ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

ELECTRICAL AND COMPUTER ENGINEERING DEPARTMENT

Pre-coding for MIMO Broadcast Channels Using Dirty Paper Coding Techniques

By

Getnet Kassa

Advisor

Dr. Eneyew Adugna

A Thesis Submitted to the School of Graduate Studies of Addis Ababa University in Partial Fulfillment of the Requirements for the Degree of Masters of Science in Electrical Engineering

July, 2011

Addis Ababa, Ethiopia

Abstract

Wireless communication systems with multiple antennas are the focus of many applications nowadays due to the higher throughput and/or more robust performance than single antenna communication. One variety of these systems is the multi-user system in which many users share the same wireless environment. In this thesis, communication strategies that use dirty paper coding (DPC) techniques to form independent spatial streams to the Multiple Input Multiple Output broadcast channel (MIMO BC) are studied. Unlike previous studies that assume ideal channel conditions, the performance of DPC algorithms under frequency selective channels having spatially correlated path gains is investigated. These algorithms are found to be equally applicable to frequency selective channels if orthogonal frequency division multiplexing (OFDM) is employed with 1.5 dB power loss due to correlation for realistic correlation values.

The second part of the thesis considers issues that arise in practical implementation of the DPC algorithms. The effect of imperfect channel estimates at the transmitter and the problem of scheduling users for transmission are considered. For time division duplex (TDD) systems, we show that errors in channel estimation result in performance degradation that can almost completely be eliminated above a SNR of 25 dB. For frequency division duplex (FDD) systems, it is shown that delay in estimation feedback results in severe performance degradation that becomes unacceptable above some delay depending on the algorithm used. User scheduling is formulated as user selection for single carrier systems. We adapt an algorithm proposed for linear pre-coding techniques and show there is an increase in sum rate in using this algorithm over uniform scheduling of users. For OFDM systems, applying the user selection algorithm proposed for single carrier systems to each subcarrier results in increased sum rate with lower fairness index compared to uniform scheduling. Allowing users to select their strongest subcarriers, on the other hand, is shown to give lower sum rate and improved fairness compared with the sum rate maximizing algorithm.

Key words: Dirty Paper Coding, MIMO Broadcast Channel, Pre-coding

Acknowledgement

I feel deeply indebted to many people from whom I received help and support in doing this work. Foremost, I would like to express my sincere gratitude to my advisor Dr. Eneyew Adugna for his full support and guidance. I am also greatly thankful to all of my friends and family whose mental support and love gave me endless comfort.

Contents

List of Tables

List of Figures

List of Acronyms

AWGN	Additive White Gaussian Noise
BC	Broadcast Channel
BER	Bit Error Rate
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSIT	Channel State Information at the Transmitter
DPC	Dirty Paper Coding
FDD	Frequency-Division Duplex
LOS	Line of Sight
MAC	Multiple Access Channel
MIMO	Multiple Input Multiple Output
MMSE	Minimum Mean Squared Error
MU-MIMO	Multi-User MIMO
QoS	Quality of Service
SDMA	Space Division Multiple Access
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single Input Single Output
SNR	Signal-to-Noise Ratio
TDD	Time-Division Duplex
TDMA	Time-Division Multiple-Access
THP	Tomlinson Harashima Precoding
VP	Vector Perturbation
ZFBF	Zero-Forcing Beamforming

Chapter 1

Introduction

Wireless communication systems are showing a rapid growth over the last few decades transforming themselves into a necessary part of our daily lives. Their functionality ranges from cellular telephony and messaging to browsing the internet and other multimedia services over wireless portable devices like PDAs and laptops. In the future, these systems are expected to deliver farther connectivity and transmission of real time multimedia content at any time and place. In spite of these developments, there is an ever increasing demand for new services and better quality for the existing ones. Meeting these demands requires an improved usage of the scarce resources, primarily bandwidth and power.

1.1 Background

The past decade saw development and utilization of several new concepts in communication theory that have helped to meet the increased data rate and QoS requirements. These include multiple-input multiple-output (MIMO) systems and multiuser diversity. In MIMO systems, multiple antennas are deployed at both the transmitter and the receiver sides. Combined with advanced signal processing and coding techniques, MIMO systems are capable of delivering higher data rates and/or more robust communications. The invention of MIMO can be viewed as the introduction of a new dimension, namely the antenna or space domain. This new dimension can be used either to improve the data rate and QoS of a single user accessing the channel by assigning all the spatial dimensions (i.e. antennas) to the user, or to serve more users by assigning the spatial dimensions to different users [1].

When a MIMO channel is used in the latter manner (i.e. when antennas are shared among multiple users in a given channel), the channel is referred to as a multiuser MIMO (MU-MIMO) channel. Among important MU-MIMO channels are MIMO broadcast channels (BCs) and MIMO multiple-access channels (MACs). A BC is an abstract model of a downlink transmission

where one transmitter (base-station or access point) serves multiple receivers (mobile users) simultaneously.

In MIMO BC, the signal sent for one user is also received by other user terminals in the system creating inter-user interference. This challenge of serving multiple users can be efficiently accomplished by assigning different spatial streams for each user [1] and/or joint encoding/decoding across users [2]. Thus, although the users may be assigned the same channel (i.e. the same time/frequency/code resource), they are separable in space. This type of antenna-aided multiuser communication scheme for MIMO BC channels is commonly called space-division multiple-access (SDMA) and will be used throughout this thesis.

In using SDMA for MIMO BC, assigning different spatial streams requires separating the signal intended for a given user from signals of other users in the system. This can't be accomplished at the users' receivers because these receivers are located at different places. Thus, the signal received by one user is not available at the other user's terminal which makes joint processing at the receiver side impossible.

This brings about the need to process the users' signals before transmission with the aim of delivering to each user the signal intended for it only, and this approach is commonly known as pre-coding. One technique that is shown to achieve the capacity of MIMO BC is dirty paper coding (DPC) [2]. This technique is analogous to writing on a dirty paper and hence its name. If the person writing on the dirty paper has perfect knowledge of the dirt to be picked up on the way to the destination, he/she can apply ink by taking this dirt into consideration. This, in effect, creates the same space for writing as a clean paper as long as the writer knows the dirt to be picked up beforehand.

The capacity of a channel with output $Y = X + S + Z$ is examined in [3] where X is the input, S is Gaussian distributed interference known at the transmitter and Z is additive white Gaussian noise. It is shown that the capacity of the channel is the same as a channel without any interference S even though S is unknown at the receiver. Exploiting the capacity promised by dirty paper coding combines all elements of signal design into one step and thus requires completely new communication protocols. But there exist suboptimal approaches that separate

pre-coding from other parts of the system [4, 5, 6] while still achieving considerable performance improvement over conventional linear pre-coding techniques [7].

1.2 Motivations

Studies in MIMO BCs mostly focused on information theoretic aspects of the system. The unavailability of algorithms that approach the capacity limit also forced researchers to consider the linear processing algorithms as an alternative to the DPC techniques. Noting that the suboptimal implementations of DPC have improved performance, they can be incorporated in practical systems if their complexity is investigated and compared with the performance improvement they attain. To determine their sensitivity to channels with non ideal conditions, the performance of the algorithms in these situations should also be studied before practical implementation.

Another concern when it comes to the capacity is its requirement of full channel knowledge (referred to as channel state information (CSI) throughout the thesis) at both communication ends. While this requirement may be feasible in communications over a wired medium as the channels therein are semi-static, obtaining accurate CSI is difficult in typical wireless communication systems due to the time-varying nature of a wireless medium and the mobility of users. This introduces an additional challenge for MIMO BCs since DPC schemes rely on the assumption of perfect CSI.

Typical application of DPC techniques like wireless LANs usually have much more users than the number of antennas that can be deployed at the transmitting station (access point in this case). Since spatial streams only as many as the number of transmit antennas can be created, the number of users that can be served at a time is limited by the number of transmitting antennas. Thus, scheduling of users to share the available resources is another issue that needs consideration in practical systems employing DPC techniques.

1.3 Objectives

General Objective

Motivated by the aforementioned limitations, this work focuses on investigating the applicability of existing DPC algorithms by considering their performance in non ideal channel conditions. It also considers other system aspects that could prohibit implementation of this technique in practical systems.

Specific Objectives

- Study the applicability of DPC algorithms for frequency selective channels having correlation between fading paths and comparing their performance with linear processing algorithms in terms of bit error rate (BER) and achievable sum rate.
- Investigate the tradeoff between complexity and performance improvement in using these algorithms.
- Analysis of the effect of imperfections in channel state information at the transmitter (CSIT) on BER performance of the DPC algorithms for time division duplex (TDD) and frequency division duplex (FDD) systems.
- Propose user scheduling techniques to be applied when there are more users than the number of transmit antennas for single carrier and multicarrier system that use DPC algorithms.

1.4 Literature Review

The name ‘dirty paper coding’ comes from the title of a 1983 paper by Max Costa, “Writing on Dirty Paper” [3] on the capacity of a Gaussian channel having interference that is known to the transmitter. Costa presented the surprising result that the capacity of this system is the same as if there were no interference present. One important application of this result is to a multiuser channel wherein a multiple-antenna transmitter is communicating with multiple users. Using Costa’s result, Yu and Cioffi [26] found the sum capacity (or highest sum of rates achievable) for this system. In addition to these theoretical results, different researchers are trying to come up

with algorithms that attain the promised capacities. One approach is based on modified version of Tomlinson Harashima pre-coder (THP), which is more like a decision feedback equalizer on the transmitter side [5], [6]. Other techniques based on quantizing to a lattice have been identified as an important tool for achieving capacity [8]. A technique that uses one dimensional lattice to perform perturbation is given in [4]. Other than the dirty paper approach, different pre-coders that use linear processing have been studied in [9]. One common feature of these works is that they all assume the channel is narrowband and elements of the channel matrix are uncorrelated with each other.

In [10], the effect of feedback delay of channel estimates on the sum rate performance of a multi-user MIMO OFDM system is studied, and it is shown that the linear algorithms are more sensitive to delay than the DPC algorithms. The effect of delay in channel estimation feedback on the BER performance of single user systems is studied in [11]. The authors also suggested a prediction technique to reduce the effect of delay in CSIT using the MMSE criterion. A numerical analysis of the effect of channel estimation error on the sum rate performance of MIMO Broadcast systems is studied in [12] for different values of error variances.

The use of orthogonality among users to perform permutation of user indices is shown to reduce the BER of Modified THP in [13]. A similar algorithm with enhanced user selection criterion is applied for ZFBF in [14]. In [15], the authors investigated the increase in sum rate of vector perturbation by using an algorithm similar to those in [14] and [13]. While these user selection algorithms are proposed for single user systems, it is shown in [16] that it is possible to reduce the total transmit power by performing subcarrier allocation over spatial and frequency domain. In [17], two approaches in subcarrier allocation are shown to result in sum rate increment and achievement of better fairness for multiplexing based MIMO OFDM systems.

1.5 Methodology and Scope

Methodology

The methodology employed to achieve the desired objectives started with a review of literatures that lay the foundation for this work. Taking the ideas obtained in the literature survey, modeling the system is carried out followed by selection of appropriate model parameters. In this phase, the models used to study the effect of fading path correlation and imperfections in CSIT are also developed. Then, implementation of the DPC algorithms and user scheduling algorithms for the selected model parameters is carried out. Using the models developed, simulation of the overall system is performed using MATLAB software as a simulation tool. Analysis of the performance of the DPC algorithms in different channel conditions and implementation impairments is performed thereafter. Finally, the results are interpreted and conclusions drawn based on the results obtained.

Scope

The performance analysis in this thesis work is based on simulation of the proposed techniques using MATLAB. Thus, hardware implementation details of the resulting systems are not considered. The MIMO broadcast system is not implemented in practical systems yet. Therefore, the simulation is not based on any existing standard although current wireless communication systems are considered in selecting simulation parameters. Other issues that are related to but not dealt in this work are listed as recommendations for future work in Section 6.2.

1.6 Outline of the Thesis

We begin the next chapter by providing an introductory overview on features and advantages of MIMO systems. We then concentrate on the multiuser type of MIMO communications with more elaborate discussion of MIMO broadcast channels (BCs). A brief discussion of the formulation of capacity for MIMO BCs will follow. Then, the system model to be used for the rest of the thesis will be introduced including models for the channel and the effect of correlated path gains. Obtaining mechanisms and causes of errors in CSI at the transmitter will be discussed and the chapter concludes with an overview of orthogonal frequency division multiplexing (OFDM).

In Chapter 3, we start with general formulation of the pre-coding problem for MIMO BCs followed by a description of a commonly used linear processing algorithm. The DPC technique will be discussed in detail with a description of existing algorithms that follow the dirty paper approach. Then, mechanism of performing pre-coding for multicarrier systems will be shown. Finally, the problem of scheduling users will be formulated and the proposed algorithms for single carrier and multicarrier systems will be discussed.

The next two chapters focus on the results of the work that has been carried out. Chapter 4 concentrates on evaluating the performance of the DPC algorithms for wideband channels with and without correlation among fading paths. An elaborate discussion of the system setup used for simulation is presented first and the results of the simulations for coded and un-coded systems follow afterwards.

In the previous chapter, knowledge of perfect and up to date channel states at the transmitter is assumed. But in Chapter 5, the results will be extended when there are imperfections in channel estimate at the transmitter. Modeling the imperfections for TDD and FDD systems is described followed by the performance of the DPC algorithms in these scenarios. The sum rate performance and fairness of the user scheduling algorithms is also investigated.

Lastly, in Chapter 6, we summarize the results and suggest future research directions that can consolidate this work.

1.7 Notations

The notations used in this thesis are as follows. We use upper case boldface letters for matrices and lower case boldface for vectors. Depending on the context, a vector can be either a column vector or a row vector. Subscripts are used to represent an element of a vector or a matrix. For example, x_i and X_{ij} represents the i^{th} element of $\boldsymbol{x}$ and the $(i,j)^{th}$ element of $\boldsymbol{X}$, respectively. Note how the scalar and vector/matrix quantities are differentiated in the usage x_i, $\boldsymbol{x}$, and $\boldsymbol{X}$. This will prevent any confusion when subscripts are used for other indexing purposes. For example, $\boldsymbol{x}_k$ represents the k^{th} vector of a set of vectors while x_k represents the k^{th} element of a vector $\boldsymbol{x}$. $|x|$ is the absolute value of x, $\|\boldsymbol{x}\|$ is the Euclidian norm of a vector $\boldsymbol{x}$. $\boldsymbol{X}^T$ ($\boldsymbol{x}^T$) stands for the transpose of a matrix $\boldsymbol{X}$ (vector $\boldsymbol{x}$), and $\boldsymbol{X}^H$ ($\boldsymbol{x}^H$) stands for the conjugate transpose of a matrix $\boldsymbol{X}$ (vector $\boldsymbol{x}$). For a random variable X, $\mathbb{E}(X)$ denotes the expected value of X.

Chapter 2

MIMO Broadcast Channels

Nowadays, wireless channel is being widely studied due to its flexibility and fully unexploited performance gains. This is partly due to the increase in signal processing power and the resulting degree of freedom in implementing complex algorithms that help in achieving information theoretic capacities in different wireless scenarios. In a wireless channel, the transmitter broadcasts electromagnetic waves that make their way to the receiver through free space propagation, reflection, diffraction and scattering. This results in multipath fading where the signals from the transmit antennas add constructively or destructively at the receive antenna resulting in burst errors. Transmitted signals may also be blocked by terrain, buildings or trees which causes shadowing of signals. In addition, noise is added to the transmitted signal. Thus, a wireless channel is mainly characterized by multipath, shadowing and additive noise [19].

While the effect of noise can be reduced by appropriate modulation and coding, techniques for reducing the effect of multipath fading have been suggested and some implemented in practical systems. Diversity combining is a powerful and commonly used technique to mitigate the effect of multipath fading and can be of space, polarization, frequency, time or directional diversity.

Space diversity is efficient and relatively easier to incorporate in practical systems. This technique uses sufficiently spaced multiple antennas at the transmitter and/or receiver to obtain independently fading signal paths which are then efficiently combined so that the fading of the resultant signal is reduced [9, 18]. Receive diversity, which use multiple antennas at the receiver, has an additional gain called power gain resulting from the coherent combining of multiple received signals [19].

2.1 Multiple Antenna Systems

In addition to increasing link reliability, multiple antennas both at the transmitter and receiver help in increasing the data rate by sending multiple streams of data at the same time and

frequency. This offers an added degree of freedom and is called ***multiplexing gain*** [20]. This is possible because the structure of the channel can be exploited to obtain independent signaling paths with the receive antennas receiving unique combinations of the transmitted data streams. Advanced digital signal processing algorithms can be used to recover the original data streams. These systems are called Multiple Input Multiple Output (MIMO) systems. A MIMO system with M transmit and N receive antennas is shown in Figure 2.1.

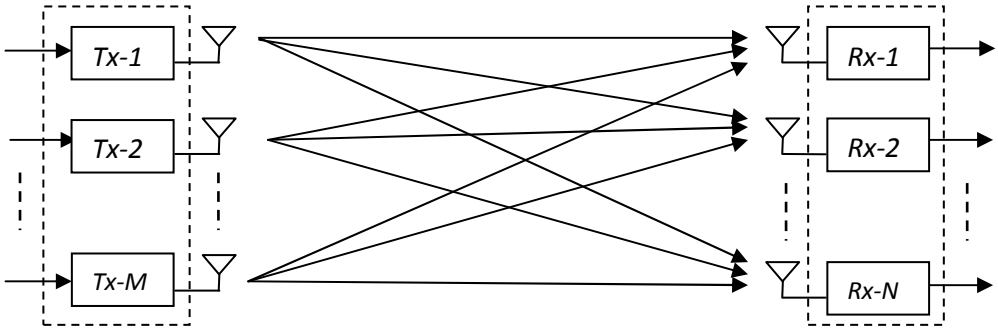

Figure 2.1: Schematic representation of MIMO System

While diversity gain doesn't necessarily need information on the state of the channel, multiplexing gain of a MIMO system can only be exploited if the channel is known at the receiver and sometimes at the transmitter too. Both diversity and multiplexing gain of MIMO systems are more pronounced in wireless channels where there is a rich scattering environment to create more spatial variation or dimensionality. The price to be paid for the performance enhancements obtained through MIMO techniques is the added cost of deploying multiple antennas, the space and power requirements of these extra antennas, and the added complexity required for multi-dimensional signal processing.

The multiple antennas in MIMO systems can be used for multiplexing gain only thereby increasing the data rate. The other mechanism is to apply beamforming techniques to coherently combine the channel gains and get a robust channel with high diversity gain. It is not a necessity to use the antennas purely for multiplexing or diversity, and some antennas can be grouped for diversity with each group is used to send independent data streams. This presents a tradeoff

between data rate, error probability and complexity in MIMO system design where the optimal approach depends on the problem at hand.

2.1.1 Single-user Systems

In single user MIMO systems, the transmit terminals work together to transmit the best signal, and the receive terminals also cooperate for demodulation and decoding [21, 22]. When both the transmitter and the receiver know the channel, an optimal transmission technique is to transform the channel into diagonal scalar channels using the singular value decomposition (SVD) of the channel matrix [20]. To achieve the capacity of single user MIMO system with CSIT and CSIR, power should be allocated to the parallel channels based on their strength. The optimal power allocation strategy is found to be the waterfilling approach [19]. Optimal decoding of the received signals requires ML demodulation which performs exhaustive search over all possible input vectors.

When there is no CSIT, the SVD approach can't be used. In this case, the optimal transmit all symbols with equal power and decode the received signal using maximum likelihood (ML) detector using CSIR. Other suboptimal approaches include V-BLAST among others that use successive interference cancellation at the receiver side [23, 24].

2.1.2 Multi-user Systems

Traditionally, in systems with multiple users, multiple access has been achieved by time-division, frequency-division, or code-division multiplexing, known respectively as TDMA, FDMA, and CDMA. The use of antenna arrays at the transmit station of such systems has resulted in Space-Division Multiple Access (SDMA), in which the spatial diversity of the signals received at the base station are used to separate signals that may be transmitted using the same time, frequency, or code sequence from different users. Unlike single-users MIMO systems, either the transmitter or the receiver terminals do not cooperate in multi-user MIMO systems.

There are two basic multi-user MIMO channel models: the MIMO multiple-access channel (MAC) and the MIMO broadcast channel (BC). In MIMO MAC, a number of users share a

common communication channel to transmit their individual signals to a receiver. Such a system is shown in Figure 2.2.

