NEAR CAPACITY OPERATING PRACTICAL TRANSCEIVERS FOR WIRELESS FADING CHANNELS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

GÖKHAN MUZAFFER GÜVENSEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

FEBRUARY 2009

Approval of the thesis:

NEAR CAPACITY OPERATING PRACTICAL TRANSCEIVERS FOR WIRELESS FADING CHANNELS

submitted by **GÖKHAN MUZAFFER GÜVENSEN** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. İsmet Erkmen
Head of Department, **Electrical and Electronics Engineering** _____

Assoc. Prof. Dr. Ali Özgür Yılmaz
Supervisor, **Electrical and Electronics Engineering Dept., METU** _____

Examining Committee Members:

Prof. Dr. Yalçın Tanık
Electrical and Electronics Engineering Dept., METU _____

Prof. Dr. Mete Severcan
Electrical and Electronics Engineering Dept., METU _____

Assoc. Prof. Elif Uysal Bıyıkoğlu
Electrical and Electronics Engineering Dept., METU _____

Assoc. Prof. Dr. Ali Özgür Yılmaz
Electrical and Electronics Engineering Dept., METU _____

Assist. Prof. Dr. Defne Aktaş
Electrical and Electronics Engineering Dept., Bilkent University _____

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: GÖKHAN MUZAFFER GÜVENSEN

Signature :

ABSTRACT

NEAR CAPACITY OPERATING PRACTICAL TRANSCEIVERS FOR WIRELESS FADING CHANNELS

Güvensen, Gökhan Muzaffer
M.S., Department of Electrical and Electronics Engineering
Supervisor : Assoc. Prof. Dr. Ali Özgür Yılmaz

February 2009, 90 pages

Multiple-input multiple-output (MIMO) systems have received much attention due to their multiplexing and diversity capabilities. It is possible to obtain remarkable improvement in spectral efficiency for wireless systems by using MIMO based schemes. However, sophisticated equalization and decoding structures are required for reliable communication at high rates. In this thesis, capacity achieving practical transceiver structures are proposed for MIMO wireless channels depending on the availability of channel state information at the transmitter (CSIT).

First, an adaptive MIMO scheme based on the use of quantized CSIT and reduced precoding idea is proposed. With the help of a very tight analytical upper bound obtained for limited rate feedback (LRF) MIMO capacity, it is possible to construct an adaptive scheme varying the number of beamformers used according to the average SNR value. It is shown that this strategy always results in a significantly higher achievable rate than that of the schemes which does not use CSIT, if the number of transmit antennas is greater than that of receive antennas.

Secondly, it is known that the use of CSIT does not bring significant improvement over capacity, when similar number of transmit and receive antennas are used; on the other hand, it reduces the complexity of demodulation at the receiver by converting the channel into non-interfering subchannels. However, it is shown in this thesis that it is still possible to achieve a performance very close to the outage probability and exploit the space-frequency diversity benefits of the wireless fading channel without compromising the receiver complexity, even if the CSIT is not used. The proposed receiver structure is based on iterative forward and backward filtering to suppress the interference both in time and space followed by a space-time decoder. The rotation of multidimensional constellations for block fading channels and the single-carrier frequency domain equalization (SC-FDE) technique for wideband MIMO channels are studied as example applications.

Keywords: Limited rate feedback, MIMO channel, diversity, iterative decision feedback equalization, block fading channel

ÖZ

KABLOSUZ SÖNÜMLEMELİ KANALLAR İÇİN KAPASİTEYE YAKIN ÇALIŞAN PRATİK ALICI-VERİCİ YAPILARI

Güvensen, Gökhan Muzaffer
Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi : Doç. Dr. Ali Özgür Yılmaz

Şubat 2009, 90 sayfa

Çok-girdili çok-çıktılı (MIMO) sistemler sahip oldukları çoğullama ve çeşitleme kapasiteleri nedeniyle ilgi çekmektedirler. MIMO tabanlı haberleşme tekniklerinin kullanılması ile birlikte, telsiz iletiminde kayda değer spektral verimlilik kazançları elde etmek mümkündür. Yüksek haberleşme hızlarında güvenilir haberleşme yapabilmek için gelişmiş denkleştirme ve kod çözme yöntemlerinin kullanılması gerekmektedir. Bu tez çalışmasında, MIMO telsiz haberleşme kanalları için, kanal bilgisinin verici tarafında bulunup bulunmamasına bağlı olarak kapasiteye yakın başarım gösteren alıcı-verici yapıları önerilmiştir.

İlk olarak, nicemlenmiş kanal bilgisinin verici tarafında kullanılmasına ve indirgenmiş ön kodlama yöntemine dayanan uyarlanır MIMO haberleşme yapısı önerilmiştir. Sınırlı hızda geribeslemeli (LRF) MIMO kanal kapasitesi için elde edilen çok sıkı bir üst sınırın kullanılması ile ön-kodlayıcı sayısını ortalama işaret-gürültü-oranına (SNR) bağlı olarak değiştiren uyarlamalı bir MIMO yapısı tasarlamak mümkündür. Verici tarafında bulunan anten sayısının alıcı tarafındaki anten sayısından fazla olması durumunda, önerilen bu yapının, kanal bilgisinin vericide kullanılmadığı sistemlere göre önemli kapasite kazançları elde ettiği görülmüştür.

İkinci olarak, kanal durum bilgisinin verici tarafında kullanılması ile MIMO kanalı, birbirleriyle girişimde bulunmayan paralel alt kanallara ayırmak ve bu sayede kip çözme karmaşıklığını azaltmak mümkündür. Fakat, kanal bilgisinin getireceği kapasite kazançları, benzer sayıda verici ve alıcı anten durumunda sınırlı kalmaktadır. Dolayısıyla, ikinci kısımda kanal bilgisini verici tarafında kullanmadan bile, alıcı karmasıklıgını makul tutarak, kanal kesinti olasılığına yakın başarım gösteren yapılar üzerinde durulmuştur. Önerilen alıcı yapıları döngülü ileri ve geri beslemeli denkleştirme ve kod çözme tekniklerine dayanmaktadır. Bu yapıların, MIMO kanalın varolan uzay ve frekans çeşitlemesini etkin bir biçimde kullandıkları ve kapasiteye yakın bir başarım elde edildiği görülmektedir. Blok sönümlemeli kanallarda çok boyutlu işaret yıldız kümelerinin döndürülmesi ve MIMO geniş bantlı kanallar için önerilen tek taşıyıcılı-frekans uzayında denkleştirme (SC-FDE) tekniği örnek uygulamalar olarak ele alınmıştır.

Anahtar Kelimeler: Sınırlı hızda geribesleme, MIMO kanal, çeşitleme, yinelemeli karar geribeslemeli denkleştirme, blok sönümlemeli kanal

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor Assoc. Prof. Dr. Ali Özgür Yılmaz for his guidance, patience, advice, encouragement, and insight throughout this thesis work. Without our frequent technical discussions and his supervision during my research, I would never have been able to complete the work presented in this thesis.

It is a great pleasure for me to thank Prof. Dr. Yalçın Tanık for his advice and motivation during my master studies. The technical discussions with him will be invaluable experiences for me during my career.

I would like to thank the other faculty members of the Telecommunication Research Team at METU supported by ASELSAN Inc. for their contribution in establishing my background in the area. The technical discussions with Prof. Dr. Yalçın Tanık, Prof. Dr. Mete Severcan, Assoc. Prof. Dr. Elif Uysal Bıyıkoğlu and Assist. Prof. Dr. Ali Özgür Yılmaz are gratefully acknowledged. I would like to appreciate the opportunity presented by ASELSAN Inc. and I gain a lot of experience as a member of the aforementioned research team. I would also like to thank Assist. Prof. Dr. Çağatay Candan for his encouragement and advice.

I am grateful to The Scientific and Technological Research Council of Turkey (TÜBİTAK) for their financial support during my master's studies.

I would also like to give big thanks to my friends Baha Baran Öztan, Ömür Özel, Tuğcan Aktaş, Selçuk Kavut, Ferit Üzer, Alper Bereketli and Ahmet Şeker for their encouragement, suggestions and for their having such friendly environment at METU over the last two years. Also, I believe that the friendship of my colleague Ümit İrgin has to be honored.

Last but not the least; I am grateful to my dear parents for their endless support during my whole life. I truly owe all my success to them.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

The wireless radio channel is a challenging medium for communication. This is not only due to its susceptibility to noise, interference, and other channel impediments but also due to the unpredictable variation of these impediments over time as a result of user movement and environment dynamics. In wireless channels, the received signal power varies over large distances due to path loss and shadowing. Also, it varies over short distances due to the constructive and destructive addition of signal components coming from different paths in a random manner [22]. This small-scale variation of the channel is called fading and it makes the wireless channel completely different from its wired counterpart where the channel impulse response is often constant and deterministic.

The main performance criterion of interest, while evaluating the different types of receiver structures, is the packet error probability or packet error rate (PER) when the focus is on data communication. Fading can cause a dramatic increase in PER and thus puts a large power penalty on receiver performance over wireless channels. One of the best techniques to mitigate the effects of fading is diversity combining of independently fading signal paths. Diversity combining exploits the fact that independent signal paths have a low probability of experiencing deep fades simultaneously. Thus, the idea behind diversity is to send the same data over independent fading paths. These independent paths can be combined in ways that fading of the resultant signal is reduced significantly [45].

Multiple-input multiple-output (MIMO) systems can be effectively utilized to improve performance through diversity since it is possible to create new independent signaling paths in space by using multiple antennas at the transmitter and/or receiver sides. In addition to their diversity capabilities, MIMO systems can also be used to increase data rates through multiplexing. That is to say, there is another mechanism for performance gain called the

multiplexing gain when both the transmitter and receiver have multiple antennas. The multiplexing gain of a MIMO system results from the fact that a MIMO channel can be decomposed into a number R of parallel independent channels, one can get an $R-$fold increase in data rate in comparison to single-input single-output (SISO) systems [45]. Therefore, MIMO systems attract much attention due to their high multiplexing and diversity capabilities mentioned. They can potentially multiply the transmission rates while preserving the desired reliability in wireless systems. MIMO based structures will function as the backbone of future wireless systems. However, sophisticated equalization and decoding structures or the knowledge of channel at the transmitter side are required in order to capitalize the diversity and multiplexing benefits of the MIMO channels.

The focus of this thesis work is in general on capacity achieving transceiver structures with reasonable complexity in wireless fading channels with special emphasis on MIMO and block fading channels. First, we consider the case where the channel knowledge is assumed to be known both at the receiver and transmitter side. In this case, the MIMO channel can be converted into non-interfering parallel subchannels. This reduces the demodulation and decoding complexity significantly. We later consider the case that the channel knowledge is available only at the receiver side. In this case, it is still possible to achieve the full diversity order of the channel without compromising the multiplexing gain and the receiver complexity as shown in the subsequent chapters of the thesis. In short, we propose various low complexity transmitter and receiver architectures, depending on the availability of the channel state information at the transmitter side, which show a very close performance to the channel capacity. Moreover, we study the tradeoffs between the spectral efficiency, diversity, accuracy of the channel state information at the transmitter (CSIT) along with issues such as constellation size and complexity of the proposed schemes.

1.2 Outline and Contribution of the Thesis

In the first part of the thesis comprising of Chapter 2, the limited rate feedback (LRF) capacity of the MIMO channel will be investigated. The capacity of the MIMO channel will be investigated when the channel knowledge is available at the transmitter with finite precision. The ergodic capacity of a channel is the probabilistic average of the channel mutual information over all possible fading states [22]. We used the ergodic capacity as the performance

measure throughout Chapter 2. It is shown that the number of beamformers used in spatial multiplexing can be adaptively varied depending on the average signal-to-noise ratio (SNR) value and equal power can be allocated to the selected sub channels. This strategy is sufficient to achieve the MIMO channel capacity and allows efficient utilization of the feedback bits required to quantize precoders (beamformers). In Chapter 2, our main contribution is the development of a very tight analytical upper bound for the LRF capacity of the MIMO channel under a wide class of vector quantization methods. This upper bound can be used to determine the SNR regions specified for the operation of studied incremental precoding scheme.

In the second part of the thesis including Chapters 3 and 4, capacity achieving receiver structures with reduced complexity are investigated for MIMO channels. It is observed that all of the sub channels (or degrees of freedom) of a MIMO channel have to be used to maximize the spectral efficiency for moderate and high SNR values since the actual capacity is not affected by the availability of CSIT. In this case, the use of CSIT is useful only for reducing the receiver complexity due to the parallelization of the channel. However, we show that one may attain without CSIT a very close performance to capacity without compromising receiver complexity. In this part, information outage probability is the capacity measure. In capacity with outage, the transmitter fixes a transmission rate and the outage probability associated with this rate is the probability that the channel has mutual information less than this given rate [22]. Outage probability is useful for performance evaluation of the proposed receiver structures in a slow fading scenario, where the channel is constant over a relatively long transmission time and then changes to a new value.

Without CSIT, the traditional transmission scheme can not convert MIMO channel into non-interfering SISO channels and the decoding complexity is exponential in the number of independent symbols transmitted over the multiple transmit antennas in this case. On the other hand, we propose a reduced complexity iterative receiver structure based on feedforward and feedback filtering process for equalization and space-time decoding for general vector channels as given in Fig. 1.1. The forward and backward filters within the equalizer are jointly optimized according to minimum mean square error (MMSE) criterion to mitigate the effect of inter-symbol-interference (ISI) and multi-array interference (MAI) since the signal streams from the multiple transmit antennas are transmitted at the same time and frequency, thus they introduce both MAI and ISI in wideband MIMO communication. The

joint optimization of the forward and backward filters by using the reliability information given by the decoder at each iteration makes our proposed structure different from the traditional iterative equalization and decoding techniques in the literature.

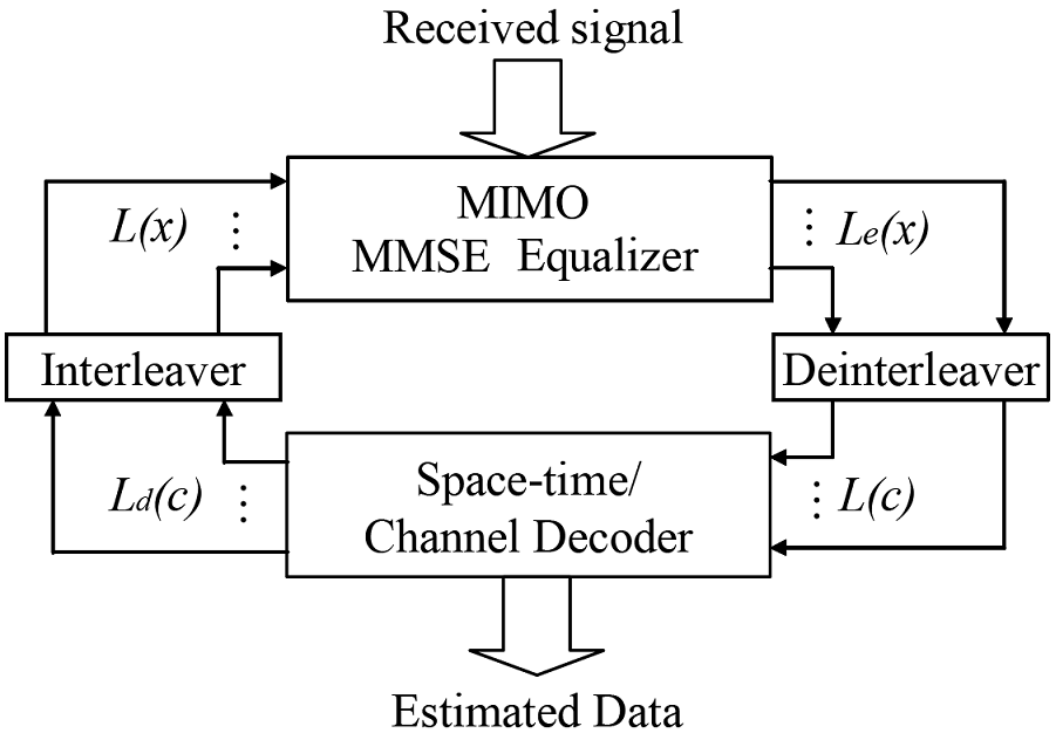

Figure 1.1: Iterative equalization and decoding

Our proposed receiver architecture is based on the information theoretically optimum receiver operation in [23]. We show that this idea can be effectively utilized for the receiver operation of practical MIMO systems. We apply our structure to different practical scenarios in which sophisticated equalization and decoding structures are necessary to achieve a performance close to capacity.

In Chapter 3, we apply our proposed scheme to the decoding of rotated constellations over block fading channels [17] which resemble the MIMO channel. In Chapter 4, we apply the same iterative equalization and decoding idea to the wideband MIMO system in conjunction with single-carrier frequency domain equalization (SC-FDE) technique [36]. In both cases,

it is observed that the proposed structure exploits the multi-path and space diversity sources of the channel effectively such that performance very close to outage probability is achieved. As a side contribution, it is noted that the rotation of multidimensional constellation is an effective technique to combat fading over block fading channels. Moreover, we can say that the SC-FDE technique poses itself as a strong alternative to existing orthogonal frequency division multiplexing (OFDM) based techniques with the use of proposed receiver architecture. Our proposed receiver structure recovers from the exponential decoding complexity by parallelizing the channel similar to the quantizing precoder's idea in Chapter 2. Therefore, the computation complexity of the likelihood information, necessary for the decoding stage, is reduced significantly. Different than the use of CSI at the transmitter with limited feedback, the proposed structure realizes this without using any CSIT but with the help of effective equalization and interference suppression techniques studied herein.

Finally, Chapter 5 is devoted to conclusions which provides a summary of our work and related future studies.

CHAPTER 2

AN UPPER BOUND FOR LIMITED RATE FEEDBACK MIMO CAPACITY

2.1 Introduction

Capacity gains promised by multi-input multi-output (MIMO) systems often require an accurate knowledge of the channel at transmitter and receiver sides especially in quest to capitalize these possible gains in practical systems. An accuracy problem arises when channel state information (CSI) has to be transmitted from the receiver to the transmitter. It is obvious that CSI cannot be transmitted with infinite precision. A limited rate feedback (LRF) channel is usually available for this communication and this sets a limit for the accuracy of CSI at the transmitter side.

It was shown that the MIMO channel is interference-limited when the channel estimation is imperfect [54]. It was further observed in [54] that instantaneous feedback, even if imperfect, gives large capacity gains in low SNR and is useful in high SNR, especially when the number of transmit antennas (n_t) is larger than that of receive antennas (n_r) [6]. In [39], quantization rules and corresponding quantizer design criteria were proposed to be used in MISO (multiple-input single-output) and MIMO channels. Quantization of beamformers were investigated under a Grassmannian line packing framework with regard to quantization codebook size, capacity-SNR loss, and outage performance in [34, 35].

We investigate the capacity of point-to-point MIMO channels in this chapter as opposed to the broadcast channel settings in aforementioned studies [25]. Although the capacity is less affected by the lack of CSI on the transmitter side at high SNR [44], its availability is very important both at low SNR and in designing practical systems that can operate close to the capacity as in adaptively modulated MIMO schemes [56], since the complex task of joint detection and decoding is avoided. Furthermore, the capacity is strictly smaller with $n_t > n_r$

if no CSI is available at transmitter [6]. We concentrate on a finite rate feedback scenario in which precoders obtained by the singular value decomposition of the MIMO channel [22] are fed back to the transmitter side.

A capacity loss bound for covariance matrix based quantization was presented in [13] and a capacity loss bound was proposed in [38] for designing matrix quantization based codebooks. We herein focus on quantizing the columns of the precoding matrix obtained from singular value decomposition (SVD). The channel is quasi-parallelized by separately quantizing precoders and well-known adaptive modulation and coding techniques can be utilized as stated in [6]. Covariance matrices generated randomly with uniform distribution on the unit sphere were used in [13], that is, random matrix quantization was studied. On the other hand, our main contribution in this chapter of the thesis is the derivation of a capacity upper bound expression that is valid for a wide range of vector based quantization schemes. The proposed upper bound turns out to be quite tight mainly due to the exact evaluation of the expected value of matrix determinant as opposed to similar studies using Hadamard inequality to upper bound the determinant as in [13] and using approximate density function of determinant expression and partition cell approximation in [38], [41]. As a byproduct, an absolute upper bound to LRF MIMO capacity using precoding based quantization is also herein derived by utilizing a bounding distribution for Grassmannian beamforming [40]. Finally, we propose a simple quantization method known as product code vector quantization (PCVQ) [21] which can be quite useful to achieve rates quite close to the channel capacity in practice, even for a small number of feedback bits.

2.2 System Model

The following notation is used throughout this chapter. Boldface lower and upper-case letters denote column vectors and matrices, respectively. Scalars are denoted by plain lower-case letters. The superscript $(\cdot)^*$ denotes the complex conjugate for scalars and conjugate transpose for vectors and matrices. The absolute value of a scalar is shown by $|\cdot|$. The $n \times n$ identity matrix is shown by $\mathbf{I}_n$. The trace operator and determinant are denoted by $tr(\cdot)$ and $|\cdot|$, respectively. The autocorrelation matrix for a random vector $\mathbf{a}$ is $R_{\mathbf{a}} = E[\mathbf{a}\mathbf{a}^*]$, where $E[\cdot]$ stands for the expected value operator. The $(i, j)^{th}$ element of a matrix $\mathbf{A}$ is denoted by $A_{i,j}$.

The general expression for a point-to-point MIMO channel with n_r receive antennas and n_t transmit antennas is given by $\tilde{\mathbf{y}} = \tilde{\mathbf{H}}\tilde{\mathbf{x}} + \tilde{\mathbf{w}}$, where $\tilde{\mathbf{y}}$ is the received vector, $\tilde{\mathbf{H}}$ is the $n_r \times n_t$ channel matrix, $\tilde{\mathbf{x}}$ is the transmitted vector, and $\tilde{\mathbf{w}}$ is the zero-mean circularly symmetric complex Gaussian (ZMCSCG) white (spatially and temporally) noise with normalized variance 1. The channel matrix $\tilde{\mathbf{H}}$ is comprised of independent ZMCSCG random variables with variance 1. Considering a block fading model, the channel matrix is assumed to be constant during a coherence interval significantly larger than symbol duration. A fixed average power is allotted for each transmission which corresponds to setting $tr(R_{\tilde{\mathbf{x}}}) \leq P$ (Sec. 10.3 in [22]).

In the case that perfect channel information is available both at the transmitter and receiver, singular value decomposition (SVD) is applied to decompose the MIMO channel into $\min(n_r, n_t)$ parallel subchannels over which multiple streams may be transmitted [32]. The following equivalent expression is obtained for the received vector when SVD is performed to attain $\tilde{\mathbf{H}} = \mathbf{U}\mathbf{D}\mathbf{V}^*$:

$$\mathbf{U}^*\tilde{\mathbf{y}} \quad = \quad \mathbf{D}\mathbf{V}^*\tilde{\mathbf{x}} + \mathbf{U}^*\tilde{\mathbf{w}}. \tag{2.1}$$

The entries of $\mathbf{D}$ are taken to be decreasing without loss of generality. The transmitted vector can be written in general as in $\tilde{\mathbf{x}} = \mathbf{P}\mathbf{\Lambda}\mathbf{x}$ where $\mathbf{P}$ is a precoding matrix, $\mathbf{\Lambda}$ is a diagonal matrix used to distribute power among subchannels, and $\mathbf{x}$ is the original information vector assumed to have $R_{\mathbf{x}} = \mathbf{I}_{\min(n_r,n_t)}$. If the precoding matrix is chosen to be $\mathbf{P} = \mathbf{V}$, by the unitary property of the precoding matrix ($\mathbf{V}^*\mathbf{V} = \mathbf{I}$)

$$\mathbf{y} = \mathbf{D}\mathbf{\Lambda}\mathbf{x} + \mathbf{w}, \tag{2.2}$$

where $\mathbf{y} = \mathbf{U}^*\tilde{\mathbf{y}}$ and $\mathbf{w} = \mathbf{U}^*\tilde{\mathbf{w}}$. Since both $\mathbf{D}$ and $\mathbf{\Lambda}$ are diagonal and $R_{\mathbf{w}} = \mathbf{I}$, the channel is decomposed into parallel subchannels. The capacity is achieved by $\mathbf{\Lambda}$ obtained through the waterfilling procedure [32] with the constraint that $tr(\mathbf{\Lambda^2}) \leq P$. We note here that the columns of matrix $\mathbf{V}$ are isotropically distributed on the n_t-dimensional complex unit circle when considered over the realizations of $\tilde{\mathbf{H}}$. When there is only a partial CSI at transmitter due to finite rate feedback, one has imperfect precoding and power distribution matrices denoted by $\mathbf{V_f}$ and $\mathbf{\Lambda_f}$, respectively. Eqn. (2.2) now becomes

$$\mathbf{y} = \mathbf{D}\mathbf{V}^*\mathbf{V_f}\mathbf{\Lambda_f}\mathbf{x} + \mathbf{w} \tag{2.3}$$

which suggests that subchannels now interfere with each other since $\mathbf{V}^*\mathbf{V_f} \neq \mathbf{I}$, in general. We will investigate the capacity of LRF MIMO channels based on (2.3).

In order to reduce the rate of the feedback channel, the idea of reduced precoding can be used [33]. In this scheme, the number of beamformers used in a spatial multiplexing system is adaptively varied in order to minimize probability of symbol vector error or to maximize capacity by allocating equal power ($\mathbf{\Lambda_f^2} = \frac{P}{n}\mathbf{I_n}$) to selected subchannels [33], [38]. Transmitting only the precoding vectors corresponding to the strongest subchannels will suffice to maximize communication rate over MIMO channels. Thus, this strategy allows efficient utilization of the feedback bits by quantizing only relevant precoders.

The analytical bound for limited rate feedback MIMO capacity to be obtained in Section 2.4 can be used to determine the number of precoders to be used at each average SNR value in order to maximize the spectral efficiency. The idea of reduced precoding and the utilization of feedback for precoders are not only useful at low SNR values but also at high SNR especially for MIMO systems with $n_t > n_r$ [6].

2.3 Capacity with Limited Rate Precoding

Eqn. (2.3) can be written with an equivalent channel matrix $\mathbf{H} = \mathbf{D}\mathbf{V}^*\mathbf{V_f}$ as in

$$\mathbf{y} = \mathbf{H}\mathbf{\Lambda_f}\mathbf{x} + \mathbf{w}. \tag{2.4}$$

One can calculate the achievable rate of the LRF scheme with reduced precoding by calculating the mutual information between $\mathbf{x}$ and $\mathbf{y}$ for a Gaussian input alphabet with equal power distribution among selected subchannels as done in [22] for MIMO channels. Throughout this thesis, we refer to the maximum achievable rate as the system capacity for a given scheme, whereas the term channel capacity is reserved for the capacity of the channel itself. Noting that $\mathbf{\Lambda_f^2} = \frac{P}{n}\mathbf{I_n}$, the capacity of this scheme which makes use of n precoders is given by

$$\mathbf{C_{n-pre}} = E\left\{\log_2 \det\left(\mathbf{I_n} + \frac{P}{n}\mathbf{H}\mathbf{H}^*\right)\right\}. \tag{2.5}$$

and can be achieved with MMSE estimation and successive interference cancellation [26, 27, 42] where the equivalent channel matrix $\mathbf{H}$ can be written as

$$\begin{aligned}\mathbf{H} &= \mathbf{D}\mathbf{V}^{*}\mathbf{V_f} \\ &= \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix} \begin{bmatrix} \mathbf{v_1^*} \\ \mathbf{v_2^*} \\ \vdots \\ \mathbf{v_n^*} \end{bmatrix} \begin{bmatrix} \mathbf{v_{1f}} & \dots & \mathbf{v_{nf}} \end{bmatrix}.\end{aligned} \tag{2.6}$$

The equivalent channel matrix $\mathbf{H}$ has its $(i, j)^{th}$ element as $H_{i,j} = d_i \mathbf{v_i^* v_{jf}}$, where $\mathbf{v_i}$ is the i^{th} column of $\mathbf{V}$ and, $\mathbf{v_{jf}}$ is the j^{th} column of $\mathbf{V_f}$. Defining $V_{ij} = \mathbf{v_i^* v_{jf}}$, one can evaluate

$$(\mathbf{H}\mathbf{H}^{*})_{i,j} = d_i d_j \left(\sum_{k=1}^{n} V_{ik} V_{jk}^{*} \right). \tag{2.7}$$

Evaluation of the capacity in (2.5) requires the probability distributions of $V_{ij} = \mathbf{v_i^* v_{jf}}$ for $i, j \in \{1, 2, \dots, n\}$ and hence the quantization rule used for limited rate feedback has to be specified.

