AN ABSTRACT OF THE THESIS OF

<u>Samia El Amrani</u> for the degree of <u>Master of Science</u> in

<u>Electrical and Computer Engineering</u> presented on <u>June 8, 2010</u>.

Title:

<u>Computationally Efficient Block Diagonalization for Downlink Multiuser MIMO-OFDM Systems</u>

Abstract approved: _____

Huaping Liu

One of the key challenges in downlink multiuser multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) systems is the mitigation of the multi-access interference when different users share the same subcarriers. In this work, the block diagonalization (BD) algorithm for inter-user interference pre-cancelation is extended to MIMO OFDM systems. However, in the context of OFDM, the complexity of the algorithm grows proportionally to the number of OFDM subcarriers. To address this issue, the inherent frequency correlation between adjacent subcarriers is exploited. Simulations show that interpolation of the BD pre/decoding matrices provides a favorable tradeoff between complexity and performance. Specifically, a 50% reduction in computational load is achieved for less than 5% throughput loss and 1 dB error performance penalty at high SNR.

Computationally Efficient Block Diagonalization for Downlink
Multiuser MIMO-OFDM Systems

by

Samia El Amrani

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented June 8, 2010
Commencement June 2011

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to my advisor Dr. Huaping Liu for his support throughout my years at OSU. I would also like to thank Dr. Kwonhue Choi, who has provided me with invaluable help and ideas, and for his contributions and guidance to this thesis. Many thanks to my committee members Dr. Hamdaoui, Dr. Raich and Dr. Ostroverkhova.

I would like to thank all my friends and colleagues for their support and encouragement. At last, thanks to my family who has always believed in me even though most of them still think I can fix cell phones.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

DEDICATION

To my family.

Chapter 1 – Introduction

1.1   Overview of MU MIMO OFDM

Wireless communications have been one of the main topics of interest in the area of communications. In the last decade, there has been an increasing demand for high data rate wireless access at high quality of service (QoS) due to the wide deployment of cellular telephony and the emergence of wireless data applications. The development of the wireless technology is supported by the high demand for flexible multimedia services integrating both voice and data communications. On the other hand, the advancement in integrated circuit (IC) design technology has enabled the implementation of sophisticated signal processors and complex systems on chip, resulting in small, low cost and power efficient handsets. Unlike wired networks, the wireless link suffers from two main phenomena that can be considered as an impairment to reliable communications. The first aspect is fading; which results in time variation of the channel strength caused by the multipath propagation as well as the signal attenuation due to the path loss and shadowing. The second challenge for wireless systems is the interference between receivers communicating with a single transmitter or the interference between signals from multiple transmitters to one receiver. This inter-user interference can be observed for example in the downlink and uplink of a cellular system. Fading and inter-user interference

are hence two main challenges for the design of wireless communication systems.

A higher link reliability and a larger network capacity are the two main design goals of next generation communication systems. However, the trend is shifting towards the design of spectrally efficient wireless systems. Important improvement of the spectral efficiency can be achieved through the deployment of multiple antennas at both the transmitter and receiver ends. Due to the physical separation between the antennas, multiantenna systems offer an additional degree of freedom - spatial domain, that is unavailable in single antenna configurations. Multiple-input multiple-output (MIMO) systems offer a tremendous advantage, which is the improvement of spectral efficiency and link reliability without any additional bandwidth or power consumption.

MIMO techniques can be divided into two categories, diversity coding and spatial multiplexing. Diversity coding increases the robustness and the reliability of the communication system by transmitting redundant copies of the data stream on different subchannels. Diversity is exploited by combining the independently faded signals at the receiver, thus resulting in an increase in performance. In the case of spatial multiplexing, independent data streams are transmitted on different antenna branches simultaneously in the same frequency band. The system's capacity is increased by using signal processing at the receiver to recover the multiple data streams. These techniques achieve the highest spectral efficiency when the spatial subchannels are independent, which is usually the case in rich scattering environments.

MIMO technology is of particular interest for next generation broadband communications which aim at delivering multimedia services at high data rates. Broadband systems in particular are subject to frequency selective fading due to the destructive interference of the signal with delayed copies of itself resulting from the multipath reflections. When the relative path delays are in the order of one symbol period or more, the signal experiences inter-symbol interference (ISI). Traditionally, channel equalization techniques are used to combat ISI. However, for high data rates communications, with shorter symbol duration, highly complex equalizers are required. Orthogonal frequency division multiplexing (OFDM) is a low complexity modulation technique that deals with ISI by splitting the high rate data stream into a number of lower rate streams. The data streams are transmitted through different orthogonal subcarriers. OFDM is a good alternative to channel equalization since it transforms the frequency selective channel into a set of parallel narrow band flat fading subchannels. The combination of MIMO and OFDM techniques is an efficient way for providing high data rate reliable communications.

In cellular networks and wireless local area networks (WLAN), MIMO systems are often used in configurations where a main base station communicates with several users simultaneously. In multiuser MIMO systems, the entire system bandwidth is used by all users all the time which is a way for achieving high bandwidth efficiency. In such configurations, the main challenge is multiple access interference (MAI). Users need to be able to mitigate the inter-user interference at

minimal cost. The most effective ways to mitigate MAI are done by pre-processing on the transmitted signal, when partial or full channel knowledge is available at the transmitter.

1.2 Organization and contributions of the thesis

As a linear processing algorithm at the transmitter for inter-user interference cancelation, block diagonalization offers a good tradeoff between computational complexity and performance. The basic idea behind block diagonalization is to transmit to each user an interference free signal given channel knowledge at the transmitter. At the base station, precoding matrices are designed for each user in order to precancel the interference of the other users' signals. On the receiver end, simple decoding is performed to retrieve the original signal. This technique does not require coordination between users or complex processing at the mobile receivers. In a MU MIMO OFDM scenario, the channel is transformed into multiple parallel MU MIMO subchannels, offering the possibility of applying the preprocessing on a per subcarrier basis. In addition to the fact that the channel on each subcarrier is flat fading, there is significant correlation between adjacent subcarriers. In this thesis, we take advantage of the existing correlation by using interpolation techniques together with the block diagonalization algorithm to cancel multiple access interference in MU MIMO OFDM systems. For the direct approach, when the precoding and decoding matrices need to be computed for all users on all subcarriers, the computational load increases linearly with the number of subcarriers. This

requires powerful processors at both ends of the link in addition to the transfer of data between the transmitter and users through feedback. We will show that using interpolation in the frequency dimension offers significant computational savings with only a slight performance penalty. We compare the error performance and throughput of two simple interpolation schemes, the piecewise constant interpolation and the linear interpolation. We show that in this application, the simple piecewise interpolation offers a good tradeoff between savings in computational cost and performance.

This thesis is organized as follows:

Chapter 2

In this chapter an overview of multiuser MIMO OFDM systems is given. The main techniques for interference cancelation in a multiuser scenario are presented in Section 2.1. In Section 2.1.2, the combination of MIMO and OFDM techniques is reviewed. In Section 2.2, the OFDM modulation is introduced and the generation of OFDM signals is developed.

Chapter 3

In this chapter we introduce the computationally efficient block diagonalization algorithm for MU MIMO OFDM systems. The system model is presented in Section 3.1. The block diagonalization scheme for interference pre-cancelation is extended to MIMO-OFDM systems and its complexity is analyzed in Section 3.2. In the last section of the chapter, interpolation

algorithms are introduced and their computational complexity is presented.

Chapter 4

In this chapter, we present simulation results of the system throughput and error performance using the interpolation schemes introduced in Chapter 3. The expression of the throughput is given when interpolation schemes are used. In Section 4.2, the error performance of the interpolation schemes is compared to that of the direct approach and the low-pass interpolation scheme.

Chapter 2 – Background

2.1 Multiuser MIMO OFDM Systems

2.1.1 MU MIMO

Point-to-point single user multiantenna communications have been well understood as an effective way to provide reliable communications and to increase per-user rates. In recent years, there has been a great interest in multiuser multiantenna networks, particularly in broadcast and multiple access scenarios. The multiple access channel, or the uplink, applies when multiple transmitters send signals to a signal receiver over the same frequency band. This setting has been extensively investigated and well understood in the literature. The broadcast channel, also referred to as the downlink, describes the case when a single base station communicates with multiple users over the same medium (i.e., at the same time and over the same frequency band). The interest in multiuser communications arises as the need for high quality wireless communications to accommodate an increasing number of users has become a priority. Hence, multiuser diversity is a key technology, for it allows for the efficient use of the available spectrum. In multiuser MIMO systems, the advantage of spatial diversity offered by multiple antennas can be exploited to improve the system capacity since multiple mobile stations are served simultaneously by means of space division multiple access (SDMA). With multiple

antenna elements, multiple access strategies require more complex processing and design but do not require any additional bandwidth, to achieve a higher throughput.

On the downlink of a multiuser spatial multiplexing system, a single base station transmits to multiple mobile stations simultaneously and over the same frequency band. The major impairment in this scenario is inter-user interference. Inter-user interference arises due to the fact that the same frequency is used to transmit data to all the users, making the signal received by the user a combination of its own signal and the signals designated for other users. Typically, mobile stations are not able to communicate with each other making any type of coordination impossible. Therefore, in order to use low complexity receivers, inter-user interference mitigation needs to be integrated at the transmitter end. This condition makes the channel knowledge at the transmitter necessary. When channel state information is available and the transmitter has knowledge of the interference between users, then it can process the signals in order to overcome the inter-user interference. In general, interference cancelation schemes are designed to suppress inter-user interference while optimizing the system performance metrics such as capacity and error rate.

A MIMO broadcast system with single antenna transmitters and single antenna receivers has been well explored in the literature. The optimal strategy for capacity maximization in this case is to transmit to the single user with the best

channel at any time. The multiuser case with multiple antennas at both ends of the system has been a topic of interest for the last two decades but major advances have occurred only recently . Results have revealed that dirty paper coding (DPC) [1] is the capacity-achieving transmit strategy for MIMO broadcast channels. Optimal strategy for maximizing the capacity of the broadcast channel was first studied in [2] for the case of single antenna users and later extended to multi-antenna receivers using game theory [3]. The basic idea behind DPC is precoding the data at the transmitter based on the knowledge of the channel interference. DPC has been proven to be the optimal strategy for sum capacity, and its capacity region was shown to be that of a MIMO broadcast channel [4]. However, even though DPC makes sense from the information theoretical point of view, it is not considered a practical solution. Implementation of DPC requires additional complexity at the transmitter and receiver and finding a practical realization of dirty paper codes has been proven to be a challenge. Practical solutions have been investigated to provide the capacity gain for multiuser MIMO systems. The key challenge is providing high link level signal quality in addition to interference can-celation. In general, the receiver design must be compact and consume low power whereas the base station is able to handle more advanced processing. Precoding is hence an attractive approach resulting in the base station performing interference pre-cancelation and making low complexity receivers viable at the mobile stations.