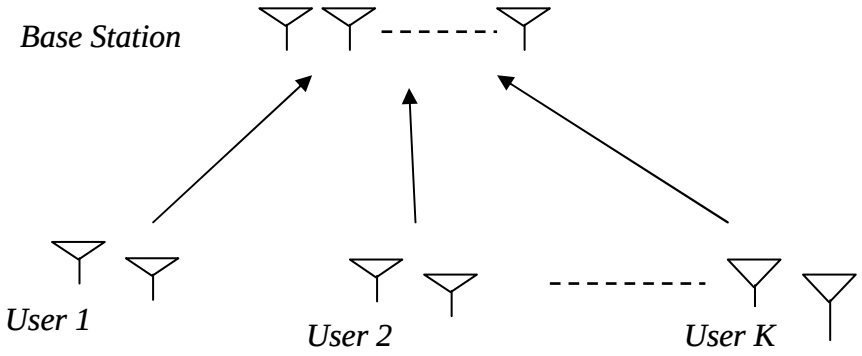

Figure 2.2: MIMO Multiple Access Channel

In the uplink of a mobile cellular communication system, the users are the mobile transmitters in any particular cell and the receiver is the base station of that cell. In MIMO BC, a transmitter sends information to multiple receivers as shown in Figure 2.3. In the downlink of a mobile cellular communication system, the transmitter is the base station and the receivers are the mobile stations. Another example of MIMO BC arises when a wireless LAN access point transmits simultaneously to several user terminals.

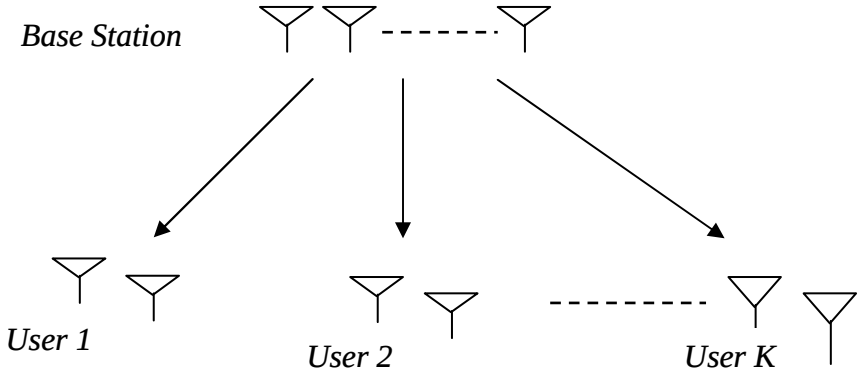

Figure 2.3: MIMO Broadcast Channel

A key difference between single-user, MAC, and BC channels is that in the single-user channel, there is a full collaboration at both sides of transmitter and the receiver, while in the MAC channel there is collaboration only at the receiver and in the BC channel collaboration exists only at the transmitter. Therefore in the MIMO BC channel joint processing between the receivers cannot be supported because the signal received by one of the receivers is not available at the others. Based on this fact, the design of MIMO BC channel has proved to be more challenging [25, 26].

2.2 MIMO Broadcast Channels

In this section, a brief overview of the system model to be used for the rest of the work will be covered. The channel models and the mechanisms of obtaining CSIT will also be discussed.

2.2.1 System Model

Consider a MIMO broadcast system with the transmitting station having M antennas that transmit data to K users each having single receive antenna. Using the discrete time, narrowband and time varying channel model, the MIMO broadcast system at a specific time can be described by:

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

In matrix form,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_K \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1M} \\ h_{21} & h_{22} & & h_{2M} \\ \vdots & & \ddots & \vdots \\ h_{K1} & h_{K2} & \cdots & h_{KM} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_K \end{bmatrix} \tag{2.2}$$

where $\boldsymbol{x} \in \mathbb{C}^{M \times 1}$ is the channel input from the transmitter, $\boldsymbol{H} \in \mathbb{C}^{K \times M}$ is the discrete-time channel matrix, $\boldsymbol{y} \in \mathbb{C}^{K \times 1}$ is the received signal, and $\boldsymbol{n} \in \mathbb{C}^{K \times 1}$ is a vector of additive noise. The $(i,j)^{th}$ element of $\boldsymbol{H}$ represents the channel gain from transmit antenna j to user i.

Throughout this thesis, we consider Gaussian channels where the additive noise $\boldsymbol{n}$ is given by a Gaussian vector. We will further assume that the noise is temporally white, i.e. the entries of $\boldsymbol{n}$ are independent, identically distributed (i.i.d) and zero-mean circularly symmetric complex Gaussian (ZMCSCG) with unit variance, i.e. $\mathbb{E}(\boldsymbol{n}\boldsymbol{n}^*) = \boldsymbol{I}$. We use a simple channel model where the channel gain from a transmit antenna to a receive antenna is described by a ZMCSCG random variable, which is an appropriate model for narrowband systems operating in a non-line-of-sight rich scattering environment [27]. We normalize the channel such that the entries of $\boldsymbol{H}$ have unit variance.

2.2.2 Channel Models

The physical characteristics of the wireless channel present a fundamental technical challenge for reliable communications. This is mainly because of the time varying multipath nature of the channel. Since the channel time variations are not predictable, the time variant multipath channel is modeled statistically. The commonly used statistical fading models are Rayleigh fading model for non LOS conditions and Rician fading model for LOS conditions [28].

The coherence bandwidth of the channel is defined as the frequency separation at which two frequency components of the signal undergo independent attenuations by the channel. If the bandwidth of the signal that is transmitted through the channel is smaller than the coherence bandwidth of the channel, the channel is called flat-fading channel in which all the frequency components of the signal undergo the same attenuation by the channel. In this case, the multipath components in the received signal are not resolvable, and the channel appears as a single fading path that can be represented by a single complex coefficient.

If the signal bandwidth is greater than the coherent bandwidth the channel, the channel is said to be frequency-selective and the signal is severely distorted by the channel. In this case, the multipath components can be resolved in the received signal and therefore the receiver is provided with several independently fading signal paths [28]. Consequently, the frequency-selective channel is modeled as a tapped delay line filter with time-variant tap coefficients. The frequency-selective fading can degrade system performance by causing inter-symbol interference

(ISI) which can be minimized by using orthogonal frequency division multiplexing (OFDM) [29]. Analytically, the frequency-flat and frequency-selective channels are also called narrowband and wideband channels when defined in comparison with the bandwidth of the signal used in the system at hand.

There are two channel models used in this work: Quasi-static wideband and Time-varying narrowband channel models. This is because each model has the particular feature of the wireless channel important for the problem at hand. The first model assumes that the channel is wideband and quasi-static, and is important to model the effect of frequency selectivity of the channel. The quasi-static assumption makes the channel static for packet transmission duration which is justifiable for indoor wireless LAN communication [24]. The second model assumes that the channel is narrowband but varying from symbol to symbol, and is important to model the effect of time variation of the channel.

In cases where there are few scatterers or there is insufficient antenna spacing, the fading paths will be correlated with each other [35]. Correlative models attempt to approximate the spatial structure of the channel by modeling the correlation between paths in a MIMO channel. These models bias a channel matrix with independent, identically distributed (i.i.d) random variables by introducing correlation matrices at the transmitter and the receiver that represent scatterers at both link ends. The correlation between all paths in a narrowband MIMO channel is defined as [30]

$$\boldsymbol{R}_H = E\{Vec(\boldsymbol{H})Vec^H(\boldsymbol{H})\} \tag{2.3}$$

where $Vec(\boldsymbol{H})$ is the mapping of all elements of $\boldsymbol{H}$ such that all columns in $\boldsymbol{H}$ are stacked on top of each other to form a vector as shown below,

$$Vec(\boldsymbol{H}) = \left[h_{11}, h_{21}, \dots, h_{N_r1}, h_{12}, \dots, h_{N_r2}, \dots, h_{N_rN_t}\right] \tag{2.4}$$

The Kronecker model is by far the most widely used correlative MIMO channel model [30]. It greatly simplifies channel analysis as it holds that scatterers around the transmitter fade independently of those around the receiver. Using the Kronecker model, the correlation matrix is

the product of the one sided correlation matrices seen from the transmitter and receiver sides [30].

$$\boldsymbol{R}_H = \boldsymbol{R}_{Tx}^T \otimes \boldsymbol{R}_{Rx} \tag{2.5}$$

where $\otimes$ represents the Kronecker product, $\boldsymbol{R}_{Tx}$ and $\boldsymbol{R}_{Rx}$ are the transmitter and receiver side correlation matrices. An ensemble of $\boldsymbol{H}$ matrices with the same spatial structure and correlation can be synthesized by spatially filtering a statistically white matrix as

$$\boldsymbol{H} = \boldsymbol{R}_{Rx}^{1/2} \boldsymbol{H}_{iid} (\boldsymbol{R}_{Tx}^{1/2})^H \tag{2.6}$$

where $\boldsymbol{R}_{Rx}^{1/2}$ and $\boldsymbol{R}_{Tx}^{1/2}$ are the matrix square roots of $\boldsymbol{R}_{Tx}$ and $\boldsymbol{R}_{Rx}$ respectively. $\boldsymbol{H}_{iid}$ consists of independent, identically distributed entries.

2.2.3 Channel State Information

In MIMO system analysis, different assumptions can be made about the knowledge of the channel gain matrix at the transmitter and the receiver, referred to as channel state information at the transmitter (CSIT) and channel state information at the receiver (CSIR), respectively. There are mainly two approaches: deterministic and statistical channel state information. The deterministic approach tries to obtain exact values for CSIR and/or CSIT. In doing so, there are different techniques of obtaining CSI depending on the underlying communication system. The statistical method use the distribution of the channel coefficients which are usually expressed in terms of channel mean and covariance matrices [31]. The deterministic type of CSI is used in this work which can be either perfect or imperfect.

For a slow-fading channel, CSIR is usually assumed since the channel gains can be obtained fairly easily by sending a pilot sequence for channel estimation. Obtaining CSI at the transmitter is more difficult than CSIR. The mechanisms of getting CSIT are different for time division duplex (TDD) and frequency division duplex (FDD) systems [32]. TDD systems offer a straightforward way for the transmitter to acquire the CSI. Since the uplink and downlink in a TDD system share the same frequency, channel reciprocity implies that the base station can learn the downlink from known pilot signals transmitted on the uplink.

In FDD systems, the base station transmits and receives on different frequencies which implies that the downlink cannot be estimated based on information about the uplink because the two channels are different. Therefore, the channel estimate at the receiver should be fed back to the transmitter through a separate channel. The accuracy of the CSIT depends on the number of feedback bits used which is limited by the capacity of the feedback path. Number of bits fed back becomes large as the number of users increases. This is because each channel vector is quantized based on a codebook known at both the mobiles and the base station, and the quantization error depends on the number of bits per codeword [33].

In practical systems, obtaining perfect CSIT is often difficult. This imperfection can be caused by

- Delay in feedback path of the channel estimates at the receiver to the transmitter. This makes the CSIT to be an outdated version of the true channel and is more problematic in channels with faster time variation.
- Error in channel estimation of the downlink (for FDD systems) and uplink (for TDD systems). This also makes the CSIR used at the receiver and the CSIT fed back to the transmitter to be erroneous.
- Quantization error and error in the feedback path for FDD systems.

2.2.4 Capacity

The channel capacity of a point to point MIMO channel is a real number: for any rate strictly smaller than capacity, there exist codes at such a rate that can achieve any desired probability of error. But any rate strictly greater than capacity is not achievable in the sense that the probability of error is bounded away from zero.

For multiuser MIMO channels, the problem is somewhat different. Given a constraint on the total transmitted power, it is possible to allocate varying fractions of that power to different users in the network, so a single power constraint can yield many different information rates. The maximum capacity for a user is achieved when all the power is allocated to it as for points A and

B in Figure 2.4. For every possible power distribution in between, there is an achievable information rate, which results in the capacity region.

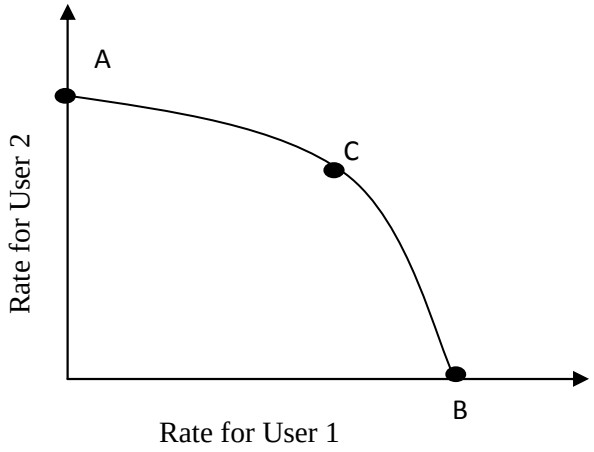

Figure 2.4: MIMO Broadcast Capacity Region for a 2 user System

The capacity region of a given MIMO broadcast channel consists of all achievable rates $\{R_1, R_2, \dots, R_k\}$ given the power constraint on the transmitted signal, i.e it is a set of rate vectors in a K dimensional space. Points on the boundary of the capacity region are typically obtained by transmitting to all users simultaneously. User rates achieved by orthogonal techniques like TDMA where at a given time only a single user communicates are contained in the capacity region, and are strictly suboptimal [34].

The maximum achievable throughput of the entire system is characterized by the point on the curve that maximizes the sum of all of the users' information rates, and is referred to as the sum capacity of the channel. This point is illustrated in Figure 2.4 by point C. A transmission technique known as Dirty Paper Coding (DPC) is shown to achieve the capacity of MIMO broadcast channels [2]. The sum capacity for the system modeled in section 2.2.1 has been formulated using the DPC framework for the case of Gaussian noise. The capacity is defined in terms of the achievable rate for each user given the set of covariance matrices for each transmitted data vector $\boldsymbol{S}_k = E\{ s_k s_k^* \}$. The achievable rate of user k is [31]:

$$R_k = \frac{\log\left|\mathrm{I} + \mathbf{H}_\mathrm{k}\left(\sum_{\mathrm{j}=1}^{\mathrm{k}} \mathbf{S}_\mathrm{j}\right)\mathbf{H}_\mathrm{k}^*\right|}{\log\left|\mathrm{I} + \mathbf{H}_\mathrm{k}\left(\sum_{\mathrm{j}=1}^{\mathrm{k}-1} \mathbf{S}_\mathrm{j}\right)\mathbf{H}_\mathrm{k}^*\right|} \tag{2.7}$$

The above equation represents the achievable rate tuples for a given permutation of users. If the ordering that maximizes sum of the rates is taken, the sum rate becomes [31]:

$$C_s = \max_{s_k \geq 0\ ;\sum tr(s_k) \leq \rho} \sum_{k=1}^{K} R_k \tag{2.8}$$

where ρ is the upper bound on the total transmit power and $tr(.)$ stands for trace of a matrix.

2.3 Orthogonal Frequency Division Multiplexing

In classical data systems that use the frequency domain to send data at higher rate, parallel transmission was achieved by dividing the total signal frequency band into non-overlapping frequency sub channels. This technique is referred to as frequency division multiplexing. This method, however, leads to inefficient use of the available spectrum. A more efficient use of bandwidth can be obtained with parallel transmission if the spectra of the individual sub-channels are permitted to partly overlap. This requires that specific orthogonality constraints are imposed to facilitate separation of the sub-channels at the receiver. This transmission technique is referred to as Orthogonal Frequency Division Multiplexing (OFDM).

The basic principle of OFDM is to split a high-rate data stream into a number of lower rate streams that are transmitted simultaneously over a number of subcarriers. The carriers in an OFDM signal are arranged so that the sidebands of the individual carriers overlap and the signals are still received without adjacent carrier interference. To do this, the carriers must be mathematically orthogonal. The receiver acts as a bank of demodulators, translating each carrier down to baseband, with the resulting signal integrated over a symbol period to recover the raw data. If the other carriers all beat down the frequencies that, in the time domain, have a whole number of cycles in the symbol period T, then the integration process results in zero contribution from all these other carriers. Thus, the carriers are orthogonal if the carrier spacing is a multiple of 1/T where T is the symbol period.

In practice, sub-carrier mapping can be implemented very efficiently by the fast Fourier transform (FFT) algorithm. The input data stream is modulated, resulting in a complex symbol stream $X[0], X[1], \ldots, X[N-1]$. The output of the OFDM modulator can be written as [29]:

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X[k] e^{j2\pi n^{k}/_{N}} \quad , \qquad 0 \leq n \leq N-1 \tag{2.9}$$

where $x[n], n = 0, 1, 2, \ldots, N-1$ are samples of the OFDM symbol, N is the number of subcarriers and $X[k]\,, k = 0, 1, 2, \ldots, N-1$ are the modulated symbols on each subcarrier.

Cyclic prefix is a crucial feature of OFDM system. The basic idea is to replicate part of the OFDM time-domain waveform from the back to the front to create a guard period. The duration of the guard period t_g should be longer than the worst-case delay spread of the target multi-path environment.

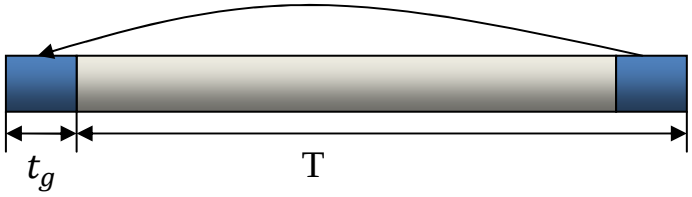

Figure 2.5: Addition of cyclic prefix

At the receiver, the starting point of sampling is chosen to satisfy the criteria:

$$t_{max} < t_{samp} < t_g \tag{2.10}$$

where t_{max} is the maximum of the delay in all reflected paths, t_{samp} is the sampling instant and t_g is duration of the guard interval.

The main advantage of OFDM is that it mitigates the frequency selectivity of the wireless channel by converting the wideband channel into a set of narrowband channels that undergo flat fading. The problem of ISI is also solved by using cyclic prefix as explained earlier.

Chapter 3

Dirty Paper Coding Techniques

As discussed in the previous chapter, a MIMO broadcast system with spatial multiplexing consists of a transmitter with multiple antennas that sends independent data to spatially separated users. In this work, systems with one receive antenna per user are discussed but, generally, users can have more than one antenna in a MIMO broadcast system. The number of independent data streams that can be transmitted is limited by $min(M, K)$ where M is the number of transmitting antennas and K is the total number of receive antennas in the system. Having multiple antennas at each user terminal can be used either to get multiple data streams per user or to have receive diversity gain. The number of users in a typical broadcast system is usually much greater than the number of antennas at the transmitter. Therefore, $min(M, K)$ usually reduces to M. This implies that having multiple streams per user is inappropriate from system design perspective.

Achieving receive diversity gain in MIMO broadcast systems is also disadvantageous since multipath fading is an advantage rather than a problem in using spatial multiplexing. This is because coding the symbols before transmission takes into account the channel vector from the transmit antennas to each receive antenna which makes having another antenna at a user terminal useless unless we have multiple data streams to that user. Having only one receive antenna at each user terminal is also easy in upgrading existing transmission strategies to the proposed ones which requires modifications at the transmitter side only. Thus, MIMO broadcast systems with one receive antenna per user terminal are considered in this work.

Using the system model described in Section 2.2.1 with users having single antenna, user j not only receives its desired signal through its channel vector $\boldsymbol{h}_j$, but also contributions from the signals destined for other users. Mathematically, this can be written as:

$$y_j = \boldsymbol{h}_{jj} d_j + \sum_{\substack{k=1 \\ k \neq j}}^{K} \boldsymbol{h}_{jk} d_k + \boldsymbol{n}_j \tag{3.1}$$

where d_j is the symbol intended for user j. The first term in the above equation represents the desired signal after passing through the channel while the second term represents the interference from signals transmitted for other users in the system.

The pre-coding problem can be formulated as: *how can we code or process signals of the different users before transmission so that each user only receives the signal intended for it.* This can be described by the block diagram shown in Figure 3.1 below for a transmitter with four antennas. During one symbol time, the input to the pre-coder is a 4 x 1 vector, $\boldsymbol{d} = [d_1 \ d_2 \ d_3 \ d_4]^T$ of symbols each intended for a single user. The pre-coder converts this vector to another vector which is then transmitted by the transmit antennas. The output vector from the pre-coder may exceed the sum power constraint of the system. Thus, it is normalized to have power equal to the required power constraint.