A set of 2^{N_f} vectors $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_{2^{N_f}}\}$ generated to construct the quantization codebook are defined where N_f stands for the number of feedback bits per precoding vector. Quantization vectors are length-n_t complex vectors on the n_t-dimensional complex unit circle and the quantization while obtaining the precoding vectors is determined by the following rule used in LRF MIMO studies [1, 25, 34, 35]:

$$\mathbf{v_{if}} = \arg \max_{\mathbf{q_j},\ j=1,\dots,2^{N_f}} |\mathbf{v_i^* q_j}|^2. \tag{2.8}$$

Any quantization codebook can be used for this purpose and there is no prerequisite on the construction of the codebook. The codebooks are available both at the transmitter and receiver. The index j, corresponding to the selected quantization vector, is sent from the receiver to the transmitter side and thus the transmitter carries out the beamforming operation by using the precoding vectors with proper indices.

The MIMO channel can be quasi-parallelized by using the given quantization rule since the precoders are separately quantized in this case and for sufficient number of feedback bits, spatial interference between the subchannels is very small. In this case, the demodulation

complexity is significantly reduced and well known adaptive modulation and coding techniques for parallel block fading channels can be easily incorporated in our LRF system.

There are two types of random variables in (2.5) whose distributions and dependence properties have to be determined in order to evaluate $\mathbf{C_{n-pre}}$. First, the cumulative distribution function (cdf) of $V_{ii} = \mathbf{v_i^* v_{if}}$ is peculiar to the given quantization codebook and rule, and we will investigate the cdf of V_{ii} for two different quantization methods in Section 2.5. Moreover, the cdf for V_{ij}'s for $i \neq j$ is needed. In [1, 34, 35], the cdf of the squared absolute inner product between two isotropically distributed length-n_t complex unit vectors is given as

$$F_o^{n_t}(x) = \begin{cases} 1-(1-x)^{n_t-1} & , 0 \leq x \leq 1 \\ 0 & , x < 0 \\ 1 & , x > 1. \end{cases} \tag{2.9}$$

The same result and hence cdf hold for the case of one fixed vector and an isotropically distributed vector, since one of them being isotropically distributed is sufficient for the result [34]. Bearing in mind that $\mathbf{v_i^* v_{jf}}$ corresponds to projection of $\mathbf{v_i}$ onto $\mathbf{v_{jf}}$, the following holds by orthogonality of $\mathbf{v_i}$'s in our problem

$$\mathbf{v_k^* v_{1f}} = \mathbf{v_k^*}\left(\mathbf{v_{1f}} - \sum_{j=1}^{k-1} (\mathbf{v_j^* v_{1f}})\mathbf{v_j}\right) \tag{2.10}$$

for $k = 2, \ldots, n$. Defining $\mathbf{v_{1f}}' = \mathbf{v_{1f}} - \sum_{j=1}^{k-1} (\mathbf{v_j^* v_{1f}})\mathbf{v_j}$, the vector $\mathbf{v_{1f}}'$ is in the null space of $\mathbf{v_i}$'s, $i = 1, \ldots, k-1$, where the null space has dimension $(n_t - k + 1)$. The squared norm of $\mathbf{v_{1f}}'$ is $(\mathbf{v_{1f}}')^* \mathbf{v_{1f}}' = 1 - \sum_{j=1}^{k-1} |\mathbf{v_j^* v_{1f}}|^2 = 1 - \sum_{j=1}^{k-1} |V_{j1}|^2$. Considering the projection of a fixed vector $\mathbf{v_k}$ onto an isotropically distributed vector $\mathbf{v_{1f}}'$ which is of dimension $(n_t - k + 1)$, one obtains the following conditional probability distribution function for $|\mathbf{v_k^* v_{1f}}|^2$ by using (2.9):

$$F_{|\mathbf{v_k^* v_{1f}}|^2}(x \mid |\mathbf{v_i^* v_{1f}}|^2 = a_i, \quad i = 1, \ldots, k-1) = F_o^{n_t-k+1}\left(\frac{x}{\left(1 - \sum_{i=1}^{k-1} a_i\right)}\right) \tag{2.11}$$

for $k = 2, \ldots, n$. The expectation in (2.5) is over the channel matrix $\mathbf{H}$ or equivalently, over V_{ij}'s. For a given channel realization, $\mathbf{v_i}$'s are fixed and the quantized precoding vectors $\mathbf{v_{if}}$'s are chosen independently of each other as the rule given in (2.8) dictates directly. Hence, $\mathbf{v_{if}}$'s are independent of each other on the condition that $\mathbf{v_i}$'s are given. One should note that $\mathbf{V^* V_f}$ product involves V_{ij} terms and the phase of V_{ij} becomes relevant in this case. Recalling that the vectors are isotropically distributed, the phases of all the random variables corresponding to the projections of the precoders onto quantized precoders are independent and uniformly

distributed in $[0, 2\pi]$, since the quantization rule in (2.8) is blind to multiplication of all the entries of a quantization vector by a complex number α of unity amplitude as $|\mathbf{a}^*\mathbf{b}|^2 = |\alpha\mathbf{a}^*\mathbf{b}|^2$. To summarize, it holds true that V_{ij} has a uniformly distributed phase in $(0, 2\pi)$ and it is independent of V_{ik} for all $k \neq j$ and V_{lj} for all $l \neq i$ for given $\mathbf{v_i}$'s, since knowing $\mathbf{v_i}$ only is sufficient to determine $\mathbf{v_{if}}$ for the given rule in (2.8).

Isotropical distribution implies that $|V_{ii}|^2 = |\mathbf{v_i^* v_{if}}|^2$'s for $i = 1, \ldots, n$ are identically distributed and independent random variables. Similarly, a corresponding distribution holds for $|V_{kj}|^2 = |\mathbf{v_k^* v_{jf}}|^2$ and its cdf has a form identical to that given in (2.11).

2.4 A Capacity Upper Bound

The capacity expression in (2.5) is impractical to be used in practical system design since it needs the distribution of V_{ij}'s and the expectation over V_{ij}'s distributions. We will obtain a very tight analytical upper bound for capacity of the n-precoder scheme and this analytical bound needs only the expectations $E_{11} = E|V_{11}|^2$ and $E_{21} = E|V_{21}|^2$ to be evaluated. The expected values will be denoted with $E_{ij} = E|V_{ij}|^2$ and E_{ii}'s are the same for all i's and can be evaluated for a given quantization codebook. When the value of E_{11} is given, the value of E_{21} can be calculated easily as follows. Due to the isotropical distribution of $\mathbf{v_i}$'s, E_{ij} is the same as E_{21} for all $i \neq j$. The value of E_{21} can be found in terms of E_{11} by using (2.11) as

$$\begin{aligned} E[|V_{21}|^2] &= E\left[E[|V_{21}|^2 \mid |V_{11}|^2]\right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{|V_{21}|^2}(x \mid |V_{11}|^2 = a) f_{|V_{11}|^2}(a) dx da \\ &= \int_{-\infty}^{\infty}\left(\int_0^{1-a} x\frac{(n_t-2)}{(1-a)}\left(1-\frac{x}{1-a}\right)^{n_t-3} dx\right) f_{|V_{11}|^2}(a) da \\ &= \int_{-\infty}^{\infty}\frac{1-a}{n_t-1} f_{|V_{11}|^2}(a) da = \frac{1-E_{11}}{n_t-1} \end{aligned} \tag{2.12}$$

This result in (2.12) is intuitive since, after projecting $\mathbf{v_{1f}}$ on $\mathbf{v_1}$, $n_t - 1$ orthonormal vectors $\{\mathbf{v_2}, \mathbf{v_3}, \ldots, \mathbf{v_{n_t}}\}$ are left. Due to isotropical distribution, the power in the part of $\mathbf{v_{1f}}$ orthogonal to $\mathbf{v_1}$ is distributed equally between $n_t - 1$ orthonormal vectors in average.

The capacity for the $n-$precoder scheme can be upper bounded by using Jensen's inequality such that

$$E\left\{\log_2\left|\mathbf{I_n} + \frac{P}{n}\mathbf{H}\mathbf{H}^*\right|\right\} \leq \log_2\left(E\left\{\left|\mathbf{I_n} + \frac{P}{n}\mathbf{H}\mathbf{H}^*\right|\right\}\right). \tag{2.13}$$

The capacity bound of the n-precoder case given in (2.13) can be written in terms of E_{11} and E_{12} by using the result of the lemma given in Appendix A as

$$\mathbf{C_{n-pre}} \leq \log_2\Big(1 + \sum_{k=1}^{n} \sum_{S_k \in P_k} \left(\frac{P}{n}\right)^k E\left[(d_{a_1} d_{a_2} \cdots d_{a_k})^2\right] \cdot \sum_{(j_1 \in S)} \sum_{(j_2 \in S,\ j_2 \neq j_1)} \cdots \sum_{(j_k \in S,\ j_k \neq j_1, \ldots, j_{k-1})} \left(E_{a_1 j_1} \cdots E_{a_k j_k}\right)\Big) \tag{2.14}$$

where

$$E_{a_i j_i} = E\left[|V_{a_i j_i}|^2\right] = \left\{ \begin{array}{ll} E_{11} = E\left[|V_{11}|^2\right] & \text{if } a_i = j_i \\ E_{12} = E\left[|V_{12}|^2\right] & \text{if } a_i \neq j_i \end{array} \right\}, \tag{2.15}$$

$S = \{1, 2, \ldots, n\}$, $S_k = \{a_1, \ldots, a_k\}$; and P_k is the set containing all possible (n, k) combinations of S. The term $\sum_{(j_1 \in S)} \cdots \sum_{(j_k \in S,\ j_k \neq j_1, \ldots, j_{k-1})} \left(E_{a_1 j_1} \cdots E_{a_k j_k}\right)$ does not depend on which $k-$element combination $(a_1, a_2, \ldots, a_k)$ chosen from set P_k is used in (2.14) since j_i's for $i = 1, 2, \ldots, k$ are chosen from the set $S = \{1, 2, \ldots, n\}$. Using this fact, one can further simplify the capacity bound given in (2.14) by selecting $(a_1, a_2, \ldots, a_k)$ as $(1, 2, \ldots, k)$ such that

$$\mathbf{C_{n-pre}} \leq \log_2\Big(1 + \sum_{k=1}^{n} \left(\frac{P}{n}\right)^k E\left\{\sum_{S_k \in P_k} (d_{a_1} d_{a_2} \cdots d_{a_k})^2\right\} \cdot \sum_{(j_1 \in S)} \sum_{(j_2 \in S,\ j_2 \neq j_1)} \cdots \sum_{(j_k \in S,\ j_k \neq j_1, \ldots, j_{k-1})} \left(E_{1 j_1} \cdots E_{k j_k}\right)\Big). \tag{2.16}$$

$E\left\{\sum_{S_k \in P_k} (d_{a_1} d_{a_2} \cdots d_{a_k})^2\right\}$'s for $k = 1, \ldots, n$ are required to evaluate the capacity bound in (2.16) and one does not need to calculate $E\left[(d_{a_1} d_{a_2} \cdots d_{a_k})^2\right]$ for all possible (n, k) combinations $(a_1, \ldots, a_k)$ from set S separately. Instead, the expectation of the sum of all possible joint moments are necessary. Hence, only the expectation of the sum of all possible joint moments $E\left\{\sum_{S_k \in P_k} (d_{a_1} d_{a_2} \cdots d_{a_k})^2\right\}$ has to be found. This can be found by using Monte Carlo simulations or analytically by using the currently available results for the joint ordered moments in a closed form presented in [24] and this makes the proposed bound completely analytical.

As will be seen in Section 2.6, the capacity bound in (2.16) is very tight and can be used for practical purposes. With the help of this analytical bound, one can design an adaptive MIMO system that can change the number of precoders and the feedback rate according to the current average SNR value. The bound given in (2.16) is valid for many vector quantization methods as long as $V_{ij} = \mathbf{v_i^* v_{jf}}$ has a distribution such that the phase of V_{ij} is uniformly distributed in $(0, 2\pi)$ which is independent of V_{ik} for all $k \neq j$ and V_{lj} for all $l \neq i$. Hence,

this uniform phase distribution of V_{ij} is a sufficient condition and the quantization rule in (2.8) satisfies it. Thus, the results of this section can be applied to a wide variety of quantization schemes. Moreover, the bound is also valid for no quantization cases and we can easily construct an upper bound for the exact capacity of a MIMO system that uses n of its strongest subchannels by simply setting $E_{11} = 1$ and $E_{12} = 0$ in (2.16).

2.5 Bounding Distribution and RVQ

In this section, we will present two example quantization methods that will be used to produce some numerical results in Section 2.6. The first method is random vector quantization (RVQ) in which a quantization codebook is generated randomly and the closest vector in the codebook is conveyed to the transmitter according to the rule given in (2.8) [1]. Quantization vectors are length$-n_t$ complex vectors which are independently chosen from the isotropic distribution on the n_t-dimensional complex unit sphere. The performance of RVQ is inspected as averaged over many codebooks that are generated randomly. The use of RVQ as a vector quantization scheme allows simpler analysis and can be helpful in designing practical limited rate feedback MIMO systems. The cdf of $|V_{11}|^2$ has been readily obtained in [1, 34, 35] by using the standard probability evaluation technique for maximum of 2^{N_f} independent random variables with the distribution given in (2.9) as

$$F_{|\mathbf{v_1^* v_{1f}}|^2}(x) = \left(F_o^{n_t}(x)\right)^{2^{N_f}} . \tag{2.17}$$

The phase of V_{ij} is independent of its magnitude and uniformly distributed in $[0, 2\pi]$. The expected value of a random variable distributed with (2.17) is evaluated in [1] as

$$E[|V_{11}|^2] = E_{11} = 1 - 2^{N_f} B(2^{N_f}, \frac{n_t}{n_t - 1}), \tag{2.18}$$

where $B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the Beta function and the Gamma function is given by $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ while $E_{12} = E|V_{12}|^2$ can be calculated from (2.12).

The second method is the one that maximizes the capacity by increasing the projection power, namely $|V_{11}|^2$. By defining a random variable z as $z = 1 - |\mathbf{v_1^* v_{1f}}|^2$, we can say that for a good beamformer, the expected value of z has to be as close as possible to zero. An upper

bound for the cdf of z, $F_z(z)$, is given in [40] such that $F_z(z) \leq \tilde{F}_z(z)$ for $0 \leq z \leq 1$ and

$$\tilde{F}_z(z) = \begin{cases} 2^{N_f}.z^{n_t-1} & , 0 \leq z < \left(\frac{1}{2^{N_f}}\right)^{\frac{1}{(n_t-1)}} \\ 1 & , z \geq \left(\frac{1}{2^{N_f}}\right)^{\frac{1}{(n_t-1)}} \end{cases} . \tag{2.19}$$

A good beamformer shall try to come as close as possible to this distribution which we refer to as the bounding distribution hereafter [40]. For a hypothetical quantization method that attains this bounding distribution, we can evaluate the E_{11} and E_{12} values in order to construct an upper bound to the capacity of limited rate feedback MIMO scheme that none of the quantization methods can exceed. E_{11} can be found as given below

$$E_{11} = 1 - E[z] = 1 - 2^{N_f}\left(\frac{n_t - 1}{n_t}\right)\left(\frac{1}{2^{N_f}}\right)^{\left(\frac{n_t}{n_t-1}\right)} \tag{2.20}$$

and E_{12} by (2.12). The value of E_{11} obtained by this bounding distribution in (2.20) is the maximum value of E_{11} that can be achieved among all quantization codebooks using the quantization rule in (2.8) for given N_f and n_t values.

2.6 Numerical Results

In this section, the capacity upper bound obtained for LRF MIMO given in (2.16) is evaluated for RVQ and the bounding distribution. Random variables are generated by the inversion method [14] in simulations and placed in (2.5) to obtain the ergodic capacity. The random variables $|V_{ii}|^2$'s are generated first, while the others are drawn based on the distribution given in (2.11) and each data point is obtained by generating $10,000$ realizations.

In Fig. 2.1, the ergodic capacity of a 6×3 MIMO scheme that uses 2 precoders with $N_f = 6, 12$ for RVQ and bounding distribution is compared with the 6×3 MIMO channel capacity obtained with waterfilling (WF) that uses a short-term power constraint (Sec. 10.3 in [22]). The equipower scheme that does not use CSIT and allocates equal power among the transmit antennas is also depicted for comparison. SNR is taken as the average received SNR. It can be observed that reduced precoding is an efficient and reasonable technique since it is better than the equipower scheme up to 14 dB and it matches the MIMO channel capacity with WF and full CSIT up to 3 dB in case of no quantization even with 2 precoders. As SNR increases, achievable rates with the $2-$precoder scheme diverge from the 6×3 capacity since all degrees of freedom are not utilized. By increasing the number of precoders used as SNR

increases, the system can operate close to the channel capacity. As N_f increases, the losses are quite tolerable for the reduced precoding scheme but the capacity of equipower scheme without CSIT does not match the ergodic channel capacity (full CSIT case) even at high SNR. In high SNR regimes, there is a $\log_{10}(\frac{n_t}{min(n_r,n_t)}) = \log_{10}(\frac{6}{3}) = 3$ dB difference between the MIMO channel capacity and the equipower scheme without CSIT and this difference becomes more significant especially for MIMO systems with $n_t > n_r$.

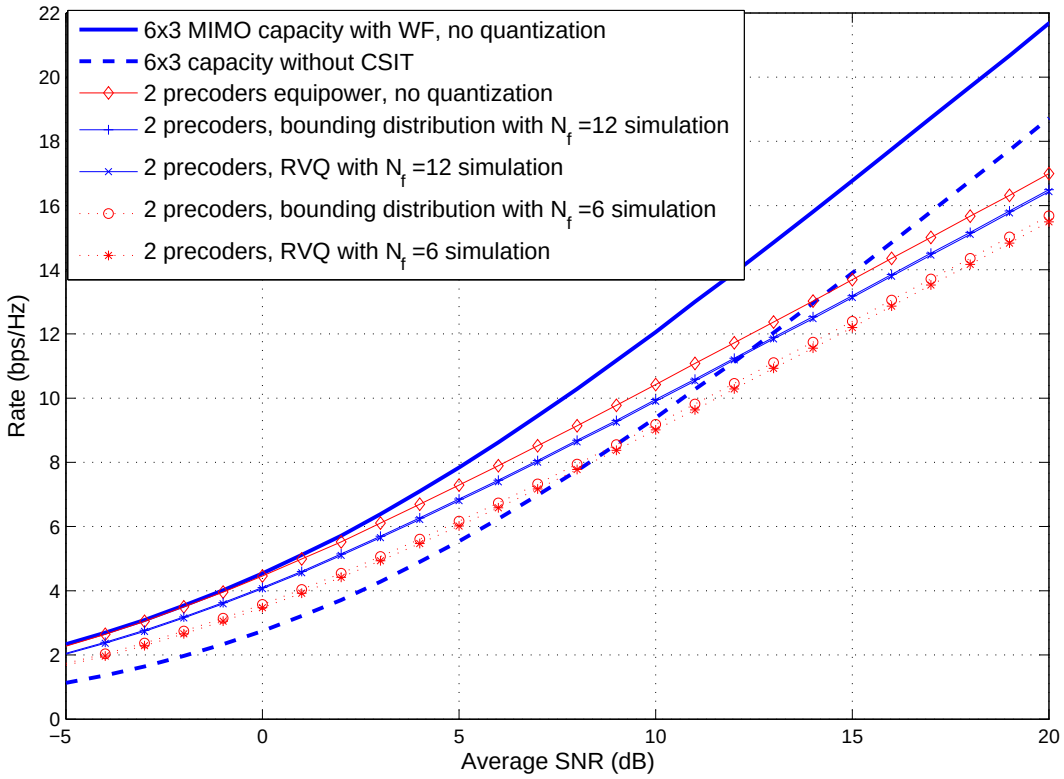

Figure 2.1: 6×3 MIMO capacities for 2-precoder LRF scheme. RVQ and bounding distribution are used with $N_f = 6,\ 12$.

Another observation is that the quantization is not too detrimental when compared to broadcast MIMO channels studied in [25]. The number of bits N_f to perform close to the capacity is relatively small when compared with the pessimistic results presented for broadcast channels in [25] with the basic statement that quite a large number N_f is necessary for performance close to capacity with ideal precoding.

In Fig. 2.2, 6×3 MIMO system capacity under limited rate feedback with $N_f = 8$ is evaluated by using RVQ and the bounding distribution for 2 and 3-precoder schemes, where the capacity bound given in (2.16) is also found by using E_{11} and E_{12} values for these schemes. It is seen in the figure that the capacity upper bound is very tight. It is 0.3-0.5 dB away from the 6×3 LRF MIMO capacities. It is further observed that the RVQ scheme is almost optimal since RVQ and bounding distribution capacities are quite close to each other. There is a 0.2 dB difference between these two capacities and hence we can say that RVQ can be used as a practical quantization technique that attains rates quite close to the capacity with tolerable N_f values. This also justifies its use in the literature for analysis purposes. Moreover, when compared to the MIMO channel capacity obtained with waterfilling (WF) that uses a short-term power constraint (Sec. 10.3 in [22]), there is a 1.3 dB loss in limited feedback incremental precoding scheme that uses RVQ with $N_f = 8$ which is well predicted by the proposed bound at high SNR.

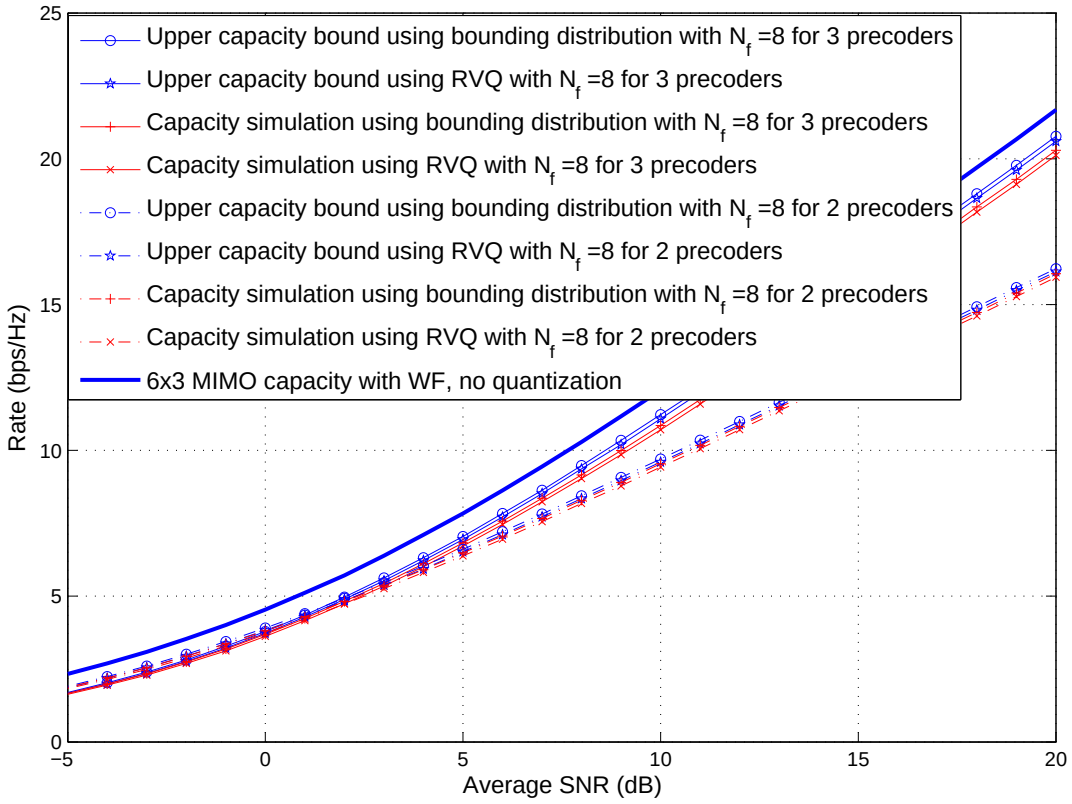

Figure 2.2: 6×3 MIMO capacities and upper bounds for 2 and 3-precoder LRF schemes. RVQ and bounding distribution are used with $N_f = 8$.

In Fig. 2.3, RVQ is used as a quantization technique and the capacities of $n-$precoder 8×4 MIMO schemes are evaluated with their corresponding bounds for $N_f = 12$ and $n = 2, 3,$ and 4. The scheme wihout CSIT that allocates equal power among the transmit antennas is also depicted for comparison. It can be observed that the capacity bounds evaluated with the E_{11} value of RVQ for $N_f = 12$ is quite fine and 0.15-0.4 dB away from the ergodic system capacity with $n = 2, 3,$ and 4. Furthermore, it is seen that the 2-precoder capacity with $N_f = 12$ is better than the scheme without CSIT up to 6dB, 3-precoder scheme is better up to 13.5 dB. The 4-precoder scheme that uses all the degrees of freedom in the system always has higher capacity than the scheme that does not use CSIT. At high SNR, there is a 1.6 dB difference between the 4-precoder and equipower schemes without CSIT. When compared to the channel capacity with WF, there is a 1.4 dB loss in limited feedback incremental precoding scheme at high SNR and this can be compensated by increasing N_f. $N_f = 12$ may seem to be large for practical systems but this large value is due to the large transmit antenna number ($n_t = 8$). In contrast, a reasonable N_f value is sufficient to reach the capacity for a small size MIMO system (small n_t value).

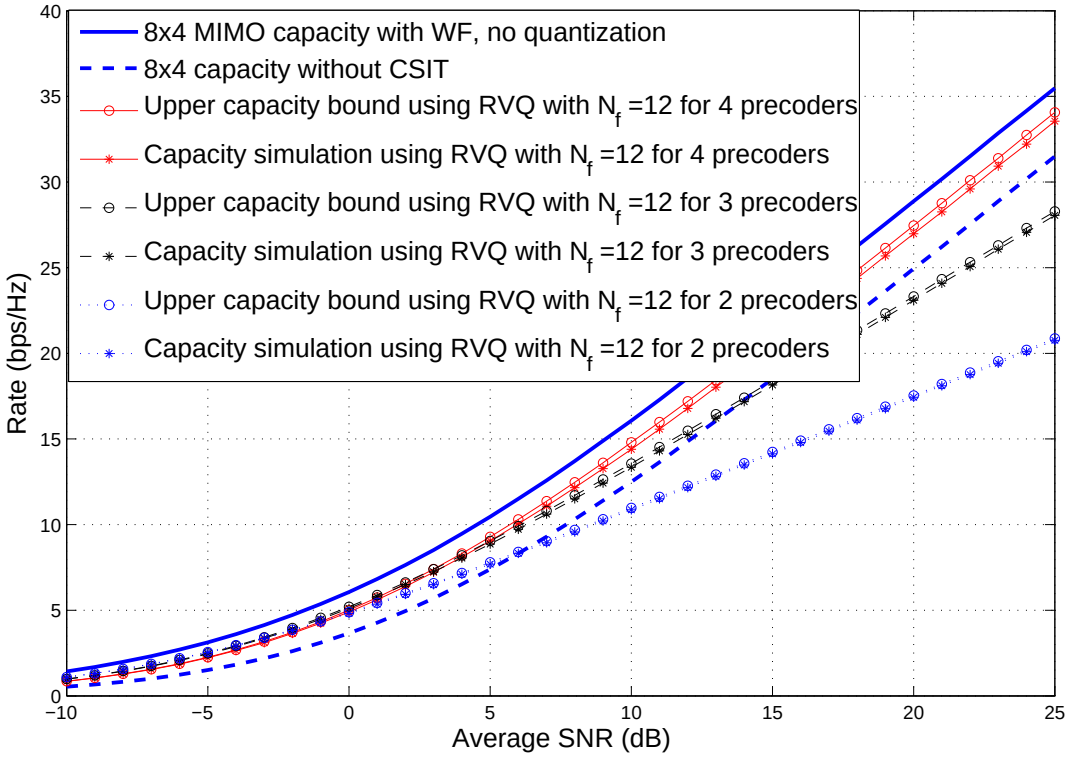

Figure 2.3: 8×4 MIMO capacities and upper bounds for 2, 3 and 4-precoder LRF schemes. RVQ is used with $N_f = 12$.

