Enhanced User Scheduling in MU-MIMO Broadcast Channels

Nikeeth Ramanathan

Submitted in total fulfilment of the requirements
for the degree of

Master of Philosohpy

March 2013

Department of Electrical and Electronic Engineering
The University of Melbourne

Abstract

Multi-user Multiple-Input Multiple-Output (MU-MIMO) systems are becoming increasingly important in wireless telecommunication networks. The multiple antennas at the base station allow for multiple users to be scheduled at the same time. This leads to multiplicative gains in transmission rates when compared to point-to-point communication. However, in order to reap the full benefits of the MU-MIMO, accurate channel state information (CSI) is required. This is a bigger issue in downlink broadcast channels (BC) as attaining accurate CSI would require users to feedback large amounts of information leading to large bandwidth requirements. As a result, this would restrict the data transmission rate in the uplink. Therefore, it is important to focus on limited or finite-rate feedback (LF/FRF) schemes.

In limited feedback, the CSI feedback is restricted by the number of feedback bits allocated by each user. By limiting the feedback bits, a more practical scenario could be modeled whereby users utilize low complexity codebooks to feedback information with minimal delay. The trade-off that comes with this is saturation in the performance of the MU-MIMO BC. As the transmit power or signal-to-noise ratio (SNR) is increased, inter-user interference becomes a major problem. While transmission schemes such as zero-forcing beamforming (ZFBF) aid to resolve the performance to an extent, the limited CSI at the transmitter (CSIT) would imply that this interference cannot be completely subdued. In order to overcome this performance saturation, in this thesis, we investigate multi-user transmission and user-scheduling schemes with limited feedback.

Scheduling of users is a critical and challenging problem. Base stations can attempt to maximize the performance of the broadcast channel by choosing the best set of users. However, in order to do so, an exhaustive search

over all possible user-sets is required. This is generally computationally infeasible. Therefore, sub-optimal algorithms are required. While classical algorithms like semi-orthogonal user scheduling (SUS) increase the performance of schemes utilizing ZFBF, the performance still saturates.

This thesis develops novel user-scheduling schemes to overcome the performance limits experienced in MU-MIMO BC. Our investigations identify the problems with classical schemes in their attempt to schedule as many users as possible. However, in the case when the channel is interference limited it is often better to switch to a smaller user-set in order to maximize performance. This raises an important research question of "how many users should the base station select?". Some attempts to answer this question have been made recently by employing multi-mode user scheduling. Here, each user approximates the system rate for different settings in order to work out its preferred number of co-scheduled users (mode of operation). The base station receives the preferred mode of operation of all the users along with CSI via limited feedback and makes a decision on the number of users to schedule. The main contribution of this thesis is the derivation of novel closed form expressions for rate approximations employed to determine their preferred mode in two different schemes. The first scheme outperforms the existing multi-mode scheduling schemes by providing improved scheduling criteria. The second is a novel scheme that combines SUS with multi-mode scheduling. Experimental simulations show the significant performance gains achieved by these schemes and reveal important considerations for future research.

Declaration

This is to certify that

(i) the thesis comprises only my original work towards the degree of Master of Philosophy,

(ii) due acknowledgement has been made in the text to all other material used,

(iii) the thesis is fewer than 50,000 words in length, exclusive of words in tables, figures and bibliography.

Nikeeth Ramanathan

Supervisors: A/Prof. Margreta Kuijper and Dr.Feng Li
External Supervisor: Prof. Jamie Evans

Acknowledgements

I wish to acknowledge the encouragement and support I have received from a number of people during my postgraduate study in The University of Melbourne. First and foremost I would like to thank Prof. Jamie Evans for suggesting the Gigabit Wireless research project to pursue and for his inspiring guidance throughout my research study. I wish to also express my gratitude to my Principal Supervisor, Associate Professor Margreta Kuijper for her dedication and encouragement rendered towards the progress and success of my research work. Special mention of appreciation also goes to my Co-Supervisor, Dr. Feng Li for his tireless effort and helpful suggestions provided from time to time. I am indebted to all my supervisors for supporting me with confidence to independently carry out my research work.

In addition to my supervisors, I would also like to thank Prof. Darryl Veitch for his constructive inputs as my MPhil advisory committee member. I also wish to acknowledge the invaluable support of Prof. Ampalavanapillai Nirmalthas and Prof. Rob Evans while serving as Department Head. I wish to thank all those I had interacted with during the advanced studies I underwent on Information Theory and Convex Optimization and the rich knowledge gained.

I wish to acknowledge the Department of Electrical and Electronic Engineering at Melbourne School of Engineering for funding my MPhil research through the Australian Postgraduate Award (APA) Industry scholarship. The funding was provided by the Australian Research Council (ARC)

Linkage Project on Gigabit Wireless: Setting the Standard for Tomorrow's Broadband with the industry partner, NEC Australia, and I am grateful to the project leader, Prof. Jamie Evans for providing this additional support.

Finally, I wish to express my sincere thanks to my parents, friends and fellow researchers for their constant support, understanding and encouraging words that had helped me to a great extent in accomplishing the various milestones of my MPhil research study.

Contents

List of Abbreviations

Abbreviation	Description
BC	Broadcast channel
BD	Block diagonalization
CDF	Cumulative density function
CDI	Channel direction information
CQI	Channel quality indicator
CSI	Channel state information
CSIR	Channel state information at the receiver
CSIT	Channel state information at the transmitter
DL	Downlink
DPC	Dirty paper coding
FDD	Frequency-division duplexing
FRF	Finite rate feedback
i.i.d.	Independent and identically distributed
LF	Limited feedback
MIMO	Multiple-input multiple-output
MISO	Multiple-input single-output
MMSE	Minimum mean square error
MSE	Mean square error
MU	Multi-user
OFDM	Orthogonal frequency-division multiplexing
PDF	Probability density function
RVQ	Random vector quantization
SINR	Signal to interference-plus-noise ratio
SISO	Single-input, single-output
SNR	Signal to noise ratio
SU	Single-user
SUS	Semi-orthogonal user scheduling
SVD	Singular-value decomposition
TDD	Time-division duplexing
VQ	Vector quantization
ZFBF	Zero-forcing beamforming

List of Symbols

Symbol	Description
a	Complex-valued scalar
$\|a\|$	Absolute value
a^*	Complex-conjugate
$\mathbf{a}$	Complex-valued vector
$\|\mathbf{a}\|$	Euclidean norm
$\mathbf{A}$	Complex-valued matrix
$\|\mathbf{A}\|$	Determinant of matrix
$\mathbf{A}^{-1}$	Matrix inverse
$\mathbf{A}^\dagger$	Pseudo-inverse (Penrose-Moore)
$\mathbf{A}^*$, $\mathbf{a}^*$	Element-wise complex conjugate
$\mathbf{A}^H$, $\mathbf{a}^H$	Complex conjugate transpose (Hermitian)
$\mathbf{A}^T$, $\mathbf{a}^T$	Transpose operator
$Tr(\mathbf{A})$	Trace of matrix
max	Maximum value
min	Minimum value
arg max	Maximizing argument
arg min	Minimizing argument
$\mathbb{E}(.)$	Expectation of random variables
$\mathbb{R}^{m\times n}$	Real space of dimensions $m \times n$ (m or n omitted if they are of dimension 1)
$\mathbb{C}^{m\times n}$	Complex space of dimensions $m \times n$
$Co(.)$	Convex hull
$\cup$	Union
$\cap$	Intersection

Chapter 1

Introduction

The ever-rising demand for broadband applications such as multimedia services has driven the need for high-speed computer networks and reliable wireless communication systems [1, 2]. These services are required to be flexible enough to be delivered anywhere anytime at home and business environments. Recent developments in cellular technology have resulted in smart-phones and other wireless portable devices that are an integral part of almost every individual's daily life. The resulting outcome being witnessed is an unprecedented increase in the traffic of cellular networks. This poses a major challenge in achieving the high quality of service (QoS) demands of the exponentially growing number of cellular users [3]. A top-level QoS could be attained predominantly via higher data transmission rates and ubiquitous connectivity for multiple users. However, QoS is fundamentally constrained by limited wireless resources, such as frequency spectrum, channel power and channel conditions which dictate fading signals and inter-cell interference [4]. These limitations have raised an immense scope for potential research contributions in this direction.

In recent years, cellular technology developments and innovative communication techniques have emerged to improve the data transmission rates in wireless systems. Among the factors that affect the performance of wire-

less transmission, the effects of fading due to fluctuations in signal strength could be addressed by deploying multiple antennas at both the base station (BS) as well as receiver units (users). The resulting configuration is a multiple-input multiple-output (MIMO) system [5]. In MIMO, the multiple transmit and receive antennas can be exploited for improving the data rate of a single-user, via effective transreceiver designs. When the transmit antennas are used to broadcast data requested by multiple users, more users in a channel are being served, thereby improving the system throughput [6]. In such a scenario Multi-User (MU) diversity can be exploited, where the user selection is done opportunistically to address fading signals. Such a channel is termed as Multi-User MIMO (MU-MIMO).

A wireless communication network typically has a downlink and an uplink [7]. In the downlink, the base station transmits information to various users in the system, via broadcast [8]. In the uplink, users transmit information to the base station. We are interested particularly in Frequency Division Duplex (FDD) where the downlink and uplink operate in separate frequency bands, and hence the base station cannot predict the downlink channel from observing the uplink. In such a scenario, each user will have to feedback information about its channel. A series of active research has investigated on MU-MIMO channels and multi-user diversity. While many of the problems involving uplink transmission have been addressed, research on the downlink broadcast channel is still progressing and the main results are still fairly recent [4, 9]. Many studies have focused on arriving at performance limits of MU-MIMO channels using information theoretic approaches and uplink-downlink duality [10, 11]. The capacity achieving transmission schemes break down in situations where the transmitter does not have full channel state information (CSI). Furthermore, these schemes are often too computationally complex to be implemented in practice. As a result, progress was made on devising simpler linear transmission schemes

that utilize beamforming to steer the user-requested data symbols in order to mitigate interference or maximize the signal power, or both. In order to improve average system throughput, user scheduling algorithms were employed [12, 13]. While not optimal, splitting the capacity optimization problem into simpler design problems reduces the computational load of the system as well as improves our understanding of the various research scopes in a MU-MIMO Broadcast Channel (BC) with limited CSI. As such, this thesis analyses existing transmission schemes and proposes an enhanced user scheduling design for a MU-MIMO BC with limited feedback and zero-forcing beamforming (ZFBF). The research is limited to a single cell, with a single base station and multiple users (receivers). Interference from external cells is not considered.

The rest of the chapter is organised as follows. Section 1.1 provides the background theory and concepts evolving with MU-MIMO. In Section 1.2, we explain the current limitations and the motivation of this research study. Next, we state the aim of the thesis and the research questions in Section 1.3. Finally, Section 1.4 provides the contributions of this research and organisation of the thesis.

1.1 Background

A pioneer in information theory, Claude Shannon is considered the father of modern communications theory due to his famous paper published in 1948 that forms the basis of communication systems design. In his paper [14], the noisy channel coding theorem provides an absolute limit on how fast it is possible to transmit error-free data within a channel of a given bandwidth, and with given noise conditions within that channel. In other words, Shannon's concept of channel capacity becomes the fundamental principle to determine the maximum data transmission rate that can be achieved over the channel. Shannon derived the channel capacity (C) with an additive

white Gaussian noise (AWGN) into a compact formula as given below:

$$C = W \log \left(1 + \frac{P}{N_0 W}\right) \tag{1.1}$$

where W is the available bandwidth in Hz, N_0 is the one-sided noise power spectral density in Watt/Hz, and P is the transmit power in Watts. The ratio of received signal to noise power, P/N_0W is called the signal-to-noise ratio (SNR).

Shannon postulated that if it is possible to separate every output in the receiver, finding the most closely matched input would yield an optimum decoding method. However, the primary implementation obstacle is that for all but the shortest bit sequences, the computational complexity in terms of memory and processing time required to decode the noisy received data is very high [15]. Hence, for all practical purposes, Shannon capacity was unreachable in reality. However, in 1990s, the introduction of Turbo Codes [16] and Gallager's Low Density Parity Check (LDPC) codes rediscovered by [17], along with increased computing power enabled ideal wireless transmission performance to approach the Shannon limit. These coding and signal processing techniques assume large bandwidth and power, which are scarce resources in many practical wireless systems, resulting in more research in optimising these resources.

Early research showed that by employing multiple antennas at both transmitter and receiver ends, the capacity can increase [18, 19] linearly by about the minimum number of antennas used between the two communication ends, as compared to using a single antenna at both ends [18, 19]. This capacity increase may not be noteworthy in practice when users typically have a single or dual antenna devices. In a single-user (SU) scenario, these multiple antennas could be exploited simultaneously through beamforming for improving the received SNR. Beamforming is a method used to steer the data symbol in the direction of the user channel, by adjusting the

power of the transmit antennas accordingly. In the case where users have only one or two antennas the data transmission rate could be improved by incorporating multiple users.

In MU-MIMO systems, where multiple users operate in the same frequency and time bands but are separated in space, the most prominent channels fall under two categories, MIMO broadcast channel (BC) and MIMO multiple-access channel (MAC). While MIMO BC models a downlink transmission from one base station to many receivers or multiple users, the MIMO MAC models an uplink transmission from multiple users to a single base station. In both cases the performance is determined as the maximum sum of each user's data rate, denoted by the *sum-rate*. In the case of MIMO MAC, successive interference cancellation at the receiver (BS) could be easily adopted to achieve the capacity region. In the case of MIMO BC, dirty paper coding (DPC) which is a well-known precoding scheme may be employed at the BS. Several research studies have shown that dirty paper coding (DPC) can achieve maximum sum capacity [20, 21, 22]. However, achieving DPC capacity has been reported to be impractical in many wireless applications due to non-availability of full channel state information (CSI) and its high complexity [23, 1, 13]. In the uplink however, the CSI can be retrieved via well designed channel estimator. This makes it more challenging to arrive at optimal transmission schemes for MIMO BC than for MIMO MAC.

When analysing the sum-rate of MU-MIMO transmission schemes two important performance measures can be noted, namely the Spatial Diversity gain and Multiplexing gain. The diversity gain is an additive performance improvement observed as the number of users in the system increases, because the probability of finding users with better channel conditions increases. As both the number of users and antennas in the base station increase, so does the steepness of the sum-rate curve at high SNR in the log

scale. This multiplicative increase in performance is known as the Multiplexing gain and is limited by whichever is lower - the number of base station antennas or the total number of antennas of the scheduled users. Further information on these performance gains can be seen in Chapter 2. While transmitting to more users results in performance gains, in the case when the base station does not have full CSI, the increase in the number of scheduled users will also result in an increase in interference. Hence, in MU-MIMO, practical techniques to suppress inter-user interference in the BC and to exploit partial CSI are being developed, utilizing downlink beamforming at the base station and limited feedback from users [5, 23]. In addition, user selection plays an important role in achieving higher throughput in MU-MIMO BC and poses more scope for discovery of efficient transmission schemes that are essential for achieving the desired QoS for the ever-growing large wireless user-base.

1.2 Motivations of the Study

After more than a decade of research in combining multiple antenna techniques with advanced signal processing schemes in MIMO technology, recent developments and interests have been moving into commercial wireless communication systems [24, 12]. Though MIMO is getting standardised in Worldwide Interoperability for Microwave Access (WiMAX) and 3rd Generation Partnership Project (3GPP) specifications, more advanced forms of MIMO are required for future releases of 3GPP Long Term Evolution (LTE) and LTE- Advanced systems [3, 2]. With the various MIMO schemes being standardised in 3GPP systems, the focus of future LTE specifications is to provide the base station with advanced capability to dynamically select an optimal MIMO scheme when channel conditions change in mobile environments.

Hence, a motivation for this study is to analyse the effect of allowing

the base station to decide on the number of users to schedule, which we will denote as "mode" of operation. The problem of "how to schedule users?" has seen developments over the past decade. However, the problem of "how many users to schedule?" has seen its importance only in recent years. The early progress in literature states that scheduling more users is better. This is justified because, assuming each user meets its QoS requirements, transmitting to several users allows for attaining spatial multiplexing gains. However, the transmission schemes that amount to these multiplexing gains highly depend on the accuracy of CSI at the transmitter (CSIT). With limited CSIT, the same arguments may hold, but only up to a certain point. As the transmit power is increased in the base station, interference between users becomes a bigger issue, and without full CSIT it is not possible to completely remove interference. While user scheduling aids in improving throughput in such scenarios, classical scheduling schemes aim to schedule as many users as possible. Because we are no longer completely cancelling interference with limited CSI, more users amount to more interference, and as the transmit power grows large, so does the interference power, causing the throughput of the system to saturate.

While one way to deal with the problem is to increase the amount of CSIT as the transmit power increases, this is often impractical. This is because the users (mobile phones), which have to feedback this information, are limited in terms of computational resources. Hence, rather than increasing feedback information, we need to limit or even reduce feedback overhead and look toward other parameters that can improve performance. The other way to deal with the problem is to reduce interference power. The base station can do this by scheduling users such that interference is reduced, based on the limited feedback information from the users. Classical scheduling algorithms such as semi-orthogonal user selection and greedy user selection can be utilized provided there is a large number of users to choose from.

With a limited number of users, the base station might not always find users that are suited to reduce interference. Furthermore, scheduling algorithms involving a large user-set are often computationally taxing.

Considering (semi-)practical scenarios with low (but error free) feedback information and a reasonable number of users, an alternative method to reduce interference is to simply reduce the number of users that are scheduled. This motivates research in the direction of designing effective schemes where the base station utilizes feedback information from the users to determine the system conditions. The base station can utilize this information to determine the number of users to schedule in order to maximize performance.

1.3 Research Questions

The primary aim of this thesis is to study the performance of transmission techniques in MU-MIMO downlink channels and to develop adaptive multi-mode transmission schemes for achieving improved performance. In order to accomplish this, we address the following research questions:

Q1. What are the popular multi-user transmission schemes and how do they compare? (Chapter 3)

Q2. How can we exploit partial CSI and perform user scheduling via limited feedback channels? (Chapter 4)

Q3. Can an adaptive multi-mode transmission scheme with user scheduling employing zero-forcing beamforming (ZFBF) and limited feedback provide significant performance gains in MU-MIMO broadcast channels? (Chapter 5)

1.4 Contributions and Organisation of the Thesis

The main contribution of this thesis is the proposition of novel user scheduling schemes that incorporate mode selection at the base station. Mode selection allows the base station to select the number of users to schedule utilizing information provided by all the users. This improves performance as the base station can utilize the user-side information to limit the number of scheduled users in high-interference scenarios, hence overcoming throughput saturation common in classical scheduling algorithms [25]. The thesis also compares different schemes and provides analysis on the important factors to consider when selecting the number of users and how they affect performance, via simulations. The initial contributions expand on previous work in this area via deriving closed form solutions for lower bounds of system performance, estimated at the user-side. We also establish improved criteria for user selection at the base station. More prominent contributions include incorporating semi-orthogonal user scheduling (SUS) in the multi-mode transmission scheme to further improve performance and novel closed form derivations of improved rate approximations at the users. The derivation process involved in the closed form expressions reveal the parameters which are important and those that are not. The derivations also highlight the similarities and differences between the rate approximations. Furthermore, these closed-form solutions can then be adapted into practical algorithms used in the base station and more importantly mobile phones, which may not have the processing capacity to perform accurate numerical integrations with minimal delay. The simulation results show significant performance improvements over existing multi-mode transmission schemes, as well as provide open pathways for future research.

The remainder of the thesis is organized as follows. Chapter 2 provides the system model, background and overview of a general downlink broadcast channel. It discusses the practical constraints as well as performance

measures that we identify as important considerations for investigating the MU-MIMO broadcast channels. Chapter 3 outlines the various full-CSI multi-user transmission schemes currently utilized in literature, establishing their practical limitations due to their ideal assumption of perfect channel knowledge. Chapter 4 describes the notion of limited feedback and the considerations required when the transmitter does not have full-CSI information in both single and multi-user systems. Furthermore, the chapter discusses how user scheduling aids in further achieving full multiplexing gain of limited feedback and establishes performance comparisons of popular user scheduling algorithms reported in literature. From these background investigations, Chapter 5 identifies the drawbacks of classical schemes with the main objective to improve on multi-mode transmission schemes present in recent literature. This chapter proposes novel user scheduling schemes for MU-MIMO broadcast channels and establishes significant sum-rate performance gains through theoretical and experimental investigations.

Chapter 2

System Overview

This chapter provides an overview of the wireless communication system being analysed. The aim of the chapter is to illustrate the various stages involved in the system in order to give a big picture of the system as a whole while also pointing out the topic of interest, where our research and contributions lie. Firstly, a schematic for a typical physical layer transmission model will be described. This will be followed by introducing the baseband equivalent system model as well as some fundamental notations. The function of precoding or beamforming in a multi-user scenario will be explained. Thereafter, the purpose of feedback and the different receiver structures will be outlined. Throughout this section the general assumptions made in the research will be stated. Finally, the chapter overviews the various performance measures that may be considered in the thesis.

2.1 Physical Downlink

A high-level view of the transmitter side operations at the physical layer is given in Fig. 2.1. The figure illustrates an extended version of Shanon's model of a general communication system [14] which traditionally consists of an information source, transmitter, noisy channel, receiver and information

sink.

Transmitter Model

The information source collects data from the upper layer(s) which can be in the form of voice signals, video, text or commands. The data will then be quantized into a set of symbols. The source encoder compresses this information and converts it into a bit stream. Redundancy (parity bits) is then added to the bit stream via the outer channel encoder, forming an error-correction code. This redundancy is used by the receiver to detect and correct errors in the bit stream. The symbol map component converts the bits from the outer channel encoder into symbols. These symbols are in the form of constellation points, in $\mathbb{C}$. The inner coder often consisting of space-time block codes, such as Alamouti's scheme, which aims to create multiple copies of the data symbols to be transmitted across the transmit antennas. Layer mapping is used to arrange the data symbols based on what is required by each user/receiver at each time slot. Layer mapping is often done in OFDM based systems to allocate data symbols into the time-frequency grid. Beamforming or precoding is then used to combine the incoming data symbols into a number of streams equal to the number of transmit antennas. In simplistic terms, the role of beamforming is to adjust the power of each antenna in order to steer the data symbol towards a user. This is mathematically done by attaching a vector (pertaining to the antenna power) that directs each user's requested data. The resulting vectors are then superimposed before being broadcast. Note that space time codes can be used without beamforming and vice-versa [26]. While beamforming provides the same diversity gain along with an array gain(received power), it requires channel state information at the transmitter in certain schemes. Finally, in order to broadcast the information, the data streams are then shaped into pulses and modulated onto carrier waves to be sent across the

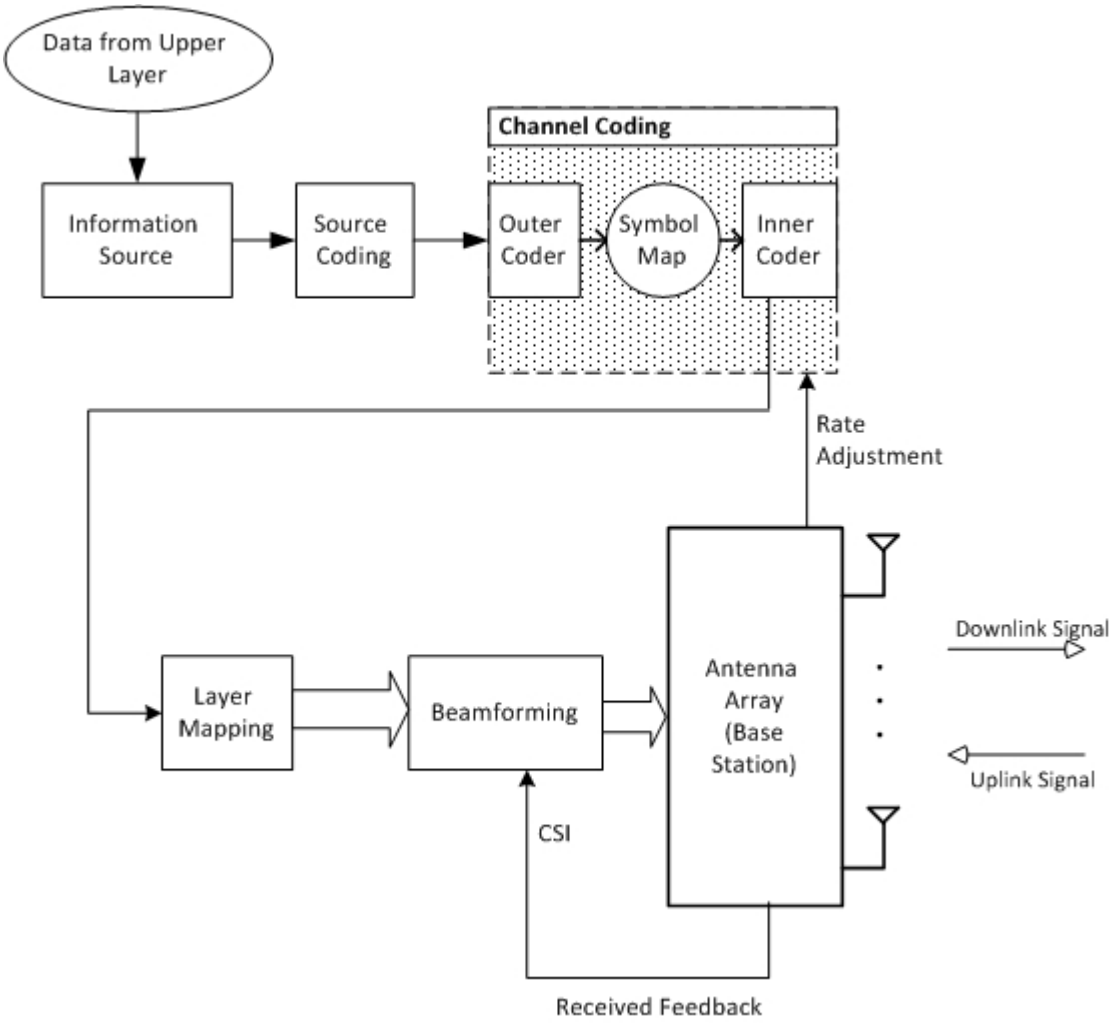

Figure 2.1: Transmitter Side Schematic

wireless channel.

The transmitter simultaneously receives signals through the uplink pertaining to feedback information needed for beamforming, from each user. Furthermore, the CSI that is received can be used for rate adjustments in the channel coders as well as user scheduling [7].

Receiver Model

On the receiver side, the antenna array receives the signal from the wireless channel and a number of operations are performed before converting it back to data symbols. Firstly, the carrier is removed via a receive filter, converting the signal into its baseband equivalent.

