, \quad \forall k, \tag{B.26c}$$

$$e^{v_k} \leq \sum_{l \neq k} \boldsymbol{b}_e^T \boldsymbol{p}_l + \boldsymbol{b}_e^T \bar{\boldsymbol{p}} + \sigma^2, \quad \forall k \tag{B.26d}$$

$$e^{\bar{t_k}}(t_k - \bar{t_k} + 1) \geq \boldsymbol{b}_e^T \boldsymbol{p}_k, \quad \forall k. \tag{B.26e}$$

$$B.18b,\ B.18c \tag{B.26f}$$

It can be seen that the exponential parameters u_k, v_e, s_k, and t_e are constrained by the expressions on the right hand sides of (B.26b)-(B.26e), respectively. The objective function in B.26a is now convex, thus it is linearised by using the monotonicity property of the exponential function. The constraints in (B.26b) and (B.26d) are convex. Next, to deal with the non-convex constraints in (B.26c) and (B.26e), we linearise them by using the first order Taylor approximation where $\bar{s_k}$, $\bar{t_e}$ are the points where the linear approximation is made. Therefore, problem (B.26) is now convex and can be solved iteratively using the CVX optimisation software.

Let us consider the case in which $\text{sign}(\Delta) = +$, the corresponding optimisation problem in (B.18) can be written as

$$\begin{aligned} \underset{\substack{\{\boldsymbol{p}_j\},\bar{\boldsymbol{p}} \\ \{u_k,v_e,s_k,t_e\}}}{\text{maximise}} \quad & \min_k \left\{ \frac{K_e}{\log_e(2)} \left(1 - \frac{K_e}{T}\right) \left(u_k + v_k - s_k - t_k\right) \right. \\ & \left. + \frac{\Delta}{\log_e(2)} \left(u_k - s_k\right) + c_k - c_e \right\} \end{aligned} \tag{B.27a}$$

subject to

$$e^{u_k} \leq \boldsymbol{b}_k^T \boldsymbol{p}_k, \qquad \forall k, \tag{B.27b}$$

$$e^{\bar{s}_k}(s_k - \bar{s_k} + 1) \geq \sum_{k \neq l} \boldsymbol{b}_k^T \boldsymbol{p}_l + \boldsymbol{b}_k^T \bar{\boldsymbol{p}} + \sigma^2, \quad \forall k, \tag{B.27c}$$

$$e^{v_k} \leq \sum_{l \neq k} \boldsymbol{b}_e^T \boldsymbol{p}_l + \boldsymbol{b}_e^T \bar{\boldsymbol{p}} + \sigma^2, \quad \forall k \tag{B.27d}$$

$$e^{\bar{t_k}}(t_k - \bar{t_k} + 1) \geq \boldsymbol{b}_e^T \boldsymbol{p}_k, \quad \forall k. \tag{B.27e}$$

$$B.18b,\ B.18c \tag{B.27f}$$

Remark 3. *In the transmit power constraints in (B.18b), it can be understood that the optimised values of the vectors of power control factors $\{\boldsymbol{p_j}\}$ and $\bar{\boldsymbol{p}}$ are done on each elements separately. For example, Fig. B.2 shows the transmit powers per AP for the case $\bar{E} = 0$ obtained by solving problem (B.26). It can be seen that the majority of the subgroup of the APs utilise large portion of their available transmit power budgets.*

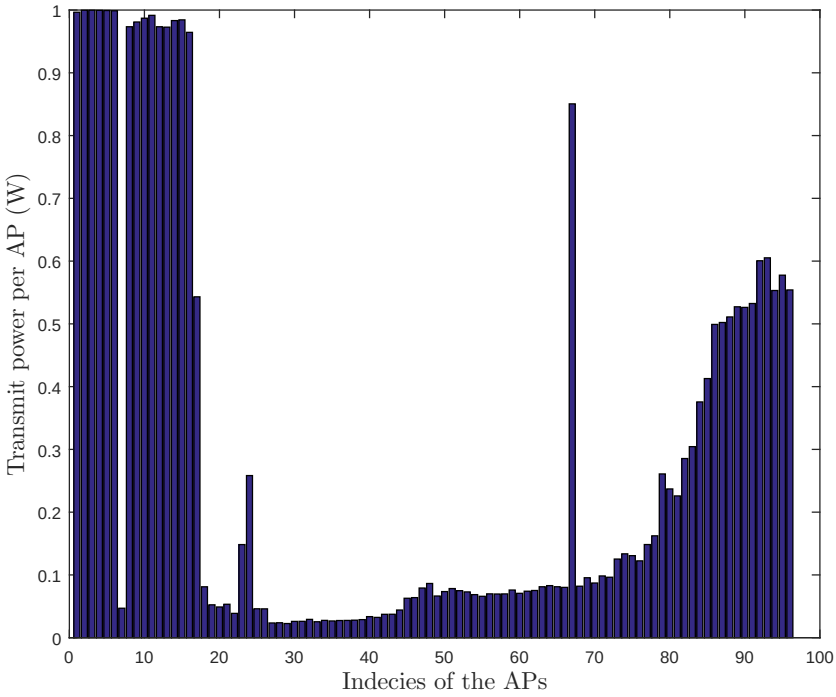

Figure B.2. An illustration of the transmit power utilization per AP.

B.3.2 Suboptimal Solution

In the optimal solution discussed in the previous subsection, the power control optimisation is per user per AP. Therefore, the solver needs to optimise $(M+1)N$ variables. This leads to large complexity when $N \to \infty$. Therefore, in this subsection, a suboptimal solution is introduced with lower complexity. The suboptimal is based on matching the vectors that contain the average transmit power to the PSP which can be explained as follows.