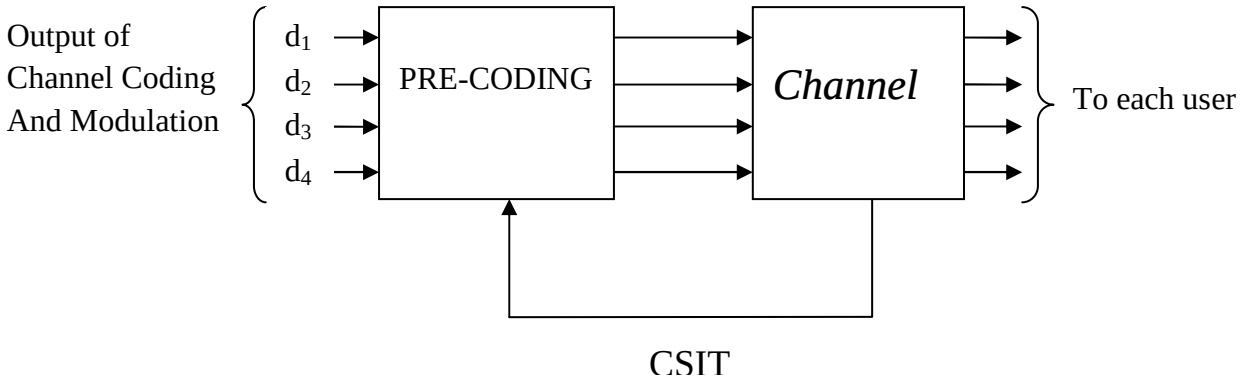

Figure 3.1: Block diagram of the Pre-coding process.

There are two fundamental approaches in pre-coding signals for MIMO broadcast channels: *linear pre-coding* and *Dirty paper pre-coding* approaches.

3.1 Linear Pre-coding Techniques

This approach relies on signal processing to pre-code signals before transmission. The pre-coding process is performed on an already channel coded and constellation mapped symbols, and the operations involved are linear. The pre-coding process only takes the channel gains into consideration and does not depend on the symbol vector to be transmitted.

Zero forcing transmit beamforming (ZFBF) is one algorithm that performs pre-coding using linear operations [7]. It multiplies the data vector $\boldsymbol{d}$ by a beamforming matrix to effectively create beams in the desired user direction while nulling the component of the signal in the direction of the other users. The transmitted vector can be expressed as:

$$\boldsymbol{x} = \sum_{j=1}^{K} \boldsymbol{b}_j d_j \tag{3.2}$$

where $\boldsymbol{b}_j$ is the beamforming vector for user j and d_j is the symbol intended for user j. The symbol received by user j then becomes:

$$y_j = \boldsymbol{h}_j \boldsymbol{x} + n_j \tag{3.3}$$

$$= \boldsymbol{h}_j \sum_{k=1}^{K} \boldsymbol{b}_k d_k + n_j \tag{3.4}$$

$$= \boldsymbol{h}_j \boldsymbol{b}_j d_j + \sum_{\substack{k=1 \\ k \neq j}}^{K} \boldsymbol{h}_j \boldsymbol{b}_k d_k + n_j \tag{3.5}$$

The algorithm forces the second term in Eq. (3.5) to be zero, i.e. $\boldsymbol{h}_j \boldsymbol{b}_k = 0$ for $k \neq j$ in selecting the beamforming vectors. The pseudo inverse of the overall channel matrix satisfies the above criterion for selection of beamforming matrix $\boldsymbol{B} = [\boldsymbol{b}_1\ \boldsymbol{b}_2\ \boldsymbol{b}_3\ \boldsymbol{b}_4]$,

$$\boldsymbol{B} = \boldsymbol{H}^{*}(\boldsymbol{H}\boldsymbol{H}^{*})^{-1} \tag{3.6}$$

which reduces to simple channel inverse when there are as many users as the number of transmitting antennas. Thus, the transmitted symbol vector can be expressed as

$$\boldsymbol{x} = \frac{1}{\sqrt{\gamma}}\boldsymbol{H}^{*}(\boldsymbol{H}\boldsymbol{H}^{*})^{-1}\boldsymbol{d} \tag{3.7}$$

where $\gamma = \|\boldsymbol{x}\|^2$ is used to normalize the transmitted signal power to unity as is true for subsequent algorithms in this work. The symbol received by user *j* and the estimated data become

$$y_j = \frac{d_j}{\sqrt{\gamma}} + n_j \tag{3.8}$$

$$\widehat{d_j} = \sqrt{\gamma}\, y_j \tag{3.9}$$

The major drawback of this algorithm is the significant signal attenuation due to the normalization constant when the channel matrix is poorly conditioned or $(\boldsymbol{H}\boldsymbol{H}^{*})^{-1}$ has large eigenvalues. This problem might be interpreted as noise amplification when the received signal is multiplied by the normalization constant.

One way to decrease the sensitivity of the Zero forcing beamforming to channel condition is to use the MMSE criterion to design the beamforming weights. The beamforming weights can now be expressed as [7]

$$\boldsymbol{B} = \boldsymbol{H}^{*}(\boldsymbol{H}\boldsymbol{H}^{*} + \zeta\boldsymbol{I})^{-1} \tag{3.10}$$

where $\boldsymbol{I}$ is the identity matrix of order equal to the number of transmitting antennas. The regularization parameter ζ is chosen so as to maximize SINR at each receive antenna and is a function of the SNR and number of transmit antennas. This algorithm adds a limited amount of interference to compensate for the behavior of the channel matrix. The amount of added interference increases with ζ, and the algorithm reduces to simple transmit beamforming

when $\zeta = 0$. The optimum value of ζ is found to be $^{M}/_{SNR}$ for a Rayleigh fading channel [36] where M is the number of transmit antennas.

3.2 Dirty Paper Pre-coding Techniques

While the previous technique relies on signal processing, the dirty paper approach is, in general, a coding approach. This is because it combines other parts of signal design like modulation and coding with the pre-coding process. This is true for the capacity achieving dirty paper coding schemes, and therefore, completely new communication protocols are needed for real world implementation. There are suboptimal algorithms that separate signal design with pre-coding which can be implemented in a similar way as the linear algorithms [5, 6, 7].

The idea behind this technique is analogous to writing on a dirty paper where the writer can apply ink by considering the dirt picked up by the paper later on. The combined effect of the ink and the dirt will make the words distinguishable by the user which may not happen if the writer didn't write this way. Consider a channel with output corrupted by interference S and noise Z as

$$Y = X + S + Z \tag{3.11}$$

The interference $S \sim \mathcal{N}(0, Q)$, the additive noise $Z \sim \mathcal{N}(0, N)$ and $\|X\|^2 \leq P$. If S is unknown at the transmitter and the receiver, the capacity of this channel is $\frac{1}{2}\log(1 + \frac{P}{Q+N})$ bits/channel use because the interference is taken as an additive noise. On the other hand, if S is known at the transmitter, the capacity per dimension is shown to be [3]

$$C = \frac{1}{2}\log(1 + \frac{P}{N}) \quad {}^{\text{bits}}/_{\text{channel use}} \tag{3.12}$$

This is the capacity of a standard Gaussian channel with SNR of $^{P}/_{N}$ and is independent of the interference S even if the interference is unknown to the receiver. Rather than attempting to fight and cancel this interference, the optimal encoder adapts to it by choosing codewords in the direction of S.

In the MIMO broadcast system, the interference signal arises from signals transmitted to users other than the intended user. Since the symbols to be transmitted to all users are available at the transmitter, this system is one variation of the side information channel discussed above. The transmitter can code information for the users such that each user experiences no interference from the signals intended for other users.

The main difference between the linear processing and the dirty paper coding approaches lies in the level of consideration of the interference. The interference encountered by a given user is a combination of the signals of the other users and the channel from each transmit antenna to the given user's receive antenna. The linear processing techniques don't take the other users' signals into consideration in performing pre-coding while the dirty paper techniques consider both the channel and the users' signals. Therefore, the linear approach incurs performance loss which is usually expressed in terms of power penalty.

The result expressed in Eq. (3.12) is an information theoretic one and the construction of practical codes that approach the capacity limit is still an active research area. There are suboptimal algorithms that separate signal design from the pre-coding step but still attain a significant portion of the promised gain with practical computational complexity relative to the optimal ones. These algorithms and their corresponding achievable sum rate are discussed in the next section.

3.2.1 Modified Tomlinson-Harashima Pre-coding (THP)

This algorithm uses the idea of the pre-coding techniques for combating ISI in single user systems in which some proportion of the previously transmitted symbol is subtracted from the symbol to be transmitted now[5,6]. It uses QR decomposition of the channel matrix to effectively make the composite channel matrix triangular, and then successively pre-subtracts interference of one user from the one encoded after it.

The overall channel of the system, $\boldsymbol{H}$, can be decomposed into a unitary matrix $\boldsymbol{Q}$ and a lower triangular matrix $\boldsymbol{R}$ using QR decomposition. Thus, Eq. (2.1) can be expressed as:

$$\boldsymbol{y} = \boldsymbol{RQx} + \boldsymbol{n} \tag{3.13}$$

By selecting the transmitted vector $\boldsymbol{x}$ to be equal to $\boldsymbol{Q}^H\boldsymbol{s}$, the received vector can be expressed as

$$\boldsymbol{y} = \boldsymbol{RQ}(\boldsymbol{Q}^H\boldsymbol{s}) + \boldsymbol{n} \tag{3.14}$$

$$= \boldsymbol{Rs} + \boldsymbol{n} \tag{3.15}$$

The channel between $\boldsymbol{s}$ and $\boldsymbol{y}$ becomes triangular. For a system with 4 transmit antennas, this can be expressed in matrix form as

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} r_{11} & 0 & 0 & 0 \\ r_{21} & r_{22} & 0 & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ r_{41} & r_{42} & r_{43} & r_{44} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{bmatrix} \tag{3.16}$$

The data symbols to be transmitted are denoted by d_i where $i = 1, 2, \ldots, K$. The pre-coding step will create an interference free subchannel as

$$y_i = r_{ii}d_i + n_i \tag{3.17}$$

In order to satisfy Eq. 3.17, the interference pre-subtraction can be expressed as

$$s_k = d_k - \sum_{i=1}^{K-1} \frac{r_{ki}}{r_{kk}} s_i \tag{3.18}$$

As shown in Figure 3.2, the effective channel between $\boldsymbol{s}$ and the users is $\boldsymbol{R}$. Since $\boldsymbol{R}$ is triangular, user 1 doesn't see interference from users 2 up to K, user 2 doesn't see interference from users 3 up to K and so on.

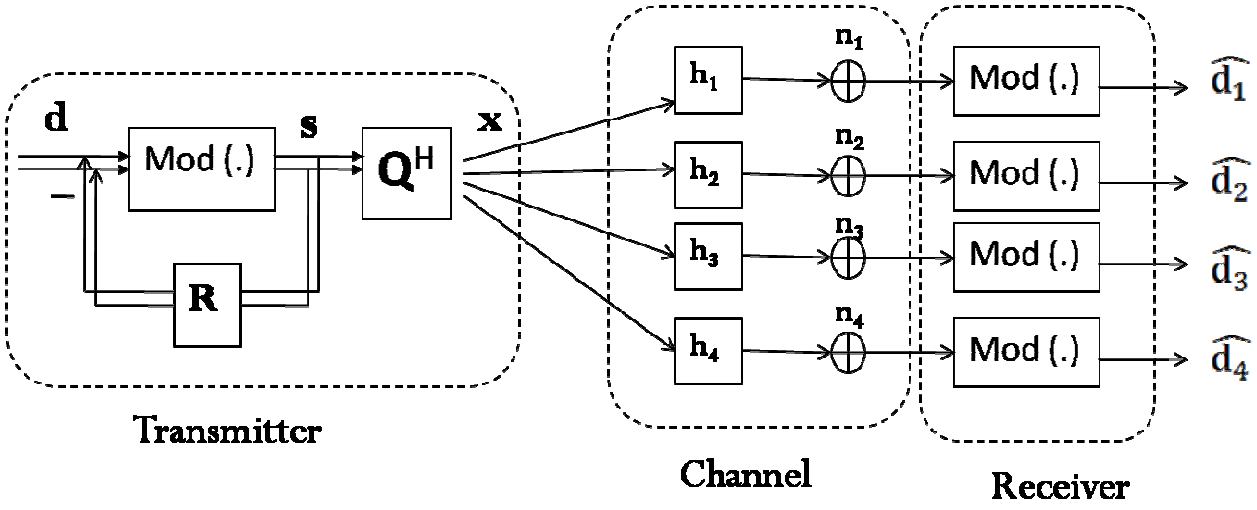

Figure 3.2: Block diagram of Modified THP

In pre-subtracting interference, power of the pre-coded signal usually exceeds the power constraint of the system. This increase in power is more pronounced in users encoded last because symbols of all users encoded before them will be subtracted making their power larger. To limit the transmit power, modulo operation is performed on the pre-coded symbols. The modulo operation with argument on a symbol is

$$\_\_\_\_\_$$

where denotes the floor operation which is largest integer less than or equal to the input. So this operation makes the real and imaginary parts of the symbol to be located within .

In effect, this operation extends the constellation of the received symbol so that demodulation is performed on the extended constellation. One step extended constellation for decoding of QPSK modulated symbols is shown in Figure 3.3 below. The dark dots belong to the original QPSK constellation.

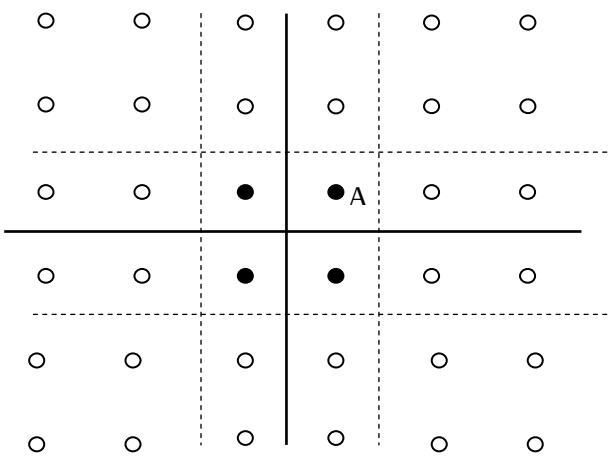

Figure 3.3: Extended constellation for QPSK modulation.

The pre-coder maps transmitted symbols that exceed the original constellation into points on the next extended constellation. Then, these symbols on the extended constellation will be mapped on to the original constellation using modulo operation to the real and imaginary parts respectively. For example, if point A has coordinates (1,1), the dotted lines which are boundaries between constellations should be located at coordinate $(\pm 2,0)$ and $(0,\pm 2)$ for the real and imaginary parts, respectively. This is to avoid ambiguity in decoding the received symbol which is located at any one of the constellations. The modulo operation will, therefore, have an argument 4 for both the real and imaginary parts. Thus, for 16 QAM modulation, the argument of the modulo operation will be 8.

At the receiver, the symbol will be located on any one constellation in the extended constellation. The receiver can perform modulo operation as shown in eq. (3.20) on the symbol to map it to the original constellation which can be demodulated accordingly.

$$\hat{d}_j = modulo_\tau\left(y_j \right) \tag{3.20}$$

where $\tau = 8$ for 16 QAM modulation with average power of $\sqrt{10}$. Aside from being a non linear operation, another problem of the modulo operation is that the additive noise may push the received symbols beyond the boundaries of the extended constellation. This makes the symbols to be decoded incorrectly.

3.2.2 Vector Perturbation

As discussed in Section 3.1, the problem with Zero Forcing Beamforming is that the pre-coded signal usually has larger power because of the large singular values in the channel inverse. The best way to reduce the transmit power is to make sure that the data vector does not lie along the singular vectors associated with the largest singular values of the channel inverse [4]. This algorithm perturbs the data vector in a data dependent way that is unknown to the receivers before multiplying it with the beamforming matrix used in Zero Forcing transmit beamforming.

After perturbation, the transmitted vector is approximately orthogonal to singular vectors associated with large singular values of $\boldsymbol{H}^{*}(\boldsymbol{H}\boldsymbol{H}^{*})^{-1}$. The simplest type of perturbation is to add an integer multiple of some predetermined number to the data vector as:

$$\hat{\boldsymbol{d}} = \boldsymbol{d} + \tau \boldsymbol{l} \tag{3.21}$$

In the above expression, τ is a positive real number which is determined by the modulation technique used. It is a design parameter selected so that the receivers can decode the data without knowledge of the perturbing vector l. So τ is similar to the argument of the modulo operation used in the Modified THP previously. A general formula for calculating τ is [4]

$$\tau = 2\left(|c|_{max} + {}^{\Delta}/_{2}\right) \tag{3.22}$$

where $|c|_{max}$ is the absolute value of the constellation symbol with largest magnitude and Δ is the spacing between constellation points.

If τ is made larger, the decoding error at the boundaries of the decoding region (on the extended constellation) will be smaller and the effect of the perturbing vector l will decrease. But the average power of the transmitted signal will increase in the process thereby degrading the overall error rate performance. If it is made smaller, error free decoding will be impossible even in the absence of additive noise because of ambiguity in the decoding regions of the extended constellation.

The vector $\boldsymbol{l}$ is K dimensional and has complex entries $a + jb$ with a and b being integers. The transmitted signal can be expressed as [4]

$$\boldsymbol{x} = \frac{1}{\sqrt{\gamma}} \boldsymbol{H}^{*} (\boldsymbol{H}\boldsymbol{H}^{*})^{-1} \hat{\boldsymbol{d}} \tag{3.23}$$

where $\gamma = \|\boldsymbol{x}\|^2$. The perturbation vector $\boldsymbol{l}$ should be chosen so that the transmitted power is minimized for the given channel matrix, and is the solution of the minimization problem

$$\boldsymbol{l} = \arg\min_{\boldsymbol{l}'} \ (\boldsymbol{u} + \tau \boldsymbol{l}')^{*} \ (\boldsymbol{H}\boldsymbol{H}^{*})^{-1} (\boldsymbol{u} \ + \ \tau \boldsymbol{l}') \tag{3.24}$$

This is a K dimensional complex integer lattice least squares problem which can be decomposed into a 2K dimensional real least squares problem. In selecting the appropriate value for the integer offset vector $\boldsymbol{l}$, we have infinite possible choice and the minimization will be over an infinite lattice.

There are different algorithms to solve this minimization problem. The simplest but complex one is to limit the search path to a predetermined interval and then try the possible complex integer values. The vector that results in minimum value of the argument will be chosen. This is analogous to maximum likelihood (ML) decoding in multiuser detection except in ML detection the search interval is already limited by the constellation points. Another approach is the Finche-Pohst algorithm [37] which limits the search interval and updates the minimum distance so as to exit a branch if its distance exceeds the minimum.

The signal received by user j and the estimated symbol can be expressed as

$$y_j = \frac{1}{\sqrt{\gamma}} \left(\hat{d}_j + \ \tau l \right) + \ n_j \tag{3.25}$$

$$\tilde{d}_j = \ modulo_{\tau} \left(\sqrt{\gamma} \, y_j \right) \tag{3.26}$$

The modulo operation is the same as the one defined in eq. (3.19) for modified THP algorithm. In eq. (3.23), the transmitter divides the pre-coded symbol by $\sqrt{\gamma}$ which is a data dependent quantity and the receivers multiply the received vector by the same quantity. The transmitter and

the receivers can instead use $E\{\sqrt{\gamma}\}$ as normalization constant as the performance difference is shown to be insignificant [36].

3.3 Pre-coding for OFDM Systems

The linear and non linear pre-coding techniques discussed in the previous sections are for single carrier systems. In order to mitigate the effect of multipath fading channels, recent wireless communications use multicarrier transmission techniques. As discussed in Section 2.3, the most common and widely used multicarrier transmission strategy is OFDM. In OFDM systems, the pre-coding techniques cannot be directly applied to the signal to be transmitted. This is because the transmitted time domain waveform is a superposition of modulated symbols mapped to different subcarrier frequencies.

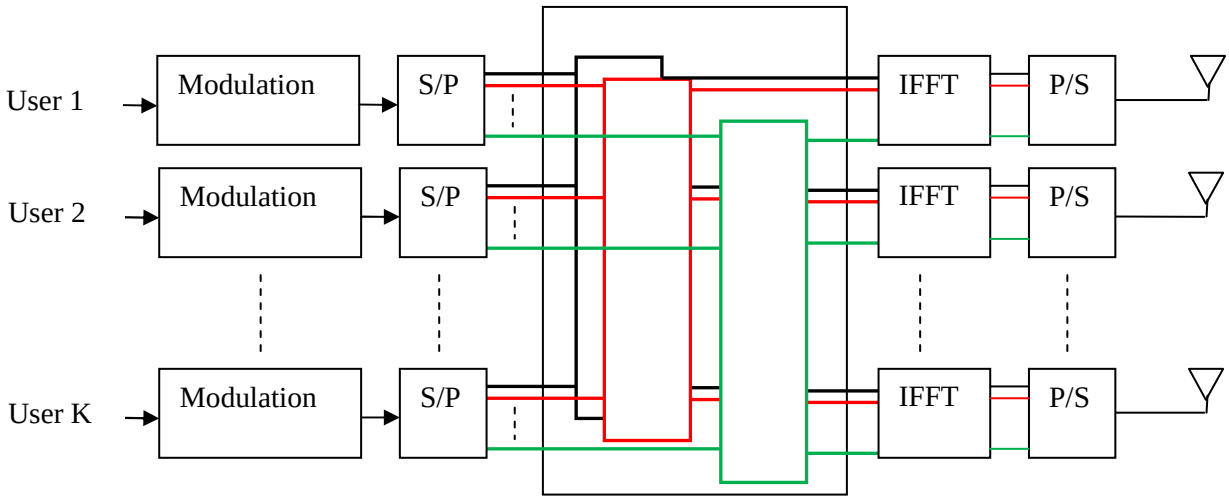

Figure 3.4: Block diagram of the pre-coding process for OFDM systems

In these systems, pre-subtraction should be performed in frequency domain. Each subcarrier, thus, can be thought of as an independent channel and data of the different users can be pre-coded on a subcarrier by subcarrier basis. As shown in Figure 3.4 above, the modulated symbol of each user is serial to parallel converted into N parallel symbols to be transmitted on the N subcarriers. User symbols on one subcarrier are taken and pre-coding is performed to get K symbols to be transmitted on that subcarrier.