Linear precoding is an alternative approach for multiuser MIMO transmission which offers a tradeoff between reduction in precoder design complexity for sub-

optimal performance. For single antenna receivers, one approach is the channel inversion technique or zero forcing (ZF) beamforming precoding. In ZF schemes, channel inversion is performed to eliminate the interference. The downside of this approach is deterioration of the signal quality. For the multiantenna case, the minimum mean squared error (MMSE) criterion has been used for transmit-receive optimization under a sum power constraint [5]. Other techniques take advantage of the uplink-downlink duality for both MSE and signal-to-interference-plus-noise ratio (SINR) [6, 7]. Other possible techniques to improve the sum rate include user or antenna selection using suboptimal strategies or iterative methods which present a high computational cost.

Another family of linear precoding is block diagonalization (BD) which is based on zero forcing [8]. Block diagonalization is a non-iterative method which transforms the multiuser downlink into parallel single user MIMO systems, thus eliminating all interference between users without inverting the channel. The sum capacity is maximized using a conventional waterfilling algorithm for power loading.

2.1.2   MU MIMO OFDM

The combination of MIMO and OFDM is a very efficient way to increase the diversity gain and to enhance the system capacity in frequency-selective channels. MIMO OFDM systems can be viewed as parallel MIMO systems at each subcar-

rier, which facilitates the application of MIMO processing on a per subcarrier basis. For multiuser MIMO OFDM systems, interference cancelation precoding schemes can then easily be applied at each subcarrier independently. A large amount of research has been done on the single user MIMO OFDM whereas the multiuser case remains fairly unexplored. Duplicy et al. [9] have extended available algorithms and studied their complexity and performance to minimize the BER. An issue with this approach is the prohibitive computational cost. Many schemes have been developed to reduce the system complexity. For channel estimation, the channel can be estimated for a subset of subcarriers then use an interpolation scheme to obtain the channel for the remaining subcarriers [10]. In the case of single user, a scheme was proposed for feedback savings, a fraction of the precoding matrices at selected subcarriers are obtained at the receiver, then sent to the transmitter where the transmitter is able to reconstruct all the precoding matrices using interpolation [11, 12]. The interpolator parameters were optimized using MSE or mutual information criterion. This method however only applies to unitary matrices and has been proven ineffective in the multiuser case [13]. Karaa et al. [13] solve the joint power allocation problem across all subcarrier using the squared mean squared error (SMSE) minimization to find the optimal precoding and decoding matrices. They also present methods to reduce the computational load by exploiting the existing correlation between closely spaced subcarriers.

In this thesis, we take advantage of the correlation between adjacent subcarriers to develop a computationally efficient block diagonalization scheme for interference

cancelation for the downlink multiuser MIMO OFDM system.

2.2 OFDM Systems

2.2.1 Overview of OFDM

One main challenge toward the realization of robust wireless communication systems is fading caused by multipath propagation. The received signal may contain a line-of-sight (LOS) component plus many delayed replicas of the signal. The delayed copies are received at different times with different phase offsets due to the reflection on the terrain features and surrounding objects. These signals interfere with the direct path, which causes ISI and degrades the network performance. Typically, to mitigate ISI, adaptive equalization is implemented at the receiver. However, for high speed systems there are practical difficulties to perform equalization with low cost, compact hardware. For high data rates, one efficient way of dealing with the effects of multipath is using parallel transmission.

The concept of OFDM dates back to the 1960s, when Chang [14] introduced the classic parallel transmission schemes where the frequency bandwidth is divided into several nonoverlapping subcarriers. A decade later, discrete Fourier transform was first used for OFDM modulation and demodulation processes. The next milestone in the history of OFDM happened in the mid 1990s, when the European Telecommunications Standard Institute (ETSI) digital audio broadcast-

ing (DAB) [15] became the first OFDM-based wireless system. In recent years, OFDM has become the core technology for a number of standards. For wired environments, it is known as Digital MultiTone (DMT) transmission and is implemented in many xDSL (digital subscriber lines) systems. For wireless applications, OFDM is the core technology of a number of standards such as WLAN (wireless local area networks) standards: IEEE 802.11a/g and IEEE 802.11n [16], which incorporates MIMO techniques, and WMAN (wireless metropolitan area networks) standard such as IEEE 802.16.

2.2.2 OFDM Principle

OFDM is a special case of multicarrier transmission. In order to deal with multipath fading, traditional parallel transmission schemes divide a single high data rate channel into several nonoverlapping lower rate subchannels. By doing so, the frequency selective channel is transformed into flat fading at each subchannel. At each subchannel, the symbol period is increased, thus reducing the sensitivity to the delay spread. To eliminate inter-carrier interference and to avoid spectral overlap, a sufficient guard space is used between adjacent subchannels. However, this method leads to an inefficient use of the available spectrum. A more efficient way to use the available bandwidth is to allow the subchannels' spectra to overlap under certain conditions.

The basic idea behind OFDM systems on the other hand is to use orthogonal subcarriers in the frequency domain. By doing so, OFDM is able to provide a high spectral efficiency in addition to interference mitigation resulting from the orthogonality property. The advantage of the orthogonality of the subchannels resides in the fact that all subchannels are linearly independent, therefore all interference from adjacent subchannels is canceled. OFDM offers also the possibility of allocating different power levels for different subcarriers.

In OFDM, the subcarrier pulse used for transmission is chosen to be rectangular. This results in the possibility of performing pulse modulation by a simple inverse discrete Fourier transform (IDFT). The use of IDFT/DFT results in significant complexity reduction [17]. When the number of subcarriers is a power of two, the IDFT/DFT can be easily and efficiently implemented using the inverse fast Fourier transform (IFFT)/fast Fourier transform (FFT).

Another main advantage of OFDM is its ability to cope with ISI through the insertion of a cyclic prefix. Channel distortion causes each OFDM symbol to spread energy into adjacent symbols, causing ISI. A guard interval is introduced to completely eliminate ISI. The guard interval is chosen to be longer than the expected delay such that the multipath from one symbol does not interfere with the next symbol. The effect of the guard interval is to absorb the delayed copies of the signal that causes interference. To prevent inter-carrier interference (ICI) and to maintain the orthogonality of the subcarriers, the guard interval is chosen to be a

cyclic extension of the signal itself; in this case it is referred to as a cyclic prefix. Each symbol is composed of two parts, the first part is a copy of the tail of the signal itself and the second part contains the active symbol as shown in Figure 2.1. The total symbol duration is $T_{total} = T_g + T_s$ where T_g is the guard time. The guard interval length depends on the application but since it reduces the data throughput T_g is usually kept less than $T_s/4$.

Figure 2.1: Guard interval insertion.

As mentioned above, the inherent structure of OFDM presents many advantages for communication systems but also has some drawbacks. Below is a summary of the main advantages and disadvantages of OFDM systems.

Advantages:

- Robustness against fading and interference: OFDM deals with the effects of multipath fading by using a cyclic prefix to absorb the ISI.

- Low complexity implementation: OFDM modulation can be easily implemented using FFT and IFFT blocks. These blocks have low complexity and power consumption.

- High spectral efficiency: Efficient use of the available spectrum by OFDM.

Drawbacks:

- Sensitivity to frequency offset and phase noise: Small errors in carrier frequency estimation might corrupt the orthogonality property of the subcarriers, thus causing ICI. Moreover, ICI can also be caused by phase noise and time varying channels.

- High average power to peak average: The OFDM signal can be viewed as a superposition of sinusoidal signal, as a result, its peak power is much larger than its mean power.

2.2.3   System Configuration

The structure of the transmitter consists of a serial-to-parallel converter that divides the data stream into parallel substreams each transmitted through a different subchannel. Each substream is first modulated, using quadrature amplitude modulation (QAM), then the time domain signal is obtained by applying the IDFT.

This is done by feeding each substream to an IFFT circuit. The next step is the insertion of the cyclic prefix. On the receiver side, the cyclic prefix is removed then the signal is fed to a FFT block to convert it back to the frequency domain. A parallel-to-serial converter is employed to obtain the original transmitted sequence. A more detailed analysis of the OFDM transmitter and receiver is presented in Section 2.2.4.

In OFDM systems, some subcarriers are used as pilot symbols. Also, for reliable detection and in order to get channel knowledge at the both ends of the communication system, training sequences are occasionally sent by the transmitter. A preamble containing the training sequence is added to the data packet. As most real life systems, the system performance is improved using channel coding, interleaving and transmit/receive filtering. Also, RF modulation/demodulation and RF amplifiers are applied to the baseband signal in order to convert it to the appropriate frequency band. Figure 2.2 shows the block diagram of an OFDM transceiver. In this thesis, the OFDM signals are processed at the baseband level. RF modulation/ demodulation and RF amplification are beyond the scope of this work.

2.2.4   OFDM Signal Generation

This section introduces the OFDM signal generation. In Section 2.2.4.1, a simple SISO OFDM system is considered. The time and frequency domain representations

of the OFDM signals are presented. The OFDM modulated signals are represented in matrix form. The OFDM demodulation process is also introduced. In Section 2.2.4.2, the OFDM modulation/demodulation principles are applied in the more general case of a MIMO system.

2.2.4.1   SISO OFDM

This section introduces the OFDM modulation and demodulation for a single antenna transmitter and single antenna receiver. Let's consider a SISO system with N_c subcarriers and suppose that N_c is a power of two. We make this assumption in order to simplify the application of the Fourier transform to the subcarrier symbols. Let S_n be a sequence of discrete time QAM-modulated data symbols. The data symbols are divided into blocks of length N_c; each block represents one OFDM symbol. The OFDM signal in the time domain is expressed as

$$s'(t) = \sum_{b=-\infty}^{+\infty} \sum_{n=0}^{N_c-1} S_n e^{j2\pi f_n(t-bT_s)} u(t - bT_s) \qquad (2.1)$$

where the function $u(t)$ is the pulse waveform for each symbol

$$u(t) = \begin{cases} 1 & 0 \leq t \leq T_s \\ 0 & \text{otherwise} \end{cases} \qquad (2.2)$$

T_s is the OFDM symbol duration and $f_n = f_c + \frac{n}{T_s}$ for $n = 0, \ldots, N_c - 1$. f_c is the carrier frequency and f_n denotes the frequency of the nth subcarrier. The spacing

between two subcarriers is $\Delta f = \frac{1}{N_c T_0} = \frac{f_0}{N_c}$, with T_0 being the sampling time and $f_0 = \frac{1}{T_0}$ the sampling rate.

For each block b, the N_c symbols are distributed over the N_c subcarriers. For a single OFDM symbol, the baseband equivalent signal is expressed as

$$s(t) = \sum_{n=0}^{N_c-1} S_n e^{j2\pi \frac{n}{T_s} t}. \tag{2.3}$$

The complex baseband OFDM signal is in fact the IDFT of the N_c symbols. By sampling the continuous time signal $s(t)$ given by (2.3) at a rate $f_0 = \frac{N_c}{T_s}$, the discrete time equivalent of the signal is expressed for each time sample m as

$$s(m) = \sum_{n=0}^{N_c-1} S_n e^{\frac{j2\pi nm}{N_c}}, \qquad 0 \le m \le N_c - 1. \tag{2.4}$$

OFDM modulation offers the possibility of representing the OFDM signals in matrix form; they can thus be manipulated using matrix algebra. OFDM matrix representation is introduced and will be used throughout this thesis.