Let the power control vector of the IU$_i$'s information signal be

$$\boldsymbol{p}_i = p_i \tilde{\boldsymbol{p}}_i \tag{B.28}$$

where

$$\tilde{\boldsymbol{p}}_i = \frac{\boldsymbol{b}_i}{\|\boldsymbol{b}_i\|}. \tag{B.29}$$

p_i is the power control factor the IU$_i$. The power control for the AN is designed in the same manner as

$$\bar{\boldsymbol{p}} = \bar{p}\breve{\boldsymbol{p}} \tag{B.30}$$

where

$$\breve{\boldsymbol{p}} = \frac{\boldsymbol{b_e}}{\|\boldsymbol{b_e}\|}. \tag{B.31}$$

$\bar{p}$ is the power control factor the AN. By this design, we matched the relative information power for a given IU at different APs to the relative entries of the PSP of that IU. Then, it is of interest to control the total information powers $\{p_1,\ p_2,\ \ldots, p_M\}$ and the total AN power $\bar{p}$. The power control in the suboptimal solution is only per user, thus the solver needs only to optimise $(M+1)$ variables. Since the computational complexity is directly proportional to the number of optimisation variables, then the suboptimal solution scheme has lower computational complexity than the optimal solution. Based on this power control design and a similar problem reformulation procedure to that used for (B.26) and (B.27), the problem is formulated as follows

$$\begin{aligned}\underset{\substack{\{p_j\},\bar{p}\\ \{u_k,v_e,s_k,t_e\}}}{\text{maximise}}\ \min_k \Bigg\{ & \frac{K}{\log_e(2)}\left(1-\frac{K}{T}\right)\Big(u_k+v_k-s_k-t_k\Big) \\ & + \frac{\Delta}{\log_e(2)}\Big(t_k-v_k\Big)+c_k-c_e\Bigg\}\end{aligned} \tag{B.32a}$$

subject to

$$e^{u_k} \le p_k \boldsymbol{b}_k^T \tilde{\boldsymbol{p}}_k, \forall k, \tag{B.32b}$$

$$e^{\bar{s}_k}(s_k - \bar{s_k} + 1) \ge \sum_{k\neq l} p_l \boldsymbol{b}_k^T \tilde{\boldsymbol{p}}_l + \bar{p}\boldsymbol{b}_k^T \breve{\boldsymbol{p}} + \sigma^2, \forall k, \tag{B.32c}$$

$$e^{v_k} \ \le \ \sum_{l\neq k} p_l \boldsymbol{b}_e^T \tilde{\boldsymbol{p}}_l + \bar{p}\boldsymbol{b}_e^T \breve{\boldsymbol{p}} + \sigma^2, \quad \forall k \tag{B.32d}$$

$$e^{\bar{t_k}}(t_k - \bar{t_k} + 1) \ge p_k \boldsymbol{b}_e^T \tilde{\boldsymbol{p}}_k, \quad \forall k. \tag{B.32e}$$

$$\sum_{i=1}^{M} p_i \left[\tilde{\boldsymbol{p}}_i\right]_j + \bar{p}\left[\breve{\boldsymbol{p}}\right]_j \le P_t, \qquad \forall j \tag{B.32f}$$

$$E_k \ge \bar{E}, \qquad \forall k \tag{B.32g}$$

where E_k is given as

$$E_k = \zeta\left(\sum_{k=1}^{M} p_k \boldsymbol{b}_e^T \tilde{\boldsymbol{p}}_k + \bar{p}\boldsymbol{b}_e^T \breve{\boldsymbol{p}}\right). \tag{B.33}$$

Problem (B.32) is now convex and can be solved iteratively by CVX.

Remark 4. *In the transmit power constraints in (B.32f), it can be understood that the optimised values of the scalars power control factors p_i and $\bar{p}$ are determined by the largest entries of $\tilde{\boldsymbol{p}}_i$ and $\breve{\boldsymbol{p}}$, i.e., $m = \arg\max\left[\tilde{\boldsymbol{p}}_i\right]_j$ and $n = \arg\max\left[\breve{\boldsymbol{p}}\right]_j$. Therefore, the rest of access points*

$\{AP_j\}_{j\notin\{m,n\}}$ have a minor contribution to the optimised solution. This is one of the weaknesses of the suboptimal formulation in (B.32) which can not be avoided since for cell-free MIMO system the power constraint is per AP. This case does not exist in the collocated MIMO systems where the power constraint can be per total transmit power at all antennas.

For example, Fig. B.3 shows the transmit powers per AP for the case $\bar{E} = 0$ obtained by solving problem (B.32). It can be seen that a minor subgroup of the APs utilise large portion of their transmit power budgets while the rest of the APs do not use their available transmit power budgets. This is because of the aforementioned weakness. In addition, this impacts the amount of the harvested energy at the EH since the APs do not transmit with their full power.

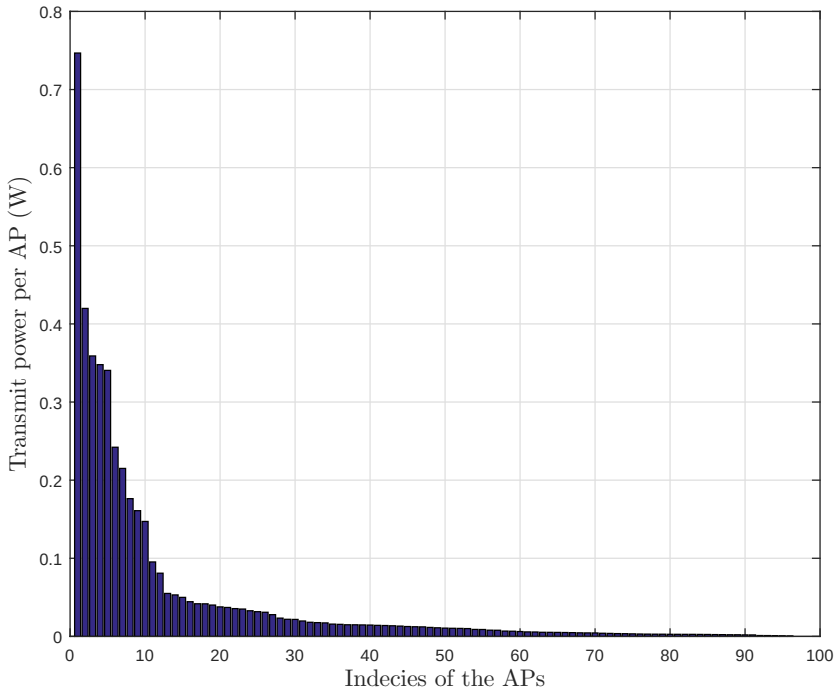

Figure B.3. An illustration of the transmit power utilisation per AP.