In Figure 3.4, the pre-coder for the different subcarriers is shown in black for subcarrier 1, red for subcarrier 2 and green for subcarrier N. The pre-coded symbols from the N pre-coders for user 1 will be inputs to the IFFT block which perform subcarrier mapping for this user. Each of the N pre-coders will be delivered with the state of the channel at that subcarrier. For a system with K transmit antennas and K single antenna users, the pre-coder for one subcarrier will have a K x K channel state matrix.

3.4 Achievable Sum Rates

The sum capacity of the MIMO broadcast system given in Eq. 2.8 is for the capacity achieving pre-coding strategy. The optimal dirty paper coding strategy performs both signal design and pre-coding simultaneously and pre-coding extends over multiple symbols. The algorithms discussed previously, on the other hand, perform pre-coding independent of the other parts of the communication system and, pre-coding is performed symbol by symbol. Therefore, these algorithms are sub-optimal and achieve a maximum sum of user rates less than the sum capacity.

In all the expressions in this section, user k has channel gain $\boldsymbol{h}_k$, the additive noise has unit variance and the total transmitting power is P. The achievable sum rate of Zero Forcing transmit beamforming (ZFBF) is achieved by allocating more power to the stronger users and less power to the weak users. Thus the waterfilling power allocation strategy achieves this rate and is given by [14]

$$R_{ZFBF} = \max_{P_i} \sum_{k=1}^{K} \log(1 + P_i) \tag{3.27}$$

$$where \qquad \sum_{k=1}^{K} \frac{P_i}{\gamma_i} \leq P \qquad \text{and} \qquad \gamma_i = \frac{1}{[(\boldsymbol{H}\boldsymbol{H}^H)^{-1}]_{i,i}} \tag{3.28}$$

The waterfilling power allocations are [14]:

$$\sum_{k=1}^{K}(\mu - \frac{1}{\gamma_i})^+ = P \quad \text{and} \quad P_i = (\mu\gamma_i - 1)^+ \tag{3.29}$$

As of our knowledge, the exact achievable sum rate of vector perturbation for arbitrary input distribution is still unknown but is found in [30] for uniformly distributed inputs. An upper bound for the achievable sum rate is also calculated as a function of channel matrix and is given by [15]

$$R_{vp} < K\log\frac{P}{K} + \log\det(W) - K\log\frac{\Gamma(K+1)^{\frac{1}{K}}e}{K+1} \tag{3.30}$$

where $\boldsymbol{W} = \boldsymbol{H}\boldsymbol{H}^H$ and $\Gamma(.)$ denotes the gamma function.

The achievable sum rate of Modified THP is attained by permutation of users before interference pre-subtraction and using the waterfilling power allocation strategy afterwards, and is given by [13]

$$R_{THP} = \sum_{k=1}^{K}(\log(\xi r_k))^+ \tag{3.31}$$

where $r_k$is the k^{th} diagonal element of the triangular matrix, R, which is the result of the QR decomposition of the channel matrix, $(x)^+ = \max(x, 0)$ and ξ is the solution of the waterfilling equation

$$\sum_{k=1}^{K}\left[\xi - {}^{1}/_{r_k}\right]^+ = P \tag{3.32}$$

In orthogonal transmission techniques, a single user is served in a given time, frequency or code slot. Thus, sum rate is maximized by always transmitting to the user with strongest channel gain. The achievable sum rate of TDMA is given by [14]

$$R_{TDMA} = \max_{k \in \{1,2,\ldots,K\}} \log(1 + P\|\boldsymbol{h}_k\|^2) \tag{3.33}$$

The Achievable sum rates of the different techniques discussed above is shown on Figure 3.5 below. The plots are for quasi-static, narrowband fading channel having Rayleigh distributed channel gains. The MIMO BC system has 4 transmitting antennas and 4 users with one receive antenna each. The results are averaged over 1000 channel realizations.

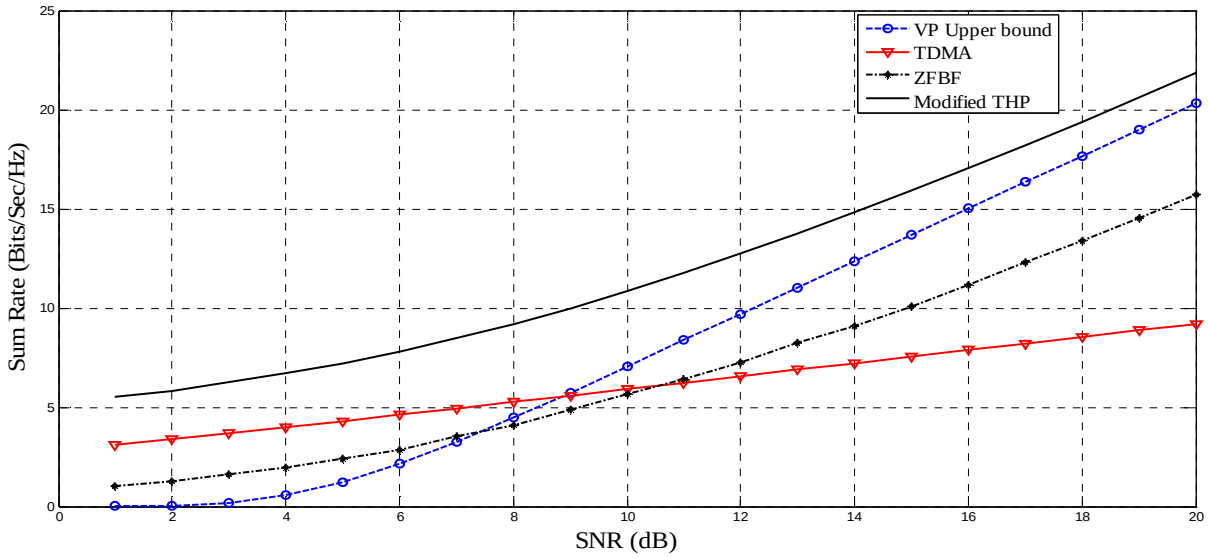

Figure 3.5: Achievable Sum Rate of different transmission techniques.

3.5 User Scheduling

As discussed previously, a typical application of dirty paper coding techniques is in wireless LAN and cellular telephony. In these systems, there are usually much more users than the number of transmit antennas. Since the number of spatial degrees of freedom is limited by $min\ (M, K)$ where M is the number of transmit antennas and K is the number of single antenna users, only M users can be served at the same time in order to exploit all M spatial streams. This brings about the need to have a criterion and technique for selecting a subset of users from all users in the system.

Besides increasing the total number of users in the system, user selection is shown to result in a significant increase in sum rate even when K is equal to M [13]. It is also necessary if different rates are to be assigned to the users both when K is equal to and greater than M [15]. In the quasi- static channel model, the best criterion for user grouping is strength of channel gains and relative orthogonality between channel gains of different users. If there are a total of K users in the system, these users will be grouped into K/M groups with M users in each group. The groups will be arranged in descending order of preference so that the first group is the best in terms of the above criteria.

After the user selection process, time slots can be assigned to each user group depending on the aim of the selection process. If we want to maximize the sum rate of the system, the first user group will be served all the time because this group results in a better sum rate than the others. If scheduling fairness is considered, varying proportion of the available time slots can be assigned to the user groups.

3.5.1 User Selection Algorithm

This algorithm performs successive Gram-Schmidt orthogonalization on the channel vectors of users with the aim of obtaining user groups with relatively good orthogonality. Algorithms similar to this one are proposed in [14] for zero forcing beamforming as a user selection algorithm and in [13] as a permutation algorithm to maximize sum rate relative to the random ordering of users.

Let U denote the set of users considered in the current iteration and i the iteration index. Denote S as the set of users that have been selected and N its cardinality. If a user is selected in the previous iteration, it will be excluded from U and becomes a member of S. Let G denote the current user group and $|X|$ is the cardinality (number of elements) of set X. The channel vector of each user is $\boldsymbol{h}_k$, $k = 1, 2, 3, \dots, K$. The algorithm performs grouping as follows:

Step 1) Initialization: $U = \{1, 2, 3, \dots, K\}$, $i = 1$, $S = \emptyset$, $G = 1$.

Step 2) For each $k \in U$, calculate $\boldsymbol{g}_k$. It is the component of $\boldsymbol{h}_k$ orthogonal to the subspace spanned by $\boldsymbol{g}_{(1)}, \boldsymbol{g}_{(2)}, \dots, \boldsymbol{g}_{(i-1)}$.

$$\boldsymbol{q}_{(j)} = \frac{\boldsymbol{g}_{(j)}}{\|\boldsymbol{g}_{(j)}\|} \tag{3.34}$$

$$\boldsymbol{g}_k = \boldsymbol{h}_k - \sum_{j=1}^{i-1} (\boldsymbol{q}_{(j)}^* \boldsymbol{h}_k)\, \boldsymbol{q}_{(j)} \tag{3.35}$$

Step 3) Select the i^{th} element of S as follows:

$$S_i = \arg \max_{k \in U} \|\boldsymbol{g}_k\| \tag{3.36}$$

$$\boldsymbol{g}_{(i)} = \boldsymbol{g}(S_i) \tag{3.37}$$

$$\boldsymbol{h}_{(i)} = \boldsymbol{h}(S_i) \tag{3.38}$$

Step 4) If $|S| < M$, then $U = \{k \in U, k \neq S_i\}$, $S = S \cup S_i$, i = i + 1 and G remains unchanged. But if $|S| = M$, then $U = \{k \in U, k \neq S_i\}$, $S = \emptyset$, $i = 1$, $G = G + 1$.

Step 5) If U is non- empty, go to step 2. Otherwise terminate the algorithm.

The algorithm initializes by considering all users in the first step. Then, it calculates the orthogonal projections of the users' channel vectors in the space spanned by the already selected users. If $i = 1$, $\boldsymbol{g}_k = \boldsymbol{h}_k$ for all users. The user with maximum orthogonal projection will be selected. This user will be added to set S and removed from set U. If the selected number of users is less than the number of transmitting antennas, the algorithm continues to the next iteration. If not, the algorithm initializes again taking the remaining users and proceeds to forming the next user group.

The proposed algorithm is different from those in [14] and [13] in that it does not terminate after forming the first group. Unlike this algorithm, the algorithm in [14] performs user shedding where the algorithm may terminate before number of users in a group equals M if the remaining users are found to have poor orthogonality with the already selected ones.

3.5.2 Subcarrier Allocation for OFDM Systems

In the previous section, user selection is performed for single carrier systems in which each user has one channel vector. Since OFDM is a multicarrier system, one user will have N channel vectors on each of the N subcarriers. Unlike the single user system, a user with good channel vector at one subcarrier may not have good channel at another subcarrier unless the two subcarriers are adjacent to one another. Thus, OFDM systems guarantee better fairness because a given user will have good channel gain on at least few subcarriers.

Subcarrier allocation can be aimed at maximizing the sum rate of the system or it may incorporate some amount of fairness. A simple measure of fairness in OFDM systems is fairness index [17]. It is related to the variance of the variable T_i which is the number of subcarriers allocated to the i^{th} user. In a MIMO broadcast system with K single antenna users,

$$fairness\ index = \frac{\left(\frac{1}{k}\sum_{i=1}^{K} T_i\right)^2}{\frac{1}{k}\sum_{i=1}^{K} T_i^2} \tag{3.39}$$

Fairness index becomes unity when all users are served equally whereas in the limit it becomes zero when only one user is served all the time. The subcarrier allocation algorithm for sum rate maximization is the same as the user selection algorithm in section 3.5.1 and is referred to as rate maximizing subcarrier allocation (RMSA) algorithm. The user selection algorithm will be applied on each subcarrier for the first group. In a K user system, each subcarrier will have K channel vectors with M elements. The subcarrier allocation algorithm will select M channel vectors with relatively good gain and orthogonality for each subcarrier. This will be repeated for all subcarriers. On each subcarrier, the first group will be served all the time because the aim is maximizing sum rate.

One way to include fairness is to make each user select its strong subcarrier and repeat the process until all subcarriers are assigned to M users. Denote the set of users selected at the n^{th} subcarrier by S_n for $n = 1,2,\ldots,N$ where N is the total number of subcarriers. Let $|S_n|$denote the number of elements of S_n. The channel gain of user k on subcarrier n is denoted by $\boldsymbol{h}_{kn}$.

Let T be the set of subcarriers with $|S_n| < M$ for $n = 1,2,\dots,N$. Denote m_k as the subcarrier index where user k has good channel strength on a given iteration and let U_k be the set of subcarrier indices for user k excluding those indices selected on previous iterations. The algorithm allocates subcarriers as follows:

Step 1) Initialization: $k = 1$, $S_n = \emptyset$ for $n = 1,2,\dots,N$, $U_k = \{1,2,\dots,N\}$ for all k and $T = \{1,2,\dots,N\}$.

Step 2) Calculate m_k, the subcarrier index for which user k has the best channel strength

$$m_k = \arg \max_{n \epsilon U_k} \|\boldsymbol{h}_{kn}\| \tag{3.40}$$

$$U_k = U_k - \{m_k\} \tag{3.41}$$

Step 3) If $|S_{m_k}| < M$, then $S_{m_k} = S_{m_k} \cup \{m_k\}$ and proceed to the next step.
But if $|S_{m_k}| = M$, then set $T = T - \{m_k\}$ and go to step 2 if $|U_k| > 0$ or proceed to the next step if $|U_k| = 0$.

Step 4) If $|T| > 0$, then set $k \leftarrow k + 1$ if $k < K$ and set $k = 1$ if $k = K$. Then go to step 2.
Otherwise, the algorithm is finished.

This algorithm is fairness oriented because each user gets its own turn and will be assigned a subcarrier if that subcarrier has less than M users assigned to it. In step 2, the subcarrier with better channel gain for the user under consideration will be selected and that subcarrier index will be excluded from U_k. If the number of selected users on the preferred subcarrier is less than M, the given user will be assigned to that subcarrier. Otherwise, the user will be given another chance to select its next best subcarrier. The algorithm finishes when all subcarriers are assigned with M users. This algorithm is thus called fairness oriented subcarrier allocation (FOSA). An algorithm similar to this one is proposed in [17] for multiplexing based MIMO OFDM systems.

Chapter 4

Performance Comparison

In this chapter, the BER performance of the different pre-coding algorithms is presented and discussed along with their complexity. The effect of fading path correlation on the BER of the algorithms is also studied. Typical application of Dirty paper Pre-coding techniques is in wireless LANs. This is because of the rich scattering properties of the environment and the broadcast nature of the service. In these applications, assuming the channel to be static for packet transmission duration is valid [24] due to lower mobility of users and objects in indoor environments. The use of OFDM minimizes frequency selectivity of the channel so that each subcarrier experiences frequency flat fading. This makes the pre-coding techniques developed for narrowband systems to be extended for wideband applications.

4.1 BER Performance Comparison

Signal to Noise ratio (SNR) is defined as the ratio of average power of the pre-coded symbols to noise variance. But the pre-coded symbols have different average power depending on the algorithm used which makes it difficult to have a fair performance comparison without normalizing the power of pre-coded symbol. Therefore, the pre-coded symbols are normalized to have unit power and SNR is the reciprocal of noise variance in all simulations.

4.1.1 Simulation Setup

Both coded and un-coded systems are used in comparing the BER performances of the algorithms. The discussion in Section 3.3 is used to perform pre-coding for OFDM systems as these systems are included in the simulation. A block diagram showing the simulation flow for coded systems is shown in Figure 4.1 below.

The simulated MIMO broadcast system has 4 antennas at the transmitter and 4 users with one receive antenna each. The random bit generator, channel coder, interleaver and modulator blocks are separate for each user, and are shown as one block for clarity of the diagram. The modulator

uses Grey encoding order so that a symbol error likely results in one bit error. Convolutional code is used as channel coding and Viterbi decoder is used to decode the received encoded bits.

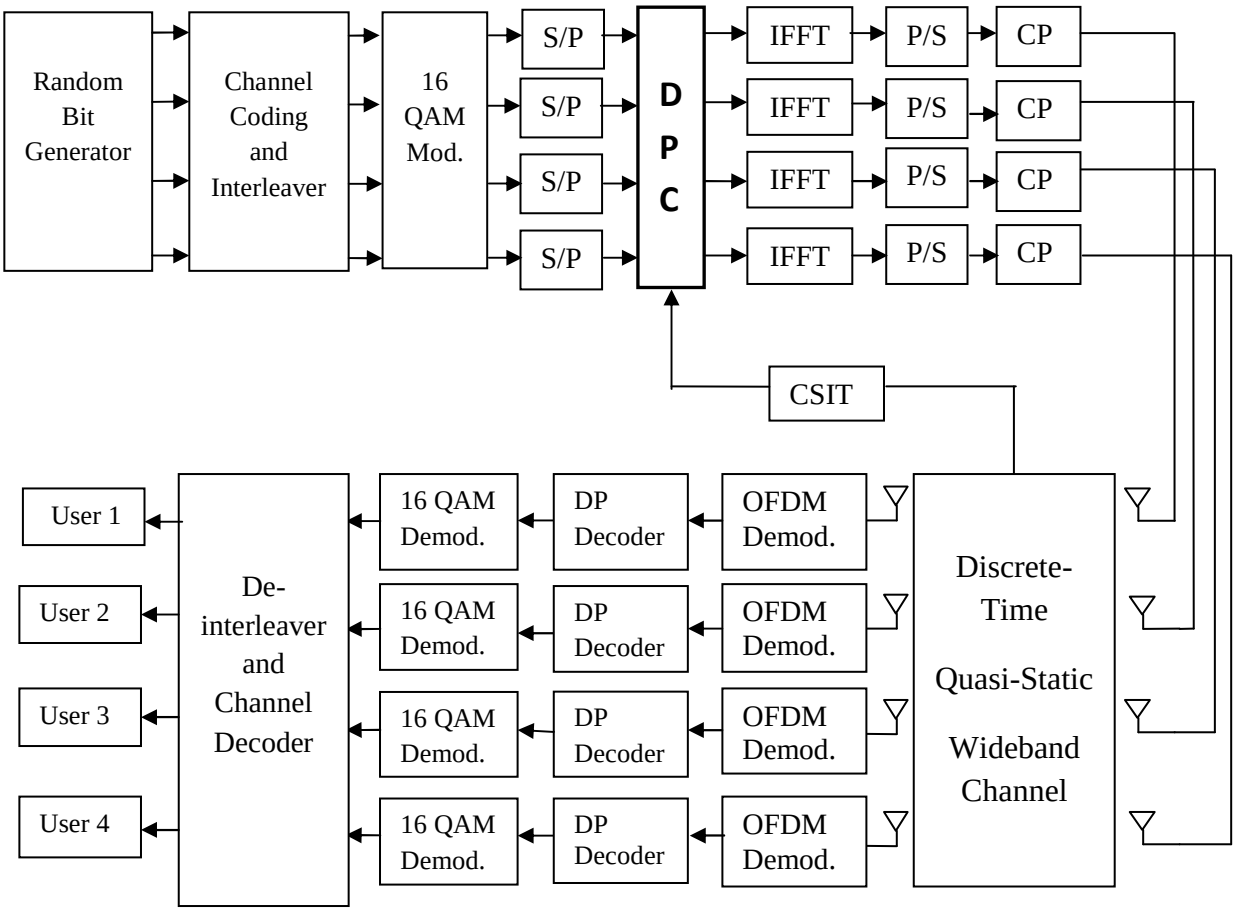

Figure 4.1: Block diagram of the Simulation Setup

The block labeled DP Decoder at the receiver side of Figure 4.1 performs simple scaling for ZFBF or modulo operation for VP and Modified THP. The OFDM demodulator performs cyclic prefix removal, serial to parallel conversion, FFT and parallel to serial conversion.