At the transmitter, the first step for OFDM modulation is feeding the data symbols to an IDFT circuit. This operation is equivalent to multiplying the data vector by a Fourier matrix. Let $\mathbf{S}(b) = [S_0(b), \dots, S_{N_c-1}(b)]^T$ be the input data vector for one OFDM symbol. The time domain representation of the bth OFDM symbol vector $\mathbf{s}$ is written as

$$\mathbf{s}(b) = \mathcal{F}^H \mathbf{S}(b) \tag{2.5}$$

where $\mathcal{F}$ is the $N_c \times N_c$ Fourier matrix. The matrix $\mathcal{F}^H$ represents the inverse Fourier transform and is expressed as

$$\mathcal{F}^H = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & W^1 & W^2 & \ldots & W^{(N_c-1)} \\ 1 & W^2 & W^4 & \ldots & W^{2(N_c-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W^{(N_c-1)} & W^{2(N_c-1)} & \ldots & W^{(N_c-1)^2} \end{bmatrix} \qquad (2.6)$$

where $W = e^{j\frac{2\pi}{N_c}}$.

The next step in the OFDM signal generation is the insertion of the cyclic prefix. For a cyclic prefix of duration T_g, let $N_g = T_g f_0$ be the number of samples within the time interval T_g. The insertion of the cyclic prefix is performed by appending the last N_g symbols of $\mathbf{s}$ to the beginning of the signal. This is equivalent to multiplying the vector $\mathbf{s}$ by the cyclic prefix insertion matrix $\boldsymbol{\Theta}$

$$\boldsymbol{\Theta} = \left[\begin{bmatrix} \mathbf{0}_{N_g \times (N_c - N_g)} & \mathbf{I}_{N_g} \end{bmatrix}^T \mathbf{I}_{N_c} \right]^T \qquad (2.7)$$

$$\bar{\mathbf{s}}(b) = \boldsymbol{\Theta} \cdot \mathbf{s}(b) \qquad (2.8)$$

where $\mathbf{0}$ is an $N_g \times (N_c - N_g)$ all zero matrix. The OFDM symbol with cyclic prefix

$\bar{\mathbf{s}}$ is an $(N_g + N_c) \times 1$ vector and can be expressed as

$$\bar{\mathbf{s}}(b) = \begin{bmatrix} \bar{s}_0(b) \\ \vdots \\ \bar{s}_{N_c+N_g-1}(b) \end{bmatrix} = \begin{bmatrix} s_{N_c-N_g}(b) \\ \vdots \\ s_{N_c-1}(b) \\ s_0(b) \\ \vdots \\ s_{N_c-1}(b) \end{bmatrix}. \tag{2.9}$$

The OFDM modulated symbol vector $\bar{\mathbf{s}}$ is transmitted through a frequency selective channel that is considered to be constant during one OFDM symbol. The bth received signal at the receiver is an $(N_c + N_g) \times 1$ vector that can be written as the product of the channel matrix and the symbol vector whose elements depend on the bth OFDM block as well as the preceding bloc $b-1$

$$\bar{\mathbf{r}}(b) = \begin{bmatrix} \bar{r}_0(b) \\ \vdots \\ \bar{r}_{Nc+N_g-1}(b) \end{bmatrix}$$

$$= \begin{bmatrix} h_{L-1} & \dots & h_0 & 0 & & \\ 0 & \ddots & & \ddots & \ddots & \\ & \ddots & \ddots & & \ddots & 0 \\ & & 0 & h_{L-1} & \dots & h_0 \end{bmatrix} \begin{bmatrix} \bar{s}_{N_c+N_g-L+1}(b-1) \\ \vdots \\ \bar{s}_{N_c+N_g-1}(b-1) \\ \bar{s}_0(b) \\ \vdots \\ \bar{s}_{N_c+N_g-1}(b) \end{bmatrix} + \bar{\mathbf{n}}. \tag{2.10}$$

Suppose that the guard interval is chosen to be greater than the excess delay of the channel, so that all interference caused by preceding subcarriers is eliminated. Since the goal of the guard interval is to absorb the interference from preceding symbols, then the first N_g samples of the received signal are contaminated. Therefore, the first step of the OFDM demodulation at the receiver is eliminating the guard interval, i.e., discarding the first N_g samples of each received symbol vector. This operation is equivalent to multiplying the received signal by the matrix $\mathbf{\Gamma}$

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{0}_{N_g \times (N_c - N_g)} & \mathbf{I}_{N_c} \end{bmatrix}. \tag{2.11}$$

The received signal after guard interval elimination can be expressed as the product of the channel matrix and the data symbol vector as shown below

$$
\mathbf{r}(b) = \begin{bmatrix} r_0(b) \\ \vdots \\ r_{Nc-1}(b) \end{bmatrix}
$$

$$
= \begin{bmatrix} h_{L-1} & \cdots & h_0 & 0 & & \\ 0 & \ddots & & \ddots & \ddots & \\ & \ddots & \ddots & & \ddots & 0 \\ & & 0 & h_{L-1} & \cdots & h_0 \end{bmatrix} \begin{bmatrix} \bar{s}_{N_g - L + 1}(b) \\ \vdots \\ \bar{s}_{N_c + N_g - 1}(b) \end{bmatrix} + \mathbf{n}. \tag{2.12}
$$

Due to the cyclic structure of $\bar{\mathbf{s}}$, the signal $\mathbf{r}(b)$ can be written as

$$
\mathbf{r}(b) =
\begin{bmatrix}
h_0 & 0 & \dots & 0 & h_{L-1} & \dots & h_1 \\
h_1 & \ddots & \ddots & & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & & \ddots & h_{L-1} \\
h_{L-1} & \ddots & h_1 & h_0 & 0 & & 0 \\
0 & \ddots & & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & & \ddots & \ddots & 0 \\
0 & \dots & 0 & h_{L-1} & \dots & h_1 & h_0
\end{bmatrix}
\begin{bmatrix}
s_0(b) \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
s_{N_c-1}(b)
\end{bmatrix}
+
\begin{bmatrix}
n_0 \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
n_{N_c-1}
\end{bmatrix} . \quad (2.13)
$$

The channel matrix is circulant, thus diagonal in the Fourier domain. Eq. (2.13) can be written in compact matrix form as

$$
\mathbf{r}(b) = \mathbf{H}\mathbf{s}(b) + \mathbf{n}(b) \tag{2.14}
$$

where $\mathbf{n} = [n_0, \dots, n_{N_c-1}]^T$ is the additive white Gaussian noise vector and $\mathbf{H}$ the channel matrix.

The second part of the demodulation process is the calculation of the Fourier transform of the received signal, which is equivalent to applying the Fourier matrix

$\mathcal{F}$ to the received signal vector $\mathbf{r}$

$$
\begin{aligned}
\mathbf{R}(b) &= \mathcal{F} \cdot \mathbf{r}(b) \\
&= \mathcal{F} \cdot \mathbf{H} \cdot \mathcal{F}^H \cdot \mathbf{S}(b) + \mathcal{F} \cdot \mathbf{n} \\
&= \begin{bmatrix} H_0 & 0 & \dots & & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & & 0 \\ 0 & \dots & 0 & & H_{N_c-1} \end{bmatrix} \cdot \mathbf{S}(b) + \mathbf{N}
\end{aligned}
\tag{2.15}
$$

with H_n being the coefficient of the channel's frequency response at the nth subcarrier expressed as

$$
H_n = \sum_{l=0}^{L-1} h_l \cdot \exp\left(\frac{-j2\pi ln}{N_c} \right).
\tag{2.16}
$$

OFDM modulation assumes the channel as flat fading on each subcarrier, even though the channel is frequency selective. This property simplifies the equalization of the signal at the receiver after the OFDM demodulation. The received signal at the nth subcarrier is then

$$
R_n = H_n S_n + N_n
\tag{2.17}
$$

where N_n represents the noise after application of the DFT on the nth subcarrier. We can see from Eq. (2.17) that the demodulated received signal on each subcarrier is not affected by ISI or ICI.

2.2.4.2 MIMO OFDM

For a MIMO OFDM system, OFDM modulation is applied at each spatial sub-channel. The $N_R \times N_T$ MIMO channel is equivalent to $N_R \times N_T$ uncorrelated SISO subchannels. We suppose that the equivalent SISO subchannels have a length less or equal to L and are constant during one OFDM symbol duration. Let h_l^{pq} be the lth coefficient of the impulse response of the channel between transmit antenna p and receive antenna q. Let S_n^p be the data symbol at the nth subcarrier transmitted from the pth antenna and n^q the noise at the input of the OFDM demodulator at the qth receive antenna. The symbol obtained after OFDM demodulation on the nth subcarrier for receive antenna q is denoted by R_n^q. If the guard interval is not less than the length of the channel L, the received OFDM symbol at the qth antenna after guard interval removal is

$$\mathbf{r}^q(b) = \sum_{p=1}^{N_T} \mathbf{H}^{qp} \mathcal{F}^H \mathbf{S}^p(b) + \mathbf{n}^q(b) \qquad (2.18)$$

where

$$\mathbf{H}^{qp} = \begin{bmatrix} h_0^{qp} & 0 & \cdots & 0 & h_{L-1}^{qp} & \cdots & h_1 \\ h_1^{qp} & \ddots & \ddots & & & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & \ddots & h_{L-1}^{qp} \\ h_{L-1}^{qp} & \cdots & h_1^{qp} & h_0^{qp} & 0 & \cdots & 0 \\ 0 & \ddots & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_{L-1}^{qp} & \cdots & h_1^{qp} & h_0^{qp} \end{bmatrix} \qquad (2.19)$$

$$= \mathcal{F}^H \mathrm{diag}\{H_0^{qp}, \ldots, H_{N_c-1}^{qp}\}\mathcal{F}$$

$$\mathbf{r}^q(b) = [r_0^q(b), \ldots, r_{N_c-1}^q(b)]^T \qquad (2.20)$$

$$\mathbf{S}^q(b) = [S_0^q(b), \ldots, S_{N_c-1}^q(b)]^T \qquad (2.21)$$

$$\mathbf{n}^q(b) = [n_0^q(b), \ldots, n_{N_c-1}^q(b)]^T. \qquad (2.22)$$

The channel matrix at each subcarrier is a circulant matrix, and is thus diagonal in the Fourier domain. The signal received at antenna q is expressed in the frequency domain as:

$$\mathbf{R}^q(b) = \sum_{p=1}^{N_T} \mathrm{diag}\{H_1^{qp}, \ldots, H_{N_c-1}^{qp}\}\mathbf{S}^p(b) + \mathbf{N}^q(b) \qquad (2.23)$$

where $\mathbf{R}^q$ and $\mathbf{N}^q$ are the OFDM symbol and noise vectors at the output of the qth receive antenna, respectively. H_n^{qp} is the nth sample of the frequency response of the subchannel between antennas p and q, given by

$$H_n^{qp} = \sum_{l=0}^{L-1} h_l^{qp} \cdot \exp\left(\frac{-j2\pi ln}{N_c}\right).$$

(2.24)

In the remainder of this thesis, for simplicity of notation especially for multiuser MIMO systems, we will primarily present the system equations in the frequency domain unless otherwise stated. Therefore a MIMO OFDM system with N_c sub-carriers can be treated as a set of N_c parallel MIMO systems. For each subcarrier n, we have a simple $N_T \times N_R$ MIMO system.
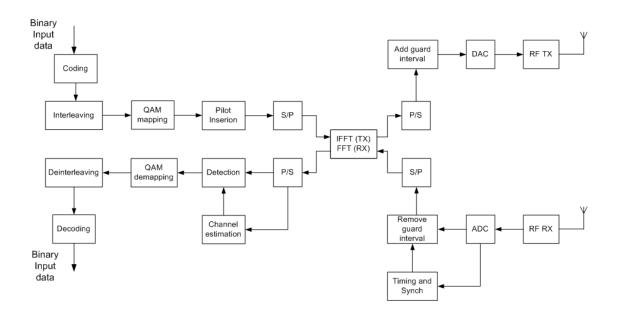
Figure 2.2: OFDM transceiver.

