The Dissertation Committee for David James Love
certifies that this is the approved version of the following dissertation:

Feedback Methods for Multiple-Input Multiple-Output Wireless Systems

Committee:

Robert W. Heath, Jr., Supervisor

Gustavo de Veciana

John E. Gilbert

Edward J. Powers

Theodore S. Rappaport

Sanjay Shakkottai

Feedback Methods for Multiple-Input Multiple-Output Wireless Systems

by

David James Love, B.S., M.S.E.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2004

Dedicated to the memory of my grandfathers,

Joe B. Stuart and Ralph D. Love

Acknowledgments

I would like to start by thanking God for the many blessings in my life. I thank my parents Jim and Priscilla Love for their support throughout my education. All of my academic successes can be traced back to them. In elementary school my parents told me, “All we expect is for you to do your best.” I am happy to say that I took this lesson to heart. My wife Elaina Love has also been instrumental to my academic life. Many people do not realize that we have been together since I was fifteen. I love her very much, and I (along with most of my important documents) would be lost without her. My younger brother Brian Love has also played a large role in my success. I always tried to present myself as a role model for Brian and doing so has made me a better person. My life has also been profoundly affected by my grandparents Joe and Ola Stuart and Ralph and Jane Love. Just like my parents, they always encouraged me to do my best. Though my grandfathers have both passed away, I hope that my success makes them proud.

I would like to thank my advisor Robert Heath for serving as a friend and mentor during my graduate studies. Robert gave me a large amount of freedom during graduate school, and I feel that this freedom was fundamental to my accomplishments. I am grateful to my committee members Gustavo de

Veciana, John Gilbert, Edward Powers, Ted Rappaport, and Sanjay Shakkottai. Their comments and insights made my research more fruitful. I would like to thank Alan Bovik for his generosity in allowing me to work with his research group early in my graduate studies. John Dollard and Lawrence Shepley have provided invaluable advice and help throughout my UT years. I would also like to thank Brian Evans and Irwin Sandberg for their support.

Thanks also goes to everyone who I interacted with at the Texas Instruments DSPS R&D Center in Dallas. Without the help of Mike Polley and Don Shaver, I would have never had the opportunity to be associated with TI. Anuj Batra, Hank Eilts, Srinath Hosur, and David Magee have all provided useful guidance.

I would also like to thank my friends. Joseph D'Auria has provided enough entertainment to supplant television. Joe and I have experienced adventures rivaling those on *Seinfeld.* My UT ECE days have been filled with wonderful memories thanks to my LIVE friends such as Umesh Rajashekar, Farooq Sabir, Hamid Sheikh, and Mehul Sampat. We have spent many hours discussing problems in signal processing, communications, ping-pong, and arranged marriages. My WSIL friends including Manish Airy, Runhua Chen, Antonio Forenza, Chun-Hung Liu, Caleb Lo, Bishwarup Mondal, Roopsha Samanta, Taiwen Tang, and Jimmy Wang have provided loads of laughs and Silicon Valley stories.

I would also like to thank The University of Texas at Austin. I have been at UT for my entire university life. I have been lucky enough to watch Ricky Williams, Major Applewhite, and T.J. Ford in person. Joe, Brian, and I were even fortunate enough to get a "Thanks guys!" from Mack Brown during the Cotton Bowl practices at SMU during the 1999 season. I got to see the highs

(Heisman, UT Final Four, Holiday Bowl comeback) and lows (UCLA, Big 12 title game meltdown, various OU shellackings). UT football and basketball have been my primary pastime and passion. I will always bleed burnt orange.

David James Love

The University of Texas at Austin
May 2004

Feedback Methods for Multiple-Input Multiple-Output Wireless Systems

Publication No. ____________________

David James Love, Ph.D.
The University of Texas at Austin, 2004

Supervisor: Robert W. Heath, Jr.

The availability and performance of wireless communication systems have grown at unprecedented rates over the last fifteen years. In order to maintain these growth rates, next generation wireless systems must supply both reliability and high data rates using a fixed amount of spectrum and limited transmit power. Multiple-input multiple-output (MIMO) wireless communication systems, which use multiple antennas at both the transmitter and receiver, are expected to be one of the enabling technologies for next generation wireless systems. MIMO wireless systems provide both data rate and reliability improvements by designing signals over space as well as time and/or frequency.

One of the key features of wireless links is the phenomenon known as fading. In a narrowband wireless link, fading can be modeled as a multiplicative channel gain. Many of the benefits promised by MIMO wireless systems will not be realized without the availability of channel gain information at the

transmitter. Current research focuses on the usage of MIMO signaling where the transmitter does not have any kind of channel information. When some form of channel information is available, however, the MIMO signal can be adapted to the current channel conditions to improve performance. Unfortunately, many wireless systems will not have any form of a priori channel knowledge without feedback from the receiver to transmitter. Because feedback will occupy a percentage of the data-rate in the reverse wireless link (i.e. the wireless link where the receiver terminal serves as the transmitter), feedback must be kept to a limited number of bits.