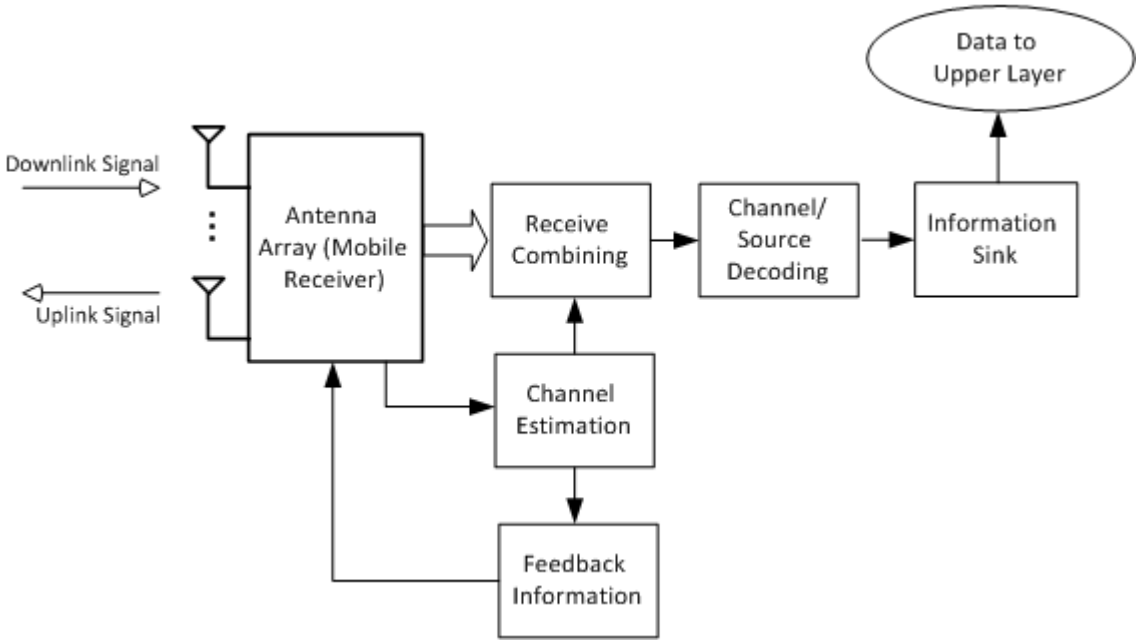

Figure 2.2: Receiver Side Schematic

Next, the signal is sampled into a discrete time series and this will be the effective received signal used for receive combining. Receive combining aims to convert the problem of joint decoding of data streams into an independent decoding problem by combining the received signals from each antenna in an optimal way to reduce inter-antenna interference as well as noise. This of course requires channel knowledge [27]. Channel decoding algorithms (e.g. ML/minimum distance decoding) may then be used to retrieve the data symbols which will then be converted into bits. These bits will be decoded again, this time to remove redudancies and detect/correct errors. Depending on the type of encoder and decoder, this two stage process can

be combined into one using soft information (e.g. distance between data point and constellation point) to decode the data stream directly into bits [7]. The decoded digital data can then be decompressed and sent back to the upper layer in the system in order to recreate its respective voice/video signal, for example.

The receiver also consists of a channel estimation block which functions to accurately estimate the wireless channel at the receiver side [28]. The channel estimation for example, could use observation of a training sequence sent by the transmitter over time in conjunction with least-squares algorithm to predict the effective channel. In the case where there is inter-symbol interference, typically an equilizer is used in conjunction to the predicted channel to reduce its effects. The channel estimate can also be fed back to the transmitter via the uplink either by finite rate digital feedback or analog feedback in order to provide the transmitter with partial CSI.

2.2 MIMO Baseband Model

One of the main models used to analyse various systems in this research is the baseband equivalent model for the MIMO communication system. This model ignores stages prior to beamforming and post receive combining. The outline of the model is depicted in Fig. 2.3. The figure depcits the transmitter-receiver model for a particular user k. Each user is assumed to have the same setup.

The input-output relationship between $\mathbf{x}$, the signal broadcast by the base station, and $\mathbf{y}_k$, the signal received at user k, is given by the expression below:

$$\mathbf{y}_k(n) = \mathbf{H}_k(n)\mathbf{x}(n) + \mathbf{z}_k(n), \quad k = 1, \ldots, K \tag{2.1}$$

Assuming that the transmitter has N_T antennas and the k^{th} user has N_{Rk} receive antennas, the input is given by $\mathbf{x}(n) \in \mathbb{C}^{N_T \times 1}$, the channel is

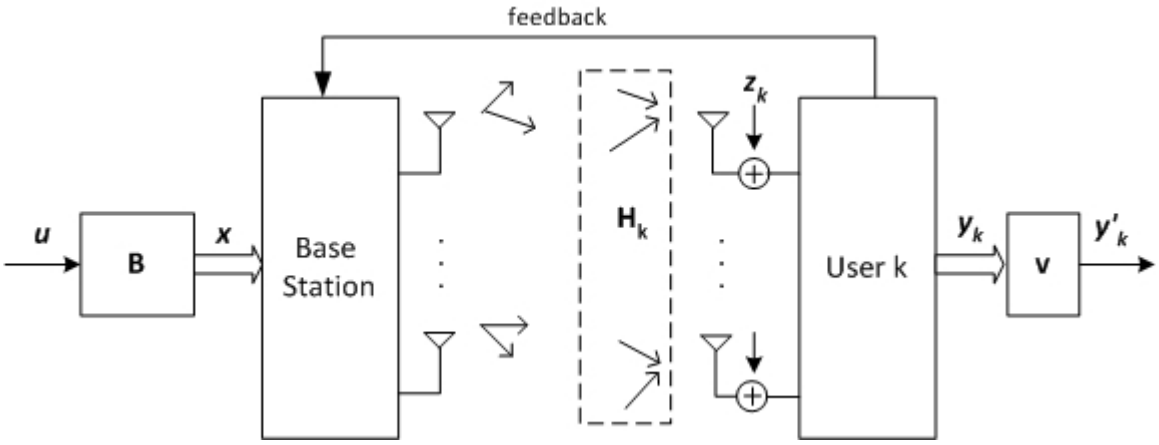

Figure 2.3: MIMO model for each user

given by $\mathbf{H}_k(n) \in \mathbb{C}^{N_{Rk} \times N_T}$, the output is given by $\mathbf{y}_k(n) \in \mathbb{C}^{N_{Rk} \times 1}$ and the AWGN is given by $\mathbf{z}_k(n) \in \mathbb{C}^{N_{Rk} \times 1}$. Here, n indexes the discrete time instant or channel use of the information transfer. In the case where there is no ISI as in the situation above, the index n can be dropped [29]. Most of the analysis in this thesis deals with a system in a particular time instant and not over multiple time instances. Unless otherwise stated, the elements of $\mathbf{H}_k(n)$ and $\mathbf{z}_k(n)$ are generated using an i.i.d. circular symmetric normal distribution denoted by $\mathcal{CN}(0, \sigma^2)$.

Beamforming

Beamforming is a type of linear precoding that superimposes the data requested by different users into a number of data streams determined by the transmit antennas. Unlike space-time coding which sends replica of the data requested by each user (over different antennas and over time), beamforming combines the user information together [26]. The other major difference is that, in certain beamforming schemes, channel knowledge is required and beamforming exploits the nature of the channel to either prevent interference or strengthen the signal. In other words, beamforming can vary depending

on the received CSI, whereas space time codes are often fixed or limited to a few different options. One of the main examples of beamforming in literature is zero-forcing. With perfect CSIT, zero-forcing is a method used to precancel the channel effects via channel inversion so that the data received at each user is interference free. The input-output relationship including linear precoding is shown below:

Firstly, we let

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{B}_k \mathbf{u}_k.$$

Assuming there is a total of K users, $\{\mathbf{B}_k\}_{k=1\ldots K}$ are the beamforming matrices for each user, $\mathbf{B}_k \in \mathbb{C}^{N_T \times N_{uk}}$. $\mathbf{u}_k \in \mathbb{C}^{N_{uk} \times 1}$ is a vector containing the N_{uk} data symbols requested by user k. The number of symbols that can be sent to each user is restricted by $N_{uk} \leq N_{Rk}$ i.e. the number of receive antennas. In practical scenarios, each user will only have one or two receive antennas. Note that a receiver with multiple receive antennas need not necessarily request multiple data symbols. For simplicity, it is usually assumed each user is limited to one data symbol per symbol transmission time ($N_{uk} = 1$), in which case $\mathbf{u}_k$ is a scalar (denoted by u_k)and $\mathbf{B}_k$ is a vector (denoted by $\mathbf{b}_k$). The input-output relationship from the base station to each user is then given by:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{B}_k \mathbf{u}_k + \sum_{j \neq k}^{K} \mathbf{H}_k \mathbf{B}_j \mathbf{u}_j + \mathbf{z}_k, \quad k = 1, \ldots, K \tag{2.2}$$

The first term consists of the actual information required by user k, the second term consists of interference from other users and the third is AWGN [28]. Clearly, the beamforming matrices have to be designed to reduce the interference, while at the same time keeping the receive power of the required information (the first term) large enough to reduce the effect of noise. A performance measure that addresses this trade-off is capacity which will be discussed in Section 2.3.

There are also non-linear precoding schemes that achieve the capacity of a MU-MIMO channel with perfect CSIT, most notably the Dirty Paper Coding (DPC)[30]. DPC attempts to successively pre-cancel interference on the transmitter side, by sending the difference between the data symbols and interference instead of the data itself [31]. In many ways DPC on transmitter side is similar to MMSE-SIC on the receiver side, in terms of achieving MIMO capacity. The duality of this relationship is discussed later in the thesis. DPC is in fact practically complex to implement, and construction and decoding of nested lattice structures required in DPC are computationally taxing [27]. DPC is seen in Tomlinson-Harashima precoding, and Costa Precoding.

Channel

Wireless communication channels consist of signals that traverse via multiple paths due to scatters and reflectors in the environment before reaching the receiver. Heavy destructive interference causes the channel to go into deep fade resulting in failure of communication. In a discrete time system the multiple paths reaching the receiver over a time period is combined into a single channel gain (in a point-to-point system) or channel matrix (in a MIMO system) per discrete time instant [12]. Typically, the channel elements can be modelled by Rayleigh fading which is a statistical model suitable for a propagation based environment with many small scatterers and reflectors. Each path is modelled as a circular symmetric random variable and the addition of these paths (by the central limit theorem) constitutes a circular symmetric Gaussian random variable. Each path in the fading channel contains the parameters Doppler Shift and Delay [32]. The maximum Delay Spread and Doppler Shift are given below:

$$D_s = \max_{i,j} f_c|\tau_i^{|}(t) - \tau_j^{|}(t)|,\ T_d = \max_{i,j}|\tau_i(t) - \tau_j(t)|$$

where $f_c\tau_i^{\cdot}(t)$ is the doppler shift of the i^{th} path and $\tau_i(t)$ is the delay of the i^{th} path The coherence time and coherence bandwidth which dictate how quickly the channel varies in time and frequency respectively, can hence be defined as follows:

$$T_c = \frac{1}{4D_s},\ W_c = \frac{1}{2T_d}$$

As a direct result of these parameters there are four possible fading scenarios that could occur, namely, slow fading, fast fading, frequency selective fading and flat fading. Fig.2.4 shows a time-frequency grid where these scenarios lie.

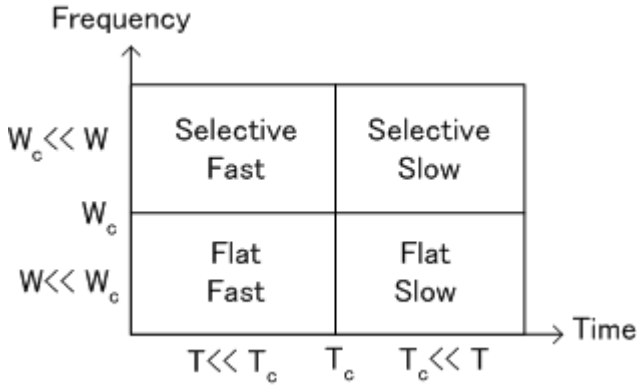

Figure 2.4: Fading Scenarios

T refers to the symbol period and $W = 1/T$. For the purposes of this thesis, we do not consider frequency selective scenarios as they require the use of equalizers to attempt to nullify the ISI and hence complicate the analysis. In the special case of under-spread channels where the delay spread is much smaller than the coherence time, pertaining to a highly time invariant channel, OFDM can be used. In this thesis we assume that the channel is in block fading, in other words the channel is invariant over a number of symbol periods. In a selective fading scenario equation (2.2) can be rewritten as follows:

$$\mathbf{y}_k(n) = \sum_{l=0}^{L-1} \mathbf{H}_k(l,n)\mathbf{x}_k(n-l) + \mathbf{z}_k(n), \quad k = 1, \ldots, K \tag{2.3}$$

Here, L corresponds to the number of channel taps or ISI terms.

Receive Combining

Receive combining, as mentioned earlier, is used to perform operations on the received data stream in order to retrieve the data required by the user with minimal noise and/or inter antenna interference. The case of receivers studied in this thesis are linear receivers.

With multiple receive antennas, each user can request more than one information symbol per channel use. In the case when the data transferred to each receive antenna is different, it is required to de-correlate the data stream into seperate streams, and in the case when the data is repeated at each receive antenna, it is required to combine them into a single stream. Fortunately, the design of the precoding/beamforming matrix addresses this issue and the classical receiver architectures can be used regardless of the situation [29]. The assumptions at this stage are that the receiver has perfect knowledge of CSI, as well as knowledge of the beamforming vectors, which in a feedback scenario can be retrieved from the vector codebook. The two main receiver architectures considered are the MMSE receiver and Zero-forcing receiver. While zero-forcing completely nulls out the interference, it does not optimally scale the noise and hence is not optimal in a low SNR scenario. The MMSE receiver is closer to optimal and is optimal amongst the linear receivers in trading off noise and interference [33]. There exist non-linear receivers such as MMSE-SIC, MMSE with successive interference cancellation which achieve capacity in a SU-MIMO scenario, but are not considered due to complexity of analysis. Other non-linear receivers such as decision-feedback equalizers are not considered as ISI in general is not considered in this thesis. The equations for working out the receive combining vector (matrix in the case of multiple data streams) $\mathbf{V}$ for the receiver are

given below:

$$\mathbf{V}^{ZF} = \left[[\mathbf{HB}]^H [\mathbf{HB}] \right]^{-1} [\mathbf{HB}]^H \tag{2.4}$$

$$\mathbf{V}^{MMSE} = \left[[\mathbf{HB}]^H [\mathbf{HB}] + \sigma^2 \mathbf{I} \right]^{-1} [\mathbf{HB}]^H \tag{2.5}$$

The additional $\sigma^2\mathbf{I}$ term in the MMSE receiver exists to account for noise. The output can then be transformed via matrix multiplication, $\mathbf{y}' = \mathbf{V}\mathbf{y}$ to form the effective received data streams at the receiver.

Feedback

Feedback is commonly used by the system to provide receiver side channel information to the transmitter via the uplink. Recalling that the system in consideration is Frequency Division Duplex (FDD), the downlink and uplink channel are in general not reciprocal [32]. Therefore, the transmitter cannot predict the users' channel by constant observation of the information it receives. For simplicity it is assumed that the feedback channel has no delay and is error free. Furthermore, the processing time of any feedback algorithm discussed will be negligible when calculating or simulating the various performance measures. However, complexity analysis of the various algorithms can be performed to compare their relative speed, albeit trade-offs between complexity and performance can be dependent on many physical factors which are not considered.

There are two main types of feedback, namely, analog and digital. In analog feedback, the receiver attempts to provide the channel information that it has estimated back to the transmitter. In one example [34], the CSI being fed back is multiplied by a spreading signal and the spreading signal for each user will be orthonormal. When the combined CSI information from each user is received, de-spreading can be done to retrieve the individual CSI, along with MMSE receiver to reduce noise. In fact, this process is very similar to beamforming and combining, the difference being that the

spreading strategy will only be effecient when the feedback link is not a fading channel. This is of course just one example of analog feedback under certain specific conditions.

Digital feedback can be split into a few stages depending on its design. In digital feedback, the CSI needs to be quantized depending on the number of bits available for feedback. To be specific there are in fact two parameters needed to be quantized; channel direction (CDI), which is useful for beamforming in order to cancel out interference and channel gain (CGI) which is useful for user selection [5]. In essence, in digital feedback the things that need to be considered are, the information being fed back and the design of the vector quantization codebook, and these differ depending on certain conditions [35]. For example, in the case where the channel is spatially correlated, it may not in fact be optimal to use the same codebook that is proven optimal for the case of an i.i.d. fading channel.

Adding further, the above feedback description focuses on a particular time slot or fading block. If we include time into the problem, it is entirely possible to create an adaptive or recursive feedback method that uses the channel information from previous slots along with the current feedback message and incrementally updates the channel state information at the transmitter (CSIT) [36]. In fact, by doing so feedback schemes dealing with spatio-temporally correlated channels can be possibly designed. Lastly, the major assumptions in this thesis are that the channel estimation is perfect, the feedback link is delay and error free and there is no processing delay in any of the stages. However, some of these assumptions can be relaxed, and the implications of doing so are discussed in later parts of this thesis.

2.3 Performance Measures

There are a few measures of performance mentioned in this thesis, capacity, outage-probability, and bit error rate [37]. While sum-rate/capacity will be

the main measure of performance, certain models presented in literature deal with other performance measures. Therefore this section will cover these as well. In a multi-user scenario capacity refers to a rate region rather than a single value and performance is evaluated via maximizing the sum of data rates of all the users. An alternative performance measure that is used more practically is minimizing the transmission power with respect to the Quality of Service (QoS) requirements of each user; often QoS is a rate request or a minimum receive SINR requirement.

Capacity

In telecommunication the main gauge of performance is speed of data transfer; in the case of a mobile phone this could be the time taken to load a webpage or file transfer or video streaming. In the downlink channel, in a physical layer, the main measure of performance is Capacity, which determines the maximum possible rate that data (bits) can be transferred per symbol time (in a given bandwidth) such that the data received has an arbitrarily small probability of having errors [28].

Capacity takes a finite value if there are power constraints on the data being sent. In the system described thus far there are two types of power constraints that can be used; Total Power Constraint (TPC) can be written as either a long term power constraint $\mathbb{E}[\|\mathbf{x}\|^2] \leq P$, or a instantaneous power constraint $\|\mathbf{x}\|^2 \leq P$; Per-antenna Power Constraint (PPC) which can be written as $\mathbb{E}[x_i] \leq P_i, \quad i = 1, \ldots, N_T$. In practical terms, the latter would be more realistic as TPC would imply that each antenna should have the capability of transmitting full, if not almost full power. In other words, TPC should not exceed the power limitations of each antenna. In PPC, however, each antenna has its own power constraint, and this is not an issue. The downside is that Capacity optimization problems with PPC take

much longer to solve, and generally do not have closed form expressions.

Under TPC and full CSI MIMO, capacity can be written as

$$\begin{aligned} C &= \max_{\mathbf{K_x}, Tr(\mathbf{K_x}) \leqslant P} I(\mathbf{x} : \mathbf{y}) \\ &= \max_{\mathbf{K_x}, Tr(\mathbf{K_x}) \leqslant P} \log_2[\det(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{K_x}\mathbf{H}^H)] \end{aligned} \tag{2.6}$$

$$\mathbf{K_x} = \mathbb{E}(\mathbf{x}\mathbf{x}^H),\ Tr(\mathbf{K_x}) = E(\|\mathbf{x}\|^2) \leqslant P$$

and the ergodic capacity can be written as

$$C = \max_{\mathbf{K_x}, Tr(\mathbf{K_x}) \leqslant P} \mathbb{E}_{\mathbf{H}}\left[\log_2[\det(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{K_x}\mathbf{H}^H)]\right] \tag{2.7}$$

With PPC, the water-filling solution will have an additional constraint, making it more complicated [23]. Capacity expressions in multi-user scenarios are shown later in the thesis. As mentioned earlier, in a multi-user scenario we are interested in maximizing the sum-rate with given power constraints or minimizing power with respect to a set of user rates. These are optimization problems that can be solved numerically given the conditions, in certain cases (e.g. zero-forcing beamforming) closed form expressions can be attained.

The definitions of capacity so far assume a certain fixed set of users; often the number of available users are far more than the number of transmit antennas and hence user scheduling as well as fairness schemes need to be implemented in order to optimize the sum-rate [10, 19]. While user selection depends on CSI feedback, fairness depends on the rate region. Multiple points in the rate region are Pareto optimal in maximizing sum-rate. In a high-level system point of view maximizing the sum-rate is sufficient. However, in a fairness point of view, it is required to allocate rates to each user as "equally" as possible so that users with low channel gains can still attain

consistent service [38]. Hence, the fairest Pareto optimal point that achieves maximum sum-rate would have to be computed.

Outage Probability

In the scenario where the transmitter does not have channel state information, ergodic capacity can still be achieved in a single-user case with appropriate receiver architecture in a fast fading scenario. This is done by coding over several time instances to combat fading using diversity. In a slow fading scenario however, without CSIT, the transmitter cannot compensate for the event that the channel is in deep fade [5]. In the event of deep fade, the decoding error probability cannot be made arbitrarily small, and the capacity of the channel is zero in the strict sense. This does not mean that information cannot be transferred. Instead, an alternative measure of performance, outage probability, is used to determine the outage capacity. Outage probability is defined as follows (MIMO).

$$P_{out}^{mimo} = \min_{\mathbf{K_x}, Tr(\mathbf{K_x}) \leqslant P} \left\{ \log_2 \det \left(\mathbf{I} + \frac{1}{N_0} \mathbf{H} \mathbf{K_x} \mathbf{H}^H \right) < R \right\} \tag{2.8}$$

$0 < \epsilon < 1$ can then be defined such that we can send information at rate R so long as $P_{out}^{mimo} < \epsilon$. ϵ-capacity or outage capacity is then the maximum R such that the inequality holds.

Bit Error Probability

Bit Error Probability(BER) is a straightforward performance measure which is often non-trivial to compute in a closed form expression, as error correcting capabilities of different channel codes can vary substantially. BER is the ratio of the number of errors to the number of information bits transferred [33]. Soft information instead of BER can also be used to gauge code performance. However, as the number of simulations grow large BER will eventually present the same outcome. BER is mostly used as a comparison

method to examine the relative performance of various channel codes under different SNR values. Analysis of the performance of the codes are limited to the AWGN channel.

Diversity and Multiplexing Gains

Diversity and multiplexing gains are high SNR performance measures derived from Capacity and error probability expressions outlined above. Spatial Multiplexing defines the number of degrees of freedom that are used in a MIMO system [37]. When the channel is i.i.d. full spatial multiplexing gains can be achieved by sending different data symbols at each transmit antenna. On the other hand, when the data symbols are the same in all the transmit antennas, full diversity gain is achieved [39]. The trade-off between Spatial Multiplexing and Diversity does implicitly come about in this thesis but will not be discussed in detail. The expressions for diversity and multiplexing for a MIMO system is given below:

$$\text{Spatial Multiplexing:} \quad \lim_{SNR\to\infty} \frac{R(SNR)}{\log(SNR)} = r$$

$$\text{Spatial Diversity:} \quad \lim_{SNR\to\infty} \frac{\log P_e(SNR)}{\log(SNR)} = -d \tag{2.9}$$

$$\text{User Diversity:} \quad \lim_{SNR\to\infty} \frac{R(SNR, K)}{r \log\log K} = m$$

User diversity comes from the ability to schedule users when a sufficiently large number of users are present. When the number of users in the system (assuming single receive antennas) equals the number of transmitting antennas, full multiplexing gain can be achieved. However, when there is a large number of users in the system, the transmitter has the choice of being able to select which users to schedule. Due to this choice, the transmitter can intelligently select users that complement each other (for example by being

orthogonal to one another) based on the channel feedback thus increasing the sum-rate when compared to a scenario where there is a small number of users and the scheduling decision is limited.

Chapter 3

Multi-User Transmission Schemes

This chapter begins by briefly introducing the single-user (or point-to-point) MIMO transmitter and receiver designs and their performance. In a practical multi-user system, each user has a limited number of antennas, and therefore full multiplexing and diversity gains cannot be achieved when transmitting to one user alone. Operating with multiple users solves this problem, however, the users cannot communicate with each other and hence receiver strategies used in single-user MIMO do not apply. Transmitter schemes in multi-user scenarios are discussed here, and these aim to reduce the effects of inter-user interference and achieve multiplexing gains. Two broad multi-user transmission strategies are prominent in literature, non-linear and linear precoding [40, 41]. The schemes in each are discussed along with performance comparisons. The complexity of the different schemes will also be shown in an intuitive sense. For the purposes of this chapter full CSIT is assumed unless otherwise stated.

3.1 Single-User MIMO

We start by reviewing the single-user MIMO system. The input-output relationship was defined in [7] and its capacity was given by the following optimization problem:

$$\begin{aligned} C &= \max_{\mathbf{K_x}, Tr(\mathbf{K_x}) \leqslant P} I(\mathbf{x}:\mathbf{y}) \\ &= \max_{\mathbf{K_x}, Tr(\mathbf{K_x}) \leqslant P} \log_2[\det(\mathbf{I} + \frac{1}{N_0}\mathbf{H}\mathbf{K_x}\mathbf{H}^H)] \end{aligned} \tag{3.1}$$

When the channel is known to both transmitter and receiver, a transreceiver architecture can be used to separate the block fading channel into parallel interference-free channels. This is done by taking the SVD of $\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H$. By multiplying a pre-processing matrix, $\mathbf{Q}$, with the data symbols, we can write the transmitted signal as $\mathbf{x} = \mathbf{Q}\mathbf{u}$. The covariance matrix of $\mathbf{x}$ can then be written as $\mathbf{K_x} = \mathbf{Q}\text{diag}(P_1, .., P_{N_T})\mathbf{Q}^H$. In the case when full CSIT is available we can choose $\mathbf{Q} = \mathbf{V}$ and multiply the received signal at the user by $\mathbf{U}^H$, thereby decomposing the received signal into parallel streams. In other words the equivalent channel can be written as

$$\mathbf{y} = \mathbf{\Lambda}\mathbf{u} + \mathbf{z}$$

the capacity expression is then simplified to

$$C = \sum_{i=1}^{N_{min}} \log\left(1 + \frac{P_i^* \lambda_i^2}{N_0}\right) \tag{3.2}$$

where λ_i are the ordered diagonal elements of $\mathbf{\Lambda}$. $N_{min} = \min(N_T, N_R)$ .The optimal power allocation is given by the well-known water-filling problem, where

$$P_i^* = \left(\mu - \frac{N_0}{\lambda_i^2}\right)^+ \tag{3.3}$$

μ is a variable in the Lagrangian form of the original optimization problem. μ or the "water-level" can then be calculated such that $\sum_{i=1}^{N_T} P_i = P^*$. A

simple analysis can then be done to show that at the high SNR regime, equal power allocation is asymptotically optimal, whereas in the low SNR regime, it is optimal to allocate power to the strongest eigenmode. In high SNR, it is important to note that in order to achieve the full spatial degrees of freedom, N_{min}, $\mathbf{H}$ has to be well-conditioned and full rank, while at low SNR, this is irrelevant as the antennas jointly transmit a single stream.

When the transmitter does not know the channel exactly, V-BLAST (Vertical-Bell Laboratories Layered Space-Time) can be used. In V-BLAST, $\mathbf{Q}$ is chosen to multiplex the incoming data symbols onto the transmit antennas. When $\mathbf{Q} = \mathbf{V}$, we achieved the capacity that was just shown above. When $\mathbf{Q} = \mathbf{I}$ (identity matrix), the data symbols are spread across the antennas without any manipulation. The latter scenario is commonly used when the transmitter has no CSI.

While V-BLAST achieves capacity, the above assumes a fast-fading channel, where there is sufficient diversity in the channel across time and scenarios involving deep fade can be avoided by well designed channel codes. In slow fading however, the channel will remain more or less constant over several symbol transmission time-slots. In V-BLAST, the data is transmitted over the antenna array over one symbol time, each stream can traverse to N_R possible receiver antennas. When the data stream from a transmit antenna to all the receive antenna goes through deep fade, that information will be lost. The MIMO channel itself provides a diversity of $N_T \times N_R$ (all possible transmitter receiver antenna combinations). The sub-optimality of V-BLAST in slow fading scenarios is because there is no coding across streams. An alternative method for outage-efficiency in slow fading was given as D-BLAST, where the coding is done diagonally over the space-time grid. Because of this parts of the codeword are scattered across different antennas over time, hence achieving the full diversity gain. (e.g. the data intended for the first receiver antenna is sent from the first transmit antenna

on the first time slot, the second transmit antenna on the second time slot and so on. In V-BLAST the data to the first receiver is always sent from the first transmit antenna, so if that particular channel is in deep fade, all the data to receiver 1 will be lost).