Discrete time baseband channel is used to model the wireless environment. This model has discrete time inputs and outputs that are sampled at the sampling frequency, f_s. In addition, the channel is modeled as wideband and the signals reflected from the different scatterers in the

channel arrive at different delay with different average powers. The main parameters used in the simulation are listed in Table 4.1 below.

Table 4.1: System parameters used for performance comparison of the DPC algorithms

Channel Model	Discrete-Time, Quasi-Static Wideband Channel
Distribution of Tap Weights	Rayleigh
RMS delay spread	100 nsec
Number of Taps	16
Modulation Type	16 QAM
Channel coding	Convolutional code, code rate = 1/2
Interleaver	Random interleaver
Number of Subcarriers	64
FFT Period	3.2 μs
Cyclic Prefix duration	0.8 μs
OFDM symbol duration	4 μs
Bandwidth	20 MHz
Subcarrier Spacing	0.3125 MHz
Bit Rate/user	32 Mb/s
Number of OFDM Symbols processed	100000

From works published in [38, 39] and other measurements, it is shown that the average received multipath power of a transmitted pulse in indoor-like environments fall off exponentially over time. Therefore, the wideband channel used in this simulation has tap weights with exponentially decaying power delay profile given by [24]

$$P[l] = \frac{\exp\left(-l/\tau_d f_s\right)}{\sum_{l=0}^{L-1} \exp\left(-l/\tau_d f_s\right)} \quad for \quad l = 0, 1, 2, \ldots, L-1 \tag{4.1}$$

where L is the number of taps of the wideband channel, τ_d is the rms delay spread and f_s is the sampling frequency. The signals from the different paths in the channel are taken to arrive at integral multiples of T_s where $T_s = {}^{1}/_{f_s}$.

In practical systems, an indoor wireless channel is static for packet transmission duration [24] which can traverse thousands of symbol periods. Assuming the channel to be static for large number of symbols results in smaller number of channel realization for the same number of bits processed. Thus, the channel is assumed to be constant for a block of 100 OFDM symbols and is independent from block to block which justifies the quasi-static assumption.

4.1.2 Simulation Results

Using the simulation parameters in the previous section, the BER performance of the algorithms discussed in Sections 3.1 and 3.2 is presented below. Figure 4.2 shows BER performance comparison of Zero Forcing Beamforming (ZFBF) and Modified THP for un-coded 16 QAM .

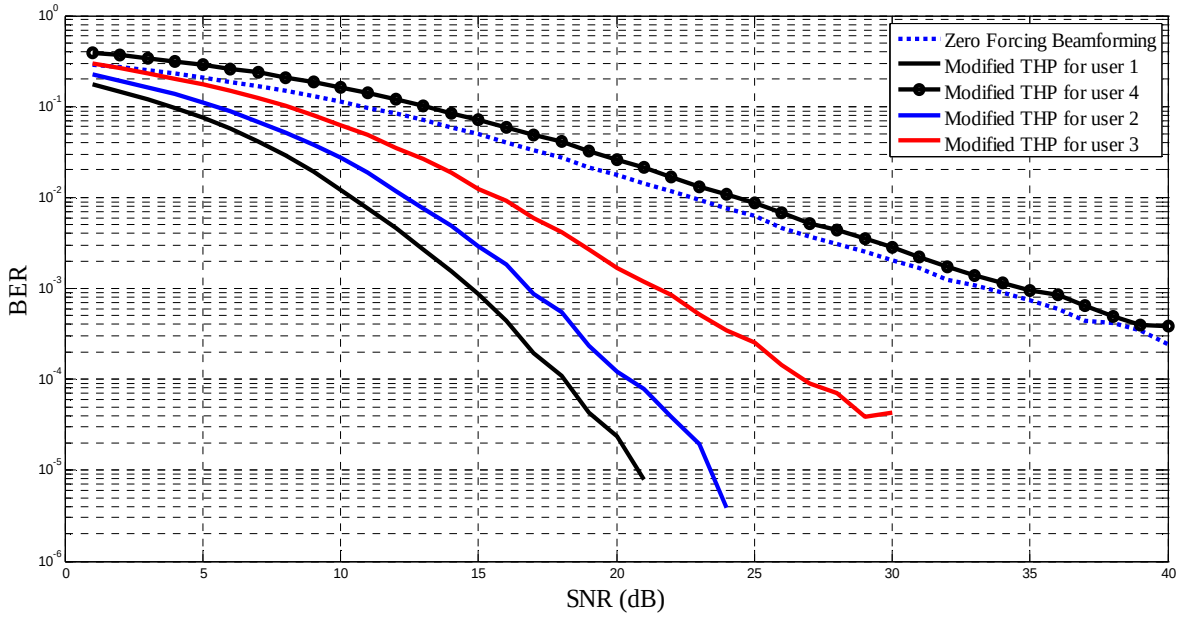

Figure 4.2: BER comparison of Zero forcing beamforming and Modified THP for un-coded 16 QAM.

ZFBF results in similar BER for the four users, and the average BER of the users is plotted for this algorithm. But Modified THP has different BER for the users based on their pre-coding order. The user pre-coded last will have higher BER and the user encoded first will have the best BER performance. This is due to the use of modulo operation in modified THP to limit the transmit power of the pre-coded symbols. The users encoded later have more symbols to be subtracted from theirs which pronounces the effect of the modulo operation. In Figure 4.2, user 1 is pre-coded first and then user 2 is pre-coded, user 4 is pre-coded last. In order to compensate the different BER performances among users, power can be allocated based on their encoding order. In this case, more power can be allocated to the user encoded last and lesser power to the user encoded first.

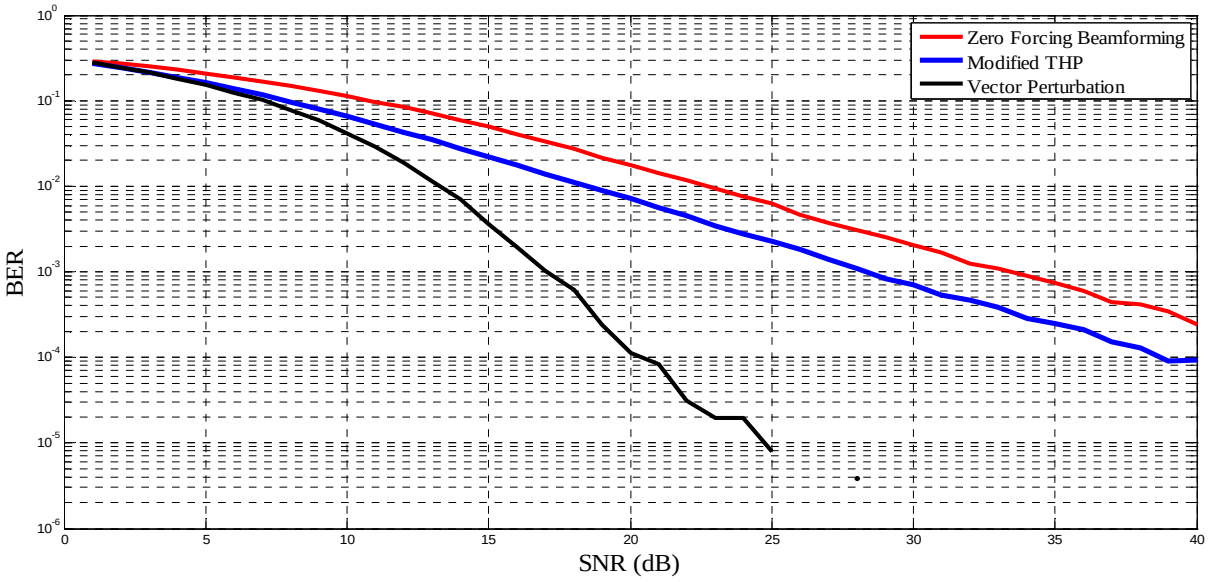

Figure 4.3: BER comparison of ZFBF, Modified THP and Vector Perturbation for un-coded 16 QAM.

Figure 4.3 shows the BER comparison of the three algorithms. The average BER of the four users is plotted for Modified THP, and so is for VP and ZFBF. To get a target BER of 10^{-3}, modified THP has 5 dB power gain in comparison with ZFBF and 11 dB power penalty in

comparison with vector perturbation. ZFBF and Modified THP show nearly the same performance difference throughout the whole SNR range while performance difference of VP relative to the two algorithms increases as SNR increases.

The performance of the pre-coding algorithms with channel coding and interleaving is shown on Figure 4.4 below. When compared with the un-coded performance, ZFBF has 17 dB improvement, modified THP has 14.6 dB improvement and vector perturbation has 7 dB improvement to get a BER of 10^{-3}.

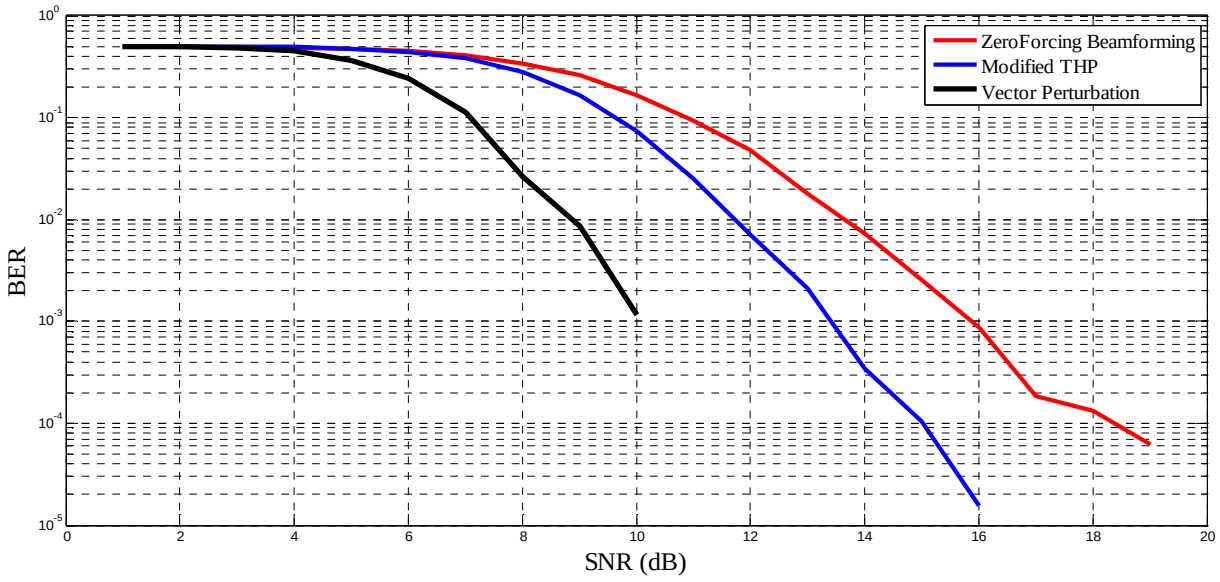

Figure 4.4: BER comparison of the three algorithms for coded and interleaved 16 QAM.

For coded systems, the performance difference between the algorithms is reduced compared to the un-coded system. The power gain of Modified THP over ZFBF has reduced from 5 dB to 2.6 dB. Similarly, the power loss of Modified THP compared to VP has reduced from 11dB to 3.4 dB to get BER of 10^{-3}. The above results are the same as the results in [4] and [7] for single carrier systems.

4.2 Complexity Analysis

The complexity of the algorithms described above can be divided into preamble processing complexity and payload processing complexity. The preamble processing uses the channel state information and is done whenever new CSIT is available. The payload processing uses the result of preamble processing and it is performed for every symbol vector ready for pre-coding.

Using the mathematical description of the algorithms given in Sections 3.1 and 3.2, their complexity is given in terms of the number of real additions and real multiplication for both preamble and payload processing. In doing so, the following basic rules of complexity of matrix related operations are used [24], [40]:

- One complex addition requires 2 real additions, and one complex multiplication requires 4 real multiplications and 2 real additions.
- Multiplication of $M\ x\ N$ matrix with $N\ x\ P$ matrix requires $2MP(2N-1)$ real additions and $4MNP$ real multiplications if both matrices have complex entries.
- Inverting an $N\ x\ N$ square matrix requires $4N^3$ real additions and $4N^3$ real multiplications.
- Multiplication of an $M\ x\ N$ matrix with an N x 1 vector requires $4MN-2M$ real additions and $4MN$ real multiplications.
- QR decomposition of an $M\ x\ N$ matrix requires $(4MN^2-2MN-N^2)$ real additions and $(4MN^2+N)$ real multiplications using the modified Gram-Schmidt algorithm.
- Finding the minimum of N real numbers requires N-1 real additions.
- Modulo arithmetic operation of a complex number requires 10 real additions and 2 real multiplications if the floor operator searches in the interval [-8, 8].

Finding the appropriate offset vector in vector perturbation using Eq. 3.24 is performed by searching through the available integer vectors within a limited radius. The exact amount of iterations on each layer is determined by counting the number of nodes visited using simulation. Therefore, a system with 4 transmitting antennas and 4 single antenna users is taken to compare the complexity of the algorithms.

Complexity of Zero Forcing Beamforming (ZFBF) and Modified THP

For the system at hand, ZFBF involves simple channel inversion in preamble processing and matrix-vector multiplication in payload processing. As described in Figure 3.2, the pre-coding process in modified THP involves QR decomposition, successive interference pre-subtraction, modulo operation and matrix-vector multiplication. Only the QR decomposition is part of the preamble processing while the rest are payload processing steps. The complexity of the pre-subtraction process when user 1 is pre-coded first is given in Table 4.2 below.

Table 4.2: Number of real additions and multiplications for the successive pre-subtraction step of Modified THP

	Real Additions	Real Multiplications
User 1	10	2
User 2	14	6
User 3	18	10
User 4	22	14

Complexity of Vector Perturbation

This algorithm first finds the integer offset vector for each symbol vector and multiplies the perturbed vector with the channel inverse as described in eq. 3.23. Finding the integer offset vector using the minimization problem of eq. 3.24 can be rewritten using QR decomposition of $\boldsymbol{H}^{-1}$ as

$$\begin{aligned} \boldsymbol{l} &= \arg\min_{\boldsymbol{l}'} \ (\boldsymbol{u} + \tau\boldsymbol{l}')^{*} \ (\boldsymbol{Q}\boldsymbol{R})^{*}(\boldsymbol{Q}\boldsymbol{R})(\boldsymbol{u} \ + \ \tau\boldsymbol{l}') \\ &= \ \arg\min_{\boldsymbol{l}'} \ \|\boldsymbol{R}(\boldsymbol{u} \ + \ \tau\boldsymbol{l}')\|^{2} \end{aligned} \tag{4.2}$$

Letting $\boldsymbol{v} = \left(\boldsymbol{u} \ + \ \tau\boldsymbol{l}^{'}\right) = [\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ \boldsymbol{v}_3 \ \boldsymbol{v}_4]^{T}$ and noting that $\boldsymbol{R}$ is upper triangular, eq. 4.2 can be written as

$$\boldsymbol{l} = \arg\min_{\boldsymbol{l}'} \left\| \begin{matrix} R_{11}v_1 + R_{12}v_2 + R_{13}v_3 + R_{14}v_4 \\ R_{22}v_2 + R_{23}v_3 + R_{24}v_4 \\ R_{33}v_3 + R_{34}v_4 \\ R_{44}v_4 \end{matrix} \right\|^2 \qquad (4.3)$$

In solving the minimization of eq. 4.3, initial radius is set to infinity and is updated by the first iteration. But in performing the next iteration, the fourth value $\|R_{44}v_4\|^2$ is computed first, and the third value is added only if $\|R_{44}v_4\|^2$ is less than the radius set by the first iteration. If not, the algorithm quits the loop and proceeds to test the next vector. The number of times each layer is visited for Rayleigh distributed channel gains is shown in Table 4.3. The result is averaged over 1000 channel realizations.

Table 4.3: Number of times layers of Vector perturbation are visited

	Layer 1	Layer 2	Layer 3	Layer 4
Number of times visited	673.52	376.07	106.33	16

After $\boldsymbol{l}$ is determined, $\boldsymbol{d} + \tau\boldsymbol{l}$ will be computed and multiplied with the channel inverse which completes the payload processing. In the pre-coding process, performing QR decomposition and inverse of the channel matrix are part of the preamble processing.

The overall complexity of the algorithms for preamble and payload processing is shown in Table 4.4. The complexity of vector perturbation is much higher than the rest two for the payload processing. In addition, payload processing of VP also has variable number of computation because the number of nodes visited at each layer depends on the data vector. This is prohibitive for implementation because it is performed for every data vector.

Table 4.4: Overall complexity of the different pre-coding Algorithms

	Preamble Processing		Payload Processing	
	Real Additions	Real Multiplications	Real Additions	Real Multiplications
Zero Forcing Beamforming	256	256	56	64
Modified THP	432	516	120	96
Vector Perturbation	464	516	9660	7328

The channel matrix used for pre-coding is updated whenever new CSIT arrives. The number of symbol periods between the arrival of two consecutive CSIT depends on the time variation of the channel and the channel estimation technique used. For comparison purposes, it is assumed that a new CSIT is available every 100 symbol periods. The number of real additions required to compute a single real multiplication depends on the number of bits used to represent numbers. If 8 bit operations are used as a basis, the complexity of a single real multiplication is roughly equivalent to 10 times that of a single real addition [24]. Using the above assumptions, the complexity of the pre-coding algorithms for one quasi-static interval in terms of equivalent additions is shown in Table 4.5.

Table 4.5: Complexity of the pre-coding algorithms for one quasi-static interval

	ZFBF	Modified THP	VP
Complexity in terms of equivalent Additions	72,416	113,592	8,299,624

As can be seen from Table 4.5, the complexity of Modified THP is about 1.56 times that of ZFBF while the complexity of VP is about 114 times that of ZFBF. Thus, Modified THP offers a better performance than ZFBF with comparable computational complexity. Although VP is

shown to perform much better than the two algorithms, less complex perturbation algorithms are necessary to practically implement this algorithm. The algorithm can also be made less complex by reducing the search radius of the perturbation equation but at the cost of reduced performance.

5.3. Effect of Fading Path Correlation

The other issue that is considered is the effect of fading path correlation on the BER performance of the pre-coding algorithms. As discussed in section 2.2.2, the Kronecker model is simple and relatively accurate [30] in modeling correlated channels. Explicitly specifying the one sided correlation matrices in this model requires 32 parameters for a system with 4 transmitting antennas and 4 single antenna users.

The number of parameters of the transmitter side correlation matrix can be decreased for a linear antenna array by assuming that the correlation between transmitting antennas decrease exponentially as their separation increases. This assumption is shown to have a unique match in capacity with the model that explicitly states the correlation for each transmit-receive antenna pair [24]. The cause for correlation might also be poor scattering nature of the channel though its value depends on spacing between transmit antennas. The transmitter side correlation matrix is symmetric and can be written as

$$R_{Tx} = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix} \tag{4.4}$$

In a MIMO broadcast system with one receive antenna per user, the users terminals are generally located at different places. This implies that the scatterers around one receiver have different spatial structure from another receiver. Therefore, the channel gains from one transmit antenna to the different receivers can be taken as uncorrelated. Therefore, the receiver side correlation matrix reduces to an identity matrix.

The above model is formulated for a narrowband system. Since the simulations in this chapter are for a wideband channels, the correlative model should be extended for these channels. The

wideband Kronecker model applies the technique explained in Section 2.2.2 for each tap of the filter channel as

$$G_{corr}[d] = R_{Rx}^{1/2}[d] G_{iid}[d] \left(R_{Tx}^{1/2}[d]\right)^{T} \tag{4.5}$$

$G_{iid}[d]$ refers to the matrix of path gains for all transmit-receive antenna pairs that arrive at delay d. $R_{Rx}[d]$ and $R_{Tx}[d]$ are the receiver and transmitter side correlation matrices of paths at delay d and $G_{corr}[d]$ is the resulting matrix of correlated path gains for paths having delay d. The wideband kronecker model reduces to the narrowband case when there is only one tap.

In Figure 4.5, transmitter side correlation, $\rho = 0.6$, is used to plot the BER performance difference of the DPC algorithms for correlated and uncorrelated channels. For both algorithms, there is a power loss of about 1.6 dB to get a BER of 10^{-3} and this loss is almost constant at all SNRs. The simulation setup of Section 4.1.1 is used along with the parameters described on Table 4.1 for un-coded system.