Chapter 3 – Computationally Efficient Block Diagonalization for Interference Cancelation

3.1 System Model

Consider a multiuser MIMO OFDM system, as illustrated in Figure 3.1, with N_T transmit antennas and K mobile users with a total of N_R receive antennas . The kth user is equipped with N_{Rk} antennas such that $\sum_{k=1}^{K} N_{Rk} = N_R$. The channel is modeled as a time-invariant, L-tap frequency-selective Rayleigh fading channel further distorted by additive white Gaussian noise (AWGN). OFDM modulation with N_c subcarriers is used at both the base station and the mobile users to counteract the frequency selectiveness of the channel and transform the channel into N_c parallel independent multiuser MIMO subchannels. In the frequency domain, the channel can be represented by a $N_T \times N_R \times N_c$ composite channel matrix $\mathbf{H}$ containing all channel coefficients for all users on all subcarriers. The entries of $\mathbf{H}$ on each subcarrier are independent and identically distributed zero-mean complex Gaussian variables with unit variance. Let $\mathbf{H}_k(n)$ be an $N_T \times N_{Rk}$ matrix, representing the channel for user k on subcarrier n. We assume that the channels $\{\mathbf{H}_k(n)\}_{k=1}^{K}$ are independent. Therefore, the matrix $\mathbf{H}(n) = [\mathbf{H}_1^T(n), \mathbf{H}_2^T(n), \cdots, \mathbf{H}_K^T(n)]^T$, which characterizes the channel for subcarrier n, has full rank.

Figure 3.1: Multiuser MIMO OFDM system.

The base station broadcasts to all K users simultaneously over all OFDM subcarriers. Each user k receives from the base station m_k data streams on every subcarrier with $m_k \leq N_{Rk}$, resulting in a total of $N_c m_k$ data streams per user. Hence, we have a transmitter equipped with N_T transmit antennas transmitting a total of $N_c m = N_c \sum_{k=1}^{K} m_k$ data streams on all N_c parallel subcarriers, to K mobile users equipped with a total of N_R receive antennas.

The data vector on subcarrier n, $\mathbf{x}_k(n)$, is an $m_k \times 1$ vector containing the data symbols for user k; the overall data vector for all users is $\mathbf{x}(n) = [\mathbf{x}_1^T(n), \mathbf{x}_2^T(n), \ldots, \mathbf{x}_K^T(n)]^T$. At the transmitter, each user's data stream is processed by the $N_T \times m_k$ precod-

ing matrix $\mathbf{M}_k(n)$, where the overall precoding matrix on subcarrier n is denoted by $\mathbf{M}(n) = [\mathbf{M}_1(n), \mathbf{M}_2(n), \cdots, \mathbf{M}_K(n)]$. The output of the transmit filter $\mathbf{M}(n)$ on subcarrier n corresponds to the input to the corresponding transmit antennas. However, these N_c symbols are OFDM modulated before transmission. At the receiver end, for each user, OFDM demodulation is applied before the signal is fed to a decoding receive filter. The output is then N_c decoded vectors $\hat{\mathbf{x}}_k(n)$, each of length m_k.

Since the focus is on precoding design, no error correction coding is used and symbol timing errors and frequency offsets are neglected. Also, it is assumed that the transmitter has perfect knowledge of the instantaneous channel state information (CSI), that is, the channel matrices $\{\mathbf{H}_k(n)\}_{k=1}^K$ for all users are obtained at the base station through a zero-delay error-free feedback channel. The output of the OFDM demodulator for the kth mobile user on subcarrier n is the superposition of the K branches' signals distorted by channel fading plus additive white Gaussian noise, and is expressed as

$$
\begin{aligned}
\mathbf{r}_k(n) &= \sum_{i=1}^{K} \mathbf{H}_k(n)\mathbf{M}_i(n)\mathbf{x}_i(n) + \mathbf{n}_k(n) \\
&= \mathbf{H}_k(n)\mathbf{M}_k(n)\mathbf{x}_k(n) + \sum_{i=1,i\neq k}^{K} \mathbf{H}_k(n)\mathbf{M}_i(n)\mathbf{x}_i(n) + \mathbf{n}_k(n) \\
&= \mathbf{H}_k(n)\mathbf{M}_k(n)\mathbf{x}_k(n) + \mathbf{c}_k(n) + \mathbf{n}_k(n).
\end{aligned}
\tag{3.1}
$$

The term $\mathbf{c}_k(n)$ corresponds to the inter-user interference of user k. The $N_{Rk} \times 1$

vector $\mathbf{n}_k(n)$ denotes the AWGN at the kth user's antenna array, which follows an i.i.d zero mean complex Gaussian distribution with variance N_0. With CSI feedback from all K users, the transmitter assigns resources and designs optimal transmit vectors. As discussed in Section 2.1, the major impairment to system performance is the presence of multiple access interference. The availability of channel knowledge at both ends of the link allows the transmitter to design the precoding matrices to precancel the interference before transmission. Block diagonalization is a precoding technique based on the orthogonalization of the signals to cancel the interference followed by waterfilling to maximize the capacity. The block diagonalization algorithm is described in the next section.

3.2   Block Diagonalization for Interference Cancelation

3.2.1   Algorithm Description

The objective of the BD approach is to find the precoding matrices $\{\mathbf{M}_k(n)\}_{k=1}^K$ for each user on all subcarriers such that

$$\mathbf{H}_k(n)\mathbf{M}_j(n) = 0, \qquad \forall n, \forall j \neq k. \tag{3.2}$$

The inter-user interference term $\mathbf{c}_k(n)$ can be expressed as $\mathbf{c}_k(n) = \mathbf{H}_k(n)\tilde{\mathbf{M}}_k(n)\tilde{\mathbf{x}}_k(n)$, where $\tilde{\mathbf{M}}_k(n)$ and $\tilde{\mathbf{x}}_k(n)$ are defined respectively as the modulation matrix and the

transmit vector for all users other than user k

$$\tilde{\mathbf{M}}_k(n) = [\mathbf{M}_1(n) \quad \cdots \quad \mathbf{M}_{k-1}(n) \quad \mathbf{M}_{k+1}(n) \quad \cdots \quad \mathbf{M}_K(n)] \qquad (3.3)$$

$$\tilde{\mathbf{x}}_k(n) = [\mathbf{x}_1(n) \quad \cdots \quad \mathbf{x}_{k-1}(n) \quad \mathbf{x}_{k+1}(n) \quad \cdots \quad \mathbf{x}_K(n)]. \qquad (3.4)$$

On each subcarrier n, the channel and modulation matrices $\mathbf{H}_S(n)$ and $\mathbf{M}_S(n)$ are defined as

$$\mathbf{H}_S(n) = [\mathbf{H}_1^T(n) \quad \mathbf{H}_2^T(n) \quad \cdots \quad \mathbf{H}_K^T(n)]^T \qquad (3.5)$$

$$\mathbf{M}_S(n) = [\mathbf{M}_1(n) \quad \mathbf{M}_2(n) \quad \cdots \quad \mathbf{M}_K(n)]. \qquad (3.6)$$

User k is free of inter-user interference if $\mathbf{H}_k(n)\mathbf{M}_j(n) = 0$ for $j \neq k$, which makes the product $\mathbf{H}_S(n)\mathbf{M}_S(n)$ block diagonal. This also translates into $\mathbf{M}_k(n)$ lying within the null space of the matrix $\tilde{\mathbf{H}}_k$ defined as

$$\tilde{\mathbf{H}}_k(n) = [\mathbf{H}_1^T(n) \quad \cdots \quad \mathbf{H}_{k-1}^T(n) \quad \mathbf{H}_{k+1}^T(n) \quad \cdots \quad \mathbf{H}_K^T(n)]^T. \qquad (3.7)$$

The zero-interference constraint is re-expressed as

$$\tilde{\mathbf{H}}_k(n)\mathbf{M}_k(n) = 0, \qquad \forall k = 1, \cdots, K. \qquad (3.8)$$

This constraint imposes a dimension condition necessary to accommodate all users. The condition guarantees that data is transmitted to each user k, which means that the dimension of the null space of $\tilde{\mathbf{H}}_k(n)$ is non-zero, i.e., $\mathrm{rank}(\tilde{\mathbf{H}}_k(n)) < N_T$. The dimension condition for all users applies to the rank of $\tilde{\mathbf{H}}_k(n)$ and can be

expressed as $\max\{\text{rank}(\tilde{\mathbf{H}}_1(n)), \cdots, \text{rank}(\tilde{\mathbf{H}}_K(n))\} < N_T$. When the dimension condition is satisfied, the algorithm starts by finding a basis of the null space of $\tilde{\mathbf{H}}_k(n)$ through the computation of its singular value decomposition (SVD), for each user k:

$$\tilde{\mathbf{H}}_k(n) = \tilde{\mathbf{U}}_k(n)\tilde{\mathbf{\Sigma}}_k(n)[\tilde{\mathbf{V}}_k^{(1)}(n) \quad \tilde{\mathbf{V}}_k^{(0)}(n)]^H. \tag{3.9}$$

Let $\tilde{L}_k(n) = \text{rank}(\tilde{\mathbf{H}}_k(n)) \leq N_R - N_{Rk}$. The matrix $\tilde{\mathbf{V}}_k^{(1)}(n)$ contains the first $\tilde{L}_k(n)$ right singular vectors and $\tilde{\mathbf{V}}_k^{(0)}(n)$ holds the last $N_T - \tilde{L}_k(n)$ right singular vectors and constitutes an orthonormal basis for the null space of $\tilde{\mathbf{H}}_k(n)$, i.e., $\mathbf{H}_k(n)\tilde{\mathbf{V}}_j^{(0)}(n) = 0$ for $j \neq k$. The modulation matrix can be written as a a linear combination of the vectors of $\tilde{\mathbf{V}}_k^{(0)}(n)$.