3.2 Multi-User MIMO

For a Multi-User MIMO scenario, we assume that fading occurs in "blocks" and the channel at each block is independent and i.i.d. In the single-user case, capacity is a scalar. In a multi-user scenario, the information sent to each user is independent, and because the users are isolated, each user has its own rate at which it receives information. The set of user rates can then be given as:

$$\mathbf{R}(\pi, \mathbf{K}_{\mathbf{x}1}, ..., \mathbf{K}_{\mathbf{x}K}) = [R_{\pi(1)}, ...R_{\pi(K)}]^T \tag{3.4}$$

$$R_{\pi(k)} = \log_2 \frac{\left|\mathbf{I} + \mathbf{H}_{\pi(k)}\left(\sum_{j\geq k} \mathbf{K}_{\mathbf{x}\pi(j)}\right)\mathbf{H}_{\pi(k)}^H\right|}{\left|\mathbf{I} + \mathbf{H}_{\pi(k)}\left(\sum_{j>k} \mathbf{K}_{\mathbf{x}\pi(j)}\right)\mathbf{H}_{\pi(k)}^H\right|}, k = 1, ..., K \tag{3.5}$$

where $\mathbf{K}_{\mathbf{x}k}$ is the covariance matrix of transmitted signal intended for user k. π is the user permutation or ordering, and $\sum_{k=1}^{K} Tr(\mathbf{K}_{\mathbf{x}k}) \leq P$ is the total power constraint. The capacity of a MU-MIMO system is then defined by its rate region. The rate region is the convex hull of the union of all the rate vectors over all possible permutations and covariance matrices. This can be written as:

$$C = Co\left(\bigcup_{\pi, \mathbf{K}_{\mathbf{x}k}\succeq 0, \sum_{k=1}^{K} Tr(\mathbf{K}_{\mathbf{x}k})\leq P} \mathbf{R}(\pi, \mathbf{K}_{\mathbf{x}1}, ..., \mathbf{K}_{\mathbf{x}K})\right) \tag{3.6}$$

$Co(.)$ is the convex hull. Clearly, comparing the performance of two schemes via the use of the rate vectors can be difficult. As such the performance is usually characterized as a simpler form via the sum-rate. The sum-rate is

maximized amongst all the possible rate vectors in the capacity region.

$$C_{SUM} = \max_{\mathbf{R} \in C} \sum_{k=1}^{K} R_k \tag{3.7}$$

In fact, for a given user ordering, the condition for maximization can be simplified to the maximization problem over all possible covariance matrices. This is the main performance measure used in most MU-MIMO block fading channels. For simplicity purposes most of the analysis in MU-MIMO systems is done assuming single antennas at the receiver $N_R^{(k)} = 1 \ \forall \ k$. The issue with maximizing the sum-rate is that while performance is maximized from a system point of view, fairness issues are not addressed. Typically, if a user has a particularly weak channel, it will be neglected for a long period of time. While fairness issues are not considered here, they provide an important scope for research especially when dealing with multi-cell scenarios.

Unlike in single-user MIMO, transreceiver designs cannot be employed here, as the users are isolated. While they have knowledge of $\mathbf{h}_k$, their own channel, they do not have knowledge of $\mathbf{H}$, the channel of the entire system, which is required for receive combining. Hence, interference cancellation has to be done at the transmitter. Therefore, perfect CSIT is essential in achieving MIMO capacity. In the next few sections non-linear and linear precoding schemes that aim to maximize the sum-rate of a MU-MIMO channel will be discussed.

3.3 Dirty Paper Coding

Dirty Paper Coding (DPC) is a non-linear precoding method that received a lot of attention in multi-user communications from Costa's paper on "Writing on Dirty Paper" [30], notably due to its ability to achieve the capacity region that maximizes the sum-rate [21, 20]. In DPC, the input signal is encoded such that the additive Gaussian interference known to the transmitter (non-causally) is pre-subtracted without increasing the transmit power. In a

multi-user scenario, DPC can be thought of a successive encoding method. Given an ordered set of users, the base station can treat the information requested by the users that come after a particular user in the set as interference. Since the base station knows the channel, it can precode the data symbol to remove this interference.

In an information theoretic perspective, let us suppose we have a point-to-point channel, with input X, interference $S \sim \mathcal{N}(0, K_S)$ (known at the transmitter non-causally), noise $Z \sim \mathcal{N}(0, N)$ and output, Y. K_S and N are arbitrary variances that are known. In the original works, dirty paper coding achieves AWGN capacity if S and Z are Gaussian, with X generated using a Gaussian codebook.

$$Y = X + S + Z$$

$$C = \sup_{p_{x,u,s}} \Big\{ I(U;Y) - I(U;S) \Big\} \tag{3.8}$$

where $x = f(u, s)$ is a deterministic function and $\mathbb{E}(X^2) \leq P$. The capacity is then given by $\frac{1}{2}\log_2(1 + P/N)$ if $P_{U|S} \sim \mathcal{N}(\alpha S, P)$. This is done by choosing $x = u - \alpha s$ and $\alpha = P/(P+N)$. The results have been generalized to arbitrary interference distribution or noise distribution provided one or the other is Gaussian.

Hence, in a multi-user scenario when we successively apply the above information theoretic concepts to design the codewords to pre-cancel interference, our effective received signal at user, for a given user ordering would look like:

$$y_k[m] = (\mathbf{h}_k^H \mathbf{b}_k) u_k[m] + \sum_{j>k}^{r} (\mathbf{h}_k^H \mathbf{b}_j) u_j[m] + z_k[m] \tag{3.9}$$

The above is when there is a single antenna at each user. Depending on the user ordering, the data symbol requested by the first user does not have any interference subtracted, each subsequent user-requested data symbol is coded to subtract the interference from the data-symbols requested by users

before it, and hence only the signal corresponding to the last user will be completely interference free. In other words, $R_k = \log(1 + SINR_k)$.

$$SINR_k = \left(\frac{P_k |\mathbf{h}_k^H \mathbf{b}_k|^2}{N_0 + \sum_{j>k} P_j |\mathbf{h}_k^H \mathbf{b}_k|^2} \right) \tag{3.10}$$

Although not immediately apparent, this is exactly the definition of the rates in the rate vector given in equation 3.5, noting that the expression here is for the single receive antenna case. The remaining problem is then to choose the beamformers (which effectively form the covariance matrix) to maximize the sum-rate.

Hence, dirty paper coding can achieve the capacity region of a MIMO broadcast channel. The downside is that the capacity region of a MIMO BC is a non-convex function of $\{\mathbf{K}_{\mathbf{x}k}\}$, which depend on the beamforming vectors and power allocation strategy. An alternative way to optimize the sum-rate according to the covariance matrices is to utilize duality. The successive decoding in the MAC is similar to DPC in BC. In fact the duality is shown in [42]. In a dual MIMO MAC, the received signal is given by

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k^H \mathbf{x}_k + \mathbf{z} \tag{3.11}$$

The information sent by the users are jumbled and superimposed at the base station. Unlike the broadcast channel however, in the MAC channel the receiver antennas can cooperate and hence with receiver CSI, successive decoding can be performed (e.g. via MMSE-SIC). The duality between BC-MAC states that the capacity regions of the MIMO BC and MIMO MAC are identical, provided that the sum-power constraints are the same. (Note that the individual powers allocated to each beamformer to achieve this capacity are in general different in the MAC and BC). The sum capacity can then be written as [10].

$$C = \max_{\mathbf{K}_{\mathbf{x}k} \succeq 0,\ \sum_{k=1}^{K} Tr(\mathbf{K}_{\mathbf{x}_k}) \leqslant P} \log_2 \left[\det \left(\mathbf{I} + \sum_{k=1}^{K} \mathbf{H} \mathbf{K}_{\mathbf{x}k} \mathbf{H}^H \right) \right] \tag{3.12}$$

An established algorithm [11] to maximize the capacity is the iterative water-filling algorithm which is described in section 3.5. Upon maximizing the capacity, the covariance vectors must also be calculated so that optimal transmission strategies can be employed. This is generally not an easy problem. While the covariance matrices that optimize the capacity can be found for the MAC channel, converting these covariance matrices for the BC is in general complicated and requires a number of matrix operations which may not be efficient in practical systems. However, once the covariance matrices are computed, the beamformers and power allocation can be extracted from them. The problem stated thus far is for a given set of users. Things become even more complicated when there is a large number of users and scheduling is required to utilize the multi-user diversity. In order to optimize the sum-rate when there is a large number of users, the above procedure must be calculated for every permutation of users.

3.4 Transmit Beamforming

While non-linear precoding schemes such as DPC are of interest due to their ability to pre cancel interference and achieve the capacity region of a MIMO BC, they are high in complexity and require perfect CSI at the transmitter [43]. When coupled with finite-rate feedback and user scheduling, the above schemes are tough to analyse or implement in practice. To this end, linear precoding (beamforming) methods are shown.

Zero-forcing

Assuming single antenna receivers, we can write the input-output relationship as follows:

$$y_k[m] = (\mathbf{h}_k^H \mathbf{b}_k)u_k + \sum_{j \neq k}^{r} (\mathbf{h}_k^H \mathbf{b}_j)u_j + z_k \tag{3.13}$$

Zero-forcing beamforming (ZFBF) is performed by selecting beamforming vectors such that the interference from the other users is nulled [5]. In other words, $\mathbf{h}_k^H \mathbf{b}_j = 0$. This is done by choosing the beamformers $\mathbf{b}_k$ in the null space of span$\{\mathbf{h}_j | 1 \leq j \leq K, j \neq k\}$.

A trivial choice is to choose the beamformers from the columns of the pseudo inverse of the channel matrix. The channel matrix is defined as $\mathbf{H} = \left[\mathbf{h}_1^T, \mathbf{h}_2^T, \ldots, \mathbf{h}_K^T\right]^T$. And the pseudo inverse is given by:

$$\mathbf{H}^\dagger = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1} \tag{3.14}$$

Let the beamforming matrix be $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_K]$. We can then choose:

$$\mathbf{B} = \mathbf{H}^\dagger \text{diag}\{P_1, \ldots, P_K\} \tag{3.15}$$

where $P_1, \ldots, P_K$ are the power allocated to each user. the effective received signals are then given by

$$y_k = P_k u_k + z_k \tag{3.16}$$

Hence, the rate maximization problem for zero-forcing beamforming can be written as:

$$\begin{aligned} R_{ZF} = \max_{\mathbf{P}=[P_1,\ldots,P_K]} \quad & \sum_{k=1}^{K} \log(1 + P_k) \\ \text{subject to:} \quad Tr(\mathbf{B}^H \mathbf{B}) = \sum_k P_k & \left[\left(\mathbf{H}\mathbf{H}^H\right)^{-1}\right]_{k,k} \leq P \end{aligned} \tag{3.17}$$

By water-filling [7], $P_k^* = \left(\frac{1}{\lambda\left[(\mathbf{H}\mathbf{H}^H)^{-1}\right]_{k,k}} - 1\right)^+$
and $\sum_k \left(\frac{1}{\lambda} - \left[\left(\mathbf{H}\mathbf{H}^H\right)^{-1}\right]_{k,k}\right)^+ = P$.

The implicit assumptions are that the channel vectors are not correlated, in other words full rank and that the dimensions are limited by $K < N_T$. The first condition is fulfilled by the choice of i.i.d. Rayleigh fading channel. As for the second, when there are multiple users in the system, user scheduling

has to be performed to limit the number of users to the number of transmit antennas. In this section, a pre-selected set of users has been assumed. Typically the optimization problem would extend over all possible subsets of the users.

Regularization

While zero-forcing solely removes interference, a regularization parameter can be added to the channel inversion to improve performance in scenarios where the channels of the users are noise-limited or ill-conditioned [40, 33]. The regularized channel inverse is given by:

$$\mathbf{H}^{\mathbf{H}}(\mathbf{H}\mathbf{H}^{H} + \alpha\mathbf{I}_K)^{-1} \tag{3.18}$$

α is a parameter that determines the ratio of noise/interference cancellation. When α is zero we have perfect interference cancellation. As α grows large, more noise is cancelled.

Zero-forcing Dirty Paper Coding

Recall that there are two stages of complexity in DPC, firstly, calculating the sum-rate and the corresponding covariance matrices that maximize it for a given user-set; secondly, the issue of implementing DPC in practice. Having introduced zero-forcing, we can give a sub-optimal solution to the first problem by choosing the beamformers (and hence covariance matrices) to null the interference while assuming DPC is perfectly implemented.

Since DPC pre-cancels interference from the users that come before the current user [27, 28], we only require the beamformers to cancel the interference from the users that are ordered after the current user. In other words, we require beamforming to convert the effective channel to a lower triangular matrix.

$\mathbf{H}$ can be decomposed as $\mathbf{H} = \mathbf{G}\mathbf{Q}$ via the Q-R decomposition method using Gram-Schmidt orthogonalization to the rows of $\mathbf{H}$. Let $m = \text{rank}(\mathbf{H})$,

then $\mathbf{G} \in \mathbb{C}^{K \times m}$ is lower triangular and $\mathbf{Q} = \mathbb{C}^{m \times N_T}$ is orthonormal in its rows. To be consistent with zero-forcing beamforming, we assume $m = K$ i.e. the channel is of full rank. Now we can choose the beamformers from $\mathbf{Q}$ from the property that $\mathbf{Q}\mathbf{Q}^H = \mathbf{I}$. This leaves the channel to equal the lower triangular matrix $\mathbf{G}$. Choose:

$$\mathbf{B} = \mathbf{Q}^H \text{diag}\{P_1, \ldots, P_K\} \tag{3.19}$$

where $\mathbf{B}$ is a matrix consisting of the beamformers in the columns. After DPC is applied, the effective channel becomes

$$y_k = P_k g_{k,k} u_k + z_k \tag{3.20}$$

$g_{k,k}$ is the k^{th} diagonal element of $\mathbf{G}$. By letting $\hat{P}_k = P_k g_{k,k}$, we can define the sum-rate as

$$\begin{aligned} R_{ZF} = \max_{\mathbf{P}} \quad & \sum_{k=1}^{K} \log(1 + \hat{P}_k) \\ \text{subject to:} \quad & Tr(\mathbf{B}^H \mathbf{B}) = \sum_k \frac{\hat{P}_k}{|g_{k,k}|^2} \leq P \end{aligned} \tag{3.21}$$

By water-filling the optimal power allocation is given by $\hat{P}_k^* = \left(\frac{|g_{k,k}|^2}{\lambda} - 1\right)^+$, where $\sum_k \left(\frac{1}{\lambda} - \frac{1}{|g_{k,k}|^2}\right)^+ = P$

Random Beamforming

The methods so far require perfect CSI to operate well. In the case when the transmitter only has partial or no CSI from the receiver, a random beamforming approach can be used [23]. In random beamforming, the beamforming vectors are chosen as random orthonormal vectors ($\mathbf{B}\mathbf{B}^H = \mathbf{I}$) according to an isotropic distribution. $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_K]$. Here, however, we assume the base station has some CSI namely the preferred beamformer and SINR of each user. A total of K beamformers are chosen where K users are scheduled at a time. We assume there exist a large number of users. Suppose

there are $n >> N_T$ users in the system, the SINR of the i^{th} user for each choice of beamforming vector is given by:

$$SINR_{i,k} = \frac{|\mathbf{h}_i \mathbf{b}_k|^2}{1 + \sum_{j \neq k} |\mathbf{h}_i \mathbf{b}_j|^2} \quad , 1 \leq k \leq K, \ i = 1, ..., n \tag{3.22}$$

The i^{th} user then sends back the index of the beamformer that maximizes this SINR, $k^*(i)$, as well as the maximum SINR given by

$$SINR_i = \max_{1 \leq k \leq K} SINR_{i,k}, \quad i = 1, \ldots, n \tag{3.23}$$

Finally, the transmitter selects the optimal set of users by going through each beamforming index and selecting the user $i^*(k)$ that has the best SINR for the corresponding beamformer.

$$i^*(k) = \arg \max_{i:k^*(i)=k} SINR_i, 1 \leq k \leq K \tag{3.24}$$

Hence, the k^{th} symbol being transmitted at a time slot will be dedicated to user $i^*(k)$. Note that there is a small probability that a single-user can have, the maximum SINR for two or more different beamformers, but this is generally ignored when the number of users in the system is large.

3.5 Beamforming with Multiple-Antenna Receivers

When dealing with multiple-antenna receivers, zero-forcing can still be applied. However, using zero-forcing via channel inversion treats the antennas at each receiver as separate users. This would mean that the total number of users that can be transmitted to at a channel use will be limited. For example, if we have a user that has $N_R = N_T$ then with zero-forcing, if that user is scheduled, no other users can be co-scheduled. However, since the receivers can effectively combine the signals from their own antennas, transmission schemes that utilize this fact can be implemented to achieve better throughputs than with channel inversion. One such scheme that uses an extended form of zero-forcing is "Block Diagonalization" [39, 36].

Block Diagaonalization

Due to multi-antenna receivers considered, the channel matrix is composed of matrices for each user rather than vectors and is given by:

$$\mathbf{H} = [\mathbf{H}_1^H \quad \mathbf{H}_2^H \quad \dots \quad \mathbf{H}_K^H]^H \tag{3.25}$$

And the corresponding beamforming matrix is given by:

$$\mathbf{B} = [\mathbf{B}_1 \quad \mathbf{B}_2 \quad \dots \quad \mathbf{B}_K] \tag{3.26}$$

The idea then is to diagonalize **HB**. This would mean that the channels to each receive antenna would be in parallel streams without interference. The zero interference criterion is then to choose $\mathbf{H}_k\mathbf{B}_j = \mathbf{0} \quad \forall k \neq j$.To do this we require the beamformers, $\mathbf{B}_k$, to lie in the null space spanned by $\tilde{\mathbf{H}}_k = [\mathbf{H}_1^H \quad \dots \quad \mathbf{H}_{k-1}^H \quad \mathbf{H}_{k+1}^H \quad \dots \quad \mathbf{H}_K^H]^H$

We note that prior to any processing the capacity maximization problem for block diagonalization can be written as:

$$C_{BD} = \max_{\mathbf{B}:\ \mathbf{H}_i\mathbf{B}_j=\mathbf{0}, i\neq j} \sum_{j=1}^{K} \log_2 |\mathbf{I} + \frac{1}{N_0}\mathbf{H}_j\mathbf{B}_j\mathbf{B}_j^H\mathbf{H}_j^H| \tag{3.27}$$

We can then decompose $\tilde{\mathbf{H}}_k$ as:

$$\tilde{\mathbf{H}}_k = \tilde{\mathbf{U}}_k\tilde{\mathbf{\Sigma}}_k[\tilde{\mathbf{V}}_k^{(1)}\tilde{\mathbf{V}}_k^{(0)}]^H \tag{3.28}$$

The right singular vectors can be split into two parts, $\tilde{\mathbf{V}}_k^{(1)}$ which is in the subspace spanned by $\tilde{\mathbf{H}}_k$ and $\tilde{\mathbf{V}}_k^{(0)}$ which is in its null space. For the second term to exist, it is required that $\text{rank}(\tilde{\mathbf{H}}_k) < N_T$ or equivalently the nullspace of $\tilde{\mathbf{H}}_k$ must have at least dimension one. Since $\tilde{\mathbf{V}}_k^{(0)}$ corresponds to the zero singular values and forms an orthonormal basis of the nullspace, it can be a candidate when choosing $\mathbf{B}_k$. The effective channel for user k then becomes

$\mathbf{H}_k\mathbf{V}_k^{(0)}$. In order for feasible transmission however, we require at least one row of $\mathbf{H}_k$ is independent of the rows of $\tilde{\mathbf{H}}_k$. Note that the equivalent channel for each user does not contain any inter-user interference and hence is equivalent to SU-MIMO that we discussed at the start of the chapter. As such, we can decompose (SVD) the effective channel as follows:

$$\tilde{\mathbf{H}}_k\tilde{\mathbf{V}}_k^{(0)} = \mathbf{U}_k \begin{bmatrix} \mathbf{\Sigma}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{V}_k^{(1)}\mathbf{V}_k^{(0)}]^H \tag{3.29}$$

Recall in SU-MIMO, for channel known at the transmitter, the precoding matrix was chosen as the right matrix of the SVD. As before, we can decompose the right singular vectors into two parts. Ideally we want to transmit the symbols along the vectors that correspond to the non-zero singular values. Therefore, the zero interference beamforming matrices can be chosen as $\mathbf{B}_k = \tilde{\mathbf{V}}_k^{(0)}\mathbf{V}_k^{(1)}$ and the capacity can be calculated via water-filling.

Sub-Optimal Strategies

A simpler and more intuitive way is to decompose $\mathbf{H}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k$ directly and multiply the received signal $\mathbf{y}_k$ by the receive filter, $\mathbf{U}_k^H$. This effectively splits the antennas at each receiver into independent streams that can be treated as separate "users". The system can then be viewed as a MIMO BC with $\sum_k N_{Rk}$ users and zero-forcing can then be applied via channel inversion.

A second way is to exploit the receiver architectures presented in chapter 2. Assuming that each user only requests one data symbol per channel use, the receive combiners will combine the signals from the receiver antennas into one stream. The transmitter can then form its beamformers based on the largest right singular value of the decomposed channel matrix of each user.

Both the above methods are suboptimal when compared to Block Diagonalization. The first is suboptimal as it treats each antenna as a user via appropriate receiver filtering. Due to this, we limit the number of effective users we can transmit to, and in scenarios where the channels are in deep fade we cannot utilize user diversity. The second is suboptimal as it only exploits information from the largest right singular value of the user's channel, in other words the "best" antenna of each user. As such when using zero-forcing beamforming, the dominant interference from other users is reduced but not completely eliminated. The advantage of the above strategies when compared to Block Diagonalization is that they take less computation time and are only required to calculate the SVD of each user channel once. In Block Diagaonlization the SVD has to be done twice along with additional matrix multiplications.

3.6 Performance and Complexity

Performance

Comparison of performance in terms of sum-rate and outage probability of the above given schemes have been shown via simulations in literature. Most of which arrive to the same conclusion. Here, we will provide an intuitive interpretation of the relative performance and discuss the basic results. The sum-capacity of a MIMO BC varies depending on SNR. At low SNR, the performance is noise limited, and at high SNR it is interference limited. The intuition behind this is that at the low SNR region both the signal and interference power are low compared to noise power, which is assumed to have fixed variance. On the other hand, at high SNR, the interference and signal power are large compared to noise, and interference becomes a bigger issue.

The most important result of the performance of MIMO BC channel

comes from its duality relationship with the MAC. In the MAC, the sum-capacity increases linearly with $\min(N_T, M)$, M represents the number of scheduled users. Using duality, it can be verified that the same linear growth is achievable for the BC. Assuming that the users are limited to a single antenna, via appropriate user scheduling and precoding the full multiplexing benefit of single-user MIMO can still be achieved, provided $K > N_T$. As discussed, DPC achieves the sum-capacity of the MIMO BC and hence also achieves the full spatial degrees of freedom of MIMO. The downside of DPC is its difficulty in implementation. Beamforming which has a much lower complexity relative to DPC in fact achieves a large portion of the capacity region.

For a large user-set, K, the expected sum-rate of DPC can be approximated as follows [41].

$$\mathbb{E}\{R_{DPC}\} \sim N_T \log\left(1 + \frac{P}{N_T}\log(K)\right) \tag{3.30}$$

In other words DPC achieves the full spatial multiplexing gain, N_T. The multi-user diversity gain is given by $\log K$.

In fact it is shown that $\mathbb{E}(R_{BF}) = \mathbb{E}(R_{DPC})$ in optimal beamforming with a large number of users, where $\mathbb{E}(R_{BF})$ is the expected sum-rate utilizing beamforming. This can be observed from the random beamforming scheme shown in section 3.4. Despite being sub-optimal, the RBF scheme is still able to achieve the MIMO multiplexing gains [41], so an optimal beamforming scheme should do so as well. While RBF achieves performance gains, its performance in practical schemes with a reasonable number of users is low and underwhelming compared to zero-forcing beamforming.

The asmyptotic optimality (as K grows large) of beamforming schemes can also be argued via ZFBF. The intuition is that when K is small, we are limited in the choice of users and in cases where condition number of the

channel is high, channel inversion is difficult to compute. More importantly, this means that some of the channel vectors are close to being aligned, hence zero-forcing would cause a huge loss in receive signal power, as most of the transmit power is used to remove the interference. This would not be the case when the user's channels are orthogonal and the beamformers would be aligned to the channels to increase the receive signal power. As K increases, the probability of finding a set of users with well conditioned channel is higher. In other words, the chances of finding users with orthogonal channels is higher [44, 45]. This improves the received SNR when zero-forcing the interference, which in turn improves sum-rate on average. And as $K \to \infty$ it is always possible to find orthogonal users, which means that zero-forcing simplifies to maximum ratio transmission and the sum-rate gap between DPC and BF goes to 0. Of course smart user scheduling algorithms are required to ensure these gains are achieved for a reasonable K, as we cannot always guarentee to have a very large number of users.

Performance Comparison

Generally,

$$R_{Coop} \geq R_{DPC} \geq R_{ZFDPC} \geq R_{ZF} \geq R_{RBF}$$

Coop stands for full cooperation, where the users know the entire channel $\mathbf{H}$, and is essentially the same as SU-MIMO. DPC, ZFDPC, ZF and RBF correspond to the various precoding schemes discussed in this chapter, where the users only know their own channel $\mathbf{h}_k$.

While there are variety of proofs justifying the above order, this is in fact quite intuitive. While practically not implementable due to size and cost constraints at the receiver, Cooperative MIMO (or SU-MIMO) has the highest sum-rate due to the receivers being able to communicate with one another. As such transreceiver designs can be constructed and are relatively easy to implement. While DPC does not perform as well as SU-MIMO, it

does achieve the capacity region of the MIMO BC channel. The remaining schemes are strictly inferior when compared to DPC [41, 23]. While they provide simplified ways of calculating the beamformers (and covariance matrices) that reduce interference or maximize signal power they are in general not optimal. Amongst these schemes ZFDPC performs better than ZF. In ZF each user interferes with all the other users, hence the beamformers have to be chosen strictly from the null space of the interfering user's channels. In ZFDPC each user only interferes with the users that come before it (for a given ordered set of users), and hence there is more freedom in choosing the beamformers, i.e the beamformers can be chosen from a larger subspace, allowing for more optimal choices. It is easy to see that RBF which only uses partial CSI at the transmitter is outperformed by all the other schemes that utilize full CSI at the transmitter.

Complexity

The complexity of DPC, ZFBF, ZFDPC and RBF is also compared in an intuitive sense. The operations of these methods are fairly common and can be broken down into, encoding, rate, and beamforming calculations.

In DPC the iterative water-filling is used to calculate the optimal covariance matrices for the transmit signals (and hence beamformers) [11]. Here, the capacity maximization problem is solved by using the dual MAC problem. However, the covariance matrices for the BC that achieve this capacity are not trivial to find. As such this iterative algorithm arbitrarily fixes the covariance matrices of the transmit signals of all but one users, and optimizes the capacity over just that one user. This is effectively a SU-MIMO channel which is trivial to optimize. In other words, at each step of the algorithm, the optimization problem simplifies to a equivalent point-to-point MIMO optimization problem where water-filling is done over a modified channel. The resulting covariance matrix that maximizes the mentioned problem is

then updated, and in the next iteration, a different user's covariance matrix is fixed, and the process is repeated. The algorithm is terminated when the maximum error between the covariance matrixes in the current and previous iteration is below a pre-designed threshold. User selection is then performed via repeating the above algorithm over different combinations of users, and by finding the optimal set.

