Precoding and Resource Allocation for Multi-User Multi-Antenna Broadband Wireless Systems

by

Ali Khanafer

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of The Edward S. Rogers Sr. Department of
Electrical and Computer Engineering
University of Toronto

Abstract

Precoding and Resource Allocation for Multi-User Multi-Antenna Broadband Wireless Systems

Ali Khanafer

Master of Applied Science

Graduate Department of The Edward S. Rogers Sr. Department of Electrical and Computer Engineering

University of Toronto

2010

This thesis is targeted at precoding methods and resource allocation for the downlink of fixed multi-user multi-antenna broadband wireless systems. We explore different utilizations of precoders in transmission over frequency-selective channels. We first consider the weighted sum-rate (WSR) maximization problem for multi-carrier systems using linear precoding and propose a low complexity algorithm which exhibits near-optimal performance. Moreover, we offer a novel rate allocation method that utilizes the signal-to-noise-ratio (SNR) gap to capacity concept to choose the rates to allocate to each data stream. We then study a single-carrier transmission scheme that overcomes known impairments associated with multi-carrier systems. The proposed scheme utilizes time-reversal space-time block coding (TR-STBC) to orthogonalize the downlink receivers and performs the required pre-equalization using Tomlinson-Harashima precoding (THP). We finally discuss the strengths and weaknesses of the proposed method.

To my Fiancée, Parents, and Siblings.

Acknowledgements

I owe my deepest gratitude to my supervisor Professor Teng Joon Lim. He provided me with knowledge on various communications topics and was always available to answer questions and explain things in detail. I also benefited from his vast experience in research which helped me succeed in my master's. Professor Lim was always very understanding and provided a healthy environment for research. I also would like to thank my co-supervisor Dr. Roya Doostnejad for her comments on my research. I owe deep thanks to Dr. Taiwen Tang for all the time he spent with me on various research problems. My warm thanks go to my master's committee: Professor Raviraj Adve and Professor Wei Yu for their valuable comments.

I am also very grateful to my colleagues and friends for making Toronto feel like home. I would like to especially thank: Hayssam Dahrouj, Masoud Mohamed, Mahmoud Abou-Beih, Ehsan Karamad, Kianoush Hosseini, Yashar Ghiassi-Farrokhfal, Gokul Sridharan, Siddarth Hari, Sachin Kadloor, Nazmul Islam, Ayman Shalaby, Amirhossein Shokouh Aghaei, Adam Tenenbaum, K V Srinivas, Hassan Masoom, Rostom Ohannessian, Elias Ferzli, Ahmad Akl, Amir Hejazi, Zhengwei Jiang, Amer Farroukh, Hagop Koulakezian, Mohammad Shahin Mahanta, and William Chou.

My deepest thanks go to my parents who have provided me with everything I needed to reach this stage in my life. I also owe all my siblings for their support and guidance especially my eldest brother Hassib Khanafer for bringing me to Canada and advising me to pursue graduate studies.

Finally, my heartfelt gratitude to my fiancée Sarah Assaad. This work would not have been possible without her love, care, and support. She paved the way to my success both in my undergraduate and graduate years. Sarah surrounded me with her love and was always there for me in each step I took towards the completion of my thesis. May He grant me the strength to support and honour her throughout our life.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The thriving need for high data rates has shaped current and next generation communications standards. Data rates up to tens of megabits per second are necessary to support applications and services requiring real-time processing. High speed Internet access, video streaming, video conferencing, and voice over IP (VoIP) are few examples of those widely used services. Nonetheless, providing economic and efficient service to users in rural areas remains a challenge faced by many communications solutions. Stretching wires or fibers over long distances represents a major economic burden to service providers. Employing fixed wireless networks is a viable solution to this problem. Fixed broadband wireless access (FBWA) represents a strong technology providing high data rates over a wide range. Different standards for FBWA have been developed such as IEEE 802.16d or fixed WiMAX (Worldwide Interoperability for Microwave Access) [1]. As the name implies, fixed WiMAX networks consist of fixed (non-mobile) transceivers with one such transceiver acting as a central controller or base station (BS). Typical deployments of fixed WiMAX are at university campuses and businesses with remote buildings. Another important example of fixed WiMAX applications is wireless backhaul solutions as in con-

necting the BS to the mobile switch center (MSC) in a cellular network. Fixed WiMAX enjoys many advantages over wired access technologies such as [1]:

- Wider coverage while guaranteeing comparable throughput up to 40 kilometers
- Lower deployment costs
- Deployment in areas where usage of copper cables is not allowed

On the other hand, transmission over wireless links faces many challenges due to the nature of wireless channels. Fading and interference are major impediments for wireless transmission. Wireless communications researchers managed to devise various schemes to mitigate fading and interference to facilitate fast data transmission over frequency-selective channels. The various standards developed enable various modes to enhance the quality of communications. IEEE 802.16 is no exception, as it possesses various mechanisms including [1]:

- Orthogonal frequency division multiple access (OFDMA)
- Adaptive modulation and coding (AMC)
- Multi-antennas and space-time coding support

Of particular interest among those techniques are multi-antenna systems. Extensive research was conducted on multiple-input multiple-output (MIMO) systems over the past decade. The importance of MIMO systems is twofold: providing improved reliability and high transmission rates. Alamouti invented a simple technique to achieve full transmit diversity by using two transmit and one receive antennas [2] – Alamouti's code. From the idea of transmit diversity, arose the concept of space-time block codes (STBC) which are used to obtain high diversity gains while maintaining low complexity receivers. Besides achieving high diversity orders, MIMO systems can be designed to achieve high data rates using spatial multiplexing. The first example of a system employing spatial

multiplexing was proposed by Foschini [3]. The increase in throughput, however, comes at the cost of a loss in diversity. In fact, there is a fundamental tradeoff between diversity and multiplexing [4]. Various standards have adopted MIMO techniques such as IEEE 802.11n (Wi-Fi) and IEEE 802.16e (Mobile WiMAX) [5].

Another important transmission scheme is OFDMA which enables transmission over frequency-selective channels with high spectral efficiency. OFDMA is the multi-user version of orthogonal frequency division multiplexing (OFDM) which divides a frequency-selective channel into parallel narrowband flat-fading channels or subcarriers [6, 7]. Multiple access is achieved in OFDMA by assigning different sets of subcarriers to different users. Besides their ability to mitigate inter-symbol interference (ISI), the use of IFFT and FFT blocks at the transmitter and receiver of OFDMA systems makes them computationally efficient. If users are allowed to share subcarriers in a multi-user network, OFDMA systems can support larger populations of users than single-carrier systems through transmitting over multiple frequencies. AMC may be applied on each subcarrier to meet quality of service (QoS) requirements or to adapt to channel conditions. OFDMA was adopted in different standards including IEEE 802.16e [5].

Nevertheless, OFDMA systems are known to contain drawbacks. Sensitivity to spectral nulls, subcarrier frequency offset, and phase noise are among these disadvantages. In addition, conventional OFDMA transceivers require expensive nonlinear power amplifiers that provide a large power back-off to overcome the high peak-to-average power ratio (PAPR) of multi-carrier signals. A possible alternative to OFDMA is using single-carrier transmission with equalization to combat multipath fading. A prominent advantage of single-carrier transmission is that the signals have low PAPR. Thus, nonlinear amplifiers are not required and the complexity of transceivers is reduced significantly. Single-carrier systems are also less sensitive to carrier frequency offset and spectral nulls. However, those advantages come at a price; the equalization required in single-carrier systems is a rather complex operation especially in severe multipath fading environments. Clearly,

there exist performance-cost trade-offs between OFDMA and single-carrier systems.

1.2 Thesis Overview: Precoding Over Multi-User ISI Channels

For broadcast channels (BC) in FBWA, where a BS communicates with multiple users, OFDMA or single-carrier modulation may be combined with MIMO systems to combat inter-symbol interference and to provide high data rates as well as improved reliability. The absence of mobility in fixed wireless systems allows channels to be modeled as time-invariant over a long period. Channel estimation at the receivers and high-quality feedback of these estimates to the BS is therefore possible. This enables the implementation of various transmitter precoding techniques. The aim of this thesis is to develop practical precoding methods for both multi and single-carrier systems over frequency-selective broadcast channels with two different objectives. For multi-carrier systems, our objective is to design precoders that maximize the information theoretic rate. We also intend to design a stream rate allocation algorithm to translate theoretical data rates achieved by the precoders into practical AMC modes for each user, for a given desired bit error probability P_b. In the case of single-carrier transmission, we explore the performance of precoders when used to tackle the signal processing problem of ISI suppression and interference cancellation resulting in a simplified receiver design.

In MIMO flat-fading BC, linear and nonlinear precoding techniques can be used to increase the throughput to each user when channel state information (CSI) is available at the BS. The capacity-achieving technique for MIMO-BC was shown to be dirty-paper coding (DPC) in [8]. Practical realizations of DPC such as Tomlinson-Harashima precoding (THP) [9, 10] involve nonlinear operations at both the transmitter and receiver which has driven researchers to explore linear precoding techniques (or beamforming), seeking lower complexity approaches for maximizing throughput [11, 12].

The weighted sum rate (WSR) maximization problem is of particular interest as the weights can be used to prioritize users based on QoS requirements. Maximizing WSR subject to a total power constraint is known to be a non-convex problem. To date, no convex representation has been found for the WSR problem; therefore, different algorithms have been reported in the literature [13, 14, 15] with the objective of converging to a local optimum of the WSR function for flat-fading MIMO-BC. A recent algorithm by Christensen *et al.* [16, 17] uses the relationship between mutual information and minimum mean squared error (MMSE). The authors show that the WSR problem can be solved as a weighted sum MMSE (WMMSE) problem with optimized MSE weights – we will refer to this algorithm by WSRBF-WMMSE. The WSRBF-WMMSE algorithm relies on closed-form expressions to iterate between WMMSE transmit filter design and MMSE receive filter design and is less complex than state-of-the-art techniques which involve nested iterations and solving geometric programs (GPs) [13, 14].

Applying [13, 14, 15, 18] to multi-carrier systems is not practical as they involve solving GPs which incur a worst case polynomial-time complexity in the number of parameters of the optimization problem [13]. Here, we extend WSRBF-WMMSE to MIMO-OFDMA systems and show that the complexity of our approach relative to WSRBF-WMMSE scales linearly with the number of subcarriers. Further, we apply the so-called clustering technique where we group subcarriers in clusters that share the same beamformer, and we study the performance-complexity trade-off offered by clustering.

Another important contribution is that we propose a generally applicable method to translate theoretical MIMO precoder designs into *practical* designs that choose the best AMC mode for each stream of each user, for a given desired P_b. To do this, we assume that a quadrature amplitude modulation (QAM) format is used and that the coding gain of the error control code employed is known. Further, we suggest a reduced complexity method that searches over a reduced number of points in the space of allowed AMC modes. Although we limit our analysis here to beamformers designed by the WSRBF-

WMMSE algorithm due to its simplicity, our technique is general and can be applied to beamformers designed with other WSR maximizing approaches. For MIMO-OFDMA systems, our AMC stream rate allocation method can be applied on each subcarrier to choose the best modulation and coding schemes without requiring expensive computations.

For the single-carrier MIMO-BC, we deal with spatial multiplexing using the time-reversal STBC (TR-STBC) technique: an extension of Alamouti's code that achieves full spatial and multipath diversity over frequency-selective channels. Instead of using TR-STBC for transmission of one users information, we use its orthogonalization feature to broadcast to multiple users. We study the performance of precoders in ISI suppression where we perform the required pre-equalization of each users data at the transmitter through THP, rather than at the individual receivers, resulting in a simple receiver structure. Moreover, we will study the performance-cost trade-off between this single-carrier scheme and the traditional way of combating ISI in broadcast channels by employing OFDMA.

1.3 Contributions

The major contributions of this thesis are the following:

- We extend WSRBF-WMMSE to MIMO-OFDMA systems and show that the complexity of our proposed algorithm increases only linearly in the number of subcarriers relative to its single-carrier counterpart.
- We apply clustering to the proposed algorithm to further reduce its complexity without introducing significant performance deterioration.
- We propose a generally applicable practical stream rate allocation technique using AMC.

- We propose a single-carrier transmission technique that overcomes known OFDMA drawbacks and makes novel use of THP and TR-STBC to suppress ISI and orthogonalize users.

- We study the performance-cost trade-offs between OFDMA-based transmission schemes and our proposed single-carrier system.

1.4 Thesis Organization

The rest of the thesis is organized as follows. In Chapter 2, we provide a literature review of known WSR maximization techniques and explain WSRBF-WMMSE in detail. We also describe the TR-STBC technique and present examples of its applications. We extend WSRBF-WMMSE to multi-carrier systems in Chapter 3 and propose a practical stream rate allocation algorithm. We also study the complexity and performance of the proposed algorithms. In Chapter 4, we introduce our single-carrier transmission scheme and compare its performance and complexity with OFDMA-based systems. We conclude the thesis and provide future directions in Chapter 5.

1.5 Notation

Matrices are denoted by upper case boldface letters, whereas column vectors are represented by lower case boldface letters. The element in the i-th row and j-th column of a matrix $\mathbf{A}$ is denoted by $[\mathbf{A}]_{ij}$. The i-th element of a vector $\mathbf{a}$ is denoted by a_i, whereas the i-th vector of a stack of vectors $\mathbf{a}$ is denoted $\mathbf{a}_i$. The operations $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, and $\otimes$ denote conjugation, transpose, Hermitian transpose, and Kronecker product, respectively. $\mathbf{I}$, $\mathbf{0}$, and $\mathbf{F}_N$ denote the identity matrix, the all-zeros matrix, and $N \times N$ orthonormal DFT matrix with $[\mathbf{F}_N]_{lm} = (1/\sqrt{N})e^{\frac{-j2\pi lm}{N}}$, respectively. A permutation matrix $\mathbf{P}_N^r$ of size $N \times N$ performs a reverse cyclic shift when applied to a $N \times 1$ vector

$\mathbf{a}$ depending on r, i.e., $[\mathbf{P}_N^r \mathbf{a}]_l = a_{((N-l+r) \bmod N)}$. The abbreviation diag{.} returns a diagonal matrix with the vector elements along its main diagonal, blkdiag{$\cdot$} returns a block-diagonal matrix, and Tr($\cdot$) is the trace operator. $Q(\cdot)$ is the tail probability of the standard Gaussian distribution and $E\{\cdot\}$ is the expectation operator.

Chapter 2

Literature Review

In our effort to devise an algorithm to maximize the WSR in MIMO-OFDMA systems using linear processing, we first review previous algorithms designed for WSR maximization in flat-fading MIMO-BC. This problem has received particular attention because the weights assigned to users can be designed to satisfy QoS requirements. Moreover, the non-convexity of the problem has stimulated researchers to propose algorithms that guarantee near-optimal performance without incurring high complexity. We give particular attention to the recently proposed WSRBF-WMMSE algorithm [16, 17] due to its simplicity. This algorithm is the keystone of the algorithms proposed in Chapter 3. We also describe the TR-STBC technique which is an extension of Alamouti's code to frequency-selective channels. TR-STBC will be utilized in the single-carrier system we propose in Chapter 4 in our investigation of precoding as an ISI suppression mechanism. In addition, we provide examples of previous applications of TR-STBC.

2.1 WSR Maximization for MIMO-BC Using Beamforming

2.1.1 Problem Formulation

Consider the downlink of a wireless system where the BS employs M transmit antennas and broadcasts to K users. User k has N_k antennas, and the total number of receive antennas is $N_r = \sum_{k=1}^{K} N_k$. Let $\mathbf{d}_k(n) \in \mathbb{C}^{N_k \times 1}$ be the symbol vector intended for the k-th user at the n-th time slot with $E\{\mathbf{d}_k(n)\mathbf{d}_k^H(n)\} = \mathbf{I}$. Let the transmit beamformer to the k-th user be $\mathbf{B}_k \in \mathbb{C}^{M \times N_k}$. The transmitted signal is thus given by $\mathbf{x}(n) = \sum_{k=1}^{K} \mathbf{B}_k \mathbf{d}_k(n)$. Let $\mathbf{H}_k \in \mathbb{C}^{N_k \times M}$ be the flat-fading channel from the BS to the k-th user where the (i,j)-th element of $\mathbf{H}_k$ is the channel gain from the j-th transmit antenna to the i-th receive antenna. The received signal at the k-th user would be

$$\mathbf{y}_k(n) = \mathbf{H}_k \mathbf{x}(n) + \mathbf{w}_k(n), \tag{2.1}$$

where $\mathbf{w}_k(n) \in \mathbb{C}^{N_k \times 1}$ is a noise vector with independent zero-mean circularly symmetric complex Gaussian (ZMCSCG) elements and covariance matrix $\mathbf{I}$. The channels to all users are assumed to be perfectly known at the BS. The transmit vectors are assumed to respect a total transmit power constraint P_{max}:

$$E\{\mathbf{x}(n)\mathbf{x}^H(n)\} = \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{B}_k \mathbf{B}_k^H) \leq P_{max}. \tag{2.2}$$

The objective is to design the beamformers $\mathbf{B}_k$ to maximize the WSR subject to the constraint (2.2). The WSR problem can thus be formulated as

$$\begin{aligned} [\mathbf{B}_1^*, \mathbf{B}_2^*, ..., \mathbf{B}_K^*] \;=\; & \underset{\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_K}{\arg\min} \sum_{k=1}^{K} -\mu_k \tilde{R}_k \\ & \text{s.t.} \quad \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{B}_k \mathbf{B}_k^H) \leq P_{max}, \end{aligned} \tag{2.3}$$

where $\tilde{R}_k$ is the rate assigned to user k, and μ_k is a pre-determined value that indicates the relative importance of user k. Assuming Gaussian transmission, $\tilde{R}_k$ is given by

$$\tilde{R}_k = \log_2 \det(\mathbf{I} + \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{C}_k^{-1} \mathbf{H}_k \mathbf{B}_k), \tag{2.4}$$

where $\mathbf{C}_k$ is the noise covariance matrix of the interference seen by user k and is given by

$$\mathbf{C}_k = \mathbf{I} + \sum_{i=1, i \neq k}^{K} \mathbf{H}_k \mathbf{B}_i \mathbf{B}_i^H \mathbf{H}_k^H. \tag{2.5}$$

2.1.2 Proposed Solutions

The literature contains many examples of beamforming design for WSR maximization. In [13], the authors utilize the MSE uplink-downlink duality [19] to propose an iterative algorithm for jointly optimizing powers, transmit filters, and receive filters. The uplink-downlink duality states that achievable MSEs in the downlink for a given set of transmit and receive beamformers and a power allocation policy, under a total power constraint, can also be achieved by a power allocation policy in the uplink. The algorithm is based on solving a GP for optimal power control. In each iteration, the algorithm performs the following optimization steps successively: i) uplink power control by solving a GP; ii) uplink MMSE receive filter design; iii) dual transformation from the uplink to the downlink; and iv) downlink MMSE receive filter design. The optimality of the power allocation step enables this algorithm to converge to a local optimum.

Codreanu *et al.* [14] proposed an iterative algorithm to jointly design transmit and receive beamformers based on WSR maximization. The authors use concepts of precoder design via conic optimization [20] and power allocation via signomial programming [21], and provide a four-step optimization algorithm that relies on standard convex optimization tools. The second step of the algorithm involves a nested iterative process that solves a sequence of GPs; those GPs approximate a signomial program which the authors coin

as a re-formulation of the WSR maximization problem. A second-order cone program (SOCP) is then solved in the third step. This algorithm also converges to a local optimum. The algorithms proposed in [13, 14] were referred to as state-of-the-art algorithms in [16, 17].

The special case of equal priorities or best effort service was considered in [15]. The authors show that maximizing the sum-rate is equivalent to minimizing the product of MSE matrix determinants. They further simplify the problem by considering a scalar version of the problem that minimizes the product of MSEs. They present a four-step iterative algorithm that converges to a local optimum of the sum-rate maximization problem. Similar to [13], the MSE uplink-downlink duality is key to this approach. The algorithm consists of the following successive steps: i) downlink precoder design; ii) downlink power allocation; iii) virtual uplink precoder design; and iv) virtual uplink power allocation. Power allocation is solved via sequential quadratic programming (SQP) [22].