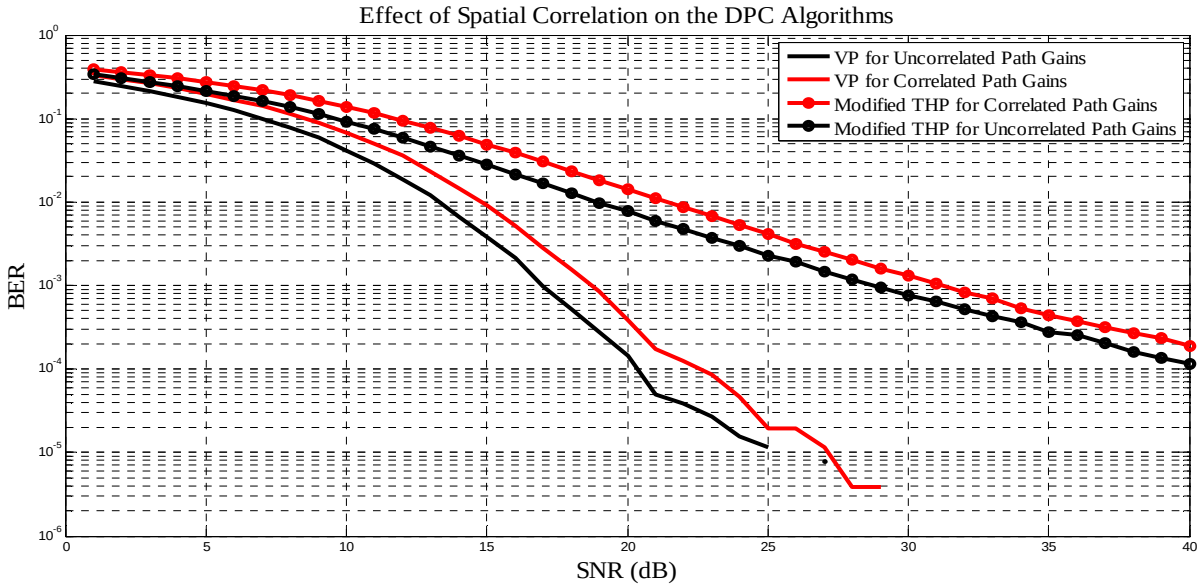

Figure 4.5: BER comparison of the DPC algorithms for correlated and uncorrelated channels.

The transmitter side correlation coefficient is likely to have different values depending on the wireless environment. Figure 4.6 shows the performance loss of the pre-coding algorithms for different values of correlation coefficient at a target BER of 10^{-3} for uncoded 16 QAM.

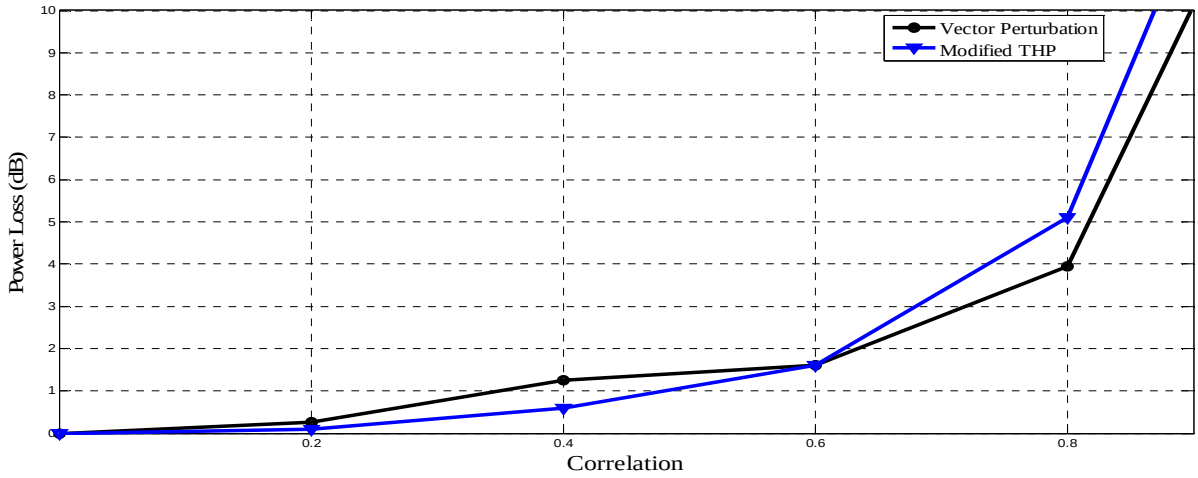

Figure 4.6: Performance loss of the DPC algorithms for un-coded 16 QAM.

As can be seen in Figure 4.6, Modified THP is less sensitive than VP for correlation values below 0.6 while it is more sensitive than VP for correlation values above 0.6. Generally, both algorithms lose 3 dB above a correlation value of about 0.7 after which the performance degradation is unacceptable.

In order to check the effect of correlation for realistic channels, the channel model proposed by TGn (Task Group n) [38] can be used. There are five models proposed and Model E which has an rms delay spread of 100 nsec is used in the simulation. The correlation matrix is generated independently for each tap so as to match experimental values [41] and the antennas are assumed to be spaced half wavelength apart. The performance of the DPC algorithms for this channel model is shown in Figure 4.7. There is a loss of about 1.5 dB for both algorithms compared to the uncorrelated performances of Figure 4.3.

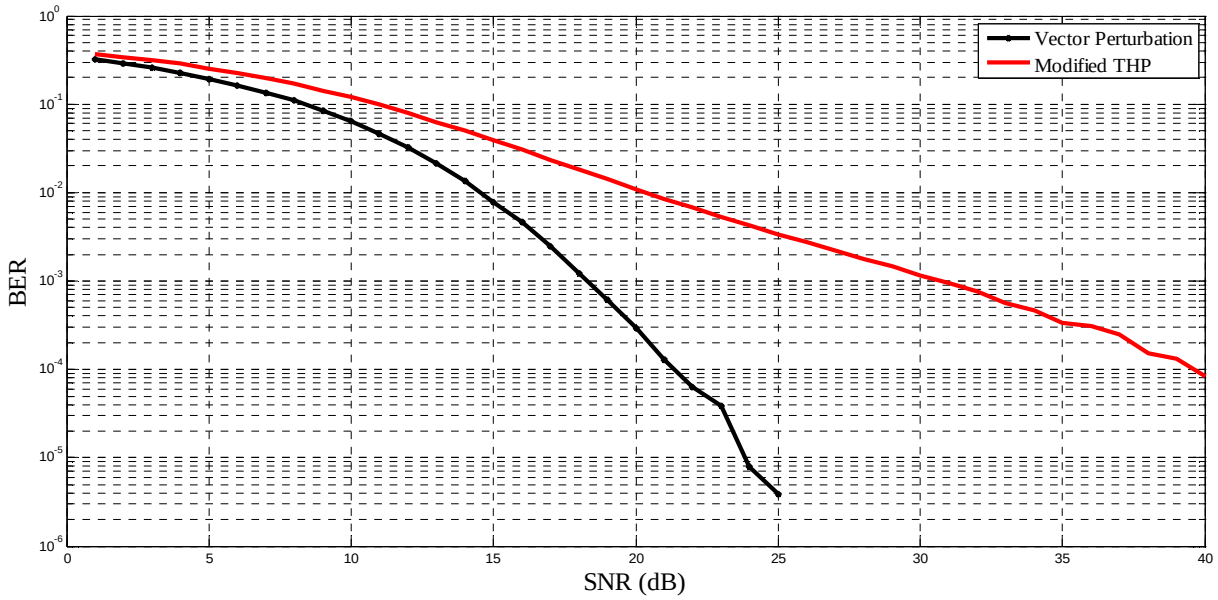

Figure 4.7: BER performance of the DPC algorithms for Model E of TGn channel models for un-coded 16 QAM.

Chapter 5

Results on Implementation Issues

In this chapter, the effect of imperfection in CSIT on performance of the pre-coding algorithms and the performance of the user scheduling algorithms will be dealt with. These issues are important when incorporating the pre-coding techniques in practical wireless systems. Obtaining up to date and error free CSI at the transmitter is not feasible using current channel estimation techniques. Therefore, it is important to investigate the performance loss incurred by modeling and simulating the imperfections in CSIT based on its obtaining mechanism. The other topic is scheduling users when the number of users is greater than the number of transmitting antennas. This is usually the case in practical systems like wireless LANs where there are much more users than the antennas on the Access point.

5.1 Effect of Imperfect CSI at the Transmitter

Among the different causes of imperfect CSIT, error in channel estimation and delay in delivering the channel estimate to the receiver are considered in this work. The other cause in FDD systems is the capacity of the feedback path. Due to the finite rate of the feedback path, the CSI to be transmitted back to the receiver should be quantized. This quantization results in quantization error which will have considerable impact on performance no matter how accurate and timely the CSIT is. Also, the feedback path in FDD systems is assumed to be error free.

5 .1.1 Model and Simulation Setup

As discussed in Section 2.2.3, the mechanisms of obtaining CSIT are different for TDD and FDD systems and so is the cause for imperfection in CSIT at the transmitter. In TDD systems, the uplink channel is estimated at the transmitter and this estimate is the same for the downlink. Therefore, imperfection in CSIT can be modeled by channel estimation error of the uplink. The channel estimation error can be modeled by the assumption that the true channel, $\boldsymbol{h}[k]$, and its estimate, $\widehat{\boldsymbol{h}}[k]$, are jointly Gaussian. This assumption is well justified for many practical

estimation techniques such as additive channel estimation, MMSE channel estimation and pilot symbol assisted estimation [42]. Using this model,

$$\widehat{\boldsymbol{h}}[k] \;=\; \rho_e \boldsymbol{h}[k] + \sqrt{(1-|\rho_e|^2)\sigma_h^2}\;\boldsymbol{\varepsilon}[k] \tag{5.1}$$

In the above equation, $\rho_e = \frac{E\{\boldsymbol{h}_i[k]\widehat{\boldsymbol{h}}_i^*[k]\}}{\sigma_h^2}$ is the correlation between the true channel and its estimate, σ_h^2 is the variance of the channel gains, $\boldsymbol{\varepsilon}[k]$ is error vector with i.i.d entries and $\boldsymbol{\varepsilon} \sim \mathcal{NC}(0,1)$. The value of the correlation coefficient, ρ_e, depends on the specific channel estimation technique used and the noise level in the channel.

As of our knowledge, an explicit relation that gives the value of the correlation as a function of SNR for the available channel estimation techniques is not found yet. The only way to get the exact effect of errors in channel estimation is to perform channel estimation which is outside the scope of this work. In [43], it has been shown that for a time varying Rayleigh fading channel with pilot symbols used for channel estimation having the same sampling rate as the data symbols, the highest value of correlation that can be attained is:

$$\rho_e = \sqrt{\frac{SNR}{1+SNR}} \tag{5.2}$$

In FDD systems, CSIR is first obtained by estimating the downlink and this channel estimate is fed back to the transmitter through a separate channel. In this case there are two causes of CSIT imperfection. First, the channel estimate of the downlink might be erroneous. This can be modeled using the approach used in TDD systems. In addition, there is a delay in delivering the channel estimate at the receiver to the transmitter. This is due to the time taken by the receiver in estimating and quantizing the estimates, and the time taken by the transmitter in decoding the estimates and performing user selection among others. This delay makes the available CSIT to be outdated because of the time variation of the wireless channel.

In the simulation, the channel is modeled as time varying and narrowband. Under the presence of delay in CSIT, the quasi-static assumption is not valid. This is because it assumes the channel to

be constant for some time and changes to an independent realization afterwards while the basis for quantifying the performance loss under delayed CSIT is the time evolution of the channel gains. The narrowband assumption is justifiable by using OFDM technique as used in Section 4.1.1.

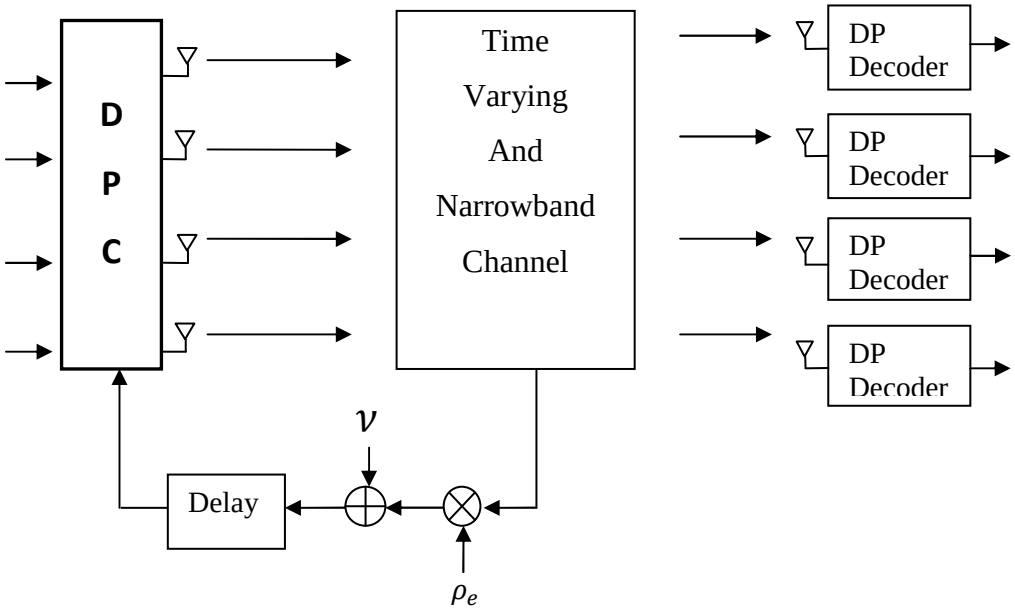

Figure 5.1: Simulation Setup for FDD systems.

Block diagram of the simulation setup for FDD systems is shown on Figure 5.1 above. Parts of the system that are not shown in the diagram are the random bit generator, channel coding, interleaving and modulator, and their receiver side counterparts. All these parts are included in this simulation but excluded from the block diagram for clarity, and are shown in Figure 4.1. The input to the pre-coder is, thus, modulated symbols and the output from the dirty paper decoder (DP decoder) is the received symbols. The simulation setup for TDD systems is exactly the same except the omission of the delay block.

The channel varies according to Jakes' model whose normalized Doppler power spectrum is analytically given as [44],

$$S(f) = \begin{cases} \dfrac{1}{\pi f_d \sqrt{1-\left(f/f_d\right)^2}} & ,|f| < f_d \\ 0 & elsewhere \end{cases} \tag{5.3}$$

where f_d is the maximum Doppler frequency. Scatterers and/or users in the system are assumed to have a maximum speed of 10 km/hr. which makes the maximum Doppler frequency 50 Hz at an operating frequency of 5 GHz.

Using the narrowband assumption, the complex channel gain from each transmit antenna to each receive antenna is generated so that both the real and imaginary parts are Gaussian distributed having zero mean and variance of $1/\sqrt{2}$. Thus, the amplitude is Rayleigh distributed and the phase is uniformly distributed in $[0, 2\pi]$. Channel estimation error as modeled in eq.(5.1) is shown in Figure 5.1 by the multiplier and the adder. The value of σ_h^2 in eq.(5.1) is unity and $\boldsymbol{\nu}[k] = \sqrt{(1-|\rho_e|^2)\sigma_h^2}\,\boldsymbol{\varepsilon}[k]$. The parameters of the system used in the simulation are listed in Table 5.1 below.

Table 5.1: System Parameters used for simulating effect of imperfect CSIT

Channel Model	Time varying and narrowband Channel
Doppler Spectrum	Jakes' model
Maximum Doppler frequency	50 Hz
Carrier Frequency	5 GHz
Sampling Frequency	20 MHz
Modulation Type	16 QAM
Channel coding	Convolutional code, code rate = 1/2
Interleaver	Random interleaver

The feedback delay denotes the time lag between measuring the CSI and using it to perform pre-coding. In current wireless systems, there is a feedback delay of up to 2-6 msec depending on the system and other resource scheduling components that contribute to increased delay [10].

5.1.2 Simulation Results

Using the simulation setup of Figure 5.1 and the parameters of Table 5.1, the performance loss of the DPC algorithms is shown in this section for TDD and FDD systems independently. Figure 5.2 shows the effect of channel estimation error on BER performance of the pre-coding algorithms for TDD system. Using the model of eq. (5.1), the BER performance of VP is shown in Figure 5.2 for different values of correlation.

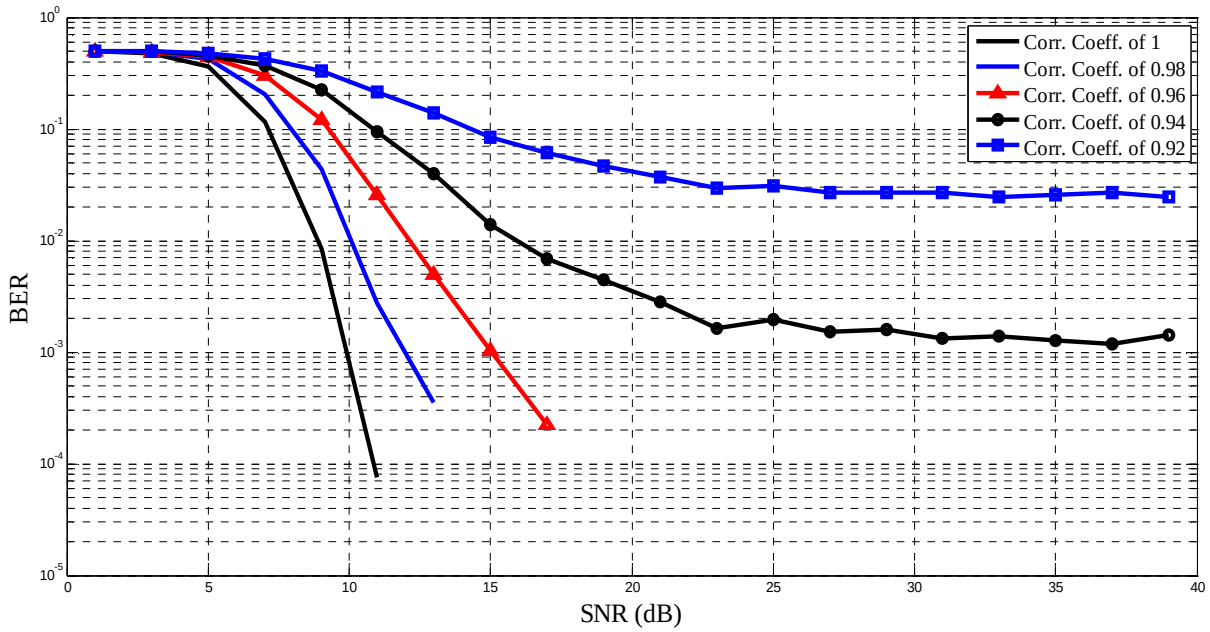

Figure 5.2: Effect of channel estimation error on VP.

As can be seen from the above figure, the BER performance of VP is acceptable for correlation values above 0.96 which is practical for current channel estimation techniques. It is also shown that there is a power loss of 5 dB to attain a BER of 10^{-3} at a correlation value of 0.96 compared to the case of perfect CSIT.

The performance of Modified THP for different values of correlation is shown in Figure 5.3 for the same simulation setup. Modified THP is shown to be more sensitive to channel estimation errors than VP as it does not achieve a BER of 10^{-3} for correlation values below 0.98.

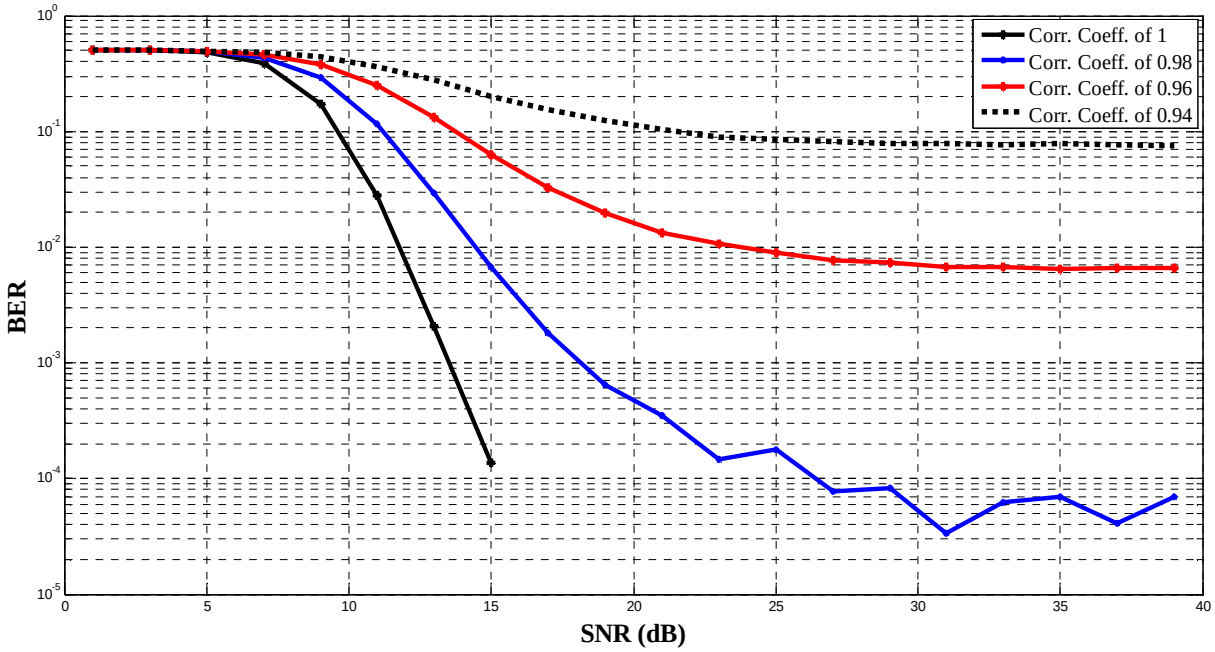

Figure 5.3: Effect of channel estimation error on Modified THP.