Furthermore, while the above can be computed via algorithms mentioned in [11], complicated methods are required to implement DPC by the means of channel coding. It is shown in [11] that nested lattice structures are used when designing symbol constellations. In short, DPC is computationally taxing and hard to implement, while the beamforming schemes are roughly similar in terms of complexity, except for ZFDPC. While, ZFDPC does not require an iterative water-filling algorithm for its beamformer choice, it still requires the implementation of DPC in a channel coding sense. The comparison of the remaining techniques along with DPC can hence be summarized in the following table:

RBF	DPC	ZF	ZFDPC
Random Vector generation SINR maximization	Iterative water-filling MAC-BC Conversion Lattice Precoding+ cancel interference	Channel Inversion water-filling	Q-R Decomposition Lattice Precoding + cancel interference

Chapter 4

Limited Feedback and User Scheduling

Transmitter knowledge of the channel information in TDD systems can be found by using various estimation schemes. In TDD, the uplink and downlink share the same frequency band and hence the respective channels are reciprocal. In FDD however, this is not the case and it is in general not simple for transmitters to attain CSI in an open loop system. As such, closed loop feedback schemes have been established in literature, where the users estimate their channels and feed this information via the uplink. The feedback link is assumed to have a low rate and negligible error and delay. Two broad ways of providing feedback are via analog and digital feedback. Analog feedback is limited by the noise in the uplink channel and digital feedback is limited by the number of feedback bits that can be used for channel quantization. Upon receiving feedback information, the base station performs its usual operations, beamforming and user scheduling to maximize throughput. The purpose of this chapter is to analyse the various feedback methods in literature and their performance. Most of this chapter focuses on digital feedback, including the design of codebooks. The performance of limited feedback and the role of CQI and CDI feedback in both single

and multi-user systems are also shown. While limited feedback, as we will observe, is able to attain the full multiplexing gain as long as the number of feedback bits per codeword is scaled appropriately with the SNR, the addition of user scheduling greatly improves the performance of the system. To this end, this chapter also provides some current user scheduling algorithms along with performance comparisons between these methods in a limited feedback scenario.

4.1 Limited Feedback for a Single-User

This section talks about the design and optimization of a limited feedback single-user system. Many of the concepts defined in this section in fact can be adapted to the multi-user scenario, which will be discussed in a later section. The design considerations for a user with a single receive antennas is first discussed before moving on to receivers with multiple antennas.

4.1.1 Single Antenna User

We start with looking at the MISO system with N_T antennas at the transmitter and 1 antenna at the receiver. The input-output relationship can be written as:

$$y = \mathbf{h}^H \mathbf{b} u + z \tag{4.1}$$

where the notation holds in consistency with that described in the system overview. Briefly, y is the output signal, $\mathbf{h}$ is the channel vector, $\mathbf{b}$ is the beamforming vector, u is the data symbol to be transmitter and z is complex AWGN. In order to utiize the available feedback bits to provide CSI, the channel directional information (CDI) can be quantized. While codebook design for vector codebooks will be discussed later, for simplicity the codewords are assumed to be randomly generated unit norm vectors. In other words, they are picked randomly from a uniform distribution on the surface

area of an isotropic complex unit sphere. We can then define the set of beamformers as $\mathcal{F} = \{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_{N_C}\}$. Assuming we have B bits of feedback, this would mean that the number of codewords per codebook would be $N_C = 2^B$. Or in the case that $\log_2 N_C$ is not an integer, $B = \lceil \log_2 N_C \rceil$. Assuming that maximal ratio transmission is used, we want to pick the beamformer that is more closely aligned to the channel vector than any other beamformer in the set. We can then define the beamformer amongst the set that minimizes the "chordal distance" with the channel as:

$$\mathbf{b}^* = \text{argmax}_{\mathbf{b}\in\mathcal{F}} |\mathbf{b}^H \mathbf{h}|^2 \tag{4.2}$$

In a single-user scenario, the channel capacity is directly related to the received SNR. As such we would want to maximize the receive SNR with respect to the beamformers in $\mathcal{F}$. Indeed, choosing the codeword that maximizes SNR is equivalent to choosing the codeword that is aligned to the channels (noting that the codewords have equal magnitude). Thus, the SNR conditioned on the knowledge of the channel at the receive side is given by:

$$\gamma = P \max_{\mathbf{b}\in\mathcal{F}} |\mathbf{b}^H \mathbf{h}|^2 \tag{4.3}$$

where, $P = \frac{E_s}{N_0}$ is the transmit SNR of each symbol. Defining $\omega = \max_{\mathbf{b}\in\mathcal{F}} \frac{|\mathbf{b}^{\mathrm{H}}\mathbf{h}|^2}{\|\mathbf{h}\|_2^2}$ we can then see that the SNR is given by the multiplication of two random variables, ω and $\|\mathbf{h}\|$, $\gamma = P\omega\|\mathbf{h}\|_2^2$. Another aspect to note is that $\|\mathbf{h}\|_2^2$ is in fact the sum of the square of normally distributed random variables. Recognizing this, it can hence be shown that $\|\mathbf{h}\|_2^2$ is Gamma distributed given by $\Gamma(N_T, \sigma^2)$, where σ^2 can be treated as the variance of each i.i.d. channel coefficient. As for ω it is first important to recognize that $0 \leq \omega \leq 1$ when ω is 1, this implies that the best codeword in the codebook is completely aligned to the channel, for a given channel. Likewise when ω is close to 0, this implies that the codebook quantizes the channel poorly. For argument's sake, if we let $x = \frac{|\mathbf{b}^{\mathrm{H}}\mathbf{h}|^2}{\|\mathbf{h}\|_2^2}$, noting that $\mathbf{b}$ is a unit vector, we

see that $x = \cos^2\theta$,where θ is the angle between the beamformer and the channel vector.

The distance between any two complex unit norm vectors $\mathbf{w}$ and $\mathbf{v}$ is given by $d(\mathbf{w}, \mathbf{v}) = \sqrt{1-x}$, $x \in [0, 1]$. If we define a space comprising of all possible unit norm channel realizations that have a distance measure less than $\sqrt{1-x}$ for a given x and beamformer $\mathbf{b}_i$, the CDF of x in a geometric sense can be found by taking the ratio of the surface area of this space of beamformers over the surface area of the total unit sphere. We can define the above mentioned space as:

$$C_{\mathbf{b}_i}(\sqrt{1-x}) = \{\mathbf{v} \in \Omega_{N_T} : \ d(\mathbf{v}, \mathbf{b}_i) < \sqrt{1-x}\} \tag{4.4}$$

where, Ω_{N_T} is the set of all unit norm vectors in a N_T dimensional complex hypersphere. Letting $\mathcal{A}(.)$ denote the surface area operator, it can be shown that [46]:

$$\frac{\mathcal{A}(C_{\mathbf{b}_i}(\sqrt{1-x}))}{\mathcal{A}(\Omega_{N_T})} = (\sqrt{1-x})^{2(N_T-1)} = (1-x)^{N_T-1} \tag{4.5}$$

The above describes geometrically the probability that two vectors are within $\sqrt{1-x}$ distance apart. The CDF of x can then be defined as [47]:

$$F_x(x) = 1 - (1-x)^{N_T-1} \tag{4.6}$$

Next, taking the maximum of x generated by each beamformer in the codebook and noting that the beamformers are equally distributed, CDF of ω can be written as:

$$F_\omega(\omega) = (1 - (1-\omega)^{N_t-1})^N \tag{4.7}$$

Now that we have defined the CDFs of ω and the channel, the closed form

expressions for Ergodic SNR and the Capacity can be derived and are written as [47]:

$$\mathbb{E}[\gamma] = \mathbb{E}_{\omega}\left[\omega\right]\mathbb{E}_{\|\mathbf{h}\|}[\|\mathbf{h}\|_2^2] = N_T P - N_T P N \mathcal{B}\left(N, \frac{N_T}{N_T - 1}\right) \tag{4.8}$$

$$\mathbb{E}[\gamma] = \mathbb{E}_{\omega}\left[\omega\right]\mathbb{E}_{\|\mathbf{h}\|}[\|\mathbf{h}\|_2^2] = N_T P - N_T P N \mathcal{B}\left(N, \frac{N_T}{N_T - 1}\right) \tag{4.9}$$

$$\begin{aligned}\mathcal{C}(P) &= \mathbb{E}_{\omega}\left[\mathbb{E}_{\|\mathbf{h}\|}\left[\log_2\left(1 + P\omega\|\mathbf{h}\|_2^2\right)\Big|\omega\right]\right] \\ &= \log_2 e\left(\sum_{k=0}^{N_T-1} E_{k+1}\left(\frac{1}{P}\right)e^{\frac{1}{P}} - \int_0^1 F_{\omega}(\omega)\frac{N_T}{\omega}e^{\frac{1}{P\omega}}E_{m+1}\left(\frac{1}{\rho\omega}\right)d\omega\right)\end{aligned} \tag{4.10}$$

In both SNR and capacity expressions the first terms (e.g. $N_T P$ for SNR) correspond to the case when the transmitter has full CSI knowledge. $E(.)$ refers to the exponential integral and $\mathcal{B}(\alpha, \beta)$ is the beta function. The second term in both expression comes from the error due to quantization; furthermore, the second terms only depend on the transmit power, P. Therefore, the quantization error reduces the gradient of the receive SNR and Capacity as P increases. This hence translates into a constant gap in performance in dB scale. While the performance of the system drops by a constant in the dB scale, the multiplexing gain is still 1. Therefore, we can still achieve the single-user multiplexing gain with a fixed number of feedback bits.

Thus far, we have concerned ourselves with the feedback of the normalized beamformers in a codebook, and it should be noted that in the analysis so far only the ergodic rate was computed. While this approach may be applicable for block fading channels that are fast varying, in block fading channels that are slow varying sending information at the ergodic rate may

cause high error rates at a particular time block due to the probability of the channel being in deep fade. In order to compute the performance and allocate an accurate rate to an user, the transmitter hence has to calculate the outage capacity of the channel, which depends on both channel direction and magnitude. This motivates feedback in such scenarios to consider both Channel Direction Indicator (CDI) and Channel Quality Indicator (CQI).

4.1.2 Multiple Antenna User

In the case of full CSI, we had established that using SVD we can form the corresponding transreceiver architecture to split the MIMO channel into several parallel SISO channels. When limited feedback is available, obviously this cannot be done perfectly, causing some interference between the antennas; and intuition tells that this would lead to performance loss depending on the accuracy of the channel estimation. The main difference between a user with single antenna and one with multiple antennas is the channel. With multiple antennas we have a channel matrix and by observing the structure of a matrix we can intuitively say that designing codebooks of matrices is firstly more complicated and secondly would require a lot of bits to maintain a similar resolution to the single antenna case. This section hence addresses simplified designs for a transreceiver structure under limited feedback.

If we take a generalized transreceiver design, we can write it as:

$$y = \mathbf{v}^H \mathbf{H} \mathbf{b} s + \mathbf{v}^H \mathbf{z} \tag{4.11}$$

This differs slightly from the single antenna case. Note that in the single antenna case the beamformers were chosen based on maximum ratio transmission. Here, we have an extra parameter, $\mathbf{v}$ that is used for "receive combining" which was discussed in Chapter 2. The overall goal in a single-user system, for example maximizing capacity, is equivalent to maximizing

the received SNR. Hence, we have to design $\mathbf{v}$ and $\mathbf{b}$ such that the received SNR is maximized. As mentioned earlier in a full CSI scenario these can be chosen based on performing the SVD of $\mathbf{H}$. For a general scenario we can write this problem as:

$$\max_{\mathbf{v},\mathbf{b}} P|\mathbf{v}^H\mathbf{H}\mathbf{b}| \tag{4.12}$$

where $P = \frac{E_s}{N_0}$ as before. The maximization problem above is too general and difficult to solve, and hence is usually given constraints depending on the implementation issues. These constraints are discussed as follows. Firstly, the above problem can be simplified by assuming a particular type of beamforming or receive combining. In other words, for a given beamformer, $\mathbf{b}$, we can perform a receive combining technique known as Maximum Ratio Combining (MRC) which can be treated as the dual to MRT. Firstly, we want to note that in order to leave the output magnitude unscaled, we can add a constraint $\|\mathbf{v}\| = 1$. Then by noting the following,

$$|\mathbf{v}^H\mathbf{H}\mathbf{b}|^2 \leq \|\mathbf{v}\|_2^2\|\mathbf{H}\mathbf{w}\|_2^2 \tag{4.13}$$

and that $\|\mathbf{v}\| = 1$, we require MRC to have $|\mathbf{v}^H\mathbf{H}\mathbf{b}|^2 = \|\mathbf{H}\mathbf{b}\|_2^2$. An obvious choice for MRC is $\mathbf{v} = \frac{\mathbf{H}\mathbf{b}}{\|\mathbf{H}\mathbf{b}\|_2^2}$. Assuming that we always use maximum ratio combining, the maximization problem then reduces to choosing a beamforming scheme. To be consistent with the single-user case, we can again use MRT. With MRT, the choosing of the transmit beamformer then reduces to the following problem:

$$\mathbf{b}^* = \arg\max_{\mathbf{x}\in\Omega_{N_T}} |\mathbf{x}^H\mathbf{H}^H\mathbf{H}\mathbf{x}| \tag{4.14}$$

The solution to the above equation is not unique, as any transformation of the solution by a multiplication of the form $e^{j\theta}$ will achieve the same maxima.

Recall in the single antenna case in calculating the maximum SNR, we required the distribution of $\|\mathbf{h}\|_2^2$. In a multiple antenna case, we need the distribution of $\mathbf{H}^H\mathbf{H}$ which follows a complex Wishart distribution. A property of Wishart distributed matrices is then that the distribution of $\mathbf{H}^H\mathbf{H}$ is equivalent to the distribution given by $\mathbf{U}\mathbf{\Sigma}\mathbf{U}^H$, where $\mathbf{U}$ is Haar distributed $N_T \times N_T$ unitary matrix, $\mathbf{\Sigma}$ is a diagonal matrix. Following this property, the Wishart distribution is invariant to unitary transformations. Another property is that for a Haar distributed $\mathbf{U}_{N_T \times N_T}$, given a unit norm vector, we have $\mathbf{x} \in \Omega_{N_T}$, $\mathbf{U}\mathbf{x} \in \Omega_{N_T}$.

This means that a solution to the maximization problem in (4.14) is to have $\mathbf{b}^* = \mathbf{U}\mathbf{e}_i$ where $\mathbf{e}_i$ is a vector, $1 \times N_T$ with 1 in the i^{th} component and 0 for every other component. Following from this and the property of Haar distributed matrices, any linear combination of the columns of $\mathbf{U}$ is uniformly distributed in the unit norm complex hypersphere. We can then write the inner terms in the maximization problem as $|\mathbf{x}^H\mathbf{U}\mathbf{\Sigma}\mathbf{U}^H\mathbf{h}|$. Knowing that $\mathbf{U}^H\mathbf{x}$ is uniformly distributed in Ω_{N_T} and that $\mathbf{\Sigma}$ is simply a scaling, it is then easy to see an important property shown in [48]:

"The problem of finding quantized beamformers for MISO systems is the same as that of finding quantized beamformers for MIMO systems".

In other words, the choice of beamformers ONLY depends on the number of transmit antennas and hence, all the analysis done in single antenna systems can be directly applied to multiple antenna systems when the receive combining is fixed as MRC.

What remains to be investigated in the context of limited feedback for single-user systems is the problem of how to determine optimal or close to optimal codebooks. Thus far we have seen that, while random codebooks are observed to perform well in an ergodic sense, in a practical sense the

transmitter would require a fixed codebook. To this end, the theory of Grassmanian line packing and using the fact that choosing a beamformer is equivalent to choosing a line across the origin in the complex space, and the results of the work in [49], together summarize that the optimal way to design a codebook is to maximize the minimum distance between any two codewords. In other words, given a distance measure,

$$\delta(\mathbf{B}) = \min_{1 \leq k < l \leq N} \sqrt{1 - |\mathbf{b}_k^H \mathbf{b}_l|^2} = \sin(\theta_{\min}) \tag{4.15}$$

the optimal codebook can be chosen by solving the following problem:

$$\mathbf{B} = \arg \max_{\mathbf{X} \in \Omega_{N_T}^N} \delta(\mathbf{X}) \tag{4.16}$$

4.2 Limited Feedback for Multiple Users

In a multi-user scenario, the base station requires feedback from each user. Recalling that the users cannot directly cooperate we need to consider two things. Firstly, interference cancellation must be done completely in the transmitter, and hence in a sense we could say that feedback becomes all the more critical in achieving the full-CSI performance gains. Secondly, a new dimension to the problem is added namely, the choice of users we are transmitting to, i.e user scheduling. While the idea of user scheduling is not unique to limited feedback, it is important when it comes to analysis of the system when the number of users is greater than the number of transmit antennas $K > N_T$, since ZFBF based transmission schemes cannot accommodate all the users. Furthermore, limited feedback not only creates quantization errors due to the limited choice of beamformers, it also makes user scheduling algorithms less optimal. It should be noted that while the differences in terms of performance analysis between multiple and single-user systems exist, the quantization process and the design of codebooks in a multi-user scenario uses much of the theory already discussed in the single-user case.

4.2.1 Performance of MU-MIMO with RVQ

In this scenario, each user quantizes the channel directions and the corresponding codebook index is fed back to the transmitter. The transmitter will contain K different codebooks, one for each user. It is easy to see that if we have the same codebook for each user, there is a possibility that two or more users will feedback the same codeword. When this happens, the interference from these users cannot be resolved by ZFBF.

Although these codebooks are randomly and independently generated, it is shown in [50] that it is sufficient to simply rotate one fixed codebook over K randomly generated unitary rotation matrices, as an equivalent way of generating these codebooks. As with the single-user case, we will now discuss the ergodic performance for limited feedback with random vector codebooks. To briefly reiterate the system model of a multi-user system in a simplified form, we can write down the input-output relationship and the SINR for the i^{th} as follows.

$$y_i = \mathbf{h}_i^H \sum_{j=1}^{K} \mathbf{b}_j s_j + z_i, \quad i = 1, \ldots, K \tag{4.17}$$

$$\gamma_i = \frac{\frac{P}{N_T}|\mathbf{h}_i^H \mathbf{b}_i|^2}{1 + \sum_{i \neq j} \frac{P}{N_T}|\mathbf{h}_i^H \mathbf{b}_j|^2} \tag{4.18}$$

Note that in the above, equal power is distributed for each of the users. Since the capacity is directly dependent on the SINR terms, we would require to analyse the distribution of the variables involved in the equation. Before doing so, recall in the single-user case, we had chosen beamformers according to equation (4.2). In a multi-user scenario we do not in fact use MRT to each user, as the interference from the different users has to be taken account of. As such the zero-forcing technique discussed in chapter 3 will be used to determine the beamformers. For zero-forcing beamforming, the beamformer is dependent on the channel and hence a fixed codebook of beamformers cannot be used. Instead, the codebooks are used to select the

quantized channel directions as opposed to the beamformers. The vectors in the codebook are still isotropically distributed in a unit norm complex sphere and techniques such as Grassmanian codebook designs will still hold. The quantized channel will then be given by:

$$\hat{\mathbf{h}}_i = \text{argmax}_{\mathbf{w}\in\mathcal{F}_i} |\mathbf{w}^H \mathbf{h}_i|^2 \tag{4.19}$$

where, $\mathcal{F}_i$ is the quantized set of channel vectors for the i^{th} and $\mathbf{w}$ is some vector in this set. Recall further, that in the single-user case we had split the SNR into two variables namely the magnitude of the channel, and the dot product of the normalized channel and the optimal beamformer. We can apply the same split in the multi-user case and rewrite the SINR received at each user as:

$$\gamma_i = \frac{\frac{P}{N_T}\|\mathbf{h}_i\|^2 |\tilde{\mathbf{h}}_i^H \mathbf{b}_i|^2}{1 + \sum_{i\neq j} \frac{P}{N_T}\|\mathbf{h}_i\|^2 |\tilde{\mathbf{h}}_i^H \mathbf{b}_j|^2} \tag{4.20}$$

The variables we require the distributions in calculating the ergodic capacity are namely, $\|\mathbf{h}_i\|^2$, $|\tilde{\mathbf{h}}_i^H \mathbf{b}_i|^2$, and $|\tilde{\mathbf{h}}_i^H \mathbf{b}_j|^2$. The distribution of the first two variables have been established in section 4.1, where the channel magnitude follows a Gamma distribution, and the second variable follows a Beta distribution with expectation given by $N\mathcal{B}(N, \frac{N_T}{N_T-1})$. As for $|\tilde{\mathbf{h}}_i^H \mathbf{b}_j|^2$, the form being similar to the second variable, we would expect it to also follow a Beta distribution. However, since the beamformer is chosen based on the nullspace of the quantized channel values of every other user, we would expect the parameters of the Beta distribution to be different. The following arguments lead to the distribution of the third variable, which we will term as the interference due to quantization error.

Recall that the quantized version of the channel vector, $\mathbf{h}_i$ is denoted $\hat{\mathbf{h}}_i$. The beamformers in zero-forcing, are chosen via channel inversion. In effect channel inversion asserts the condition that the beamformer for user i are chosen

in the nullspace of the channel vectors of the other users: $\mathbf{b}_i \in \text{Null}(\{\hat{\mathbf{h}}_j\}_{j\neq i})$. Keeping this in mind, we can resolve $\tilde{\mathbf{h}}_j$ (the normalized channel vector) as:

$$\tilde{\mathbf{h}}_j = \alpha_1 \hat{\mathbf{h}}_j + \alpha_2 \mathbf{s}_j \tag{4.21}$$

where $\mathbf{s}_j$ is orthogonal to the quantized channel direction and α_1 and α_2 are scaling factors such that the terms $\alpha_1\hat{\mathbf{h}}_j$, $\tilde{\mathbf{h}}_j$ and $\mathbf{s}_j$ geometrically form as a right angled triangle. The angle between $\alpha_1\hat{\mathbf{h}}_j$ and $\tilde{\mathbf{h}}_j$ can be denoted by θ.

It is easy to see that $\alpha_1 = |\tilde{\mathbf{h}}_\mathbf{j}^H \hat{\mathbf{h}}_j|$, since both the channel variables are of unit norm. Noting the structure of the right angle triangle, we can work out α_2 from $\alpha_2^2\|\mathbf{s}_j\|^2 = \|\tilde{\mathbf{h}}_j\|^2 - \alpha_1^2\|\hat{\mathbf{h}}_j\|^2$. Noting that $\tilde{\mathbf{h}}_j$ and $\hat{\mathbf{h}}_j$ are unit norm and using the expression for α_1, we can write $\alpha_2 = \sqrt{1 - |\tilde{\mathbf{h}}_\mathbf{j}^H \hat{\mathbf{h}}_j|^2} = \sin\theta$. Now if we consider $|\tilde{\mathbf{h}}_j^H \mathbf{b}_i|^2, \quad i \neq j$, and substitute the expression for $\tilde{\mathbf{h}}_j$, noting that $\mathbf{b}_i$ is chosen to be orthogonal to $\hat{\mathbf{h}}_j$, we have:

$$|\tilde{\mathbf{h}}_j^H \mathbf{b}_i|^2 = \sin^2\theta \; |\mathbf{s}_j^H \mathbf{b}_i|^2 \tag{4.22}$$

The distribution of $\sin\theta$ is the same as that of the quantization error and is given by the minimum of N, $\mathcal{B}(N_T - 1, 1)$ distributed random variables as in section 4.1. Since $\mathbf{s}_j$, $\mathbf{b}_j$ are unit norm vectors in the nullspace of $\hat{\mathbf{h}}_j$ of dimension $(N_T - 1)$, $|\mathbf{s}_j^H \mathbf{b}_j|^2$ is Beta distributed according to: $\mathcal{B}(1, N_T - 2)$.

Now that the distributions of the various variables within the SINR expressions are established, these can then be used to generate upper bounds on the ergodic capacity of the limited feedback system. The performance measures that are interesting to look at are the bounds on the ergodic capacity, the rate difference between full CSI and limited CSI and the bit allocation law required to maintain the multiplexing gains. While the details of the proofs can be found in [5] the outcomes will be stated in the

following set of equations. An important set of bounds used on most of the proofs are the set of bounds on the quantized error, which are given as follows:

A simple property to note is that Beta random variables are limited to $[0, 1]$, which means that the quantization error variables can be upper bounded by:

$$|\mathbf{h}_j^H \mathbf{b}_i|^2 \leq \sin^2(\theta_j), \quad \forall j \neq i \tag{4.23}$$

Using this, the expectation of the log upper bound of the quantization error can be bounded by:

$$\frac{B + \log_2 e}{N_T - 1} \leq E_{F,H}\left[\log_2\left(\sin^2(\angle\hat{\mathbf{h}}_j, \hat{\mathbf{h}}_j)\right)\right] \leq \frac{B}{N_T - 1} \tag{4.24}$$

where the expectation is taken over all the possible channel realizations and codebook sets. B is the number of feedback bits allocated to each user (assumed to be the same for all users), and $2^B = N_C$ codewords exist in each random vector codebook. The first performance measure we can analyse is in fact the performance drop in terms of throughput due to the above quantization error, and this is given by:

$$\triangle R(P) \triangleq \frac{1}{N_T}[R_{CSI}(P) - R_{LF}(P)] \tag{4.25}$$

where, $R_{CSI}(P)$ is the throughput from zero-forcing with full CSI and R_{LF} is from zero-forcing with limited feedback (random vector codebooks). Using the properties of expectation, Jensen's equality and the bounds shown above, it can be shown that the average expected throughput loss due to quantization is upper bounded as follows:

$$\triangle R(P) < \log_2\left(1 + 2^{-\frac{B}{N_T - 1}} P\right) \tag{4.26}$$

What we can say from this is that the quantized error is a function the SNR and in order to keep the upper bound on the error constant, we

would need to manipulate the feedback bits, B according to SNR. The exact relationship will be shown in discussions that follow. For now the second performance measure we are interested in is the upper bound on the ergodic capacity of the limited feedback scheme. This can be shown to be [5]:

$$R_{LF}(P) \leq N_T \left(1 + \frac{B + \log_2 e}{N_T - 1} + \log_2(N_T - 2) + \log_2 e \right) \tag{4.27}$$

Here we see something interesting in that the upper bound is in fact not dependent on the SNR and for a fixed number of transmit antennas and feedback bits we see that the capacity in fact asymptotically approaches this upper bound as the SNR increases. This reveals an important point that unlike in single-user MIMO scenarios, the multiplexing gain is in fact not achieved for high SNR in the case of limited feedback, provided the number of feedback bits are fixed. In other words the performance saturates. Thus, we can see the upper bounds on both $\triangle R$ and R_{LF} come to the same conclusion that the number of feedback bits has to be a function of SNR in order for full multiplexing gain to be achieved. To this end, we can simply compute this relationship between feedback bits and SNR by forcing the upper bound on $\triangle R$ to be a constant.

By setting $\log_2 \left(1 + 2^{-\frac{B}{N_T - 1}} P \right) = \log_2 b$, where $\log_2 b$ is some maximum possible rate gap we would have for any SNR value, the expression for feedback bits required to maintain the multiplexing gain can then be written as:

$$B = (N_T - 1) \log_2 P - (N_T - 1) \log_2(b - 1) \tag{4.28}$$

Setting $\log_2 b$ to 3dB simplifies the above expression to the one in equation (4.29), with P_{dB} being the power in dB scale.

$$B = \frac{N_T - 1}{3} P_{dB} \quad bits/mobile \tag{4.29}$$

From the analysis above, the implications are that for a fixed number of transmitters, the throughput in a limited feedback scenario can indeed achieve the multiplexing gain of a full CSI scenario using zero-forcing and random vector codebooks. However, the number of bits required scales linearly (in dB scale) with SNR. When operating at a highly interference limited SNR regime, the feedback bits required to maintain the multiplexing gain can be impractical.

Note that for the discussion so far, we have assumed that the set of users are completely independent and randomly generated. Furthermore, the number of users in the system is equal to the number of transmit antennas. If we are looking at an urban environment with a large set of users within the cell range of a base station, intuitively we can choose users based on feedback information to be orthogonal or semi-orthogonal to one another. As such, in order to improve the performance of a limited feedback system, user scheduling or selection algorithms must also be considered. These algorithms will be discussed in section 4.3.

4.3 User Scheduling

Recall in Chapter 3 that the capacity maximization problem for a multi-user scenario via dirty paper coding was two-fold. Firstly, given a set of users, the optimal beamforming and power allocation strategy has to be chosen to maximize capacity. Secondly, this process is repeated over all possible sets of users to find the "global" (in the sense of the best user-set) maximum capacity. While the first problem is simplified via ZFBF in limited feedback scenarios, in most of the analysis so far we have assumed a fixed set of users. In this section we address the second problem by discussing the recent user selection algorithms, focusing on the Semi-orthogonal User Scheduling (SUS) algorithm.