B.4 Simulation Results and Discussion

In this section, the asymptotic performance of cell-free noncoherent massive MIMO system with passive eavesdropper is evaluated. The APs are randomly deployed on area A_a using a homogeneous PPP Φ_a with an intensity λ_a. On the other hand, the IUs and the EH are randomly deployed in area A_u where the origins is the same for both A_a and A_u, yet $A_a > A_u$. The PSP large-scale fading coefficient $\{b_{i,j}, b_{e,j}\}$ are modeled according to the standard-based model.

Table B.1. The values of PSP parameters to be used in (5.14).

Parameter	Value
A_a, λ_u	$1 \times 1Km^2$, $1.6 \times 10^{-4}m^{-1}$
A_a, M, N	$300 \times 300m^2$, 2, 100
α, σ	2.6, 5dB
$p_{j,i}$, $\bar{p}_i$	1 W, 1W
ζ	0.5

$$b_{i,j} = d_{i,j}^{-\alpha} 10^{\frac{\nu_{i,j}}{10}}, \quad \text{and} \qquad b_{e,j} = d_{e,j}^{-\alpha} 10^{\frac{\nu_{e,j}}{10}}. \tag{B.34}$$

where $d_{i,j}$ denotes the distance from the IU$_i$ to the AP$_j$, $d_{i,j}$ denotes the distance form EH to the AP$_j$, α is the pathloss exponent, and $\{\nu_{i,j}, \nu_{e,j}\} \sim \mathcal{CN}(0, \sigma^2)$ are the shadow fading coefficients with standard division σ^2. The system parameters are chosen such that in Table B.1.

Fig. B.4 shows geometry deployment of AP, IU, and EH, where the number of APs is $N = 100$, number of IU is $M = 2$, and number of energy harvester EH=1.

Fig. B.5 shows the secrecy performance for the achievable worst-case SR. In this scenario, The APs is 100, the IU is 2 with K=5 antennas, the EH is 1 with K$_e$=4 antennas. It can be seen that as the HE increases, more downlink transmission resources are optimised, thus SR decreases.

In Fig. B.6, the experiment is repeated to test the secrecy performance for the achievable worst-case SR, but with The APs is 100, the IU is 2 with K=4 antennas, the EH is 1 with K$_e$=4 antennas. Now as the the HE increases, less downlink transmission resources are optimised, thus SR is less decreasing and thanks to the equality in the number of receive antennas for IU and EH.

More cases will be study and investigated.

B.5 Summary

This chapter proposed a secure system design and the performance evaluation of noncoherent cell-free massive MIMO with SWIPT when the CSI is not known at the APs. In particular,

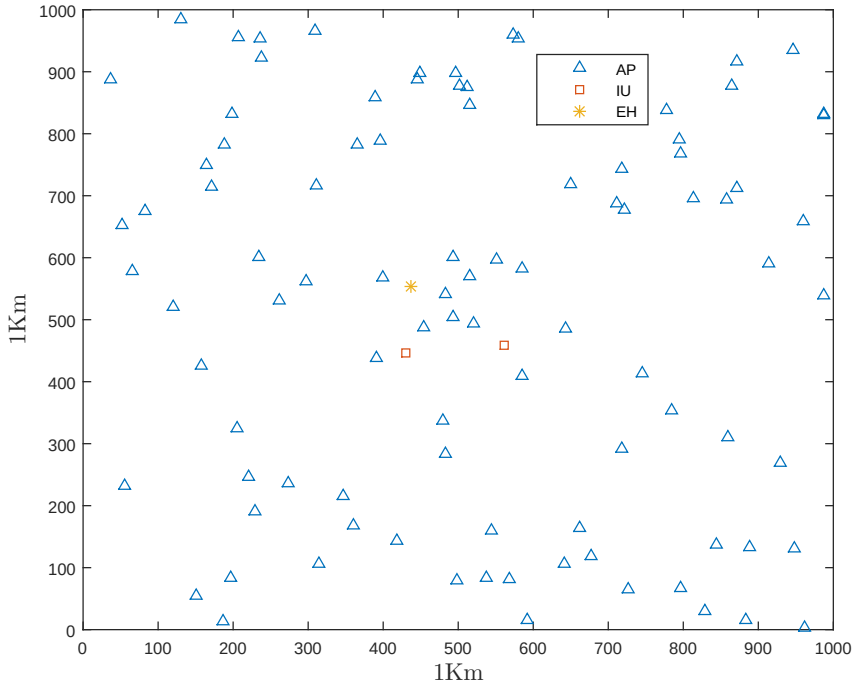

Figure B.4. An illustration of the transmit power utilization per AP in Cell-Free. $N = 100$, $M = 2$, $EH = 1$

the transmitted information intended to the legitimate IU is to be kept secure and confidential from the eavesdropper, where the wireless communication links between the APs and the IUs are vulnerable to be overheard by an information untrusted passive EH. To achieve this, an achievable noncoherent rate at high average SINR is provided then a statistical closed form expression is derived for the ergodic SR and average HR in cell-free massive MIMO which can be optimised jointly to overcome eavesdropping attacks and improve the overall system performance. In contrast, based on the statistical CSI, an optimal solution is proposed by optimising the downlink power allocation dedicated for information and AN signals at all APs and for all users to maximise the worst-case secrecy rate. Further, for low complexity systems, the suboptimal optimisation solution is provided.