A different approach to the WSR maximization problem is taken in [16, 17]. The authors exploit recent results emphasizing the relationship between mutual information and MMSE to establish a relationship between the WSR maximization and the WMMSE minimization problems. They propose an iterative algorithm, which is called WSRBF-WMMSE in [16, 17], that converges to a local optimum. In its three optimization steps, the algorithm iterates between WMMSE transmit filter design, MMSE receive filter design, and MSE weights update. A prime advantage of this algorithm over the above algorithms is its low complexity due to its sole dependence on closed-form expressions without the need to solve GPs.

As mentioned in Chapter 1, our goal is to maximize the WSR in MIMO-OFDMA systems. The large number of subcarriers employed in those systems highlights the need for low complexity optimization algorithms. Due to its low complexity and near optimal performance, WSRBF-WMMSE is an attractive candidate for this objective. We will

present the details of this algorithm in the following section.

WSRBF-WMMSE

It was shown in [23] that $\tilde{R}_k$ can be expressed in terms of the error covariance matrix after MMSE receive filtering. The MMSE receiver of user k is given by

$$\begin{aligned} \mathbf{A}_k^{\text{MMSE}} &= \arg\min_{\mathbf{A}_k} E\{\| \mathbf{A}_k\mathbf{y}_k - \mathbf{d}_k \|^2\} \\ &= \mathbf{B}_k^H\mathbf{H}_k^H(\mathbf{B}_k\mathbf{H}_k\mathbf{B}_k^H\mathbf{H}_k^H + \mathbf{C}_k)^{-1}. \end{aligned} \tag{2.6}$$

When $\mathbf{A}_k^{\text{MMSE}}$ is applied, the MMSE matrix for user k can be written as [23]

$$\begin{aligned} \mathbf{E}_k &= E\{(\mathbf{A}_k^{\text{MMSE}}\mathbf{y}_k - \mathbf{d}_k)(\mathbf{A}_k^{\text{MMSE}}\mathbf{y}_k - \mathbf{d}_k)^H\} \\ &= (\mathbf{I} + \mathbf{B}_k^H\mathbf{H}_k^H\mathbf{C}_k^{-1}\mathbf{H}_k\mathbf{B}_k)^{-1}. \end{aligned} \tag{2.7}$$

From (2.4) and (2.7), we can re-write $\tilde{R}_k$ as

$$\tilde{R}_k = \log_2\det(\mathbf{E}_k^{-1}). \tag{2.8}$$

In [16, 17], the authors establish a relationship between the WSR maximization and the WMMSE minimization problems. The Karush-Kuhn-Tucker (KKT) conditions of both problems are shown to be similar by comparing the gradients of their Lagrangians. This enables solving the WSR maximization problem as a WMMSE minimization one using optimized MSE weights.

The objective of the WMMSE minimization problem is to design transmit filters based on minimizing the WMMSE assuming that MMSE receive filtering is applied. The problem can be formulated as follows:

$$[\mathbf{B}_1^*, \mathbf{B}_2^*, ..., \mathbf{B}_K^*] = \underset{\mathbf{B}_1,\mathbf{B}_2,...,\mathbf{B}_K}{\arg\min} \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_k\mathbf{E}_k) \quad \text{s.t.} \quad \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{B}_k\mathbf{B}_k^H) \leq P_{max}. \tag{2.9}$$

The Lagrangian of problem (2.3) is given by

$$\mathcal{L}(\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_K, \lambda) = \sum_{k=1}^{K} -\mu_k \tilde{R}_k + \lambda \left(\sum_{k=1}^{K} \mathrm{Tr}(\mathbf{B}_k\mathbf{B}_k^H) - P_{max} \right), \tag{2.10}$$

and the Lagrangian of problem (2.9) is

$$\mathcal{G}(\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_K, \lambda) = \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{W}_k\mathbf{E}_k) + \lambda \left(\sum_{k=1}^{K} \mathrm{Tr}(\mathbf{B}_k\mathbf{B}_k^H) - P_{max} \right), \tag{2.11}$$

where λ is the Lagrange multiplier. The derivatives of (2.10) and (2.11) (see Appendix A for details) are given by

$$\begin{aligned} \nabla_{\mathbf{B}_k}\mathcal{L} &= -\mu_k \mathbf{H}_k^H \mathbf{C}_k^{-1} \mathbf{H}_k \mathbf{B}_k \mathbf{E}_k \\ &\quad + \left(\sum_{i=k, i\neq k}^{K} \mu_i \mathbf{H}_i \mathbf{C}_i^{-1} \mathbf{H}_i \mathbf{B}_i \mathbf{E}_i \mathbf{B}_i^H \mathbf{H}_i^H \mathbf{C}_i^{-1} \mathbf{H}_i \right) \mathbf{B}_k + \lambda \mathbf{B}_k, \qquad (2.12) \\ \nabla_{\mathbf{B}_k}\mathcal{G} &= -\mathbf{H}_k^H \mathbf{C}_k^{-1} \mathbf{H}_k \mathbf{B}_k \mathbf{E}_k \mathbf{W}_k \mathbf{E}_k \\ &\quad + \left(\sum_{i=k, i\neq k}^{K} \mathbf{H}_i^H \mathbf{C}_i^{-1} \mathbf{H}_i \mathbf{B}_i \mathbf{E}_i \mathbf{W}_i \mathbf{E}_i \mathbf{B}_i^H \mathbf{H}_i^H \mathbf{C}_i^{-1} \mathbf{H}_i \right) \mathbf{B}_k + \lambda \mathbf{B}_k. \qquad (2.13) \end{aligned}$$

Also, notice that

$$\nabla_\lambda \mathcal{L} = \nabla_\lambda \mathcal{G} = \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{B}_k\mathbf{B}_k^H) - P_{max}. \tag{2.14}$$

By comparing (2.12) and (2.13), and given (2.14), it is evident that the two problems are closely related. In particular, for a given set of beamformers $\{\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_K\}$ and the corresponding set of MMSE matrices $\{\mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_K\}$, the gradients of the two problems can be made identical by choosing the MSE weights to be

$$\mathbf{W}_k = \mu_k \mathbf{E}_k^{-1}, \quad \forall k. \tag{2.15}$$

An optimal point for the WSR maximization problem is achieved when $\nabla_{\mathbf{B}_k}\mathcal{L} = \mathbf{0}$ and $\nabla_{\lambda}\mathcal{L} = 0$ for all k. Let the set of transmit beamformers at this point be $\{\mathbf{B}_1^*, \mathbf{B}_2^*, ..., \mathbf{B}_K^*\}$ and the corresponding set of MMSE matrices be $\{\mathbf{E}_1^*, \mathbf{E}_2^*, ..., \mathbf{E}_K^*\}$. If this set of MMSE matrices is used to compute the MSE weights using (2.15), then the KKT conditions for the MMSE minimization problem are satisfied for the same set of transmit beamformers. Therefore, a WSR maximum is also a WMMSE minimum. This enables solving the WSR problem via the WMMSE problem using optimized MSE weights.

Using this correspondence, the authors proposed an iterative algorithm that alternates between the WMMSE optimization of the beamformers $\mathbf{B}_k$ and the MSE weights $\mathbf{W}_k$ through (2.15). Starting with an initial set of transmit beamformers, the algorithm computes MMSE receive filters, updates MSE weights, and designs a new set of WMMSE transmit filters. The algorithm repeats these steps until convergence. All computations are performed using closed-form expressions and the algorithm contains a single loop only making it less complex than state-of-the-art algorithms [13, 14, 15]. The WSRBF-WMMSE algorithm is outlined in Table 2.1.

Table 2.1: WSRBF-WMMSE.

1: set $t = 0$
2: set $\mathbf{B}_k^t = \mathbf{B}_k^{init}\ \forall k$
3: iterate
4: update $t = t + 1$
5: compute $\mathbf{A}_k^t|\mathbf{B}_i^{t-1}\ \forall i$ for all k using (2.6))
6: compute $\mathbf{W}_k^t|\mathbf{B}_i^{t-1}\ \forall i$ for all k using (2.15)
7: compute $\mathbf{B}^t|\mathbf{A}^t, \mathbf{W}^t$ using (2.17)
8: until convergence

Computing the WMMSE transmit filter $\mathbf{B} = [\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_K]$ in step 7 of WSRBF-

WMMSE is carried out using

$$\bar{\mathbf{B}} = (\mathbf{H}^H\mathbf{A}^H\mathbf{W}\mathbf{A}\mathbf{H} + \frac{\mathrm{Tr}(\mathbf{W}\mathbf{A}\mathbf{A}^H)}{P_{max}}\mathbf{I})^{-1}\mathbf{H}^H\mathbf{A}^H\mathbf{W}; \quad (2.16)$$

$$\mathbf{B} = b\bar{\mathbf{B}}, \quad (2.17)$$

where $b = \sqrt{\frac{P_{max}}{\mathrm{Tr}(\bar{\mathbf{B}}\bar{\mathbf{B}}^H)}}$ is a scaling factor to satisfy the total transmit power constraint, $\mathbf{H} = [\mathbf{H}_1^T, \mathbf{H}_2^T, ..., \mathbf{H}_K^T]^T$, $\mathbf{A} = \mathrm{blkdiag}\{\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_K\}$, and $\mathbf{W} = \mathrm{blkdiag}\{\mathbf{W}_1, \mathbf{W}_2, ..., \mathbf{W}_K\}$. Convergence analysis of this algorithm is shown in Appendix A.

2.2 TR-STBC

In [24], Lindskog and Pulraj proposed the so-called TR-STBC scheme which is a generalization of Alamouti's code to channels with ISI. For a frequency-selective channel of order L, TR-STBC achieves a diversity order of $MN_r(L+1)$ [25, 26], where M is the number of transmit antennas, and N_r is the number of receive antennas. Also, TR-STBC enables independent detection of data streams yielding a simple receiver structure. However, equalization is required after orthogonalizing the received data streams to recover the original transmitted symbols.

2.2.1 System Model

Consider the discrete-time channel model of a point-to-point system employing two transmit antennas and one receive antenna. Let the channel from the j-th transmit antenna to the i-th receiver antenna be

$$h_{ij}(q^{-1}) = h_{ij}[0] + h_{ij}[1]q^{-1} + ... + h_{ij}[L-1]q^{-L+1}, \quad (2.18)$$

where q^{-1} is the unit delay element. Let the symbol sequence to be transmitted be $d(n)$ which is split into two blocks $d_1(n)$ and $d_2(n)$ each containing $J+1$ symbols, $J \geq L$.

The transmission is done over two time slots. During the first time slot, the block $d_1(n)$ is transmitted through antenna 1 whereas $d_2(n)$ is transmitted through antenna 2. The received signal is then given by

$$r_1(n) = h_{11}(q^{-1})d_1(n) + h_{12}(q^{-1})d_2(n) + w_1(n), \tag{2.19}$$

where $w_1(n)$ is a sequence of independent ZMCSCG elements and covariance matrix $\sigma_n^2\mathbf{I}$. In the next time slot, $d_2(n)$ is first time reversed, conjugated, and negated then transmitted through antenna 1. Through antenna 2, a time reversed and conjugated version of $d_1(n)$ is transmitted. The received symbol sequence can be written as

$$\check{r}_2(n) = h_{12}(q^{-1})d_1^*(J-n) - h_{11}(q^{-1})d_2^*(J-n) + \check{w}_2(n), \tag{2.20}$$

where $\check{w}_2(n)$ is also a sequence of independent ZMCSCG elements and covariance matrix $\sigma_n^2\mathbf{I}$. Fig. 1 shows the functional blocks of a transmitter employing TR-STBC.

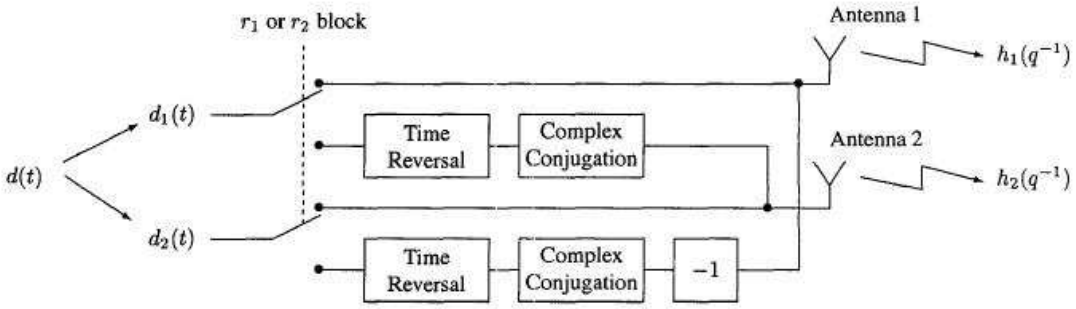

Figure 2.1: Transmission using TR-STBC [24].

At the receiver, the signal received at the second time slot is time-reversed and conjugated yielding

$$r_2(n) = \check{r}_2^*(J-n) = h_{12}^*(q)d_1(n) - h_{11}^*(q)d_2(n) + w_2(n), \tag{2.21}$$

where $w_2(n) = \check{w}_2^*(J-n)$. Using (2.19) and (2.21), we can re-formulate the signal model

as follows:

$$\begin{aligned}\mathbf{r}(n) &= \begin{bmatrix} r_1(n) \\ r_2(n) \end{bmatrix} = \begin{bmatrix} h_{11}(q^{-1}) & h_{12}(q^{-1}) \\ h_{12}^*(q) & -h_{11}^*(q) \end{bmatrix} \begin{bmatrix} d_1(n) \\ d_2(n) \end{bmatrix} + \begin{bmatrix} w_1(n) \\ w_2(n) \end{bmatrix} \\ &= \mathbf{H}(q^{-1})\mathbf{d}(n) + \mathbf{w}(n). \end{aligned} \tag{2.22}$$

Note that the channel matrix is orthogonal where

$$\mathbf{H}^H(q)\mathbf{H}(q^{-1}) = (h_{11}^*(q)h_{11}(q^{-1}) + h_{12}^*(q)h_{12}(q^{-1}))\mathbf{I}. \tag{2.23}$$

Thus, filtering the received vector $\mathbf{r}(n)$ with the matched filter $\mathbf{H}^H(q)$ yields the orthogonal sequences

$$z_1(n) = (h_{11}^*(q)h_{11}(q^{-1}) + h_{12}^*(q)h_{12}(q^{-1}))d_1(n) + n_1(n); \tag{2.24}$$

$$z_2(n) = (h_{11}^*(q)h_{11}(q^{-1}) + h_{12}^*(q)h_{12}(q^{-1}))d_2(n) + n_2(n), \tag{2.25}$$

where $\mathbf{H}^H(q)\mathbf{w}(n) = [n_1(n), n_2(n)]^T$. The detection problem of the two streams therefore decouples. Moreover, the channel in (2.24) and (2.25) is the same as in a system with one transmit antenna and two receive antennas. Thus, TR-STBC achieves full diversity. The orthogonal sequences $z_1(n)$ and $z_2(n)$ both depict single-input single-output ISI channels. They can therefore be fed *independently* to a Maximum-Likelihood (ML) sequence estimator (MLSE) or a suboptimal equalizer, in time or frequency domain, to produce the estimates $\hat{d}_1(n)$ and $\hat{d}_2(n)$ – see [26, 27, 28, 29, 30] for examples of different equalization techniques used with TR-STBC. The functional blocks of the TR-STBC receiver are shown in Fig. 2.2.

2.2.2 Extensions and Applications

TR-STBC was explored thoroughly and was used in various applications. Due to ISI, this signaling scheme suffers from edge-effects when transmitting more than one frame of data. In [26], the use of zero-padding (ZP) was proposed to overcome this problem, and

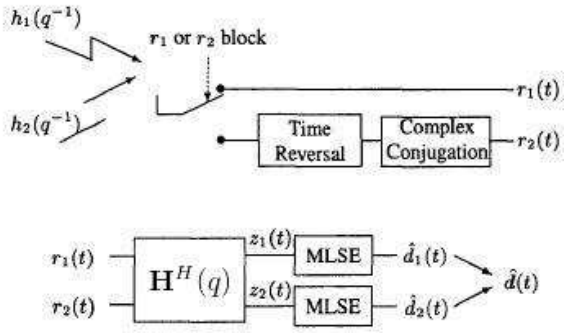

Figure 2.2: Detection using TR-STBC [24].

the authors showed that trailing each symbol stream with zeros guarantees achieving full spatial and multipath diversity.

In [31], Flore and Lindskog compared TR-STBC with a transmit delay diversity (TDD) scheme. TDD creates artificial multipath channels by delaying the data streams, and hence it increases the memory of the channel. It was shown that TR-STBC achieves higher diversity orders relative to TDD. Also, equalization complexity for TDD increases by an order of magnitude in order to achieve the same diversity order as TR-STBC.

Although TR-STBC requires equalization at the receiver, it was shown in [26, 32] that the overall complexity can be made comparable to that of single-antenna transmissions over ISI channels. Besides time-reversing the transmitted vectors, the authors in [26] showed that any permutation of the symbols of the transmitted blocks can still lead to full diversity[1]. The TR-STBC technique was extended to support any number of transmit and receive antennas in [25, 26].

An interesting application of TR-STBC was recently presented in [33, 34] where the authors extended TR-STBC to a cooperative transmission scenario with amplify-and-forward relaying. Similar to point-to-point systems, the inherited orthogonality in distributed TR-STBC results in decoupled data streams at the receivers yielding low-complexity implementations. In a single-relay network, under the assumption of perfect

[1]The authors point out an exception for codes of rate 3/4 when $M \in \{3, 4\}$ where only a specific permutation is allowed to guarantee full diversity.

power control for the relay terminal and high signal-to-noise-ratio (SNR) for the underlying links, the authors show that distributed TR-STBC achieves a maximum diversity order of $\min\{L_1, L_3\} + L_2 + 2$ where L_1, L_2, and L_3 are the channel orders of the source-to-relay, source-to-destination, and relay-to-destination links, respectively.

Jonietz *et al.* [35, 36] combined TR-STBC with transmit beamforming while taking the reliability of the CSI at the transmitter into consideration. The beamformers are designed to minimize the mean symbol error probability subject to a unit transmit power constraint. Their proposed scheme was shown to provide the same diversity gains as TR-STBC in the case of unreliable CSI and significant beamforming gains in addition to the TR-STBC diversity gains when reliable CSI is available.

2.3 Summary

In this chapter, we presented different approaches to WSR maximization using linear processing for MIMO-BC. We gave particular attention to WSRBF-WMMSE due to its low complexity. The algorithm is based on the relationship between the WMMSE minimization and WSR maximization problems where their KKT conditions can be made identical with optimized MSE weights. WSRBF-WMMSE is the basis for the algorithms presented in Chapter 3.

We have also explained the TR-STBC technique and discussed a few of its applications. TR-STBC achieves joint spatial-multipath diversity and enables the orthogonalization of data streams at the receiver which simplifies the detection process. Those advantages motivate the use of TR-STBC coupled with pre-equalization techniques in single-carrier transmission as an alternative to OFDMA. We will present this single-carrier based system in Chapter 4.

Chapter 3

WSR Maximization With AMC for MIMO-OFDMA-BC

3.1 Overview

Few attempts have been made to design linear precoders to maximize WSR for the MIMO-BC of OFDMA systems (MIMO-OFDMA-BC). The state-of-the-art algorithms devised for flat-fading systems require nested optimization or GP solvers, making them not suitable for multi-carrier systems, as they would incur prohibitive complexity. In this chapter, we extend the algorithm proposed in [16, 17] to MIMO-OFDMA and show that the complexity of the proposed method increases only linearly with the number of subcarriers relative to its single-carrier counterpart. Moreover, we offer a novel practical rate adaptation technique to translate theoretical data rates into allowed practical rates. The method first uses a known algorithm to select users and design linear precoders to maximize WSR, and then utilizes the SNR gap to capacity concept to choose the rates to allocate to each data stream based on the available AMC modes. The method is particularly useful for the case when downlink receivers have more than one antenna each, in which case WSR maximization often results in the transmission of more than

one independent data stream to each user. The main idea is that if N_k data streams are transmitted to the k-th user, the channel from the base station to the k-th receiver will effectively be a multiple access channel (MAC) with N_k "users", in which a polymatroid of rate vectors are achievable. The method then maps available AMC modes to the space of allowed theoretical rates, using the SNR gap to capacity concept, and selects the operating point with the largest sum-rate.