In case of limited feedback, the losses are approximately equal to the SNR loss due to quantization $10\log_{10}(\frac{1}{E_{11}})$ dB. Intuitively speaking, the instantaneous effective SNR of the channel for single precoder case is $P \cdot |V_{11}|^2$ and hence there occurs a $-10\log_{10} E_{11}$ dB SNR loss in the LRF scenario. This holds approximately true for the general $n-$precoder scheme. As N_f increases, E_{11} asymptotically becomes 1 as observed in (2.18) and (2.20) so that the capacity of LRF MIMO approaches to the channel capacity with no quantization. The capacity upper bound in (2.16) is maximized at $E_{11} = 1$ and $E_{12} = 0$, thus a good quantization scheme should have a E_{11} value which is as close to 1 as possible for a given number of feedback bits. Among the given vector quantization techniques and N_f, the best quantization codebook is the one that gives the highest E_{11} value and thus, it has the greatest capacity bound value in (2.16).

As a result, we can use the capacity bound given in (2.16) to evaluate the performance of different quantization schemes. For a given quantization scheme, it can be used to determine the number of precoders to be used at each average SNR value. In other words, for a given MIMO system and the quantization technique with N_f value, one calculates the E_{11} value of the quantization and the expectation of the sum of possible joint moments $E\left\{\sum_{S_k \in P_k}(d_{a_1} d_{a_2} \cdots d_{a_k})^2\right\}$ in (2.16) only once. After that, these two values can be used to construct the bound in (2.16) easily. With the help of this bound, it is possible to determine the SNR regions in which the number of precoders to be used to maximize system capacity are specified during the operation of the studied incremental precoding scheme. For our example quantization in Fig. 2.3, it is seen that using 2 precoders up to 0 dB, using 3 precoders between 0 and 5 dB, and using 4 precoders above 5 dB maximizes the capacity and this strategy always results in significantly higher capacity than the equipower MIMO scheme without CSIT, especially for $n_t > n_r$.

The main conclusion is that the tightness of the proposed bound is established. The upper bound can be applied to many LRF MIMO schemes and also to other problems where the exact evaluation of the expected value of the determinant is needed.

2.7 A Vector Quantization Method Based on Product Codes

We observed in the previous section that good performance can be attained by the RVQ method for quantization of precoders. However, RVQ is not a practical scheme since an averaging over the quantization codebooks have to be performed to enjoy its good performance. Meanwhile there are various vector quantization methods in literature which can be used for quantization of precoders [21]. In such methods, a precoder is usually compared to a large codebook and the best matching vector is declared as the quantized vector. This matching operation which corresponds to the search of the vector in the codebook closest to the given vector constitutes a major disadvantage especially for large codebooks. In contrast, each entry of the precoder could be quantized with a scalar quantizer. When all the quantized entries are combined, an overall quantized vector comes up which represents the quantization of the precoder. This sort of quantization which is based on scalar quantization is referred to as product codes in quantization literature [21]. Allocating a fixed number of quantization bits for each dimension, the search time in the latter method is linear in the number of dimensions, whereas it is exponential for vector quantization.

Since each entry of a precoder will be quantized independently of all the others in this method, the probability density function for this complex random variable is required. The probability density function obviously depends on the dimensionality n_t. As an example, all entries should have absolute value equal to 1 when $n_t = 1$, whereas this is not necessarily so for $n_t = 2$. Due to lack of an expression for the aforementioned probability density function, we resort to the widely used technique of training-based Lloyd algorithm [21] with the distortion measure of squared absolute value. The training set is generated by first generating isotropically distributed n_t-dimensional unit-norm vectors and then placing all the entries of vectors in a sequence. The initial codebook is generated randomly and after a sufficient number of iterations, the Lloyd algorithm produces a codebook close to the optimum quantizer belonging to the probability distribution of a single entry of the isotropically distributed n_t-dimensional unit norm vectors for a given n_t and bits per dimension.

Specifying the number of codewords in the scalar quantization, the algorithm is run and results similar to in Fig. 2.4 are obtained. The codebooks are rotation invariant since phases of entries are uniformly distributed by the isotropic distribution of the precoders and hence the codebook is a constellation with an arbitrary rotation. No optimization for the initial

codebook or the resultant codebook has been performed since the aim is to show that even a non-optimally designed codebook will perform quite good from a capacity point of view. We refer to this coding scheme as product code vector quantization (PCVQ).

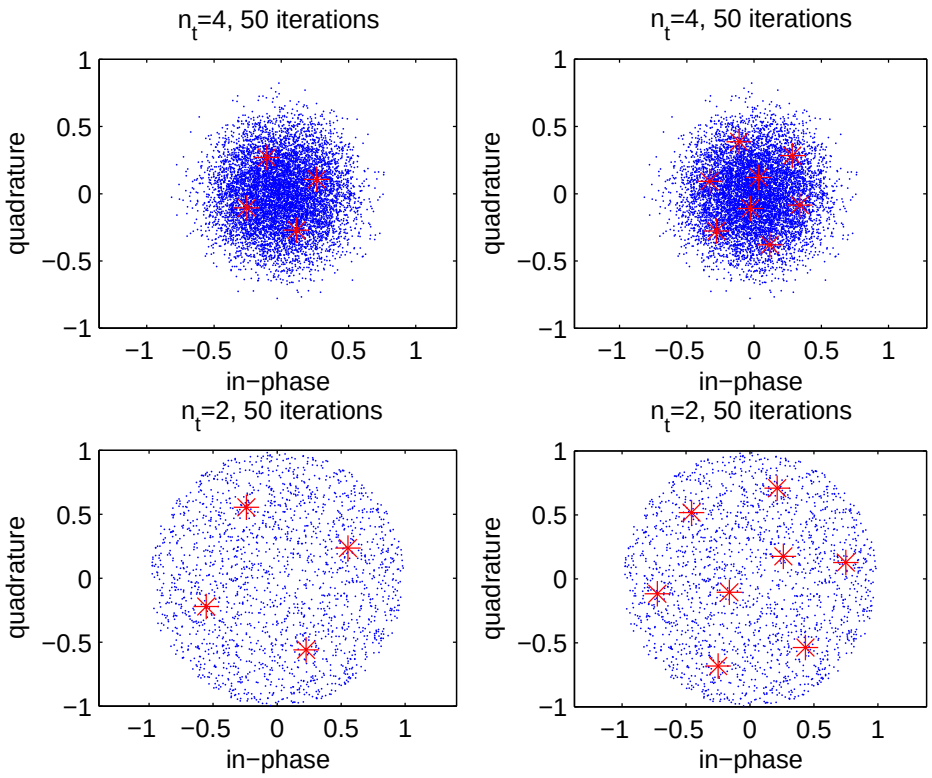

Figure 2.4: Codewords generated for different codebook sizes and dimensionality. Each codebook is generated by 50 iterations of the Lloyd algorithm.

Let the quantized form of the precoder be denoted by $\mathbf{Q(V_1)}$. This quantized vector is not necessarily a unit-norm vector. Hence, the finite rate precoder vector fed back to the transmitter side is scaled first to keep the transmitted power constant and then used as the precoder vector. Hence, in PCVQ

$$\mathbf{V_{1f}} = \frac{\mathbf{Q(V_1)}}{\|\mathbf{Q(V_1)}\|}. \tag{2.21}$$

The method of quantization does not naturally alter the system model and (2.3) holds in the case of PCVQ. Then, for different realizations of the channel matrix $\tilde{\mathbf{H}}$, one can obtain the quantization of the precoding matrix by using the codebook. Using these quantized precoders for each realization, one can evaluate the $n-$precoder capacity given in (2.5).

In Fig. 2.5 and 2.6, The capacity of RVQ and PCVQ MIMO systems are compared for different number of used precoders. The capacity curves are obtained by 10,000 point Monte Carlo simulation. When the scalar codebooks of sizes (number of vectors in the codebook) 2 and 4 are closely inspected, one can see that the constellation is symmetric around 0 and hence there will be a reduction of 2 bits for the codebook of size 4 and a reduction of 1 bit for the codebook of size 2 in the total number of bits used for quantizing a precoder. This is due to the fact that the precoders are invariant under multiplication by scalars. This enables the entry of the first (or one of the other) dimension to be fixed. Therefore, we have to compare PCVQ using 2 bits per dimension with $N_f = 14$ RVQ for 8×8 MIMO and PCVQ using 1 bit per dimension with $N_f = 7$ RVQ for 8×8 MIMO. For 8×8 MIMO scheme, there is a loss by PCVQ about 0.5 dB when only two bits are alloted to each dimension. By increasing the bits alloted to each dimension, the loss can be significantly reduced.

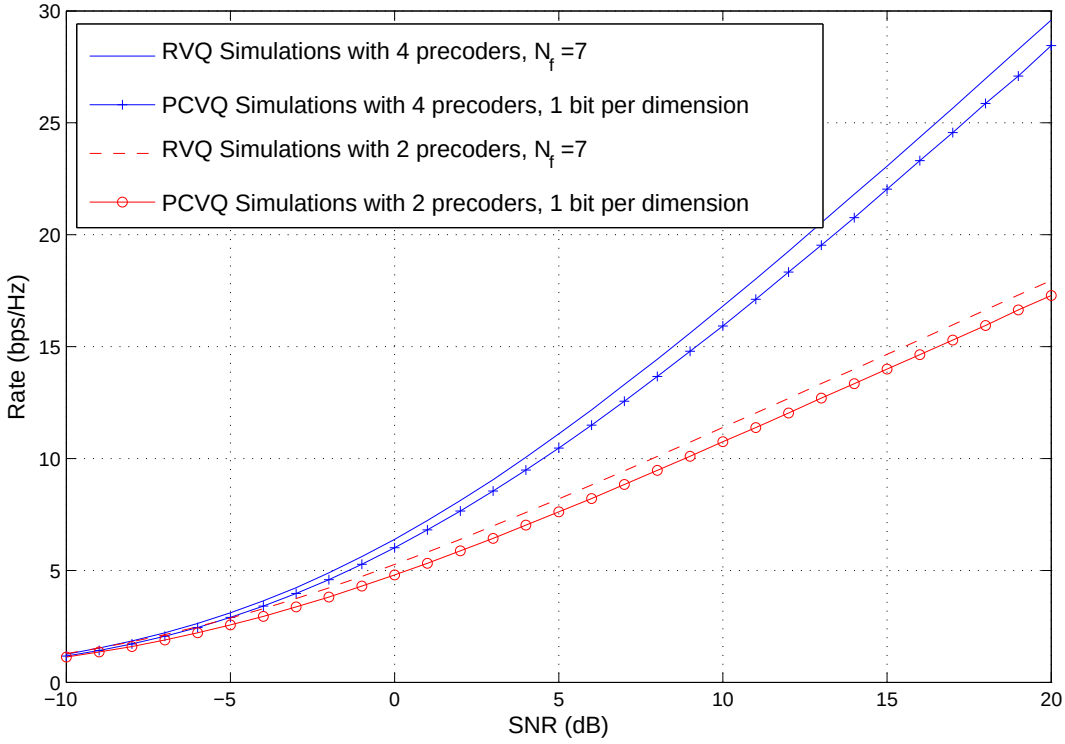

Figure 2.5: Average communication rates for a n-precoder RVQ and PCVQ 8×8 MIMO systems for $n = 2$ and 4 with $N_f = 7$ bits.

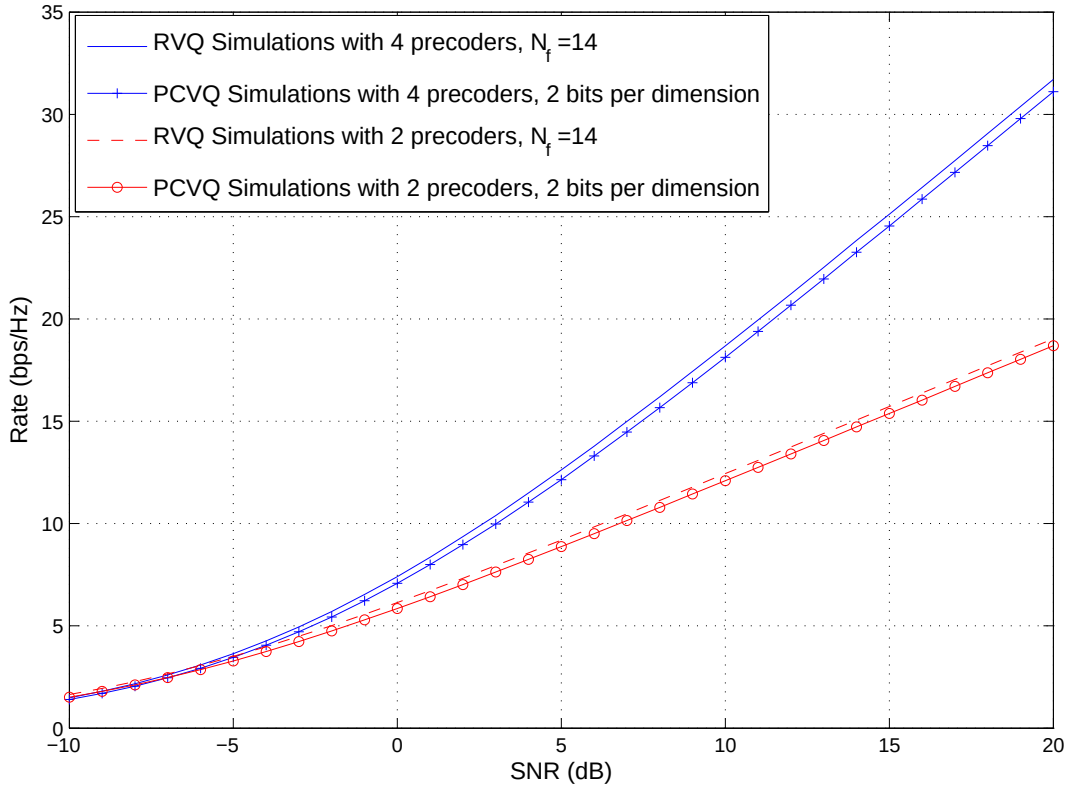

Figure 2.6: Average communication rates for a n-precoder RVQ and PCVQ 8×8 MIMO systems for $n = 2$ and 4 with $N_f = 14$ bits.

Actually, one does not need to operate Lloyd algorithm for the case of 2 bits per dimension since the distribution of the precoding vectors are isotropic, quantization points must be symetric around 0 and hence the optimal codebook is $[1; j; -1; -j]$, when 2 bits are allotted to each dimension in PCVQ. However, for the general case, the use of Lloyd algorithm is necessary to obtain optimum quantizer for PCVQ.

In order to evaluate the upper bound given in (2.16) for PCVQ, we need to know E_{11} and E_{12} values. E_{11} is found by simulation, i.e., by using different realizations of isotropically distributed n_t-dimensional unit norm vectors, a codebook is constructed by using Lloyd algorithm described above and 10,000 point Monte Carlo simulation gives the expectation of $|\mathbf{v}_1^* \mathbf{v}_{1f}|^2$. For practical system design, we only need to calculate E_{11} once, then we can use it as a design parameter. In Fig. 2.7, PCVQ MIMO system capacity bound obtained by using (2.16) is shown and it is very tight (0.1 dB different from the PCVQ MIMO system capacity). Also it is observed that PCVQ operates close to the upper capacity bound obtained by using bounding distribution given in (2.19) (within 0.6 dB).

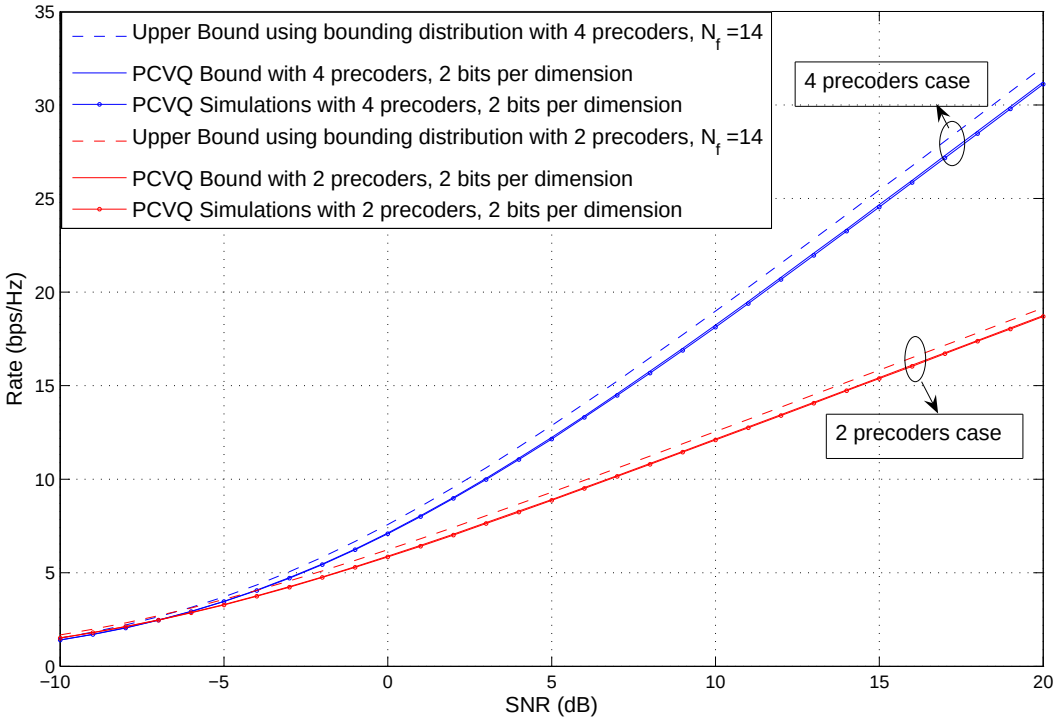

Figure 2.7: Average communication rates for a n-precoder 8×8 PCVQ MIMO system with its bound and the upper capacity bound for $n = 2$ and 4.

According to the results, we can say that the fine performance of random vector quantization can be matched even with very simple quantization methods such as PCVQ. Therefore, PCVQ can be quite helpful to achieve rates quite close to channel capacity in practice.

2.8 Conclusion

We developed a tight upper bound to point-to-point LRF MIMO capacity that is valid for a large set of vector quantization schemes. Using the upper bound developed, the number of precoders to be used in a practical system for any given value of average SNR can be determined. We furthermore evaluated the upper bound using a bounding distribution from Grassmannian beamforming which resulted in the observation that the simple RVQ technique performs quite close to capacity upper bound. Moreover, we proposed a simple quantization method in which each entry of a precoder is quantized independently. Along with its practicality, PCVQ performs very close to RVQ with regard to achieving rates close to capacity and hence poses a viable solution in enabling high rate MIMO communications systems. Perfect channel estimation at the receiver is assumed in this study. Future studies will include the consideration of imperfect channel estimation and its delayed transmission to the transmitter side within a limited rate feedback scenario in a mobile system. Also, practical quantization methods for MIMO systems will be investigated within the framework developed in this thesis.

CHAPTER 3

ITERATIVE DECISION FEEDBACK EQUALIZATION AND DECODING FOR ROTATED MULTIDIMENSIONAL CONSTELLATIONS IN BLOCK FADING CHANNELS

3.1 Introduction

The block-fading model is a useful model for transmission over slowly varying channels, such as orthogonal frequency division multiplexing (OFDM) or slow time-frequency-hopped systems [45]. In this chapter, we investigate block-fading channel with rotated constellations as example study and propose a practical receiver architecture having performance close to the capacity.

Rotated multidimensional constellations with uncoded modulation has been studied and shown to be an effective way to attain full-rate and full-diversity transmission in fading channels [9], [19], [31]. Even random multidimensional rotations are shown to exhibit good diversity distributions to combat channel fading for uncoded transmission in [31]. The problem of constructing general coded modulation schemes over multidimensional signal sets obtained by rotating classical complex-plane signal constellations has recently been studied in [17] for block fading channels with B fading blocks.

Despite the benefits of rotation over all B fading blocks, a large decoding complexity is imposed due to the inter-stream interference (ISI) caused by rotated constellations. A problem here is related to the complexity of optimum decoding, i.e., maximum likelihood (ML) receiver interfaces exhibit a complexity that grows exponentially as $O(|S|^B)$ with the modulation size $|S|$ and the dimension of rotation B, and becomes quickly unpractical when either parameter is large. In [31], a suboptimal MMSE equalizer with decision feedback is proposed and is shown to achieve good performance without destroying the high diversity order in the rotated constellation. In [49], sphere decoding is employed to avoid exhaustive

search over all candidate points. However, the structures in [31] and [49] were proposed for uncoded rotations. When coded modulation is used, the code trellis structure has to be incorporated and soft information should be provided to the decoder, which further complicates the problem. As a remedy to this problem, in [17], the use of rotations with a dimension smaller than the number of fading blocks was considered. The intuition behind this idea is that the channel code itself can help to achieve full diversity and sometimes rotations of a smaller dimension might be sufficient. However, for some rate values and constellation sizes, using rotations with small dimensions may not be sufficient to achieve the optimal rate-diversity tradeoff, i.e., the rotations of large dimensions might be necessary to attain full diversity order and the decoding complexity has still exponential dependence on the dimension of rotation. Soft-output sphere decoding technique for rotated constellation was proposed in [8], but it still shows some undesirable limitations in practice.

In this chapter, we propose an iterative receiver structure with reasonable complexity for coded modulation schemes with rotated constellations. The proposed detector is based on iterative forward and backward filtering followed by a channel decoder that works by using preliminary soft values of the coded symbols. Since the reliability of coded symbols from the decoding process are used in deriving the jointly optimal forward and backward filters, the filters employed in this work have a different structure from those of previous interference-cancellation based turbo equalizers, such as [7], [29], [51]. It has been observed that the proposed scheme yields a very close performance to the outage probability with reasonable complexity for rotated constellations. The benefits that rotation brings in terms of diversity exponent is justified without compromising the decoding complexity when compared to the optimal ML based structures with exponential complexity.

3.2 System Model

The following notation is used throughout Chapters 3 and 4. Boldface upper-case letters denote matrices and scalars are denoted by plain lower-case letters. The superscript $(\cdot)^*$ denotes the complex conjugate for scalars and $(\cdot)^H$ denotes the conjugate transpose for vectors and matrices. The $n \times n$ identity matrix is shown with $\mathbf{I}_n$. The autocorrelation matrix for a random vector $\mathbf{a}$ is $\mathbf{R_a} = E\{\mathbf{a}\mathbf{a}^H\}$ where $E\{\cdot\}$ stands for the expected value operator. The $(i, j)^{th}$ element of a matrix $\mathbf{A}$ is denoted by $\mathbf{A}(i, j)$ and the i^{th} element of a vector $\mathbf{a}$ is denoted

by a^i.

In this chapter, we consider block-based transmission as in [5], [7], [36]. During the transmission of one block, the channel is assumed to be constant and it changes independently from block to block. Without dealing with the channel estimation problem, the channel is assumed to be perfectly known at each block transmission.

Assuming symbol rate sampling, the discrete time baseband equivalent model of the point-to-point single-input single-output block fading channel with B fading blocks can be written as [16], [30],

$$\mathbf{y}_k = \mathbf{D}\mathbf{a}_k + \mathbf{n}_k, \qquad k = 0, 1, \ldots, N-1, \tag{3.1}$$

where N is the codeword length (block length) and $\mathbf{D}$ is a diagonal $B \times B$ matrix with main entries, d_i, $i = 1, \ldots, B$, i.e., $\mathbf{D} = \text{diag}(d_1, \ldots, d_B)$. $\mathbf{a}_k = [a_k^1, \ldots, a_k^B]^T$ is the portion of the transmitted codeword at time k and $\mathbf{y}_k = [y_k^1, \ldots, y_k^B]^T$ is the corresponding received vector at time k. The main diagonal entries of $\mathbf{D}$, d_i's, are the fading coefficients which are independent zero-mean circularly symmetric complex Gaussian (ZMCSCG) random variables with variance 1. The block fading model is considered and thus the channel matrices are assumed to be constant during a coherence interval significantly larger than a duration needed for the transmission of one block [22] and channel state information at transmitter (CSIT) is not available. Noise vectors $\mathbf{n}_k$ are also taken as ZMCSCG white (spatially and temporally) noise with variance N_0. The block-fading channel is a useful model for transmission over slowly varying channels, such as orthogonal frequency division multiplexing (OFDM) or slow time-frequency-hopped systems [45].

We consider that $\mathbf{a}_k$'s are obtained via the rotation of the symbols, i.e.,

$$\mathbf{a}_k = \mathbf{V}\mathbf{x}_k, \qquad k = 0, 1, \ldots, N-1, \tag{3.2}$$

where $\mathbf{x}_k = [x_k^1, \ldots, x_k^B]^T$ is the vector of complex-plane signal constellation symbols that is rotated by the $B \times B$ rotation matrix $\mathbf{V}$. The rotation matrix is unitary, i.e., $\mathbf{V}\mathbf{V}^H = \mathbf{I}_B$ and applied uniformly throughout a transmitted block.

The codewords $\mathbf{X} = [\mathbf{x}_0, \ldots, \mathbf{x}_{N-1}]$ form a coded modulation scheme $\chi \subset \mathbb{C}^{B \times N}$. In particular, we consider that χ is obtained as the concatenation of a binary code of rate r and a modulation over the signal constellation $S \in \mathbb{C}$ with $M = \log_2 |S|$ (See Fig. 3.1). The rate in bits per channel use of this scheme is $R = rM$. After the transmitted signal block has been

rotated, one obtains an equivalent model from (3.1) and (3.2) as in

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{n}_k, \qquad k = 0, 1, \ldots, N-1, \tag{3.3}$$

where $\mathbf{H} = \mathbf{DV}$. This form resembles the baseband equivalent form of a MIMO channel. Therefore, we will call our structure as space-time decoder hereafter and construct our receiver based on (3.3) in Section 3.3.

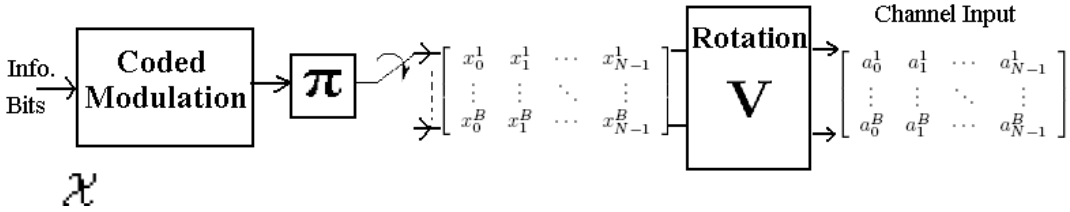

Figure 3.1: Block diagram for coded modulation with rotated constellations with rotation matrix **V** (Transmitter side).