$$\tilde{\mathbf{M}}_k(n) = \tilde{\mathbf{V}}_k^{(0)}(n)\mathbf{A}_k(n) \tag{3.10}$$

where $\mathbf{A}_k(n)$ is the $\tilde{L}_k(n) \times m_k$ transmit beamformer for the equivalent single-user MIMO system. To maximize the sum capacity, $\mathbf{A}_k(n)$ can be found through the SVD of the projection of the channel of the kth user on the null space of $\tilde{\mathbf{H}}_k(n)$, resulting in the product $\mathbf{H}_k(n)\tilde{\mathbf{V}}_k^{(0)}(n)$. The SVD of the product is expressed as

$$\mathbf{H}_k(n)\tilde{\mathbf{V}}_k^{(0)}(n) = \mathbf{U}_k(n) \begin{bmatrix} \mathbf{\Sigma}_k(n) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{V}_k^{(1)}(n) \quad \mathbf{V}_k^{(0)}(n)]^H \tag{3.11}$$

where $\mathbf{\Sigma}_k(n)$ is an $\bar{L}_k(n) \times \bar{L}_k(n)$ matrix of singular values of $\mathbf{H}_k(n)\tilde{\mathbf{V}}_k^{(0)}(n)$, with $\bar{L}_k(n) = \text{rank}\{\mathbf{H}_k(n)\tilde{\mathbf{V}}_k^{(0)}(n)\}$. $\mathbf{V}_k^{(1)}(n)$ is then the matrix that holds the first

$\bar{L}_k(n)$ right singular vectors of $\mathbf{H}_k(n)\tilde{\mathbf{V}}_k^{(0)}(n)$. The product $\tilde{\mathbf{V}}_k^{(0)}(n)\mathbf{V}_k^{(1)}(n)$ builds an orthonormal basis of dimension $\bar{L}_k(n)$ and represents the transmission vectors that maximize the rate for user k on subcarrier n. The precoding matrix $\mathbf{A}_k(n)$ can hence be written as

$$\mathbf{A}_k(n) = \mathbf{V}_k^{(1)}(n)\mathbf{\Lambda}^{1/2}(n) \tag{3.12}$$

where $\mathbf{\Lambda}^{1/2}(n)$ is the power loading matrix on subcarrier n given by the waterfilling algorithm applied to the diagonal elements of the matrix $\mathbf{\Sigma}(n)$, which can be expressed as

$$\mathbf{\Sigma}(n) = \begin{bmatrix} \mathbf{\Sigma}_1(n) & & \\ & \ddots & \\ & & \mathbf{\Sigma}_K(n) \end{bmatrix}. \tag{3.13}$$

Waterfilling algorithm is explained in Appendix A.

For each user k, the precoding matrix is then $\mathbf{M}_k(n) = \tilde{\mathbf{V}}_k^{(0)}(n)\mathbf{V}_k^{(1)}(n)\mathbf{\Lambda}^{1/2}(n)$. The modulation matrix on each subcarrier becomes

$$\mathbf{M}_S(n) = \begin{bmatrix} \tilde{\mathbf{V}}_1^{(0)}(n)\mathbf{V}_1^{(1)}(n) & \tilde{\mathbf{V}}_2^{(0)}(n)\mathbf{V}_2^{(1)}(n) & \cdots & \tilde{\mathbf{V}}_K^{(0)}(n)\mathbf{V}_K^{(1)}(n) \end{bmatrix} \mathbf{\Lambda}^{1/2}(n) \tag{3.14}$$

On each subcarrier n, the waterfilling algorithm is applied to the diagonal elements of $\mathbf{\Sigma}(n)$, the matrix of singular values of $\mathbf{H}_k(n)\tilde{\mathbf{V}}_k^{(0)}(n)$ for all users. Waterfilling is used for capacity maximization under a total power constraint P.

At the receiver, post-processing is performed by applying a decoding matrix

$\mathbf{D}(n)$ to the received signal. The output at each subcarrier is an $m_k \times N_R$ vector of estimated data symbols

$$\hat{\mathbf{x}}(n) = \mathbf{D}(n)(\mathbf{H}_S(n)\mathbf{M}_S(n)\mathbf{x}(n) + \mathbf{n}(n)). \qquad (3.15)$$

The overall decoding matrix on subcarrier n, $\mathbf{D}(n)$, is block diagonal and can be written as

$$\mathbf{D}(n) = \begin{bmatrix} \mathbf{D}_1(n) & & \\ & \ddots & \\ & & \mathbf{D}_K(n) \end{bmatrix}. \qquad (3.16)$$

The decoding matrix for user k on subcarrier n, $\mathbf{D}_k(n)$, is

$$\mathbf{D}_k(n) = \mathbf{U}_k^H(n). \qquad (3.17)$$

To implement the block diagonalization algorithm for sum capacity maximization through waterfilling, the knowledge of $\mathbf{U}_k^H(n)$ is required at each receiver. For each user k, the decoding matrix $\mathbf{U}_k^H(n)$ depends not only on the user's channel matrix $H_{k,}(n)$ but also on the nulling matrix $\tilde{\mathbf{V}}_k^{(0)}(n)$. The calculation of the nulling matrix requires knowledge of all users' CSI. Since we assume that no coordination is possible between users, the decoding matrices cannot be calculated directly. One way is for each receiver to calculate the decoding matrices from an estimate of its

effective channel [18]. The effective channel for user k on subcarrier n is

$$\mathbf{H}_{\mathrm{eff},k}(n) = \mathbf{H}_k(n)\tilde{\mathbf{V}}_k^{(0)}(n) \tag{3.18}$$

which is the equivalent single user MIMO channel for user k on the nth OFDM subcarrier and the equivalent transmit preprocessing matrix is $\mathbf{A}_k(n)$. Another technique to solve this problem is by using coordination information from the transmitter. The base station has more computational resources, so the receiver postprocessing matrices are calculated at the transmitter and the quantized beamforming information is sent to the receivers, through limited feedforward [19].

The total achievable capacity of the system resulting from the block diagonalization on each subcarrier is expressed as

$$\begin{aligned} C_{BD}(n) &= \max_{\mathbf{M}_S(n),\mathbf{H}_k(n)\mathbf{M}_j(n)=0,j\neq k} \log_2 \left| \mathbf{I} + \frac{1}{N_0}\mathbf{H}_S(n)\mathbf{M}_S(n)\mathbf{M}_S^H(n)\mathbf{H}_S^H(n) \right| \\ &= \max_{\mathbf{H}_k(n)\mathbf{M}_j(n)=0,j\neq k} \sum_{k=1}^{K} \log_2 \left| \mathbf{I} + \frac{1}{N_0}\mathbf{H}_k(n)\mathbf{M}_k(n)\mathbf{M}_k^H(n)\mathbf{H}_k^H(n) \right|. \end{aligned} \tag{3.19}$$

The waterfilling algorithm maximizes the system's overall capacity. With $\mathbf{M}_S(n)$ chosen as in Eq. (3.14), the capacity of the BD method becomes

$$C_{BD}(n) = \max_{\mathbf{\Lambda}(n)} \log_2 \left| \mathbf{I} + \frac{\mathbf{\Sigma}^2(n)\mathbf{\Lambda}(n)}{N_0} \right|. \tag{3.20}$$

The capacity is given under total power constraint P, such that $P = \sum_{n=0}^{Nc-1} P_n$,

where P_n is the power allocated to subcarrier n. For uniform power loading on all subcarriers for $n = 0, \cdots, N_c - 1$, the allocated power is $P_n = \frac{P}{N_c}$. In [13] it was shown that optimal power allocation across subcarriers provides only slight performance gain. Since the main focus of our work is complexity reduction, we consider an equal power allocation across all subcarriers.

The block diagonalization algorithm applied is summarized in Table 3.1.

1: For $n = 0, \ldots, N_c - 1$, for $k = 1, \ldots, K$. Compute SVD of $\tilde{\mathbf{H}}_k$

$$\tilde{\mathbf{H}}_k(n) = \tilde{\mathbf{U}}_k(n)\tilde{\boldsymbol{\Sigma}}_k(n)[\tilde{\mathbf{V}}_k^{(1)}(n) \quad \tilde{\mathbf{V}}_k^{(0)}(n)]^H$$

2: Compute the SVD of the projection of $\mathbf{H}_k(n)$ on the right null space of $\tilde{\mathbf{H}}_k(n)$

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

3: Use waterfilling on the diagonal elements of $\boldsymbol{\Sigma}(n) = \text{diag}(\boldsymbol{\Sigma}_1(n), \ldots, \boldsymbol{\Sigma}_K(n))$ to find the optimal power loading matrix $\Lambda(n)$.

4: Set

$$\mathbf{M}_S(n) = [\tilde{\mathbf{V}}_1^{(0)}(n)\mathbf{V}_1^{(1)}(n) \quad \tilde{\mathbf{V}}_2^{(0)}(n)\mathbf{V}_2^{(1)}(n) \quad \cdots \quad \tilde{\mathbf{V}}_K^{(0)}(n)\mathbf{V}_K^{(1)}(n)]\boldsymbol{\Lambda}^{1/2}(n)$$

Table 3.1: Block diagonalization algorithm.

3.2.2   Complexity Analysis

The primary focus of this work is to reduce the computational load of the block diagonalization algorithm in OFDM systems. In this section, we quantify the complexity of the algorithm in terms of floating-point operations (flop). A flop

is defined as one addition, substraction, multiplication or division of two real floating-point numbers. One real addition or multiplication counts as one flop. A complex addition and multiplication have two and six flops, respectively [20]. The complexity of the algorithm can be estimated as a polynomial of the problem dimension. Although flop counting does not provide an accurate prediction of the computational complexity of the algorithm, it is a useful measure to capture the computational load. An approximation of the order of number of flop counts of the SVD algorithm for a complex valued $p \times q$ matrix with $p \geq q$ is shown in Table 3.2

Required	Flop count
Σ	$24pq^2 - 8q^3$
Σ, V	$24pq^2 + 48q^3$
Σ, U, V	$24p^2q + 48pq^2 + 54q^3$

Table 3.2: Complexity of complex valued SVD.