3.2 WSR Maximization for MIMO-OFDMA-BC Using Beamforming

Extending many of the WSR maximization algorithms that were developed for single-carrier systems to multi-carrier systems is not practical as they often involve highly complex operations. For instance, it would be impractical to extend [13, 14, 15] to N-carrier systems as this would involve solving GPs with N times the number of optimization variables. In [18], a sum-rate maximizing approach for MIMO-OFDM-BC was proposed, but it also exhibits high complexity as it requires solving a GP or a signomial program which requires prohibitive computations to yield the globally optimal solution [21]. The main obstacle is that optimization for multi-user MIMO-OFDMA systems must be performed jointly across all subcarriers and users subject to a total power constraint. This optimizes subcarrier allocation to users and beamforming design on each subcarrier but makes WSR maximization computationally expensive, especially for systems employing a large number of subcarriers. Here, we extend WSRBF-WMMSE to MIMO-OFDMA systems due to its low complexity and its reliance on closed form expressions which makes the computations very efficient – we will refer to this algorithm as WSRBF-WMMSE-OFDMA. We will show that WSRBF-WMMSE-OFDMA can allocate subcarriers to users automatically. A critical point we will highlight is that the complexity of WSRBF-WMMSE-OFDMA compared to WSRBF-WMMSE scales only linearly

with the number of subcarriers as opposed to the worst case polynomial-time complexity in the number of parameters of the optimization problem incurred by algorithms requiring a GP to be solved. In order to further reduce the complexity of the algorithm, we apply the so-called clustering technique where we group subcarriers in clusters that share the same beamformer, and we study the effect of clustering on the WSR performance through simulations.

3.2.1 Signal Model

We consider a similar model to that of Section 2.1.1, where the BS employs M transmit antennas and broadcasts to K users. We assume the channels are frequency-selective with order L, and OFDMA is employed with N subcarriers. User k has N_k antennas, and the total number of receive antennas is $N_r = \sum_{k=1}^{K} N_k$. Let $\mathbf{d}_k[n] \in \mathbb{C}^{N_k \times 1}$ be the symbol vector intended for the k-th user on the n-th subcarrier[1] with $E\{\mathbf{d}_k[n]\mathbf{d}_k^H[n]\} = \mathbf{I}$. Let the transmit beamformer to the k-th user on the n-th subcarrier be $\mathbf{B}_k[n] \in \mathbb{C}^{M \times N_k}$. The transmitted signal on the n-th subcarrier is thus given by $\mathbf{x}[n] = \sum_{k=1}^{K} \mathbf{B}_k[n]\mathbf{d}_k[n]$. Let $\mathbf{H}_k[n] \in \mathbb{C}^{N_k \times M}$ be the channel from the BS to the k-th user on the n-th subcarrier where the (i,j)-th element of $\mathbf{H}_k[n]$ is the channel gain from the j-th transmit antenna to the i-th receive antenna. The received signal at the k-th user on the n-th subcarrier would be

$$\mathbf{y}_k[n] = \mathbf{H}_k[n]\mathbf{x}[n] + \mathbf{w}_k[n] \tag{3.1}$$

where $\mathbf{w}_k[n] \in \mathbb{C}^{N_k \times 1}$ is a noise vector with independent ZMCSCG elements and covariance matrix $\mathbf{I}$.

To maximize the WSR, the WSRBF-WMMSE algorithm, described in Chapter 2, selects a subset of users $\mathcal{K}$ to serve from among all available users. It also finds the beamforming matrices $\forall k \in \mathcal{K}$. The channels to all users are assumed to be perfectly

[1] To simplify the discussion, the number of elements in $\mathbf{d}_k[n]$ is assumed to be N_k – this assumption will be dropped later.

known at the BS. The channels were assumed to be flat-fading in [16, 17], but we will assume the channels are frequency-selective and extend WSRBF-WMMSE to MIMO-OFDMA systems.

3.2.2 Outline of the Solution

There are two possible approaches to extend WSRBF-WMMSE to MIMO-OFDMA systems:

1. Perform resource allocation *jointly* over all subcarriers by extending the definition of the beamformers and channel matrices of each user to contain the frequency dimension;

2. Run WSRBF-WMMSE for each subcarrier *independently* subject to a per subcarrier power constraint (uniform power allocation across subcarriers).

Clearly, Approach 1 is the optimal approach to allocate subcarriers to different users or, equivalently, perform user selection across subcarriers. On the other hand, Approach 2 appears to be simpler as no joint processing is required. However, we will show that both approaches have the same complexity.

To adopt Approach 1, we need to re-write the received signal (3.1) to span all subcarriers. This is done by extending the definitions of the channels and beamforming matrices to include both the spatial and frequency dimensions as follows:

$$\begin{aligned} \mathbf{y}_k &= [\mathbf{y}_k[1]^T, \mathbf{y}_k[2]^T, ..., \mathbf{y}_k[N]^T]^T \\ &= \mathbf{H}_k \sum_{k=1}^{K} \mathbf{B}_k \mathbf{d}_k + \mathbf{w}_k, \end{aligned} \tag{3.2}$$

where

$$\begin{aligned}\mathbf{H}_k &= \text{blkdiag}\{\mathbf{H}_k[1], \mathbf{H}_k[2], ..., \mathbf{H}_k[N]\} \\ \mathbf{B}_k &= \text{blkdiag}\{\mathbf{B}_k[1], \mathbf{B}_k[2], ..., \mathbf{B}_k[N]\} \\ \mathbf{d}_k &= [\mathbf{d}_k^T[1], \mathbf{d}_k^T[2], ..., \mathbf{d}_k^T[N]]^T. \end{aligned} \tag{3.3}$$

WSRBF-WMMSE relies on computing a set of closed-form expression in each iteration until convergence. We introduced these expressions in Chapter 2, and we re-state them here for reference:

$$\text{MMSE Receiver Filter: } \mathbf{A}_k = \mathbf{B}_k^H \mathbf{H}_k^H (\mathbf{H}_k \mathbf{B}_k \mathbf{B}_k^H \mathbf{H}_k^H + \mathbf{C}_k)^{-1} \tag{3.4}$$

$$\text{MSE Matrix: } \mathbf{E}_k = (\mathbf{I} + \mathbf{B}_k^H \mathbf{H}_k^H \mathbf{C}_k^{-1} \mathbf{H}_k \mathbf{B}_k)^{-1} \tag{3.5}$$

$$\text{Weight Matrix: } \mathbf{W}_k = \mu_k \mathbf{E}_k^{-1} \tag{3.6}$$

$$\text{WMMSE Transmit Filter: } \bar{\mathbf{B}} = (\mathbf{H}^H \mathbf{A}^H \mathbf{W} \mathbf{A} \mathbf{H} + \frac{\text{Tr}(\mathbf{W}\mathbf{A}\mathbf{A}^H)}{P_{max}} \mathbf{I})^{-1} \mathbf{H}^H \mathbf{A}^H \mathbf{W} \tag{3.7}$$

$$\text{Normalized Transmit Filter: } \mathbf{B} = b\bar{\mathbf{B}} \tag{3.8}$$

Note that because $\mathbf{H}_k$ and $\mathbf{B}_k$ are block-diagonal matrices, the term $\mathbf{H}_k \mathbf{B}_k \mathbf{B}_k^H \mathbf{H}_k^H + \mathbf{C}_k$ in (3.4) is also block-diagonal. Therefore, computing the MMSE receiver $\mathbf{A}_k$ can be carried out using N matrix-inverse operations on $N_k \times N_k$ matrices, rather than a full-blown inversion of a $NN_k \times NN_k$ matrix. Similarly, computing $\mathbf{E}_k$ and $\mathbf{W}_k$ requires N matrix-inverse operations on $N_k \times N_k$ matrices. The term $\mathbf{H}^H \mathbf{A}^H \mathbf{W} \mathbf{A} \mathbf{H}$ in (3.7) can be written using the definitions in (3.3) as follows:

$$\mathbf{H}^H \mathbf{A}^H \mathbf{W} \mathbf{A} \mathbf{H} = \sum_{k=1}^{K} \mathbf{H}_k^H \mathbf{A}_k^H \mathbf{W}_k \mathbf{A}_k \mathbf{H}_k. \tag{3.9}$$

The summation in (3.9) results in a block-diagonal matrix with N matrices along the diagonal each of size $M \times M$. Hence, we can compute $\bar{\mathbf{B}}$ using $K \cdot N$ matrix-inverse operations on $M \times M$ matrices. The term $\text{Tr}(\mathbf{W}\mathbf{A}\mathbf{A}^H)$ can be computed efficiently since the trace of a block-diagonal matrix is simply the trace of the summation of all the

matrices along its diagonal. After calculating $\bar{\mathbf{B}}$, the beamforming matrices must be scaled by the factor $b = \sqrt{\frac{P_{max}}{\mathrm{Tr}(\bar{\mathbf{B}}\bar{\mathbf{B}}^H)}}$ to satisfy the total transmit power constraint in each iteration of the algorithm. The total number of matrix-inverse operations *per iteration* needed to execute WSRBF-WMMSE-OFDMA is:

- $K \cdot N$ matrix-inverse operations on $N_k \times N_k$ matrices to compute $\mathbf{A}_k$ using (3.4);
- $K \cdot N$ matrix-inverse operations on $N_k \times N_k$ matrices to compute $\mathbf{W}_k$ using (3.5), (3.6); and
- $K \cdot N$ matrix-inverse operations on $M \times M$ matrices to compute $\mathbf{B}$ using (3.7), (3.8);

The complexity of WSR maximization in an N-subcarrier system is therefore only N times that of the single-carrier case.

In Approach 2, we use the expressions (3.4), (3.5), (3.6), (3.7), and (3.8) on each subcarrier so that the complexity is N times that of the single-carrier case, just as in the approach described above. However, the WSR performance deteriorates compared to Approach 1, because Approach 2 assigns equal powers to each subcarrier. We therefore propose to optimize jointly over all the subcarriers in this work as an extension of WSRBF-WMMSE to MIMO-OFDMA systems.

3.2.3 Discussion

We conjecture that WSRBF-WMMSE-OFDMA will exhibit the same performance as WSRBF-WMMSE in terms of user selection on each subcarrier. In other words, the number of users selected by WSRBF-WMMSE-OFDMA on each subcarrier should be equal to the rank of the channel at high SNR. At low SNR, WSRBF-WMMSE-OFDMA should try to allocate all the transmit power to the user with the channel of the highest gain. User selection is performed automatically due to the MSE weight update in each iteration of the algorithm. We will verify this behavior by simulations.

Some of the beamforming matrices may be poorly conditioned, e.g., if $M < \sum_{k\in\mathcal{K}} N_k$. WSRBF-WMMSE-OFDMA splits the total transmit power among all the columns of $\mathbf{B}_k$ $\forall\ k \in \mathcal{K}$ which can make the power assigned to certain columns very small making the corresponding beamformer poorly conditioned. Specifically, M columns out of all the beamforming matrices of the users in $\mathcal{K}$ will contain almost all the total transmit power. In this case, although WSRBF-WMMSE-OFDMA assigns non-zero power to all beamformers of all selected users, we can consider the streams corresponding to the beamformers with very small power to have zero rate. We therefore set those low-power beamformers to zero.

In terms of WSR performance, WSRBF-WMMSE-OFDMA converges to a local optimum of the WSR maximization problem as do state-of-the-art algorithms such as [13, 14, 15], if they are to be applied to multi-carrier systems. Changing the definitions of the channel and beamforming matrices does not affect the WSRBF-WMMSE property of monotonic convergence, and the convergence analysis presented in [16, 17] still applies to WSRBF-WMMSE-OFDMA. Moreover, our WSR maximization approach for MIMO-OFDMA systems exhibits much lower complexity than those prior approaches, because the latter involve solving GPs which have a worst case polynomial-time complexity in the total number of parameters $N \cdot N_r$ [13]. Hence, WSRBF-WMMSE-OFDMA is a practical WSR maximizing algorithm that relies on closed-form expressions and avoids GPs while guaranteeing convergence to a local optimum of the WSR maximization problem.

Fig. 3.1 shows the number of iterations required for the convergence of WSRBF-WMMSE-OFDMA for different parameters. It can be seen that the algorithm needs about 5 iterations to converge without the need to solve GPs as in [13, 14, 15, 18].

Fig. 3.2 compares the achieved throughput per subcarrier (WSR/N) by DPC and WSRBF-WMMSE-OFDM for a system with $M = 4$, $K \in \{4, 20\}$, $N_k = 1$, $N = 32$, and $L = 10$. The DPC bound was obtained by applying the algorithm in [37] to parallel channels. We chose the transmit matched filter as an initialization for the beamformers,

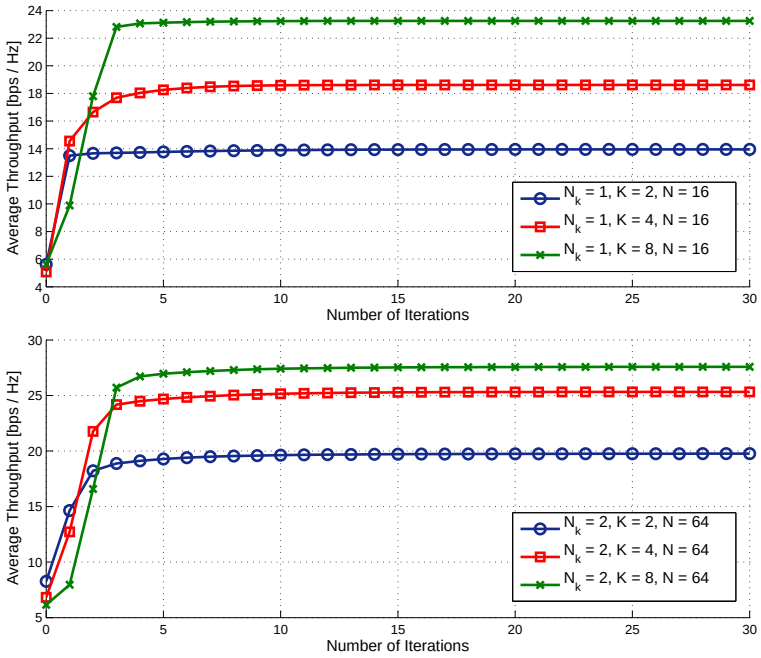

Figure 3.1: Number of iterations needed for WSRBF-WMMSE-OFDMA to converge for a fixed channel realization under different parameters. $M = 4$, $L = 10$, SNR = 20 dB.

i.e., $\mathbf{B}_k = \alpha \mathbf{H}_k^H$ where α is selected to satisfy the transmit power constraint. We allowed WSRBF-WMMSE-OFDMA to run for 10 iterations only which are enough for convergence, as suggested by Fig. 3.1. The channel gains are assumed to be i.i.d. ZMCSCG with variance σ_h^2. Quasi-static fading is assumed where the channels are constant over the duration of a codeword and change independently between codewords. The curves are generated by averaging over 1000 channel realizations. SNR is equal to σ_h^2 since both the signal and noise powers are normalized to have unit power. We observe that WSRBF-WMMSE-OFDMA traces the optimal bound closely. In Fig. 3.2(a), it is seen that WSRBF-WMMSE-OFDMA sees a loss starting at SNR = 15 dB compared to DPC. This loss is less severe in Fig. 3.2(b) due to multi-user diversity.

Fig. 3.3 illustrates the performance of WSRBF-WMMSE-OFDMA for networks with $M = 4$, $N = 32$, $L = 10$, and a different number of users each employing $N_k = 2$. In the single-user case, it can be seen that WSRBF-WMMSE-OFDMA achieves the DPC

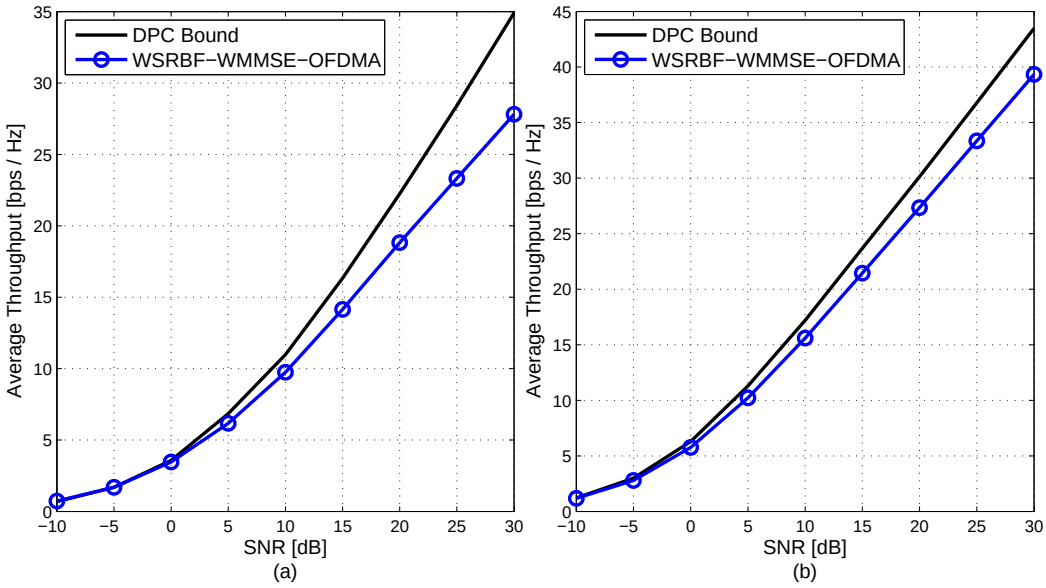

Figure 3.2: Comparing the average throughput achieved by WSRBF-WMMSE-OFDMA and DPC. $M = 4$, $N_k = 1$, $N = 32$, and $L = 10$. (a) $K = 4$ (b) $K = 20$.

bound. This is because the single-user problem is convex; hence, WSRBF-OFDMA-MMSE is guaranteed to converge to the global optimal solution regardless of the transmit filter initialization. For $K \in \{2, 7\}$, WSRBF-WMMSE-OFDMA traces the DPC bound up to 10 dB, but the performance worsens slightly for higher SNR values. It is worth mentioning that a different initializing filter might lead to an improved performance. This is because the WSR maximization problem is non-convex and the initial filter used determines whether WSRBF-WMMSE-OFDMA will converge to a local or global solution. In [16, 17], the authors choose among 10 runs of WSRBF-WMMSE each using a different initializing filter and obtain an improved performance over using the simple matched filter as an initialization.

Our simulations confirmed that WSRBF-WMMSE-OFDMA behaves as WSRBF-WMMSE in assigning all the transmit power to the user with the best channel on each subcarrier at low SNR. Some users have very low channel gains over all subcarriers that

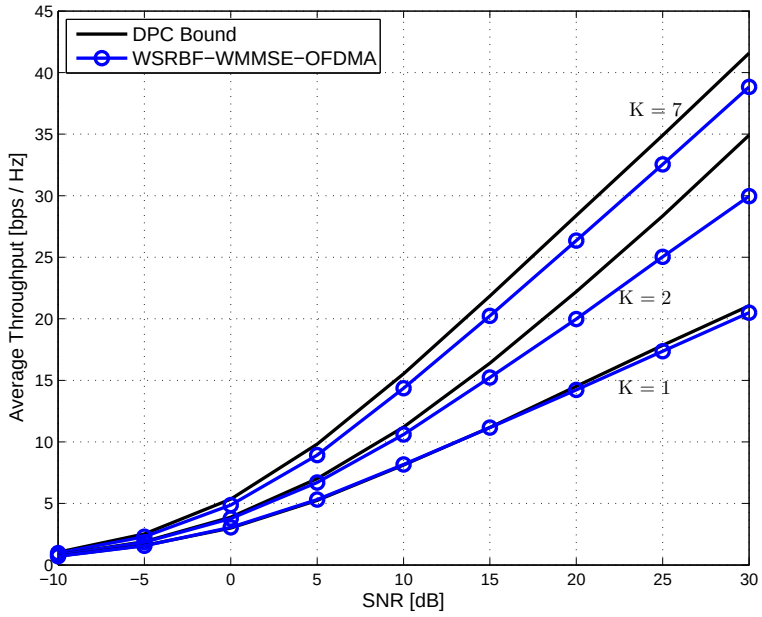

Figure 3.3: Comparing the average throughput achieved by WSRBF-WMMSE-OFDMA and DPC. $M = 4$, $K \in \{1, 2, 7\}$, $N_k = 2$, $N = 32$, and $L = 10$.

no power is assigned to them. Also, when $M < K$, WSRBF-WMMSE-OFDMA selects M users that are almost orthogonal with a good set of channels on each subcarrier at high SNR. More importantly, the algorithm is capable of assigning subcarriers to users automatically. This behaviour is demonstrated in Table 3.1 which shows the subcarrier allocation and user selection at SNR = 20 dB for a fixed channel realization of a system with the $M = 4$, $K = 18$, $N_k = 1$, $N = 16$, and $L = 2$.