Block diversity d_χ of a coded modulation scheme $\chi \in \mathbb{C}^{B\times N}$ consisting of B blocks of length N symbols from an alphabet S is defined as the minimum number of nonzero rows of $\mathbf{X}-\mathbf{X}'$ for any pair of codewords $\mathbf{X}' \neq \mathbf{X} \in \chi$. As it is shown in [30], this distance metric is the principle asymptotic indicator of pairwise-error-probability for any coding scheme and it determines the slope of the error-rate curve. When one views the N symbol transmitted in the same block as a super-symbol over S^N, d_χ is simply the Hamming distance in S^N for the non-binary block code with a block length B.

When no rotations are used, the optimal diversity reliability exponent is given by the Singleton bound for a given rate R (bits per channel use) and signal constellation M as [16], [30]

$$d_\chi^* = 1 + \left\lfloor B\left(1 - \frac{R}{M}\right)\right\rfloor \tag{3.4}$$

for B Rayleigh faded blocks. This value is an upper bound for the block diversity of any coded modulation scheme $\chi \subset \mathbb{C}^{B\times N}$ with rate R and constellation $S \in \mathbb{C}$ with $M = \log_2 |S|$. A code is block-wise maximum-distance separable (MDS) if it achieves the maximum diversity order given in (3.4) [16], [17]. As can be seen from (3.4), very large symbol alphabets may be required in order to achieve high asymptotic diversity for high rates ($R > 2$ bits/dim).

For example, with $B = 8$ and $R = 3$ bits/dim, a 16−point constellation can only achieve a diversity of $d^*_\chi = 3$. To achieve $d^*_\chi = 7$ a constellation with 4096 points is needed.

Different from the finite constellation case, the diversity order is not affected by the rate when codebooks from Gaussian alphabets are used [30]. Moreover, it is still possible to attain full diversity order without compromising rate or increasing constellation size by rotating traditional complex-plane signal constellations such as QAM over fading blocks. The idea of rotating a finite constellation is shown to increase the diversity order by spreading the information contained in each symbol over several independent fading blocks [31], and thus it can be seen as an effective way to combat channel fading. It was shown in [17] that the optimal diversity reliability exponent achieved by random Gaussian codes can also be achieved by random coded modulation schemes concatenated with a full-diversity rotation of dimension B when $R < M$. In this case, the optimal reliability exponent is given by

$$d^* = B \tag{3.5}$$

which is the available degrees of freedom in the channel. The B-dimensional rotation takes care of achieving full diversity while the coding gain is left to the outer coded modulation scheme over S and the MDS constraint on the code is relaxed for rotated schemes [17].

The diversity order of a coded modulation scheme $\chi \subset \mathbb{C}^{B\times N}$ with rotated constellations is the minimum number of nonzero rows of $\mathbf{V}(\mathbf{X} - \mathbf{X}')$ matrix for any pair of codewords $\mathbf{X}' \neq \mathbf{X} \in \chi$ and unitary rotation $\mathbf{V}$ from (3.2). If the rotation matrices $\mathbf{V}$ is full-diversity rotation, i.e., $\mathbf{V}$ satisfies

$$\mathbf{V}(\mathbf{s} - \mathbf{s}') \neq \mathbf{0}, \qquad \forall s, s' \in S^B, s \neq s' \tag{3.6}$$

componentwise like Krüskemper or cyclotomic rotations [17], all the rows of $\mathbf{V}(\mathbf{X} - \mathbf{X}')$ are nonzero for all distinct $\mathbf{X}, \mathbf{X}'$ pairs and the scheme achieves the optimal reliability exponent B. However, as it will be seen in Section 3.5, simple rotations like DFT which is not a full-diversity rotation may be sufficient to reach the optimal diversity order B in coded schemes, since the code itself helps achieve maximum reliability exponent. In other words, the optimal diversity order in (3.5) is achieved by a combination of coded modulation and rotation in this case.

3.3 Iterative Decision Feedback Equalization (DFE) for Rotated Constellations

Iterative equalization and decoding (also known as turbo equalization) was well studied in the literature for SISO ISI channels such as in [46], [47] and it was applied to multiuser and MIMO systems in [51], [29] and [43]. This technique yields much better performance than that of traditional equalization methods as shown in these papers. The equalizer and decoder exchange soft information in terms of likelihood values of the transmitted data iteratively to improve their performance in this case. The soft-in soft-out decoder produces likelihood information of each coded bit and it can be in the form of a convolutional, block or space-time trellis decoder depending on the encoding structure. The equalizer coefficients can be updated by using the likelihood information of transmitted data given by the decoder at each iteration.

We consider iterative space-time decoder with soft decision feedback in this chapter. Since both equalization and decoding processes can be performed at each iteration, turbo principle can be applied as done in [7], [29], [51]. In Fig. 3.2, an exemplary receiver structure is shown for iterative decision feedback equalizer (DFE).

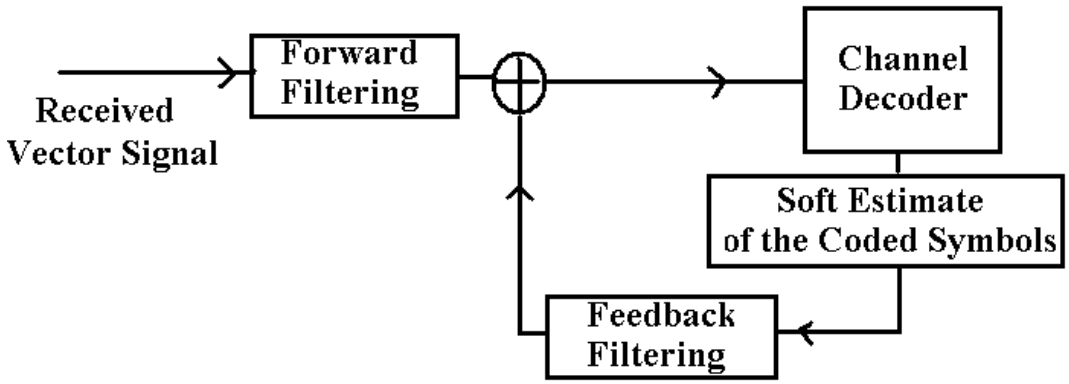

Figure 3.2: Iterative Decision Feedback Equalization (DFE) and decoding for rotated constellations

One can write the output from the DFE for the k^{th} vector in the block in the i^{th} iteration as

$$\tilde{\mathbf{x}}_k^{(i)} = (\mathbf{W}^{(i)})^H \mathbf{y}_k - (\mathbf{F}^{(i)})^H \hat{\mathbf{x}}_k^{(i-1)} \tag{3.7}$$

for $k = 0, \ldots, N-1$. $\mathbf{W}^{(i)}$'s and $\mathbf{F}^{(i)}$'s are forward and feedback filters both of size $B \times B$ respectively and $\hat{\mathbf{x}}_k^{(i-1)}$'s are soft decisions from the previous iteration. When the filters are designed based on the MMSE criterion and the information bearing signals are Gaussian, this structure is information theoretically optimum as stated in [23]. The first term in (3.7) is actually the feedforward estimate of the k^{th} transmitted vector. In (3.7), $\hat{\mathbf{x}}_k^{(i-1)}$'s are the soft feedback decisions from the previous iteration and they are utilized at the feedback filtering process to improve the estimate of $\mathbf{x}_k$. The forward and backward filter matrices are jointly optimized and found according to the MMSE criterion given by $E\left\{\sum_{k=0}^{N-1} \|\tilde{\mathbf{x}}_k^{(i)} - \mathbf{x}_k\|^2\right\}$ presented in [5], [36].

The n^{th} component of the estimation is not used in the feedback equalization of the n^{th} component of the received vector, and so we impose the following condition on the feedback filter

$$\mathbf{F}^{(i)}(n,n) = 0, \ n = 1, \ldots, B \tag{3.8}$$

since, by imposing this constraint, one can avoid self-subtraction of the desired symbol by its previous estimate. The joint optimization of the forward and backward filters at each iteration by taking the reliability of the decoded symbols used in feedback into account makes our proposed structure different from the traditional iterative equalizers. Also, the mitigation of inter-stream interference induced by the rotation is done optimally with this structure. This differs from previous studies which use spatial interference suppression techniques based on successive interference cancellation (SIC).

The Lagrange multiplier method can be used to obtain the optimal filter coefficients. Lagrangian vectors and the corresponding scalar constraints (Lagrangian function) can be written as

$$\mathbf{\Gamma}^{(i)} = \text{diag}\left[\Gamma_1^{(i)}, \ldots, \Gamma_B^{(i)}\right]_{(B\times B)}, \quad \textit{Lagrangian}(\mathbf{\Gamma}^{(i)}) = \sum_{n=1}^{B} (\mathbf{F}^{(i)}(n,n))^* \Gamma_n^{(i)}. \tag{3.9}$$

Due to an interleaving operation both in time and space, we can assume that

$$E\{\mathbf{x}_k(\mathbf{x}_l)^H\} = E_s \mathbf{I}_{n_t} \delta_{kl}, \ \text{for } k, l = 0, \ldots, N-1, \tag{3.10}$$

where δ_{kl} is the delta function which is 0 for all k but $k = l$. Some important correlation matrices used by the forward and feedback filters are defined for the i^{th} iteration as

$$\mathbf{P}^{(i)} = E\{\mathbf{x}_k(\hat{\mathbf{x}}_k^{(i-1)})^H\}, \quad \mathbf{B}^{(i)} = E\{\hat{\mathbf{x}}_k^{(i-1)}(\hat{\mathbf{x}}_k^{(i-1)})^H\} \tag{3.11}$$

for $k = 0, \ldots, N-1$. To simplify the computation of the filter coefficients, feedback decisions are assumed to be independent. Furthermore, due to the interleaving operation of the coded symbols, feedback decisions are assumed to be uncorrelated with the symbols transmitted at different block or symbol time. It is further assumed that the reliability matrices of the decision feedback are the same for all k, i.e.,

$$E\{\mathbf{x}_k(\hat{\mathbf{x}}_l^{(i-1)})^H\} = \mathbf{0}, \quad E\{\hat{\mathbf{x}}_k^{(i-1)}(\hat{\mathbf{x}}_l^{(i-1)})^H\} = \mathbf{0}, \text{ for } k \neq l \tag{3.12}$$

$$E\{x_k^m(\hat{x}_k^n)^*\} = \rho_m\delta_{mn}, \; E\{\hat{x}_k^m(\hat{x}_k^n)^*\} = \beta_m\delta_{mn} \tag{3.13}$$

for $m, n = 1, \ldots, B$ and the expectations are independent of symbol index k. Then, we can write

$$\mathbf{P}^{(i)} = \text{diag}\left[\rho_1, \ldots, \rho_B\right], \; \mathbf{B}^{(i)} = \text{diag}\left[\beta_1, \ldots, \beta_B\right]. \tag{3.14}$$

This assumption makes the forward and backward filters independent of time index k, so the block processing on each received signal can be implemented effectively. This can be achieved by simply averaging the correlations of soft feedback decisions from the previous iteration as will be done in Section 3.4. These are standard and reasonable assumptions as stated in [7], [36], [51], since the average symbol error probability is approximately the same for each symbol in a large block with quasi-static fading. Calculation of the correlation matrices $\mathbf{P}^{(i)}$ and $\mathbf{B}^{(i)}$ will be done in Section 3.4.

After taking the gradient of the MMSE cost function and the Lagrangian with respect to the rows of $(\mathbf{W}^{(i)})^H$ and $(\mathbf{F}^{(i)})^H$, equating the gradients to the zero vector, taking expectations and combining vectors into single matrix equations for $n = 1, \ldots, B$, one can obtain the following matrix equations giving the optimal forward and backward filter matrices

$$\mathbf{R}_\mathbf{y}\mathbf{W}^{(i)} = \mathbf{H}\left[E_s\mathbf{I}_B + \mathbf{P}^{(i)}\mathbf{F}^{(i)}\right] \tag{3.15}$$

$$\mathbf{B}^{(i)}\mathbf{F}^{(i)} = (\mathbf{P}^{(i)})^H\left[\mathbf{H}^H\mathbf{W}^{(i)} - \mathbf{I}_B\right] - \mathbf{\Gamma}^{(i)} \tag{3.16}$$

where

$$\mathbf{R}_\mathbf{y} = E\{\mathbf{y}_k(\mathbf{y}_k)^H\} = \left(\mathbf{H}\mathbf{H}^H E_s + N_0\mathbf{I}_B\right) \tag{3.17}$$

and $\mathbf{\Gamma}^{(i)}$ can be obtained from the constraint in (3.8). By substituting $\mathbf{W}^{(i)}$ into (3.16) and using the constraint, the Lagrangian terms given in (3.9) and backward filter matrices can be readily found after some calculations as

$$\Gamma_n^{(i)} = \frac{\left[\mathbf{A}^{(i)}(n,:)\mathbf{D}^{(i)}(:,n)\right]}{\mathbf{A}^{(i)}(n,n)},\ n = 1,\ldots,B \tag{3.18}$$

$$\mathbf{F}^{(i)} = \mathbf{A}^{(i)}\left[\mathbf{D}^{(i)} - \mathbf{\Gamma}^{(i)}\right], \tag{3.19}$$

where

$$\mathbf{A}^{(i)} = \left[\mathbf{B}^{(i)} - (\mathbf{P}^{(i)})^H\mathbf{H}^H\mathbf{R}_\mathbf{y}^{-1}\mathbf{H}\mathbf{P}^{(i)}\right]^{-1}, \tag{3.20}$$

$$\mathbf{D}^{(i)} = (\mathbf{P}^{(i)})^H\mathbf{H}^H\mathbf{R}_\mathbf{y}^{-1}\mathbf{H}E_s - (\mathbf{P}^{(i)})^H, \tag{3.21}$$

$\mathbf{A}^{(i)}(n,:)$ is the $n-$th row of $\mathbf{A}^{(i)}$, $\mathbf{D}^{(i)}(:,n)$ is the $n-$th column of $\mathbf{D}^{(i)}$ and forward filter $\mathbf{W}^{(i)}$ can be obtained from (3.15).

3.4 Iterative Decoding

In this section, we will calculate the log-likelihood ratios (LLR) and soft decisions of the coded symbols to be used in decision feedback. BPSK modulation is assumed for simplicity, but the extension to other M-ary or M-PSK modulations is straightforward. At each iteration, extrinsic information is extracted from detection and decoding stages and is then used as a priori information in the next iteration, just as in turbo decoding. The soft output from the DFE in the i^{th} iteration after (3.7) can be written as,

$$\tilde{x}_k^{m\,(i)} = \mu_m^{(i)} x_k^m + \eta_k^{m\,(i)} \tag{3.22}$$

for $k = 0,\ldots,N-1$ and $m = 1,\ldots,B$. In this case, the equalized channel in (3.22) can be considered as a quasi-parallelized channel and the LLR for the m^{th} component of the k^{th} transmitted symbol can be written as

$$L_k^{m\,(e)} = \log_e \frac{P(\tilde{x}_k^{m\,(i)}|x_k^m = +1)}{P(\tilde{x}_k^{m\,(i)}|x_k^m = -1)}. \tag{3.23}$$

The LLR term $L_k^{m\,(e)}$ is the extrinsic information that can be obtained from the equalizer output. An a priori probability ratio $L_k^{m\,(p)}$ ($\log_e \frac{P(x_k^m=+1)}{P(x_k^m=-1)}$) is given by the decoder as the intrinsic information obtained from the previous iteration [36], [51] and used to construct

a soft estimate of the coded symbol x_k^m. The extrinsic information given in (3.23) can be expressed as,

$$L_k^{m\,(e)} = \frac{4\mathrm{Re}\{(\mu_m^{(i)})^* \tilde{x}_k^{m\,(i)}\}}{E\{|\eta_k^m|^2\}} \tag{3.24}$$

by using the equivalent complex amplitude, $\mu_m^{(i)}$ of x_k^m at the output of the equalizer and the residual interference power, $E\{|\eta_k^{m\,(i)}|^2\}$. These values can be easily found in terms of channel matrices, forward and backward filter coefficients and correlation matrices as done for the SISO systems in [36], [51]. While computing the LLRs, we resort to simplification of the decoding algorithm by neglecting the correlation existing between the residual noise terms, i.e., the η_k^m's are taken as uncorrelated for $m = 1, \ldots, B$ as done in the decoding stage of [7] for flat fading MIMO channel and the residual interference is further approximated by a Gaussian distribution as in [36], [51]. It can be shown that $\mu_m^{(i)}$ and $E\{|\eta_k^{m\,(i)}|^2\}$ values do not depend on symbol time index k, so these values are calculated only once for the decoding of one block in each iteration, which reduces the complexity significantly.

Soft feedback decisions, $\hat{x}_k^m$ for the DFE equals $\tanh\left(\frac{1}{2}L_k^{m\,(p)}\right)$ for $E_s = 1$, $m = 1, \ldots, B$ and $k = 0, \ldots, N-1$ as in [7], [36], [51]. The non-zero diagonal entries of the correlation matrices $\mathbf{P}^{(i)}$ and $\mathbf{B}^{(i)}$ in (3.11) used by the forward and backward filters can be calculated by using the following approximation,

$$\rho_{k,m} \triangleq E\{x_k^m(\hat{x}_k^m)^*\} = E\{E\{x_k^m\}(\hat{x}_k^m)^*\} = |\hat{x}_k^m|^2 \tag{3.25}$$

$$\rho_m = \beta_m = \frac{1}{N}\sum_{k=0}^{N-1}\rho_{k,m}. \tag{3.26}$$

$E\{x_k^m\}$ was taken as $\hat{x}_k^m$ and this is a common assumption in various turbo detection techniques as done in [36], [51] and [46].

Correct estimation of $\mathbf{P}^{(i)}$ and $\mathbf{B}^{(i)}$'s is important, since our proposed DFE takes into account the reliability of the feedback decisions and therefore alleviates the error propagation problem different than the original DFE studies assuming perfect feedback decisions. In the first iteration, $\mathbf{P}^{(i)}$ and $\mathbf{B}^{(i)}$ can be taken as $\mathbf{0}_B$, i.e., reliable feedback decisions are not available. As the number of iterations increases, both metrics approach the asymptotic value: $E_s\mathbf{I}_B$.

The proposed decoding architecture posses the same level of complexity as the one for MIMO channels in [7]. A detailed complexity comparison between the iterative MMSE and ML-based receiver architectures was made in [7]. It was stated in [7] that although there may be variations between the complexity of two receivers in terms of required addition and

multiplications during equalization, decoding stages, soft-decision feedback computations; the overall complexity is moderately sensitive to these values and the most critical part that affects the complexity is the computation of LLR for each coded bit in the codeword before the decoding stages. It was evaluated that the ML receiver has $O(NMB|S|^B)$ complexity and the MMSE receiver has $O(NMB|S|)$ complexity [7]. Then, one can say that due to the exponential dependence of the complexity on the number of fading blocks B, the complexity of ML-based receiver can not be affordable for moderate values of constellation size $|S|$ and number of fading blocks B. However, the proposed MMSE based receiver here recovers from this problem by parallelizing the rotated channel and thus reduces the complexity of LLR computation significantly. Therefore, it shows linear dependence on $|S|$ and B. Even for moderate values of $|S|$ and B, the complexity difference between two schemes is tremendous. For example, let $|S| = 8$ (8-PSK modulation and $M = \log_2|S| = 3$) and $B = 6$, the complexity of ML-based receiver is approximately $\frac{|S|^B}{|S|} = \frac{8^6}{8} \approx 3.3 \times 10^4$ times larger than that of the MMSE receiver for the same block length N.

Our analysis and proposed decoding architecture are also valid for the schemes in which the rotations of dimension smaller than B are used. Rotation with smaller sizes was proposed in [17] in order to establish the tradeoff between the transmission rate, diversity, constellation size and complexity induced by the rotations. In this case, the decoding complexity is still exponential with the dimension of rotation, thus our proposed structure will be useful in reducing the complexity.

3.5 Simulation Results

3.5.1 Outage Probability Calculations

For sufficiently large block length N, the packet error probability of any coding scheme is lower bounded by the information outage probability [22]. In this section, we will compare the performance of our proposed decoding structure with the corresponding constrained outage probability of rotated and unrotated schemes. The constrained capacity can be found for the system model in (3.3) given the complex vector set χ of cardinality $|S|^B = (2^M)^B$ (e.g., M-ary or M-PSK modulations) similar to the derivations for block fading channels in [16]

and rotated schemes in [17] as

$$\begin{aligned} C^{\chi}_{rotated} &= \frac{1}{N}\sum_{k=0}^{N-1}\frac{1}{B}I(\mathbf{x}_k;\mathbf{y}_k|\mathbf{H}) \\ &= \log_2|S| - \frac{1}{B}E_{\mathbf{n}}\left\{\sum_{\mathbf{x}_k\in\chi}\frac{1}{|S|^B}\log_2\sum_{\mathbf{x}_i\in\chi}\exp\left(\frac{-\|\mathbf{H}(\mathbf{x}_k-\mathbf{x}_i)+\mathbf{n}\|^2+\|\mathbf{n}\|^2}{N_0}\right)\right\} \end{aligned} \quad (3.27)$$

where $\mathbf{n}$ is a ZMCSCG vector and the corresponding outage probability can be written as

$$P^{rotated,\,\chi}_{out}(R) = \mathbb{P}\left\{C^{\chi}_{rotated} < R\right\}. \quad (3.28)$$

Constrained outage probabilities will be used for performance evaluation in the next part.

3.5.2 Performance Results

In Fig. 3.3, simulation results are depicted for block fading channels with 3 fading blocks. Each block is Rayleigh faded with unity power. The packet error probability of rotated and unrotated systems with QPSK modulation and their corresponding outage probabilities are shown. Also, the outage probability for Gaussian input under average power constraint is drawn. A full block diversity attaining blockwise concatenated convolutional code (BCCC) based on bit-interleaved coded modulation (BICM) is used for encoding for both rotated and unrotated cases as adapted from [16]. The outer code is a rate-$\frac{1}{2}$ convolutional code and the inner codes are 3 trivial rate-1 accumulators. The information block length, i.e., the information bits entering the outer encoder is taken as 148 per frame and the rate in bits per channel use of this scheme is $R = rM = \frac{1}{2}2 = 1$. A DFT matrix with size 3 is used to rotate discrete QPSK inputs. Number of iterations inside the Turbo BCCC decoder is set to 10 and the number of equalizer iterations at which the forward and backward filters are updated by using the reliability matrices is taken as 3.

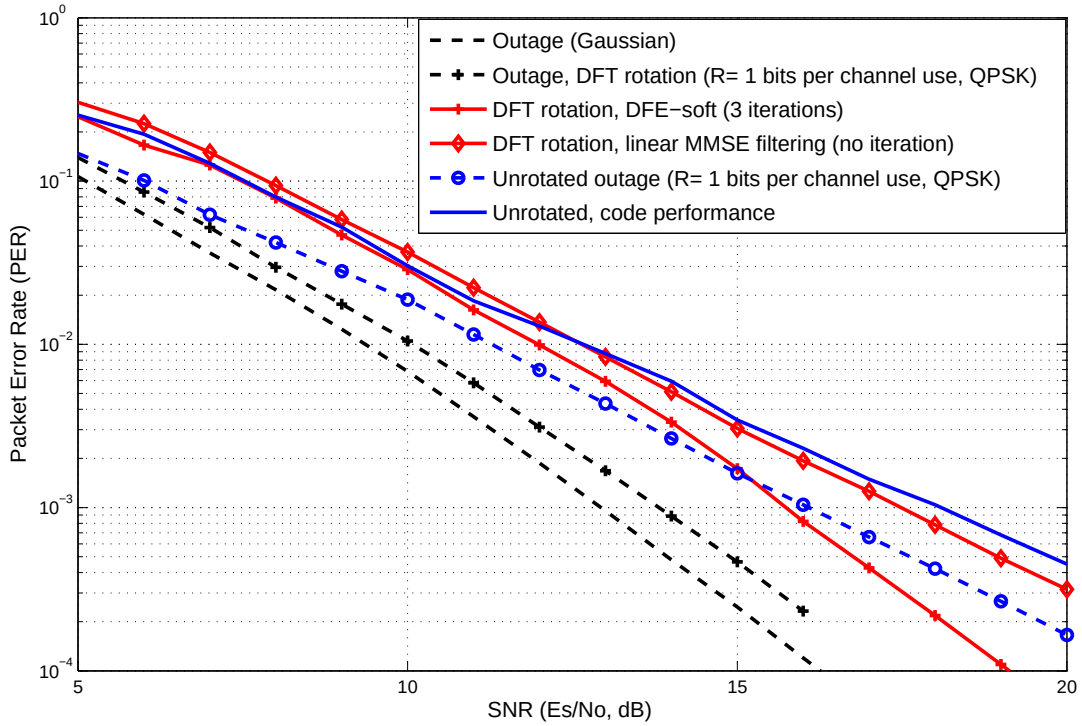

Figure 3.3: Performance comparison of iterative DFE and outage for rotated and unrotated constellations, $B = 3$

As it is seen from the outage probabilities, rotation enables to capture the largest possible reliability exponent achieved by Gaussian inputs, namely $d^* = B = 3$, while unrotated inputs have $d^*_\chi = 2$. It has been observed that there is approximately 2 dB difference between the outage probability with rotated inputs and the performance of decision feedback equalizer (DFE) with 3 iterations. This gap from the outage is similar to the gap between the outage and coded performance of unrotated inputs and iterative DFE has also similar performance gap from the outage with the ML-based receiver given in [17]. Then, one can say that the spatial interference and the error propagation problem inherent in decision feedback are almost eliminated and it is possible to attain optimal diversity of the block fading channel by using the proposed space-time equalizer. These results show that the theoretical benefit of rotation can be materialized by the proposed practical decoding structure with significantly reduced complexity. Moreover, it is seen that the simple DFT rotation is sufficient to attain the optimal diversity order in coded schemes since the code itself helps achieve full diversity different than the uncoded rotations in which the full diversity rotations are necessary to get the optimal exponent.

Furthermore, it is interesting to note that the performance improvement of the iterative DFE with soft feedback over the linear MMSE filtering without decision feedback is about 3 dB at PER=0.0001. There is also a loss in diversity as observed in the reduced PER slope without decision feedback. The suboptimality of linear equalizer prevents the system achieving high diversity orders. One can say that the proposed equalizer gains more diversity in comparison to the linear forward MMSE filtering by a careful design of both the forward and backward filters.

In Fig. 3.4, simulations are repeated for 6 fading blocks and outage probabilities are constructed for Gaussian inputs, BPSK inputs and rotated BPSK inputs with DFT rotation of size 6. The same BCCC structure is used with a rate-$\frac{1}{2}$ outer convolutional encoder and 6 inner rate-1 accumulators. The information block length, i.e., the number of information bits entering the encoder is taken as 238. Similar results to those in Fig. 3.3 are obtained and the optimal reliability exponent $d^* = 6$ is achieved by coded modulation scheme with simple DFT rotation, while unrotated inputs have $d^*_\chi = 4$ from the singleton bound. The optimal diversity order and a close performance to outage probability of rotated scheme at rate $R = 0.5$ bits per channel use within 2 dB are achieved by our practical decoding structure.