The number of flop counts for each step of the algorithm computed for each of the following operations:

1. SVD of the $(N_R - N_{Rk}) \times N_T$ matrix $\tilde{\mathbf{H}}_k$

2. Product $\mathbf{H}_k(n)\tilde{\mathbf{V}}_k^{(0)}(n)$ and SVD of $\mathbf{H}_k(n)\tilde{\mathbf{V}}_k^{(0)}(n)$

3. Waterfilling for $(\sum_k \bar{L}_k(n))$ real eigenmodes

4. Product of $\tilde{\mathbf{V}}_k^{(0)}(n)\mathbf{V}_k^{(1)}(n)$ and product by power loading matrix

The table below shows an order of the number of flop counts for every step of Algorithm 3.1 for a single subcarrier for user k. For block diagonalization to com-

Step	Flop counts
1	$24N_T^2(N_R - N_{Rk}) + 48N_T(N_R - N_{Rk})^2 + 54(N_R - N_{Rk})^3$
2	$6N_T N_{Rk}(N_T - \tilde{L}_k(n)) + 24(N_T - \tilde{L}_k(n))N_{Rk}^2 + 48N_{Rk}^3$
3	$2(\sum_k \bar{L}_k(n))^2 + 6(\sum_k \bar{L}_k(n))$
4	$12N_T\bar{L}_k(n)(N_T - \tilde{L}_k(n)) + 12N_T(\sum_k \bar{L}_k(n))^2$

Table 3.3: BD algorithm complexity

pletely cancel the interference, the system must satisfy the dimension condition, $N_T > \text{rank}(\tilde{\mathbf{H}}_k(n))$. To get an idea of the maximum flop counts needed for the block diagonalization algorithm, we assume that all channel matrices $\{\mathbf{H}_k(n)\}_{k=1}^K$ are independent, and for each user the matrix $\tilde{\mathbf{H}}_k(n)$ is full rank, i.e., $\tilde{L}_k(n) = N_R - N_{Rk}$. The product matrix $\mathbf{H}_k(n)\tilde{\mathbf{V}}_k^{(0)}(n)$ is of also full rank, $\bar{L}_k(n) = N_{Rk}$. The number

of flop counts in this case is an order of

$$
\begin{aligned}
c \ = \ & 24N_T^2(N_R - N_{Rk}) + 6N_T N_{Rk}(N_T - \tilde{L}_k(n)) + 12N_T \bar{L}_k(n)(N_T - \tilde{L}_k(n)) \\
& + 48N_T(N_R - N_{Rk})^2 + 24(N_T - \tilde{L}_k(n))N_{Rk}^2 + 12N_T \left(\sum_k \bar{L}_k(n) \right)^2 + 54(N_R - N_{Rk})^3 \\
& + 48N_{Rk}^3 + 2 \left(\sum_k \bar{L}_k(n) \right)^2 + 6 \left(\sum_k \bar{L}_k(n) \right) \\
= \ & 24N_T^2(N_R - N_{Rk}) + 6N_T N_{Rk}(N_T - N_R + N_{Rk}) + 12N_T \bar{L}_k(n)(N_T - N_R + N_{Rk}) \\
& + 48N_T(N_R - N_{Rk})^2 + 24(N_T - N_R + N_{Rk})N_{Rk}^2 + 12N_T N_R^2 + 54(N_R - N_{Rk})^3 \\
& + 48N_{Rk}^3 + 2N_R^2 + 6N_R.
\end{aligned} \tag{3.21}
$$

The complexity of the block diagonalization algorithm is then $O(N_T^2(N_R - N_{Rk}))$ on each subcarrier. The overall complexity when block diagonalization is applied to the OFDM system is $O(N_c N_T^2(N_R - N_{Rk}))$.

3.3   Interpolation for Complexity Reduction

In order to reduce the computational load of the interference pre-cancelation at the transmitter, we user the block diagonalization algorithm combined with inter-polation of the processing matrices. Various interpolation methods have been used for OFDM systems particularly for channel estimation purposes. Colieri et al. [10] have used interpolation for pilot based channel estimation. The performance of lin-ear, second-order, low-pass, spline and time-domain interpolation algorithms was compared for comb-type based channel estimation. Interpolation was also used

for complexity reduction for the design of linear precoding matrix based on SMSE minimization using uplink-downlink duality [13]. The authors in [11] and [12] propose a scheme to limit feedback requirements for a MIMO OFDM system using interpolation. A fraction of the precoding matrices for chosen subcarriers are obtained at the receiver, quantized and fed back to the transmitter. The complete set of matrices is then recovered using interpolation while assuming the precoding matrices are unitary. The interpolator parameters are optimized using MSE or mutual information criterion. However, the proposed method only applies to unitary matrices.

In this work, block diagonalization algorithm is applied for all users on a per-subcarrier basis for precancelation of the inter-user interference. As discussed in Section 3.2.2, the complexity of the algorithm for each user is $O(N_c N_T^2 (N_R - N_{Rk}))$, which grows linearly with the number of OFDM subcarriers N_c. With a fixed number of users and antenna dimension, the computational load can be efficiently reduced without changing the number of subcarriers by taking advantage of the correlation between subcarriers and applying interpolation. The computational load and feedback requirements can be reduced by adopting interpolation techniques along the frequency dimension. Instead of computing the precoding matrices at all subcarriers, the matrices $\mathbf{M}_S(n)$ are calculated for a subset of N_p subcarriers. Matrices for the subcarriers not in the subset are calculated by interpolating the precoding matrices corresponding to the adjacent subcarriers in the subset. For $n = 0 : L_i : Nc - 1$, where $L_i = \frac{N_c}{N_p}$ is the interpolation factor, the matrices

are obtained by performing the direct calculation of the BD matrices as shown in the algorithm in Table 3.1. Each element of all directly calculated matrices are interpolated along the frequency dimension for all subcarriers, as summarized in Table 3.4. It is important to have an interpolation algorithm that provides a good tradeoff between simplicity of implementation and estimation errors. In this work,

---

1: For user k, for $n = 0 : L_i : N_c - 1$
    Compute precoding matrix $\mathbf{M}_k(n)$ and decoding matrix $\mathbf{D}_k(n)$
2: For all entries of $\mathbf{M}_k(n)$, $\mathbf{D}_k(n)$
    Interpolate the samples obtained in step 1 along the frequency domain.

Table 3.4: Interpolation for block diagonalization

we focus on the two simplest interpolation algorithms, the piecewise constant interpolation and the linear interpolation. These two simple techniques were chosen among the many mathematical interpolation algorithms for their simplicity and low complexity implementation. The interpolation algorithms are described below.

3.3.1 Piecewise Constant Interpolation

This is the simplest of the interpolation methods available. It attributes to each intermediate point the value of the nearest point in the direct calculation subset. In other words, the interpolated precoding matrix $\mathbf{M}_S(n)$ and decoding matrix $\mathbf{D}(n)$ for subcarrier n, are equal to $\mathbf{M}_S(mL_i)$ and $\mathbf{D}(mL_i)$, where $m = 0 : N_p - 1$ and mL_i is the nearest subcarrier in the direct calculation subset.

Piecewise constant interpolation is therefore very simple to implement and for each

subcarrier, only the values of the nearest neighbor subcarrier is required and does not take into account any other points. This method does not require any additional computations as it assigns previously calculated matrices to intermediate subcarriers.

3.3.2   Linear Interpolation

Linear interpolation is a simple interpolation algorithm that uses first degree polynomials to approximate the intermediate points. Let $\mathbf{M}_S$ be the precoding matrix to be interpolated and $m = 0 : N_p-1$. For subcarrier n, with $mL_i < n < (m+1)L_i$, using linear interpolation $\mathbf{M}_S(n)$ is given by

$$
\begin{aligned}
\mathbf{M}_S(n) &= \mathbf{M}_S(mL_i + l) \quad 0 \leq l \leq L_i \\
&= \frac{1}{L_i} \left(\mathbf{M}_S(m+1) - \mathbf{M}_S(m) \right) + \mathbf{M}_S(m).
\end{aligned}
\tag{3.22}
$$

The procedure above is applied to all intermediate subcarriers for all values of the precoding matrix $\mathbf{M}_S$ and decoding matrix $\mathbf{D}$. For N_c total subcarriers, if N_p are calculated using the direct method, interpolation is applied $N_c - N_p$ times. Let $M \times N$ be the dimension of the matrix to be interpolated, the total complexity of the linear interpolation is then $O\left(2(N_c + N_p)MN\right)$. The complexity of the interpolation of the $N_T \times m$ precoding matrix M_S for all subcarriers is $O(2(N_c + N_p)N_Tm)$. For the decoding matrix D, which is block diagonal as shown in Eq. (3.16), the interpolation is applied to the $m_k \times N_{Rk}$ decoding matrices for each user k. The off-diagonal zero matrices do not need to be interpolated, making

the number of required operations even lower.

Chapter 4 – Simulation Results

4.1   Data Throughput

In the block diagonalization process with waterfilling for power allocation, the precoding matrices are designed for capacity maximization. On each subcarrier, the achievable throughput is maximized as shown in Eq. (3.19). In general, for multiuser systems with K mobile users, the throughput on subcarrier n is expressed as

$$C_{BD}(n) = \max_{\mathbf{M}_k(n)} \log_2 \left| \mathbf{I} + \frac{\mathbf{H}_k(n)\mathbf{M}_k(n)\mathbf{M}_k^H(n)\mathbf{H}_k^H(n)}{N_0\mathbf{I} + \sum_{j=1,j\neq k}^{K} \mathbf{H}_k(n)\mathbf{M}_j(n)\mathbf{M}_j^H(n)\mathbf{H}_k^H(n)} \right|. \quad (4.1)$$

The term $\sum_{j=1,j\neq k}^{K} \mathbf{H}_k(n)\mathbf{M}_j(n)\mathbf{M}_j^H(n)\mathbf{H}_k^H(n)$ represents the covariance of the effective noise (i.e., noise plus interference) at receiver k. The precoding matrices $\mathbf{M}_k(n)$ are computed by the block diagonalization algorithm for throughput maximization such that the interference term is null. In the study of the system throughput, we make the following assumptions: (1) each channel use corresponds to at least one OFDM symbol, (2) the channel stays constant during each channel use, and (3) uniform power loading is applied across all subcarriers. For an OFDM system with block diagonalization, the overall achievable throughput is the sum of the rates on all subcarriers [21].

$$C_{BD} = \max_{\mathbf{M}_k(n)} \sum_{n=0}^{N_c-1} \sum_{k=1}^{K} \log_2 \left| \mathbf{I} + \frac{\mathbf{H}_k(n)\mathbf{M}_k(n)\mathbf{M}_k^H(n)\mathbf{H}_k^H(n)}{N_0\mathbf{I} + \sum_{j=1,j\neq k}^{K} \mathbf{H}_k(n)\mathbf{M}_j(n)\mathbf{M}_j^H(n)\mathbf{H}_k^H(n)} \right| \quad (4.2)$$

We propose the use of interpolation techniques to estimate the precoding matrices on a subset of the OFDM subcarriers. For the subset of subcarriers where interpolation is performed, the precoding and power loading matrices are not optimal due to the interpolation errors. This results in a decrease of the system's total throughput that can be attributed to the effect of the interference from other users. With that in mind, we can write the average throughput per subcarrier as

$$\hat{C}_{BD} = \frac{1}{N_c} \sum_{n=0}^{N_c-1} \sum_{k=1}^{K} \log_2 \left| \mathbf{I} + \frac{\mathbf{H}_k(n)\mathbf{M}_k(n)\mathbf{M}_k^H(n)\mathbf{H}_k^H(n)}{N_0\mathbf{I} + \sum_{j=1,j\neq k}^{K} \mathbf{H}_k(n)\mathbf{M}_j(n)\mathbf{M}_j^H(n)\mathbf{H}_k^H(n)} \right| \quad (4.3)$$

Eq. (4.3) represents the average throughput per subcarrier in bits/s/Hz for one channel realization. In our simulations, the throughput in Eq. (4.3) is averaged over R channel realizations. To get the overall data throughput in bits/s, the average throughput is scaled by the system's bandwidth. We ignore the system losses due to the the introduction of the cyclic prefix as well as the pilot symbols for channel estimation and synchronization.