3.2.4 Reduced Complexity WSRBF-WMMSE-OFDMA: Clustering

Practical MIMO-OFDMA systems employ a large number of subcarriers ranging from 48 subcarriers in IEEE 802.11a to 6817 subcarriers in Digital Video Broadcasting - Terrestrial (DVB-T). We can further reduce the complexity of the algorithm by considering the fact that fading on closely spaced subcarriers is correlated [38] which makes the beam-

Table 3.1: WSRBF-WMMSE-OFDMA subcarrier allocation. X indicates subcarrier allocation to users. $M = 4$, $K = 18$, $N_k = 1$, $N = 16$, $L = 2$, and SNR = 20 dB.

n \ k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1					X		X			X					X			
2					X		X								X			X
3	X						X						X					X
4	X						X						X			X		
5	X						X						X			X		
6	X						X						X			X		
7	X			X									X		X			
8	X			X				X					X					
9	X			X		X							X					
10	X			X		X							X					
11					X	X						X	X					
12					X							X	X				X	
13										X		X	X				X	
14										X		X	X				X	
15				X						X					X		X	
16					X		X			X					X			

formers correlated as well. This can be exploited using various methods in order to reduce the complexity of designing beamformers for a large number of subcarriers. Here, we employ clustering where we group subcarriers in clusters containing ν adjacent subcarriers. The number of subcarriers we expect to be correlated is equal to the coherence bandwidth of the channel N/L. Clustering is a common complexity reduction technique in the OFDMA literature. It has been used in [39, 40, 41] and was also adopted in the Long

Term Evolution (LTE) standard. For each cluster, we only need to design a beamformer for the center subcarrier, and use the same beamformer for all the subcarriers in that cluster. The computational savings obtained via clustering are therefore proportional to the cluster size ν.

The clustering effect on the average throughput achieved by WSRBF-WMMSE-OFDMA for $M = 4$, $K = 4$, $N_k = 1$, $N = 64$, and $L = 10$ can be seen in Fig 3.4. As the size of the clusters ν increases, more subcarriers use the same beamformer which leads to increasing the interference among users and decreasing the throughput as can be seen in the figure. However, complexity of the beamforming design decreases with increasing ν. Hence, there is a trade-off between throughput performance and computational complexity, and ν can be tuned according to the needs of the specific application.

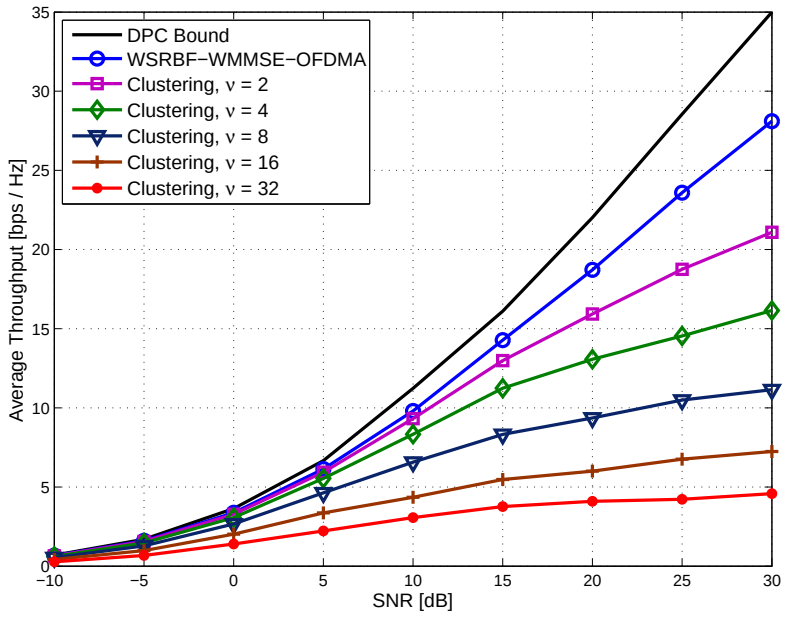

Figure 3.4: Effect of the cluster size ν on throughput of WSRBF-WMMSE-OFDMA. $M = 4$, $K = 4$, $N_k = 1$, $N = 64$, and $L = 10$.

3.3 Stream Rate Allocation

We now shift our attention to the problem of rate allocation to the streams of all users. With all of the WSR maximizing algorithms, a user with N_k receive antennas will generally be allocated a beamforming matrix of rank N_k, i.e., the BS will transmit N_k independent data streams to that user. The rate to be transmitted to the k-th user is $\tilde{R}_k$, which has to be split among N_k streams. This rate-splitting problem does not have a unique solution, because the channel between the BS and user k, with beamforming vectors fixed by a certain WSR maximization algorithm, is effectively a MAC with N_k "users" (as shown later) and any point on the maximum sum-rate surface (the -45 degree face for the two-user MAC) achieves the desired rate $\tilde{R}_k$.

An important contribution of this chapter is that we propose a generally applicable method to translate theoretical MIMO precoder designs, such as the one in [16, 17], into *practical* designs that choose the best AMC mode for each user, for a given desired bit error probability P_b. To do this, we assume that a QAM format is used, and that the coding gain of the error control code employed is known. Further, we suggest a reduced complexity method that searches over a reduced number of points in the space of allowed AMC modes. Although we limit our analysis here to beamformers designed by the WSRBF-WMMSE-OFDMA algorithm, our technique is general and can be applied to beamformers designed with other WSR maximizing approaches. For MIMO-OFDMA systems, our AMC method can be applied on each subcarrier to choose the best modulation and coding schemes without requiring expensive computations.

3.3.1 Problem Description

Assuming the proposed algorithm has selected the subset of users $\mathcal{K}$ for a given subcarrier, the BS transmits up to N_k independent data streams to the k-th user, so that the total data rate for user k, $\tilde{R}_k$, is split among N_k streams, or $\tilde{R}_k = \sum_{l=1}^{N_k} \tilde{R}_{k,l}$ where $\tilde{R}_{k,l}$ is the

rate allocated to stream l of user k. Therefore, we have the problem of choosing the best AMC mode for each stream.

Without loss of generality, we will let $N = 1$ and assume the channels to be flat-fading, as each subcarrier experiences flat fading in OFDMA. We will therefore drop the subcarrier index n for the rest of the section to simplify notation. Also, define $\mathbf{d}_k = [d_{k,1}, ..., d_{k,N_k}]^T$ and $\mathbf{G}_k = [\mathbf{g}_{k,1}, ..., \mathbf{g}_{k,N_k}]$. Then, (3.2) can be re-written as

$$\mathbf{y}_k = \sum_{l=1}^{N_k} \mathbf{g}_{k,l} d_{k,l} + \sum_{j=1, j\neq k}^{K} \mathbf{H}_k \mathbf{B}_j \mathbf{d}_j + \mathbf{w}_k. \tag{3.10}$$

By considering the multi-user interference (MUI) term $\sum_{j\neq k} \mathbf{H}_k \mathbf{B}_j \mathbf{d}_j$ as part of the Gaussian noise, we can further write

$$\mathbf{y}_k = \sum_{l=1}^{N_k} \mathbf{g}_{k,l} d_{k,l} + \mathbf{v}_k = \mathbf{G}_k \mathbf{d}_k + \mathbf{v}_k, \tag{3.11}$$

where $\mathbf{v}_k$ is coloured and Gaussian if the number of interfering terms is large. Expression (3.11) is identical to that of a vector Gaussian MAC [42], in which user l transmits the symbol $d_{k,l}$ over a vector channel $\mathbf{g}_{k,l}$. This resemblance of the downlink channel between the base station and a receiver with multiple antennas to a Gaussian MAC is a key realization used in the rate allocation algorithm to follow.

Given the equivalence between (3.11) and the vector Gaussian MAC with N_k users, all rate vectors $(\tilde{R}_{k,1}, \ldots, \tilde{R}_{k,N_k})$ in the interior of the polymatroid MAC capacity region defined by the channels $(\mathbf{g}_{k,1}, \ldots, \mathbf{g}_{k,N_k})$ are achievable [43]. All points on the maximum sum-rate surface yield the designed total rate $\tilde{R}_k$ for user k, and in theory any of these rate vectors can be used (with time sharing if the point is not one of the vertices of that surface) with the beamforming matrix $\mathbf{B}_k$ designed by the algorithm in [16, 17].

In practice, there are two problems with the above method:

1. The rates $\tilde{R}_{k,l}$ are not achievable using practical modulation and coding – there is a gap to capacity which is not negligible;

2. Even with AMC, a system cannot transmit at arbitrary rates. Instead only a finite number of rate values are available.

The algorithm described in the next section seeks to address both problems. In the remainder of this section, we outline the basis for the proposed method.

3.3.2 Outline of the Solution

Depending on the structure of the receiver, the solution to the rate-splitting problem may differ. In this work, we have considered two receiver structures: ML and successive interference cancellation (SIC).

ML Receiver

By defining the SNR gap to capacity in the l-th stream of the k-th user as $\Gamma_{k,l}$ [44, 45], its practically achievable rate is given by

$$R_{k,l} = \log_2 \left(1 + \frac{1}{\Gamma_{k,l}} \mathbf{g}_{k,l}^H \mathbf{C}_{k,l}^{-1} \mathbf{g}_{k,l} \right), \tag{3.12}$$

where $\mathbf{C}_{k,l}$ is the covariance matrix of the interference seen by $d_{k,l}$, which includes AWGN, MUI, and intra-user (inter-stream) interference. Therefore we have

$$\mathbf{C}_{k,l} = \mathbf{I} + \sum_{j=1, j\neq k}^{K} \mathbf{H}_k \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_k^H + \sum_{i=1, i\neq 1}^{N_k} \mathbf{g}_{k,i} \mathbf{g}_{k,i}^H, \tag{3.13}$$

At high SNR, (3.12) can be approximated as [45]:

$$\begin{aligned} R_{k,l} &\approx \log_2(1 + \mathbf{g}_{k,l}^H \mathbf{C}_{k,l}^{-1} \mathbf{g}_{k,l}) - \log_2(\Gamma_{k,l}) \\ &= \tilde{R}_{k,l} - \log_2(\Gamma_{k,l}), \end{aligned} \tag{3.14}$$

and hence

$$\tilde{R}_{k,l} \approx R_{k,l} + \log_2(\Gamma_{k,l}). \tag{3.15}$$

Equation (3.14) says that if the theoretical achievable rate is $\tilde{R}_{k,l}$, the practical achievable rate is lower by about $\log_2(\Gamma_{k,l})$ bits; equation (3.15) tells us that if we want to transmit

at the rate $R_{k,l}$, we need to design a system that has a theoretical achievable rate about $\log_2(\Gamma_{k,l})$ bits higher.

Problem 2 above is addressed in conjunction with Problem 1 by using (3.15) to derive the $\tilde{R}_{k,l}$ corresponding to each allowable practical rate $R_{k,l} = r \cdot \log_2 \mathcal{M}$, where r is the code rate and $\mathcal{M}$-ary modulation is used. Assume the transmitter can choose among P $\mathcal{M}$-ary square QAM constellations with $\mathcal{M} \in \{2^{b_1}, 2^{b_2}, ..., 2^{b_P}\}$ and Q coding schemes $\{d_1, d_2, ..., d_Q\}$ each associated with a code rate $r(d_q) \in (0, 1]$ and a coding gain $G^c(d_q)$. The SNR gap due to the use of 2^{b_p}-ary QAM inputs rather than Gaussian inputs is known to be [44]

$$\Gamma_{k,l}^{mod} = \frac{1}{3}\left(Q^{-1}\left(\frac{b_p P_b}{4(1-2^{-b_p/2})}\right)\right)^2, \text{where } P_b \text{ is the bit error rate (BER).} \tag{3.16}$$

If the coding gain provided by a coding scheme d_q at a certain desired error probability for stream l of user k is denoted $G_{k,l}^c(d_q)$, then we can write the gap to capacity for each feasible choice of coding and modulation formats as

$$\Gamma_{k,l} = \frac{\Gamma_{k,l}^{mod}}{G_{k,l}^c(d_q)}. \tag{3.17}$$

The literature contains tables of coding gains achieved by different coding schemes at a given error probability. For example, coding gains achieved by convolutional codes with soft-decision Viterbi decoding are given in [46].

Having defined the gap to capacity, the constellation $\mathcal{V}_k$ of allowable theoretical rate vectors $(\tilde{R}_{k,1}, \tilde{R}_{k,2}, \ldots, \tilde{R}_{k,N_k})$ can be found from the corresponding set of allowable practical rate vectors $(R_{k,1}, R_{k,2}, \ldots, R_{k,N_k})$. By finding the members of $\mathcal{V}_k$ contained within the MAC capacity region with channels $\mathbf{g}_{k,1}, \mathbf{g}_{k,2}, \ldots, \mathbf{g}_{k,N_k}$, denoted as $\mathcal{C}_k^{MAC}(\mathbf{H}_k, \mathcal{B})$ to reveal the dependence of the capacity region on the k-th downlink channel and the set of designed precoders $\mathcal{B} = \{\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_K\}$, we obtain the subset of allowable rate vectors. From this subset of $\mathcal{V}_k$, we choose the operating point that yields the highest practical sum rate $\sum_l R_{k,l}$. If there are more than one operating point giving this rate, then we

choose the one that is furthest from the capacity boundary, in order to provide the largest margin of error.

To illustrate the above concepts, consider Fig. 3.5 which simulates the rate region corresponding to a given channel realization of the first user in a system with $M = 4$, $K = 2$, and $N_k = 2$ employing uncoded modulation at SNR = 20 dB and $P_b = 10^{-3}$. We assume the BS has access to four modulation schemes $\{4, 16, 64, 256\}$-QAM. The points shown are the possible operating points with the given modulation schemes. The coordinates of the operating points were calculated using (3.15). The set $\mathcal{X}_1 \cup \mathcal{X}_2$ contains all rate maximizing points in $\mathcal{V}_1$. The point in the set $\mathcal{X}_1$ corresponds to $R_{1,1} = 2$ and $R_{1,2} = 4$, while the point in $\mathcal{X}_2$ corresponds to $R_{1,1} = 4$ and $R_{1,2} = 2$. The maximum practical rate available to this user is therefore 6 bits. In this example, the algorithm would choose the point belonging to $\mathcal{X}_1$ as it is the furthest from the boundaries of the rate region which would allow for the largest error margin.

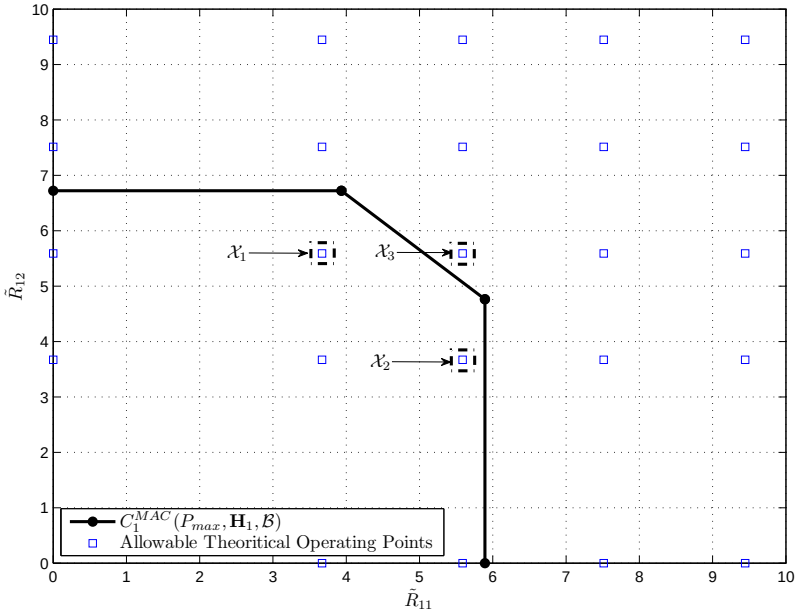

Figure 3.5: An example with multiple solutions. $M = 4$, $K = 2$, $k = 1$, $N_k = 2$, SNR = 20 dB, and $P_b = 10^{-3}$.

SIC Receiver

A SIC receiver allows for different possible decoding orders. Therefore, for each practical rate vector, there exists $N_k!$ theoretical rate vectors corresponding to all the possible decoding orders.

In this case, the covariance matrix of the interference seen by $d_{k,l}$ is

$$\mathbf{C}_{k,l} = \mathbf{I} + \sum_{j=1, j\neq k}^{K} \mathbf{H}_k \mathbf{B}_j \mathbf{B}_j^H \mathbf{H}_k^H + \sum_{i=l+1}^{N_k} \mathbf{g}_{k,i} \mathbf{g}_{k,i}^H, \tag{3.18}$$

assuming a SIC receiver acting on streams 1 through N_k successively.

We note that the last term on the RHS of (3.18) depends on decoding order. Suppose in (3.12), we replaced $\mathbf{C}_{k,l}$ with $\mathbf{C}_{k,N_k}$, and introduced another SNR gap parameter $\Gamma_{k,l}^{dec}$ so that

$$R_{k,l} = \log_2 \left(1 + \frac{1}{\Gamma_{k,l}} \mathbf{g}_{k,l}^H \mathbf{C}_{k,N_k}^{-1} \mathbf{g}_{k,l} \right), \tag{3.19}$$

where now

$$\Gamma_{k,l} = \frac{\Gamma_{k,l}^{mod} \Gamma_{k,l}^{dec}}{G_{k,l}^c(d_q)}. \tag{3.20}$$

The new parameter $\Gamma_{k,l}^{dec}$ represents the increase in SNR required to transmit rate $R_{k,l}$ as the l-th decoded stream, rather than the last decoded stream.

From (3.12) and (3.17),

$$R_{k,l} = \log_2 \left(1 + \frac{G_{k,l}^c(d_q)}{\Gamma_{k,l}^{mod}} \mathbf{g}_{k,l}^H \mathbf{C}_{k,l}^{-1} \mathbf{g}_{k,l} \right), \tag{3.21}$$

and therefore, from (3.19) and (3.20),

$$\Gamma_{k,l}^{dec} = \frac{\mathbf{g}_{k,l}^H \mathbf{C}_{k,N_k}^{-1} \mathbf{g}_{k,l}}{\mathbf{g}_{k,l}^H \mathbf{C}_{k,l}^{-1} \mathbf{g}_{k,l}} \tag{3.22}$$

is the SNR gap at stream l due to its position in the decoding order.

It is worth mentioning that certain decoding orders can result in operating points lying outside the achievable rate region. This emphasizes the importance of the decoding order gap to capacity concept. Fig. 3.6 shows the MAC rate region corresponding to the first user in a system with $K = 2$ using no error control coding and operating at

SNR = 15 dB and $P_b = 10^{-6}$. Consider the points corresponding to both streams having non-zero rate in which case SIC is applicable. Here, all operating points corresponding to decoding order $\boldsymbol{\xi}_2$ (detect stream 2 then stream 1), with non-zero rate on both streams, lie outside the rate region and may violate the target BER. We therefore conclude that when a SIC receiver is applied, it is necessary to determine the decoding order at the transmitter, as not all decoding orders guarantee achieving the target BER.