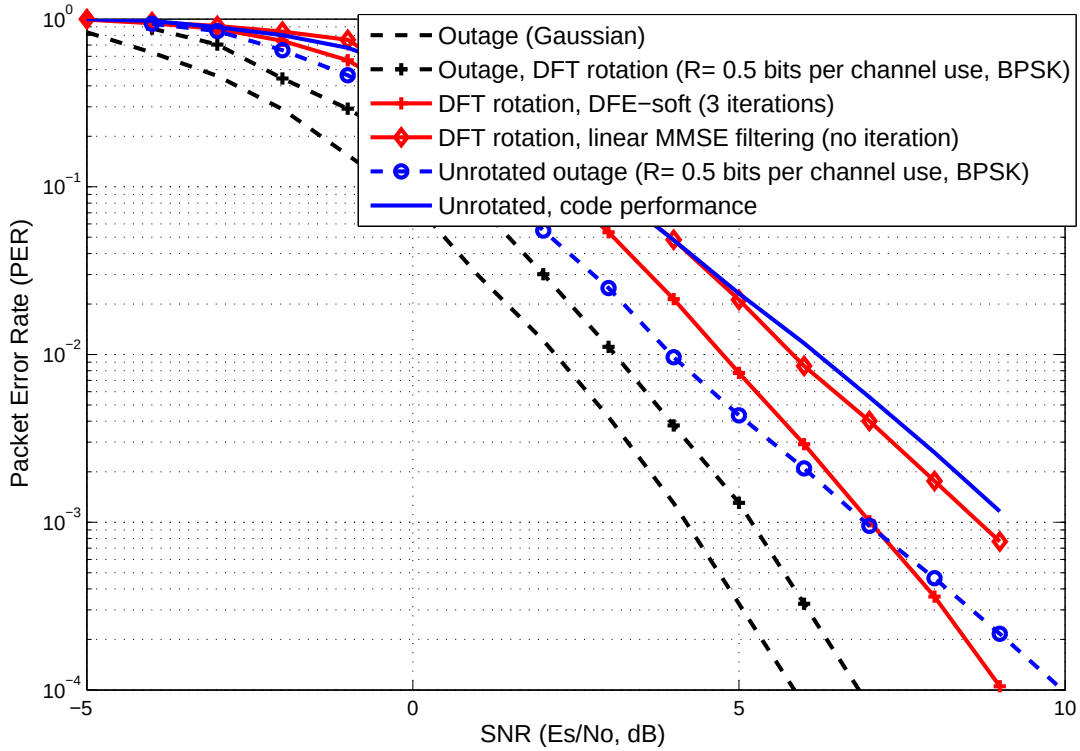

Figure 3.4: Performance comparison of iterative DFE and outage for rotated and unrotated constellations, $B = 6$

Fig. 3.5 shows the benefits of rotations by comparing the performance of the proposed iterative DFE for rotated QPSK inputs and unrotated code performances for 8 fading blocks. DFT rotation and BCCC structure with rate-$\frac{1}{2}$ outer convolutional encoder and 8 inner rate-1 accumulators are used. The information block length is taken as 318. The maximum diversity order, namely $d^* = 8$ is achieved by the iterative DFE with soft feedback, since the performance of iterative DFE shows the same slope as outage with Gaussian inputs, while the code performances with unrotated inputs can get $d_X^* = 5$. However, the gap between rotated and unrotated schemes may not be so significant at moderate PER values and even performance of the rotated scheme with the use of suboptimal non-iterative MMSE equalizer is below the performance of unrotated schemes. Therefore, for channels with large diversity order, one may not observe a considerable benefit of rotated constellations over some PER values. One has to be careful in choosing the decoding architecture, since the use of non-iterative suboptimal structures may destroy the high diversity benefits induced by rotated constellations due to residual spatial interference.

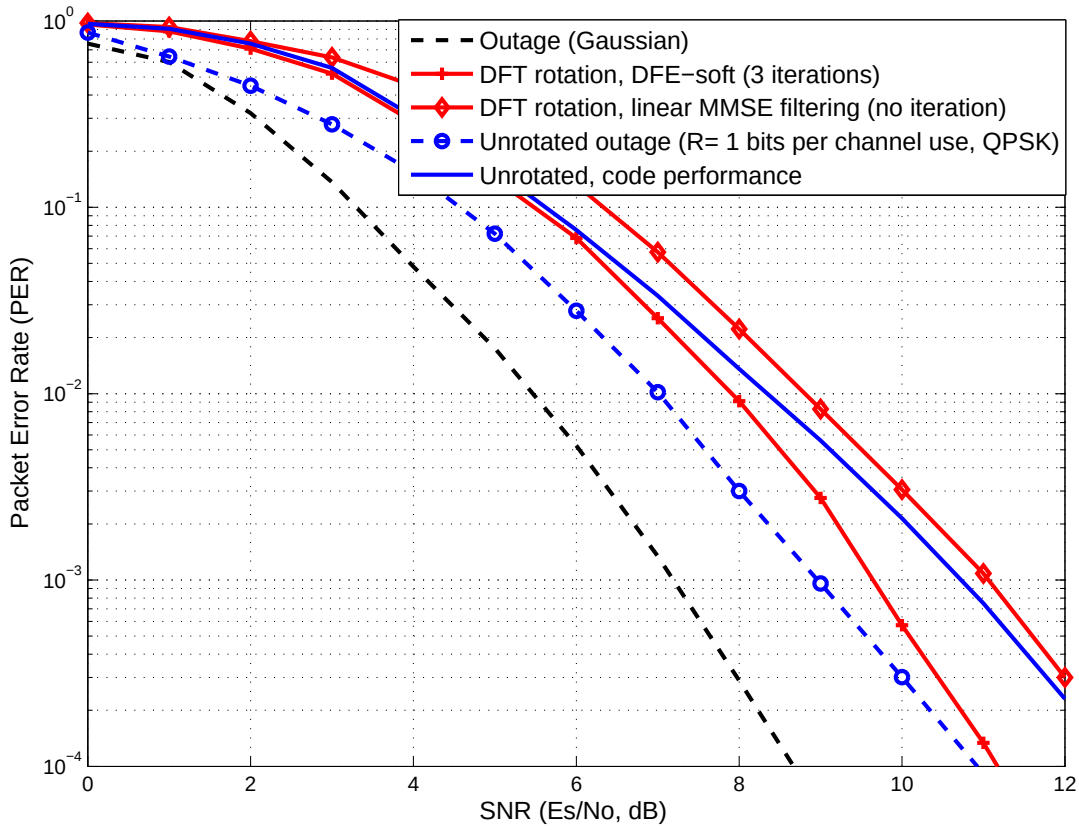

Figure 3.5: Performance comparison of iterative DFE and outage for rotated and unrotated constellations, $B = 8$

3.5.3 Diversity Comparison

In Table 3.1 and Table 3.2, the distribution of the diversity reliability exponent of the coded modulation schemes with rotated and unrotated constellations are shown for different number of fading blocks B and information block length K. Monte Carlo simulations were run, where 10^5 and 10^6 codewords are randomly generated among the 2^K different codewords and the number of nonzero rows of $\mathbf{V}(\mathbf{X}-\mathbf{X}_0)$ matrix, namely, block diversity, is recorded for each generated codeword. The BCCC encoding structure with rate$-\frac{1}{2}$ outer encoder and inner rate-1 accumulators with QPSK modulation are used again and $\mathbf{X}_0$ is the corresponding codeword of all zero information bits. The block length can be written as $N = \frac{1}{2}\left[\frac{2(K+2)}{B} + 1\right]$ including termination bits in the outer convolutional encoder and inner accumulators.

Table 3.1: Empirical diversity distribution for rotated and unrotated schemes for different information block length K values, number of fading blocks B is 4

$B = 4$				
diversity	1	2	3	4
$K = 12$				
DFT rotation and coded modulation	0	0	68	99932
Unrotated coded modulation	0	25	3060	96915
$K = 20$				
DFT rotation and coded modulation	0	0	31	999969
Unrotated coded modulation	0	5	2435	997560
$K = 32$				
DFT rotation and coded modulation	0	0	0	1000000
Unrotated coded modulation	0	0	31	999969

diversity	1	2	3	4
Uncoded DFT rotation	425	4300	18508	76767

Table 3.2: Empirical diversity distribution for rotated and unrotated schemes for different information block length K values, number of fading blocks B is 6

$B = 6$						
diversity	1	2	3	4	5	6
$K = 13$						
DFT rotation and coded modulation	0	0	0	20	93	99887
Unrotated coded modulation	0	0	52	1629	15290	83029
$K = 19$						
DFT rotation and coded modulation	0	0	0	0	69	999931
Unrotated coded modulation	0	2	48	1214	44403	954333
$K = 31$						
DFT rotation and coded modulation	0	0	0	0	0	1000000
Unrotated coded modulation	0	0	0	9	2940	997051

diversity	1	2	3	4	5	6
Uncoded DFT rotation	2	62	444	1741	7904	89847

Table 3.1 and 3.2 show the distribution of the number of nonzero rows indicating the diversity order of the given schemes. It is observed that there is a significant difference between the diversity order of rotated and unrotated schemes. For block fading channels, it is known that both the block diversity and the product distance between the codeword pairs affect the packet error probability, i.e., the pairwise error probability between the codeword matrices $\mathbf{X}_i$ and $\mathbf{X}_j$ is determined by the number of nonzero rows of $\mathbf{V}\left(\mathbf{X}_i - \mathbf{X}_j\right)$ and the squared

product distance between $\mathbf{X}_i$ and $\mathbf{X}_j$ [45]. Morover, it was observed experimentally that the codeword pairs with smaller diversity order also have poorer product distance values. Thus, one can say that the codeword pairs with minimum block diversity determines the packet error rate and maximizing diversity order is the best way to reduce the error probability on Rayleigh fading channels.

At first glance, it is seen that the number of codewords with smaller diversity order is diminished for unrotated case as the block length increases; however, the number of randomly generated codewords (10^6) during the simulations is very small compared to the actual number of codewords (2^K) for $K = 31$ and $K = 32$. Hence, there exist a lot of codeword pairs with smaller diversity gain and the product distance which cause significant performance degradation when compared to the rotated schemes achieving optimal diversity order B as confirmed by the simulation results.

Moreover, it is seen from the tables that DFT rotation shows poorer diversity distribution for uncoded transmissions, but it achieves the performance of full diversity rotations (Krüskemper) and the full reliability exponent B, when used with the coded modulation schemes, since the code itself helps rotation achieve the maximum diversity order. Therefore, we can say that both coded modulation and rotation cooperate to attain the maximum achievable diversity order in practical situations, where the full diversity rotations are not used, i.e., simple rotations like DFT or rotations of dimension smaller than B are used.

3.6 Conclusion

We have studied the block-fading channels with rotated signal constellations. Although rotated schemes can provide large diversity to combat fading, demodulation is prohibitive for large number of fading blocks and combined with coded modulations. We have proposed an iterative MMSE type decoding structure based on soft decision feedback without exponential complexity (with linear complexity). The proposed architecture shows a very close performance to the outage probability with rotated inputs and achieves the optimal diversity order attained by Gaussian inputs. Therefore, the theoretical benefit of rotated constellations is captured by the proposed structure with significantly reduced complexity.

CHAPTER 4

ITERATIVE FREQUENCY DOMAIN EQUALIZATION FOR SINGLE-CARRIER WIDEBAND MIMO CHANNELS

4.1 Introduction

In this chapter, we investigate wideband MIMO channel as example study and propose practical receiver structure in conjunction with single-carrier frequency domain equalization (SC-FDE) technique.

Multiple-input multiple-output (MIMO) systems have received much attention due to their multiplexing and diversity capabilities and potentially can offer tens of megabits per second transmission rates in future wireless systems. However, sophisticated equalization and decoding schemes are required for reliable communication at such high rates. While OFDM based schemes are well recognized candidates as a broadband wireless technology, single-carrier (SC) technology based on frequency domain equalization (FDE) has also started to gain considerable attention due its comparable complexity with OFDM. It has been shown in [18] that frequency domain equalization (FDE) can be readily applied to SC transmission to yield similar performance as OFDM. Since OFDM suffers from high peak-to-average power ratio (PAPR), SC techniques leading to more efficient use of power amplifiers are more suitable for uplink channels [18], [4]. It is known that OFDM and SC techniques are similar in terms of spectral efficiency and that OFDM only shifts the multipath fading problem from the time domain to the frequency domain [22]. Actually, OFDM breaks the frequency diversity and channel coding is needed to reclaim it as opposed to the SC-based systems, where multipath diversity can be attained without channel coding.

Due to the attractive features of SC-FDE, it has been viewed as a strong alternative to OFDM-based systems recently and its importance is clear for wideband channels. Block iterative FDE was proposed for uncoded single-input single-output (SISO) multipath chan-

nels in [5] and, block iterative FDE was considered in [36] together with channel decoding. Reliability metrics for uncoded and coded symbols are utilized in SISO systems to prevent the error propagation problem of decision feedback process in [5] and [36], respectively.

Iterative equalization schemes for wideband MIMO channels were considered in [29]. Authors consider minimum mean squared error (MMSE) type forward filtering and successive interference cancellation (SIC) to mitigate the interference in time domain. Turbo equalization with MMSE type filtering in frequency domain was studied in [53], but it does not consider the use of decision feedback filters or SIC operation and thus can not achieve the total multipath diversity gain of the channel as will be shown later. Recently, iterative frequency domain equalization techniques have been considered in [28] and [48]. They are both based on soft interference cancellation and MMSE forward filtering followed by a maximum a posteriori probability (MAP) detector. In [57], a hybrid equalization scheme is proposed, where forward filtering is performed in frequency domain, backward filtering is performed in time domain, and a further SIC operation is utilized to mitigate inter-stream interference. All these studies consider V-BLAST type architectures, where each stream transmitted over different antennas is coded separately without paying any regard to the possible diversity gains by careful coding across transmit antennas as in [30], [16] and D-BLAST based structures.

In this chapter, iterative FDE with decision feedback is applied to MIMO wideband systems. Actually, our analysis and results are not only valid for MIMO schemes but also can be applied to other multipath vector channels such as asynchronous multipath multi-user CDMA or systems which have the same discrete time model after sampling as MIMO schemes. Our work here is a generalization of the FDE technique from SISO to more general system models. A novel low-complexity iterative frequency domain equalizer utilizing soft frequency domain decision feedback from the multi-stream space-time MIMO decoder is proposed and the performance of this equalizer is shown to achieve the hypothetical matched filtering bound (MFB) performance [3] that upper bound the performance of any MIMO receiver. It is also shown that the proposed equalization scheme combined with the capacity achieving coding-multiplexing based techniques for parallel block fading channels [30], [16] has a close performance to the outage probability of the MIMO-OFDM scheme and the maximum diversity of the multipath vector channel can be attained by the proposed equalization and decoding structure.

Iterative frequency domain equalizers proposed in our work contain two separate filters, namely; the forward and the backward filters, which are jointly optimized in each iteration to minimize both the inter-symbol-interference (ISI) within streams and interference across streams. Since reliability of coded symbols from the decoding process are used in deriving the optimal forward and backward filters, the filters employed in this work have a different structure from that of previous interference-cancellation based MIMO turbo equalizers, such as [28], [48] and [57]. Our proposed equalization technique for multivariate ISI channels here is actually mathematically equivalent to the information theoretically optimum approach stated in [23]. Hence, the proposed structure is optimum if the channel symbols are drawn from a Gaussian alphabet and otherwise performs very close to the corresponding input alphabet constrained outage probability of MIMO-OFDM systems.

The contribution of the thesis in this chapter is threefold. We first show that SC-FDE with both forward and backward filters can be generalized from SISO to vector channels, which includes MIMO as a special case. We, furthermore, derive the jointly optimal forward and backward filters in the frequency domain so that the complexity advantage of FDE is not compromised. The error propagation problem in decision feedback is eliminated by taking the reliability of the decisions into account in each iteration. Third, the MIMO wideband channel can be quasi-parallelized with the help of our proposed space-time equalizer and so the code construction techniques achieving optimal rate-diversity tradeoff given by the singleton bound for block-fading channels [30], [16] can be effectively used such that the proposed equalization scheme combined with this type of coding structures yields a very close performance to the MIMO-OFDM outage probability. Therefore, we can say that the proposed SC-FDE based scheme could be a promising candidate for wideband MIMO systems as an alternative to MIMO-OFDM schemes.

4.2 System Model

In this chapter, we consider block based transmission as in [36] and [5]. During the transmission of one block, the channel is assumed to be constant and it changes independently from block to block. Without dealing with the channel estimation problem, the channel is assumed to be perfectly known at each block transmission. Cyclic prefix (CP) is used to prevent inter-block interference with length larger or equal to maximum channel length (L)

as explained in [22]. The signal for a transmitted block with CP is a sequence of vectors: $[\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{N-1}, \mathbf{x}_0, \ldots, \mathbf{x}_{L-1}]_{n_t \times (N+L)}$.

Assuming symbol rate sampling, the discrete time baseband equivalent model of the point-to-point MIMO wideband channel with n_r receive antennas and n_t transmit antennas can be written as [6],

$$\mathbf{y}_k = \sum_{l=0}^{L-1} \mathbf{H}_l \mathbf{x}_{k-l} + \mathbf{n}_k, \qquad k = 0, 1, \ldots, N-1, \tag{4.1}$$

where $\mathbf{H}_l$'s, $l = 0, \ldots, L-1$, are $n_r \times n_t$ complex channel matrices comprised of independent zero-mean circularly symmetric complex Gaussian (ZMCSCG) random variables with variance given by the power delay profile of each channel [37]. Block fading model is considered and thus the channel matrices are assumed to be constant during a coherence interval significantly larger than a duration needed for the transmission of one block [22] and channel state information at transmitter (CSIT) is not available. Noise vectors $\mathbf{n}_k$ are also taken as ZMCSCG white (spatially and temporally) noise with variance N_0. Only BPSK modulation is considered during the analysis and simulation studies. Extension to other M-ary or M-PSK modulations is straightforward.

If we define the DFT operation as $A_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} a_n e^{-j2\pi nk/N}$ for $k = 0, \ldots, N-1$, where a_n and A_k are the time domain sequence and its frequency domain sequence, respectively, then after the DFT operation to each element of $\mathbf{y}_k$ in (4.1), we can obtain the following expression in the frequency domain as done in [57]

$$\mathbf{Y}_k = \overline{\mathbf{\Lambda}}_k \mathbf{X}_k + \mathbf{N}_k, \qquad k = 0, \ldots, N-1 \tag{4.2}$$

where $\overline{\Lambda}_k$ is an $n_r \times n_t$ matrix representing the channel frequency response at the k^{th} tone with the entries [57]

$$\lambda_m^i(k) = \sum_{l=0}^{L-1} \mathbf{H}_l(i, m) e^{-j2\pi kl/N}, \tag{4.3}$$

for $i = 1, \ldots, n_r$ and $m = 1, \ldots, n_t$ and $\mathbf{H}_l(i, m)$ is a scalar and defined as the $(i, m)^{th}$ element of the channel matrix $\mathbf{H}_l$. In (4.2), $\mathbf{X}_k$ is the DFT of a vector sequence $\{\mathbf{x}_k\}$ with

$$\mathbf{X}_k = [X_k^1, \ldots, X_k^{n_t}]^T \tag{4.4}$$

for $k = 0, \ldots, N-1$ as adapted from [57]. X_k^i is the DFT of the sequence transmitted at the i^{th} antenna at the k^{th} frequency bin such that

$$X_k^i = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} x_l^i e^{-j2\pi kl/N}, \quad i = 1, \ldots, n_t. \tag{4.5}$$

Similarly, Y_k^i and N_k^i are the DFT of the corresponding received and noise sequences at the i^{th} receive antenna at the k^{th} frequency bin for $i = 1, \ldots, n_r$. The expression in (4.2) is the frequency domain equivalent of the channel in (4.1) and will be frequently used in the remainder of this chapter. As observed in (4.2), the channel gains of an OFDM system are converted from scalars in SISO to matrices in MIMO due to multiple transmit and receive antennas. Also, DFT operation is performed by using $q_n^m = \frac{1}{\sqrt{N}} e^{-j2\pi mn/N}$ for $m, n = 0, 1, \ldots, N-1$ hereafter.

4.3 Iterative Frequency Domain Equalization for Wideband MIMO channels

We consider iterative frequency domain equalization (FDE) with both time and frequency domain decision feedback in this chapter. Since both equalization and decoding processes can be performed in each iteration, the turbo principle can be applied as done in [36], [51]. According to the turbo principle, log-likelihood ratios (LLR) of the coded bits can be obtained from the channel equalizer and this information is used by the decoder. The decoder produces LLR of the coded symbols and the soft estimates of the coded symbols are constructed based on them to be used in feedback process at next iteration as explained in Chapter 3.

The decoding scheme based on the BCJR algorithm [2] with convolutional codes may be used as the decoding scheme for each substream and V-BLAST or D-BLAST type architectures on which MIMO systems are built can be used for transmission of each coded substream. V-BLAST and D-BLAST are widely used in MIMO schemes since they do not require CSIT and streams are separately demodulated by some sort of filtering such as MMSE and then decoded [22], [52], [20]. In Fig. 4.1, an exemplary receiver structure is shown for frequency domain decision feedback (FDDF) case.

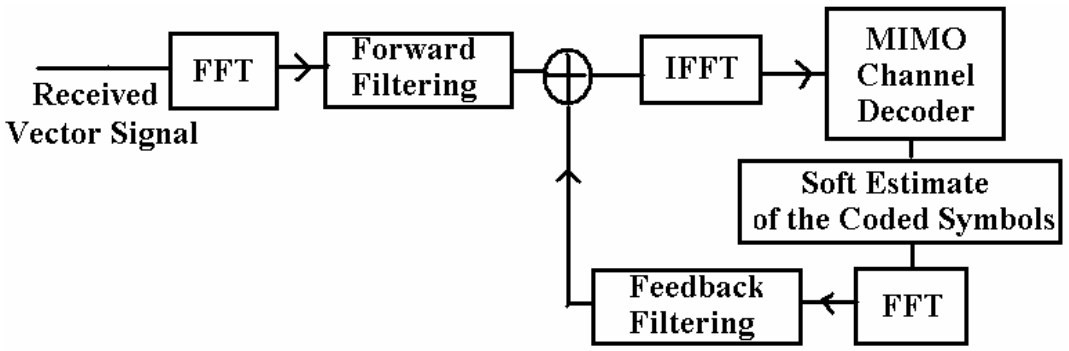

Figure 4.1: Iterative FDE with frequency domain decision feedback (FDDF)

As opposed to previous works in MIMO systems [28], [48] and [57], forward and feedback filters are jointly optimized in our approach to minimize both inter-symbol-interference (ISI) and interference from other streams. As it will be observed in Section 4.6, the combined multipath and space enriched diversity of the channel is exploited by decision feedback equalizer effectively such that the performance obtained by the matched filtering bound (MFB) is approximately achieved and a close performance to the outage of MIMO-OFDM can be obtained when the proposed equalization scheme is combined with coding structures that achieve the optimal rate-diversity tradeoff [16].

4.3.1 Frequency Domain Equalization with Time Domain Decision Feedback (FDE-TDDF)

To start with, $\mathbf{Y}_j$'s for $j = 0, \ldots, N-1$ can be easily shown to be uncorrelated from (4.2). We can further say that $\mathbf{Y}_j$ is a sufficient statistic to estimate $\mathbf{X}_j$ when it is Gaussian. As similar to the SISO case in [36], we can write the output from the FDE-TDDF for the k^{th} vector in the block in the i^{th} iteration as,

$$\tilde{\mathbf{x}}_k^{(i)} = \sum_{j=0}^{N-1} (q_j^k)^* (\mathbf{W}_j^{(i)})^H \mathbf{Y}_j - \sum_{j=0}^{N-1} (\mathbf{F}_j^{(i)})^H \hat{\mathbf{x}}_{(k+j) \bmod N}^{(i-1)} \tag{4.6}$$

for $k = 0, \ldots, N-1$. $\mathbf{W}_j^{(i)}$'s and $\mathbf{F}_j^{(i)}$'s are forward and feedback filters with sizes $n_r \times n_t$ and $n_t \times n_t$, respectively. When the filters are designed on the MMSE criterion, this structure is information theoretically optimum as stated in [23] when $\mathbf{x}_k$'s are Gaussian. The first summation is actually the feedforward estimate of the k^{th} transmitted vector. First, the feedforward equalization is done at the j^{th} frequency tone by multiplying an $n_t \times n_r$ matrix $(\mathbf{W}_j^{(i)})^H$ with the input vector $\mathbf{Y}_j$. Then, one can get the feedforward estimate at the time domain by converting the equalized frequency tones to the time domain with inverse DFT as done in the first summation in (4.6). In (4.6), $\hat{\mathbf{x}}_k^{(i-1)}$'s are the soft feedback decisions from the previous iteration and they are utilized at the feedback filtering process to improve the estimate of $\mathbf{x}_k$.

For the iterative FDE-TDDF operation, we consider feeding back the entire block of interfering vectors. Thus, there are N feedback filter matrices but it is possible to use only a number of L nonzero feedback filters to reduce computational complexity to find $\mathbf{F}_j^{(i)}$'s. On the contrary, our scheme here is operating on a block basis, therefore it cancels both the

pre-cursor and the post-cursor ISI while eliminating interference from other streams.

The forward and backward filter matrices are jointly optimized and found according to the MMSE criterion presented in [36], [5]. The total mean square error in one block, conditioned on the channel matrices and the results of the previous iteration, can be expressed as,

$$J = E\left\{\sum_{k=0}^{N-1} \|\tilde{\mathbf{x}}_k^{(i)} - \mathbf{x}_k\|^2\right\} = \sum_{k=0}^{N-1}\sum_{m=1}^{n_t} E\{(\tilde{x}_k^{m\,(i)} - x_k^m)(\tilde{x}_k^{m\,(i)} - x_k^m)^*\} \tag{4.7}$$

where $\tilde{x}_k^{m\,(i)}$ is the estimate of the k^{th} symbol transmitted at m^{th} antenna at i^{th} iteration.

The following constraint on backward filter $\mathbf{F}_j^{(i)}$ in (4.6) is necessary in order to avoid self-subtraction of the desired symbol by its previous estimate.

$$\mathbf{F}_0^{(i)}(n,n) = 0,\ n = 1,\ldots,n_t. \tag{4.8}$$

Since the signal streams are transmitted from the multiple transmit antennas at the same time and frequency, they introduce both multi-array interference (MAI), i.e., other antenna stream's spatial interference and inter symbol interference (ISI) in wideband MIMO communication. However, the mitigation of inter-stream interference originated from other antenna stream's spatial interference and ISI resulted from frequency selectivity is done optimally with this structure. This differs from previous MIMO studies which use spatial interference suppression techniques based on successive interference cancellation (SIC) in [28], [48] and [57]. In SIC based techniques, the decoded streams are subtracted from the received signal without any regard to their reliability and use of feedback filtering. Moreover, the joint update of the forward and backward filters at each iteration by using the reliability information given by the decoder makes our proposed structure different from the traditional iterative MIMO equalizers.

The Lagrange multiplier method can be used to obtain the optimal filter coefficients. Lagrangian vectors and the corresponding scalar constraints (Lagrangian function) can be written as

$$\mathbf{\Gamma}^{(i)} = \mathrm{diag}\left[\Gamma_1^{(i)},\ldots,\Gamma_{n_t}^{(i)}\right]_{(n_t\times n_t)},\quad Lagrangian(\mathbf{\Gamma}^{(i)}) = \sum_{n=1}^{n_t}(\mathbf{F}_0^{(i)}(n,n))^*\Gamma_n^{(i)} \tag{4.9}$$

By taking the gradient of the cost function and the Lagrangian with respect to rows of $(\mathbf{W}_j^{(i)})^H$, the following is obtained

$$\nabla_{(\mathbf{W}_l^{(i)}(n))^H} J = E\{\sum_{k=0}^{N-1}(q_l^k)^*\mathbf{Y}_l[\sum_{j=0}^{N-1}\mathbf{Y}_j^H\mathbf{W}_j^{(i)}(n)q_j^k - \sum_{j=0}^{N-1}(\hat{\mathbf{x}}_{(k+j)\ mod\ N}^{(i-1)})^H\mathbf{F}_j^{(i)}(n) - (x_k^n)^*]\} \tag{4.10}$$

for $l = 0, \ldots, N-1$ and $n = 1, \ldots, n_t$. When the gradient is taken with respect to rows of $(\mathbf{F}_j^{(i)})^H$,

$$\nabla_{(\mathbf{F}_l^{(i)}(n))^H} J = E\{\sum_{k=0}^{N-1} -\hat{\mathbf{x}}_{(k+l) \bmod N}^{(i-1)} [\sum_{j=0}^{N-1} \mathbf{Y}_j^H \mathbf{W}_j^{(i)}(n) q_j^k - \sum_{j=0}^{N-1} (\hat{\mathbf{x}}_{(k+j) \bmod N}^{(i-1)})^H \mathbf{F}_j^{(i)}(n) - (x_k^n)^*]\} + \sum_{k=0}^{N-1} \Gamma_n^{(i)} \mathbf{e}_n \delta_l \tag{4.11}$$

for $l = 0, \ldots, N-1$ and $n = 1, \ldots, n_t$ where $\mathbf{W}_j^{(i)}(n)$ and $\mathbf{F}_j^{(i)}(n)$ are the n^{th} column of $\mathbf{W}_j^{(i)}$ and $\mathbf{F}_j^{(i)}$ respectively. δ_l is the delta function which is 0 for all l but $l = 0$ and $\mathbf{e}_n$ denotes a n_t-dimensional vector of all zeros except for the n^{th} element which is 1.