B.6 Appendix

B.6.1 Proof of Theorem 2

To calculate the expected value in B.10, it is first expanded as follows

$$\boldsymbol{H}_i \boldsymbol{H}_i^H = \left[\boldsymbol{h}_{i,1}^H \boldsymbol{h}_{i,1}, \cdots, \boldsymbol{h}_{i,1}^H \boldsymbol{h}_{i,K} \right], \tag{B.35}$$

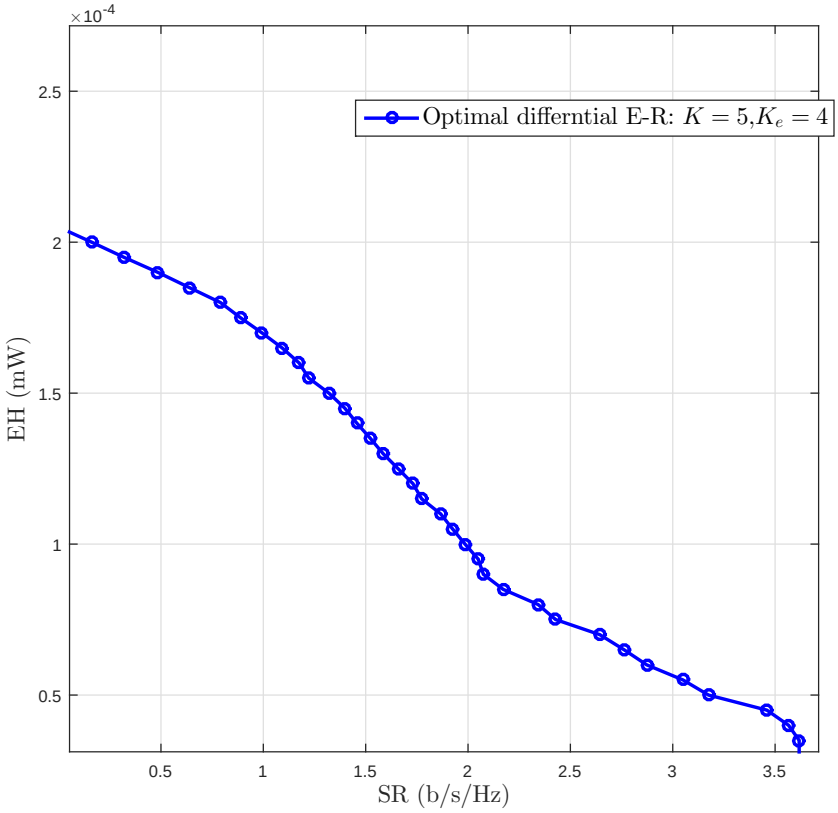

Figure B.5. An illustration of E-R regions of cell-free differential MIMO. $N = 100$, $M = 2$, $EH = 1$, $K = 5$, $K_e = 4$

where

$$\boldsymbol{h}_{i,j}^H \boldsymbol{h}_{i,k} = \bar{\boldsymbol{h}}_{i,j}^H \boldsymbol{B}_i^{\frac{1}{2}} \boldsymbol{B}_i^{\frac{1}{2}} \bar{\boldsymbol{h}}_{i,k} \overset{N\to\infty}{\to} \begin{cases} \text{trace}\left(\boldsymbol{B}_i\right), & j = k \\ 0, & j \neq k \end{cases}. \tag{B.36}$$

Therefore from (B.36), it follows

$$\boldsymbol{H}_i \boldsymbol{H}_i^H = \text{trace}\left(\boldsymbol{B}_i\right) \boldsymbol{I}_N. \tag{B.37}$$

From (B.36) and (B.37), the expected value in (B.10) can be calculated as

$$\mathbb{E}\left[\log_2 \det \boldsymbol{H}_i \boldsymbol{H}_i^H\right] \overset{N\to\infty}{\to} \mathbb{E}\left[\log_2 \det \left(\text{trace}\left(\boldsymbol{B}_i\right) \boldsymbol{I}_N\right)\right] \tag{B.38}$$

$$\overset{N\to\infty}{\to} \mathbb{E}\left[\log_2 \left(\text{trace}\left(\boldsymbol{B}_i\right) \det(\boldsymbol{I}_N)\right)\right] \tag{B.39}$$

$$\overset{N\to\infty}{\to} \log_2 \text{trace}\left(\boldsymbol{B}_i\right). \tag{B.40}$$

The proof for the value of $\mathbb{E}\left[\log_2 \det \boldsymbol{G}\boldsymbol{G}^H\right]$ in (B.11) is omitted since it can be obtained by following the same previous steps used to obtain (B.10). This concludes the proof.