For the subcarriers for which the precoding and decoding matrices are calculated using direct calculation, the interference term is equal to zero. For the subset of interpolated subcarriers, the inter-user interference is non negligible and is caused by the interpolation errors. When using interpolation, we expect a decrease of the system's total throughput as well as the average throughput per subcarrier.

As mentioned earlier, a substantial reduction in complexity is expected, but that comes at the expense of lower performance. In this section, we will investigate the throughput penalty due to the use of interpolation. We will simulate the system's throughput for different interpolation factors and compare the throughput of our algorithm and the direct calculation.

In the simulations, we consider a downlink multiuser MIMO OFDM system with $N_T = 4$ transmit antennas and $K = 2$ users, each is equipped with $N_{Rk} = 2$ receive antennas. The configuration considered is for OFDM systems with a transmission bandwidth of 1.25 MHz and $N_c = 128$ OFDM subcarriers unless mentioned otherwise. The symbol time is $T = 66.67\mu s$ and the subcarrier spacing is

$$\Delta f = \frac{1.25 \times 10^6}{128} = 15 \text{ kHz} \tag{4.4}$$

The system parameters are summarized in Table 4.1

Bandwidth	1.25 MHz
FFT size	128
Symbol duration	$66.7\mu s$
Subcarrier spacing	15 kHz
Cyclic prefix	1/4
Modulation	16QAM

Table 4.1: Frequency selective channel parameters

The multipath fading channel is modeled by an L_t-tap impulse response, with

an exponential power delay profile and a root mean square (rms) delay spread τ_{rms}. Specifically, we only consider the paths with a maximum delay of $5\tau/T$ [22]. The coherence bandwidth of the system is defined by the rms delay spread. The coherence bandwidth for a frequency correlation greater than 0.9 can be approximated by [23]

$$B_c \approx \frac{1}{50\tau_{\mathrm{rms}}}. \tag{4.5}$$

The number of correlated subcarriers with a correlation factor greater than 0.9 is defined as

$$N_L = \frac{B_c}{\Delta f}. \tag{4.6}$$

We have then N_L adjacent subcarriers that fade coherently. For a typical value of rms delay spread $\tau_{\mathrm{rms}} = 100ns$, the number of highly correlated subcarriers is approximately $N_L = 13$. The value of N_L helps determine the range of appropriate interpolation factors that would minimize the throughput loss. Theoretically, for the above parameters, we can use an interpolation factor up to 13 and only observe a slight reduction in system throughput.

As mentioned above, there are $N_L = 13$ subcarriers with a correlation coefficient of at least 0.9. To take full advantage of the correlation, we choose an interpolation interval size smaller than 13, such as $L_i = 2, 4, 8$. By allowing fewer correlated subcarriers to be interpolated and choosing an interpolation factor larger than N_L, $L_i = 16$ for example, we can allow more performance degradation and further reduce the complexity.

To demonstrate the tradeoff between the complexity reduction and performance of our proposed algorithm, Figure 4.1 shows the data throughput of the system for the piecewise interpolation and linear interpolation algorithms for interpolation factors $L_i = 4, 8, 16$.
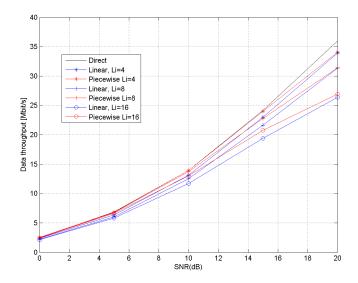

Figure 4.1: Data throughput of the direct approach, piecewise and linear interpolation versus SNR.

The simulation results show that at low to moderate SNR, the throughput gap between the direct calculation and piecewise interpolation is less than 0.5 dB. For very low SNR, the throughput for the piecewise interpolation is equal to the direct approach independent of the interpolation interval length. In the high SNR region, the throughput loss is more significant and is dependent on the interpolation factor. The higher the interpolation factor, the higher the throughput penalty. From

Figure 4.1, we can see that the piecewise constant interpolation leads to a better performance than the linear interpolation, with a throughput gain of up to 2 dB for high SNR.

Using either interpolation algorithm results in substantial computational load reduction as shown in Table 4.2.

Figure 4.2 and Figure 4.3 show the data throughput versus the number of operations for the direct calculation, piecewise and linear interpolation for interpolation factors $L_i = 2, 4, 8, 16$ for SNR=10 dB and SNR=20 dB, respectively. We can see that piecewise interpolation has a low throughput penalty with high complexity reduction, as it does not require any additional operations.

L_i	Piecewise	Linear
2	50%	48%
4	75%	72%
8	87.5%	84%
16	93.7%	90%

Table 4.2: Complexity savings as a percentage of the direct calculation approach for interpolation factors $L_i = 2, 4, 8, 16$.

The piecewise constant interpolation induces less throughput reduction for higher complexity savings than linear interpolation. We can see from Table 4.3, the throughput penalty is approximately 0.5% for SNR=10 dB and 5% when SNR=20 dB for interpolation factor of 4 with a complexity reduction of 75%.
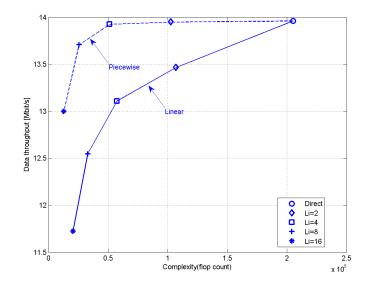
Figure 4.2: Data throughput of the direct approach, piecewise and linear interpolation for SNR=10 dB.

(a) Capacity penalty for SNR=10dB

L_i	Piecewise	Linear
2	0.15%	3.62%
4	0.48%	6.32%
8	1.66%	9.98%
16	5.79%	15.04%

(b) Capacity penalty for SNR=20dB

L_i	Piecewise	Linear
2	2.03%	3.42%
4	5.03%	5.61%
8	12.02%	12.19%
16	24.29%	25.92%

Table 4.3: Capacity penalty of the piecewise and linear interpolation for interpolation factors $L_i = 2, 4, 8, 16$

The above simulations are obtained by averaging the data throughput of the system over 1000 channel realizations. The results presented are useful in predict-
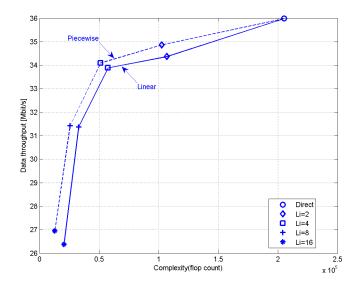
Figure 4.3: Data throughput of the direct approach, piecewise and linear interpolation for SNR=20 dB.

ing the performance of the the piecewise constant and linear interpolation schemes. The simple piecewise interpolation provides an advantage over the linear algorithm in terms of throughput, that is more significant at low to moderate SNR regime, which is a reasonable operating region for wideband communications systems when channel coding is applied. The piecewise method also requires less computational complexity and hence it provides a good tradeoff between throughput maximization and complexity reduction.

4.2   Error Performance

In this section, the error performance of the interpolation schemes is investigated. The same system configuration as in Section 4.1 is used here. The average BER performance illustrates the results of Monte Carlo simulations of 2000 QAM symbols per user per subcarrier per channel realization over 1000 realizations. A uniform power allocation is applied across the subcarriers. The error performance of the interpolation schemes for different interpolation factors is compared to that of the direct calculation method. In Section 4.1 it has been observed that the simple piecewise interpolation outperforms the linear interpolation in terms of system data throughput. Since the piecewise method also offers lower complexity, it will be considered as a method of choice for throughput maximization. Figure 4.4 shows the average BER of the direct calculation approach and the piecewise constant interpolation for interpolation factors $L_i = 2, 4, 8, 16$.

For an interpolation factor of 2, the piecewise interpolation results in a penalty of less than 0.5 dB at a BER of 10^{-3}. At the same BER, the power loss of $L_i = 4$ is approximately 1.8 dB. For interpolation factor 8 and 16, the BER penalty is 4 dB and 7 dB, respectively. Interpolation factors of 2 and 4 result in reasonable BER penalty for complexity reduction of 50% and 75%, respectively. Linear interpolation performance is compared to piecewise interpolation in Figure 4.5.

The linear interpolation scheme has a penalty loss of approximately 6 dB for
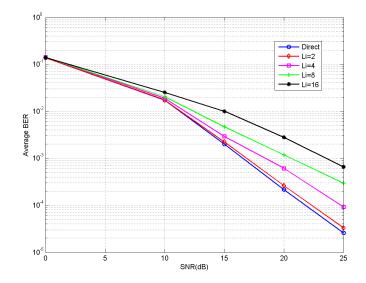
Figure 4.4: Comparison of BER performance of direct approach and piecewise interpolation for $L_i = 2, 4, 8, 16$.

interpolation factor 2. This is about 12 times higher than the piecewise constant interpolation performance penalty. As much as it offers a reasonable tradeoff between throughput loss and complexity savings, the linear interpolation results in poorer error performance.

The performance of the piecewise interpolation is compared to the low-pass interpolation scheme used in [13] and [10]. Low-pass interpolation is performed by applying a $2L_iL + 1$ length FIR filter after inserting zeros in the original sequence. The filter allows the original data to be unchanged and interpolates between the original values such as the mean-squared error of the interpolated points and the ideal values is minimized. Figure 4.6 shows the performance of the low-pass in-
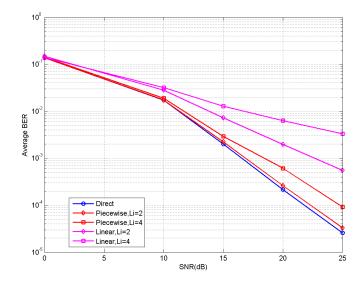
Figure 4.5: Comparison of BER performance of direct approach, piecewise and linear interpolation schemes for $L_i = 2, 4$.

terpolation scheme compared with the direct approach. For a BER of 10^{-3}, the performance penalty for the low pass-interpolation is approximately 8 dB independently of the interpolation factor. Even though it results in poor error performance, the advantage of the low-pass interpolation is its robustness to the interpolation interval length. Figure 4.7 shows the error performance of the direct calculation, the piecewise, linear and low-pass interpolation for interpolation factors $L_i = 4, 8$.

From Figure 4.7, the piecewise constant interpolation outperforms other interpolation schemes and results in lower performance penalty for $L_i \leq N_L$.