We now define some important metrics used in equalization process. Similar definitions are made in Chapter 2, but we give the same definitions again here in order not to break the completeness of this chapter. Due to an interleaving operation both in time and space, we can assume that,

$$E\{\mathbf{x}_k(\mathbf{x}_l)^H\} = E_s \mathbf{I}_{n_t} \delta_{kl}, \text{ for } k, l = 0, \ldots, N-1. \tag{4.12}$$

Some important correlation matrices used by the forward and feedback filters are defined for the i^{th} iteration as

$$\mathbf{P}^{(i)} = E\{\mathbf{x}_k(\hat{\mathbf{x}}_k^{(i-1)})^H\}, \quad \mathbf{B}^{(i)} = E\{\hat{\mathbf{x}}_k^{(i-1)}(\hat{\mathbf{x}}_k^{(i-1)})^H\} \tag{4.13}$$

for $k = 0, \ldots, N-1$. To simplify the computation of the filter coefficients, feedback decisions are assumed to be independent. Furthermore, due to interleaving operation of the coded symbols, feedback decisions are assumed to be uncorrelated with the symbols transmitted at different antenna or symbol time. It is further assumed that the reliability matrices of the decision feedback are same for all k, i.e.,

$$E\{\mathbf{x}_k(\hat{\mathbf{x}}_l^{(i-1)})^H\} = \mathbf{0}, \quad E\{\hat{\mathbf{x}}_k^{(i-1)}(\hat{\mathbf{x}}_l^{(i-1)})^H\} = \mathbf{0}, \text{ for } k \neq l \tag{4.14}$$

$$E\{x_k^m(\hat{x}_k^n)^*\} = \rho_m \delta_{mn}, \; E\{\hat{x}_k^m(\hat{x}_k^n)^*\} = \beta_m \delta_{mn} \tag{4.15}$$

for $m, n = 1, \ldots, n_t$ and the expectations are independent of symbol index k. Then, we can write

$$\mathbf{P}^{(i)} = \text{diag}\,[\rho_1, \ldots, \rho_{n_t}], \; \mathbf{B}^{(i)} = \text{diag}\,[\beta_1, \ldots, \beta_{n_t}]. \tag{4.16}$$

These are standard and reasonable assumptions as stated in [36], [51], since the average symbol error probability is approximately the same for each symbol in a large block with quasi-static fading. This approximation should not lead to a significant loss in performance.

The vector filters $\mathbf{W}_j^{(i)}$ and $\mathbf{F}_j^{(i)}$ can be calculated by equating the gradients to zero vectors in (4.10) and (4.11). Expectations required to find filter coefficients can be calculated easily from vector DFT operation and (4.2) as follows,

$$E\{\mathbf{X}_k(\mathbf{X}_l)^H\} = E\{\mathbf{x}_k(\mathbf{x}_l)^H\} = E_s\mathbf{I}_{n_t}\delta_{kl} \tag{4.17}$$

$$E\{\mathbf{N}_k(\mathbf{N}_l)^H\} = E\{\mathbf{n}_k(\mathbf{n}_l)^H\} = N_0\mathbf{I}_{n_r}\delta_{kl} \tag{4.18}$$

$$E\{\mathbf{Y}_k(\mathbf{Y}_l)^H\} = \left(\overline{\mathbf{\Lambda}}_k\overline{\mathbf{\Lambda}}_k^H E_s + N_0\mathbf{I}_{n_r}\right)\delta_{kl} = \mathbf{R}_{\mathbf{Y}_k}\delta_{kl} \tag{4.19}$$

$$E\{\mathbf{Y}_l(\hat{\mathbf{x}}_{(k+j)\ mod\ N}^{(i-1)})^H\} = \overline{\mathbf{\Lambda}}_l\mathbf{P}^{(i)}q_{(k+j)\ mod\ N}^l \tag{4.20}$$

$$E\{\mathbf{Y}_l(\mathbf{x}_k)^H\} = \overline{\mathbf{\Lambda}}_l E_s q_k^l \quad . \tag{4.21}$$

After equating the gradients to zero vector, taking expectations by using the above equations and combining vectors into single matrix equations for $n = 1, \ldots, n_t$, one can obtain the following matrix equations providing the optimal forward and backward filter matrices

$$\mathbf{R}_{\mathbf{Y}_j}\mathbf{W}_j^{(i)} = \overline{\mathbf{\Lambda}}_j\left[E_s\mathbf{I}_{n_t} + \sqrt{N}\mathbf{P}^{(i)}\sum_{m=0}^{N-1} q_m^j\mathbf{F}_m^{(i)}\right] \tag{4.22}$$

for $j = 0, \ldots, N-1$,

$$\mathbf{B}^{(i)}\mathbf{F}_l^{(i)} = \frac{1}{\sqrt{N}}\sum_{m=0}^{N-1}(q_l^m)^*(\mathbf{P}^{(i)})^H\overline{\mathbf{\Lambda}}_m^H\mathbf{W}_m^{(i)} \tag{4.23}$$

for $l = 1, \ldots, N-1$ and

$$\mathbf{B}^{(i)}\mathbf{F}_0^{(i)} = \frac{1}{\sqrt{N}}\sum_{m=0}^{N-1}(q_0^m)^*(\mathbf{P}^{(i)})^H\overline{\mathbf{\Lambda}}_m^H\mathbf{W}_m^{(i)} - (\mathbf{P}^{(i)})^H - \mathbf{\Gamma}^{(i)}. \tag{4.24}$$

By substituting $\mathbf{W}_m^{(i)}$'s into (4.23) and (4.24), the backward filter matrices can be readily found after some manipulations. After some calculus, one obtains the following matrix equations and $\mathbf{F}_l^{(i)}$'s can be found from the solution of the given equations.

$$\mathbf{B}^{(i)}\mathbf{F}_l^{(i)} - \sum_{n=0}^{N-1}\mathbf{T}(l,n)\mathbf{F}_n^{(i)} = \mathbf{V}(l),\ l = 1, \ldots, N-1 \tag{4.25}$$

$$\mathbf{B}^{(i)}\mathbf{F}_0^{(i)} - \sum_{n=0}^{N-1}\mathbf{T}(0,n)\mathbf{F}_n^{(i)} = \mathbf{V}(0) - (\mathbf{P}^{(i)})^H - \mathbf{\Gamma}^{(i)} \tag{4.26}$$

where

$$\mathbf{T}(l,n) = \sum_{m=0}^{N-1}(q_l^m)^*(\mathbf{P}^{(i)})^H\overline{\mathbf{\Lambda}}_m^H\mathbf{R}_{\mathbf{Y}_m}^{-1}\overline{\mathbf{\Lambda}}_m\mathbf{P}^{(i)}q_n^m, \tag{4.27}$$

$$\mathbf{V}(l) = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} (q_l^m)^* (\mathbf{P}^{(i)})^H \overline{\mathbf{\Lambda}}_m^H \mathbf{R}_{\mathbf{Y}_m}^{-1} \overline{\mathbf{\Lambda}}_m E_s \tag{4.28}$$

for $l, n = 0, \ldots, N-1$. Optimal feedback filter matrices, $\mathbf{F}_j^{(i)}$'s can be found as a solution of the above equations by using the constraint given in (4.8) and forward filters, $\mathbf{W}_j^{(i)}$'s can be obtained from (4.22).

As seen in (4.25) and (4.26), the computation of feedback matrices requires the inversion of a Hermitian block Toeplitz matrix with size $Nn_t \times Nn_t$ similar to the SISO cases in [36]. However, this matrix inversion burden for time domain decision feedback filters will be significantly reduced when we switch to the frequency domain decision feedback case and thus, the computational simplicity of SC FDE technique will be observed.

Soft feedback decisions of the coded symbols can be obtained by using the information given by the decoder. Using these soft decisions, it is possible to approximate correlation matrices $\mathbf{P}^{(i)}$ and $\mathbf{B}^{(i)}$ as done for the SISO case in [36]. Correct estimation of $\mathbf{P}^{(i)}$ and $\mathbf{B}^{(i)}$'s is important, since FDE-TDDF and FDE-FDDF take into account the reliability of the feedback decisions and therefore alleviates the error propagation problem different than the original FDE studies in [18] and [4] assuming perfect feedback decisions. In the first iteration, $\mathbf{P}^{(i)}$ and $\mathbf{B}^{(i)}$ can be taken as $\mathbf{0}_{n_t}$, i.e, reliable feedback decisions are not available. As the number of iterations increases, both metrics approach the asymptotic value: $E_s \mathbf{I}_{n_t}$. Calculation of these correlation matrices will be done in Section 4.4.

4.3.2 Frequency Domain Equalization with Frequency Domain Decision Feedback (FDE-FDDF)

The iterative frequency domain equalizer with hard and soft decision feedback in the frequency domain is studied in [4], [5] and [36] for the SISO systems. We derive the filter matrices based on the MMSE criterion like the FDE-TDDF case. Since, FDE-TDDF described in the previous section and FDE-FDDF are both based on the same MMSE criterion in the time domain, both structures are actually equivalent. Output of the FDE-FDDF for the k^{th} vector in the block (for the i^{th} iteration) can be expressed as,

$$\tilde{\mathbf{x}}_k^{(i)} = \sum_{j=0}^{N-1} (q_j^k)^* \left[(\mathbf{W}_j^{(i)})^H \mathbf{Y}_j - (\mathbf{C}_j^{(i)})^H \hat{\mathbf{X}}_j^{(i-1)} \right] \tag{4.29}$$

for $k = 0, \ldots, N-1$. $\mathbf{W}_j^{(i)}$'s and $\mathbf{C}_j^{(i)}$'s are forward and feedback filters both in frequency domain with sizes $n_r \times n_t$ and $n_t \times n_t$ respectively and $\hat{\mathbf{X}}_j^{(i-1)}$'s are the DFT's of soft decisions from the previous iteration. Since, the proposed FDE-TDDF and FDE-FDDF structures are equivalent, one can find a relation between the time domain feedback filters $\mathbf{F}_j^{(i)}$ and the frequency domain feedback filters $\mathbf{C}_j^{(i)}$. It can be shown that

$$(\mathbf{F}_k^{(i)})^*(m,n) = \sum_{l=0}^{N-1} (\mathbf{C}_l^{(i)})^*(m,n) e^{j2\pi kl/N} \tag{4.30}$$

for $m, n = 1, \ldots, n_t$ and $k = 0, \ldots, N-1$ where $(m,n)^{th}$ element of $(\mathbf{C}_j^{(i)})^H$ and $(\mathbf{F}_j^{(i)})^H$ are defined as $(\mathbf{C}_j^{(i)})^*(m,n)$ and $(\mathbf{F}_j^{(i)})^*(m,n)$ respectively. Since the optimization problem for FDE-FDDF case is mathematically equivalent to FDE-TDDF (Section 4.3.1) with the constraint $\mathbf{F}_j^{(i)}(n,n) = 0, \ n = 1, \ldots, n_t$, we can set the constraint for frequency domain feedback filters from (4.30) as

$$(\mathbf{F}_0^{(i)})^*(n,n) = \sum_{l=0}^{N-1} (\mathbf{C}_l^{(i)})^*(n,n) = 0, \ n = 1, \ldots, n_t \tag{4.31}$$

With this constraint, one can avoid self-subtraction of the desired symbol by its previous estimate. The Lagrange multiplier method can be used once again to obtain optimal forward and backward frequency domain filters. Lagrangian vectors and the corresponding scalar constraints (Lagrangian function) can be written as

$$\mathbf{\Gamma}^{(i)} = \text{diag}\left[\Gamma_1^{(i)}, \ldots, \Gamma_{n_t}^{(i)}\right]_{(n_t \times n_t)}, \quad Lagrangian(\mathbf{\Gamma}^{(i)}) = \sum_{n=1}^{n_t} \sum_{j=0}^{N-1} (\mathbf{C}_j^{(i)}(n,n))^* \Gamma_n^{(i)}. \tag{4.32}$$

By taking the gradient of the cost function and the Lagrangian with respect to the rows of $(\mathbf{W}_j^{(i)})^H$ and $(\mathbf{C}_j^{(i)})^H$, the following are obtained

$$\nabla_{(\mathbf{W}_l^{(i)}(n))^H} J = E\{\sum_{k=0}^{N-1} (q_l^k)^* \mathbf{Y}_l [\sum_{j=0}^{N-1} (\mathbf{Y}_j^H \mathbf{W}_j^{(i)}(n) q_j^k - (\hat{\mathbf{X}}_j^{(i-1)})^H \mathbf{C}_j^{(i)}(n) q_j^k) - (x_k^n)^*]\} \tag{4.33}$$

$$\nabla_{(\mathbf{C}_l^{(i)}(n))^H} J = E\{\sum_{k=0}^{N-1} -(q_l^k)^* \hat{\mathbf{X}}_l^{(i-1)} [\sum_{j=0}^{N-1} (\mathbf{Y}_j^H \mathbf{W}_j^{(i)}(n) q_j^k - (\hat{\mathbf{X}}_j^{(i-1)})^H \mathbf{C}_j^{(i)}(n) q_j^k) - (x_k^n)^*]\} + \sum_{k=0}^{N-1} \Gamma_n^{(i)} \mathbf{e}_n \tag{4.34}$$

for $l = 0, \ldots, N-1$ and $n = 1, \ldots, n_t$ where $\mathbf{W}_j^{(i)}(n)$ and $\mathbf{C}_j^{(i)}(n)$ are the n^{th} column of $\mathbf{W}_j^{(i)}$ and $\mathbf{C}_j^{(i)}$ respectively. Expectations required to find filter coefficients can be calculated easily from vector DFT operation and (4.2) as follows

$$E\{\mathbf{Y}_l (\hat{\mathbf{X}}_k^{(i-1)})^H\} = \overline{\mathbf{\Lambda}}_l \mathbf{P}^{(i)} \delta_{kl} \tag{4.35}$$

$$E\{\hat{\mathbf{X}}_k^{(i-1)} (\hat{\mathbf{X}}_l^{(i-1)})^H\} = \mathbf{B}^{(i)} \delta_{kl} \tag{4.36}$$

$$E\{\hat{\mathbf{X}}_k^{(i-1)} (\mathbf{x}_l)^H\} = q_l^k (\mathbf{P}^{(i)})^H \quad . \tag{4.37}$$

After equating the gradients to zero vector, taking expectations by using the equations above and combining vectors into single matrix equations for $n = 1, \ldots, n_t$, one can obtain the following matrix equations giving out the optimal forward and backward filter matrices in the frequency domain

$$\mathbf{R}_{\mathbf{Y}_j}\mathbf{W}_j^{(i)} = \overline{\mathbf{\Lambda}}_j\left[E_s\mathbf{I}_{n_t} + \mathbf{P}^{(i)}\mathbf{C}_j^{(i)}\right] \tag{4.38}$$

$$\mathbf{B}^{(i)}\mathbf{C}_j^{(i)} = (\mathbf{P}^{(i)})^H\left[\overline{\mathbf{\Lambda}}_j^H\mathbf{W}_j^{(i)} - \mathbf{I}_{n_t}\right] - \mathbf{\Gamma}^{(i)} \tag{4.39}$$

for $j = 0, \ldots, N-1$, and $\mathbf{\Gamma}^{(i)}$ can be obtained from the constraint:

$$\sum_{j=0}^{N-1}\mathbf{C}_j^{(i)}(n,n) = 0,\ n = 1, \ldots, n_t \tag{4.40}$$

By substituting $\mathbf{W}_j^{(i)}$'s into (4.39) and using the constraint, the Lagrangian terms given in (4.32) and backward filter matrices can be readily found after some calculus as,

$$\Gamma_n^{(i)} = \frac{\left[\sum_{j=0}^{N-1}\mathbf{A}_j^{(i)}(n,:)\mathbf{D}_j^{(i)}(:,n)\right]}{\left[\sum_{j=0}^{N-1}\mathbf{A}_j^{(i)}(n,n)\right]},\ n = 1, \ldots, n_t \tag{4.41}$$

$$\mathbf{C}_j^{(i)} = \mathbf{A}_j^{(i)}\left[\mathbf{D}_j^{(i)} - \mathbf{\Gamma}^{(i)}\right], \tag{4.42}$$

where

$$\mathbf{A}_j^{(i)} = \left[\mathbf{B}^{(i)} - (\mathbf{P}^{(i)})^H\overline{\mathbf{\Lambda}}_j^H\mathbf{R}_{\mathbf{Y}_j}^{-1}\overline{\mathbf{\Lambda}}_j\mathbf{P}^{(i)}\right]^{-1}, \tag{4.43}$$

$$\mathbf{D}_j^{(i)} = (\mathbf{P}^{(i)})^H\overline{\mathbf{\Lambda}}_j^H\mathbf{R}_{\mathbf{Y}_j}^{-1}\overline{\mathbf{\Lambda}}_j E_s - (\mathbf{P}^{(i)})^H, \tag{4.44}$$

$\mathbf{A}_j^{(i)}(n,:)$ is the $n-$th row of $\mathbf{A}_j^{(i)}$, $\mathbf{D}_j^{(i)}(:,n)$ is the $n-$th column of $\mathbf{D}_j^{(i)}$ and $\mathbf{W}_j^{(i)}$'s are obtained from (4.38) for $j = 0, \ldots, N-1$.

It is seen that the computational complexity to obtain forward and feedback filters is considerably reduced for SC FDE-FDDF case in comparison to FDE-TDDF, since only $n_r \times n_r$ and $n_t \times n_t$ matrix inversions are needed as can be seen from (4.41)-(4.44) and sizes of these matrices are independent of the block length (N) like OFDM based systems. Therefore, the complexity of SC FDE-FDDF technique for MIMO wideband channels is comparable to MIMO-OFDM systems.

4.4 Iterative Decoding

In this section, we will calculate the log-likelihood ratios (LLR) and soft decisions of the coded symbols for FDE with frequency domain decision feedback (FDDF). BPSK modulation is assumed for simplicity, but the extension to other M-ary or M-PSK modulations is

straightforward in principle. At each iteration, extrinsic information is extracted from detection and decoding stages and is then used as a priori information in the next iteration, just as in turbo decoding. The soft output from the FDE-FDDF in the i^{th} iteration after (4.29) can be written as,

$$\tilde{x}_k^{m\,(i)} = \mu_m^{(i)} x_k^m + \eta_k^{m\,(i)} \tag{4.45}$$

for $k = 0, \ldots, N-1$ and $m = 1, \ldots, n_t$. In this case, the equalized MIMO channel in (4.45) can be considered as a quasi-parallelized channel and the LLR for the k^{th} symbol transmitted at m^{th} antenna can be written as

$$L_k^{m\,(e)} = \log_e \frac{P(\tilde{x}_k^{m\,(i)} | x_k^m = +1)}{P(\tilde{x}_k^{m\,(i)} | x_k^m = -1)}. \tag{4.46}$$

The LLR term $L_k^{m\,(e)}$ is the extrinsic information that can be obtained from the equalizer output. An a-priori probability ratio $L_k^{m\,(p)}$ ($\log_e \frac{P(x_k^m=+1)}{P(x_k^m=-1)}$) is given by the decoder as the intrinsic information obtained from the previous iteration [36], [51] and used to construct a soft estimate of the coded symbol transmitted at m^{th} antenna for k^{th} vector.

The equivalent complex amplitude, $\mu_m^{(i)}$ of the symbol transmitted from the m^{th} antenna at the output of the equalizer and the residual interference power, $E\{|\eta_k^{m\,(i)}|^2\}$ can be computed by using (4.29) as follows,

$$\begin{aligned} \mu_m^{(i)} &= E\{\tilde{x}_k^{m\,(i)} (x_k^m)^*\}/E_s = \sum_{j=0}^{N-1} \frac{1}{N} [(\mathbf{W}_j^{(i)}(m))^H \overline{\mathbf{\Lambda}}_j - (\mathbf{C}_j^{(i)}(m))^H (\mathbf{P}^{(i)})^H] \mathbf{e}_m \\ &= \sum_{j=0}^{N-1} \frac{1}{N} (\mathbf{W}_j^{(i)}(m))^H \overline{\mathbf{\Lambda}}_j \mathbf{e}_m \end{aligned} \tag{4.47}$$

and $E\{|\eta_k^{m\,(i)}|^2\} = E\{|\tilde{x}_k^{m\,(i)}|^2\} - E_s|\mu_m^{(i)}|^2$ where

$$\begin{aligned} E\{|\tilde{x}_k^{m\,(i)}|^2\} = &\sum_{j=0}^{N-1} \frac{1}{N} (\mathbf{W}_j^{(i)}(m))^H \mathbf{R}_{\mathbf{Y}_j} \mathbf{W}_j^{(i)}(m) + \sum_{j=0}^{N-1} \frac{1}{N} (\mathbf{C}_j^{(i)}(m))^H \mathbf{B}^{(i)} \mathbf{C}_j^{(i)}(m) \\ &- \sum_{j=0}^{N-1} \frac{2}{N} \mathrm{Re}\left\{ (\mathbf{W}_j^{(i)}(m))^H \overline{\mathbf{\Lambda}}_j \mathbf{P}^{(i)} \mathbf{C}_j^{(i)}(m) \right\} \end{aligned} \tag{4.48}$$

for $m = 1, \ldots, n_t$, where $\mathbf{W}_j^{(i)}(m)$ and $\mathbf{C}_j^{(i)}(m)$ are the m^{th} column of $\mathbf{W}_j^{(i)}$ and $\mathbf{C}_j^{(i)}$ respectively.

It is important to note that $\mu_m^{(i)}$ and $E\{|\eta_k^{m\,(i)}|^2\}$ values do not depend on symbol time index k, so these values are calculated only once for the decoding of one block in each iteration. The inputs to the decoder in terms of the LLR for each coded stream can be calculated by knowing the optimal filter coefficients. The residual interference at the output from the

equalizer is well approximated by a Gaussian distribution as in [36], [51]. Then, the extrinsic information given in (4.46) can be expressed as

$$L_k^{m\,(e)} = \frac{4\mathrm{Re}\{(\mu_m^{(i)})^* \tilde{x}_k^{m\,(i)}\}}{E\{|\eta_k^m|^2\}} \tag{4.49}$$

Soft feedback decisions for the FDE-FDDF can be expressed in terms of the extrinsic information provided by the decoder as follows [51]:

$$P[x_k^j] \triangleq P[x_k^j = b_j] = \frac{1}{2}\left[1 + b_j \tanh\left(\frac{1}{2}L_k^{j\,(p)}\right)\right],\ b_j \in \{+1, -1\} \tag{4.50}$$

$$\hat{x}_k^m = E\{x_k^m\} = \sum_{x_k^m \in \{+1,-1\}} x_k^m P[x_k^m] = \tanh\left(\frac{1}{2}L_k^{m\,(p)}\right) \tag{4.51}$$

for $E_s = 1$, $m = 1, \ldots, n_t$ and $k = 0, \ldots, N-1$. The non-zero diagonal entries of the correlation matrices $\mathbf{P}^{(i)}$ and $\mathbf{B}^{(i)}$ in (4.13) used by the forward and backward filters can be calculated by using the following approximation,

$$\rho_{k,m} \triangleq E\{x_k^m(\hat{x}_k^m)^*\} = E\{E\{x_k^m\}(\hat{x}_k^m)^*\} = |\hat{x}_k^m|^2 \tag{4.52}$$

$$\rho_m = \beta_m = \frac{1}{N}\sum_{k=0}^{N-1} \rho_{k,m} \tag{4.53}$$

$E\{x_k^m\}$ was taken as $\hat{x}_k^m$ and this is a common assumption in various turbo detection techniques as done in [36], [51] and [46].

This structure recovers from the exponential decoding complexity of the optimal maximum likelihood (ML) receiver by parallelizing the channel at the equalization stage. Therefore, the complexity of LLR computation which is the most computationally demanding part of the receiver is significantly reduced in this case and it shows linear dependence on the number of transmit antennas and constellation size.

4.5 Asymptotic Performance Analysis

At each iteration, forward and feedback filters approach the optimal coefficients in case of perfect feedback with the help of improved log a-posteriori probability (APP) ratio of each coded symbol obtained from the decoder. At later iterations, feedback decisions become more and more reliable and correlation matrices approach the asymptotic values: $\mathbf{P}^{(i)} \rightarrow E_s\mathbf{I}_{n_t}$ and $\mathbf{B}^{(i)} \rightarrow E_s\mathbf{I}_{n_t}$. Signal-to-interference-noise-ratio's (SINR) of each parallelized

channel in (4.45) after equalization are evaluated in Appendix B for the asymptotic case and given as

$$SINR_m = \sum_{l=0}^{L-1} \sum_{i=1}^{n_r} |H_l(i,m)|^2 \frac{E_s}{N_0}, \quad \text{for } m = 1, \ldots, n_t \tag{4.54}$$

It is seen from (4.54) that one can achieve the full diversity gain ($n_r \times L$) at each of the parallelized channels. If transmit diversity schemes in the form of coding across antennas such as universal space-time codes [45] or other coding-multiplexing based techniques [16] are utilized, the maximum potential diversity gain of ($n_r \times n_t \times L$) can be achieved by the proposed equalization scheme here.

A scheme is approximately universal if it is in deep fade only when the channel itself is in outage [45]. D-BLAST architecture satisfies this criterion by appropriate choice of code-words and being approximately universal is sufficient for a scheme to achieve the diversity-multiplexing tradeoff of the channel [55], [45]. Transmission schemes based on D-BLAST can achieve the full diversity gain of the flat fading MIMO channel ($n_r \times n_t$), if the temporal coding with stream rotation is capacity-achieving (Gaussian code books with infinite block size T). Moreover, the D-BLAST system can achieve the maximum capacity with outage, if the wasted space-time dimensions along the diagonals are neglected [45]. Therefore, by incorporating these types of coding structures in our proposed equalizer, one can achieve the diversity-multiplexing tradeoff of the frequency selective MIMO channel. However, this pursuit is beyond the scope of this thesis.

In our case, we have used simple coding structures that achieve the optimal rate-diversity tradeoff given by the singleton bound for block-fading channels in [30]. As it will be seen in Section 4.6, one can get a very close performance to the outage probability of MIMO-OFDM scheme.

We can calculate the asymptotic outage of our MIMO SC-FDE structure approximately by assuming independent identically distributed channel taps in (4.1) and using (4.54) as

follows,

$$
\begin{aligned}
P_{out}^{asymp}(R) &= \mathbb{P}\left\{C_{FDE-FDDF}^{asymp} < R\right\} = \mathbb{P}\left\{\sum_{m=1}^{n_t} \log_2\left(1 + SINR_m\right) < R\right\} \\
&\leq \mathbb{P}\left\{\max_{m,\, m\in\{1,\ldots,n_t\}} \log_2\left(1 + \frac{E_s}{N_0}\sum_{l=0}^{L-1}\sum_{i=1}^{n_r}|H_l(i,m)|^2\right) < R\right\} \\
&= \left(\mathbb{P}\left\{\log_2\left(1 + \frac{E_s}{N_0}\sum_{l=0}^{L-1}\sum_{i=1}^{n_r}|H_l(i,1)|^2\right) < R\right\}\right)^{n_t} \\
&\leq \left(\mathbb{P}\left\{\max_{l,\, i}\frac{E_s}{N_0}|H_l(i,1)|^2 < 2^R - 1\right\}\right)^{n_t} = \left(\mathbb{P}\left\{\frac{E_s}{N_0}|H_0(1,1)|^2 < 2^R - 1\right\}\right)^{n_t n_r L} \\
&= \left(1 - \exp^{-\frac{2^R-1}{(E_s/N_0)}}\right)^{n_t n_r L} \approx \frac{(2^R-1)^{n_t n_r L}}{(E_s/N_0)^{n_t n_r L}} \qquad (4.55)
\end{aligned}
$$

where standard probability evaluation techniques for maximum of independent random variables are employed. It is observed from (4.55) that the outage probability has a decay rate in the order of $(n_r \times n_t \times L)$ similar to MIMO-OFDM scheme and this outage is approximately achieved by our practical SC-FDE scheme as will be seen in Section 4.6.