This dissertation describes new, practical methods for limited feedback space-time signaling. It develops space-time techniques that provide improved performance when channel information, in the form of a fixed number of bits, is conveyed from the receiver to the transmitter. Limited feedback techniques are developed for linearly precoded orthogonal space-time block codes and linearly precoded spatial multiplexing. A new adaptive modulation technique for linearly precoded spatial multiplexing called multi-mode precoding is presented. Past research in the area of space-time signaling is also overviewed.

Contents

Chapter 1

Introduction

This chapter presents an introduction and overview of this dissertation. Section 1.1 gives a general introduction to wireless. In Section 1.2, the basic ideas behind multiple-input multiple-output (MIMO) wireless systems are reviewed. Current and past research into closed-loop MIMO wireless is presented in Section 1.3. The contributions are briefly summarized in Section 1.4. A summary of the organization of the remaining chapters is presented in Section 1.5.

1.1 Wireless Communications

Wireless communications have forever changed the way we live because of their ability to supply ubiquitous voice and data communications. Internet applications and appliances have fueled research in wireless systems that supply data rates much larger than those found in wireless voice communications. With second and a half generation (2.5G) and third generation (3G) communications becoming commercially viable, it is clear that the future of wireless communications is in wireless appliances that can provide both voice and data communications over a reliable and high rate link.

In designing a wireless system, designers must deal with the scarcity of the available bandwidth. Purchasing a licensed band requires a several hundred billion dollar investment. Unlicensed bands, in contrast, are free to use but are unregulated. Both licensed and unlicensed bands are limited in the amount of bandwidth allotted. Thus system designers can not look to an increase in bandwidth in designing next generation wireless communication systems.

Operating over a fixed bandwidth means that a high spectral efficiency is required. Spectral efficiency is defined as the bits per second transmitted divided by the system bandwidth (in Hertz). The capacity of a wireless link is the largest spectral efficiency that can be reliably obtained for a given amount of bandwidth. Thus, we would like to choose a wireless link with a large capacity so that we can obtain a good spectral efficiency when operating.

Wireless systems must also operate reliably. They must maintain the same link quality that is found in wireline communications. In communication theory, reliability quantitatively is described by the probability of error at the receiver. A reliable link will have a low probability of error. Variations in the received signal strength due to mobile motion and the propagation environment make link reliability a challenging problem.

Transmit power is another commodity in wireless system design. Unfortunately, transmit power is also limited because of issues such as battery power, Federal Communications Commission (FCC) regulations, and health concerns. Transmit power, therefore, can not be used as a solution to the spectral efficiency and reliability design challenges. System designers must instead look to smart communications and signal processing algorithms for next generation communications.

1.2 Multiple-Input Multiple-Output Wireless Systems

One solution to the problem of simultaneously maximizing spectral efficiency and reliability is the use of wireless systems employing multiple antennas at both the transmitter and receiver, also known as multiple-input multiple-output (MIMO) wireless systems. MIMO wireless systems provide gains in Shannon capacity and link reliability over single-antenna wireless systems by signaling over the spatial dimension of a wireless link [1]–[3].

The concept behind the MIMO capacity gain can be understood from the use of sufficiently spaced multiple antennas. Multiple antenna receivers have been used for over fifty years in receive combining systems (see for example [4], [5]). Multiple antenna transmitters were studied in [6] for use with single antenna receivers. Paulraj and Kailath proposed using a multiple antenna transmitter and receiver [7], however, it was not until Telatar's work in [1], [8] that the effect of a multiple antenna transmitter and receiver on the capacity was fully understood.

Multiple antenna technology has been standardized in wireless metropolitan access networks [9], wideband code division multiple access (WCDMA) [10] and is expected to be a critical component of next-generation wireless local area networks (LANs) [11]. Most of the major industrial focus, however, has been limited to two transmit and/or receive antennas due to practical considerations. Because of the performance gains that larger transmit and receive antenna arrays can provide, it is of utmost importance for researchers to find practical methods to take advantage of the MIMO capacity.

1.3 Closed-Loop Space-Time Signaling

The concept of space-time coding was introduced in [12]–[14]. Space-time codes work using sophisticated signal processing at both the transmitter and receiver to leverage the MIMO spatial diversity advantage. Space-time coding has become synonymous with open-loop* MIMO transmission. Space-time codes improve performance over single-antenna modulation schemes but lose array gain compared with other space-time techniques because they do not use any kind of channel information [15], [16].