4.3.1 Semi-Orthogonal User Selection (SUS)

The SUS algorithm can be thought of as a tailor-made algorithm for zero-forcing beamforming. The main idea of the algorithm is to choose a set of users that are nearly orthogonal to one another, at the same time having selected user channels that guarantee a reasonably high receive SNR. Since, the user channels are nearly orthogonal to one another, this simplifies the channel inversion procedure used in zero-forcing, by keeping the channel well conditioned. Furthermore, balancing channel magnitude and directional feedback information becomes important in ensuring high receive SINR at the scheduled users. The SUS algorithm based on the material in [41], is then summarized as follows:

Firstly, the initialization parameters are defined

$$\mathcal{T}_i = \{1, \ldots, K\}$$

$$\mathcal{S} = \phi$$

$$i = 1$$

$\mathcal{T}_i$ is the set of users yet to be scheduled, $\mathcal{S}$ is the set of selected users so far and i is the iteration count.

Secondly, Gram-Schmidt orthogonalization is performed. For each user, k that is yet to be scheduled, i.e $k \in \mathcal{T}_i$, the component of its channel vector, $\mathbf{h}_k$, that is orthogonal to the space spanned by the orthogonalized versions of the channels of the users previously scheduled is computed. Here, $\mathbf{g}_k$ is the component of $\mathbf{h}_k$ that is orthogonal to the subspace spanned by $(\mathbf{g}_{(1)}, \ldots \mathbf{g}_{(i-1)})$; the set denotes the orthogonalized channel vectors of the users scheduled so far and the subscript in parentheses, $\mathbf{g}_{(.)}$, denote the

iteration rather than the user index.

$$\begin{aligned}\mathbf{g}_k &= \mathbf{h}_k - \sum_{j=1}^{i-1} \frac{\mathbf{h}_k^H \mathbf{g}_{(j)}}{\|\mathbf{g}_{(j)}\|^2} \mathbf{g}_{(j)} \\ &= \mathbf{h}_k \left(\mathbf{I} - \sum_{j=1}^{i-1} \frac{\mathbf{g}_{(j)}^H \mathbf{g}_{(j)}}{\|\mathbf{g}_{(j)}\|^2} \right) \\ \mathbf{h}_k &= \mathbf{g}_k, \text{when } i = 1\end{aligned}$$

Thirdly, now that $\mathbf{g}_k$ has been computed for all the users, the process of chosing the next user is to simply chose the user with the largest $\|\mathbf{g}\|$. Assuming $\pi(i) = k$ is the user selected at the i^{th} iteration, we can then update the set of scheduled users.

$$\pi(i) = \arg\max_{k \in \mathcal{T}_i} \|\mathbf{g}_k\|$$

$$\mathcal{S} \leftarrow \mathcal{S} \cup \{\pi(i)\}$$

$$\mathbf{h}_{(i)} = \mathbf{h}_{\pi(i)}$$

$$\mathbf{g}_{(i)} = \mathbf{g}_{\pi(i)}$$

The remaining step is to update the set of users yet to be scheduled. Normally, we can leave the algorithm in the previous step until N_T users are selected. However, the concept of semi-orthogonality introduced in this step serves to force some amount of orthogonality between the selected users, by filtering out the remaining users via the following condition.

$$\mathcal{T}_{i+1} = \left\{ k \in \mathcal{T}_i, k \neq \pi(i) | \frac{|\mathbf{h}_k^H \mathbf{g}_{(i)}|}{\|\mathbf{h}_k\| \|\mathbf{g}_{(i)}\|} < \epsilon \right\}$$

$$i \leftarrow i + 1$$

Adaptation to Limited Feedback

While the above scheduling algorithm is based on full CSIT information, a similar scheduling algorithm can be performed based on the CQI and CDI feedback. In the limited feedback case, the first user is chosen based on the maximum CQI parameter that was fed back. The details of this algorithm can be seen in Chapter 5.

4.3.2 Other User Selection Algorithms

SUS can be computationally taxing especially with a large number of users. In this section, we will go through some simplified user selection algorithms.

Channel Magnitude Quantization

As inferred from the title, here only the norm of the channel is used for selection and orthogonality is ignored. As such the user selection is much simpler, as the entire Gram-Schmidt procedure is not required, and neither is the forcing of semi-orthogonality. The algorithm can be summarized as follows:
The algorithm is initialized by selecting the user with the highest channel magnitude as follows:

$$\pi(1) = \arg\max_{k \in \mathcal{T}_1} \|\mathbf{h}_k\|$$

$$S = k$$

The remaining users are selected in a similar process until $K = N_T$ users are selected as follows:

$$\pi(i) = \arg\max_{k \in \mathcal{T}_i} \|\mathbf{h}_k\|$$

$$S \leftarrow S \cup k$$

$$\mathcal{T}_{i+1} \leftarrow \mathcal{T}_i / k$$

Angle based User Selection

The flipside of magnitude based selection is angular selection (which is effectively selection based on user CDI). The initialization is similar to the previous case, i.e. the first user is chosen based on channel magnitude maximization. The subsequent users are then selected based on the following criteria:

$$\pi(i) = \arg\max_{k \in \mathcal{T}_i} \sum_{j \in S} \frac{\theta_{(\mathbf{h}_k, \mathbf{h}_j)}}{|S|}$$

$$S \leftarrow S \cup k$$

$$\mathcal{T}_{i+1} \leftarrow \mathcal{T}_i / k$$

Recall in the SUS algorithm, the Gramm-Schmidt algorithm was used to select the orthogonal set of users, here a more simplistic maximum average angle approach is used.

Random User Selection

Random User Selection (RUS) with RVQ based beamforming would provide similar results to that discussed in section 4.2 and hence is a useful scheduling algorithm for comparison purposes.

$S = \pi(1), ...\pi(K)$, where $\pi(i)$ is chosen randomly from $\mathcal{T}$, the set of all users

None of these algorithms perform notably well, as they do not incorporate both CQI and CDI in their design. With CDI alone, the base station might be choosing semi-orthogonal users and simplifying the computation of the channel inverse. However, in such a scenario, the users being selected may not have good channels and hence the system throughput might suffer. With CQI alone, the base station will select users with the best channels, but this would not overcome the performance saturation experienced in limited feedback if the users are interfering with one another. On the upside,

these algorithms do simplify the computational time for the base-station to schedule the users.

4.4 Analog vs Digital Feedback

While this research focuses on Digital Feedback, this section provides a brief introduction to analog and hybrid feedback methods. The basic result from RVQ limited feedback was that the throughput saturates with SNR, and the number of feedback bits have to be scaled with SNR in order to achieve multiplexing gains. In other words, the design of the codebook depends on the SNR. In practice however, the SNR is only known to lie in a certain range i.e. the base station does not have an accurate estimate. Due to this, performance loss in practical systems will be incurred. Furthermore, when operating at the interference limited (high SNR) region the number of feedback bits required to maintain the multiplexing gain are high. This is also not very practical as it complicates the problem of codebook design and increases the complexity of the design procedure. The motivation for digital feedback in wireless communication is the same as in any other field that incorporates coding - to employ error correction procedures in order to minimize the effect of noise from the feedback channel. Error-free feedback is indeed an assumption that can be relaxed and schemes that take into account of error and delay in the feedback link are being researched. However, it is intuitive to state that error-free feedback holds more value in digital feedback than in analog. Having stated the practical limitations of the limited feedback schemes, this section will discuss ideas shown in [27] and compare digital, analog and hybrid feedback schemes.

4.4.1 Analog Feedback

A popular technique in analog feedback is that users estimate and feedback their channel vector using orthogonal feedback channels. In order to do this, the downlink channel coefficients for user k is multiplied by a spreading signal $\mathbf{s}_k$ known at both the transmitter and receiver. For simplicity the following analysis considers only AWGN feedback channels (as opposed to the usual fading channel). A term useful later on for comparing the digital and analog feedback performance is given by β_{fb}. β_{fb} is known as the bandwidth expansion factor. In order to feedback the N_T channel coefficients, each user requires $\beta_{fb}N_T$ channel uses. The received feedback signal at the base station over $\beta_{fb}N_T$ channel uses from user i can then be written as:

$$\mathbf{y}_i = \sqrt{\beta_{fb}P_{fb}}\sum_{k=1}^{N_T}\mathbf{s}_k h_{i,k} + \mathbf{z}_i$$

where the spreading sequences, $\{\mathbf{s}_k\}_{k=1}^{N_T}$ (Assuming $K = N_T$ users have been scheduled) are orthonormal, and $h_{i,k}$ represents the channel gain from the k^{th} antenna in the base station to the i^{th} user. The noise vector instead of being $N_T \times 1$ in a single channel user is now $\beta_{fb}N_T \times 1$. The base station then performs de-spreading to retrieve the individual streams of channel information. The effective received CSI from base station antenna k to user i can then be written as:

$$r_{i,m} = \sqrt{\beta_{fb}P_{fb}}h_{i,k} + \hat{z}_{i,m}$$

Recall, in the downlink case, in a multi-user MIMO system, an important difference to single-user MIMO is that the receiver antennas cannot cooperate as they are from different users. In the case of the feedback link, the base station antennas can cooperate and hence receiver side methods such as MMSE-SIC are still viable in the estimation of the received information. Due to the spreading sequence, the information received does not

have inter-antenna interference. As such, MMSE estimation is sufficient and can be employed to reduce the effects of noise when retrieving the CSI.

A theoretical interpretation about the comparison of the upper and lower rate bounds using analog and digital feedback for ZFDPC was given in [28]. These results also hold for ZF. For analog feedback, the multiplexing gain is preserved as long as SINR in the feedback link scales as follows:

$$\beta_{fb} P_{fb} \ \sim \ a \left(\frac{P}{N_0} \right)^b$$

For some pre-defined constants, $a, b > 0$, where a multiplexing gain of N_T is achieved if and only if $b \geq 1$. As per usual, without loss of generality, $N_0 = 1$ can be assumed. An important point of note is that in order to compare the results of digital feedback with analog feedback, it is required to relate the feedback bits per user with the feedback SNR and the number of channel uses in the feedback link, β_{fb}. Assuming the availability of an error free code, the maximum rate at which feedback bits can be transmitted is $\log_2(1 + P_{fb})$. A proposed comparison in [28] is that $\beta_{fb} N_T$ channel uses in analog feedback corresponds to $\beta_{fb}(N_T - 1)$ channel uses in digital feedback, assuming no feedback bits are required for CQI. As such, the effective "bits" used in analog feedback is given by $B = \beta_{fb}(N_T - 1) \log_2(1 + P_{fb})$. Furthermore, noting that a multiplexing gain of N_T is achieved if B scales as $(M - 1) \log_2(P/N_0)$. The SNR in analog feedback hence needs to scale with that of digital feedback based on the following condition for comparison to be valid:

$$P_{fb} = \left(\frac{P}{N_0} \right)^{\frac{1}{\beta_{fb}}} - 1$$

The theoretical result hence shows that when $\beta_{fb} > 1$ digital feedback outperforms analog feedback asymptotically, and full multiplexing gain is achieved if and only if $\beta_{fb} < 1$. Having compared analog and digital feedback, there in fact exist methods that accommodate both types of feedback

for attaining CSI at the transmitter. The next section will go through a particular "hybrid" feedback scheme.

4.4.2 Hybrid Feedback

The hybrid algorithm [34] is based on quantizing the CSI given a fixed number of feedback bits. The quantized CSI is then sent through to the transmitter along with a short error correction code. Along with this, the quantization error (calculated at the user) is sent via analog feedback. The steps of the algorithm are then as follows:

Firstly, the channel is quantized using a B bit codebook. While the quantization can be done depending on the scenario, i.e. split between CQI and CDI and choice of CQI, for simplicity we assume only CDI quantization is performed and the criteria as before is to choose $\hat{\mathbf{h}}_k = \arg\max_{\mathbf{w}\in\mathcal{F}} \|\mathbf{w} \,.\, \tilde{\mathbf{h}}_k\|$, with $\mathcal{F}$ being a set of 2^B, N_T dimensional vectors chosen from a random or Grassmanain codebook.

Secondly, a unitary transformation is performed on the channel vectors, to align $\tilde{\mathbf{h}}$ to $\hat{\mathbf{h}}$. Note that this is just an equal phase rotation on each element of the channel and does not affect the channel direction. The point of this algorithm is to find the unitary transformation matrix that minimizes the variance given by:

$$\Phi_k = \arg \min_{\Phi\in\mathcal{U}_{N_T}} \|\tilde{\mathbf{h}}_k\Phi - \hat{\mathbf{h}}_k\|^2$$

where Φ denotes the unitary rotation matrix and $\mathcal{U}_{N_T}$ denotes the set of all $N_T \times N_T$ unitary matrices. The aligned version of the channel can be denoted by $\tilde{\mathbf{h}}_k$. Thirdly, following up from the digital feedback, the analog feedback which is the quantization error can be defined as follows:

$$E_k = \tilde{\mathbf{h}}_k - \hat{\mathbf{h}}_k$$

Fourthly, the digital information (representing the index of the quantized channel) is encoded by a short algebraic code e.g. a Reed Muller code, and the resulting codeword can be denoted by $\mathbf{c} = (c1, \ldots, c2)$. Finally, each user transmits E_k and $\mathbf{c}_k$ via an appropriate split of resources between the digital and analog feedback (e.g. a split of power or bandwidth between the two parts to the feedback). The base station then uses a ML decoder to retrieve the digital information and the analog error is retrieved via a MMSE estimator as was the case in the previous section.

Chapter 5

Novel User Scheduling with Mode Selection

The chapters so far have developed the foundations of the research topic, as well as outlined some recent developments in literature. The downlink physical layer operations have been illustrated along with the baseband communication model and its performance measures. Various transmission designs have been analysed in the full-CSI case for both single and multiple users. Zero-forcing beamforming, which balances simplicity of implementation and performance gains, is implemented in this research. We then proceeded to examine the effects of limited feedback, again for both single and multi-user scenarios. The performance losses due to the limited feedback information was improved by employing user scheduling. Despite this, high SNR multiplexing gains were not achieved.

In order to overcome the performance losses of limited feedback, this chapter considers the concept of mode selection, whereby the base station, depending on various parameters, decides how many users to select for transmission. The notion of the number of users the base station selects is referred to as the "mode" of operation. To this end, this chapter explores the main research area by outlining and analysing the existing schemes in literature.

Following this, the chapter will describe the system design in section 5.2 and outline the main operations at the transmitter and receiver in section 5.3. Section 5.3 onwards, we propose novel improvements and alternatives to the existing schemes and compare the results via Monte Carlo simulations.

5.1 Introduction to Multi-Mode Scheduling

The prime motivation for enabling the Base Station to select the number of users is to improve on the saturation in performance experienced in multi-user transmission schemes, as the transmit power increases. To justify this, we can refer to the single-user transmission scheme discussed in Chapter 4. Even with limited feedback, single-user schemes see continual increase in performance as the SNR grows. The limiting factor in the performance in a multi-user scenario is that due to quantization errors, zero-forcing cannot completely eliminate interference. As the transmit power increases, both the signal and interference power also increase, causing the performance to saturate. An intuitive way to improve on the performance is to switch from multi-user to single-user mode depending on the transmit SNR (i.e. switch from transmitting to N_T users to 1 user). Unfortunately, the choice between single and multi-user modes is not that simple and the following will explain why.

Firstly, the choice of single-user or multi-user essentially depends on the SNR region where the rate of single-user transmission outperforms the sum-rate of multi-user transmission. Because the channel varies in time and so do the selected users, the SNR region where the mode of operation switches would also vary in each channel use. Therefore, the choice of single or multi-mode transmission depends not only on the transmit power, but also the channels of all the users, the quantized channels and the beamformers, all of which make up the received SINR of each user. Not all of these parameters are readily available at each user and hence approximations have to be made

on the rates of single and multi-mode transmission before this information is fed back to the base station.

Secondly, we can also utilize this concept of switching modes of operation when we are transmitting to multiple users. Majority of the scheduling algorithms focus on selecting as many users as possible, limited by the number of transmit antennas. However, it is arguable that depending on channel conditions it may not be optimal to always transmit to N_T users where multi-mode transmission is preferred. Hence, we can expand the problem of choosing between 1 and N_T users to a general problem of "How many users should the base station select?".

5.1.1 Existing Schemes

This subsection will give an overview of some of the more prominent multi-mode scheduling schemes in literature.

Mode Selection with Opportunistic Beamforming

Mode selection is a fairly recent topic of research and one of the initial methods proposed was by [51]. In this work, opportunistic beamforming was used, where both the base station and all the users have a single common orthonormal codebook. Each user chooses the codeword that is most closely aligned to the channel direction, as its beamformer. The scheduled users would have to have unique codewords so as to reduce interference, so only one user, out of all the users that chose the same codeword, can be selected. A further improvements c ation was added, namely the number of codewords equals the number of antennas at the Base Station.

Each user firstly selects its preferred beamformer from the codebook, and estimates its SINR. The added stipulation in [51] that $N_C = N_T$ is used to approximate the SINR for each mode of operation via a linear approximation on the SINR when N_T users (mode N_T) are scheduled. What the authors

had deduced from the SINR approximations was that, it was sufficient to transmit to either 1 user or N_T users. Each user then selects whether it prefers single or multi-user mode ($K = N_T$) and sends this information along with its preferred beamformer, and its SINR to the base station.

The base station uses the feedback information to select either 1 or N_T users to transmit to. In the case of single-user transmission, the base station chooses the user with the highest receive SNR amongst the users that had preferred single-user transmission. For N_T users, the base station chooses the users by selecting the best user (based on SINR feedback) from each group of users pertaining to a codeword. This is repeated for all the codewords, after which N_T users will be selected, one for each beamformer. Depending on the channel conditions and SINR, what we will see is that, on an average, as the SNR increases, the system will transition from multi-user transmission to single-user transmission, hence overcoming the saturation in performance.

The problem with this scheme is that it depends on opportunistic beamforming and $N_T = N_C$. Opportunistic beamforming is not always the best choice for beamformers. Zero-forcing performs better in the interference limited region (mid to high SNR). Furthermore, zero-forcing can perform as well as opportunistic beamforming in lower SNR regions by the use of regularization and water-filling. In order to improve on the scheme, firstly, we need to expand this design to ZFBF. Secondly, with ZFBF and random vector codebooks, each user will now have an independently generated codebook. This will allow for $N_C > N_T$, i.e increase in resolution of the feedback information. Thirdly, to make the scheme holistic, the limitation of transmitting to 1 or N_T users can be lifted by allowing any mode $(1, 2, \ldots, N_T)$ to be chosen by the base station.

Mode Selection with ZFBF

The work in [9] considered the mentioned improvements and focused on advancing multi-mode transmission in order to utilize ZFBF. To this end, a ZFBF-based rate approximation method was derived for each user. Along with ZFBF, [9] incorporated RVQ for its codebooks, and did not limit the feedback to $N_T = N_C$. The increased feedback resolution and RVQ came with the trade-off that the rate approximations no longer have information on the choice of beamformers as the codebooks are random and independently generated at each user. The scheme however took a step back in terms of user scheduling. The base station used round robin scheduling which has essentially the same ergodic sum-rate as random user selection. Furthermore, in order to select the overall mode at the base station, the users do not send back their preferred mode of operation as in the previous scheme. Instead, they feedback the rate approximations of all possible modes of operation. The base station then uses this information and determines the best users and the overall mode to transmit at. The problem however, is that the feedback load is high if we apply the design to a practical scenario, where CQI (the rate approximations) would have to be quantized, or utilize extra resources in the case of analog transmission. [13] took this issue into consideration and developed a closed form expression for the rate approximation that each user can use to determine its preferred mode. This way, each user does not have to feedback multiple rate approximations to the base station. Furthermore, a novel scheduling scheme was designed such that only the selected users would need to feedback CDI information. While [13] outperformed the previously mentioned schemes, further improvements can be made. To this end, this research aims to analyse and improve on the two main aspects of the multi-mode scheduling scheme, namely the mode selection process at the base station and the rate approximation at the user.

5.2 Multi-Mode Scheduling Model

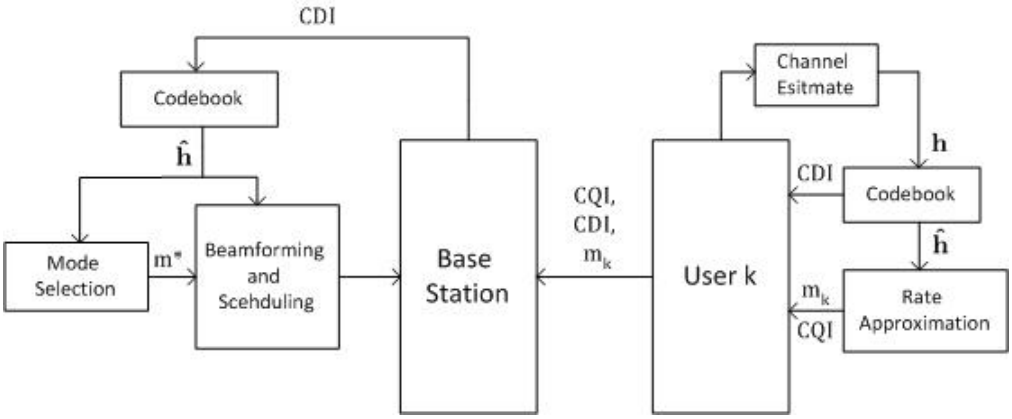

Figure 5.1: Block Diagram of the Multi-Mode Scheduling Process

In Multi-Mode scheduling, each user attempts to approximate the sum-rate by utilizing the information known to it. This is by no means as accurate as the sum-rate calculated at the base station, as the user does not know which users are co-scheduled or their channel coefficients. Each user, however, can approximate the inter user interference by utilizing the channel statistics (in our case we have a Rayleigh channel with AWGN noise) and by knowing the scheduling process at the base station. This can be done for a single user scenario, or when there is one co-scheduled user, two co-scheduled users and so on. In other words, the user approximates the sum-rate for each posisble mode of operation. The reason for doing this is that, the user can then work out which mode of operation (or assumed number of scheduled users) produces the highest approximated sum-rate. This information can then be fed back to the base station, where computations are made to compare the preferred modes provided by each user in order to make an informed decision on how many users to schedule. For example, in the event that interference is a big issue, it is likely that most of the users would prefer lower modes of operations (one or two users). This would hint

to the base station that it should transmit to either one or two users, and by doing so, throughput can be improved when compared to transmitting to several users, since the channel is interference limited. The details of the operations performed at the base station and users are provided in Section 5.3.

Fig. 5.1 illustrates the multi-mode scheduling scheme used in our research. Each user first estimates its own channel and quantizes it using a pre-designed vector codebook. The codebook index that has been selected would be converted to binary and this is the Channel Directional Information (CDI). The codeword vector itself is used for rate approximation, where the user approximates the sum-rate for each possible mode of operation. The user then compares these rates and selects the mode corresponding to the best sum-rate approximation as its preferred mode of operation. Along with the CDI and preferred mode, the user also sends a Channel Quality Indicator (CQI) back to the base station. CQI can typically be channel magnitude or SINR or the Rate approximation itself and the choice of CQI depends on the scenario.

The base station, which has its own copy of each user's codebook, accumulates the feedback information from all the users and retrieves the quantized channel vector using the CDI. The feedback information can then be used for mode selection, where the base station decides the overall mode of operation. The overall mode (m^* along with the feedback information is then used for scheduling and beamforming.

Majority of the processes in the above system model have been discussed and built up in the previous chapters. Omitting the Rate approximation and Mode selection, the above design is simply that of a limited feedback MU-MIMO Channel with user scheduling. As a recap, the following will give the

details of the various stages in a limited feedback ZFBF scheme.

Channel Estimate: Each user approximates the block fading channel vectors at each channel use. Estimation error is not considered in this model and hence we assume the user knows the exact channel coefficients.

Codebook Generation: Each user has randomly and independently generated codebook, $\mathcal{C}_k = \{\mathbf{c}_1, \ldots, \mathbf{c}_{N_C}\}$. Each vector $\mathbf{c}_i$ is chosen from a unit-norm N_T dimensional complex unit sphere.

CDI Quantization is done via choosing the codeword that is most closely aligned to the channel vector. The index of this codebook vector is then fed back to the base station in the uplink channel. Assuming that each user uses B feedback bits for quantization of channel direction, then $N_C = 2^B$ is the total number of codewords per user. The criteria for choosing the quantized channel vector is given by

$$\hat{\mathbf{h}}_k = \arg \max_{\mathbf{c}_i \in \mathcal{C}_k} \frac{|\mathbf{h}_k^H \mathbf{c}_i|^2}{\|\mathbf{h}_k\|_2^2}, \quad i = 1, \ldots, N_C \tag{5.1}$$

The CDI information is the index, i, of the codeword that attains this maximum.

Feedback Channel: The user sends feedback information through the uplink(feedback) channel, which is assumed to have no delay and no error. Furthermore, while CDI is quantized, CQI information and m_k, the preferred mode of a given user k, are assumed to be perfectly known at the base station.

Base Station Operations: The user scheduling process somewhat depends on the mode selection and will be discussed further in the chapter. Beamforming however is done using zero-forcing on the quantized channel

directions. At the transmitter, the set of unit norm beamforming vectors, $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_K]$ can be evaluated as follows.

$$[\hat{\mathbf{b}}_1, \ldots, \hat{\mathbf{b}}_K] = \hat{\mathbf{H}}(\hat{\mathbf{H}}^H\hat{\mathbf{H}})^{-1}$$

$$\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_K] = \left[\frac{\hat{\mathbf{b}}_1}{\|\hat{\mathbf{b}}_1\|_2}, \ldots, \frac{\hat{\mathbf{b}}_K}{\|\hat{\mathbf{b}}_K\|_2}\right] \tag{5.2}$$

Where, $\hat{\mathbf{H}}$ is the matrix consisting of the quantized channel vectors of the scheduled users in its columns.

Performance Measure: For a given set of scheduled users, the performance of the system can then be evaluated via its sum-rate. Assuming the total power available at the transmitter is P and equal power is distributed amongst the scheduled users, the received Signal to Interference plus Noise Ratio (SINR) at user k (with unit noise variance) is given as follows.

$$\gamma_k = \frac{\frac{P}{K}\|\mathbf{h}_k\|_2^2|\tilde{\mathbf{h}}_k^H\mathbf{b}_k|^2}{1 + \frac{P}{K}\sum_{j\in S, j\neq k}\|\mathbf{h}_k\|_2^2|\tilde{\mathbf{h}}_k^H\mathbf{b}_j|^2} \tag{5.3}$$

where $\tilde{\mathbf{h}}_k = \mathbf{h}_k/\|\mathbf{h}_k\|_2$, $\mathcal{S}$ is the set of scheduled users and $K = |\mathcal{S}|$. The achievable sum-rate for a limited-feedback zero-forcing scheme is then:

$$C_{LF-ZFBF} = \sum_{k\in\mathcal{S}} \log_2(1 + \gamma_k) \tag{5.4}$$

This is the performance measure utilized in analysing how well the ideas and designs presented in this chapter fair in comparison to one another and the existing schemes, via Monte Carlo simulations. In addition to the standard design of the limited feedback ZFBF based broadcast channel, we have mode selection/scheduling at the base station and rate approximation at the user. If we look at the performance of the system, there are two things that need to be optimized to achieve the maximum sum-rate for a given set of users; namely the number of users, K, and the set of scheduled users, $\mathcal{S}$. As the number of users grow relatively large, optimizing the sum-rate becomes non-trivial. To this end, we aim to simplify this problem by

compounding scheduling schemes proposed in literature, and by deriving novel rate approximations that each user can utilize to provide information to the base station on its preferred mode as well as its channel quality.