4.6 Simulation Results

4.6.1 Outage Probability and MFB Calculations

In this section, we will compare the performance of our proposed equalizer with the hypothetical matched filtering bound (MFB) performance and the corresponding constrained outage probability of MIMO-OFDM system. The constrained capacity can be found for the system model in (4.2) given the complex vector set χ of cardinality M^{n_t} (e.g., M-ary or

M-PSK modulations) as follows

$$
\begin{aligned}
C^{\chi}_{MIMO-OFDM} &= \frac{1}{N}\sum_{j=0}^{N-1} I(\mathbf{X}_j;\mathbf{Y}_j|\overline{\mathbf{\Lambda}}_j) \\
&= \frac{1}{N}\sum_{j=0}^{N-1}\left(H(\mathbf{X}_j|\overline{\mathbf{\Lambda}}_j) - H(\mathbf{X}_j|\mathbf{Y}_j,\overline{\mathbf{\Lambda}}_j)\right) \\
&= \frac{1}{N}\sum_{j=0}^{N-1}\left(\log_2|\chi| - E_{\mathbf{X}_j,\mathbf{Y}_j}\left\{\log_2\left[\frac{1}{p(\mathbf{X}_j|\mathbf{Y}_j)}\right]\right\}\right) \\
&= \log_2|\chi| - \frac{1}{N}\sum_{j=0}^{N-1} E_{\mathbf{X}_j,\mathbf{Y}_j}\left\{\log_2\left[\frac{\sum_{\mathbf{X}_i\in\chi} p(\mathbf{Y}_j|\mathbf{X}_j=\mathbf{X}_i)}{p(\mathbf{Y}_j|\mathbf{X}_j)}\right]\right\} \\
&= \log_2|\chi| - \frac{1}{N}\sum_{j=0}^{N-1} E_{\mathbf{X}_j}\left\{E_{\mathbf{Y}_j|\mathbf{X}_j}\left\{\log_2\left[\frac{\sum_{\mathbf{X}_i\in\chi}\exp\left(-\|\mathbf{Y}_j-\overline{\mathbf{\Lambda}}_j\mathbf{X}_i\|^2/N_0\right)}{\exp\left(-\|\mathbf{Y}_j-\overline{\mathbf{\Lambda}}_j\mathbf{X}_j\|^2/N_0\right)}\right]\right\}\right\} \\
&= n_t\log_2(M) - \frac{1}{N}\sum_{j=0}^{N-1} E_{\mathbf{N}_j}\left\{\sum_{\mathbf{X}_k\in\chi}\frac{1}{M^{n_t}}\log_2\sum_{\mathbf{X}_i\in\chi}\exp\left(\frac{-\|\overline{\mathbf{\Lambda}}_j(\mathbf{X}_k-\mathbf{X}_i)+\mathbf{N}_j\|^2+\|\mathbf{N}_j\|^2}{N_0}\right)\right\}
\end{aligned}
\tag{4.56}
$$

where N is the number of OFDM subcarriers. Then, the corresponding outage probability can be written as

$$
P_{out}^{MIMO-OFDM,\,\chi}(R) = \mathbb{P}\left\{C^{\chi}_{MIMO-OFDM} < R\right\}. \tag{4.57}
$$

It is well known that the SC-MMSE receiver reduces to a channel matched filter if the perfect a priori information of the all transmitted symbols leading to ISI and inter-stream interference is available at the receiver and all the interference is cancelled [51]. Therefore, an upper bound to the packet error rate (PER) referred to as the matched filter bound (MFB) of the receiver can be obtained by assuming perfect decision feedback [3].

Alternatively, we can define a different MFB for wideband MIMO channels. This second MFB for wideband MIMO channels can be constructed such that the system is kept silent after one symbol vector transmission for a time duration exceeding the delay spread until the last echo is received. The receiver optimally combines all observed echoes coming from different multipaths belonging to the same symbol. Then, the equivalent MFB channel can be written as,

$$
\mathbf{y}_k = \mathbf{H}_{eq}\mathbf{x}_k + \mathbf{n}_k \tag{4.58}
$$

where

$$
\mathbf{H}_{eq} = \left[\mathbf{H}_0^T\ \mathbf{H}_1^T\ \ldots\ \mathbf{H}_{L-1}^T\right]^T. \tag{4.59}
$$

In this case, one can construct a nearly optimal iterative MIMO receiver structure based on the channel in (4.58). The nearly optimal decoder consists of producing the posterior probabilities of the binary coded symbols, and then feeding these probabilities to an ML decoder for the given binary code over the resulting binary-input continuous-output channel. One can calculate the extrinsic information given by the MIMO multi-stream detector to the SISO channel decoder exactly such that

$$
\begin{aligned}
L_k^{m\,(e)} &= \log_e \frac{p(\mathbf{y}_k|x_k^m=+1)}{p(\mathbf{y}_k|x_k^m=-1)} && (4.60)\\
&= \log_e \frac{\sum_{\mathbf{x}_k\in\chi_m^+} p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|x_k^m=+1)}{\sum_{\mathbf{x}_k\in\chi_m^-} p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|x_k^m=-1)} && (4.61)\\
&= \log_e \frac{\sum_{\mathbf{x}_k\in\chi_m^+} \exp\left[-\|\mathbf{y}_k-\mathbf{H}_{eq}\mathbf{x}_k\|^2/N_0\right]\prod_{j\neq m} P[x_k^j]}{\sum_{\mathbf{x}_k\in\chi_m^-} \exp\left[-\|\mathbf{y}_k-\mathbf{H}_{eq}\mathbf{x}_k\|^2/N_0\right]\prod_{j\neq m} P[x_k^j]} && (4.62)
\end{aligned}
$$

$m=1,\ldots,n_t$, where BPSK modulation is used for simplicity and

$$\chi_m^+ \triangleq \left\{(b_1,\ldots,b_{m-1},+1,b_{m+1},\ldots,b_{n_t}) : b_j\in\{+1,-1\},\ j\neq m\right\}, \tag{4.63}$$

$$P[x_k^j] \triangleq P[x_k^j=b_j],\ b_j\in\{+1,-1\}. \tag{4.64}$$

Similarly χ_m^- is defined. This scheme is particularly effective if used in conjunction with bit-interleaved coded modulation (BICM) [10], but it suffers from tremendous computational complexity which is exponential in constellation size, channel length and the number of transmit antennas as explained in [51] for multiuser schemes. $P[x_k^j]$'s are the a priori probabilities of the coded symbols obtained by the decoder as in (4.50) for $k=0,\ldots,N-1$ and $j=1,\ldots,n_t$. Constrained outage probability and two MFBs obtained here will be used for evaluation purposes in the next part.

4.6.2 Code Construction and Performance Results

The code construction used in our work is similar to the structure for random-like codes adapted to the block-fading channel based on blockwise concatenation and on bit-interleaved coded modulation (BICM) in [16]. The presented coded modulation construction in [16] systematically yields singleton-bound achieving turbo-like codes defined over an arbitrary signal set. As such, any other coding architecture that performs well in parallel block fading channels can be used in our system. We have used the same encoding and decoding structures as in [16] and [30] in simulations.

A sample coding structure used in simulations is shown in Fig. 4.2. Here, the outer code is a simple repetition code of rate $r = 1/n_t$ and the inner codes are rate-1 accumulators which is referred to as the repeat and blockwise accumulate (RBA) code [16]. Fig. 4.3 shows the performance of the proposed FDE-FDDF for a 4×4 MIMO system with the use of full block diversity attaining RBA code of rate $r = 1/4$. The channel model described in Section 4.2 is assumed and typical COST207 channel with exponential power delay profile for suburban and urban areas [37] is used. BPSK modulation is used for simplicity, but other M-ary or M-PSK modulations combined with BICM [10] can be applied to our proposed structure. Symbol duration is taken as 1 μ second, and the channel length L equals 8. The first channel tap is taken as unity power. The information block length, i.e., the information bits entering the outer encoder is taken as $K = 250$, then the block length N is equal to $K/(r \cdot n_t)+1 = 251$ including termination bits. The number of iterations inside the turbo RBA decoder is set to 10 and the number of equalizer iterations at which the forward and backward filters are updated by using the reliability matrices is taken as 3.

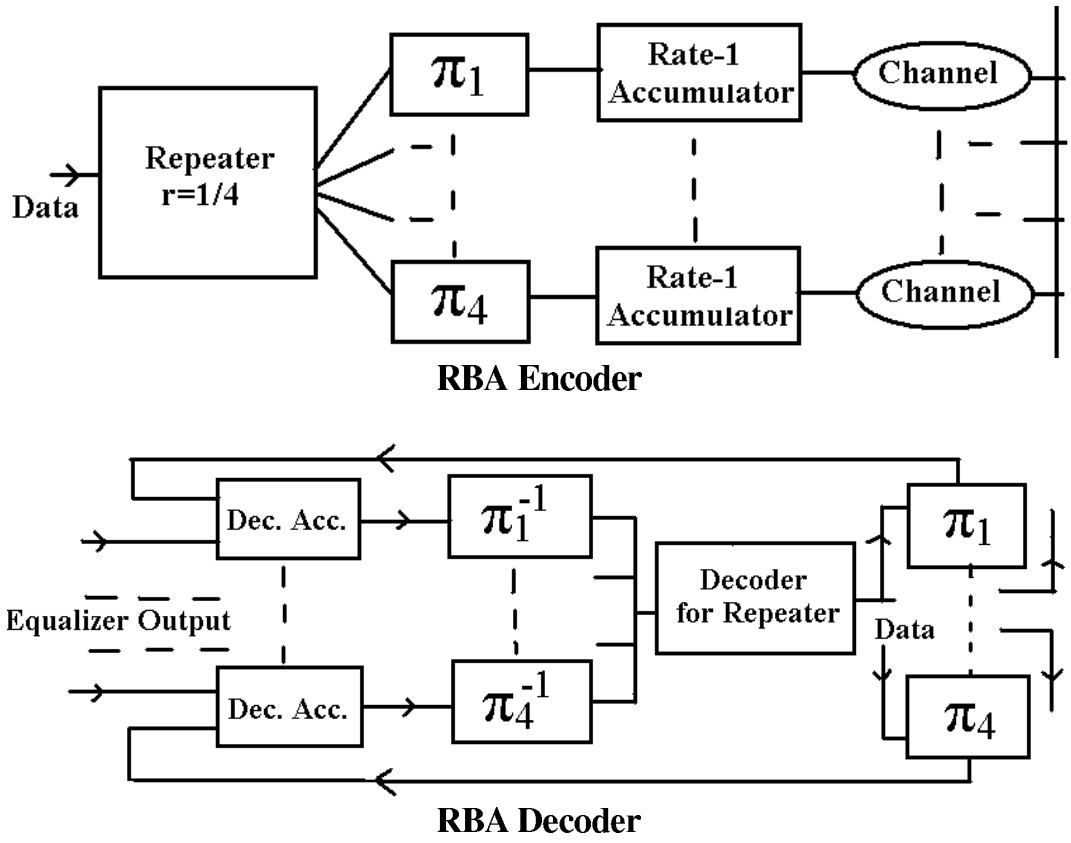

Figure 4.2: RBA Encoding and Decoding structure for 4×4 MIMO channel with $r = 1/4$

It is seen from Fig. 4.3 that the performance of FDE-FDDF is 0.3 dB away from MFB. There is approximately 1.5 dB difference between the outage probability of the MIMO-OFDM at rate $R = n_t \cdot r = 1$ bits/sec/Hz and this gap from the outage is similar to the gaps obtained with RBA in parallel block fading channels in [16]. Then, one can say that the ISI, the substream interference, and the error propagation problem in decision feedback is almost eliminated since the perfect decision feedback performance (MFB) is approximately achieved. Moreover, it is seen that the performance of FDE-FDDF shows the same slope as MIMO-OFDM outage and so it is possible to attain the maximum diversity of the MIMO broadband channel by using the proposed space-time equalizer and coding across transmit antennas. Furthermore, SC-FDE based schemes could be a promising candidate for wideband MIMO systems as an alternative to MIMO-OFDM schemes and if one takes the loss due to PAPR problem in OFDM based systems into consideration, the performance difference between SC-FDE based MIMO schemes and the MIMO-OFDM systems will be more significant.

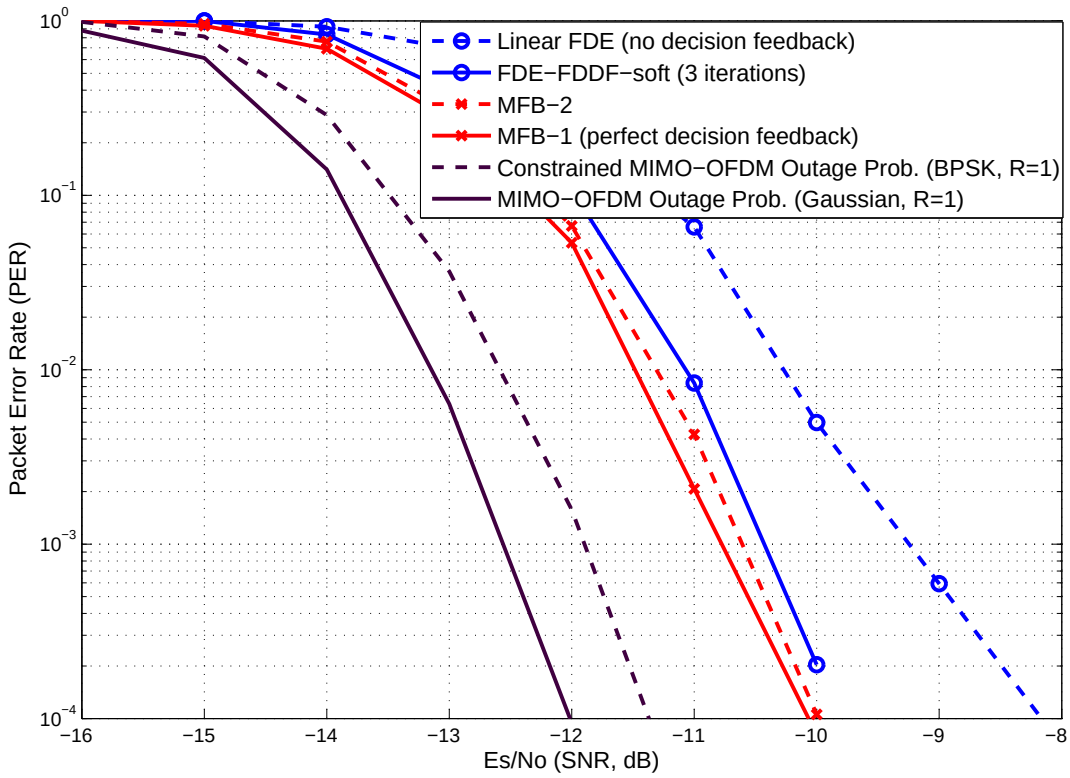

Figure 4.3: Performance comparison of frequency domain equalization techniques with matched filter bound and with outage probability for 4×4 MIMO system, $T_s = 1\mu$ sec, COST 207 typical suburban exponential channel, $L = 8$ and RBA code is used with $\frac{1}{4}$ code rate.

In Fig. 4.4, simulation results are depicted for code rate, $r = 1/2$. A full block diversity attaining blockwise concatenated code (BCC) is used for encoding as adapted from [16]. The outer code is a rate$-\frac{1}{2}$ convolutional code and the inner codes are n_t trivial rate-1 accumulators. The information block length K is taken as 248. Similar results are obtained and a close performance to MIMO-OFDM outage at rate $R = n_t \cdot r = 2$ bits/sec/Hz is achieved within 2 dB.

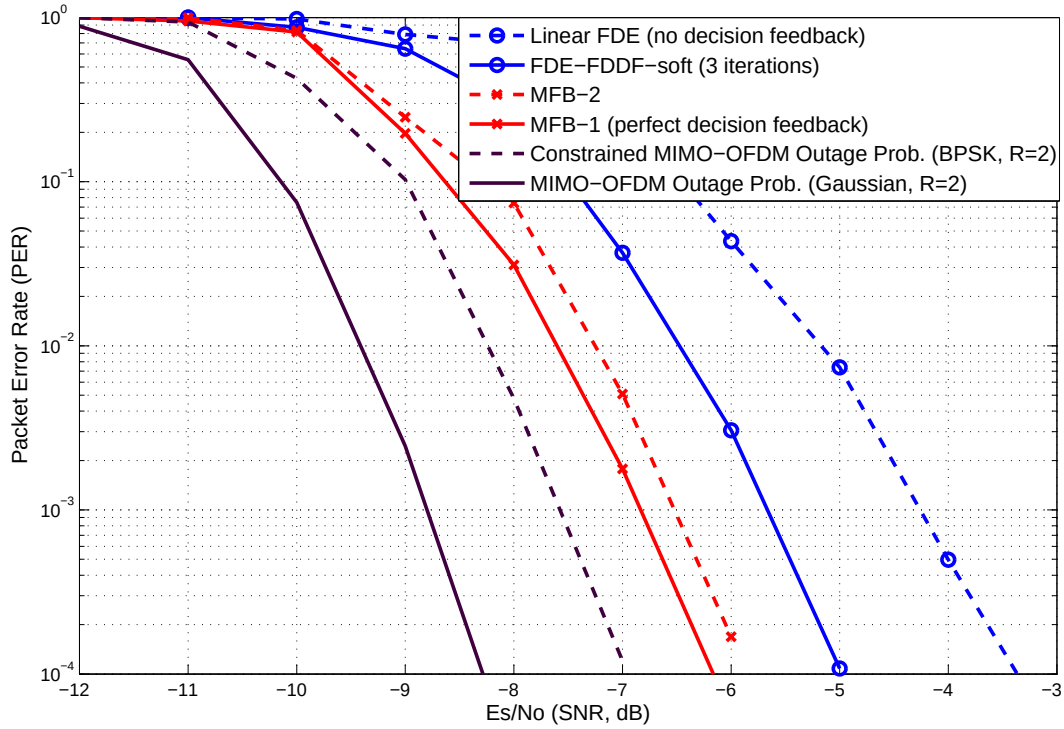

Figure 4.4: Performance comparison of frequency domain equalization techniques with matched filter bound and with outage probability for 4×4 MIMO system, $T_s = 1\mu$ sec, COST 207 typical suburban exponential channel, $L = 8$ and BCC is used with $\frac{1}{2}$ code rate.

Our proposed SC-FDE can also be applied to classical SISO ISI channels. In Fig. 4.5, we compared the performance of iterative SC-FDE-FDDF-soft feedback with that of the outage of an OFDM scheme and BPSK modulation is used. A convolutional encoder with $r = 1/2$ serially concatenated (SC) to a rate-1 accumulator is used with information block length $K = 123$. At first glance, it is surprising to note that the constrained OFDM outage probability is surpassed by the iterative FDE-FDDF, but as stated in [50], the capacity of wideband channels under non-Gaussian alphabets is an open problem and OFDM is not the capacity achieving scheme for non-Gaussian input alphabets. Maximum diversity order achieved by OFDM is given by the singleton-bound and this diversity is below the diversity order of the channel for BPSK modulation at $r = 1/2$. However, SC-FDE does not necessitate coding or Gaussian alphabet to attain maximum potential diversity order of the channel in SISO systems.

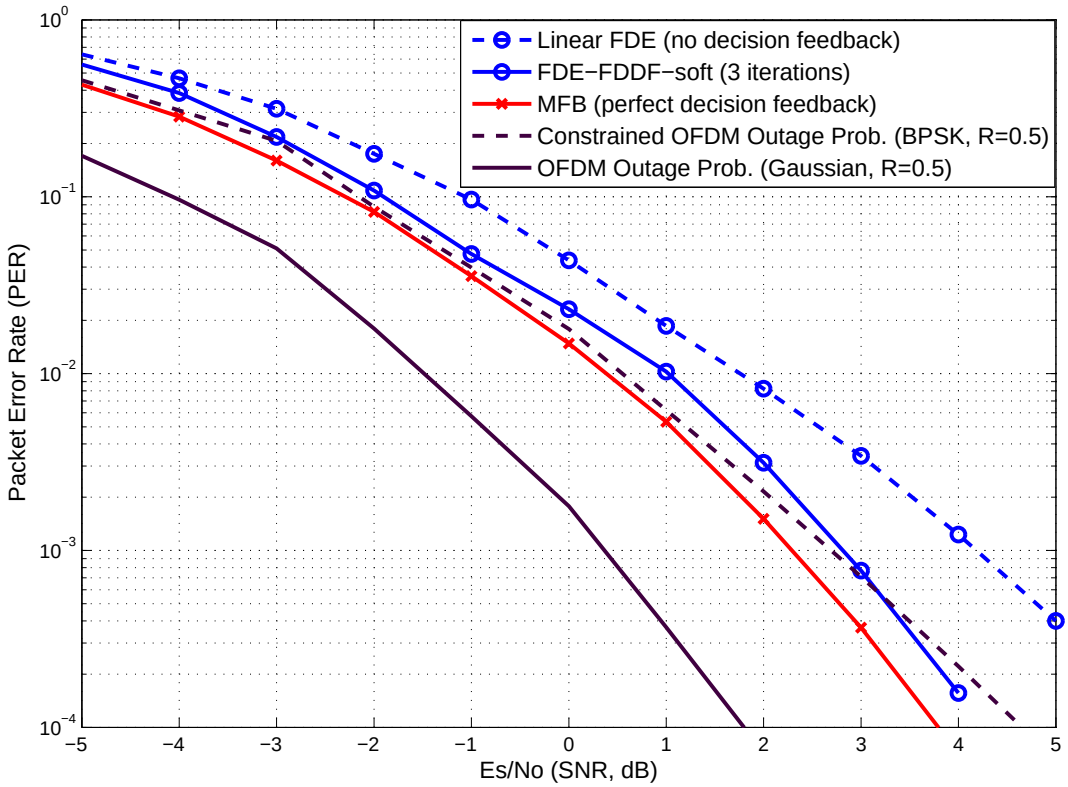

Figure 4.5: Performance comparison of frequency domain equalization techniques with matched filter bound and with outage probability for OFDM system, $T_s = 0.5\mu$ sec, COST 207 typical suburban exponential channel, $L = 15$ and SC code is used with $\frac{1}{2}$ code rate.

Furthermore, it is interesting to note that the performance improvement of the FDE-FDDF scheme over the linear FDE without decision feedback is about 2 dB at PER=0.0001 for all simulation results. There is also a loss in diversity as observed in the reduced PER slope without decision feedback. One can say that the proposed space-time equalizer gains more diversity in comparison to linear FDE by a careful design of both the forward and backward filters.

4.7 Conclusion

In this chapter, we extended the SC-FDE mechanism from SISO channels to more general vector-based models which include MIMO as a special case. We have also shown that capacity-achieving jointly optimal forward and backward filtering operations can be effectively performed in the frequency domain. It is observed that error performance close to the outage probability can be attained by careful coding across transmit antennas without compromising computational complexity. Therefore, our proposed iterative SC FDE technique for MIMO wideband channels can be viewed as a strong alternative to MIMO-OFDM schemes with similar complexity.

CHAPTER 5

CONCLUSIONS

5.1 Conclusions

In this thesis, we proposed various transmitter and receiver architectures which yield a close performance to channel capacity for the cases with and without CSIT. In Chapter 2, an adaptive scheme based on the reduced precoding idea is proposed for the limited rate feedback (LRF) MIMO channel. Our main contribution in Chapter 2 is the derivation of an upper bound expression for the ergodic capacity that is valid for a wide range of vector based quantization schemes. The tight analytical bound can be used to determine the number of precoders to be used at each average SNR value in order to maximize the spectral efficiency for a given accuracy of the CSIT. This strategy brings large capacity improvements especially when $n_t > n_r$ and decoding complexity reduction due to the parallelization of the MIMO channel when compared with no CSIT schemes.

In Chapters 3 and 4, we propose a practical receiver architecture over coded modulation schemes that exhibits a close performance to the information outage for MIMO and block fading channels, when CSIT is not available. It is known that the availability of CSIT does not have an observable effect over ergodic capacity, when n_t is not much larger than n_r. Thus, the only advantage coming from CSIT is the ease of decoding due to the conversion of the MIMO channel into non-interfering SISO channels. However, we show that it is still possible to exploit the total diversity benefits of the channel, when the channel knowledge is only available at receiver, without compromising the receiver complexity.

Our proposed receiver structure is actually information theoretically optimal and based on the iterative equalization and decoding idea. The feedback and feedforward filters in the equalization stage are jointly optimized to mitigate the interference by using the reliability information given by the decoder.

In Chapter 3, the proposed receiver architecture is applied to the decoding of rotated multidimensional constellation over block fading channels. It is observed that the optimal reliability exponent attained by Gaussian inputs and the corresponding outage probability is achieved by the proposed structure with finite input alphabets. Therefore, we can say that the theoretical benefits of multidimensional rotations are effectively materialized by the proposed practical receiver. As a side contribution, we can say that simple rotations like DFT are sufficient to get full diversity performance when the proposed structure is used. In Chapter 4, we applied our proposed receiver architecture to wideband MIMO channels without CSIT in cooperation with the SC-FDE technique. It is observed that the proposed receiver exploits the multipath and space diversity sources of the channel effectively such that very close performance to the outage probability is possible. As a side conclusion, we can say that this structure can be viewed as a strong alternative to MIMO-OFDM based schemes. Both equalizer structures in Chapter 3 and 4 reduce the receiver complexity, mainly, the complexity of LLR computation before the decoding stage, significantly by parallelizing the channel effectively.

In Chapter 2, we considered only the adaptive precoding scheme in which the number of beamformers used is changed based on the average SNR value to maximize spectral efficiency. However, adaptive modulation and coding structures, in which the constellation size for non-Gaussian alphabets and code rates are changed adaptively, can easily be incorporated into our proposed scheme.

In Chapters 3 and 4, the proposed schemes based on rotated constellations and SC-FDE technique are studied for fixed rate and coding scenarios. Various power or rate adaptation schemes useful for block fading channels can be applied to our system easily. Therefore, the proposed schemes are suitable for slow time-varying channels, in which the adaptive modulation techniques are used to keep the desired outage probability constant.

To sum up, SC-FDE and rotated constellation based schemes can achieve a very close performance to the outage and exploit full diversity by using effective equalization and interference suppression techniques different than the CSIT-based schemes. In case of CSIT, the quantized form of channel realization at the receiver is fed back to the transmitter in order to parallelize the channel.

One has to take the number of antennas at the receiver and the transmitter, the number of

feedback bits used to quantize beamformers, availability of the feedback channel, imperfections due to quantization, mobility, the channel estimation issues and required complexity into account while choosing between the proposed schemes and deciding whether to use CSIT or not.

5.2 Future Studies

In the slow fading channel model assumed throughout the study, we mentioned the power adaptation based on short-term power constraint. Although, it is known that the selection of short-term and long-term power constraints does not have an observable effect on ergodic capacity, its effect on the outage is remarkable for the slow fading channels in which one codeword sees only a finite number of channel states [11]. Therefore, power adaptation schemes based on long-term power constraint will be investigated and incorporated into our proposed structures.

Future studies will also include the consideration of imperfect channel estimation at the receiver on the performance of the proposed structures and the delayed transmission of the CSI to the transmitter side within a limited rate feedback scenario.