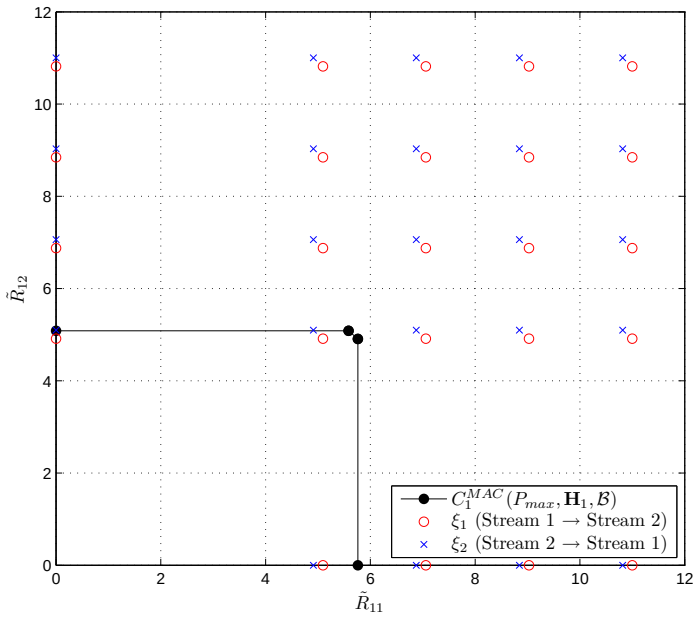

Figure 3.6: Points corresponding to one of the two possible decoding orders for the first user lie outside the rate region. $M = 4$, $K = 2$, $k = 1$, $N_k = 2$, SNR = 15 dB, and $P_b = 10^{-6}$.

The next section provides the finer details of this algorithm.

3.3.3 Finer Points of the Proposed Rate Allocation Algorithm

No Feasible Rate Vectors

Because only a finite number of rate values are available in a practical system, a feasible rate vector might not exist for some $k \in \mathcal{K}$ for a given channel realization. We handle

this case by removing the user with the worst channel, i.e., having the lowest channel gain, from $\mathcal{K}$ and redesigning the beamformers of the remaining users in the set. We repeat this process until a feasible rate vector is found for each remaining user in $\mathcal{K}$. It should be noted that this case usually does not occur when coding is employed.

Allowing One or More Streams to Have Zero Rate

The number of cases where no feasible solutions exist can be reduced, even if coding is not used, if we allow one or more streams to have zero rate. Although the beamformers designed by WSRBF-WMMSE-OFDMA assume N_k streams transmitted to each user, it can be seen from Fig. 3.11, discussed below, that allowing some streams to have zero rate yields a higher WSR in the uncoded case. When powerful coding is used, this approach does not offer major improvements as feasible rate vectors usually exist.

Operating Points not Achievable by SIC

Some operating points falling inside the rate region are not achievable by a SIC receiver. In Fig. 3.7, the point in $\mathcal{X}_1$ is not achievable by either decoding order of the SIC receiver and it requires time-sharing between both decoding orders. Also, the points in $\mathcal{X}_1$ in Fig. 3.8 correspond to decoding order $\boldsymbol{\xi}_1$ but exceed the vertex corresponding to this decoding order; therefore, they are not achievable by a SIC receiver. A similar argument applies to the points in $\mathcal{X}_2$. Hence, the rate achievable by the stream rate allocation algorithm should see a drop if a SIC receiver is used rather than a ML receiver which can achieve any point inside the rate region.

Our proposed rate allocation algorithm is shown in Table 3.2, where $\mathcal{R}_k$, $\mathcal{D}_k$, and $\boldsymbol{\pi}_k$ are the modulation formats, the coding schemes, and the SIC decoding order of the streams of user k, respectively. In the table, $S_k \leq N_k$ denotes the total number of streams transmitted to the k-th user.

One should note that if a SIC receiver is used, points such as those in Fig. 3.7 and

Table 3.2: Stream Rate Allocation.

1: Run WSRBF-WMMSE-OFDMA and obtain $\mathcal{B}$ and $\mathcal{K}$
2: **for all** $k \in \mathcal{K}$ **do**
3: **for all** $l \in \{1, ..., S_k\}$, $\boldsymbol{\pi}_k \in \{\boldsymbol{\xi}_1, ..., \boldsymbol{\xi}_{S_k!}\}$, and $R_{k,l} \in \{b_1 r(d_1), b_1 r(d_2), ..., b_P r(d_Q)\}$ **do**
4: $\tilde{R}_{k,l} = R_{k,l} + \log_2(\Gamma_{k,l})$
5: **end for**
6: $(\mathcal{R}_k^*, \mathcal{D}_k^*, \boldsymbol{\pi}_k^*) = \text{argmax}_{\mathcal{R}_k, \mathcal{D}_k, \boldsymbol{\pi}_k} \{\sum_{l=1}^{S_k} R_{k,l}\}$ s.t. $(\tilde{R}_{k,1}, ..., \tilde{R}_{k,S_k}) \in \mathcal{C}_k^{MAC}(\mathbf{H}_k, \mathcal{B})$
7: **if** $\mathcal{R}_k^* = \{\emptyset\}$ **then**
8: Remove the user with the weakest channel from $\mathcal{K}$
9: Run WSRBF-WMMSE-OFDMA over the reduced set of selected users
10: Go to step 2
11: **end if**
12: **if** $\mathcal{R}_k^*$ is not unique **then**
13: Pick operating point furthest from the boundary of $\mathcal{C}_k^{MAC}(\mathbf{H}_k, \mathcal{B})$
14: **end if**
15: **end for**

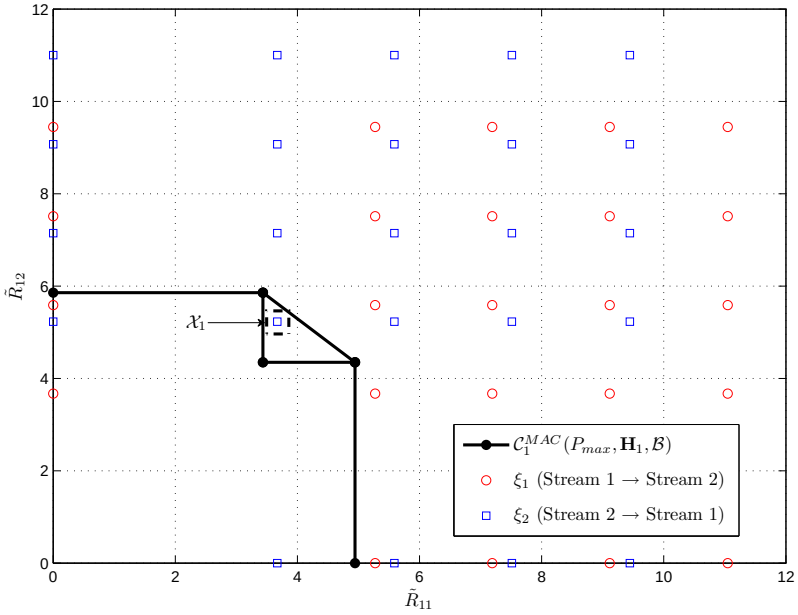

Figure 3.7: An example with multiple solutions. $M = 4$, $K = 2$, $k = 1$, $N_k = 2$, SNR = 20 dB, and $P_b = 10^{-3}$.

Fig. 3.8 are not feasible solution for the maximization of step 6.

Reduced Complexity Algorithm

The total number of points the above algorithm searches over to find the best AMC mode for user k is $(P \cdot Q)^{S_k} \cdot S_k!$. Searching the entire set $\mathcal{V}_k$ (step 6 in Table 3.2) is a computationally expensive operation especially for large P, Q, or S_k. Because the optimal solution is most likely to exist in the vicinity of the maximum sum-rate face of the MAC region, a reduced complexity search would consider the points closest to that face of the rate region only. We can do this by ordering the elements of $\mathcal{V}_k$ according to the $\tilde{R}_k$ they represent and performing a bisection search on the ordered list to pick the one with the sum-rate lower than the desired $\tilde{R}_k$ by the smallest amount.

For the SIC receiver case, and since points such as $\mathcal{X}_1$ in Fig. 3.7 are not achievable, a solution can be reached by finding the coordinates of the $N_k!$ vertices of the rate

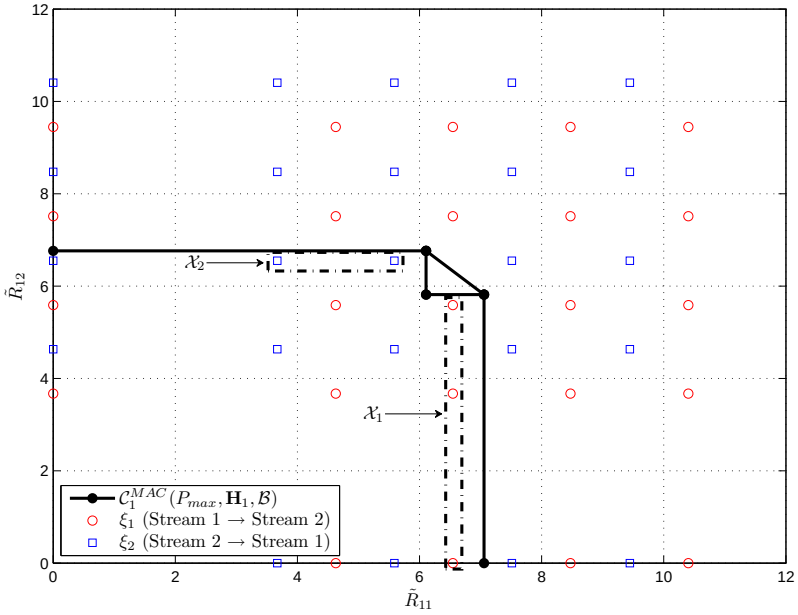

Figure 3.8: An example with multiple solutions. $M = 4$, $K = 2$, $k = 1$, $N_k = 2$, SNR $= 20$ dB, and $P_b = 10^{-3}$.

region, rounding them to the next available theoretical rate values, and choosing the point corresponding to the decoding order that produces the highest data rate. The problem is therefore simplified, and we can avoid searching within the rate region.

We now verify that the target BER is achieved when transmitting using the parameters specified by the proposed method. The transmitter employs the beamformers designed by WSRBF-WMMSE-OFDMA, and either ML or SIC receivers are used at each user. When coding is employed, a feasible solution exists almost always. Therefore, we limit the simulations here to uncoded systems in order to observe cases with no feasible rate vectors.

We first estimate the BER for a system with ML receivers. The point in the set $\mathcal{X}_1$ of Fig. 3.5 is inside the rate region, whereas $\mathcal{X}_3$ contains an infeasible solution to the AMC problem as it falls outside the rate region. Table 3.3 shows the estimated BER of each stream for both operating points. Clearly, the target error rate of 10^{-3} is achieved when

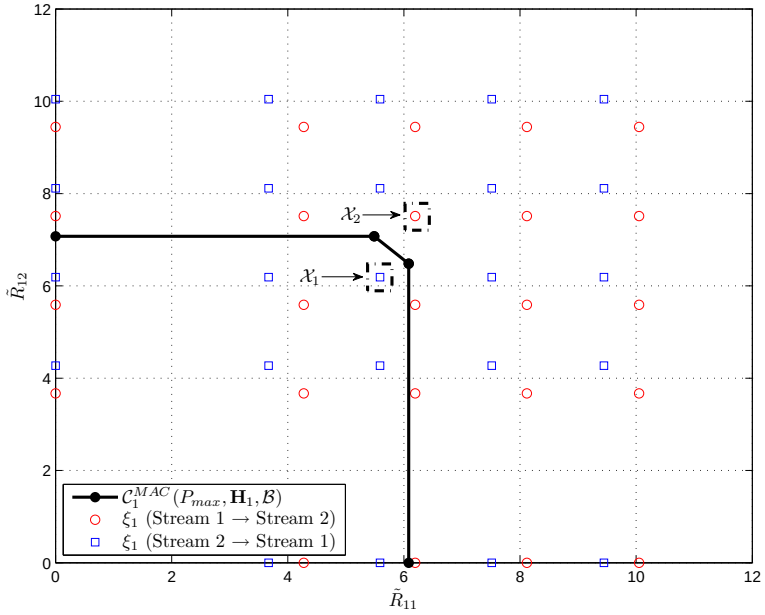

Figure 3.9: Rate region corresponding to a fixed channel realization. $M = 4$, $K = 2$, $k = 1$, $N_k = 2$, SNR $= 20$ dB, and $P_b = 10^{-3}$.

operating with a feasible point, and not achieved otherwise.

Table 3.3: Average BER per stream for a system with ML receivers. $M = 4$, $K = 2$, $k = 1$, $N_k = 2$, SNR $= 20$ dB, and $P_b = 10^{-3}$.

	$\mathcal{X}_1$	$\mathcal{X}_2$
$\hat{P}_{b,11}$	0.2×10^{-4}	0.34×10^{-2}
$\hat{P}_{b,12}$	0.1×10^{-4}	2.3×10^{-2}

We repeat the same experiment for a system with SIC receivers. Fig. 3.9 shows the rate region corresponding to a channel realization of the first user in a system with $K = 2$ operating at SNR $= 20$ dB and $P_b = 10^{-3}$. The BER for the streams of the points in $\mathcal{X}_1$ and $\mathcal{X}_2$ are shown in Table 3.4, and we can see that the target error rate is also achieved for the feasible operating point when a SIC receiver is employed.

We now compare the WSR achieved by WSRBF-WMMSE-OFDMA and our rate

Table 3.4: Average BER per stream for a system with SIC receivers. $M = 4$, $K = 2$, $k = 1$, $N_k = 2$, SNR $= 20$ dB, and $P_b = 10^{-3}$.

	$\mathcal{X}_1$	$\mathcal{X}_2$
$\hat{P}_{b,11}$	0.6×10^{-3}	1.2×10^{-3}
$\hat{P}_{b,12}$	0.1×10^{-3}	4.3×10^{-3}

allocation algorithm. Fig. 3.10 shows the average WSR achieved by both algorithms for $M = 4$, $K = 20$, and $P_b = 10^{-3}$. For the stream rate allocation algorithm, we assume that two coding schemes d_1 and d_2 are available with $r(d_1) = r(d_2) = 1/2$ and $G^c(d_1) = 4$ dB, $G^c(d_2) = 5.5$ dB. We estimate the rate for systems employing ML and SIC receivers. We observe that our algorithm traces the ideal rate of WSRBF-WMMSE-OFDMA for both type of receivers. In the ML receiver case, the throughput of the proposed algorithm caps at around 32 bits for SNR ≥ 15 dB. This can be explained by noting that at high SNR with $M < K$, WSRBF-WMMSE-OFDMA selects M users on each subcarrier, that are nearly orthogonal and have high channel gains. Also, the largest modulation scheme available is 256-QAM, and each user can receive a maximum of $N_k = 2$ symbols. Each stream carries a maximum of 4 information bits due to the rate of the employed coding schemes. Hence, the maximum throughput possible at high SNR is 32 bits. For the SIC case, the receiver is not able to achieve all the points in the rate region, and the gap to capacity due to the decoding order reduces the number of points in the region which limits the number of total received streams that are assigned practical rates to about 4 data streams, and hence the sum-rate reaches a maximum of 16 bits.

Fig. 3.11 compares the average WSR achieved by WSRBF-WMMSE-OFDMA and the rate allocation algorithm for uncoded modulation and using ML receivers. Method 1 handles the case of no feasible vectors by redesigning the precoders for a smaller set of users as explained in Section 3.3.3. Method 2 does not handle this case and the rate yielded by our algorithm would be zero. Method 3 does not allow any stream to have

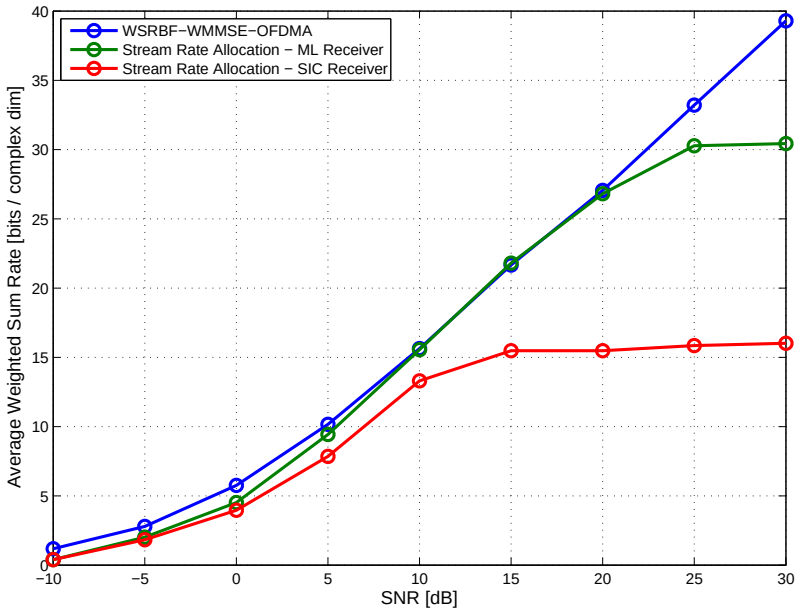

Figure 3.10: Comparing the average WSR achieved by WSRBF-WMMSE-OFDMA and the proposed rate allocation algorithm with ML and SIC receivers. $M = 4$, $K = 20$, $N_k = 2$, and $P_b = 10^{-3}$.

zero rate, but handles this case using the same method of Method 1. As can be seen from the figure, Method 1 outperforms Method 2, and they both outperform Method 3. The insignificant difference between Methods 1 and 2 implies that this case does not occur frequently when some streams are allowed to have zero rate, even when coding is not used, and almost never occurs at high SNR values. However, Method 3 is at a disadvantage, because not allowing some streams to have zero rate increases the cases with no feasible solutions.

3.4 Summary

In this chapter, we proposed a practical WSR maximization algorithm using linear processing for MIMO-OFDMA systems. The increase of complexity induced by WSRBF-

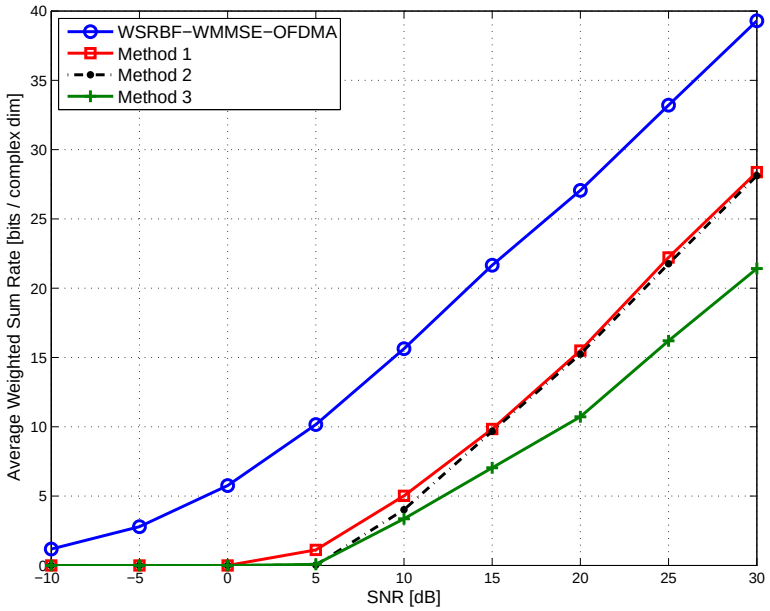

Figure 3.11: Averaged WSR achieved when allowing some streams to have zero rate and the effect on the number of cases where no feasible rate vector exists. $M = 4$, $K = 20$, $N_k = 2$, and $P_b = 10^{-3}$.

WMMSE-OFDMA relative to its single-carrier counterpart was shown to scale only linearly with the number of subcarriers offering an attractive solution for the WSR maximization problem for multi-carrier systems. We also proposed a rate allocation algorithm to translate theoretical data rates achieved by the beamformers designed using a WSR maximizing algorithm into practical AMC modes for each user, for a given desired bit error probability. The SNR gap to capacity concept was extended to include the effect of order of detection in SIC and coding gain of the employed coding scheme; it was then used to map available AMC modes to the space of allowed theoretical rates. This chapter attempts to bridge the gap between theory and practice with regards to WSR maximization, and our proposed algorithms can be utilized in future research in order to have new insights on the WSR maximization problem for MIMO-OFDMA systems as well as arrive at practical realizations of MIMO beamformers designed from information

theoretic principles.