5.3 Multi-Mode User Scheduling

Mode based scheduling schemes in literature are somewhat different compared to classical scheduling schemes. The preferred mode of the k^{th} user, m_k, adds an additional parameter to consider in the scheduling process. The advantage of mode based scheduling algorithms is that they group the users depending on their preferred modes. This is evident in [51] where the base station treats the users that preferred mode 1 and mode N_T separately in the scheduling process, for example if the base station decides to transmit to multiple users, the users corresponding to mode 1 are not even considered in the scheduling. While mode selection schemes primarily aim to improve performance, they also simplify the scheduling process compared to popular scheduling schemes like Semi-orthogonal User Scheduling (SUS), by reducing the feedback load of the system. The user scheduling process we utilize in the preliminary simulations is adapted from [13], which expanded the user grouping process in [51] for the ZFBF based Multi-Mode transmission scheme proposed in [48]. In this section, the Multi-mode User Scheduling process is described, and analysis is performed on the validity of the scheme and assumptions used in [13].

The user scheduling process in [13] has two stages. Firstly, each user approximates its SINR as well as its preferred mode of transmission and sends this information to the base station. As mentioned earlier, we assume that no quantization is performed on the SINR (CQI) and preferred mode, and these are fed back without error and delay. The base station groups the users according to their preferred modes and chooses the mode with the best performance based on the SINR feedback from the users within the

mode. Secondly, supposing mode K has the best performance, the K users that have the highest SINR within the group send their quantized channel directions to the base station for ZFBF. The operations performed at the transmitter and receivers can be summarized in the following subsections.

5.3.1 Operations at the Receiver

Firstly, the user has to approximate its rate. In fact, one of the most important aspects of the given design is the rate approximation, as the entire mode selection procedure depends on the accuracy of this approximation. While sections 5.4 and 5.6 provide the derivation and comparisons of the rate approximations, the following describes the mode selection process. The estimated capacity by each user is a function of P, $\mathbf{h}_k$, $\hat{\mathbf{h}}_k$ and K i.e. the parameters known at the user. The unknown parameters in γ_k have to be approximated. Since $\mathbf{h}_k$ and $\hat{\mathbf{h}}_k$ are common in every mode, the preferred mode of user k from [13] is written as:

$$m_k = \arg \max_{1 \leq K \leq N_T} K\tilde{C}_k(P/K) \tag{5.5}$$

$\tilde{C}_k(P/K)$ is the approximate rate of the user k at mode K. Since user k has no means to retrieve the channel information or SINR of the other users in the system, [13] has adopted a scheme where user k assumes that the other $K-1$ users have the same approximate rate. In an ergodic sense it can be shown via simulations that the above approximation is almost as accurate as the actual ergodic capacity when selecting K users randomly. However, it is intuitive that the above approximation can be quite far from the actual sum-rate at a particular channel instantiation. It is also intuitive to argue that instead of replicating its own rate K times to estimate the sum-rate, a scheme that uses an expected rate for the interfering users would perform at least as well as the above criterion. Hence, we propose the following mode

choice for user k:

$$\grave{C}_k(P/K) = \begin{cases} \tilde{C}_k(P) & \text{if } K = 1; \\ \tilde{C}_k(P/K) + (K-1)\mathbb{E}[\tilde{C}(P/K)] & \text{if } K > 1. \end{cases}$$

$$m_k = \arg \max_{1 \leq K \leq N_T} \grave{C}_k(P/K) \tag{5.6}$$

Regardless of the mode selection criteria used, the user feeds back, m_k and $\tilde{C}_k(P/K)$ to the base station for processing. After the base station selects the best mode and the K users that correspond to that mode, each user in $\mathcal{S}$ feeds back its quantized channel direction, $\tilde{\mathbf{h}}_k$.

5.3.2 Operations at the Transmitter

The transmitter groups the set of received modes, m_k, into sets denoted $\mathcal{G}^m = \{k \in M \mid m_k = m\}$. Let M denote the total number of users. The best transmission mode is then chosen via the following expression:

$$m^* = \arg \max_{1 \leq m \leq N_T} \sum_{k \in \mathcal{S}_m} \grave{C}_k(P/m) \tag{5.7}$$

$\mathcal{S}_m$ can be recursively defined as:

$$\mathcal{S}_m = \phi$$

$$\mathcal{S}_m \leftarrow \left\{ \arg \max_{k \in \mathcal{G}^m / \mathcal{S}_m} \grave{C}_k(P/m) \right\} \cup \mathcal{S}_m, \text{ while } |\mathcal{S}_m| < m$$

The scheduled set of users can then be denoted as $\mathcal{S} = \mathcal{S}_{m^*}$. In [13], it was argued that it is better to select mode m for users in $\mathcal{G}^m$ as opposed to any other mode, $m' \neq m$. However, it was not shown that choosing mode m provides the best performance (based on the given rate approximations) iff all users are from $\mathcal{G}^m$. In fact, it can be shown that there are cases when users from $\mathcal{G}^{m'}$, $m' \neq m$ perform better than $m^* = m$. Fig. 5.2 plots the rate approximations under different values of K (Mode) and Z (accuracy of quantization) for different received power $P\|\mathbf{h}_k\|_2^2$. The square pointers refer to the rate approximations made by users 1,2 and 3 at mode 2. The

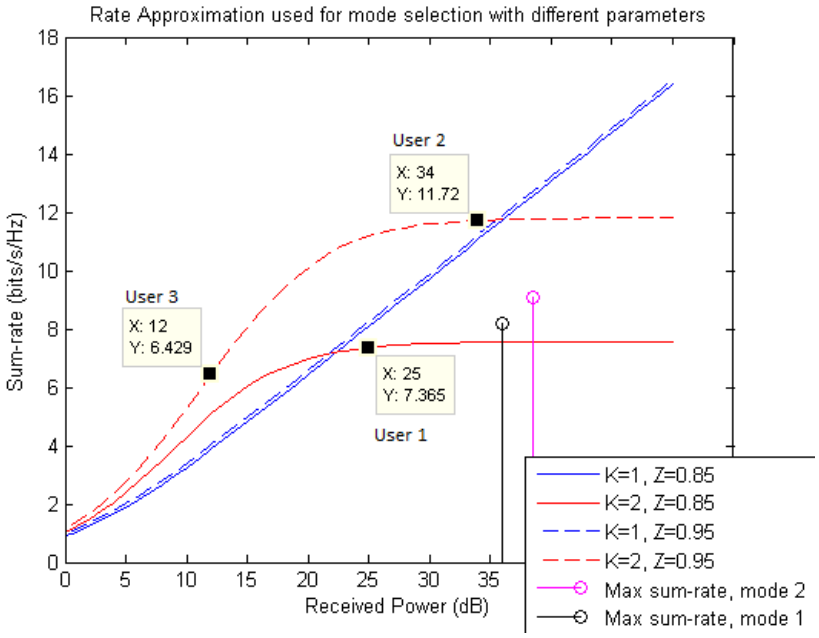

Figure 5.2: Scenario illustrating sub-optimality in the user selection

blue curves represent rate approximations for mode 1 and red curves for mode 2. From the plots it is clear that User 1 $\in \mathcal{G}^1$ and User 2, 3 $\in \mathcal{G}^2$. We assume the users in question have the highest rate approximations in their respective modes. The black stem shows the max sum-rate approximation for mode 1 and magenta for mode 2. The sum-rate for mode 2 being higher than mode 1 is hence chosen. Note however, if we instead transmit to User 1 and User 3, we would have a sum-rate that outperforms the proposed scheme.

Clearly, the user grouping procedure can be improved. However, the idea behind keeping the user groups independent in the scheduling process is twofold. Firstly, the feedback information is scaled differently in each mode, and hence the base station cannot directly compare the rate approximations across different user groups. Secondly, if the base station does have this information, the optimal set of users can only be found via trying all combinations of users. Although grouping users is sub-optimal, it

does perform better than random selection proposed in previous works [52] and requires less feedback bits when compared to semi-orthogonal scheduling(SUS) methods [4], while providing some performance gains via multi-mode selection.

5.4 Proposed Rate Approximation

Rate approximation plays a vital role in multi-mode scheduling as it gives each user an idea of how many other users can be co-scheduled. The accuracy of the rate approximation not only affects the effective mode that the base station chooses but also determines the users that lie within each mode. Different rate approximations will in general not contain the same users in $\mathcal{G}^m$. In the end, the actual performance of the system at a given time is determined solely on $\mathcal{S}$ and m^*. Since each user will neither have the channel information of the co-scheduled users, nor can it predict which users will be co-scheduled, an accurate estimate of the rate at each time instance is not possible. It is however, possible to evaluate the expected interference within each user's SINR. We start off by reformulating γ_k. Firstly we note that $\tilde{\mathbf{h}}_k$ can be decomposed as:

$$\tilde{\mathbf{h}}_k = \sqrt{Z}\hat{\mathbf{h}}_k + \sqrt{1-Z}\mathbf{s}_k \tag{5.8}$$

where $Z = |\tilde{\mathbf{h}}_k^H \hat{\mathbf{h}}_k|^2$, $\mathbf{s}_k$ is unit norm and orthogonal to $\hat{\mathbf{h}}_k$. If we take the dot product of equation (5.8) with $\mathbf{b}_k$ and take the magnitude squared, owing to the fact that $(1-Z)|\mathbf{s}_k\mathbf{b}_k|^2 > 0$, we can state the following lower bound - $|\tilde{\mathbf{h}}_k^H \mathbf{b}_k|^2 \geq Z|\hat{\mathbf{h}}_k^H \mathbf{b}_k|^2$. Furthermore, since $\mathbf{b}_j$ is orthogonal to $\hat{\mathbf{h}_k}$ for $j \neq k, j \in \mathcal{S}$, we can also state the following $|\tilde{\mathbf{h}}_k^H \mathbf{b}_j|^2 = (1-Z)|\mathbf{s}_k^H \mathbf{b}_j|^2$, $j \neq k$. We can then form a lower bound on the SINR given by:

$$\gamma_k \geq \frac{ZX}{\frac{K}{Pa} + (1-Z)\sum_{j\in S, j\neq k} Y_j} = \hat{\gamma}_k$$

where, $\|\mathbf{h}_k^H\| = a$, $|\hat{\mathbf{h}}_k^H \mathbf{b}_k|^2 = X$, and $|\mathbf{s}_k^H \mathbf{b}_j|^2 = Y_j$. Y_j are i.i.d. [5], and we can set $Y_1 = Y_2 = ... = Y$. Before proceeding, we will discuss certain prerequisites for formulating the rate approximation.

5.4.1 Distribution Of Parameters

In [13, 5, 9] it is shown that each of the variables are distributed as follows, assuming that the channel elements are circular symmetric complex Gaussian:

$X \sim \mathcal{B}(N_T - K + 1, K - 1)$

$Y \sim \mathcal{B}(1, N_T - 2)$

$Z : F_Z(z) = (1 - (1 - z)^{N_T - 1})^{2^B}$ (Also the minimum of 2^B, $\mathcal{B}(1, N_T - 1)$ distributed random variables

$a \sim \Gamma(N_T, \sigma^2)$ σ^2 is the variance of the channel coefficients

Here, $\mathcal{B}(\alpha, \beta)$ corresponds to the beta distribution, $\Gamma(k, \sigma^2)$ corresponds to the gamma distribution, and $F_Z(z)$ is the cumulative distribution function (CDF) of the random variable Z.

5.4.2 Convexity Analysis

Essentially, we now have the lower bound of the SINR to be a function of a,X,Y and Z.

$$\hat{\gamma}_k = f_k(a, X, Y, Z)$$

We can split up the above parameters based on whether or not they are known at the user. For the parameters unknown to the user it is intuitive to take their expectation in order to approximate the SINR. To this end, Z and a are known to the user, whereas X and Y, which depend on the beamformers, are not. However, the sum-rate approximation performed at each user also depends on the channel details of the co-scheduled users, and this is not available. A solution proposed in [13], was to simply assume

the co-scheduled users have the same rate as the user in question. We propose a more intuitive solution whereby, we take the expectation over all four parameters when computing the rates of the co-scheduled users. While the differences between the two solutions will be highlighted again when formulating the rate approximations, the main reasoning to perform convexity analysis on the variables is that the joint distributions required to compute the mentioned expectations are not trivial to formulate. With convexity, Jensen's inequality can be utilized while maintaining the lower bound to drag the expectations within the function. i.e $\mathbb{E}_X[f_k(X,Y,Z)] \geq f_k(\mathbb{E}_X[X],Y,Z)$ if f_k is convex in X. Another underlying assumption is that the parameters are independent in order to split up the joint expectation. This will be discussed in the next subsection.

Convexity in Y

Convexity in Y is simple to address due to the well known fact that $\log(1+1/x)$ is convex. Now if we look at $f_k(Y;X,Z)$, it is of the form:

$$f(Y) = \frac{\alpha_1}{\alpha_2 + \alpha_3 Y}, \quad \alpha_1, \alpha_2, \alpha_3 > 0$$

$$C(Y) = \log_2(1 + f(Y))$$

Since all the constants are greater than 0, the above operation is simply a scaling and a shifting operation of $\log(1+1/x)$ and hence $C(Y)$ is convex. As an example setting $\alpha_i = 0.5, i = 1,2,3$ fig. 5.3 plots $C(Y)$.

Convexity in Z

Convexity in Z is similar to the above, however there are extra conditions to be enforced before convexity can be proved. As such, the second deravitive approach is used here.

$$f(Z) = \frac{\alpha_1 Z}{\alpha_2 + \alpha_3(1-Z)}, \quad \alpha_1, \alpha_2, \alpha_3 > 0$$

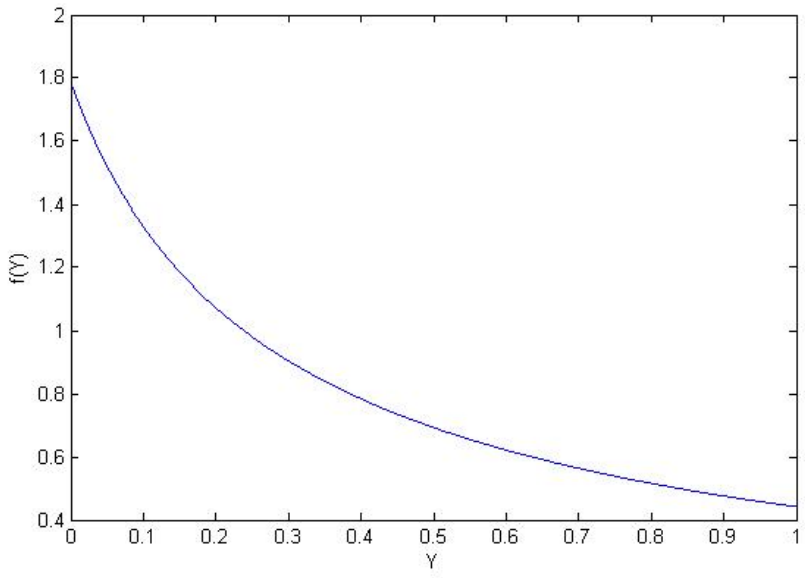

Figure 5.3: Convexity in Y

First we note that the above can be rewritten as:

$$f(Z) = \frac{\alpha_1 Z}{\alpha_4 - Z}, \quad \alpha_1 > 0, \alpha_4 > 1, 1 \geq Z \geq 0$$

$$C(Z) = \log_2(1 + f(Z))$$

$$\begin{aligned}
\frac{d}{dx}C(x) &= \frac{\alpha_1\alpha_2}{(\alpha_2 - x)^2(1 + \frac{\alpha_1}{\alpha_2 - x})} \\
&= \frac{\alpha_1\alpha_2}{(\alpha_2 - x)^2 + \alpha_1(\alpha_2 - x)} \\
\frac{d^2}{dx^2}C(x) &= \frac{\alpha_1\alpha_2}{((\alpha_2 - x)^2 + \alpha_1(\alpha_2 - x))^2}(2(\alpha_2 - x) + \alpha_1) \\
&\geq 0
\end{aligned}$$

The last inequality comes from the fact that $\alpha_2 \geq 1 \geq x$ and $\alpha_2, \alpha_1, x > 0$. This implies that $C(Z)$ is convex as well. Again, if desired, the above can be verified via simulations.

Convexity in a and X

$f(.)$ is a non-convex function of a and X. Noting that $\log(1+x)$ is not convex, we see that $f(X) = a_1 X$, for $a_1 > 0$ and $f(a) = \frac{b_1}{b_2/a+b_3}$, for $b_1, b_2, b_3 \geq 0$. Clearly $\log(1+f(X))$ is non-convex. $\log(1+f(a))$ would be non-convex for all a if $b_3 = 0$; a quick check by double differentiating $\log(1+f(a))$ would show that this not always convex when $b_3 > 0$. Since a can take any value in $[0, \infty]$, Jensen's inequality will fail at certain regions depending on the coefficients.

It is however important to note that $\log(1+f(a))$ is a convex function of $\frac{1}{a}$, by the same arguments used in proving the convexity in Y. So for future reference, instead of taking $\mathbb{E}_a(\log(1+f(a)))$ we can instead perform $\mathbb{E}_{\frac{1}{a}}(\log(1+f(\frac{1}{a}))) \geq \log(1+f(\mathbb{E}_{\frac{1}{a}}[\frac{1}{a}]))$ in order to maintain the lower bound on the rate-approximation.

5.4.3 Independence Analysis

As mentioned earlier, we can use convexity to simplify a lot of the derivations while maintaining a lower bound on the rate approximation. This however depends on an important assumption that X, Y, Z and a are mutually independent. It is easy to argue that a is independent of the other three simply because a refers to the channel magnitude while the other terms depend on the quantized channel directions and the beamformers which do not utilize the channel magnitude.

It is also arguable that Z is independent of X. Z corresponds to the quantization error of the channel. X is simply the dot product of two vectors that are independent of the original channel, and hence are also independent of the quantization error. Y is the dot product of the channel of the user and the beamformers of the other users and it is hard to picture intuitively the independence or the dependence of Z and Y. While a decrease in quantization error decreases teh magnitude of the interference terms in the SINR,

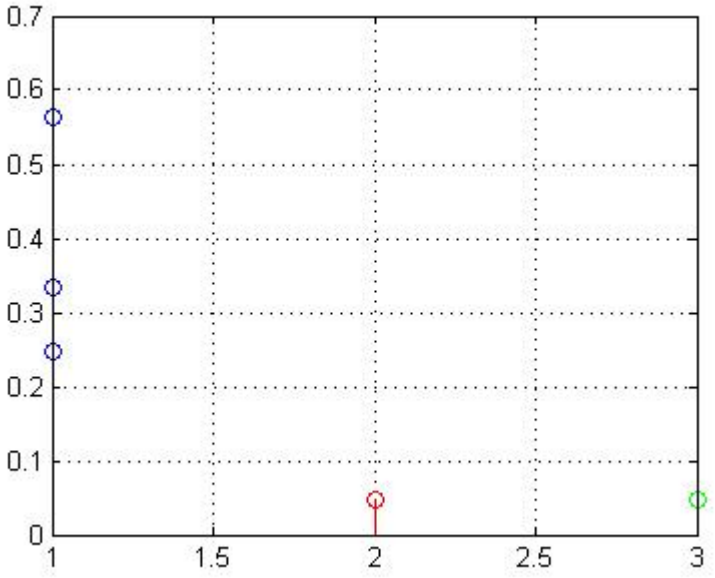

Figure 5.4: Demonstrating Independence property of X, Y and Z

this is taken care of by the $(1-Z)$ term in the expression. In order to prove independence we have to show the following.

$$p(X,Y,Z) = p(X|Y,Z)f(Y|Z)f(Z) = p(X)p(Y)p(Z)$$

$p(.)$ represents the joint pdf of X, Y, Z

It is not trivial however to work out the joint distributions of X,Y and Z and we cannot directly show independence, therefore when using independence in the proof we use it solely as an assumption to simplify the proof. In order to give an idea if the variables could possibly be independent is by using a property of independence. In other words we are testing by simulation of the following holds:

$$\mathbb{E}[XYZ] = \mathbb{E}_Z[Z]\mathbb{E}_X[X]\mathbb{E}_Y[Y]$$

Figure 5.4 simulates X, Y and Z by performing ZFBF and channel quantization over randomly selected users. The blue stems correspond to

the individual expectations of X, Y and Z. The red stem corresponds to the product of these expectations and the green stem correspond to the expectation of the product of X, Y and Z. As we note, the green and red stems are equal.

5.4.4 Rate Approximation Formulation

Here our proposed rate approximations are derived. The problem with the rate approximation proposed in [13] was that firstly, while they mentioned it was closed form, upon careful analysis the solution contained a hyper-geometric function, which is essentially an infinite sum. Furthermore, the sum did not converge under certain scenarios, even so it was not simplified for the cases where it does converge. As such we derive an alternative closed form expression that can be verified to be accurate via using the integral function in MATLAB.

Mode 1

In the case of a single-user scenario, the SINR can be written as:

$$\gamma = P\|\mathbf{h}\|^2|\tilde{\mathbf{h}}^H\mathbf{b}|^2$$

$$\tilde{C}_k(1) = \log_2(1+\gamma)$$

Notice that the SINR is a function of the channel vector, and the beamformer is based solely on the user's channel as there are no interferers. This means that there is no approximation needed, assuming the user knows the transmission scheme being used (maximum ratio transmission in this case).

Mode K

When approximating the sum-rate of $1 < K \leq N_T$ co-scheduled users, we need to work out the rate of the user as well as the co-scheduled users. The

rate of the user k can be approximated as follows.

$$\begin{aligned}\mathbb{E}\left[\log_2(1+\gamma_k)\right] &\geq \mathbb{E}_{X,Y}\left[\log_2\left(1+\hat{\gamma}_k(X,Y)\right)\right] \\ &\geq \mathbb{E}_X\left[\log_2\left(1+\hat{\gamma}_k\left(X,\mathbb{E}_Y[Y]\right)\right)\right] \\ &= \tilde{C}_k(K)\end{aligned}$$

It should be noted that the above expression is an approximate lower bound. We used the discussions from earlier that the above expression is a convex function of Y and therefore maintained the lower bound and simplifying the expression via Jensen's inequality. Furthermore, it was assumed that X and Y are independent. The closed form expression for the rate approximation can then be given by:

$$\begin{aligned}\tilde{C}_k(P,K;\delta_k) &= \frac{\Gamma(N_T)}{\Gamma(N_T-K+1)\Gamma(K-1)}\sum_{i=0}^{K-2}\binom{K-2}{i} \\ &\frac{(-1)^i}{N_T-K+i+1}\Bigg\{\left[1-\frac{(-1)^{N_T-K+i+1}}{{\delta_k}^{N_T-K+i+1}}\right]\log(1+\delta_k) \\ &+\sum_{l=1}^{N_T-K+i+1}\frac{(-1)^l}{(N_T-K+i-l+2){\delta_k}^{l-1}}\Bigg\}\times\frac{1}{\log(2)}\end{aligned} \tag{5.9}$$

where $\Gamma(.)$ is the gamma function, and δ_k is given by:

$$\delta_k = \frac{Z}{\frac{K}{Pa}+(K-1)(1-Z)\mathbb{E}_Y[Y]}$$

A brief account of the derivation is given below:

$$\tilde{C}_k(P/K) = \int_0^1 \log_2(1+\delta_k x) f_X(x)dx \tag{5.10}$$

The probability density function (pdf) of X can then be substituted in the above equation to form:

$$\tilde{C}_k(P/K) = A\int_0^1 \log_2(1+\delta_k x)x^{N_T-K}(1-x)^{K-2}dx \tag{5.11}$$

where A is a constant dependent on N_T and K. In order to simplify the integral, a binomial expansion can be performed on $(1-x)^{K-2}$. Having

done so, the integral simplifies to a sum of several integrals of the form:

$$\int_0^1 \log(a+bx)x^m dx = \frac{1}{m+1}\left[x^{m+1} - \frac{(-a)^{m+1}}{b^{m+1}}\right] \times \log(a+bx) + \frac{1}{m+1}\sum_{l=1}^{m+1} \frac{(-1)^l x^{m-l+2} a^{l-1}}{(m-l+2)b^{l-1}} \tag{5.12}$$

The above involves the substitution $e^y = a + bx$, as well as integration by parts and a binomial expansion. Equations (5.10) (5.11), and (5.12) can be used in conjunction to form equation (5.9). Recall in equation (5.6), we proposed an alternative sum-rate expression that used an expected estimate for the rate of the interfering users. Unfortunately, the closed form expression for $\mathbb{E}_{a,Z}[\tilde{C}_k(P,K;\delta_k)]$ is complicated to evaluate and might be intractable. To this end, we use convexity again to our advantage to maintain the lower bound and find a simpler expression for the co-scheduled users. To be explicit, we can define the user rate by the following function:

$$f_k(a,X,Y,Z) = \log_2\left(1 + \frac{ZX}{\frac{K}{Pa} + (1-Z)\sum_{j\in S, j\neq k} Y_j}\right) \tag{5.13}$$

We had already described that $f_k(a,X,Y,Z)$ is convex in Y. Through convexity tests, e.g. double derivative test, it can be shown that $f_k(.)$ is also convex in Z, and concave in a (but convex in $1/a$). This means that for the interfering terms, have an expression similar to (5.9) , except instead of δ_k, we use δ_j:

$$\delta_j = \frac{\mathbb{E}_Z[Z]}{\frac{K}{P}\mathbb{E}_{\frac{1}{a}}[\frac{1}{a}] + (K-1)\left(1-\mathbb{E}_Z[Z]\right)\mathbb{E}_Y[Y]}$$

$$\mathbb{E}(Y) = \frac{1}{N_T - 1}$$

$$\mathbb{E}(1/a) = 1/(N_T-1)$$

$$\mathbb{E}(Z) = 1 - \sum_{i=0}^{N_C} \frac{\binom{N_C}{i}(-1)^i}{i(N_T-1)+1}$$

The approximated sum-rate at user k used for mode selection can then be given by:

$$m_k = \arg \max_{1\leq K\leq N_T} \tilde{C}_k(P,K;\delta_k) + (K-1)\tilde{C}_k(P,K;\delta_j) \tag{5.14}$$

For future reference, the corrected version of the rate approximation proposed in [13] is given by:

$$\text{RA1} : \arg \max_{1 \leq K \leq N_T} K\tilde{C}_k(P, K; \delta_k)$$

and the proposed rate approximation is given by:

$$\text{RA2} : \arg \max_{1 \leq K \leq N_T} \tilde{C}_k(P, K; \delta_k) + (K-1)\tilde{C}_k(P, K; \delta_j)$$

Fig. 5.5 compares the ergodic sum-rate (via simulations) with the two rate approximations for the case when $K = 4$. We will note that as expected, both of the rate approximations are lower bounds and in fact the gap between RA1 and RA2 is relatively small. RA1 is labelled as biased, as the interference terms in the approximation are based on the user's channel at a particular symbol period. RA2 is labelled as unbiased, as the interference terms are calculated using the ergodic channel statistics and are not biased to any particular user.