REFERENCES

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

[2] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Trans. Inform. Theory*, 20:284–287, Mar. 1974.

[3] J. R. Barry, E. A. Lee, and D. G. Messerschmitt. *Digital Communication*. Springer, Third Edition.

[4] N. Benvenuto and S. Tomasin. On the comparison between OFDM and single carrier modulation with a DFE using a frequency-domain feedforward filter. *IEEE Trans. Commun.*, 50:947–955, June 2002.

[5] N. Benvenuto and S. Tomasin. Iterative design and detection of a DFE in the frequency domain. *IEEE Trans. Commun.*, 53(11):1867–1875, Nov. 2005.

[6] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. Vincent Poor. *MIMO Wireless Communications*. Cambridge University Press, 2007.

[7] E. Biglieri, A. Nordio, and G. Taricco. Doubly iterative decoding of space time turbo codes with a large number of antennas. *IEEE Trans. Commun.*, 53(5):773–779, May 2005.

[8] J. Boutros, N. Gresset, L. Brunel, and M. Fossorier. Soft-input soft-output lattice sphere decoder for linear channels. *IEEE Global Commun. Conf*, Dec. 2003.

[9] J. Boutros and E. Viterbo. Signal space diversity: a power and bandwidth-efficient diversity technique for the rayleigh fading channel. *IEEE Trans. Inform. Theory*, 44(4):1453 1467, Jul. 1998.

[10] G. Caire, G. Taricco, and E. Biglieri. Bit-interleaved coded modulation. *IEEE Trans. Inform. Theory*, 44(3):927946, May 1998.

[11] G. Caire, G. Taricco, and E. Biglieri. Optimum power control over fading channels. *IEEE Trans. Inform. Theory*, 45(5):1468 – 1489, Jul. 1999.

[12] H. Campbell. *Linear Algebra With Applications*. Appleton Century Crofts, 1971.

[13] A. D. Dabbagh and D. J. Love. Feedback rate-capacity loss tradeoff for limited feedback MIMO systems. *IEEE Transactions on Information Theory*, 52(5):2190–2202, May 2006.

[14] L. Devroye. *Non-Uniform Random Variate Generation*. New York: Springer-Verlag, 1986.

[15] J. D. Dixon and B. Mortimer. *Permutation Groups, Graduate Texts in Mathematics*. Springer-Verlag, 1996.

[16] A. G. Fabregas and G. Caire. Coded modulation in the block-fading channel: Coding theorems and code construction. *IEEE Trans. Inform. Theory*, 52(1):91–114, Jan. 2006.

[17] A. G. Fabregas and G. Caire. Multidimensional coded modulation in block-fading channels. *IEEE Trans. Inform. Theory*, 54(5):2367–2372, May 2008.

[18] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson. Frequency domain equalization for single-carrier broadband wireless systems. *IEEE Commun. Mag.*, 40(4):58–66, April 2002.

[19] E. B. Fluckiger, F. Oggier, and E. Viterbo. New algebraic constructions of rotated lattice constellations for the rayleigh fading channel. *IEEE Trans. Inform. Theory*, 50(4):702 714, Apr. 2004.

[20] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multiple antennas. *Bell Laboratories Technical Journal*, 1(2):41–59, 1996.

[21] A. Gersho and R. M. Gray. *Vector Quantization and Signal Compression*. Kluwer Academis Press/Springer, 1992.

[22] A. Goldsmith. *Wireless Communications*. Camridge University Press, 2005.

[23] T. Guess and M. K. Varanasi. An information-theoretic framework for deriving canonical decision-feedback receivers in Gaussian channels. *IEEE Trans. Inform. Theory*, 51(1):173–187, Jan. 2005.

[24] S. Jin, M. R. McKay, X. Gao, and I. B. Collings. MIMO multichannel beamforming: SER and outage using new eigenvalue distributions of complex noncentral wishart matrices. *IEEE Transactions on Communication*, 56(3):424–434, March 2008.

[25] N. Jindal. MIMO broadcast channels with finite rate feedback. *Proc. IEEE Globecom*, 2005.

[26] G.D. Forney Jr. On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener. *Proc. 2003 Allerton Conf.*, pages 430–439, Oct. 2003.

[27] G.D. Forney Jr. Shannon meets Wiener II: On MMSE estimation in successive decoding schemes. *Proc. 2004 Allerton Conf.*, 2004.

[28] J. Karjalainen, N. Veselinovic, K. Kansanen, and T. Matsumoto. Iterative frequency domain joint-over-antenna detection in multiuser MIMO. *IEEE Trans. Wireless Comm.*, 6(10):3620–3631, Oct. 2007.

[29] K.Kansanen and T. Matsumoto. A computationally efficient MIMO turbo-equaliser. *Proc. IEEE VTC*, 1:277–281, 2003.

[30] R. Knopp and P. Humblet. On coding for block fading channels. *IEEE Trans. Inform. Theory*, 46(1):189–205, Jan. 2000.

[31] C. Lamy and J. Boutros. On random rotations diversity and minimum MSE decoding of lattices. *IEEE Trans. Inform. Theory*, 46(4):1584–1589, Jul. 2000.

[32] E.G. Larsson and P. Stoica. *Space-Time Block Coding for Wireless Communications*. Cambridge University Press, 2003.

[33] D. J. Love and R. W. Heath Jr. Multi-mode precoding using linear receivers for limited feedback MIMO systems. *2004 IEEE International Conference on Communications*, 1:448– 452, June 2004.

[34] D.J. Love, R.W. Heath Jr., and T. Strohmer. Grassmannian beamforming for multiple-input multiple-output wireless systems. *IEEE Trans. Inform. Theory*, 49(10):2735–2747, Oct. 2003.

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

[36] B. Ng, C. Lam, and D. Falconer. Turbo frequency domain equalizer for single-carrier broadband wireless systems. *IEEE Trans. Wireless Comm.*, 6(2):759–767, Feb. 2007.

[37] J. G. Proakis. *Digital Communications*. McGRAW-HILL, 2001.

[38] J. C. Roh and B. D. Rao. Design and analysis of MIMO spatial multiplexing systems with quantized feedback. *IEEE Transactions on Signal Processing*, 54(8):2874–2886, August 2006.

[39] J.C. Roh and B.D. Rao. Channel feedback quantization methods for MISO and MIMO systems. *Proc. IEEE PMIRC*, 2:805–809, Sept. 2004.

[40] B. Giannakis S. Zhou, Z. Wang. Quantifying the power loss when transmit beamforming relies on finite-rate feedback. *IEEE Trans. on Wireless Communications*, 4(4):1948–1956, July 2005.

[41] W. Santipach and M. L. Honig. Asymptotic performance of MIMO wireless channels with limited feedback. *Military Communications Conference, 2003. MILCOM 2003. IEEE*, 1:141–146, October 2003.

[42] P. Stoica, Y. Jiang, and J. Li. On MIMO channel capacity: An intuitive discussion. *IEEE Signal Processing Mag.*, pages 83–84, May 2005.

[43] Y. Sun, M. Yee, and M. Sandell. Iterative channel estimation with MIMO MMSE-turbo equalisation. *Proc. IEEE VTC*, 2:1278– 1282, Oct. 2003.

[44] I. E. Telatar. Capacity of multi-antenna Gaussian channels. *Europ. Trans. Telecommun.*, 10:585–598, Nov./Dec. 1999.

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

[46] M. Tuchler and J. Hagenauer. Linear time and frequency domain turbo equalization. *Proc. IEEE VTC*, 2:1449–1453, 2001.

[47] M. Tuchler, R. Koetter, and A. C. Singer. Turbo equalization: principles and new results. *IEEE Trans. Commun.*, 50(5):754–767, May 2002.

[48] R. Visoz, A. O. Berthet, and S. Chtourou. Frequency-domain block turbo-equalization for single-carrier transmission over MIMO broadband wireless channel. *IEEE Trans. Commun.*, 54(12):2144–2149, Dec. 2006.

[49] E. Viterbo and J. Boutros. A universal lattice decoder for fading channels. *IEEE Trans. Inform. Theory*, 45(4):1639 1642, Jul. 1999.

[50] P. O. Vontobel, A. Kavcic, D. M. Arnold, and H. A. Loeliger. A generalization of the BlahutArimoto algorithm to finite-state channels. *IEEE Trans. Inform. Theory*, 54(5):1887–1918, May 2008.

[51] X. Wang and H. V. Poor. Iterative (turbo) soft interference cancellation and decoding for coded CDMA. *IEEE Trans. Commun.*, 47:1046–1061, July 1999.

[52] P. W. Wolniansky, G. J. Foschini, and G. D. Golden. V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel. *Bell Laboratories, Lucent Technologies, Crawford Hill Laboratory*, Oct. 1998.

[53] M. S. Yee, M. Sandell, and Y. Sun. Comparison study of single carrier and multicarrier modulation using iterative based receiver for MIMO system. *Proc. IEEE VTC*, pages 1275–1279, 2004.

[54] T. Yoo, E. Yoon, and A. Goldmisth. MIMO capacity with channel uncertainty: Does feedback help? *Proc. IEEE Globecom*, pages 96–100, 2004.

[55] L. Zheng and D. N. C. Tse. Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels. *IEEE Trans. Inform. Theory*, 49(5):10731096, May 2003.

[56] Z. Zhou, B. Vucetic, M. Dohler, and Y. Li. MIMO systems with adaptive modulation. *IEEE Trans. Vehicular Tech.*, 54(5):1828–1842, Sept. 2005.

[57] Y. Zhu and K. B. Letaief. Single-carrier frequency-domain equalization with noise prediction for MIMO systems. *IEEE Trans. Commun.*, 55(5):1063–1076, May 2007.

APPENDIX A

PROOF OF (2.14)

Define a matrix $\mathbf{B} = \frac{P}{n}\mathbf{H}\mathbf{H}^*$ with elements $B_{i,j} = \frac{P}{n} d_i d_j \left(\sum_{k=1}^{n} V_{ik} V_{jk}^*\right)$ and another matrix $\mathbf{A} = \mathbf{I_n} + \mathbf{B}$. For the capacity bound of the $n-$precoder MIMO scheme, the determinant of the $n \times n$ matrix $\mathbf{A}$ is necessary in evaluating (2.14) and it can be found by Leibniz formula [12]:

$$\det(\mathbf{A}) = \sum_{(\Sigma_k \in S^*,\ k=1,\ldots,n!)} \left(\prod_{i=1}^{n} A_{i,\sigma^k(i)} \right) . sgn(\Sigma_k), \tag{A.1}$$

where $\sigma^k(i)$ is the i^{th} element of Σ_k which is the k^{th} element of the permutation group S^*, and S^* includes all possible permutations of the set $S = \{1, 2, \ldots, n\}$. There are $n!$ different permutations of S and hence S^* is composed of $n!$ permutations. The function $'sgn'$ of permutations in the permutation group S^* returns $+1$ or -1 for even and odd permutations, respectively [15].

Recalling that A_{ij} equals B_{ij} for $i \neq j$ and $1 + B_{ij}$ otherwise, one can write the determinant expression in a compact form in terms of B_{ij}'s directly from (A.1) after a careful inspection as

$$\det(\mathbf{A}) = 1 + \sum_{k=1}^{n} \sum_{S_k = (a_1,\ldots,a_k) \in P_k} \left(\sum_{(\Sigma_l^k \in S_k^*,\ l=1,\ldots,k!)} \left(\prod_{i=1}^{k} B_{a_i,\sigma^l(i)} . sgn\left(\Sigma_l^k\right) \right) \right), \tag{A.2}$$

where P_k is the set containing all possible (n, k) combinations of $\{1, 2, \ldots, n\}$. S_k^* includes all possible permutations of $S_k = \{a_1, a_2, \ldots, a_k\}$ and there are $k!$ different permutations in S_k^*. In the above expression, the set of $k-$element combination from $S = \{1, 2, \ldots, n\}$ is determined first. Considering the $\Sigma_{(a_1,\ldots,a_k) \in P_k}$ term in the summation in (A.2), $\sigma^l(i)$ is the i^{th} element of Σ_l^k which is the l^{th} element of the permutation group S_k^*. Recalling that $B_{ij} = \frac{P}{n} d_i d_j \left(\sum_{k=1}^{n} V_{ik} V_{jk}^*\right)$, one can put B_{ij} into $\det(\mathbf{A})$ expression given in (A.2) and obtain

the following:

$$\det(\mathbf{A}) = 1 + \sum_{k=1}^{n} \sum_{S_k \in P_k} \Big(\sum_{\Sigma_l^k \in S_k^*} \prod_{i=1}^{k} \left(\frac{P}{n} d_{a_i} d_{\sigma^l(i)} \sum_{m=1}^{n} (V_{a_i m})(V_{\sigma^l(i) m})^* \right) . sgn\left(\Sigma_l^k\right) \Big). \tag{A.3}$$

The above expression can be simplified as follows. Defining $S_k = \{a_1, \ldots, a_k\}$ and a partition $S_1^k, S_2^k, \ldots, S_p^k \subset S_k$ which are disjoint sets that satisfy $S_1^k \cup S_2^k \cup \ldots \cup S_p^k = S_k = \{a_1, \ldots, a_k\}$, $s_i^k(j)$ is the j^{th} element in S_i^k so that $s_i^k(j),\ j = 1, \ldots, |S_i^k|$, are the elements belonging to S_i^k where $|S_i^k|$ is the cardinality of S_i^k for $i = 1, 2, \ldots, p$ $(1 \leq p \leq k)$. Although there are many terms in (A.3), the only terms that have nonzero mean are the ones all composed of squared forms $|V_{ij}|^2$. This is due to the reason that V_{ij} has a uniformly distributed phase in $(0, 2\pi)$, which is independent of V_{ik} for all $k \neq j$ and V_{lj} for all $l \neq i$. This uniform phase distribution will result in a zero expected value for any term which has a non-squared form of V_{ij}. Using this fact and after a few straightforward steps, the following simplified expression can be obtained from (A.3) after taking expectation

$$\begin{aligned} E\{\det(\mathbf{A})\} \ = \ E\{1 + \sum_{k=1}^{n} \sum_{S_k \in P_k} \Big(\Big(\frac{P}{n}\Big)^k (d_{a_1} d_{a_2} \cdots d_{a_k})^2 . \\ \sum_{((S_1^k, S_2^k, \ldots, S_p^k) \in [S_k]^{p=1,\ldots,k})} \ \sum_{(\Sigma_l^k(t) \in S_t^{k*},\ t=1,\ldots,p)} \\ \left(\prod_{i=1}^{p} \sum_{z_i=1}^{n} \prod_{j=1}^{|S_i^k|} \left| V_{s_i^k(j) z_i} \right|^2 \right) . sgn\left(\Sigma_l^k\right) \Big)\}, \end{aligned} \tag{A.4}$$

where $z_1 \neq z_2 \neq \ldots \neq z_p$, $\Sigma_l^k = (\Sigma_l^k(1), \ldots, \Sigma_l^k(p))$, and S_t^{k*} is the permutation group that includes all possible permutations of the elements in S_t^k. In the above equation, $[S_k]^{p=1,\ldots,k}$ is the set that includes all possible $p-$element partitions $\left(S_1^k, S_2^k, \ldots, S_p^k\right)$ of the set S_k for $p = 1, \ldots, k$. Therefore, the summation $\sum_{((S_1^k, S_2^k, \ldots, S_p^k) \in [S_k]^{p=1,\ldots,k})} \sum_{(\Sigma_l^k(t) \in S_t^{k*},\ t=1,\ldots,p)}$ in (A.4) is over all different permutations of all possible $\left(S_1^k, S_2^k, \ldots, S_p^k\right)$ partitions, i.e., the summation shows that $p-$element partition is chosen from $[S_k]^{p=1,\ldots,k}$ first and then, the second summation $\sum_{(\Sigma_l^k(t) \in S_t^{k*},\ t=1,\ldots,p)}$ is taken over all different permutations of the chosen partition of S_k. Actually, the summation $\sum_{(\Sigma_l^k(t) \in S_t^{k*},\ t=1,\ldots,p)}$ is the shorthand notation of $\sum_{(\Sigma_l^k(1) \in S_1^{k*})} \sum_{(\Sigma_l^k(2) \in S_2^{k*})} \cdots \sum_{(\Sigma_l^k(p) \in S_p^{k*})}$ and $sgn\left(\Sigma_l^k\right)$ takes values $+1$ or -1 depending on whether $\Sigma_l^k = (\Sigma_l^k(1), \ldots, \Sigma_l^k(p))$ is an even or odd permutation.

***Lemma* 1** *The expression given in (A.4) can be simplified as*

$$E\{\det(\mathbf{A})\} = E\{1 + \sum_{k=1}^{n} \sum_{S_k \in P_k} \left(\frac{P}{n}\right)^k (d_{a_1} d_{a_2} \cdots d_{a_k})^2 . \sum_{(j_1 \in S)} \sum_{(j_2 \in S,\ j_2 \neq j_1)} \cdots \sum_{(j_k \in S,\ j_k \neq j_1, \ldots, j_{k-1})} (|V_{a_1 j_1}|^2 \cdots |V_{a_k j_k}|^2)\}. \tag{A.5}$$

Proof: The lemma suggests that the only remaining terms in (A.4) are the terms resulting from the partition of $(S_1^k, \ldots, S_p^k)$ with $p = k$. The only possible partition is then $S_1^k = \{a_1\}, S_2^k = \{a_2\}, \ldots, S_k^k = \{a_k\}$, and $|S_i^k| = 1$ for all i's. Eqn. (A.4) reduces to (A.5) for $p = k$ and $z_1 \neq z_2 \neq \ldots \neq z_k$. Any term included within the summation $\sum_{((S_1^k, S_2^k, \ldots, S_p^k) \in [S_k]^{p=1,\ldots,k})} \sum_{(\Sigma_l^k(t) \in S_t^{k*},\ t=1,\ldots,p)}$ in (A.4) but not in (A.5) can also be written as

$$\prod_{i=1}^{p} \prod_{j=1}^{|S_i^k|} \left| V_{s_i^k(j) z_i} \right|^2 \tag{A.6}$$

for any given $z_1, z_2, \ldots, z_p$ between 1 and n where $z_1 \neq z_2 \neq \ldots \neq z_p$ (the terms with the same z_i's are already in another partition of S_k). $s_i^k(j)$ is the j^{th} element of S_i^k as defined before. For the above term, at least one of the S_i^k's has cardinality $|S_i^k|$ greater than or equal to 2 for some i since $p < k$, i.e., $max_{1 \leq i \leq p}(|S_i^k|) \geq 2$. Note that the terms in (A.5) have $|S_i^k| = 1$ for $i = 1, 2, \ldots, p$ and $p = k$.

The terms in the form given in (A.6) originate from the term given below in (A.4)

$$E\left\{\left(\prod_{i=1}^{p} \sum_{z_i=1}^{n} \prod_{j=1}^{|S_i^k|} \left| V_{s_i^k(j) z_i} \right|^2\right) . sgn\left(\Sigma_l^k\right)\right\}. \tag{A.7}$$

Now focus on the summation in (A.4) over permutations $\Sigma_l^k(t) \in S_t^{k*}$ for $t = 1, \ldots, p$. Fix $z_1, \ldots, z_p$ ($z_1 \neq z_2 \neq \ldots \neq z_p$ mandated by (A.6)) and all the $\Sigma_l^k(1), \Sigma_l^k(2), \ldots, \Sigma_l^k(p)$ permutations except for one of the $\Sigma_l^k(c)$ with $|S_c^k| \geq 2$. Then, (A.6) can be written as

$$\left(\prod_{i \neq c}^{p} \prod_{j=1}^{|S_i^k|} \left| V_{s_i^k(j) z_i} \right|^2\right) . \prod_{j=1}^{|S_c^k|} \left| V_{s_c^k(j) z_c} \right|^2 . sgn\left(\Sigma_l^k\right). \tag{A.8}$$

along with the permutation sign. By taking the summation over different $\Sigma_l^k(c)$ permutations,

one can get

$$sgn\left(\Sigma_l^k(t),\ t = 1,\ldots,p \text{ and } t \neq c\right).\prod_{i\neq c}^{p}\prod_{j=1}^{|S_i^k|}\left|V_{s_i^k(j)z_i}\right|^2.$$

$$\sum_{\Sigma_l^k(c)\in S_c^{k*}}\left(\prod_{j=1}^{|S_c^k|}\left|V_{s_c^k(j)z_c}\right|^2\right).sgn\left(\Sigma_l^k(c)\right) = 0, \tag{A.9}$$

since there are equal number of even and odd permutations of S_c^k set if $|S_c^k| \geq 2$ with opposite signs [15]. In other words, in the above equation $sgn(\Sigma_l^k(c))$ is +1 for $\frac{|S_c^k|!}{2}$ times and −1 for $\frac{|S_c^k|!}{2}$ times again. Therefore, the expression in (A.9) goes to zero and this can be done for other $\Sigma_l^k(1),\ldots,\Sigma_l^k(p)$ permutations. The terms presented in (A.6) cancel each other in (A.4) and (A.4) reduces to the equation given in (A.5). ***Q.E.D.***

The simplified form of the capacity bound in (A.5) is important, since

$$E\left[|V_{a_1 j_1}|^2\cdots|V_{a_k j_k}|^2\right] = E\left[|V_{a_1 j_1}|^2\right]\cdots E\left[|V_{a_k j_k}|^2\right] \tag{A.10}$$

by the independence of $|V_{a_1 j_1}|^2,\ldots,|V_{a_k j_k}|^2$. Since the distribution of $|V_{a_i j_i}|^2$'s are identical and same as the distribution of $|V_{11}|^2$ if $a_i = j_i$ and equal to the distribution of $|V_{12}|^2$ if $a_i \neq j_i$, one need only the expected values $E\left[|V_{11}|^2\right] = E_{11}$ and $E\left[|V_{12}|^2\right] = E_{12}$ in order to calculate the expression given in (A.5).

APPENDIX B

ASYMPTOTIC SINR CALCULATION

If $\mathbf{P}^{(i)} = \mathbf{B}^{(i)} = E_s\mathbf{I}_{n_t}$, one can write $\mathbf{A}_j$ in (4.43) as

$$\mathbf{A}_j = \left[E_s\mathbf{I}_{n_t} - E_s^2\overline{\mathbf{\Lambda}}_j^H\mathbf{R}_{\mathbf{Y}_j}^{-1}\overline{\mathbf{\Lambda}}_j\right]^{-1} \tag{B.1}$$

By using Matrix Inversion Lemma, one can get

$$\mathbf{A}_j = \frac{1}{E_s}\mathbf{I}_{n_t} + \frac{1}{E_s}\mathbf{I}_{n_t}E_s^2\overline{\mathbf{\Lambda}}_j^H\left[\mathbf{R}_{\mathbf{Y}_j} - \overline{\mathbf{\Lambda}}_j\frac{1}{E_s}\mathbf{I}_{n_t}E_s^2\overline{\mathbf{\Lambda}}_j^H\right]^{-1}\overline{\mathbf{\Lambda}}_j\frac{1}{E_s} = E_s^{-1}\mathbf{I}_{n_t} + \overline{\mathbf{\Lambda}}_j^H N_0^{-1}\mathbf{I}_{n_r}\overline{\mathbf{\Lambda}}_j \tag{B.2}$$

and

$$\mathbf{D}_j = -(\mathbf{A}_j)^{-1} \tag{B.3}$$

is written from (4.44). Lagrangian terms in (4.32) can be found as

$$\Gamma_n = \frac{-N}{\sum_{j=0}^{N-1}\mathbf{A}_j(n,n)} = \frac{-E_sN}{\sum_{j=0}^{N-1}\left[1 + \frac{E_s}{N_0}\overline{\mathbf{\Lambda}}_j^H(n)\overline{\mathbf{\Lambda}}_j(n)\right]},\ n = 1,\ldots,n_t \tag{B.4}$$

by using (4.41) and (B.2) and further noting that $\mathbf{A}_j^{(i)}(n,:)\mathbf{D}_j^{(i)}(:,n) = (\mathbf{A}_j^{(i)}\mathbf{D}_j^{(i)})(n,n) = -1$ by (B.3). Defining

$$\Sigma_n = \sum_{j=0}^{N-1}\left[1 + \frac{E_s}{N_0}\overline{\mathbf{\Lambda}}_j^H(n)\overline{\mathbf{\Lambda}}_j(n)\right], \tag{B.5}$$

feedback filter matrices can be obtained as

$$\mathbf{C}_j = -\mathbf{I}_{n_t} - \mathbf{A}_j\mathbf{\Gamma} \tag{B.6}$$

from (4.42) and (B.3). One can obtain the columns of feedback filter matrices by putting (B.2) and (B.4) into (B.6) such that

$$\mathbf{C}_j(n) = -\mathbf{e}_n + \frac{N}{\Sigma_n}\left[\mathbf{e}_n + \frac{E_s}{N_0}\overline{\mathbf{\Lambda}}_j^H\overline{\mathbf{\Lambda}}_j(n)\right] \tag{B.7}$$

and forward filter can be found as

$$\mathbf{W}_j(n) = \frac{\frac{E_s}{N_0}N\overline{\mathbf{\Lambda}}_j(n)}{\Sigma_n} \tag{B.8}$$

for $j = 0,\ldots,N-1$, $n = 1,\ldots,n_t$ from (4.38) and (B.7).

In this case, the soft estimate of x_k^m is a scaled version of the matched filter output after ideal interference cancellation in the frequency domain. SINR's of each parallelized channels after equalization can be found after some manipulation by using (4.47) and (4.48) for the asymptotic case as follows,

$$\mu_n = \frac{\sum_{j=0}^{N-1}\left[\frac{E_s}{N_0}\overline{\mathbf{\Lambda}}_j^H(n)\overline{\mathbf{\Lambda}}_j(n)\right]}{\Sigma_n} \tag{B.9}$$

and

$$E\{|\tilde{x}_k^n|^2\} = \frac{E_s}{N}\sum_{j=0}^{N-1}\left[1+\frac{N^2}{\Sigma_n^2}-2\frac{N}{\Sigma_n}\right]+\frac{\sum_{j=0}^{N-1}\left[\frac{E_s^2}{N_0}N\overline{\mathbf{\Lambda}}_j^H(n)\overline{\mathbf{\Lambda}}_j(n)\right]}{\Sigma_n^2} \tag{B.10}$$

$$E\{|\eta_k^n|^2\} = \frac{E_s[\Sigma_n - N]^2}{\Sigma_n^2}+\frac{\sum_{j=0}^{N-1}\left[\frac{E_s^2}{N_0}N\overline{\mathbf{\Lambda}}_j^H(n)\overline{\mathbf{\Lambda}}_j(n)\right]}{\Sigma_n^2}-E_s|\mu_n|^2 = \frac{\sum_{j=0}^{N-1}\left[\frac{E_s^2}{N_0}N\overline{\mathbf{\Lambda}}_j^H(n)\overline{\mathbf{\Lambda}}_j(n)\right]}{\Sigma_n^2} \tag{B.11}$$

for $n = 1, \ldots, n_t$.

The SINR can then be evaluated as

$$SINR_m = \frac{|\mu_m|^2 E_s}{E\{|\eta_k^m|^2\}} = \frac{1}{N}\sum_{j=0}^{N-1}\overline{\mathbf{\Lambda}}_j^H(m)\overline{\mathbf{\Lambda}}_j(m)\frac{E_s}{N_0} \tag{B.12}$$

from (4.45), and one can obtain the following based on Parseval's relation by using (4.3)

$$SINR_m = \frac{1}{N}\sum_{j=0}^{N-1}\frac{E_s}{N_0}\sum_{i=1}^{n_r}|\lambda_m^i(j)|^2 = \frac{1}{N}\sum_{i=1}^{n_r}\frac{E_s}{N_0}\sum_{j=0}^{N-1}N|H_j(i,m)|^2 = \sum_{l=0}^{L-1}\sum_{i=1}^{n_r}|H_l(i,m)|^2\frac{E_s}{N_0} \tag{B.13}$$

for $m = 1, \ldots, n_t$.