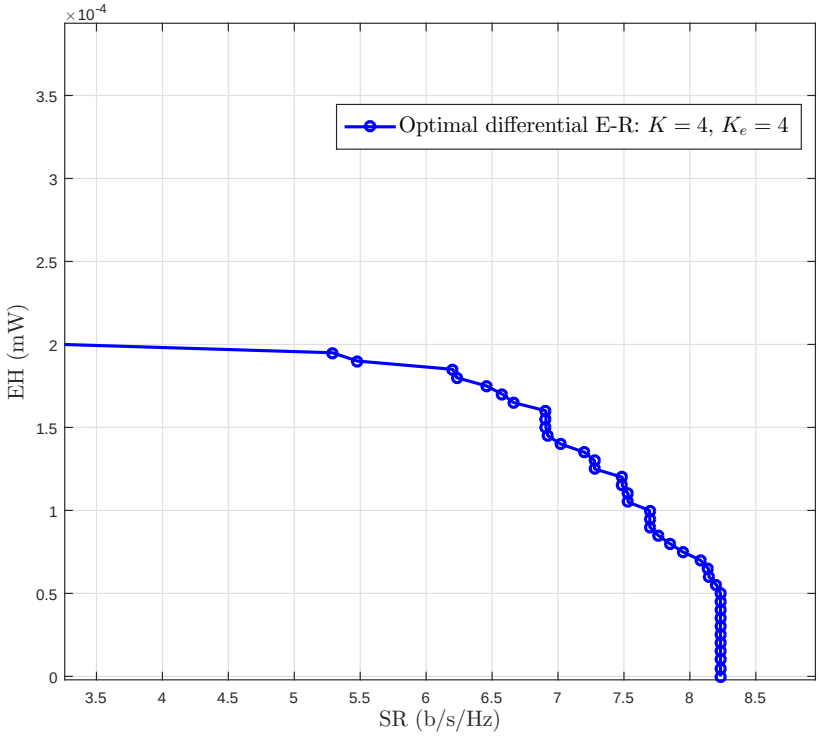

Figure B.6. An illustration of E-R regions of cell-free differential MIMO. $N = 100$, $M = 2$, $EH = 1$, $K = 4$, $K_e = 4$.

B.6.2 Proof of Theorem 3

The average SINR value in (B.5) is an expectation of a fractional value. By using the Taylor series expansion, the SINR_i can be expanded as

$$\begin{aligned}\text{SINR}_i &= \mathbb{E}\left[\frac{X_i}{Y_i}\right] \\ &= \frac{\mathbb{E}\left[X_i\right]}{\mathbb{E}\left[Y_i\right]} - \frac{\text{cov}\left(X_i, Y_i\right)}{\left(\mathbb{E}\left[Y_i\right]\right)^2} + \frac{\text{var}\left(Y_i\right)\mathbb{E}\left[X_i\right]}{\left(\mathbb{E}\left[Y_i\right]\right)^2 \mathbb{E}\left[Y_i\right]} + R\end{aligned} \tag{B.41}$$

where X_i and Y_i are the numerator and denominator of the fraction in B.5, respectively. R is the remainder of the series expansion. Now, let us compute $\mathbb{E}[X_i]$ as[1]

$$\begin{aligned}&\mathbb{E}\left[X_i\right] = \\ &\mathbb{E}\left[\left|\boldsymbol{h}_{i,j}^H \boldsymbol{x}_{i,t}\right|^2\right] = \mathbb{E}\left[\bar{\boldsymbol{h}}_{i,j}^H \boldsymbol{B}_i^{\frac{1}{2}} \boldsymbol{P}_i^{\frac{1}{2}} \bar{\boldsymbol{x}}_{i,t} \bar{\boldsymbol{x}}_{i,t}^H \boldsymbol{P}_i^{\frac{1}{2}} \boldsymbol{B}_i^{\frac{1}{2}} \bar{\boldsymbol{h}}_{i,j}\right] \\ &= \mathbb{E}\left[\bar{\boldsymbol{h}}_{i,j}^H \boldsymbol{B}_i^{\frac{1}{2}} \boldsymbol{P}_i^{\frac{1}{2}} \mathbb{E}\left[\bar{\boldsymbol{x}}_{i,t} \bar{\boldsymbol{x}}_{i,t}^H\right] \boldsymbol{P}_i^{\frac{1}{2}} \boldsymbol{B}_i^{\frac{1}{2}} \bar{\boldsymbol{h}}_{i,j}\right] \\ &= \text{trace}\left(\boldsymbol{B}_i \boldsymbol{P}_i\right).\end{aligned} \tag{B.42}$$

[1]Please note that the expectation $\mathbb{E}[\cdot]$ is over the small-scale fading randomness. This average value is equal for any antenna or time slot.

In the third equality, the expectation is moved to the $\bar{\boldsymbol{x}}_{i,t}\bar{\boldsymbol{x}}_{i,t}^H$ based on the statistical independence between the vectors $\bar{\boldsymbol{h}}_{i,j}$ and $\bar{\boldsymbol{x}}_{i,t}$. The forth equality follows from $\mathbb{E}[\bar{\boldsymbol{x}}_{i,t}\bar{\boldsymbol{x}}_{i,t}^H] = \boldsymbol{I}_N$ and $\mathbb{E}[\bar{\boldsymbol{h}}_{i,j}\bar{\boldsymbol{h}}_{i,j}^H] = \boldsymbol{I}_N$. By following the same steps, it is easy to prove that

$$\mathbb{E}\left[\left|\boldsymbol{h}_{i,j}^H\boldsymbol{x}_{l\neq i,t}\right|^2\right] = \text{trace}\left(\boldsymbol{B}_i\boldsymbol{P}_l\right),\ \mathbb{E}\left[\left|\boldsymbol{h}_{i,j}^H\bar{\boldsymbol{x}}_t\right|^2\right] = \text{trace}\left(\boldsymbol{B}_i\bar{\boldsymbol{P}}\right).$$

Therefore

$$\mathbb{E}\left[Y_i\right] = \sum_{i\neq l}\text{trace}\left(\boldsymbol{B}_i\boldsymbol{P}_l\right) + \text{trace}\left(\boldsymbol{B}_i\bar{\boldsymbol{P}}\right) + \sigma^2. \tag{B.43}$$

To calculate the values of var(Y_i), let us first expand $\left|\boldsymbol{h}_{i,j}^H\boldsymbol{x}_{l,t}\right|^2$ as follows

$$\begin{aligned}\left|\boldsymbol{h}_{i,j}^H\boldsymbol{x}_{l,t}\right|^2 &= \bar{\boldsymbol{h}}_{i,j}^H\boldsymbol{B}_i^{\frac{1}{2}}\boldsymbol{P}_i^{\frac{1}{2}}\bar{\boldsymbol{x}}_{l,t}\bar{\boldsymbol{x}}_{l,t}^H\boldsymbol{P}_i^{\frac{1}{2}}\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{h}}_{i,j}\\ &= \bar{\boldsymbol{h}}_{i,j}^H\boldsymbol{B}_i^{\frac{1}{2}}\boldsymbol{P}_i^{\frac{1}{2}}\left(\boldsymbol{A}_{l,t} + \bar{\boldsymbol{A}}_{l,t}\right)\boldsymbol{P}_i^{\frac{1}{2}}\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{h}}_{i,j}\\ &\overset{N\to\infty}{\to} \text{trace}\left(\boldsymbol{B}_i\boldsymbol{P}_l\right) + \bar{\boldsymbol{h}}_{i,j}^H\boldsymbol{B}_i^{\frac{1}{2}}\boldsymbol{P}_i^{\frac{1}{2}}\bar{\boldsymbol{A}}_{l,t}\boldsymbol{P}_i^{\frac{1}{2}}\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{h}}_{i,j}\end{aligned} \tag{B.44}$$