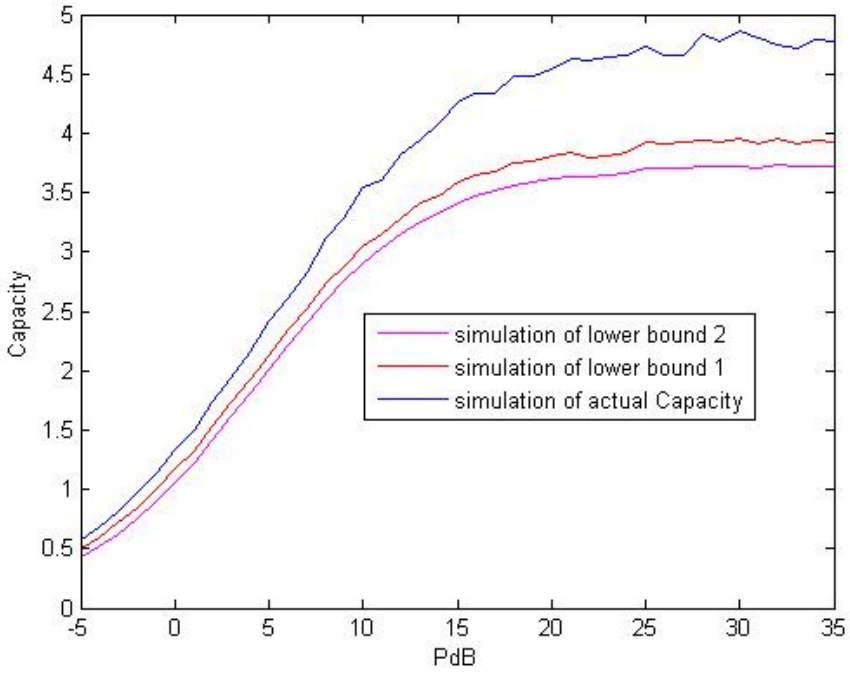

Figure 5.5: Ergodic Sum-Rate of Random Scheduling (Blue) vs Biased Rate Approximation, RA1 (Red) vs Unbiased Rate Approximation RA2

5.5 Preliminary Simulations and Analysis

Most of the following simulations compare the mode selection strategies used in equation (5.5) vs that in equation (5.6). The simulations are done over the range of powers -10 to 50 dB, and are repeated for 1000 iterations for each power level. The simulation in Fig. 5.6 shows the achieved sum-rate on average for the two different schemes which we will denote Rate Approximation 1 (RA1), equation (5.5), and Rate Approximation 2 (RA2), equation (5.6). Fig. 5.7 shows the modes that have been selected on average at each power level; $N_T = 4$, $M = 25$. The solid curves correspond to 4 feedback bits, and the dashed curves correspond to 15 feedback bits.

For the 4 bit case, we can see that RA2 outperforms RA1 at all SNR

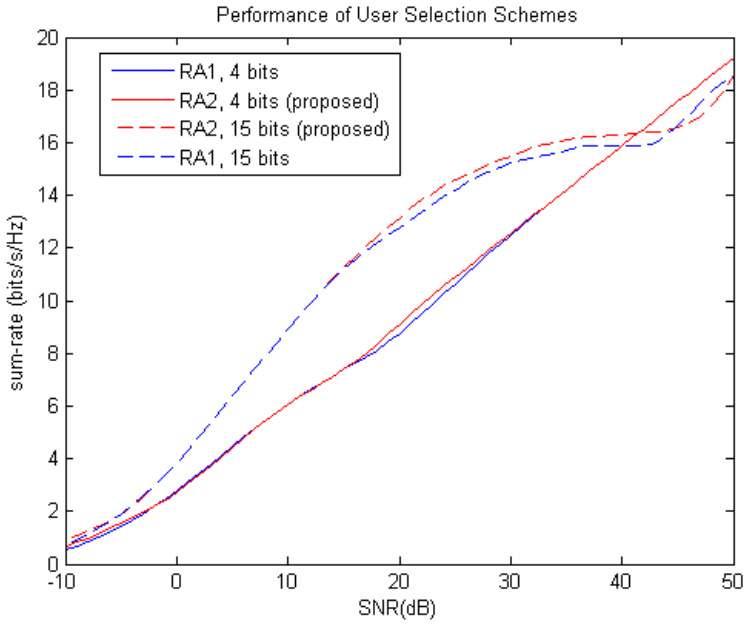

Figure 5.6: Sum-rate of RA1 vs RA2

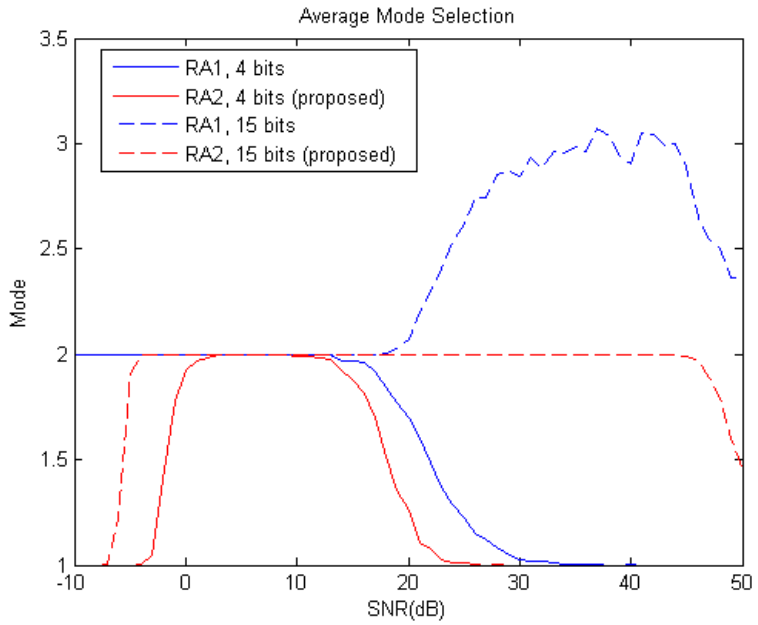

Figure 5.7: Average mode selected in RA1 and RA2

levels. On the other hand, for 15 bits, while RA2 performs well at a certain range, RA1 briefly outperforms RA2 at high SNR. One major discrepancy with respect to intuition is that the 15 bit curves should in theory outperform the 4 bit curves simply because of the improved quantization. Whereas at around 40-50dB we see that the 4 bit curves in fact perform better. The reason for the difference in the simulation can be partly explained at follows. If we look at Fig. 5.7, we note that at the 40-50dB range, the 15 bit curves are transitioning from modes 2 and 3 down to a lower mode. In the 4 bit case, the transition to mode 1, which is optimal at high SNR, has already been made.

5.5.1 Varying the Number of Users

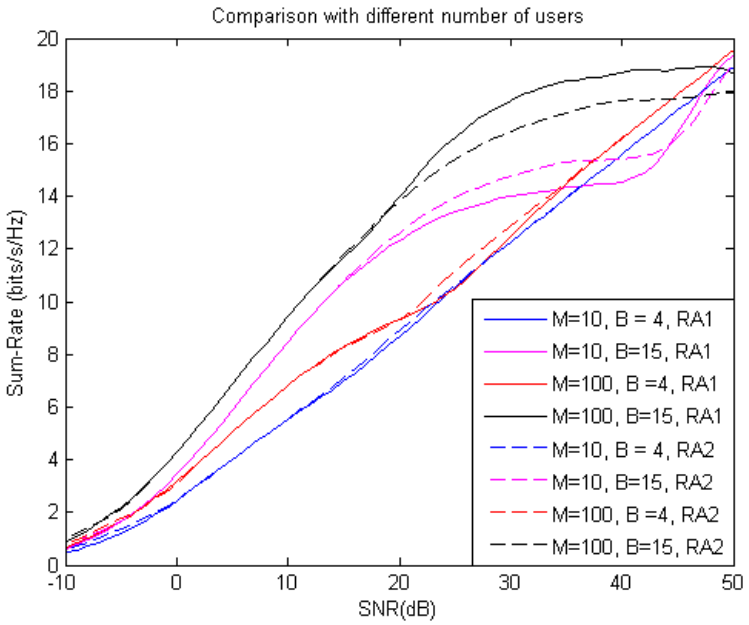

Figure 5.8: Sum-rate of RA1 vs RA2 for 10 and 100 users

Here, we study the effects of varying the number of users in the system. In Fig. 5.8, the solid curves represent the sum-rate under different values of M and B for RA1. The dashed curves repeat the same simulations for

RA2. It is apparent that, except in the black curves and parts of the magenta curve, RA2 outperforms RA1. The black curves correspond to the case where B=15, M=100. In other words, when there is a large number of users and high resolution of feedback. In such a case, RA1 outperforms RA2 substantially. The reasoning behind this is as follows. With a large number of users, multi-user diversity gains come into play. With the choice of more users, there is a higher likelihood of finding users that fit in higher modes of operation. With a higher resolution of feedback zero-forcing can be done more accurately, and hence the system is able to support more users. In other words, this improves multiplexing gains. With higher multiplexing and diversity gains, they system is expected to prefer higher modes of transmission for the given SNR values. Recalling that although RA2 provides a more unbiased rate approximation, it is essentially a lower bound compared to RA1, and hence would prefer lower modes of operation. As such, it is not surprising that RA1 performs better in a scenario where higher modes of operation are preferred. Therefore, we can conclude from the above observation that RA2 performs better than RA1 in a scenario where there is a small-medium number of users and a low resolution of feedback, in the case where the transmitter has four antennas. Furthermore, we also note that when the number of feedback bits is low, increasing the number of users does not have as much of an effect as if the number of feedback bits is high.

5.5.2 Mean Square Error using Rate Approximation

This section aims to explain the observation that RA2 performs better than RA1 in scenarios involving few feedback bits and a small to moderate number of users. Recall, RA1 approximates the rate of user k and assumes the rate of the co-scheduled users is the same. More often than not the assumption will not hold and will underestimate or overestimate the actual rate. Hence, by intuition we would expect the mean squared error between RA1 and the

actual rate to be significant. On the other hand, RA2 is a lower bound to RA1, which means that its rate approximation on average is worse than RA1. However, this is outweighed by the way RA2 is constructed. In RA2 the rate of the co-scheduled users follows an average approximation, causing it to be less biased than RA1. Therefore, while RA2 might expect a larger deviation from the actual sum-rate because it is a lower bound to RA1, this is outweighed by its more stable approximation. In order to validate the

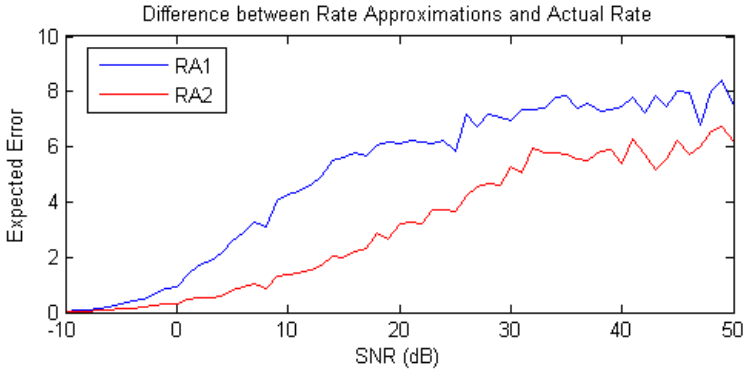

Figure 5.9: Mean squared error

above intuition, Fig. 5.9, plots the average error experienced between each of the rate approximations and the actual sum-rate. The above is plotted for $B = 4$ and $M = 25$, operating at mode 2. For the given scenario, RA2 has less error on average than RA1.

In the cases involving large number of users and high feedback resolution, the base station would prefer to operate at higher modes of operation. With more co-scheduled users, the average rate-gap between RA1 and RA2 grows larger. Hence, the stability of the approximation in RA2 is unable to offset the error caused from being a lower bound to RA1. A potential improvement would then to exploit the properties of RA1 and RA2. Firstly, noting that the above scheme does not assume a QoS requirement at the user, it is not necessary to stick to lower bounds, and hence some of the convexity assumptions in the derivations can be relaxed. Furthermore, the

rate approximations are constructed without the use of scheduling knowledge. With user scheduling, we would expect the sum-rate to be higher than with random selection. Considering this, it may actually be more useful to incorporate upper bounds on the rate approximations instead.

5.6 Proposed Multi-Mode Scheduling with SUS

We left the previous section with the idea that we need not restrict ourselves to a lower bound when computing the rate approximations. What this means is that we can relax some of the convexity properties and even possibly attempt to establish a more accurate rate approximation via higher order integrals. While these may improve the performance slightly in an ergodic sense, in the end, tweaking the rate approximation to match the actual ergodic sum-rate neglects a bigger problem that inhibits substantial improvement in performance. The reason for this restraint is as follows. Recall when we discussed scheduling in chapter [Ch 4], we established that in literature the scheduling algorithms that seen most improvement in performance incorporate both the CQI feedback, as selecting users with higher CQI would mean that they would attain higher receive power, and CDI feedback, as knowledge of direction means we can attempt to avoid users that have channels aligned to one another as this would produce interference. What we note however, is that the multi-mode scheduling algorithms seen in literature incorporate only CQI when attempting to schedule the users and use CDI purely for zero-forcing. As such, this section proposes a novel scheme to incorporate mode selection with a modified Semi-orthogonal User Scheduling (SUS) utilizing limited feedback, thus hopefully achieving the best of both schemes.

5.6.1 Incorporating SUS at the Base Station

The mode selection procedure is done as per the norm by the base station. The base station has knowledge of each user's preferred mode (m_k), the CQI ($\tilde{C}_k(P/K)$), and the CDI ($\hat{\mathbf{h}}_k$). The base station then computes m^*. Now instead of selecting the m^* best users from $\mathcal{G}^{m^*}$ based on their CQI, we perform SUS on all the users within $\mathcal{G}^{m^*}$ to select as many users as possible that are semi-orthogonal to one another, up to m^* users. Once m^* and its corresponding group of users are established by the base station the SUS algorithm can then be performed as follows.

Let $\pi(i)$ denote the index of the user selected at the i^{th} iteration of the algorithm (also known as permutation).$\omega(.)$ is the CQI(in our case the rate approximation $\tilde{C}_k(P/K)$) used in the algorithm and is usually the function of the channel. $\mathcal{T}_i$ is the set of users available for scheduling at the i^{th} iteration. ϵ is a tuning parameter used to determine the tightness of the semi-orthogonality between the users; in other words it restricts how aligned the user channels can be. Supposing there is a total of M users in the system and we require to choose up to m^* users, we can then formulate the algorithm as shown.

Initialization:

$$\mathcal{T}_1 = \mathcal{G}^{m^*}$$

$$\pi(1) = \arg\max_{k \in \mathcal{T}_1} \ \omega(\mathbf{h}_k)$$

Repeat:

$$\mathcal{T}_i = \left\{ m \in \mathcal{T}_{i-1} : |\hat{\mathbf{h}}_m^H \hat{\mathbf{h}}_{\pi(j)}| \leq \epsilon, \text{ for } 1 \leq j \leq i \right\}$$

$$\pi(i) = \arg\max_{k \in \mathcal{T}_1} \ \omega(\mathbf{h}_k)$$

$$i \leftarrow i + 1$$

Termination Conditions:

$$\mathcal{T}_i = \emptyset \text{ or } i > m^*$$

The set $\mathcal{S} = \{\pi(1), \ \ldots, \ \pi(K)\}$ where $1 \leq K \leq m^*$ is the total number of scheduled users. In the initialization, the set of available users is simply $\mathcal{G}^{m^*}$. The best user is chosen based on the rate approximation feedback. The remaining users that are available are then filtered and the users that are not semi-orthogonal (determined by ϵ) to the scheduled users so far are removed from the set $\mathcal{T}_i$. The next user is again selected based on rate approximation feedback. This process is repeated until there are no more available users in $\mathcal{T}$ or when m^* users have been scheduled.

A common problem in this scheme is that when the base station chooses m^* users, but the SUS algorithm only decides to choose $< m^*$ users (possibly because the users in group $\mathcal{G}^{m^*}$ have channels that are mostly aligned to one another). Intuitively, one might think in such a case it might be better to switch to a different group, for example the group with the second highest sum-rate approximation or the group corresponding to the number of users the SUS algorithm decides to select. The problem with these solutions is that it cannot be clearly justified whether making a transition to a different group would improve the performance. If we transition to a different group, say $\mathcal{G}^{m_1}$, the scheduling algorithm might pick up m_1 users. However, it is not necessarily true that the m_1 users from $\mathcal{G}^{m_1}$ would out perform the $< m^*$ users from $\mathcal{G}^{m^*}$. While it is true that, based on the rate approximations, transmitting to m^* semi-orthogonal users in group $\mathcal{G}^{m^*}$ is better than transmitting to $< m^*$ or even $> m^*$ users from the same group, the same does not hold when comparing different groups. Therefore, there is not enough motive for us to implement a fix for the scenario where the SUS algorithm chooses $< m^*$ users. Nonetheless, the alternative methods are described in the following; while these are not implemented, their advantages/disadvantages can be left for future work.

Next Best Group

Supposing the SUS algorithm could not find m^* suitable users from $\mathcal{G}^{m^*}$, the base station can instead look into the group with the next best sum-rate approximation. Let $[m_1, m_2, \ldots, m_{N_T}] = \mathbf{m}$ be a vector of the various modes ordered based on the sum-rate approximation of their corresponding groups, $\mathcal{G}^{m_i}$. Note that the subscripts on the elements of $\mathbf{m}$ do not refer to the actual mode, they are simply used to identify the best (m_1) and worst (m_{N_T}) mode groups in terms of rate approximation. Further suppose that $\mathcal{S}$ is the set of users scheduled by the SUS algorithm. Then the Next Best Group scheduling can be written as follows.

Initialization:

$$i = 1$$

Repeat:

$$\mathcal{S} = \text{SUS}(\mathcal{G}^{m_i})$$

$$\text{if } |\mathcal{S}| = m_i : \ m^* = m_i, \text{ terminate}$$

$$i \leftarrow i + 1$$

Add Closest Group

A common reason why the SUS algorithm chooses less than m^* users is simply due to the lack of users in the group. In such a scenario, we can add a set of users from a different group to $\mathcal{G}^{m^*}$ before performing SUS. An interesting question then is which group should be added? The most intuitive choices are the group with the next best rate approximation, as this would ensure that the users have decent channel quality, or it could be the group corresponding to $m^* - 1$ or $m^* + 1$ (i.e. the closest groups), if they exist. While this would increase the number of users to choose from in the SUS algorithm thereby achieving higher user diversity, there are two

drawbacks of this scheme. Firstly, because we are combining two groups worth of users in the scenario where SUS schedules $< m^*$ users, we may end up with a substantial increase in users which may increase the number of operations within the SUS algorithm. This issue is however, a minor one compared to the second. The second issue is that, the rate approximations from different groups cannot be directly compared since they correspond to different preferred modes, i.e. the feedback information is different. To this end, the base station would have to request the users in the closest group or next best group to send their rate approximation for the case when there are m^* scheduled users. This would cause feedback overheads. The research problem would then be whether the performance increase from this scheme outweights the feedback overhead and the increase in computational complexity, and again this is left for future work.

5.6.2 Simulations and Analysis

So how does incorporating SUS into the multi-mode scheduling scheme affect the performance? The following figures plot the performance of SUS based mode scheduling on top of the performance of RA1 and RA2 as in section 5.5. The black curve refers to the SUS scheme and the blue and red curves as before refer to RA1 and RA2 respectively. Fig. 5.10 corresponds to the case where there are 4 feedback bits and 50 users. Fig. 5.11 has 8 feedback bits and 100 users.

We can see consistent increase in performance by using SUS especially in the low-mid SNR regions. This improvement however is only slightly more prominent as we increase the number of feedback bits and users. So why are the improvements only marginal? The following subsections aim to explain the observations and propose improvements.

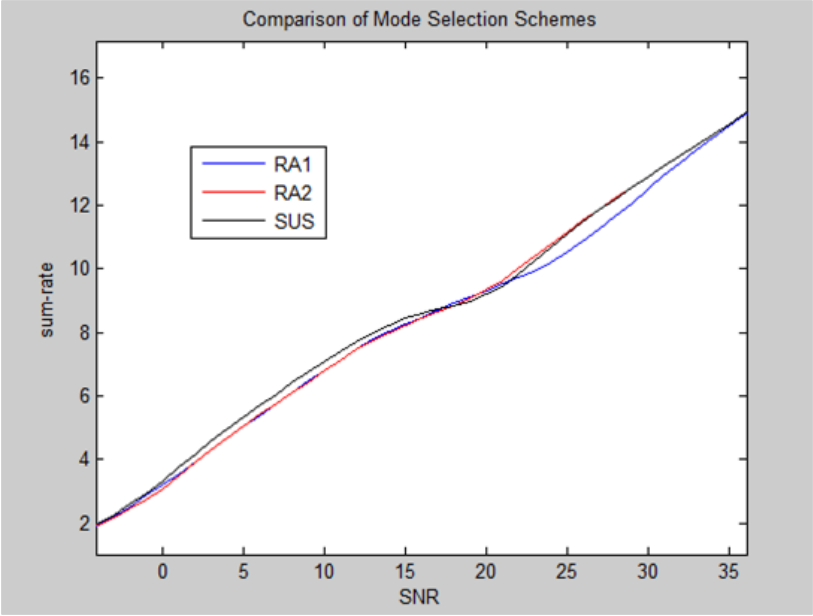

Figure 5.10: Mode selection with SUS + RA2 (Black) vs RA1 (Blue) vs RA2 (Red) : 4 feedback bits 50 users

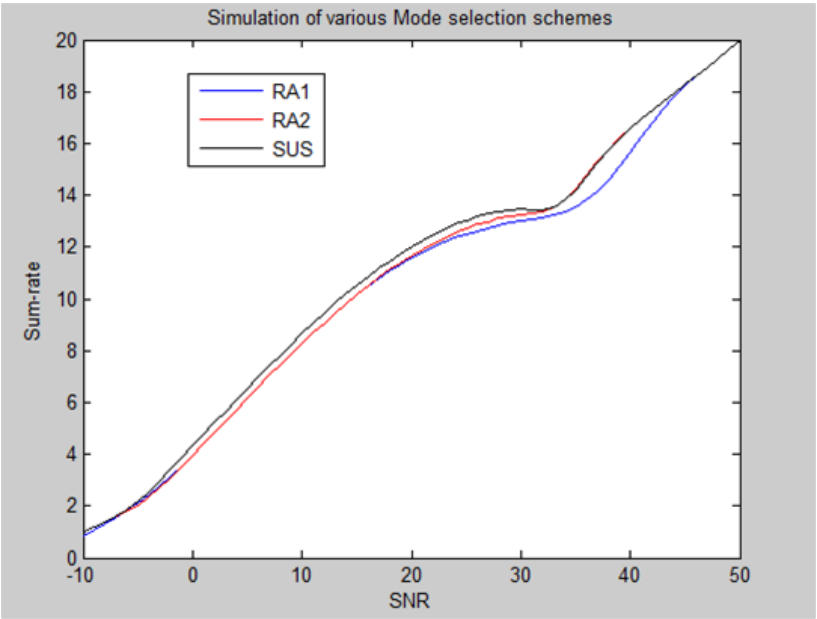

Figure 5.11: Mode selection with SUS + RA2 (Black) vs RA1 (Blue) vs RA2 (Red) : 8 feedback bits 100 users

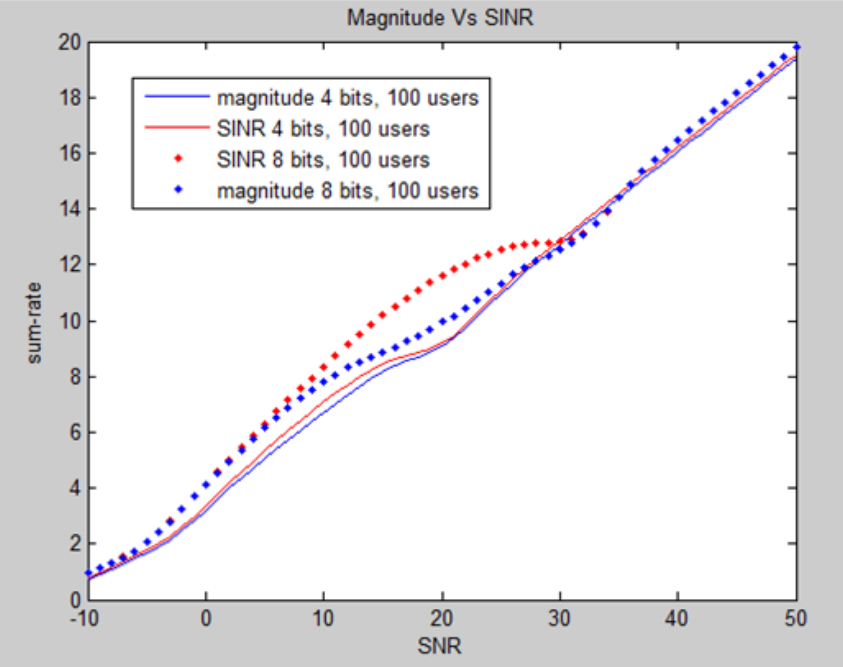

Figure 5.12: Mode Selection(SUS) SINR feedback vs Magnitude feedback

SINR vs Magnitude feedback

The first way in which we can attempt to increase the performance gains is to look at whether SINR feedback is better than magnitude feedback. As we know rate approximations from each user are essentially the log of their SINR. SUS uses this to select the best user in each iteration. However, this SINR is merely an approximation, whereas the channel magnitude which is assumed to be known perfectly at the receiver need not be approximated. So, if instead we use magnitude feedback instead would we see an improvement? Fig 5.12 shows the results. The red curves correspond to the sum-rate of the system with SINR feedback in the SUS algorithm and blue curves correspond to magnitude feedback. This is plotted for 4 and 8 bits of feedback. What we see is that SINR feedback consistently outperforms magnitude feedback. The reasoning for this is simple. While SINR is an approximation, it provides rough information on how much the quantization

errors and interference affect the given user. If we simply send magnitude feedback, while the user might have a strong channel, it may be the case that it suffers from quantization errors and we cannot infer this from $\|\mathbf{h}_k\|$.

User Diversity

Even though the graphs simulate schemes with a relatively average number of users, the fact that we are dividing these users into different groups mean that we are only dealing with a few users when performing SUS. So while performing SUS without mode selection may give a more significant gain in the low-mid SNR regimes, mode switching is essential in attaining multiplexing gains at high SNR. Essentially by using SUS along with mode selection apart from simply using SUS, we are trading off some of the user diversity for multiplexing gains at high SNR regimes. As such, the performance gains we see are marginal, and the reason why we are seeing these gains is that SUS assures some degree of semi-orthogonality between the co-scheduled users thereby reducing interference. In order to see a more substantial improvement in performance, one approach would be to perform mode selection as per usual, but perform SUS over all the users, not just the selected group. In such a scenario we are increasing the computational complexity as a trade-off for performance improvements.

Mode Selection Does Not Account for SUS

As the heading points out, we are selecting the best group based on the highest rate approximations from each group and do not take into account SUS. The problem with this is that the users that correspond to these high rate approximations may not even be selected by the SUS algorithm as they may not be semi-orthogonal to the other users in the group. Arguably it might be better to decide on the best mode of transmission based on the average sum-rate approximation of each group. This way the users with

low rate-approximations will also have some weight to the sum-rate approximations in case they are chosen by SUS for lack of better choices. Doing so however produces another problem. While the mentioned solution works for multi-user scenarios, when deciding between single-user and the other modes an important point can be raised. In single-user mode, there is no user scheduling required and the best user is always chosen. Furthermore, recall that our rate approximations are lower bounds to the ergodic rates but this doesn't apply to the single-user rates which can be computed accurately. Hence, the group corresponding to single-user mode will have a higher weightage in terms of being selected. However, in the event that one of the multi-user modes is chosen, the use of the average rate approximation in this selection may improve the overall performance when using SUS as opposed to simply selecting the group based on the best m^* users.

Rate Approximation Does Not Account for SUS

By far the biggest limiting factor on the performance improvement however, is that the rate approximations is a lower bound when compared to the actual performance and the rate gap between the approximation and the actual sum-rate is in fact larger when we utilize SUS. In other words, in order to see more of an improvement in performance the rate approximation has to be modified to take into account that SUS is being used at the base station. To this end, section 5.6.3 derives and analyses a modified version of the rate approximation that incorporates SUS.