Among the interpolation schemes investigated, the simple piecewise constant
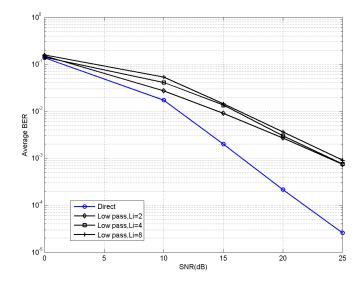
Figure 4.6: Comparison of BER performance of direct approach and low-pass interpolation for $L_i = 2, 4, 8$.

interpolation provides the best throughput and error performance. Since piecewise interpolation does not require any additional complexity and it has less than 0.5 dB performance loss, it provides a good tradeoff between complexity reduction and error performance.
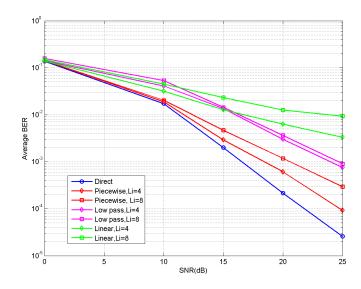
Figure 4.7: Comparison of BER performance of direct approach, piecewise, linear and low-pass interpolation for $L_i = 4, 8$.

Chapter 5 – Conclusion

One of the main challenges for the multiuser MIMO downlink is the inter-user interference mitigation. The block diagonalization scheme is a linear precoding algorithm for interference cancelation at the transmitter. With channel knowledge at the base station, precoding is applied to the transmitted signal so that each user receives an interference-free signal. The waterfilling algorithm is used for power allocation optimization across the MIMO subchannels. For downlink MU MIMO-OFDM systems, the block diagonalization can be extended to the parallel OFDM subcarriers. Practical MIMO-OFDM systems usually have a large number of OFDM subcarriers, which makes the computation of the block diagonalization precoding and decoding matrices computationally prohibitive. By exploiting the inherent frequency correlation between adjacent subcarriers, interpolation of the pre/decoding matrices and power allocation can applied without much loss in performance. By using low complexity algorithm, important power savings can be achieved at both the base station and the mobile users in addition to the reduction of the amount of information exchanged through feedback.

In this thesis, the computational complexity of the block diagonalization algorithm is computed and interpolation schemes are investigated. The throughput loss and error performance of the interpolation algorithms are compared to that of

the direct approach. Since the focus is on the computationally efficient solutions for block diagonalization, two simple interpolation schemes are investigated: the piecewise constant interpolation and the linear interpolation. Simulations show that the piecewise constant interpolation outperforms the linear interpolation in terms of throughput and error performance. For a 50% complexity saving, the piecewise interpolation throughput loss is 5% at high SNR and the error performance penalty is 0.5 dB at a BER of 10^{-3}. The simple piecewise interpolation appears to provide a favorable tradeoff between computational load reduction and performance. This tradeoff can be further improved by using an optimal interpolation scheme.

Bibliography

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

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

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

[4] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936 –3964, sept. 2006.

[5] A. Tenenbaum and R. Adve, "Joint multiuser transmit-receive optimization using linear processing," *IEEE International Conference on Communications*, vol. 1, pp. 588 – 592, june 2004.

[6] M. Schubert, S. Shi, E. Jorswieck, and H. Boche, "Downlink sum-MSE transceiver optimization for linear multi-user MIMO systems," *Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers*, pp. 1424 – 1428, October 2005.

[7] A. Khachan, A. Tenenbaum, and R. Adve, "Linear processing for the downlink in multiuser MIMO systems with multiple data streams," *IEEE International Conference on Communications*, vol. 9, pp. 4113 –4118, june 2006.

[8] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461 – 471, feb. 2004.

[9] J. Duplicy, J. Louveaux, and L. Vandendorpe, "Optimization of linear predecoders for multi-user closed-loop MIMO-OFDM," *IEEE International Conference on Communications*, vol. 3, pp. 2026 – 2030 Vol. 3, may 2005.

[10] S. Colieri, M. Ergen, A. Puri, and B. A, "A study of channel estimation in OFDM systems," *IEEE 56th Vehicular Technology Conference*, vol. 2, pp. 894 – 898 vol.2, 2002.

[11] J. Choi and J. Heath, R.W., "Interpolation based unitary precoding for spatial multiplexing MIMO-OFDM with limited feedback," *IEEE Global Telecommunications Conference*, vol. 1, pp. 214 – 218 Vol.1, nov.-3 dec. 2004.

[12] J. Choi, B. Mondal, and R. Heath, "Interpolation based unitary precoding for spatial multiplexing MIMO-OFDM with limited feedback," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4730 –4740, dec. 2006.

[13] H. Karaa, R. Adve, and A. Tenenbaum, "Linear precoding for multiuser MIMO-OFDM systems," *IEEE International Conference on Communications*, pp. 2797 –2802, june 2007.

[14] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell Systems Technical Journal*, vol. 45, pp. 1775–1796, December 1966.

[15] E. T. S. Institute, "Radio broadcasting systems; digital audio broadcasting (DAB) to mobile, portable and fixed receivers," *ETSI ETS*, vol. 300 401, Feb 1995.

[16] IEEE Std 802.11n-2009, "Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications - Amendment 5: Enhancements for Higher Throughput," pp. c1 –502, oct. 2009.

[17] L. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Transactions on Communications, [legacy, pre - 1988]*, vol. 33, no. 7, pp. 665–675, 1985.

[18] L.-U. Choi and R. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 20 – 24, jan. 2004.

[19] C.-B. Chae, D. Mazzarese, T. Inoue, and R. Heath, "Coordinated beamforming for the multiuser MIMO broadcast channel with limited feedforward," *IEEE Transactions on Signal Processing*, vol. 56, no. 12, pp. 6044 –6056, dec. 2008.

[20] G. H. Golub and C. F. Van Loan, *Matrix Computations (Johns Hopkins Studies in Mathematical Sciences)(3rd Edition)*, 3rd ed. The Johns Hopkins University Press, October 1996.

[21] P. Chan and R. Cheng, "Reduced-complexity power allocation in zeroforcing MIMO-OFDM downlink system with multiuser diversity," pp. 2320 –2324, 4-9 2005.

[22] K.-K. Wong, R.-K. Cheng, K. Letaief, and R. Murch, "Adaptive antennas at the mobile and base stations in an OFDM/TDMA system," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 195 –206, jan 2001.

[23] T. Rappaport, *Wireless Communications: Principles and Practice.* Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.

APPENDICES

Appendix A – Waterfilling Algorithm

Waterfilling is an iterative procedure for power allocation. It is considered an optimal strategy for transmit power adaptation when channel state information is available at the transmitter. When the channel can be transformed into parallel independent subchannels, the principle of waterfilling is to allocate more power to better subchannels, with a high signal-to-noise ratio, in order to maximize the sum rate of all subchannels. The weaker subchannels are assigned lower power or no power at all. The inputs of the waterfilling algorithm are the channel eigenvalues and the total power. Subchannels with higher eigenvalues are allocated a high corresponding power, subject to the sum of the power at all subchannels is less than or equal to the total power.

We apply the waterfilling to $\boldsymbol{\Sigma}(n)$ shown in Eq. (3.13) on all subcarriers $n = 0, \ldots, N_c - 1$, under a total power constraint P. The waterfilling procedure outputs a diagonal matrix whose elements scale the power transmitted into each vector of the modulation matrix that maximizes the throughput $C_{BD}(n)$ on each subcarrier with

$$C_{BD}(n) = \max_{\boldsymbol{\Lambda}(n)} \log_2 \left| \mathbf{I} + \frac{\boldsymbol{\Sigma}^2(n)\boldsymbol{\Lambda}(n)}{N_0} \right|. \tag{A.1}$$

The matrix of singular values $\mathbf{\Sigma}(n)$ on subcarrier n can be written as

$$
\mathbf{\Sigma}(n) = \begin{bmatrix} \mathbf{\Sigma}_1(n) & & \\ & \ddots & \\ & & \mathbf{\Sigma}_K(n) \end{bmatrix}
$$

$$
= \begin{bmatrix} \sigma_1(n) & & \\ & \ddots & \\ & & \sigma_{L_n}(n) \end{bmatrix}
$$

$$
= \begin{bmatrix} \sigma_{1,1}(n) & & & & & & \\ & \sigma_{1,2}(n) & & & & & \\ & & \ddots & & & & \\ & & & \sigma_{1,\bar{L}_1}(n) & & & \\ & & & & \ddots & & \\ & & & & & \sigma_{K,1}(n) & \\ & & & & & & \ddots & \\ & & & & & & & \sigma_{K,\bar{L}_K}(n) \end{bmatrix} \tag{A.2}
$$

where $\sigma_{k,i}(n)$ denotes the ith singular value of the kth user on the nth subcarrier and $L(n) = \sum_{k=1}^{K} \bar{L}_k(n)$. $\sigma_{k,i}(n)$ represents the channel gains of the kth user's ith eigenmode on subcarrier n.

For subcarrier n, let

$$\mathbf{\Lambda}(n) = \begin{bmatrix} \lambda_1(n) & & \\ & \ddots & \\ & & \lambda_L(n) \end{bmatrix} \tag{A.3}$$

where the $\lambda_l(n)$'s are the power loading coefficients subject to:

$$\sum_{l=1}^{L} \lambda_l(n) = P_n, \qquad \lambda_l(n) > 0, \qquad l = 1 \quad , \cdots , \quad L \tag{A.4}$$

where P_n is the power allocated to the nth subcarrier.

The BD capacity at each subcarrier can then be written as

$$C_{BD}(n) = \max_{\lambda_1(n)...\lambda_L(n)} \sum_{l=1}^{L} \log_2 \left(1 + \frac{\lambda_l(n)\sigma_l^2(n)}{N_0} \right). \tag{A.5}$$

To find the optimal power loading matrix, the Lagrangian method is used

$$\mathcal{L}(u, \lambda_1(n), \ldots, \lambda_L(n)) = \sum_{l=1}^{L} \log_2 \left(1 + \frac{\lambda_l(n)\sigma_l^2(n)}{N_0} \right) - u \sum_{l=1}^{L} \lambda_l(n) \tag{A.6}$$

where u is the Lagrange multiplier. We then apply the Kuhn-Tucker condition to optimize power allocation

$$\frac{\partial \mathcal{L}}{\partial \lambda_l(n)} \begin{cases} = 0 & \text{if } \lambda_l(n) > 0 \\ \leq 0 & \text{if } \lambda_l(n) = 0 \end{cases} \tag{A.7}$$

Define $x^+ := \max(x, 0)$. The optimal power allocation is then

$$\lambda_l(n) = \left(\frac{1}{u} - \frac{N_0}{\sigma_l^2(n)} \right)^+ \tag{A.8}$$

with the Lagrange multiplier u chosen to be such that the power constraint is met

$$\sum_{l=1}^{L} \left(\frac{1}{u} - \frac{N_0}{\sigma_l^2(n)} \right)^+ = P_n. \tag{A.9}$$