Using the upper limit on correlation for pilot symbol assisted estimation of Eq. (5.2), the power loss of the DPC algorithms as a function of SNR is shown in Figure 5.4. The power loss is relative to the case of perfect CSIT to achieve a target BER of 10^{-3}. The effect of channel estimation error is severe for SNR below 15 dB but is insignificant as SNR increases. In order to reduce the effect of estimation error, the SNR should be kept above 15 dB for both algorithms.

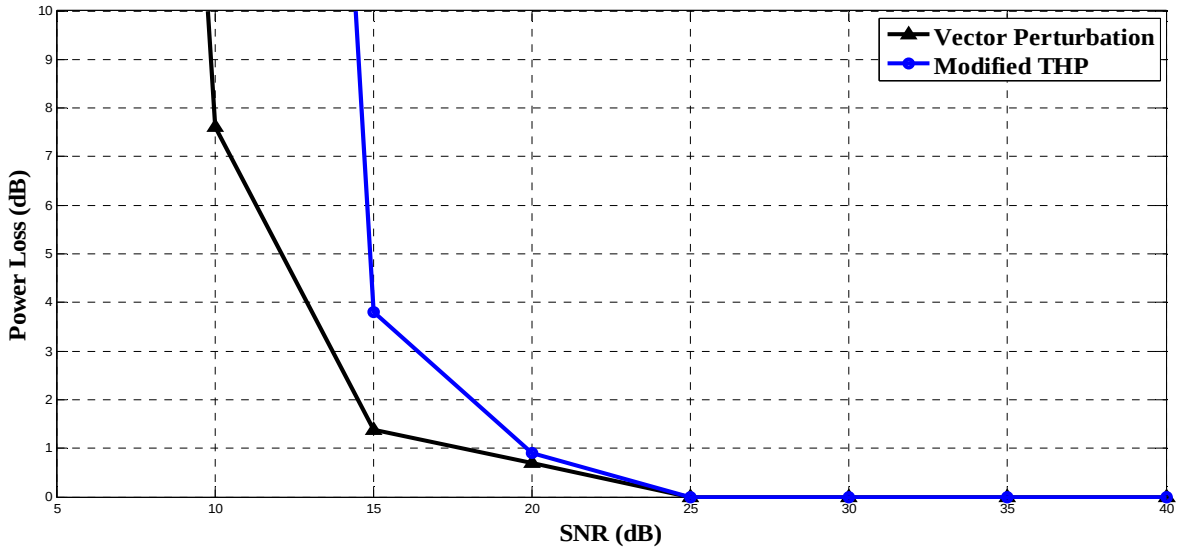

Figure 5.4: power loss of the DPC algorithms due to channel estimation error.

For FDD systems, the correlation between the true channel and its estimate is taken to be 0.995 for all simulations in order to stress on the effect of delay. At a sampling frequency of 20 MHz, delays in feedback of 2-6 msec translate to delays of 40000-120000 samples. Figure 5.5 shows the BER performance of VP for feedback delays ranging from 2-6 msec for the simulation parameters of Table 5.1.

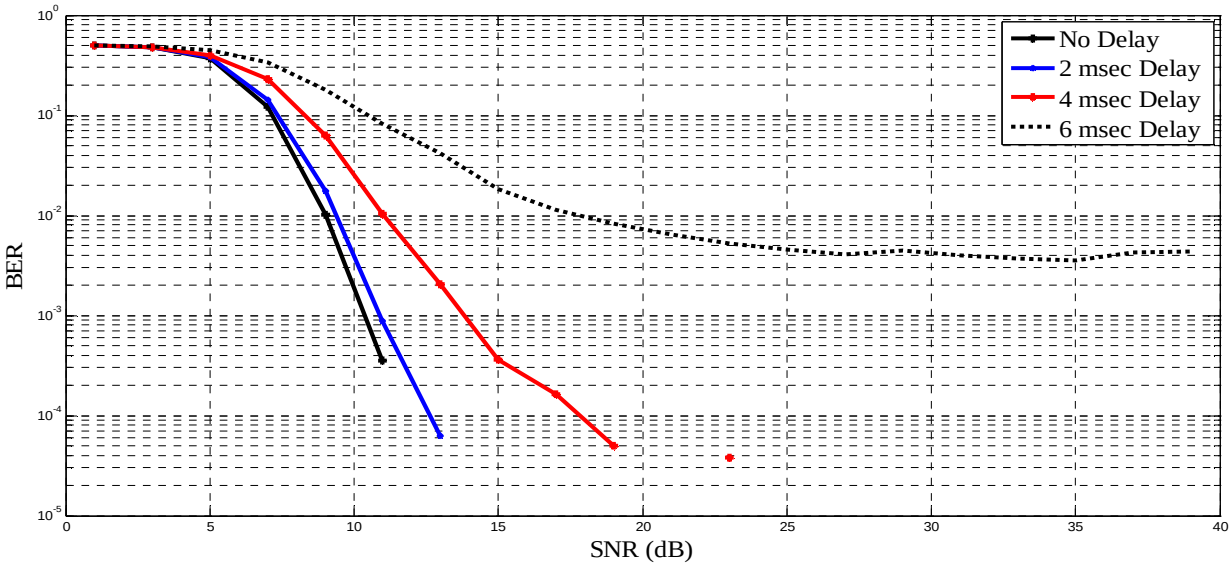

Figure 5.5: Effect of feedback delay on BER performance of VP.

Compared with the case of perfect CSIT, the BER performances of the pre-coding algorithms is considerably degraded. It can be seen that VP do not attain a BER of 10^{-3} for feedback delays of about 6 msec and above but it is applicable if the delay can be made smaller than 4 msec.

Similarly, the BER performance of Modified THP for different values of delay in CSIT is shown in Figure 5.6 below. Unlike VP, Modified THP do not attain a BER of 10^{-3} for feedback delays of about 4 msec and above. It is applicable for delays below 2 msec which is a difficult constraint for practical implementation.

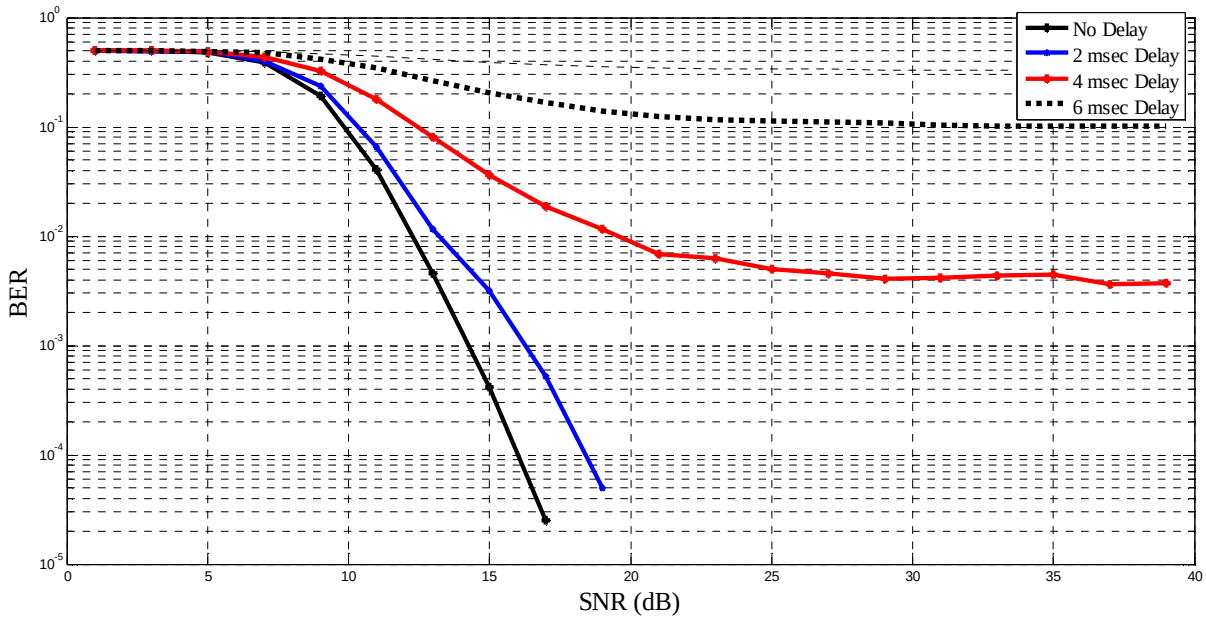

Figure 5.6: Effect of feedback delay on BER performance of Modified THP.

To get a clear view of the sensitivity of these algorithms to feedback delay, Figure 5.7 shows the power loss of the pre-coding algorithms to achieve a target BER of 10^{-3} for different values of delay in feedback when the correlation between the true channel and its estimate is $\rho_e = 0.995$.

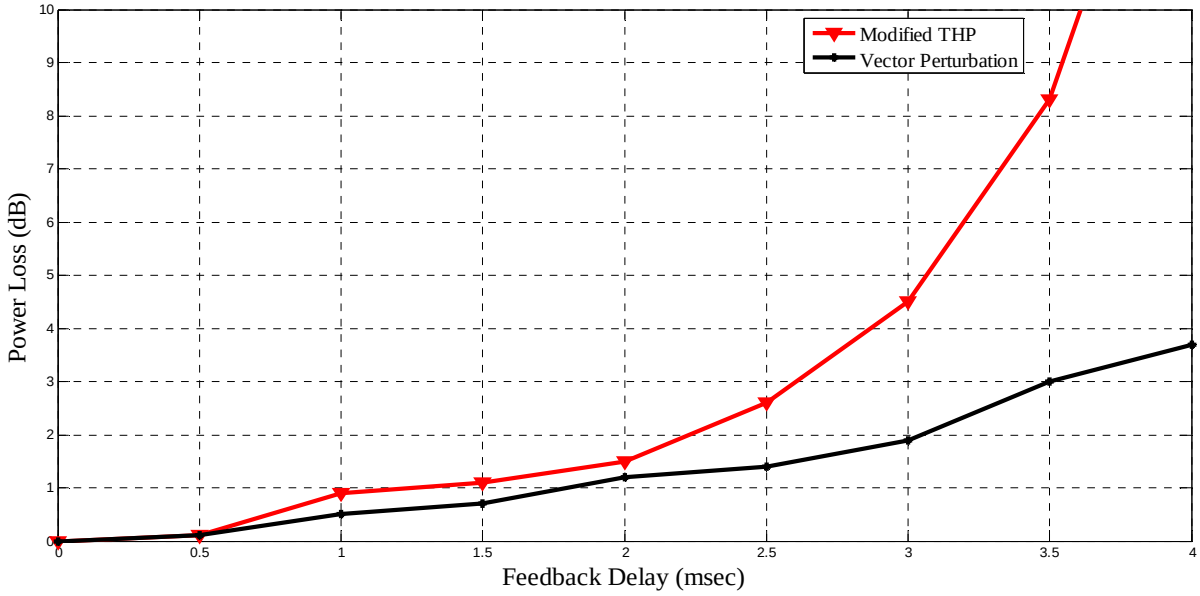

Figure 5.7: power loss for different values of feedback delay to achieve a target BER of 10^{-3}.

As can be seen from Figure 5.7, Modified THP is more sensitive than VP for the whole range of delay values and results in a 3 dB power loss at a delay of 2.5 msec. Vector perturbation, on the other hand, has much lower sensitivity and has acceptable performance for delay values up to 4 msec.

It can be concluded that feedback delay in FDD systems has considerable effect on the performance of DPC algorithms. This indicates the need to have a means for fast delivery of CSIT to reduce its effect to the minimum. Other approach to reduce the effect of delayed CSIT is to use techniques that estimate the current state of the channel using the CSIT delivered and the statistics of the channel. This is possible because the state of the channel can be modeled as an autoregressive (AR) process whose coefficient can be determined using the autocorrelation function of the channel gains thus reducing the effect of feedback delay. The autocorrelation function can in turn be determined from the Doppler spectrum of the wireless environment.

5.2 Performance of User Scheduling Algorithms

In addition to its uses in cases where there are more users than the number of transmit antennas, user scheduling is also the first step in rate allocation. Rate allocation in multiuser systems consists of user scheduling and assigning variable rates to the selected users/subcarriers depending on the channel strength. The problem of user scheduling has different formulations for single carrier and multicarrier systems as discussed next.

5.2.1 Performance of User Selection Algorithm

In single carrier systems, user scheduling reduces to user selection since each user has one channel vector. After the user selection step, variable rates can be assigned to users based on their channel gains. The improvement in sum rate due to user selection is discussed here while assigning variable rate is left as a future work.

The elements of the channel vectors of the different users are complex valued random numbers having Rayleigh distributed magnitude with unit variance and uniformly distributed phase. The system has 40 users with single transmit antenna and transmitting station with 4 antennas. The results are averaged over 1000 channel realizations. The sum rate of different user groups using the proposed algorithm in Section 3.5.1 is shown in Figure 5.8.

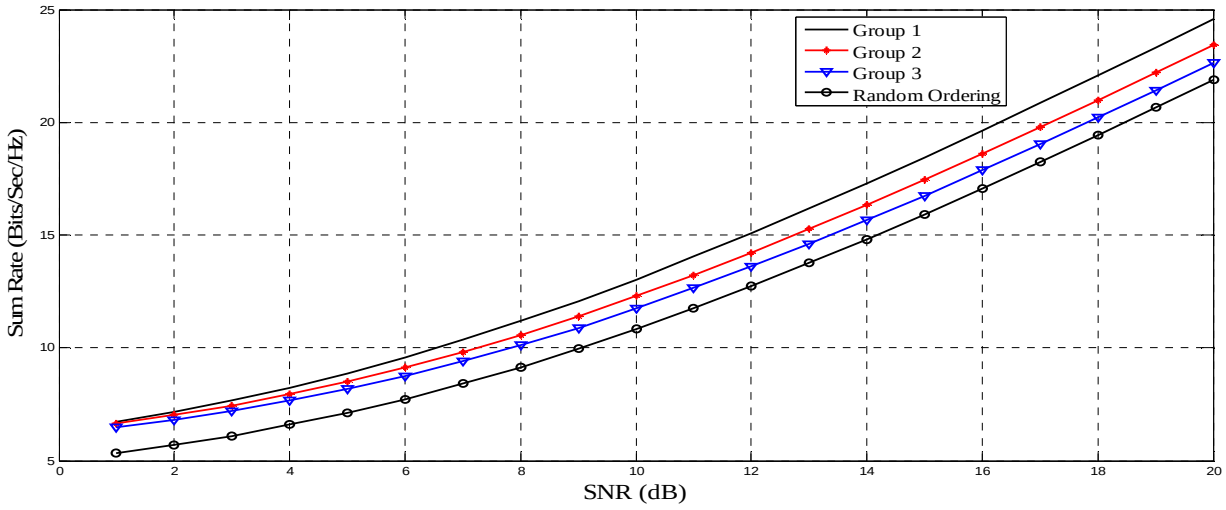

Figure 5.8: Achievable sum rate of user groups for the proposed algorithm using Modified THP.

Also shown is the sum rate of the system if the users are grouped randomly and time slots are assigned equally to all user groups. This scheme is labeled random ordering. If the first group is always served, the sum rate is shown to be greater than that of random ordering. The same is true for the second and third user groups. The other groups achieve sum rates below that of random ordering.

To include fairness, time slots can be assigned proportional to the strength of the user groups but the sum rate of the resulting system will obviously be lower than that of group 1. The sum rate performance of Vector Perturbation using the proposed algorithm is shown in Figure 5.9. Similar to modified THP, the sum rate of the first three user groups is above that of random ordering while the other groups have lower sum rates than random ordering.

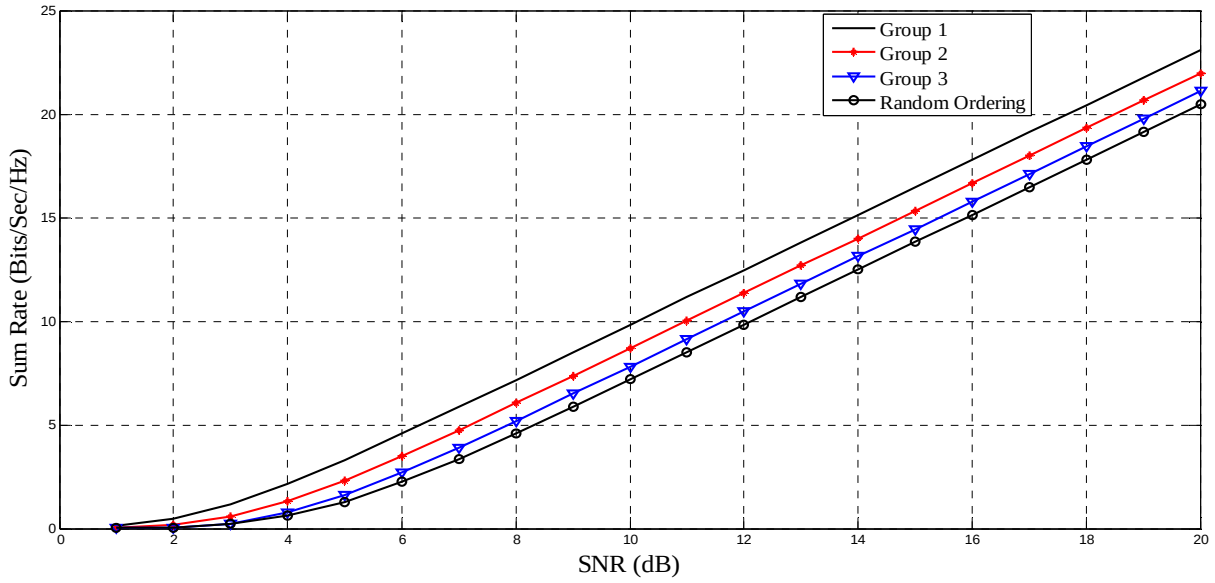

Figure 5.9: Achievable sum rate of user groups for the proposed algorithm using VP.

The fairness index of the system as a function of number of realizations is shown in Figure 5.10 below. For the random ordering case, the fairness index is always unity because all users are served equally. The fairness index of the system when the first group is always served is also shown. The fairness index of proposed algorithm is much lower than that of random ordering although it approaches random ordering as the number of channel realizations increase.

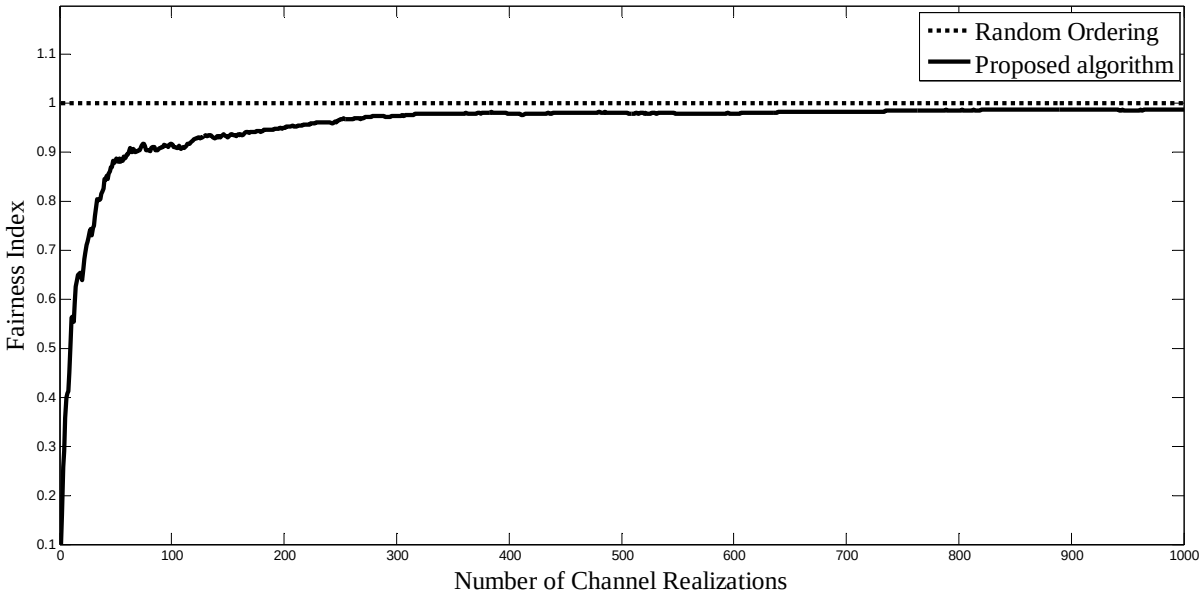

Figure 5.10: Fairness index of random ordering and the proposed algorithm.