where $\boldsymbol{A}_{l,t} + \bar{\boldsymbol{A}}_{l,t} = \bar{\boldsymbol{x}}_{l,t}\bar{\boldsymbol{x}}_{l,t}^H$, $\boldsymbol{A}_{l,t}$ is a diagonal matrix whose entries are exponential random variable (RV) with parameter 1, and $\bar{\boldsymbol{A}}_{l,t}$ is the equal to $\bar{\boldsymbol{x}}_{l,t}\bar{\boldsymbol{x}}_{l,t}^H$ but with zeros along the main diagonal. $\bar{\boldsymbol{h}}_{i,j}^H\boldsymbol{B}_i^{\frac{1}{2}}\boldsymbol{P}_i^{\frac{1}{2}}\boldsymbol{A}_{l,t}\boldsymbol{P}_i^{\frac{1}{2}}\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{h}}_{i,j} = \text{trace}\left(\boldsymbol{B}_i\boldsymbol{P}_l\right)$ follows from applying Corollary 1 in [88]. In the same way, it follows

$$\left|\boldsymbol{h}_{i,j}^H\bar{\boldsymbol{x}}_t\right|^2 \overset{N\to\infty}{\to} \text{trace}\left(\boldsymbol{B}_i\bar{\boldsymbol{P}}\right) + \bar{\boldsymbol{h}}_{i,j}^H\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{P}}^{\frac{1}{2}}\bar{\boldsymbol{A}}_t\bar{\boldsymbol{P}}^{\frac{1}{2}}\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{h}}_{i,j}. \tag{B.45}$$

Based on (B.44) and (B.45), yields

$$\begin{aligned}Y_i \overset{N\to\infty}{\to} \mathbb{E}\left[Y_i\right] + \sum_{l\neq i}\bar{\boldsymbol{h}}_{i,j}^H\boldsymbol{B}_i^{\frac{1}{2}}\boldsymbol{P}_i^{\frac{1}{2}}\bar{\boldsymbol{A}}_{l,t}\boldsymbol{P}_i^{\frac{1}{2}}\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{h}}_{i,j} +\\ \bar{\boldsymbol{h}}_{i,j}^H\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{P}}^{\frac{1}{2}}\bar{\boldsymbol{A}}_t\bar{\boldsymbol{P}}^{\frac{1}{2}}\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{h}}_{i,j},\end{aligned} \tag{B.46}$$

then

$$\begin{aligned}\text{var}\left(Y_i\right) = \mathbb{E}\left[\left|Y_i - \mathbb{E}\left[Y_i\right]\right|^2\right] \overset{N\to\infty}{\to}\\ \mathbb{E}\Bigg[\Bigg|\sum_{l\neq i}\bar{\boldsymbol{h}}_{i,j}^H\boldsymbol{B}_i^{\frac{1}{2}}\boldsymbol{P}_i^{\frac{1}{2}}\bar{\boldsymbol{A}}_{l,t}\boldsymbol{P}_i^{\frac{1}{2}}\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{h}}_{i,j}\\ +\ \bar{\boldsymbol{h}}_{i,j}^H\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{P}}^{\frac{1}{2}}\bar{\boldsymbol{A}}_t\bar{\boldsymbol{P}}^{\frac{1}{2}}\boldsymbol{B}_i^{\frac{1}{2}}\bar{\boldsymbol{h}}_{i,j}\Bigg|^2\Bigg] \overset{N\to\infty}{\to} 0\end{aligned} \tag{B.47}$$

The result in (B.47) can be justified as follows. Since the diagonal entries of $\bar{\boldsymbol{A}}_{l,t}$ and $\bar{\boldsymbol{A}}_t$ are zeros, then the quantity inside the second expectation in (B.47) is a summation of zero mean RVs, therefore, their expectation tends to zero.

The independence between X_i and Y_i implies $\text{cov}(X_i,\ Y_i) = 0$. By substituting this result and the results from (B.42), (B.43), and (B.47) in (B.41), we get the same result in (B.12). The proof for the value of SINR_{e_k} in (B.15) is omitted since it can be obtained by following the same previous steps used to obtain SINR_i. This concludes the proof.

References

[1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[2] K. R. Liu, *Cooperative communications and networking*. Cambridge University Press, 2009.

[3] Y. Jing, *Distributed space-time coding*, ser. SpringerBriefs in Computer Science. Springer New York, Apr. 2013.

[4] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge University Press, 2003.

[5] Y. Mehmood, W. Afzal, F. Ahmad, U. Younas, I. Rashid, and I. Mehmood, "Large scaled multi-user MIMO system so called massive MIMO systems for future wireless communication networks," in *2013 19th International Conference on Automation and Computing (ICAC)*, London, UK, Sept. 2013, pp. 1–4.

[6] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471., Feb. 2004.

[7] Runhua Chen, J. G. Andews, and R. W. Health, "Multiuser space-time block coded MIMO with downlink precoding," in *2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577)*, vol. 5, Paris, France, June 2004, pp. 2689–2693 Vol.5.

[8] R. F. Fischer, C. Stierstorfer, and J. B. Huber, "Point-to-point transmission over MIMO channels with intersymbol interference: Comparison of single-and multicarrier modulation," in *9th Int. OFDM-Workshop*, Dresden, Germany, sep. 2004, pp. 1–5.