Chapter 4

Downlink Transmission Using TR-STBC

4.1 Overview

Having studied beamforming design with the objective of maximizing the WSR, we now explore the use of precoding as a pre-equalization mechanism. In this chapter, we present an alternative to multi-carrier transmission on the downlink of a multi-user multi-antenna system, in which the base station has M antennas and downlink receiver k has N_k antennas. The proposed system makes novel use of the single-carrier TR-STBC method for orthogonal multiplexing of information to up to M users. Instead of using TR-STBC for transmission of one user's information, we use its orthogonalization feature to transmit to a small number of users in the downlink, e.g., in a broadband wireless backhaul solution. TR-STBC was shown to achieve the full diversity gain of $MN_r(L+1)$ [25, 26] for a point-to-point system using N_r receive antennas and transmitting over frequency-selective channels of order L. Assuming perfect channel knowledge at the BS, we will perform the required pre-equalization of each user's data at the transmitter through the zero-forcing (ZF) implementation of THP, rather than at the individual

receivers. The resulting system is a single-carrier one, not based on OFDMA, which we call THP-TR-STBC. In comparison with OFDMA-based techniques, the proposed method (i) does not require coding to extract the frequency diversity in the channel, and (ii) exhibits similar complexity to MIMO-OFDMA under certain conditions. THP-TR-STBC is geared towards achieving spatial and multipath diversity; this makes STBC-OFDMA a suitable OFDMA-based system for comparison with our proposed single-carrier system. STBC-OFDMA achieves spatial diversity using space-time coding and achieves multipath diversity using error control coding which introduces redundancy among subcarriers. A comparative analysis between STBC-OFDM and TR-STBC for point-to-point communication was presented in [28, 29]. In this chapter, we present a detailed comparison of STBC-OFDMA and THP-TR-STBC in terms of performance and complexity for point-to-multi-point networks. We also discuss the strengths and weaknesses of these two classes of transmission methods.

4.2 Pre-equalization Methods for MIMO-BC

We consider a fixed broadband wireless system where the BS employs M transmit antennas, and it broadcasts to $K \leq M$ users with user k having N_k antennas. Let $\mathbf{u}_k = [u_k(0), u_k(1), ..., u_k(N-1)]^T$ be the symbol block intended for the k-th user where $u_k(n)$ is an information symbol drawn from a certain coded constellation's alphabet, with variance σ_u^2. Also, let $\mathbf{h}_{k,ij} = [h_{k,ij}(0), h_{k,ij}(1), ..., h_{k,ij}(L)]^T$ be the equivalent discrete-time ISI channel impulse response (CIR) between the i-th receive antenna and the j-th transmit antenna of the k-th user, where L is the channel order. Quasi-static fading is assumed where the CIRs are considered constant over α consecutive time slots – α depends on the underlying space-time code to be used [25, 47] – and may vary independently after each α time slots. Zero-padding is used in order to remove inter-block interference (IBI) in the single-carrier case. A block of all-zeros of length $C \geq L$ is appended to each sym-

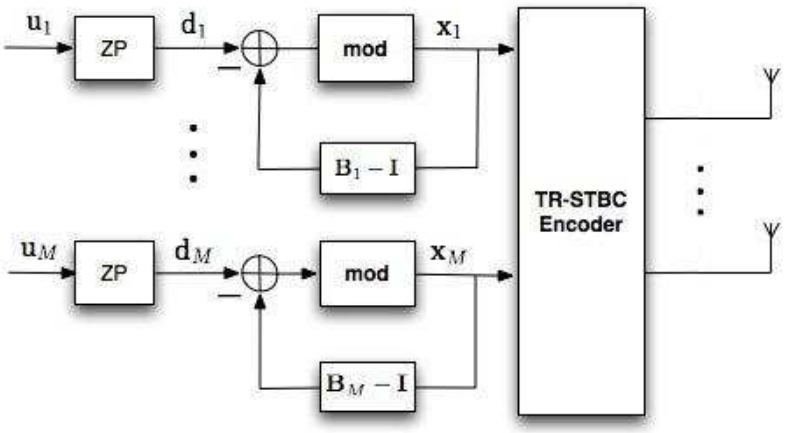

Figure 4.1: System diagram of the BS of THP-TR-STBC.

bol block yielding the length $J = N + C$ blocks $\mathbf{d}_k = [\mathbf{u}_k^T, \mathbf{0}^T]^T$. In the OFDMA-based system, a cyclic-prefix (CP) with length C is appended to each symbol block to combat IBI. It is then removed at the receivers.

We now present the system models of THP-TR-STBC and STBC-OFDMA which perform equalization and spatial multiplexing and achieve diversity differently. Throughout this chapter, we assume that the BS and the receivers have full CSI.

4.2.1 THP-TR-STBC

The system block diagrams of the BS and the k-th user in a THP-TR-STBC system are shown in Fig. 2.1 and Fig. 2.2. The TR-STBC encoding and decoding blocks were shown in Fig. 4.1 and Fig. 4.2. The zero-padded symbol blocks $\mathbf{d}_k$ are passed to a precoder in order to pre-cancel ISI. Different pre-equalization methods can be applied at the BS. Linear pre-equalization techniques, such as ZF, are known to suffer from noise enhancement. In this work, we utilize ZF-THP which is a nonlinear precoding technique that overcomes the noise-enhancement problem of linear zero-forcing precoders. From Fig. 4.1, the precoded signal intended for the k-th user is given by [48]

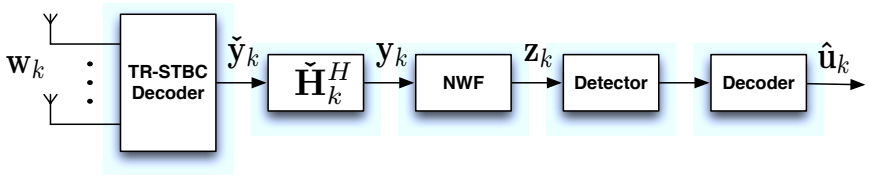

Figure 4.2: System diagram of the k-th user in a THP-TR-STBC system.

$$\mathbf{x}_k = f(\mathbf{d}_k - (\mathbf{B}_k - \mathbf{I})\mathbf{x}_k). \tag{4.1}$$

The matrix $\mathbf{B}_k$ will be defined later. The function $f(.)$ is designed to limit the power of the transmitted signal. Particularly, $f(.)$ is the modulo operation, where $\mathcal{M}$ is a square integer equal to the size of the constellation:

$$f(\mathbf{x}) = \begin{cases} \mathbf{x} & \text{if } -\sqrt{\mathcal{M}} \leq x_n \leq \sqrt{\mathcal{M}} \\ \mathbf{x} + (2\sqrt{\mathcal{M}}p)\mathbf{e}_n & \text{if } x_n \notin \mathbb{T} \\ & p \in \mathbb{Z} \text{ s.t. } x_n + 2\sqrt{\mathcal{M}}p \in \mathbb{T} \end{cases} \tag{4.2}$$

where $\mathbb{T} = [-\sqrt{\mathcal{M}}, \sqrt{\mathcal{M}})$, $n \in \{0, 1, ..., N-1\}$, and $\mathbf{e}_n$ is a unit-vector with the n-th entry equal to one and zeros elsewhere. A linearized description of $f(.)$ can be obtained by noticing that $f(\mathbf{x}_k) = \mathbf{x}_k + \mathbf{b}_k$ where $b_{k,n} = 2\sqrt{\mathcal{M}}p$, $p \in \mathbb{Z}$ [48]. We can therefore write

$$\begin{aligned} \mathbf{x}_k &= \mathbf{d}_k - (\mathbf{B}_k - \mathbf{I})\mathbf{x}_k + \mathbf{b}_k \\ \therefore \mathbf{x}_k &= \mathbf{B}^{-1}(\mathbf{d}_k + \mathbf{b}_k). \end{aligned} \tag{4.3}$$

The precoded signal $\mathbf{x}_k$ is then fed to a TR-STBC encoder. The encoding is done according to a certain encoding rule which depends on the number of antennas employed. Time-reversal processing includes conjugating, negating, and time-reversing the symbol blocks at certain time slots (see Chapter 2 for details). The encoded vectors are then transmitted over the MIMO channel from the BS to the k-th user which is denoted by the $((J+L)N_k) \times JM$ matrix $\mathbf{H}_k$ with:

$$[\mathbf{H}_k]_{ij} = \mathbf{H}_{k,ij}. \tag{4.4}$$

Here, $\mathbf{H}_{k,ij}$ is a $(J+L) \times J$ Toeplitz banded matrix whose first row is $[h_{k,ij}(0), 0, ..., 0]$, and first column is $[h_{k,ij}(0), ..., h_{k,ij}(L), 0, ..., 0]^T$.

At the receiver, a TR-STBC decoder performs the necessary time-reversal processing at specific time slots. The number of signaling time slots α required for the orthogonalization of the transmitted blocks depends on M or equivalently the rate of the underlying TR-STBC. Let $\check{\mathbf{H}}_k$ be a $(\alpha(J+L)N_k) \times JM$ TR-STBC matrix which encapsulates the effects of the MIMO channel as well as TR-STBC decoding over α time slots. Following the explanation of TR-STBC decoding in Chapter 2, $\check{\mathbf{H}}_k$ for $M = 2$ and $N_k = 1$ would be

$$\check{\mathbf{H}}_k = \begin{bmatrix} \mathbf{H}_{k,11} & \mathbf{H}_{k,12} \\ \mathbf{P}^0_{J+L}\mathbf{H}^*_{k,12}\mathbf{P}^0_J & -\mathbf{P}^0_{J+L}\mathbf{H}^*_{k,11}\mathbf{P}^0_J \end{bmatrix}, \tag{4.5}$$

where $\mathbf{P}_J$ and $\mathbf{P}_{J+L}$ perform the necessary time-reversal operations at the BS and the receiver, respectively. Note that (4.5) is equivalent to the channel matrix in (2.22). This matrix representation of the TR-STBC signal model that replaces the use of the operator q^{-1} with matrices was first introduced in [26]. The general form of $\check{\mathbf{H}}_k$ for $M > 4$ can be found in [25]. The signal at the output of the k-th user's TR-STBC decoder after α time slots is given by

$$\begin{aligned} \check{\mathbf{y}}_k &= [\mathbf{y}^{(1)}_{k,1}, ..., \mathbf{y}^{(1)}_{k,N_k}, ..., \mathbf{y}^{(\alpha)}_{k,1}, ..., \mathbf{y}^{(\alpha)}_{k,N_k}] \\ &= \check{\mathbf{H}}_k \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_M \end{bmatrix} + \tilde{\mathbf{w}}_k, \end{aligned} \tag{4.6}$$

where $\mathbf{y}_{k,i}^{(n)}$ is the signal received at the i-th antenna of the k-th user at the n-th time slot. Assuming the noise added at the terminal of the k-th user $\mathbf{w}_k$ to be white with samples of variance σ_w^2, the noise vector $\tilde{\mathbf{w}}_k$ produced from the TR-STBC decoder is still white with the same variance, as time-reversal processing does not introduce any correlation among the noise samples.

A key property of space-time codes is the ability to decouple the received symbols, or *streams* in the case of TR-STBC, by exploiting the orthogonality of the channel matrix at the receiver. The orthogonalization property of TR-STBC (2.23) was shown to hold when circulant matrices are used to express the TR-STBC signals [26]. In this work, however, we have expressed the signals using Toeplitz banded matrices. Using the following two properties:

$$\textbf{Property 1}\ [\mathbf{P}_{J+L}^{r}\mathbf{H}_{k,ij}\mathbf{P}_{J}^{r}]_{lm} = h_{k,ij}(L-l+m);$$

$$\textbf{Property 2}\ [\mathbf{P}_{J}^{r}\mathbf{H}_{k,ij}^{T}\mathbf{P}_{J+L}^{r}]_{lm} = h_{k,ij}(L+l-m),$$

which we prove in Appendix B, it can be easily shown that $\check{\mathbf{H}}_k$ is orthogonal, and the following holds:

$$\check{\mathbf{H}}_k^H\check{\mathbf{H}}_k = \text{blkdiag}(\bar{\mathbf{H}}_k, ..., \bar{\mathbf{H}}_k). \tag{4.7}$$

Here, $\bar{\mathbf{H}}_k$ is a $J \times J$ Hermitian and Toeplitz matrix whose first column is

$$[\gamma_{k,0}, \gamma_{k,1}, ..., \gamma_{k,L}, 0, ..., 0],$$

where

$$\gamma_{k,m} = \sum_{i=1}^{N_k}\sum_{j=1}^{M}\sum_{l=0}^{L} h_{k,ij}^{*}(l)h_{k,ij}(l+m), \quad \forall m \in \{0,1,...,L\},$$

$$h_{k,ij}(l) = 0\ \forall l \notin \{0,1,...,L\}. \tag{4.8}$$

The elements of $\{\gamma_{k,m}\}_{m=0}^{L}$ correspond to the coefficients of the double sided complex conjugate symmetric polynomial that was derived in [24]. The polynomial corresponding

to $M = 2$ and $N_k = 1$ was derived in (2.23). After matched filtering with $\check{\mathbf{H}}_k^H$, we therefore get

$$\mathbf{y}_k = \check{\mathbf{H}}_k^H \check{\mathbf{y}}_k = \begin{bmatrix} \bar{\mathbf{H}}_k \mathbf{x}_1 \\ \bar{\mathbf{H}}_k \mathbf{x}_2 \\ \vdots \\ \bar{\mathbf{H}}_k \mathbf{x}_M \end{bmatrix} + \bar{\mathbf{n}}_k. \tag{4.9}$$

The matrix $\bar{\mathbf{H}}_k$, however, represents a noncausal channel. Also, the noise vector becomes colored after matched filtering with covariance matrix $\sigma_w^2 \check{\mathbf{H}}_k^H \check{\mathbf{H}}_k$. Hence, a noise-whitening filter (NWF) should follow in order to yield an equivalent ISI channel that is causal and stable. The matched filter and the NWF together form a whitening matched filter (WMF) front-end. By defining the Cholesky decomposition $\bar{\mathbf{H}}_k^H \bar{\mathbf{H}}_k = \mathbf{L}_k \mathbf{L}_k^H$, where $\mathbf{L}_k$ is a unique non-singular lower-triangular matrix, we can filter the received vector $\mathbf{y}_k$ to get the noise-whitened model:

$$\mathbf{z}_k = (\mathbf{I} \otimes \mathbf{L}_k^{-1}) \mathbf{y}_k = \begin{bmatrix} \mathbf{L}_k^H \mathbf{x}_1 \\ \mathbf{L}_k^H \mathbf{x}_2 \\ \vdots \\ \mathbf{L}_k^H \mathbf{x}_M \end{bmatrix} + \mathbf{n}_k, \tag{4.10}$$

where the respective channel of each symbol block is now causal, and the samples of the noise vector $\mathbf{n}_k$ are independent with variance σ_w^2, because:

$$\begin{aligned} E\{\mathbf{n}_k \mathbf{n}_k^H\} &= E\{((\mathbf{I} \otimes \mathbf{L}_k^{-1}) \bar{\mathbf{n}}_k)((\mathbf{I} \otimes \mathbf{L}_k^{-1}) \bar{\mathbf{n}}_k)^H\} \\ &= \sigma_w^2 E\{(\mathbf{I} \otimes \mathbf{L}_k^{-1}) \check{\mathbf{H}}_k^H \check{\mathbf{H}}_k (\mathbf{I} \otimes \mathbf{L}_k^{-H})\} \\ &= \sigma_w^2 \mathbf{I}. \end{aligned} \tag{4.11}$$

By letting $\mathbf{V}_k = \text{diag}\{\frac{1}{\ell_{k,11}}, \frac{1}{\ell_{k,22}}, ..., \frac{1}{\ell_{k,JJ}}\}$, where $\ell_{k,ll} = [\mathbf{L}_k^H]_{ll}$, then $\mathbf{V}_k \mathbf{L}_k^H$ has a diagonal of all-ones. Also, let $\mathbf{z}_k = [\mathbf{z}_{k,1}^T, \mathbf{z}_{k,2}^T, ..., \mathbf{z}_{k,M}^T]^T$ and $\mathbf{n}_k = [\mathbf{n}_{k,1}^T, \mathbf{n}_{k,2}^T, ..., \mathbf{n}_{k,M}^T]^T$. Then, the

equivalent ISI channel of the symbol block intended for the k-th user becomes

$$\mathbf{z}_{k,k} = \mathbf{V}_k \mathbf{L}_k^H \mathbf{x}_k + \mathbf{n}_{k,k}. \tag{4.12}$$

From (4.3), we can write

$$\mathbf{z}_{k,k} = \mathbf{V}_k \mathbf{L}_k^{-1} \bar{\mathbf{H}}_k \mathbf{B}_k^{-1} (\mathbf{d}_k + \mathbf{b}_k) + \mathbf{n}_{k,k}. \tag{4.13}$$

For the ZF implementation of THP, we define

$$\mathbf{B}_k = (\mathbf{V}_k \mathbf{L}_k^H)^{-1}. \tag{4.14}$$

Therefore, (4.13) can be re-written as

$$\mathbf{z}_{k,k} = \mathbf{d}_k + \mathbf{b}_k + \mathbf{n}_{k,k}. \tag{4.15}$$

A modulo device removes $\mathbf{b}_k$ before a hard decision is made by a slicer followed by a decoder producing $\hat{\mathbf{u}}_k$, after discarding the last C symbols of $\mathbf{d}_k$.

4.2.2 STBC-OFDMA

In OFDMA, ISI is mitigated by the inherent multi-carrier structure where the frequency-selective channel is subdivided into parallel flat-fading subcarriers. Therefore, equalization is not required at the receiver. We employ STBC on each subcarrier in order to achieve spatial diversity. However, coding is required to exploit the frequency diversity – unlike TR-STBC which achieves spatial and frequency diversity without the need for coding.

At the transmitter, the $N \times 1$ block of coded symbols $\mathbf{u}_k$ is passed to a STBC encoder where the encoding is done according to a certain encoding rule which depends on the number of antennas employed. The STBC processing includes conjugating and negating the symbol blocks at certain time slots [47, 49]. An IFFT operation is performed on each

encoded subcarrier block and a CP is inserted. The encoded vectors are transmitted over the MIMO channel from the BS to the k-th user. At the receiver, the CP is removed from each block and a FFT operation follows. A STBC decoder performs the necessary processing over α time slots, where α depends on the rate of the underlying STBC. The insertion of a CP and its removal at the receiver makes the channel matrix from the j-th transmitter to the i-th receiver of the k-th user a $N \times N$ circulant matrix [50]; we denote this matrix by $\tilde{\mathbf{H}}_{k,ij}$ where:

$$[\tilde{\mathbf{H}}_{k,ij}]_{lm} = h_{k,ij}((l-m) \mod N). \tag{4.16}$$

Being a circulant matrix, $\tilde{\mathbf{H}}_{k,ij}$ has the following eigen-decomposition:

$$\tilde{\mathbf{H}}_{k,ij} = \mathbf{F}_N^H \mathbf{\Lambda}_{k,ij} \mathbf{F}_N, \tag{4.17}$$

where $\mathbf{\Lambda}_{k,ij}$ is a diagonal matrix whose (l,l)-th element is equal to the l-th DFT coefficient of the CIR vector $\mathbf{h}_{k,ij}$. By including the IFFT and FFT operations performed at the transmitter and the receiver, in addition to the CP concatenation and removal operations, in the definition of the equivalent MIMO channel from the BS to the k-th user, we can define a $NN_k \times NM$ matrix $\tilde{\mathbf{H}}_k$ with elements:

$$[\tilde{\mathbf{H}}_k]_{ij} = \mathbf{F}_N \tilde{\mathbf{H}}_{k,ij} \mathbf{F}_N^H = \mathbf{\Lambda}_{k,ij}. \tag{4.18}$$

In addition, define $\breve{\mathbf{H}}_k$ to be a $\alpha NN_k \times NM$ STBC matrix which encapsulates the effects of the MIMO channel as well as STBC decoding over α time slots – see [47] for examples. The decoded received signals are passed through a STBC matrix matched filter $\breve{\mathbf{H}}_k^H$. It can be easily shown that

$$\breve{\mathbf{H}}_k^H \breve{\mathbf{H}}_k = \text{blkdiag}(\mathbf{\Gamma}_k, ..., \mathbf{\Gamma}_k), \tag{4.19}$$

where

$$\mathbf{\Gamma}_k = \left\{ \sum_{i=1}^{N_k} \sum_{j=1}^{M} \mathbf{\Lambda}_{k,ij}^{*} \mathbf{\Lambda}_{k,ij} \right\}^{\frac{1}{2}}. \tag{4.20}$$

Hence, the $NM \times 1$ matched filter output $\mathbf{z}_k$ is:

$$\mathbf{z}_k = \begin{bmatrix} \mathbf{\Gamma}_k \mathbf{x}_1 \\ \mathbf{\Gamma}_k \mathbf{x}_2 \\ \vdots \\ \mathbf{\Gamma}_k \mathbf{x}_M \end{bmatrix} + \mathbf{n}_k, \tag{4.21}$$

where the samples of the noise vector $\mathbf{n}_k$ are independent with variance σ_w^2. As can be seen from (4.21), the subcarrier blocks are orthogonalized and the equivalent channel for each block is $\mathbf{\Gamma}_k$, which is a diagonal matrix. Hence, a block can be detected by simply filtering the vector of interest $\mathbf{z}_{k,k}$ by $\mathbf{\Gamma}_k^{-1}$ and making a hard decision on the resultant.