5.2.2 Performance of Subcarrier Allocation Algorithm

In multicarrier systems like OFDM, each user has as many channel vectors as the number of subcarriers in the system. Thus, user scheduling can be formulated as a subcarrier allocation problem because users have different channel gains at different subcarriers. The parameters of the simulation are summarized in Table 5.2 below.

Table 5.2: Simulation Parameters for performance comparison of subcarrier allocation algorithms

Number of Transmitting Antennas	4
Number of users	40
Number of Subcarriers	64
Channel Model	Discrete-Time, Quasi-Static Wideband Channel
Distribution of Tap Weights	Rayleigh
Number of Taps of the channel	18
Number of Channel Realizations	100

In Figure 5.11, the achievable sum rate of the different algorithms for subcarrier allocation are plotted for Modified THP. The sum rate of RMSA is shown to exceed both FOSA and random ordering of subcarriers and users. At SNR of 10dB, there is a gain of about 1.67 bits/sec/Hz in using RMSA instead of FOSA. FOSA has improved gain in sum rate when compared with that of random ordering for increasing SNR.

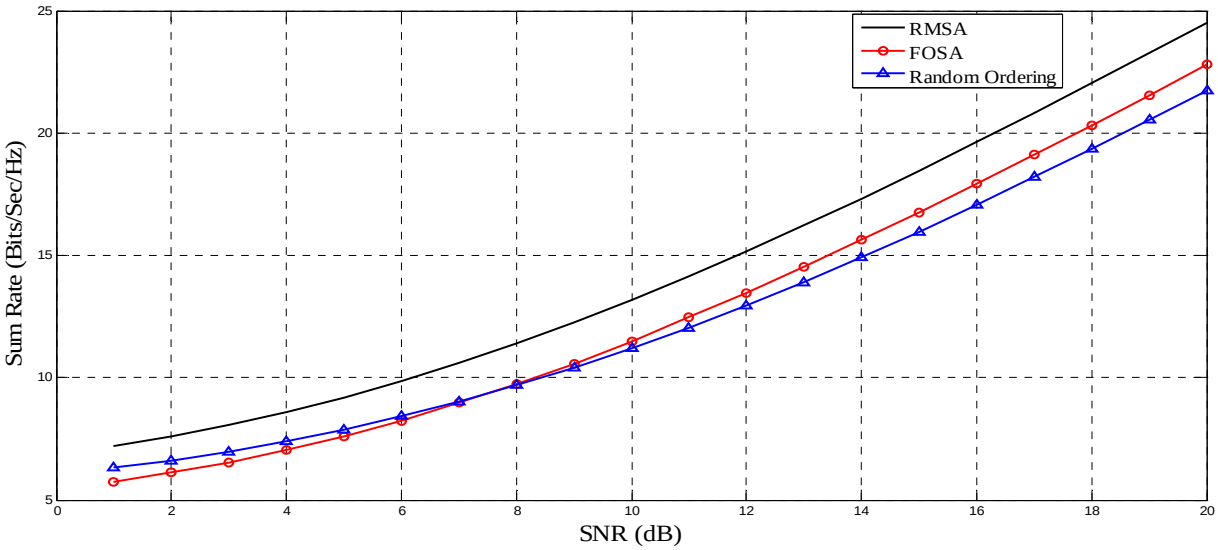

Figure 5.11: Achievable sum rate of RMSA and FOSA algorithms for Modified THP.

The achievable sum rate of Vector Perturbation for the subcarrier allocation algorithms is shown in Figure 5.12. Similar to Modified THP, RMSA algorithm has about 2 bits/sec/Hz improvement over FOSA at SNR of 10 dB. But the performance of FOSA is shown to be better than random ordering at all SNRs unlike the case of Modified THP.

Figure 5.13 shows the fairness index of RMSA and FOSA for 100 OFDM symbol periods. The FOSA algorithm is a modified round robin which assigns nearly the same amount of subcarriers for all users. Therefore the fairness index is shown to be 0.9942 for almost all of the symbol indices. The RMSA algorithm has a much lower fairness index which varies from one symbol period to the other but improves as number of OFDM channel realizations increase.

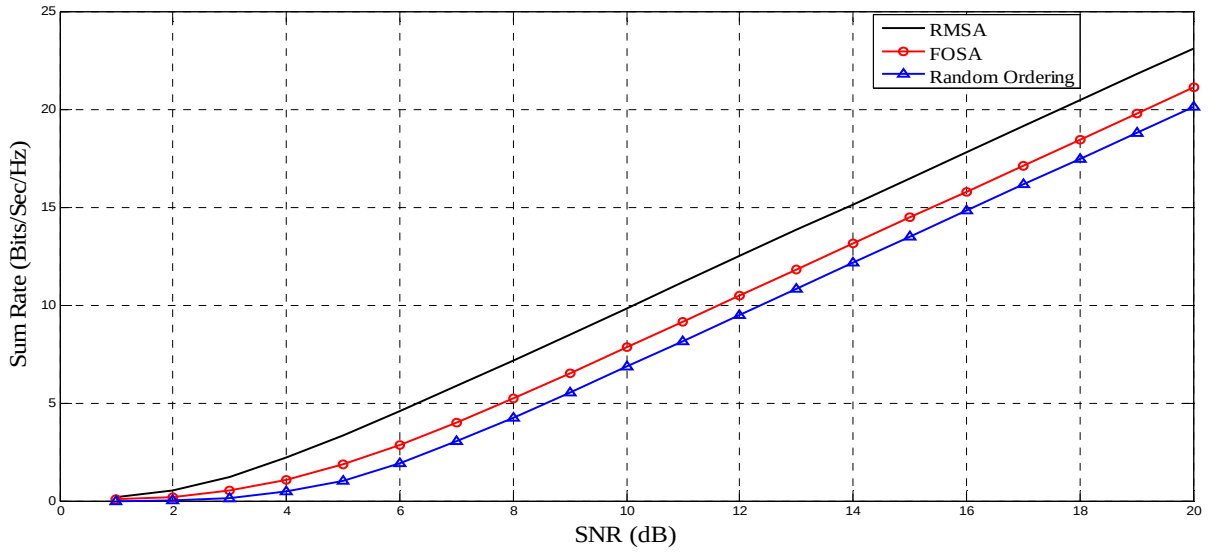

Figure 5.12: Achievable sum rate of RMSA and FOSA algorithms for VP.

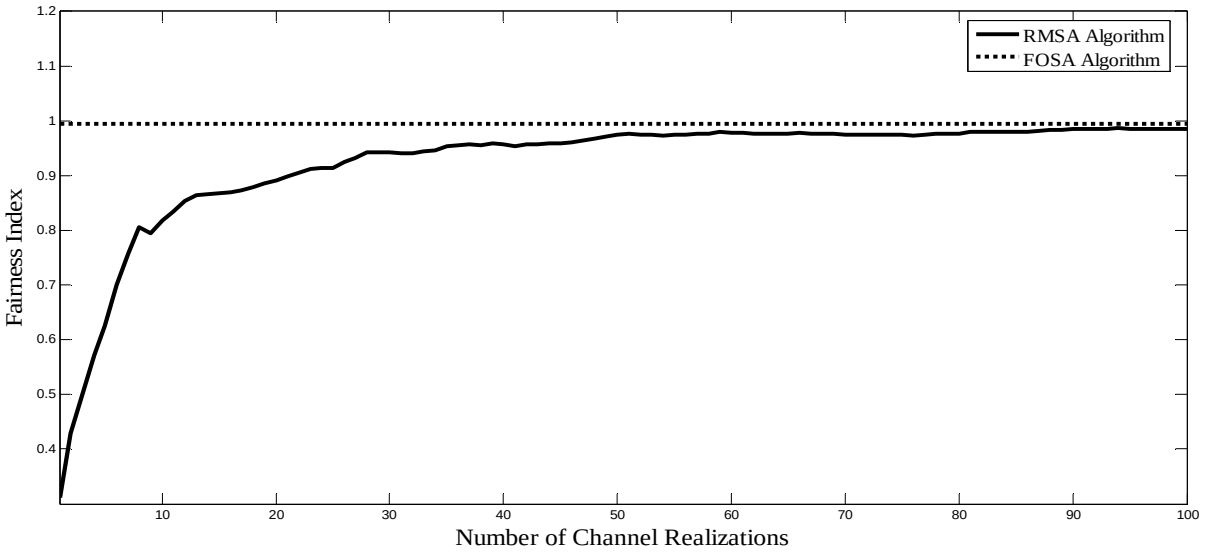

Figure 5.13: Fairness index of RMSA and FOSA algorithms for OFDM Systems.

Chapter 6

Conclusion and Future Work

In this chapter, concluding remarks about the findings of the thesis and recommendations for future works based on the ideas discussed here or possible alternatives that could strengthen the issues in this thesis are presented.

6.1 Conclusion

A theoretical analysis of DPC has been analyzed for quite some time although practical techniques for achieving the limits set by theoretical findings are still a challenge. In this work, issues concerning practical implementation of the suboptimal algorithms under the DPC framework are dealt with. These algorithms, Modified THP and Vector Perturbation (VP), are found to achieve considerable improvement over the linear pre-coding techniques. While Modified THP has comparable complexity with ZFBF, the complexity of VP is very high compared with both Modified THP and ZFBF but can be made less complex by limiting the search radius depending on the level of complexity that can be handled.

Even though application of the DPC algorithms for MIMO BC is formulated for narrowband systems in previous works, they are found to be equally applicable for frequency selective channels. This is possible by using multi carrier system like OFDM and performing pre-coding in frequency domain for each subcarrier independently. In MIMO BC, the user terminals are generally located at different places which makes the receiver side correlation between fading paths to be insignificant. The effect of transmitter side correlation is shown to have little influence on the performance of the DPC algorithms for realistic channel parameters.

Obtaining perfect transmitter side CSI is probably the most difficult task in implementing pre-coding for MIMO BC. Realistic channel estimation techniques have imperfections due to channel estimation errors and/or delay in feedback of channel estimates. These causes are shown to have considerable influence on the performance of the DPC algorithms, and Modified THP is more sensitive to CSIT imperfections than VP. For TDD systems, error in channel estimation is

the cause of CSIT imperfection and the performance degradation is shown to have less and less influence as SNR increases. For FDD systems, CSIT imperfection is more severe as it is caused by a delayed and erroneous feedback of channel estimates. The performance of MIMO BC employing FDD is degraded as the delay in estimation feedback increases because the channel varies with time. This results in unacceptable BER performance above some value depending on the DPC algorithm used.

When it comes to practically implementing the DPC algorithms, user scheduling is also important because all users can not be served simultaneously. This is due to the limited number of spatial streams that can be created which is equal to the number of transmitting antennas. In single user systems, one criterion is to select users with relatively good channel gains and orthogonality among themselves. This is shown to result in increased sum rate of the system when compared with the scheme that assigns time slots equally to all users. In effect, the BER performance is will get better for the same spectral efficiency if this scheme is implemented. Fairness among users can be achieved by allocating time slots proportional to the strength of the user groups.

For OFDM systems, user scheduling reduces to subcarrier allocation that aims at either maximizing sum rate or achieving better fairness. Implementing the proposed user selection algorithm for each subcarrier results in increased sum rate and hence named Rate Maximizing Subcarrier Allocation (RMSA). Allowing users to select their favorite subcarrier turn by turn results in a lower sum rate than RMSA but improved fairness and hence named Fairness Oriented Subcarrier Allocation (FOSA).

6.2 Recommendations for Future Work

In addition to investigating issues that arise in the practical implementation of the DPC algorithms, this work opened numerous areas to be explored in the future. Some of these ideas are listed below:

- In this thesis, user terminals with only single receive antenna are considered. The case of having more than one antenna per user raises the question of whether or not transmitting independent data streams to each user is optimal. Thus, this scenario will be a significant extension to this work.
- Modified THP is shown to have BER performance that varies from user to user depending on their encoding order. This is not appropriate for practical implementation. Assigning different power or modulation order to users based on encoding order makes the BER performance of the users less varying. Investigating the optimal strategy to do this is an important contribution to the practicality of this algorithm.

- For ZFBF, regularization of the channel inverse using the MMSE criterion is shown to improve the BER performance if there is a means of estimating the SNR at each receiver [27]. Since Vector Perturbation performs beamforming after perturbation, analyzing the improvement obtained by regularizing the channel inverse using the MMSE criterion can be studied.

- For FDD systems, the CSI estimated at the receiver should be quantized and transmitted through a feedback channel that has limited capacity. This results in quantization error as the CSI is quantized into a finite number of codewords where the number of bits per codeword is limited by the capacity of the feedback channel. The effect of the number of bits used for quantization and the capacity of the feedback path are important issues that are not dealt in this work.

- Feedback delays in CSIT are shown to have significant influence on the performance of FDD systems. In single user systems, estimating the future state of the channel based on the Doppler spectrum of the channel with Autoregressive modeling is shown to improve

estimation errors. Achieving better quality of CSIT using similar approach will reduce the challenge of fast delivery of CSIT and is worth considering.

- In this thesis, selecting users or subcarriers to get improved sum rate or fairness is investigated. The issue of assigning variable rates to users in the same group based on their channel vector can improve the sum rate of the system and needs further study.

- DPC algorithms that take statistical properties of the channel called Channel Distribution Information (CDI) instead of deterministic CSI will relieve the transmitter of burdens of getting perfect channel estimates. Exploring new approaches or modifying the existing algorithms using this idea will ultimately be a new challenge that makes implementation of DPC one step closer.

References

[1] G. Caire and S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. on Information Theory*, vol. 43, pp. 1691–1706, July 2003.

[2] H.Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel", In *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, 2004.

[3] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.

[4] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multi-Antenna Multi-User Communication—Part II: Perturbation," *IEEE Trans. on communications,* Vol. 53, No. 3, March 2005.

[5] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, Vol. 7, pp. 138–139, March 1971.

[6] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. On Communications*, vol. 20, No. 4, pp. 774–780, August 1972.

[7] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multi-Antenna Multi-User Communication—Part I: Channel Inversion and Regularization," *IEEE Trans. on communications,* Vol. 53, No. 1, January 2005.

[8] U. Erez, S. Shamai, and R. Zamir, "Capacity and lattice strategies for cancelling known interference," in *Proceedings International Symposium on Information Theory and its Applications*, Honolulu, Hawaii, pp. 681–684, Nov. 2000.

[9] S. Alamouti, "Space time block coding: A simple transmitter diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 8, pp. 1451–1458, October 1998.

[10] M. Goldenbaum, R. A. Akl, S. Valentin, S. Stanczak, "On the Effect of Feedback Delay in the Downlink of Multiuser OFDM Systems", *45th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1-6, 2011.

[11] K. Kobayashi, T. Ohtsuki, T. Kaniko, "MIMO Systems in the Presence of Feedback Delay", *IEEE International Conference on Communications*, Vol. 9, pp. 4102-4106.

[12] W. Huang, K. Sun, F. Bai and Y. Zhao, "Comparison of Downlink Multi-User Multiple Input Multiple Output Schemes in Wireless Communication Systems", *2010 International Conference on Computer Application and System Modeling (ICCASM 2010),* College of Electronic Information Engineering, Inner Mongolia University, Hohhot, China.

[13] Z. Tu and R. S. Blum, "Multiuser Diversity for a Dirty Paper Approach," *IEEE Communications Letters*, vol. 7, No. 8, pp. 370–372, August 2003.

[14] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[15] A. Razi, D. Ryan, I.B. Collings, J. Yuan, "Sum Rates, Rate Allocation, and User Scheduling for Multi-User MIMO Vector Perturbation Precoding", *IEEE Transactions on wireless communications,* Vol. 9, No. 1, January 2010.

[16] Y. Shin, T. S. Kang, H. M. Kim," An Efficient resource allocation for multi-user MIMO OFDM systems with Zero-forcing Beamformer", *the 18th annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communication,* 2007.

[17] Li-chun Wang, "Multi-user MIMO OFDM and MIMO Broadcast system", National Chiao Tung University, Taiwan, 2009.

[18] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. on Information Theory*, Vol. 44, No. 2, pp. 744–765, March 1998.

[19] A. Goldsmith, "*Wireless Communications,*" Cambridge University Press, 2005.

[20] G. G. Raleigh and J. M. Cioffi, "Spatio-Temporal Coding for Wireless Communication," *IEEE Transactions on Communications*, vol. 46, no. 3, pp. 357–366, March 1998.

[21] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585-595, Nov. 1999.

[22] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311-335, 1998.

[23] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi element antennas", *Bell Labs Technical Journal*, Vol. 1, No. 2, pp. 41–59, Aug. 1996.

[24] A. Van Zelst, "MIMO OFDM for Wireless LANs", Ph.D. dissertation, Eindhoven University of Technology, Eindhoven, The Netherlands, April 2004.

[25] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints", *IEEE Trans. On Vehicular Technology*, Vol. 53, No. 1, pp. 18–28, Jan. 2004.

[26] W. Yu and J. Cioffi, "Sum capacity of a Gaussian vector broadcast channel," in *Proc. IEEE International Symposium on Information Theory*, p. 498, July 2002.

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

[28] J. G. Proakis, "*Digital Communications*", 4th Edition, New York, McGraw-Hill, McGraw-Hill Series in Electrical and Computer Engineering, 2000.

[29] A. R. S. Bahai, B. R. Saltzberg, and M. Ergen, "*Multi-carrier digital communications: theory and applications of OFDM"*. Springer, 2004.

[30] Nelson Costa and Simon Haykin, "*Multiple-Input Multiple-Output Channel Models: Theory and Practice*", John Wiley & Sons ,New Jersey, 2010.

[31] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity Limits of MIMO Channels," *IEEE J. Sel. Areas in Comm.*, vol. 21, No. 5, pp. 684–702, June 2003.

[32] T. L. Marzetta and B. M. Hochwald, "Fast Transfer of Channel State Information in Wireless Systems", *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. l268-1278, April 2006.

[33] Nihar Jindal, "MIMO Broadcast Channels with Finite-Rate Feedback," *IEEE Transactions on Information Theory*, Vol. 52, No. 11, November 2006.

[34] N. Jindal & A. Goldsmith, "Dirty Paper Coding vs. TDMA for MIMO Broadcast Channels," *IEEE International Conference on Communications*, June 2004.

[35] J. P. Kermoal, L. Shumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation", *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, August 2002, pp. 1211-1216.

[36] C. B. Peel, "Studies in Multiple-Antenna Wireless Communications," Ph.D. dissertation, Brigham Young Universty, Provo UT (USA), January 2004.

[37] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short lengths in a lattice, including a complexity analysis," *Mathematics of Computation*, vol. 44, pp. 463–471, April 1985.

[38] V. Erceg et. al. "TGn Channel Models," IEEE 802.11 document 11-03/0940r4, May 2004.

[39] J. Hansen and M. Nold, "Analytic calculation of the power delay profile for single room wireless LAN environments", *IEEE Global Telecommunications Conference (GLOBECOM)* 2000, vol. 1, pp. 98-102.

[40] P. Luethi, C. Studer, S. Duetsch, E. Zgraggen, H. Kaeslin,N. Felber, W. Fichtner, "Gram-Schmidt-based QR decomposition for MIMO detection: VLSI implementation and comparison", *IEEE Asia Pacific Conference on Circuits and Systems*, pp. 830-833, 2008.

[41] L. Schumacher "WLAN MIMO Channel Matlab program," Accessed on 03, May 2011. http://www.info.fundp.ac.be/~lsc/Research/IEEE_80211_HTSG_CMSC/distribution_terms.html,

[42] R. Annavajjala, P. C. Cosman, and L. B. Milstein, "Performance analysis of linear modulation schemes with generalized diversity combining on Rayleigh fading channels with noisy channel estimates," *IEEE Transactions on Information Theory*, Vol. 53, No. 12, December 2007.

[43] Taesang Yoo, "Sum capacity, Scheduling, and Multi-user Diversity in MIMO Broadcast Systems", Ph.D dissertation, Stanford University, September 2007.

[44] W. C. Jakes, "*Microwave mobile communications*", IEEE ed. Piscataway, N.J.: IEEE Press, 1994.