[9] S. Hua, H. Liu, M. Wu, and S. S. Panwar, "Exploiting MIMO antennas in cooperative cognitive radio networks," in *2011 Proceedings IEEE INFOCOM*, Shanghai, China, April 2011, pp. 2714–2722.

[10] W. Chen and L. Hong, "Cooperative MIMO paradigms for cognitive radio networks," in *2013 IEEE International Symposium on Parallel Distributed Processing, Workshops and Phd Forum*, Cambridge, MA, USA, May 2013, pp. 734–740.

[11] J. N. Laneman, D. N. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, 2004.

[12] L. Rugini, P. Banelli, H. A. Suraweera, and C. Yuen, "Performance of Alamouti space-time coded OFDM with carrier frequency offset," in *2011 IEEE Global Telecommunications Conference - GLOBECOM 2011*, Houston, TX, USA, Dec. 2011, pp. 1–5.

[13] H. Q. Ngo, *Massive MIMO: Fundamentals and system designs*. Linköping University Electronic Press, 2015, vol. 1642.

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

[15] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 186–195, 2014.

[16] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Transactions on Information Theory*, vol. 48, no. 2, pp. 359–383, 2002.

[17] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 543–564, March 2000.

[18] B. Hassibi and B. M. Hochwald, "Cayley differential unitary space-time codes," *IEEE Transactions om Information Theory*, vol. 48, no. 6, pp. 1485–1503., Aug. 2002.

[19] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. Prentice Hall PTR New Jersey, Jan. 1996, vol. 2.

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

[21] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: 10 myths and one grand question," *arXiv preprint arXiv:1503.06854*, 2015.

[22] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2921–2935, May 2017.

[23] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[24] L. Sanguinetti, E. Bjornson, and J. Hoydis, "Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 232–257, Jan 2020.

[25] N. Beigiparast, G. M. Guvensen, and E. Ayanoglu, "The effect of antenna correlation in single-carrier massive MIMO transmission," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, Porto, Portugal, June 2018, pp. 1–7.

[26] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 7, pp. 1169–1174., Jul. 2000.

[27] J. Proakis and M. Salehi, *Digital communications*, ser. McGraw-Hill International Edition. McGraw-Hill, 2008. [Online]. Available: https://books.google.co.uk/books?id=ksh0GgAACAAJ

[28] B. L. Hughes, "Differential space-time modulation," *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2567–2578., Nov. 2000.

[29] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Transactions on Communications*, vol. 48, no. 12, pp. 2041–2052., Dec. 2000.

[30] B. Hassibi and T. L. Marzetta, "Multiple-antennas and isotropically random unitary inputs: the received signal density in closed form," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1473–1484, June 2002.

[31] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467., 1999.

[32] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1804–1824., Aug. 2002.

[33] M. O. Damen, K. Abed-Meraim, and J.-C. Belfiore, "Diagonal algebraic space-time block codes," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 628–636., Aug. 2002.

[34] R. W. Heath and A. J. Paulraj, "Linear dispersion codes for MIMO systems based on frame theory," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2429–2441., Nov. 2002.

[35] H. El Gamal and M. O. Damen, "Universal space-time coding," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1097–1119., May 2003.

[36] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space-time block codes," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3885–3902., Sept. 2006.

[37] L. T. Berger and L. Schumacher, "Modified space-time transmission in DS-CDMA downlink facilitating MISO channel equalization," in *Proceedings IEEE 56th Vehicular Technology Conference*, vol. 2, Vancouver, BC, Canada, Sept. 2002, pp. 941–945 vol.2.

[38] Y. Hong, E. Viterbo, and J. Belfiore, "A space-time block coded multiuser MIMO downlink transmission scheme," in *2006 IEEE International Symposium on Information Theory*, Seattle, WA, USA, July 2006, pp. 257–261.

[39] F. Alsifiany, A. Ikhlef, and J. Chambers, "On differential modulation in downlink multiuser MIMO systems," in *2017 25th European Signal Processing Conference (EUSIPCO)*, Kos, Greece, Aug. 2017, pp. 558–562.

[40] F. Alsifiany, A. Ikhlef, and J. Chambers, "Exploiting high rate differential algebraic space-time block code in downlink multiuser MIMO systems," *IET Communications*, vol. 12, no. 17, pp. 2188–2197, Oct. 2018.

[41] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[42] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving "Massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3226–3239, Sept. 2012.

[43] 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, Apr. 2003.

[44] 3GPP. (2015) Preliminary observations and proposals for DMRS enhancement for EBF/FD-MIMO. Accessed: 2019-04-08. [Online]. Available: https://www.3gpp.org/DynaReport/TDocExMtg–R1-81–31256.htm

[45] M. K. Simon and M.-S. Alouini, "Multiple symbol differential detection with diversity reception," *IEEE Transactions on Communications*, vol. 49, no. 8, pp. 1312–1319, Aug. 2001.

[46] V. Lottici and Z. Tian, "Multiple symbol differential detection for UWB communications," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, pp. 1656–1666, May 2008.

[47] A. Schenk and R. F. H. Fischer, "Noncoherent detection in massive MIMO systems," in *WSA 2013; 17th International ITG Workshop on Smart Antennas*, Stuttgart, Germany, March 2013, pp. 1–8.

[48] R. M. Buehrer and N. A. Kumar, "The impact of channel estimation error on space-time block codes," in *Proceedings IEEE 56th Vehicular Technology Conference*, vol. 3, Vancouver, BC, Canada, Sep. 2002, pp. 1921–1925 vol.3.

[49] G. H. Golub and C. F. Van Loan, *Matrix computations.* The Johns Hopkins University Press, 1996, vol. 3.

[50] K. Zu, R. C. de Lamare, and M. Haardt, "Generalized design of low-complexity block diagonalization type precoding algorithms for multiuser MIMO systems." *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4232–4242., Oct. 2013.