4.3 THP-TR-STBC Versus STBC-OFDMA

4.3.1 Complexity

Transmitter Complexity

The complexity of the THP-TR-STBC transmitter lies in the feedback filter of the THP block. In order to get an insight about the complexity, we revert to the q-transform representation of the system model. Using (2.18), we can re-write (4.9) as

$$y_k[n] = \underbrace{\sum_{i=1}^{N_k} \sum_{j=1}^{M} (h_{k,ij}^{*}(q) h_{k,ij}(q^{-1}))}_{\gamma_k(q^{-1})} x_k[n] + \bar{n}_k[n]. \tag{4.22}$$

Let the spectral factorization of the double sided complex conjugate polynomial, whose coefficients are defined in (4.8), be $\gamma_k(q^{-1}) = \eta^2 \mathcal{F}_k(q^{-1}) \mathcal{F}_k^{*}(q)$ where $\mathcal{F}_k(q^{-1})$ is minimum-phase, monic, and stable. The feedback filter is therefore given by $\mathcal{F}_k(q^{-1}) - 1$ [48, 51].

Then, we can re-write (4.3) in the q-domain in terms of the sequence $\{f_k[l]\}_{l=0}^{L}$ which corresponds to $\mathcal{F}_k(q^{-1})$ as

$$\begin{aligned} x_k[n] &= d_k[n] + b_k[n] - \sum_{l=0}^{L} f_k[l]x_k[n-l] - x_k[n] \\ &= d_k[n] + b_k[n] - \underbrace{\sum_{l=1}^{L} f_k[l]x_k[n-l]}_{\mathcal{I}_k[n]}, \end{aligned} \tag{4.23}$$

where $n \in \{0, 1, ..., J-1\}$. Since the length of the interference sequence $\mathcal{I}_k[n]$ is L, we conclude that L multiplications and $L-1$ additions are performed per information symbol by the feedback filter. Therefore, the overall complexity of the transmitter in THP-TR-STBC is $\mathcal{O}(L)$[1]. For STBC-OFDMA, the N-IFFT block at the transmitter incurs a complexity of $\mathcal{O}(\log_2 N)$ per information symbol.

Receiver Complexity

The receiver of each user of THP-TR-STBC contains a NWF which is given by the noncausal filter $\frac{1}{\eta^2 \mathcal{F}_k^*(q)}$ [51]. A practical implementation would truncate the ideal impulse response of the NWF to an acceptable number of samples. In a ZF implementation of THP, the number of taps in the feedforward filter, the NWF, is equal to the channel order L. Hence, the NWF accounts for $L+1$ multiplications and L additions per information symbol, and the complexity of the THP-TR-STBC receiver is $\mathcal{O}(L)$ per information symbol. For STBC-OFDMA, The receiver of each user contains a N-FFT block which has a complexity of $\mathcal{O}(\log_2 N)$ per information symbol.

The overall complexity per information symbol of THP-TR-STBC is therefore $\mathcal{O}(2L)$, and that of STBC-OFDMA is $\mathcal{O}(2\log_2 N)$. We did not consider the complexity incurred by matched filtering at the receiver, because it is a common operation for the receivers

[1]Multiplication is known to be more complex than addition; thus, we base the complexity analysis on the number of multiplication operations.

of the two systems. Also, STBC processing in STBC-OFDMA and TR-STBC processing in THP-TR-STBC include operations such as conjugating and time-reversing which do not account for any significant complexity and were neglected in our complexity analysis. The complexity of STBC-OFDMA depends on N and will incur prohibitive complexity in a system employing a large number of subcarriers. In particular, the complexity of STBC-OFDMA would exceed that of THP-TR-STBC when

$$\begin{aligned} 2\log_2 N &> 2L \\ N &> 2^L \end{aligned} \tag{4.24}$$

For example, THP-TR-STBC would be less complex than an OFDMA-based system employing 2048 subcarriers over frequency-selective channels with $L < 11$. However, the complexity of THP-TR-STBC would be greater than that of an STBC-OFDMA system employing a small number of subcarriers over channels with very large orders. It is worth noting that the complexity of THP-TR-STBC can be made similar to STBC-OFDMA by using frequency-domain equalization (FDE) [52]. This can be obtained by replacing the THP block by an equalizer preceded by a N-FFT block. The use of FDE makes the complexity of THP-TR-STBC independent of L leading to computational savings.

In summary, the relative complexity of THP-TR-STBC and STBC-OFDMA depends on the relative values of N and L, and which system should be utilized depends on the specific application.

4.3.2 BER Performance

We now compare the BER performance of THP-TR-STBC and STBC-OFDMA. In all the following simulations, we let $L = C$, $N = 128$, and $N_k = 1$. The modulation adopted is quadrature phase-shift keying (QPSK). The channel taps $\{h_{k,ij}(l)\}_{l=0}^{L}$ are assumed to be independent ZMCSCG with unit variance, or $E\{\mathbf{h}_{k,ij}\mathbf{h}_{k,ij}^H\} = \frac{1}{L+1}\mathbf{I}_{L+1}$. In all the figures presented, we plot the average BER versus SNR which we define as the energy-

per-bit to the noise power spectral density or E_b/N_o in dB, where E_b is the energy per information bit and $N_o = \sigma_w^2$. The curves represent the average performance of the K users.

We first compare the performance of THP-TR-STBC and STBC-OFDMA without using coding. We let $L = 1$, $M = 2$, and $K = 2$. For the single-carrier system, we use a rate 1/2 TR-STBC; a space-time code of the same rate is used in STBC-OFDMA. Fig. 4.3 depicts this comparison. We observe that THP-TR-STBC outperforms STBC-OFDMA. This is because TR-STBC exploits joint spatial and frequency diversity, whereas STBC-OFDMA achieves spatial diversity only.

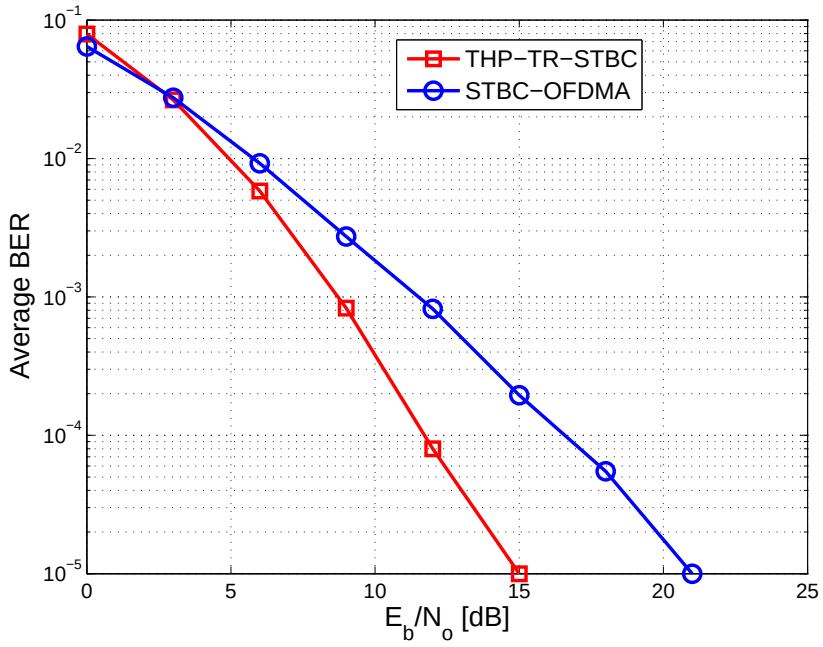

Figure 4.3: Comparison between THP-TR-STBC and STBC-OFDMA when no coding is used. $M = 2$, $K = 2$, and $L = 1$.

In Fig. 4.4 we simulate the same systems, but with the use of a convolutional code of rate 1/2 to allow STBC-OFDMA to achieve frequency diversity. We observe that both systems have comparable performance, and STBC-OFDMA outperforms THP-TR-STBC by 1/2 dB only for a bit error rate of 10^{-3}.

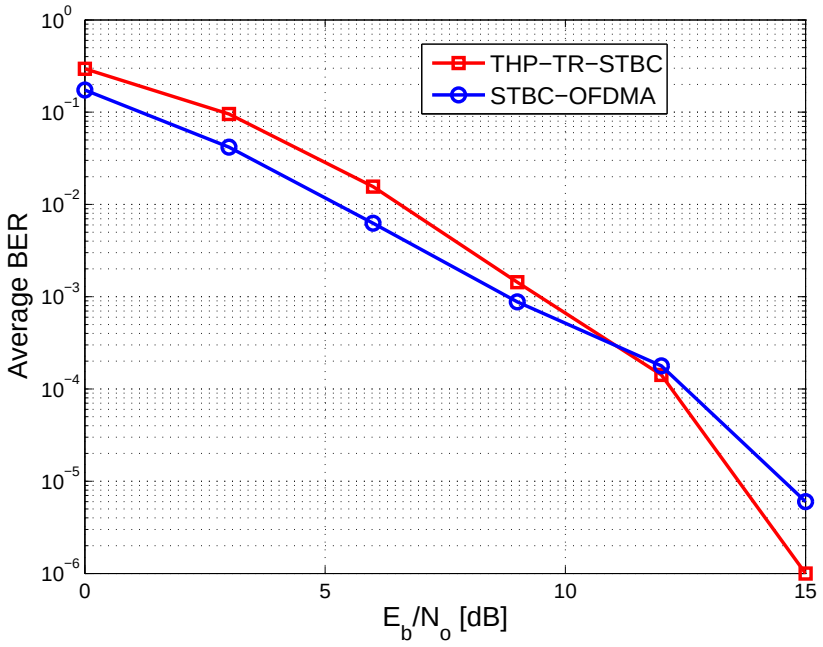

Figure 4.4: Comparison between THP-TR-STBC and STBC-OFDMA when a convolutional code of rate 1/2 is employed in both systems. $M = 2$, $K = 2$, and $L = 1$.

Fig. 4.5 shows the comparison between THP-TR-STBC and STBC-OFDMA for $L = 2$, $M = 4$, and $K = 3$. The rate of the space-time codes used is 3/4. For both systems, we employ a convolutional code of rate 1/2. We note that THP-TR-STBC outperforms STBC-OFDMA, and this is due to fully exploiting the frequency diversity of the channel, which STBC-OFDMA might achieve with the help of a more powerful error control code or frequency interleaving techniques. We have used random frequency interleaving, but we did not observe noticeable improvements in the performance of STBC-OFDMA. This was also observed in [29] in a point-to-point system. Practical OFDMA systems employ powerful codes, such as Turbo coding, and AMC across subcarriers; hence, we expect both systems to exhibit the same performance in practice.

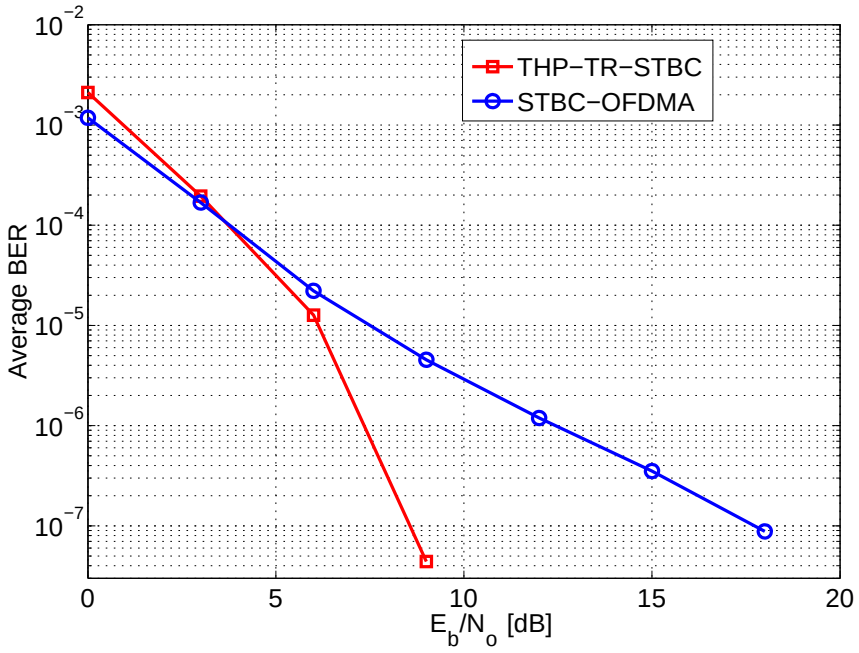

Figure 4.5: Comparison between THP-TR-STBC and STBC-OFDMA when a convolutional code of rate 1/2 is employed in both systems. $M = 4$, $K = 3$, and $L = 2$.

4.3.3 Data Rate and Number of Users Supported

The availability of CSI at the BS enables MIMO-OFDMA systems to implement powerful transmission schemes. For example, combining linear and non-linear precoding techniques with MIMO-OFDMA allows for achieving high data rates. However, in this chapter we are trying to compare the performance of MIMO-OFDMA in mitigating ISI with a possible single-carrier alternative whose structure depends on space-time processing. The main objective of using space-time processing is to achieve high diversity orders to provide improved reliability and not to maximize the data rate. In fact, the rate of space-time codes drops when the number of transmit antennas increase. Therefore, we have considered STBC-OFDMA which is a suitable contender to THP-TR-STBC.

Due to its multi-carrier structure, STBC-OFDMA is capable of supporting more users than THP-TR-STBC, allowing it to provide a higher sum-rate, whereas the number of users supported by THP-TR-STBC is limited by the number of transmit antennas.

Nonetheless, THP-TR-STBC remains an attractive solution in networks with a limited number of users, such as a wireless backhaul systems, as it provides comparable performance to STBC-OFDMA while providing the benefits of single-carrier transmission.

4.4 Summary

We have investigated the possible use of single-carrier transmission in the downlink of MIMO-BC in fixed broadband wireless systems with a small number of users ($K \leq M$) as an alternative to MIMO-OFDMA. We made novel use of TR-STBC in multi-user systems to multiplex data streams to different users over frequency-selective channels. We pre-equalized the channels using ZF-THP assuming full CSI is available at the BS. We compared THP-TR-STBC and STBC-OFDMA based on complexity and performance. THP-TR-STBC systems have lower complexity in systems employing a large number of subcarriers over channels with low order. They also outperform STBC-OFDMA when no error control codes are employed, and exhibit comparable performance to coded STBC-OFDMA. In addition, THP-TR-STBC enjoys all the inherent advantages of the single-carrier structure such as low PAPR and insensitivity to spectral nulls.

Chapter 5

Conclusion and Future Work

5.1 Summary of Contributions

In this thesis, we have tackled two classes of precoding problems over multi-user MIMO-BC: beamforming design for WSR maximization in OFDMA systems and THP design for ISI suppression in single-carrier systems.

In Chapter 3, we have extended the WSRBF-WMMSE algorithm [16, 17] to MIMO-OFDMA systems by extending the definition of the beamforming and channel matrices to include both the spatial and frequency dimensions. We showed that the complexity of the proposed algorithm scales only linearly with the number of subcarriers compared to that of WSRBF-WMMSE. The proposed algorithm relies entirely on closed-form expressions and does not require nested optimization or GP solvers. The algorithm performs joint optimization across all subcarriers and is capable of performing subcarrier allocation automatically. We have shown that WSRBF-WMMSE-OFDMA converges to a local optimum of the WSR maximization problem and exhibits near-optimal performance. We have also applied the clustering technique to further reduce the complexity of WSRBF-WMMSE-OFDMA where we design a single beamformer for a cluster of adjacent subcarriers. We have demonstrated the trade-off between the size of the cluster

and the WSR performance through simulations.

We have also devised a generally applicable stream rate allocation algorithm that assigns practical modulation and coding schemes to each data stream of each user on each subcarrier. The method translates beamformers designed by WSR maximizing algorithms, such as WSRBF-WMMSE-OFDMA, to practical AMC modes for a given desired bit error probability. The algorithm utilizes the SNR gap to capacity concept which we have extended to include the effect of the decoding order when a SIC receiver is employed at each downlink user. The proposed method is particularly useful for the case when the users have more than one antenna each and receive up to N_k data streams. In this case, the channel from the base station to the k-th receiver will effectively be a MAC, in which a polymatroid of rate vectors are achievable. The method then maps available AMC modes to the space of allowed theoretical rates, using the SNR gap to capacity concept, and selects the operating point with the largest sum-rate. We have also described a method to decrease the complexity of the algorithm by reducing the search space using a bisection search. Simulation results show that the proposed method traces the theoretical WSR achieved by WSRBF-WMMSE-OFDMA.

In Chapter 4, we described a potential single-carrier alternative to multi-carrier communications for fixed wireless networks with a limited number of users $K \leq M$. The proposed scheme exploits the ability of TR-STBC to orthogonalize data streams over frequency-selective channels to orthogonalize downlink users. We have utilized ZF-THP to perform the required pre-equalization. In comparison to OFDMA-based transmission, THP-TR-STBC performs better in terms of BER when coding is not employed, due to achieving full spatial-multipath diversity, and exhibits similar performance when coding is used. Besides the inherited single-carrier advantages over OFDMA, such as having low PAPR, the complexity of THP-TR-STBC is comparable to that of OFDMA-based systems when the number of subcarriers employed is large and the order of the frequency-selective channels is low.

5.2 Future Work

In this research, we proposed precoding methods that can lead to interesting applications. The following are possible future directions and extensions to the work we presented:

- Throughout this thesis we have assumed the availability of full CSI at the BS. This assumption might not be possible for certain applications. Therefore, it would be instructive to analyze the effect of imperfect CSI on the performance of the proposed algorithms.

- We have considered two possible approaches when extending WSRBF-WMMSE to OFDMA systems: optimizing jointly or independently across subcarriers. A third possibility would be an alternating optimization approach that alternates between power allocation across subcarriers and the design of unit norm beamformers. The complexity and performance of this approach can then be compared to WSRBF-WMMSE-OFDMA.

- The practical rate assigned to a given data stream by the stream rate allocation algorithm might be lower than the theoretical rate. This translates to that stream not using all its available power. Hence, one can design an extension to our algorithm which collects the spare power and use it to support users with low channel gains.

- The stream rate allocation algorithm is targeted at practical systems. The implementation cost of this approach using digital logic can be of interest for future research.

- Using FDE in conjunction with TR-STBC is an attractive alternative to THP as it enables single-carrier schemes to have comparable complexity to OFDMA-based systems.