5.6.3 Incorporating SUS in the Rate Approximation

So how can we incorporate the fact that the base station is using SUS to improve on the rate approximations performed by each user. A quick search in literature will provide a common large user approximation of the SINR that the base station uses as CQI. This is given as follows:

$$\gamma_k \approx \frac{\rho \|\mathbf{h}_k\|^2 \cos_2^2 \theta_k}{1 + \rho \|\mathbf{h}_k\|_2^2 \sin^2 \theta_k} \tag{5.15}$$

Here, $\cos^2 \theta_k$ is simply Z, $\rho = P/K$ where K is the total number of scheduled users by the SUS algorithm, and $\|\mathbf{h}_k\|^2$ is simply a. We can then rewrite the above as follows:

$$\gamma_k \approx \frac{Z}{\frac{N_T}{Pa} + (1 - Z)} \tag{5.16}$$

Compared to the rate approximation in equation (5.13), we notice that a lot of terms are in fact missing from the SUS-based approximation. Firstly, $X = 1$ is due to the fact that this is a large user approximation, as the number of users grow large it becomes more probable to find a set of users whose quantized channels are almost entirely orthogonal to one another. In order to ensure this, as the number of users grow large, ϵ can be decreased. In such a scenario, assuming the quantized channels are indeed orthogonal, this would mean that beamformers will be aligned with the quantized channels. In other words, $\|\hat{\mathbf{h}}_k \mathbf{b}_k\|^2 = X = 1$. Secondly, we note that the term $(K-1)\mathbb{E}_Y[Y] = (K-1)/(N_T - 1) = 1$. This is because, classically SUS algorithm attempts to schedule as many users as possible, and in a ZFBF based broadcast channel the maximum users possible is $K = N_T$. In a large user scenario, it is almost always possible to find N_T candidates and thus the SINR approximation does not consider the case where $K < N_T$. For our purposes however, we need to modify the above expression to account for cases where $K < N_T$, since we are incorporating mode selection. Lastly we notice that in order to approximate the rate of the co-scheduled users, we are taking the expectation over Z and a, whereas previously we had used convexity to simplify this procedure. Since X is no longer a variable in the large user approximation we can afford to do this without overcomplicating the derivation. The generalization of the above for any $K > 1$ can be given

as follows.

$$C_k^{SUS}(K) = \log_2(1+\gamma_k) \approx \log_2\left(1 + \frac{Z}{\frac{K}{Pa} + \frac{K-1}{N_T-1}(1-Z)}\right) \tag{5.17}$$

This is derived simply by taking $X = 1$ and the expectation on Y via Jensen's inequality in equation (5.13). Now since Z and a are known to the user, it can approximate its own rate by the given equation. For the co-scheduled users however, we need to take the expectation over Z and a. In other words we need to solve the following:

$$\int_0^\infty \int_0^1 \log_2\left(1 + \frac{z}{\frac{K}{Pa} + \frac{K-1}{N_T-1}(1-z)}\right) p_Z(z).dz \; p_a(a).da \tag{5.18}$$

Where, as before, $p_X(.)$ is the probability density function of variable X. While the above double integral can be evaluated, its process is tedious and is in the form of exponential integrals, some of which do not converge. Hence, in order to simplify the proof, we replace $1/a$ with $\mathbb{E}[1/a]$ by utilizing Jensen's inequality as we did in the previous rate approximations. Equation can then be simplified to the following expression.

$$\bar{C}_k^{SUS}(K) = \int_0^1 \log_2\left(1 + \frac{z}{\frac{K}{P(N_T-1)} + \frac{K-1}{N_T-1}(1-z)}\right) p_Z(z).dz \tag{5.19}$$

In order to evaluate the above, we can perform a simple integration by parts. Let:

$$g(z) = \log_2\left(1 + \frac{z}{\frac{K}{P(N_T-1)} + \frac{K-1}{N_T-1}(1-z)}\right)$$

$$F_z(z) = \left(1-(1-z)^{N_T-1}\right)^{2^B} = \sum_{k=0}^{N_C} \sum_{i=0}^{k(N_T-1)} \binom{N_C}{k} \binom{k(N_T-1)}{i} (-1)^{i+k} z^i$$

F_z above is expanded using the binomial theorem, and $N_C = 2^B$ as before. Then, the integral in 5.19 can be expressed as:

$$\Big[F_z(z)g(x)\Big]_0^1 - \int_0^1 g'(z)F_z(z)dz \tag{5.20}$$

The first part of the expression can be evaluated simply by substitution and is given by:

$$\Big[F_z(z)g(x)\Big]_0^1 = \sum_{k=0}^{N_c}\sum_{i=0}^{k(N_T-1)}\binom{N_c}{k}\binom{k(N_T-1)}{i}(-1)^{i+k}\log\left(1+\frac{P(N_T-1)}{K}\right) \tag{5.21}$$

As for the integral term, we need to first evaluate $g'(z)$ using a combination of the chain rule and the quotient rule, we can work this out to be:

$$\begin{aligned} g'(z) &= \frac{\alpha_2}{(\alpha_1+\alpha_2(1-z))^2+z(\alpha_1+\alpha_2(1-z))}, \text{ where} \\ \alpha_1 &= \frac{K}{P(N_T-1)},\ \alpha_2 = \frac{K-1}{N_T-1} \end{aligned} \tag{5.22}$$

We can in fact write the denominator as a trinomial, $R = a + bz + cz^2$. a, b, and c in this case are constants. We can then express the integral term as shown:

$$\int_0^1 g'(z)F_z(z)dz = \frac{K}{P(N_T-1)}\sum_{k=0}^{N_C}\sum_{i=0}^{k(N_T-1)}\binom{N_C}{k}\binom{k(N_T-1)}{i}(-1)^{i+k}\int_0^1\frac{z^i}{R}.dz \tag{5.23}$$

$$R = a+bz+cz^2,\ a=(\alpha_1+\alpha_2)^2,\ b=(\alpha_1+\alpha_2)(1-2\alpha_2),\ c=\alpha_2(\alpha_2-1)$$

Now from 2.175 of the Table of Integrals, Series, and Products [53], we note that the integral can be solved as:

$$\int\frac{z^i}{R}.dz = \frac{z^{m-1}}{(1-i)c} - \frac{b}{c}\int\frac{z^{i-1}}{R}.dz - \frac{a}{c}\int\frac{z^{i-2}}{R}.dz \tag{5.24}$$

Supposing $A(i) = \int_0^1 \frac{z^i}{R}.dz$ The above can be solved recursively as follows:

$$\begin{aligned} A(i) &= \frac{z^{m-1}}{(1-i)c} - \frac{b}{c}A(i-1) - \frac{a}{c}A(i-2) \\ A(1) &= \frac{1}{2c}\log R - \frac{b}{2c}A(0) \\ A(0) &= \frac{1}{\sqrt{-\Delta}}\log\frac{\sqrt{-\Delta}-(b+2cz)}{\sqrt{-\Delta}+b+2cz} = A_0(\Delta) && \text{for } \Delta<0 \\ &= \frac{-2}{b+2cz} && \text{for } \Delta=0,\ b \text{ and } c\neq 0 \\ &= -A_0(-\Delta) && \text{for } \Delta>0 \end{aligned}$$

Where $\Delta = 4ac - b^2$. A simple program can than be executed to compute the solution to the above recursive relation. The special case when $K = N_T$ in fact simplifies the above leading to a closed form expression. It should be noted that closed form expressions are possible to be derived for $K < N_T$ as well, though they are tedious to derive. Finally, combining the solution of the recursive relation in equation (5.24) with (5.23), (5.21), and (5.20), we can then work out $\bar{C}_k^{SUS}(K)$.

Our new rate approximation performed by user k in order to determine its preferred mode can than be given by:

$$m_k = \arg \max_{1 \leq K \leq N_T} C_k^{SUS}(K) + (K-1)\bar{C}_k^{SUS}(K) \tag{5.25}$$

5.6.4 Simulation and Analysis

Now that SUS has been incorporated in the base-station and a large user SUS approximation has been incorporated at the user, how will the system perform? For the following simulations we still use a similar number of users and bits as we did when analysing SUS without the new rate approximation in order to keep the comparisons feasible. Intuitively what we would expect is that we would see benefits from using a SUS-based approximation. This is because, as we have mentioned earlier, the previous rate approximation underestimates the performance of the system when the base station utilizes SUS. In the SUS approximation we are no longer limiting ourselves to a lower bound when constructing the rate approximation and by doing so more users will prefer higher modes of operation which, when aided by SUS at the base station, would improve the performance of the system. The drawback however, is that we are using a large user approximation and analysing the system with moderate number of users. Therefore, In certain scenarios the SUS-approximation may overestimate the capability of the system to handle higher modes and this may cause performance drops when compared to the previous schemes. The following figures show the performance of the SUS-

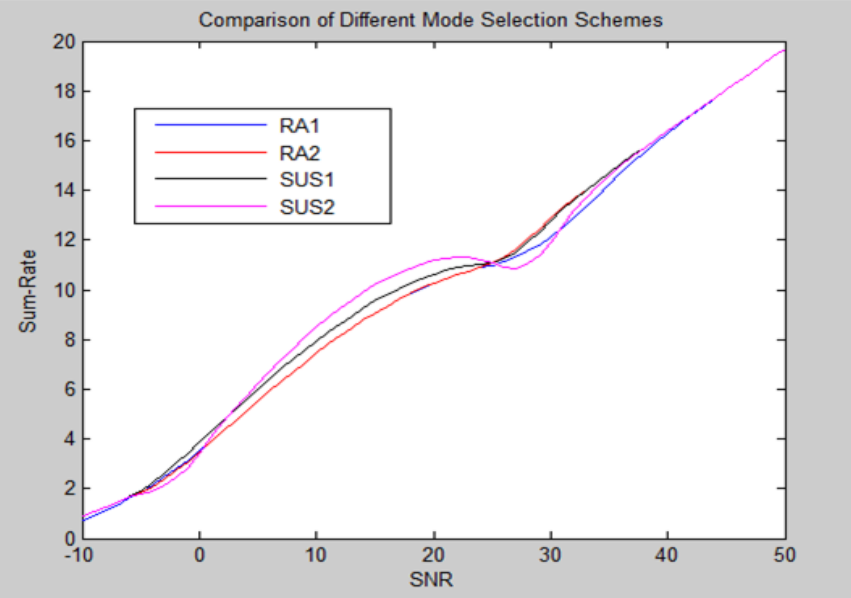

Figure 5.13: Comparison of the different schemes (latest in magenta) for 8 feedback bits 100 users

based rate approximation, overlayed on the previous three schemes that have been discussed thus far.

Fig 5.13 refers to the case with 8 feedback bits and 100 users, whereas Fig 5.14 is for 8 feedback bits and 200 users. The magenta curve represents the performance of the latest scheme. As per the intuition, we do see significant performance improvements in the 5-25 dB SNR regions for both scenarios. However we do see a noticeable performance drop especially in relation to the schemes using RA2 (black curve and red curve).

We can attempt to explain this performance drop as follows. Firstly, we notice for all the curves, the 30 dB mark shows an obvious change in the preferred mode of operation on average. This is noted by the change in the curves' behaviour. A similar transition is seen in the 40 dB mark where all the curves transition to single-user mode. So what causes the drop in performance in our latest scheme? In the transition regions, mode selection becomes of prime importance, as at these SNR regions the users are more

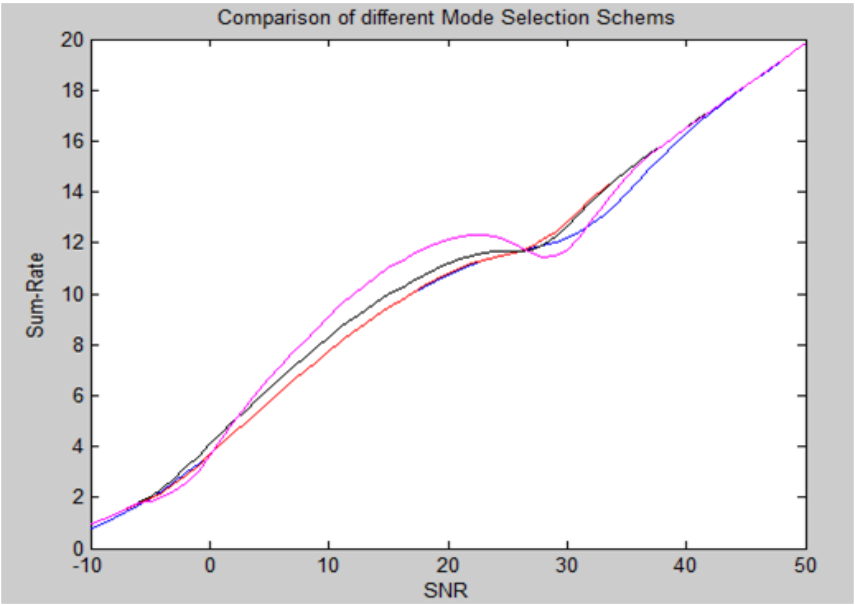

Figure 5.14: Comparison of the different schemes (latest in magenta) for 8 feedback bits 200 users

likely to be split in to groups by the base station. In other regions, the users unanimously prefer one mode group. Note that the SUS rate approximation works well when the number of users are large. In the transition regions, since the users are split into several groups, even if the total number of users is increased from 100 to 200, the number of users in each group is still relatively small. Therefore, the large user approximation will in fact overestimate the performance of the system, and allow for higher modes an opportunity to be selected, even though they are not optimal. With higher mode groups being selected, we experience more interference, since the base station is transmitting to more users and 30 dB is a relatively high SNR regime. SUS is used to help this situation, but again any selected group only has a few number of users in it. These reasons coupled together explain why the performance drops in the transition region.

Overall, RA1 which is adapted from [13] out-performs pre-existing similar schemes found in literature. The comparison between existing schemes

and the one proposed in [13] can be found in the respective paper. With the modified rate approximation, RA2, a slight improvement can be seen in the throughput, especially in the mid-SNR range. With the addition of SUS at the base station, this improvement is more apparent. However, the addition of SUS increases the computational complexity of the scheduling, and the slight throughput increase may or may not justify the trade-off. The main reason why the increase in throughput is not very drastic with the addition of SUS is that the rate approximation at the user side in SUS1 is not altered to take into account that the base station is using SUS. Upon using a large user approximation at the user-side, and rederiving an alternate rate approximation, we see that the improvement in SUS2 is much more significant in the mid-SNR regions. However, there are some performance drops in the transition regions owing to the fact that we are using a large user approximation for a setting with a relatively small number of users. The following sub section provides potential future research directions that could improve upon our results.

Proposed Future Improvements

A brief account of research directions that can be undertaken is summarized here.

Firstly, the base station can perform mode selection as per normal, with the SUS-based large user rate approximation. However, after chosing the preferred mode of operation, instead of performing SUS on $\mathcal{G}^{m^*}$ we could perform SUS on all the users. While this might increase the computational complexity of the SUS algorithm, in return we expect to see improvements in performance in the transition region due to the increase in the number of users.

Secondly, once the transition regions have been identified for a given system, a system specific solution would be to tune ϵ, the semi-orthogonality

forcing parameter. If we decrease ϵ in the transition regions, then the base station would be more inclined to schedule users from lower modes.

Thirdly, it would be good to derive a small user approximation instead, to suit our scenario, which is in fact simpler said than done. However, [54] gives some good leads on how to go about doing this. Suppose we start of with a unit-norm complex N_T-sphere, each point in the surface area of the sphere represents a possible channel direction. In our analysis, the channel direction has equal probability to point toward anywhere on this surface. Supposing, the SUS algorithm chooses its first user with a random channel direction, the next user has to be semi-orthogonal to the user that has just been selected. What this means is that we can no longer select the next user from the entire sphere, and must limit ourselves to a smaller strip, determined by ϵ. As the SUS process continues the surface area from which we choose users becomes significantly smaller. Therefore, the average SINR approximation computed by each user in the first iteration of SUS would not be the same at each subsequent iteration of the algorithm. This does not matter in a large user-setting, as even though after every iteration the pool of users we can choose from gets significantly smaller, it is still large enough to provide diversity in the selection. This does matter, however, in a small user-setting where we might quickly jump from choosing between 100 users in the first iteration to 20 in the second and maybe 5 on the third. In fact, supposing $N_T = 4$, we may not see very many instances where the base station schedules 4 users with a total of a 100 users. This is even more the case when we use mode selection in the transition regions where the 100 users are split into smaller groups before SUS is performed. Thus, a small user approximation is imperative in taking the next step toward more accurately approximating the rate of a SUS-based mode selection scheme.

Chapter 6

Conclusions

This thesis has proposed novel schemes to improve on the limitations of classical user scheduling in MIMO-broadcast channels with limited feedback. The main idea incorporated in order to overcome the performance limits of the classical schemes was to employ multi-mode transmission. Here, the technique involved users to approximate the sum-rate of the system by utilizing parameters known locally and treating unknown parameters as random variables and taking the expectation over them. By adopting the rate approximations for different modes of operation, the user can then specify its preferred mode of operation (the number of users co-scheduled). The base station receives not only the channel quality indicator (CQI) and channel directional information (CDI) information from the users, via limited feedback, but also their preferred mode of operation. The information received can then be utilized by the base station to determine how many users to schedule.

The idea of multi-mode scheduling is quite recent and some research has been done in this topic. This thesis has proposed improvements over existing noteworthy schemes, and more importantly formulated novel schemes. These schemes combine the grouping procedure used in [13] with the semi-orthogonal user scheduling algorithm proposed in [4]. The design in [13]

aims to reduce computational complexity at the base station by grouping users according to their "preferred mode" feedback. On the other hand, semi-orthogonal user selection (SUS) [4] aims to improve performance by avoiding users with aligned or poor channels. However, by incorporating these two schemes into the scheduling process, the base station addresses only one part of the problem. The other part of the problem is to formulate rate approximations at the user-side that could coincide with the given operations at the base station. To this end, we have derived two new closed form expressions for rate approximations at the user-side. When utilized with the scheduling process at the base station these rate approximations improve on the performance of pre-existing schemes. The two rate approximations correspond to the cases with and without SUS at the base station. An important observation via simulations is that rate approximations have to be tailored to the scheduling process at the base station. For instance, using the SUS-based rate approximation at the user when the base station does not use SUS will result in poor performance.

While the enhanced multi-mode scheme proposed in the thesis does improve performance in most of the mid-high SINR region, it experiences performance drops in certain regions when compared to earlier schemes proposed in literature. The reason for this is that the SUS-based rate approximation that was used was for a large user-set, while the target user-set in the simulations consisted of only twenty to about two hundred users. Deriving closed form expressions for rate approximations that accommodate such smaller user-sets is of prime importance in overcoming the slight performance deteriorations as well as further increasing the performance gains. This is left for future research along with the many design considerations at the base station, discussed in Chapter 5. In summary, this thesis has covered the evolution of the basic point-to-point wireless communication model to the MIMO broadcast channel. It has explained the shortcomings of limited

feedback schemes currently used in literature and has proposed improvements on existing multi-mode transmission schemes as well as shown closed form derivations of rate approximations at the user-side. The contributions of the thesis is of practical value since these closed form derivations not only aid to model and compare the discussed rate approximations but from a practical perspective, they also form gateways to algorithms that can be implemented in real-life devices.

Bibliography

[1] M. Katz and F.H.P. Fitzekl. *WiMAX Evolution: Emerging Technologies and Applications.* Chichester, U.K.: Wiley & Sons, 2009.

[2] P. Rysavy. *Mobile Broadband Explosion.* Rysavy Research White Paper, 4G Americas, 2012.

[3] M. Mustaqim, K. Khan, and M. Usman. LTE-advanced: Requirements and technical challenges for 4G cellular network. *Journal of Emerging Trends in Computing and Information Sciences*, 3(5):665 – 671, May 2012.

[4] T. Yoo, N. Jindal, and A. Goldsmith. Multi-antenna downlink channels with limited feedback and user selection. *IEEE Journal of Selected Areas in Communications*, 25(7):1478 – 1491, 2007.

[5] N. Jindal. MIMO broadcast channels with finite rate feedback. *IEEE Transactions on Information Theory*, 52(11):5045–5060, 2006.

[6] D.Gesbert, M. Kountouris, Jr. R.W. Heath, C. Chae, and T. Salzer. From single user to multiuser communications: Shifting the MIMO paradigm. *IEEE Signal Processing Magazine*, 24(5):36 – 46, 2007.

[7] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication.* Cambridge University Press, New York, 2005.

[8] T. Cover. Broadcast channels. *IEEE Transactions of Information Theory*, 18(1):2 – 14, Jan. 1972.

[9] C. Zhang, W. Xu, and M. Chen. Hybrid zero-forcing beamforming/orthogonal beamforming with user selection for MIMO broadcast channels. *IEEE Communication Letters*, 13(1):10 – 12, 2009.

[10] P. Viswanath and D.N.C. Tse. Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality. *IEEE Journal on. Selected Areas in Communications*, 49(8):1912 – 1921, August 2003.

[11] N. Jindal, W. Rhee, S. Vishwanath, S.A. Jafar, and A. Goldsmith. Sum power iterative water-filling for multi-antenna Gaussian broadcast channels. *IEEE Transactions on Information Theory*, 51:1570 – 1580, April 2005.

[12] X. She and L. Chen. Precoding and scheduling techniques for increasing capacity of MIMO channels. *NTT DOCOMO Technical Journal*, 10(4):38 – 44, 2009.

[13] W. Xu and X. Dong. Enhanced multi-mode transmission by user scheduling in MISO broadcast channels with finite-rate feedback. *Wireless Personal Communications*, 65(1):103 – 123, 2012.

[14] C.E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379 – 423, July, Oct. 1948.

[15] R. G. Gallager. Low density parity check codes. *Monograph, M.I.T. Press, Cambridge, Massachusetts*, 1963.

[16] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon limit error-correcting coding and decoding: Turbo-codes. In *IEEE International Communications Conference*, 1993.

[17] D. J. C. MacKay and R. M. Nealt. Near Shannon limit performance of low density parity check codes. *IEEE Transactions on Communications*, 32(18):1645 – 1646, 1996.

[18] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6(3):311 – 335, Mar. 1998.

[19] E. Telatar. Capacity of multi-antenna Gaussian channels. *Transactions on Telecommunications*, 10(6):585 – 598, Nov. 1999.

[20] H.Weingarten, Y. Steinberg, and S. Shamai. The capacity region of the Gaussian MIMO broadcast channel. In *38th conference information sciences and systems (CISS04)*, 2004.

[21] G. Caire and S. Shamai. On the achievable throughput of a multi-antenna Gaussian broadcast channel. *IEEE Transactions on Information Theory*, 49(7):1691 – 1706, 2003.

[22] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath. Capacity limits of MIMO channels. *IEEE Journal on Selected Areas in Communications*, 21(5):684 – 702, 2003.

[23] M. Sharif and B. Hassibi. On the achievable rates of multiple antenna broadcast channels with feedback-link capacity constraint. *EURASIP Journal on Wireless Communications and Networking*, 2011(21):1–16, 2011.

[24] LTE World. *MIMO Transmission Schemes for LTE and HSPA Networks*. 3G Americas Report, 2009.

[25] N. Ramanathan, F. Li, M. Kuijper, and J. Evans. Performance of multi-mode transmission with finite rate feedback in MIMO broad-

cast systems. In *IEEE Australian Communications Theory Workshop AusCTW2013*, 140-145, 2013.

[26] B. Khoshnevis and W. Yu. A limited-feedback scheduling and beamforming scheme for multi-user multi-antenna systems. In *IEEE Global Telecommunications Conference (GLOBECOM 2011)*, 2011.

[27] P. Ding, D. J. Love, and M.D. Zoltowski. Multiple antenna broadcast channels with shape feedback and limited feedback. *IEEE Transactions on Signal Processing*, 55(7):3417–3428, 2007.

[28] X. Chen, W. Miao, Y. Li, S. Zhou, and J. Wang. On the achievable rates of multiple antenna broadcast channels with feedback-link capacity constraint. *EURASIP Journal on Wireless Communications and Networking*, 21(1):1–16, 2011.

[29] R. D. Wesel and J. M. Cioffi. Achievable rates for Tomlinson-Harashima precoding. *IEEE Transactions on Information Theory*, 44(2):824–831, Mar. 1998.

[30] M. H. M. Costa. Writing on dirty paper. *IEEE Transactions on Information Theory*, 29(3):439 – 441, 1983.

[31] Z. Tu and R. S. Blum. Multiuser diversity for a dirty paper approach. *IEEE Communications Letters*, 7(8):370 – 372, 2003.

[32] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran. Multiuser MIMO achievable rates with downlink training and channel state feedback. *IEEE Signal Processing Magazine*, 56(6):2845 – 2866, 2010.

[33] A.S. Burashed, H. E. A. Hassan, and M. M. Salahe. Performance evaluation of the parallel interference cancellation with a regularized zero forcing in the frequency domain. *International Journal of Research and Reviews in Computer Science (IJRRCS)*, 2(5):1193–1198, 2011.

[34] M. M. Shanechi, R. Porat, and U. Erez. Comparison of practical feedback algorithms for multiuser MIMO. In *IEEE Vehicular Technology Conference (VTC Spring)*, 2009.

[35] B. Clerckx, G. Kim, and S. Kim. MU-MIMO with channel statistics-based codebooks in spatially correlated channels. In *IEEE Global Telecommunications Conference (GLOBECOM 2008)*, 2008.

[36] N. Ravindran and N. Jindal. Limited feedback-based block diagonalization for the MIMO broadcast channel. *IEEE Journal on Selected Areas in Communications*, 26(8):1473–1482, 2008.

[37] D. Tse, P. Viswanath, and L. Zheng. Diversity-multiplexing tradeoff in multiple-access channels. *IEEE Transactions on Information Theory*, 50(9):1859 – 1874, 2004.

[38] M. Kountouris, R. de Francisco, D. Gesbert, D. T. M. Slock, and T. Slzer. Exploiting multiuser diversity in MIMO broadcast channels with limited feedback. *France Telecom R&D Internal Report*, March 2006.

[39] 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*, 52(2):461 – 471, 2004.

[40] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst. A behavioral approach to linear exact modeling. *IEEE Transactions on Communications*, 53(3):537 – 544, Mar. 2005.

[41] T. Yoo and A. Goldsmith. On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming. *IEEE Journal on Selected Areas in Communications*, 24(3):528 – 541, Mar. 2006.

[42] S. Vishwanath, N. Jindal, and A. Goldsmith. Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels. *IEEE Transactions on Information Theory*, 49(6):2658 – 2668, 2003.

[43] R. Zamir, S. Shamai, and U. Erezi. Nested linear/lattice codes for structured multiterminal binning. *IEEE Transactions on Information Theory*, 48(6):1250 – 1276, June 2002.

[44] R. Knopp and P.A. Humblet. Information capacity and power control in single-cell multiuser communications. In *IEEE International Conference on Communications (ICC)*, June 1995.

[45] P. Viswanath, D.N.C. Tse, and R. Laroia. Opportunistic beamforming using dumb antennas. *IEEE Transactions on Information Theory*, 48(6):1277 – 1294, 2002.

[46] K.K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang. On beamforming with finite rate feedback in multiple-antenna systems. *IEEE Transactions on Information Theory*, 49(10):2562–2579, 2003.

[47] C. K.Au-Yeung and D.J. Love. On the performance of random vector quantization limited feedback beamforming in a MISO system. *IEEE Transactions on Wireless Communications*, 6(2):458–462, 2007.

[48] D.J. Love. Grassmannian beamforming for multiple-input multiple-output wireless systems. *IEEE Transactions on Information Theory*, 49(10):2735 – 2747, 2003.

[49] D. J. Love, R. W. Heath, and T. Strohmer. Grassmannian beamforming for multiple-input multiple-output wireless systems. *IEEE Transactions on Information Theory*, 49(10):2735 – 2746, 2003.

[50] E. Bjornson, K. Ntontin, and B. Ottersten. Channel quantization design in multiuser MIMO systems: Asymptotic versus practical conclusions.

In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2011.

[51] C. Au-Yeung, S.Y. Park, and D.J. Love. A simple dual-mode limited feedback multiuser downlink system. *IEEE Transactions on Communications*, 57(5):1514 – 1522, 2009.

[52] J. Zhang, M. Kountouris, J. G. Andrews, and R. W. Heath Jr. Multimode transmission for the MIMO broadcast channel with imperfect channel state information. *IEEE Transactions in Communications*, 59(3):803 – 814, 2011.

[53] I.S. Gradshteyn and I.M. Ryzhikl. *Table of Integrals, Series, and Products.* A.l. Jeffrey and D. Zwillinger(editors), New York: Academic Press, 2007.

[54] W. Ni, Z. Chen, H. Suzuki, and I. B. Collings. On the performance of semi-orthogonal user selection with limited feedback. *IEEE Communications Letters*, 15(12):1359 – 1361, 2011.