[51] P. Nagaradjane and T. Muthu, "Performance of feedback multiple-input multiple-output system aided by polarisation multiplexing and pre-processing," *IET Communications*, vol. 9, no. 16, pp. 1988–1998., 2015.

[52] C. Xu, S. Sugiura, S. X. Ng, P. Zhang, L. Wang, and L. Hanzo, "Two decades of MIMO design tradeoffs and reduced-complexity MIMO detection in near-capacity systems," *IEEE Access*, vol. 5, pp. 18 564–18 632., May 2017.

[53] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765., Mar. 1998.

[54] J. Boutros and E. Viterbo, "Signal space diversity: a power-and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Transactions om Information Theory*, vol. 44, no. 4, pp. 1453–1467., Jul. 1998.

[55] X. Giraud, E. Boutillon, and J.-C. Belfiore, "Algebraic tools to build modulation schemes for fading channels," *IEEE Transactions om Information Theory*, vol. 43, no. 3, pp. 938–952., May 1997.

[56] B. M. Hochwald and S. Ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 389–399., Apr. 2003.

[57] K. Zu and R. C. de Lamare, "Low-complexity lattice reduction-aided regularized block diagonalization for MU-MIMO systems," *IEEE Communications Letter*, vol. 16, no. 6, pp. 925–928., Apr. 2012.

[58] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[59] K. N. R. Prasad, E. Hossain, and V. K. Bhargava, "Energy efficiency in massive MIMO-based 5G networks: Opportunities and challenges," *IEEE Wireless Communications*, vol. 24, no. 3, pp. 86–94, June 2017.

[60] X. Wang and C. Zhai, "Simultaneous wireless information and power transfer for downlink multi-user massive antenna-array systems," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 4039–4048, Sep. 2017.

[61] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[62] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission with an active eavesdropper," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3880–3900, July 2016.

[63] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264–273, February 2013.

[64] Y. Wu, S. Jin, X. Gao, M. R. McKay, and C. Xiao, "Transmit designs for the MIMO broadcast channel with statistical CSI," *IEEE Transactions on Signal Processing*, vol. 62, no. 17, pp. 4451–4466, Sep. 2014.

[65] M. Alageli, A. Ikhlef, and J. Chambers, "Optimization for maximizing sum secrecy rate in MU-MISO SWIPT systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 1, pp. 537–553, Jan. 2018.

[66] M. Alageli, A. Ikhlef, and J. Chambers, "Optimal transmit power minimization in secure MU-MISO SWIPT systems," in *2017 IEEE 17th International Conference on Ubiquitous Wireless Broadband (ICUWB)*, Salamanca, Spain, Sept. 2017, pp. 1–7.

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

[68] A. Tarighat, M. Sadek, and A. H. Sayed, "A multi user beamforming scheme for downlink MIMO channels based on maximizing signal-to-leakage ratios," in *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 3, Philadelphia, PA, USA, March 2005, pp. iii/1129–iii/1132 Vol. 3.

[69] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, Jan. 2001.

[70] D. Warrier and U. Madhow, "Spectrally efficient noncoherent communication," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 651–668, March 2002.

[71] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *1998 URSI International Symposium on Signals, Systems, and Electronics Conference Processing (Cat. No.98EX167)*, Pisa, Italy], year=1998, volume=, number=,

pages=295-300, keywords=indoor radio;radiowave propagation;electromagnetic wave scattering;multipath channels, doi=10.1109/ISSSE.1998.738086, ISSN=, month=Oct.,.

[72] 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, March 2017.

[73] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: a new next-generation paradigm," *IEEE Access*, vol. 7, pp. 99 878–99 888, 2019.

[74] M. Alageli, A. Ikhlef, and J. Chambers, "SWIPT massive MIMO systems with active eavesdropping," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 1, pp. 233–247, Jan 2019.

[75] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3899–3911, July 2015.

[76] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "A distributed SON-based user-centric backhaul provisioning scheme," *IEEE Access*, vol. 4, pp. 2314–2330, 2016.

[77] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: challenges and research advances," *IEEE Network*, vol. 28, no. 6, pp. 6–11, Nov 2014.

[78] D. Kapetanovic, G. Zheng, and F. Rusek, "Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 21–27, June 2015.

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

[80] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Stockholm, Sweden, June 2015, pp. 201–205.

[81] J. Zhang, C. Yuen, C. Wen, S. Jin, K. Wong, and H. Zhu, "Large system secrecy rate analysis for SWIPT MIMO wiretap channels," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 1, pp. 74–85, Jan 2016.

[82] J. Zhu, Y. Li, N. Wang, and W. Xu, "Wireless information and power transfer in secure massive MIMO downlink with phase noise," *IEEE Wireless Communications Letters*, vol. 6, no. 3, pp. 298–301, June 2017.

[83] L. Wang, K. Wong, M. Elkashlan, A. Nallanathan, and S. Lambotharan, "Secrecy and energy efficiency in massive MIMO aided heterogeneous C-RAN: A new look at interference," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 8, pp. 1375–1389, Dec 2016.

[84] Y. Zhu, L. Wang, K. Wong, S. Jin, and Z. Zheng, "Wireless power transfer in massive MIMO-aided HetNets with user association," *IEEE Transactions on Communications*, vol. 64, no. 10, pp. 4181–4195, Oct 2016.

[85] N. Nguyen, H. Q. Ngo, T. Q. Duong, H. D. Tuan, and K. Tourki, "Secure massive MIMO with the artificial noise-aided downlink training," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 4, pp. 802–816, April 2018.

[86] G. Chen, Z. Tian, Y. Gong, Z. Chen, and J. A. Chambers, "Max-ratio relay selection in secure buffer-aided cooperative wireless networks," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 4, pp. 719–729, Apr. 2014.

[87] A. M. Tulino, S. Verdú *et al.*, "Random matrix theory and wireless communications," *Foundations and Trends® in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[88] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2059–2078, Sep. 2000.