Appendix A

Details of WSRBF-WMMSE

This appendix provides the derivation of the gradients of the WSR maximization and WMMSE minimization problems as well as the convergence analysis of the WSRBF-WMMSE algorithm. These details are provided in [16, 17] and are shown here for completeness.

A.1 Gradients of the WSR and WMMSE Problems

The WSR problem gradient (2.12) is derived by dividing the summation in (2.10) into multiple parts and applying the chain rule on each part. The (l, m)-th element of the gradient matrix is given by $[\nabla_{\mathbf{B}_k}\mathcal{L}]_{lm} = \nabla_{[\mathbf{B}_k]lm}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial [\mathbf{B}_k^*]_{lm}}$. Then, we can write

$$\begin{aligned}
\nabla_{[\mathbf{B}_k]lm}\tilde{R}_k &= \mathrm{Tr}\left(\left(\frac{\partial \tilde{R}_k}{\partial \mathbf{E}_k^{-1}}\right)^T \left(\frac{\partial \mathbf{E}_k^{-1}}{\partial [\mathbf{B}_k^*]lm}\right)\right) \\
&= \mathrm{Tr}(\mathbf{E}_k \mathbf{e}_m \mathbf{e}_l^T \mathbf{H}_k^H \mathbf{C}_k^{-1} \mathbf{H}_k \mathbf{B}_k) \\
&= \mathbf{e}_l^T \mathbf{H}_k^H \mathbf{C}_k^{-1} \mathbf{H}_k \mathbf{B}_k \mathbf{E}_k \mathbf{e}_m, \qquad \text{(A.1)}
\end{aligned}$$

where $\mathbf{e}_l$ is a unit-vector with the l-th element being one and zeros elsewhere. Because $\nabla_{[\mathbf{B}_k]_{lm}}\tilde{R}_k = [\nabla_{\mathbf{B}_k}\tilde{R}_k]_{lm}$, we conclude that

$$\nabla_{\mathbf{B}_k}\tilde{R}_k = \mathbf{H}_k^H \mathbf{C}_k^{-1} \mathbf{H}_k \mathbf{B}_k \mathbf{E}_k. \qquad \text{(A.2)}$$

To compute $\nabla_{\mathbf{B}_k}\tilde{R}_i$ for $i \neq k$, define the real-valued scalar function h which depends on a matrix $\mathbf{K}$ through $\mathbf{S} = \mathbf{X} + \mathbf{L}\mathbf{K}\mathbf{C}\mathbf{C}^H\mathbf{K}^H\mathbf{L}^H$, where $\mathbf{X}$, $\mathbf{L}$, and $\mathbf{C}$ are fixed matrices independent of $\mathbf{K}$. It can be shown that $\nabla_{\mathbf{K}}h = \mathbf{L}^H(\nabla_{\mathbf{S}}h)\mathbf{L}\mathbf{K}\mathbf{C}\mathbf{C}^H$. By considering the noise covariance $\mathbf{C}_i$ to be the matrix $\mathbf{S}$, we get

$$\nabla_{\mathbf{B}_k}\tilde{R}_i = \mathbf{H}_i^H(\nabla_{\mathbf{C}_i}\tilde{R}_i)\mathbf{H}_i\mathbf{B}_k. \tag{A.3}$$

By taking similar steps to the derivation of (A.1), we can write

$$\nabla_{\mathbf{C}_i}\tilde{R}_i = -\mathbf{C}_i^{-1}\mathbf{H}_i\mathbf{B}_i\mathbf{E}_i\mathbf{B}_i^H\mathbf{H}_i^H\mathbf{C}_i^{-1}. \tag{A.4}$$

Combining (A.3) and (A.4), we obtain

$$\nabla_{\mathbf{B}_k}\tilde{R}_i = -\mathbf{H}_i^H\mathbf{C}_i^{-1}\mathbf{H}_i\mathbf{B}_i\mathbf{E}_i\mathbf{B}_i^H\mathbf{H}_i^H\mathbf{C}_i^{-1}\mathbf{H}_i\mathbf{B}_k. \tag{A.5}$$

Finally, combining (A.2) and (A.5) yields the gradient of the WSR problem (2.12).

The gradient of the WMMSE problem is computed in a similar manner. We first compute

$$\nabla_{\mathbf{B}_k}\mathrm{Tr}(\mathbf{W}_k\mathbf{E}_k) = -\mathbf{H}_k^H\mathbf{C}_k^{-1}\mathbf{H}_k\mathbf{B}_k\mathbf{E}_k\mathbf{W}_k\mathbf{E}_k. \tag{A.6}$$

Then, we compute the term

$$\nabla_{\mathbf{B}_k}\mathrm{Tr}(\mathbf{W}_i\mathbf{E}_i) = \mathbf{H}_i^H\mathbf{C}_i^{-1}\mathbf{H}_i\mathbf{B}_i\mathbf{E}_i\mathbf{W}_i\mathbf{E}_i\mathbf{B}_i^H\mathbf{H}_i^H\mathbf{C}_i^{-1}\mathbf{H}_i\mathbf{B}_k. \tag{A.7}$$

By combining (A.6) and (A.7), we obtain the gradient of the WMMSE problem (2.13).

A.2 Convergence Analysis

Convergence of WSRBF-WMMSE is shown by proving the monotonic convergence of an equivalent problem. The equivalent problem is formed by modifying the objective function of the WSR problem $-\mu_k \log_2 \det(\mathbf{E}_k^{-1})$ to include the MSE weights and the receive beamformers as optimization variables, in addition to the transmit beamformers.

Define the MSE matrix

$$\tilde{\mathbf{E}} = E[(\mathbf{A}_k\mathbf{y}_k - \mathbf{d}_k)(\mathbf{A}_k\mathbf{y}_k - \mathbf{d}_k)^H], \tag{A.8}$$

which is different than the matrix $\mathbf{E}_k$ defined in (2.7). Consider the modified optimization problem

$$\begin{aligned}[\mathbf{B}_1^*, \mathbf{B}_2^*, ..., \mathbf{B}_K^*] &= \underset{\mathbf{B}_k, \mathbf{A}_k, \mathbf{W}_K\ \forall k}{\arg\min} \sum_{k=1}^{K} \tilde{l}_k(\mathbf{W}_k, \mathbf{A}_k, \mathbf{B}_i\ \forall i) \\ &\text{s.t.} \quad \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{B}_k\mathbf{B}_k^H) \leq P_{max},\end{aligned} \tag{A.9}$$

where the cost function is

$$\tilde{l}_k(\mathbf{W}_k, \mathbf{A}_k, \mathbf{B}_i\ \forall i) = \mathrm{Tr}(\mathbf{W}_k\tilde{\mathbf{E}}_k) - \mu_k \log_2 \det(\mu_k^{-1}\mathbf{W}_k) - \mu_k N_k. \tag{A.10}$$

We first show that optimizing the transmit filters $\mathbf{B}_k$ using (A.9) is equivalent to optimizing using the original WSR problem (2.3). By fixing the transmit filters and the MSE weights, the minimizing value of (A.10) with respect to $\mathbf{A}_k$ is unique and is denoted by $\mathbf{A}_k^{\mathrm{MMSE}}(\mathbf{B}_i\ \forall i)$, which is the same as (2.6). The variable $\mathbf{A}_k$ intervenes only in $\tilde{\mathbf{E}}_k$ and substituting $\mathbf{A}_k^{\mathrm{MMSE}}(\mathbf{B}_i\ \forall i)$ in (A.8) we get $\mathbf{E}_k$. We can therefore define a new cost function involving the transmit beamformers and the MSE weights only:

$$\tilde{l}_k(\mathbf{W}_k, \mathbf{B}_i\ \forall i) = \mathrm{Tr}(\mathbf{W}_k\mathbf{E}_k) - \mu_k \log_2 \det(\mu_k^{-1}\mathbf{W}_k) - \mu_k N_k. \tag{A.11}$$

Similarly, minimizing (A.11) with respect to $\mathbf{W}_k$ gives $\mathbf{W}_k^{\min} = \mu_k\mathbf{E}_k^{-1}(\mathbf{B}_i\ \forall i)$ which when substituted in (A.11) yields the original WSR objective function $-\mu_k \log_2 \det(\mathbf{E}_k^{-1})$.

Having shown the equivalence between the original WSR problem and optimizing the cost function (A.10), it remains to show that alternating minimization of (A.10) corresponds to steps 5 to 7 of the WSRBF-WMMSE algorithm which is shown in Table 2.1. Fixing $\mathbf{W}_k\ \forall k$, the cost function becomes $\mathrm{Tr}(\mathbf{W}_k\tilde{\mathbf{E}}_k(\mathbf{A}_k, \mathbf{B}_i\ \forall i))$. Then, finding $\mathbf{A}_k$ given $\mathbf{B}_i^{t-1}\ \forall i$ gives the same result as in step 5, and finding $\mathbf{B}_k\ \forall k$ given $\mathbf{A}_k^t\ \forall k$ and $\mathbf{W}_k^t\ \forall k$ gives the same result as step 7. Finally, fixing $\mathbf{A}_k$ and $\mathbf{B}_i\ \forall i$ and optimizing with

respect to $\mathbf{W}_k$ yields $\mathbf{W}_k = \mu_k \tilde{\mathbf{E}}_k^{-1}(\mathbf{A}_k^t, \mathbf{B}_i^{t-1}\ \forall i) = \mu_k \mathbf{E}^{-1}(\mathbf{B}_i^{t-1}\ \forall i)$ which is the same result as step 6.

The cost in (A.9) decreases monotonically as a result of the alternating minimization process. Assuming a minimal value of the cost function exists, the cost function is lower bounded and therefore the algorithm converges to a local optimum. It is important to note that the original WSR objective function does not necessarily experience a monotonic convergence; however, simulation results show that this is often the case.

Appendix B

Proofs of Property 1 and Property 2

B.1 Property 1

The (l, m)-th element of a $(J + L) \times J$ Toeplitz banded channel matrix is given by

$$[\mathbf{H}_{k,ij}]_{lm} = h_{k,ij}(l - m), \tag{B.1}$$

where $l \in \{1, 2, .., J + L\}$ and $m \in \{1, 2, ..., J\}$. Pre-multiplying the channel matrix $\mathbf{H}_{k,ij}$ by the permutation matrix $\mathbf{P}^n_{J+L}$ performs a reverse cyclic row shit, or, equivalently, moves the $(J + L - (l - 1) + n)$-th row to l-th row:

$$\begin{aligned} [\mathbf{P}^n_{J+L}\mathbf{H}_{k,ij}]_{lm} &= h_{k,ij}((J + L - (l - 1) + n - 1) - (m - 1)) \\ &= h_{k,ij}(J + L - l + m + n - 1). \end{aligned} \tag{B.2}$$

Similarly, post-multiplying $\mathbf{P}^n_{J+L}\mathbf{H}_{k,ij}$ by $\mathbf{P}^n_J$ performs a reverse cyclic column shift, or, equivalently, moves the $(J - (m - 1) + n)$-th column to the m-th column:

$$\begin{aligned} [\mathbf{P}^n_{J+L}\mathbf{H}_{k,ij}\mathbf{P}^n_J]_{lm} &= h_{k,ij}((J + L - (l - 1) + n - 1) - (J - (m - 1) + n - 1)) \\ &= h_{k,ij}(L - l + m), \end{aligned} \tag{B.3}$$

which proves Property 1.

B.2 Property 2

The proof of Property 2 is similar to that of Property 1. The (l, m)-th element of $\mathbf{H}_{k,ij}^T$ is given by

$$[\mathbf{H}_{k,ij}^T]_{lm} = h_{k,ij}(m - l), \tag{B.4}$$

where $l \in \{1, 2, .., J\}$ and $m \in \{1, 2, ..., J + L\}$. Pre-multiplying the channel matrix $\mathbf{H}_{k,ij}^T$ by the permutation matrix $\mathbf{P}_J^n$ performs a reverse cyclic row shit, or, equivalently, moves the $(J - (l - 1) + n)$-th row to l-th row:

$$\begin{aligned} [\mathbf{P}_J^n \mathbf{H}_{k,ij}^T]_{lm} &= h_{k,ij}((m-1) - (J - (l-1) + n - 1)) \\ &= h_{k,ij}(-J + l + m + n - 1). \end{aligned} \tag{B.5}$$

Similarly, post-multiplying $\mathbf{P}_J^n \mathbf{H}_{k,ij}^T$ by $\mathbf{P}_{J+L}^n$ performs a reverse cyclic column shift, or, equivalently, moves the $(J + L - (m - 1) + n)$-th column to the m-th column:

$$\begin{aligned} [\mathbf{P}_J^n \mathbf{H}_{k,ij}^T \mathbf{P}_{J+L}^n]_{lm} &= h_{k,ij}((J + L - (m-1) + n - 1) - (J - (l-1) + n - 1)) \\ &= h_{k,ij}(L + l - m), \end{aligned} \tag{B.6}$$

which proves Property 2.

Bibliography

[1] K. Burrows and L. Fielbrandt, "Overcoming the limitations of todays fixed wired access technologies," *Siemens Information and Communication Networks Solution Management Access*, White Paper, 2004. Available at: `http://www.docstoc.com/docs/827413/siemens_WiMAX_whitepaper0304`.

[2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451-1458, October 1998.

[3] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, pp. 41-59, Autumn 1996.

[4] L. Zheng and D. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Channels", *IEEE Transactions on Information Theory*, 2002

[5] K. Etemad, "Overview of mobile WiMAX technology and evolution," *IEEE Communications Magazine*, vol.46, no.10, pp. 31-40, October 2008.

[6] G. L. Stuber, J. R. Barry, S. W. McLaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proceedings of the IEEE*, vol. 92, no.2, pp. 271-294, February 2004.

[7] Y. Li, J. H. Winters, and N. R. Sollenberger, "MIMO-OFDM for wireless communications: signal detection with enhanced channel estimation," *IEEE Transactions on Communication*s, vol. 50, no. 9, pp. 1471-1477, September 2002.

[8] 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, September 2006.

[9] M. Tomlinson, "New automatic equalizer employing modulo arithmetic," *Electronic Letters*, vol. 7, pp. 138-139, March 1971.

[10] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Transactions on Communications*, vol. 20, pp. 774-780, August 1972.

[11] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646-2660, June 2007.

[12] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528-541, March 2006.

[13] S. Shi, M. Schubert, and H. Boche, "Rate optimization for multiuser MIMO systems with linear processing," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 4020-4030, August 2008.

[14] M. Codreanu, A. Tolli, M. Juntti, and M. Latva-aho, "MIMO downlink weighted sum rate maximization with power constraint per antenna Groups," *Proceedings of 65th Vehicular Technology Conference (VTC 2007-Spring)*, pp. 2048-2052, April 2007.

[15] A. J. Tenenbaum and R. S. Adve, "Linear processing and sum throughput in the multiuser MIMO downlink," *IEEE Transactions on Wireless Communications*, vol. 8, no. 5, pp. 2652-2661, May 2009.

[16] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Transactions Wireless Communications*, vol. 7, no. 12, pp. 4792-4799, December 2008.

[17] ———, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 1-6, June 2009.

[18] G. Zheng, K. Wong, and T. Ng, "Throughput maximization in linear multiuser MIMO-OFDM downlink systems," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 3, pp. 1993-1998, May 2008.

[19] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE transceiver optimization for multiuser MIMO systems: Duality and sum-MSE minimization," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5436-5446, November 2007.

[20] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 161-176, January 2006.

[21] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," *Stanford University EE Technical Report*, 2005. Available at `http://www.stanford.edu/~boyd/papers/pdf/gp_tutorial.pdf`.

[22] P. T. Boggs and J. W. Tolle, "Sequential quadratic programming," *Acta Numerica*, Cambridge University Press, Cambridge, UK, pp. 1-51, 1995.

[23] D. Palomar and S. Verdu, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 141-154, January 2006.

[24] E. Lindskog and A. Paulraj, "A transmit diversity scheme for channels with intersymbol interference," in *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 307-311, June 2000.

[25] P. Stoica and E. Lindskog, "Space-time block coding for channels with intersymbol interference," *Conference Record of the Thirty-Fifth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 252-256, October 2001.

[26] S. Zhou and G. B. Giannakis, "Single-carrier spacetime block-coded transmissions over frequency-selective fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 1, pp. 164-179, January 2003.

[27] N. Al-Dhahir, "Single-carrier frequency-domain equalization for space-time block-coded transmissions over frequency-selective fading channels," *IEEE Communications Letters*, vol. 5, no. 7, pp. 304-306, July 2001.

[28] N. Al-Dhahir, "Overview and comparison of equalization schemes for space-time-coded signals with application to EDGE," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2477-2488, October 2002.

[29] H. Mheidat, M. Uysal, and N. Al-Dhahir, "Comparative analysis of equalization techniques for STBC with application to EDGE," *Proceedings of 59th Vehicular Technology Conference (VTC 2004-Spring)*, vol. 1, pp. 555-559, May 2004.

[30] Y. Zhu and K. B. Letaief, "Single-carrier frequency-domain equalization with decision-feedback processing for time-reversal spacetime block-coded systems," *IEEE Transactions on Communications*, vol. 53, no. 7, pp. 1127-1131, July 2005.

[31] D. Flore and E. Lindskog, "Time reversal space-time block coding vs. transmit delay diversity – A comparison based on a GSM-like system," in *Proceedings of the 9th Digital Signal Processing Workshop*, Hunt, TX, 2000.

[32] E. G. Larsson, P. Stoica, E. Lindskog, and J. Li, "Space-time block coding for frequency-selective channels," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, pp. III-2405 - III-2408, May 2002.

[33] H. Mheidat, M. Uysal, and N. Al-Dhahir, "Time-reversal space-time equalization for amplify-and-forward relaying," *Proceedings of the IEEE International Conference on Communications (ICC)*, vol.4, pp.1705-1711, June 2006.

[34] ———, "Equalization techniques for distributed space-time block codes with amplify-and-forward relaying," *IEEE Transactions on Signal Processing*, vol.55, no.5, pp. 1839-1852, May 2007.

[35] C. Jonietz, W.H. Gerstacker and R. Schober, "Robust transmit processing for frequency-selective fading channels with imperfect channel feedback," *IEEE Transactions on Wireless Communications*, vol.7, no.12, pp. 5356-5368, December 2008.

[36] ———, "Combined Time-Reversal Space-Time Block coding and Transmit Beamforming for Frequency-Selective Fading Channels," *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 3955-3960, May 2008.

[37] N. Jindal, W. Rhee, S. Vishwanath, S.A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1570-1580, April 2005.

[38] S. Colieri, M. Ergen, A. Puri, and A. Bahai, "A study of channel estimation in OFDM systems," *Proceedings of 56th Vehicular Technology Conference (VTC 2002-Fall)*, vol. 2, pp. 894-898, September 2002.

[39] K. Wong, R. 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, January 2001.

[40] J. Choi and R. W. Heath Jr., "Interpolation based unitary precoding for spatial multiplexing for MIMO-OFDM with limited feedback," *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, pp. 214-218, December 2004.

[41] H. Karaa, A. Tenenbaum, and R. S. Adve, "Linear precoding for multiuser MIMO-OFDM systems," *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 2797-2802, June 2007.

[42] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, Cambridge, UK, May 2005.

[43] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York, NY, USA, 2nd ed., 2006.

[44] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, Kluwer Academic, Norwell, MA, USA, 3rd ed., 2004.

[45] C. F. Fung, W. Yu, and T. J. Lim, "Precoding for the multi-antenna downlink: multiuser SNR gap and optimal user ordering," *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 188-197, January 2007.

[46] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 5th ed., 2008.

[47] V. Tarokh, H. Jafarkhani and A.R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456-1467, July 1999.

[48] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Transmission*, John Wiley & Sons, Inc., New York, NY, 2002.

[49] H. Jaafarkhani, *Space-Time Coding: Theory and Practice*, Cambridge University Press, 2005.

[50] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29-48, May 2000.

[51] T. J. Lim, *Lecture Notes of ECE 1521S: Statistical Signal Processing*, University of Toronto, 2009.

[52] D. Falconer, S. L. Ariyarisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems, *IEEE Communications Magazine*, vol. 40, no. 4, pp. 58-66, April 2002.