An alternative signaling method to space-time coding is spatial multiplexing [3], [16], [17]. In spatial multiplexing systems, a single symbol stream is demultiplexed into multiple symbol streams. Each symbol stream is assigned to a transmit antenna, and the symbols are then transmitted without the benefit of any transmit spatial diversity. Spatial multiplexing is often not thought of as a space-time code because there is no spatial or temporal redundancy (i.e. each symbol is only sent over one antenna during one transmission) [16].

Closed-loop† MIMO transmission originated with the extension of beamforming and combining to MIMO channels [18]. In beamforming, a transmitted symbol is projected onto a data vector that is then sent from a multiple-antenna transmitter (i.e. each entry of the vector is sent to a different transmit

*Open-loop transmission is defined as signaling without any knowledge of the channel. This means that the space-time signals are designed without respect to the current channel conditions, average channel parameters, or receiver feedback.

†Closed-loop transmission is defined as signaling with some form of channel knowledge or feedback from the receiver.

antenna). Combining works by linearly weighting and summing the outputs of multiple receive antennas to yield an estimate of the transmitted signal. The idea with this early work was to convert the MIMO matrix channel into a single-input single-output channel that was resilient to fading. MIMO beamforming and combining were later studied in [19]–[27] for the case where exact channel state information (CSI) was available at the transmitter.

The ideas behind the single substream (i.e. only one symbol transmitted at each channel use) beamforming and combining methods were extended to multiple substream (or equivalently spatial multiplexing [3]) linear precoding systems in [28]–[35]. Linear precoding can be understood as sending a vector obtained from right-multiplying a symbol vector, a multi-dimensional complex column vector, by a precoding matrix. These precoding methods were based on a variety of assumptions about the transmitter's channel knowledge such as full CSI [28], [29], knowledge of the first-order statistics [30]–[32], or knowledge of the second order statistics of the channel [33]–[35].

Precoded spatial multiplexing is actually a combination of beamforming and spatial multiplexing because the number of substreams is reduced to allow for spatial redundancy. Systems that make tradeoffs such as this will be defined as diversity-multiplexing tradeoff systems. Diversity gain and multiplexing gain are two parameters of a space-time signaling strategy that can be given explicit mathematical definitions [36]. Beamforming and space-time coding are two transmission techniques that achieve maximum diversity gain [12], [14], [37], [38], while spatial multiplexing achieves full multiplexing gain [3]. It was shown in [39] that the number of substreams transmitted can be adapted based on current CSI to simultaneously maximize both parameters. A multi-mode diversity-multiplexing tradeoff was studied in [40] by varying the number of antennas used for transmission. Because of the precoding interpretation of antenna selection, this can be thought of as an adaptive precoder.

Closed-loop space-time block coding was first introduced in [41] using complete channel knowledge to decompose the multiple antenna channel into parallel sub-channels. More practical closed-loop space-time codes were proposed in [32], [42]–[51]. All of these references use some form of linear precoding by applying a linear transformation to the spatio-temporal block (or matrix) before transmission.

Most of these closed-loop methods unfortunately require either complete CSI or statistics of the channel. These assumptions are unrealistic in many wireless systems such as those using frequency division duplexing since the forward and reverse channels are approximately independent. For this reason, more practical methods of closed-loop signaling have been studied where the receiver sends back a limited number of bits to the transmitter based on current CSI. These were studied in great detail for beamforming and combining in [18], [19], [22], [27], [52]–[63].

Introductory work on limited feedback, closed-loop space-time coding was done in [43]–[48], [51]. Each of these papers restricted the space-time code to an orthogonal space-time block code (OSTBC) (see [13], [14]) and studied the problem as an antenna selection problem, where $\lceil \log_2 \binom{M_t}{M} \rceil$ bits are allocated for feedback per channel realization for an M_t transmit antenna system and M antenna space-time code, or as a problem of directly quantizing the channel. Quantizing the channel, however, should in general be avoided because of the sensitivity of the eigen-structure to quantization error [64] and the large amount of feedback necessary for even coarse quantization of a matrix channel. For example, a four-by-four MIMO system with entry-by-entry quantization according to two bits per real part and two bits per imaginary part would require at least 64 bits of quantization! This is a dramatic amount of feedback for this extremely coarse quantization.

Limited feedback precoding for spatial multiplexing systems has been addressed only in the context of antenna selection [65]–[71]. While antenna selection is easily implemented, it suffers in performance because of the restrictions on the form of the precoding matrices [72], [73]. Furthermore, antenna selection does not allow the allocation of more feedback to further improve performance.

1.4 Contributions

Previous work on closed-loop techniques for MIMO communications shows that there is great promise in sending partial CSI[‡] to the transmitter. Unfortunately, prior work fails to address the fundamental problem of how to provide CSI to the transmitter optimally using a limited amount of feedback. For this reason, this dissertation studies limited feedback in MIMO communications. The work provides a practical solution for implementing high performance, closed-loop space-time signaling techniques. The proposed techniques could find application in wireless LANs (ex. the IEEE 802.11N study group), wireless fixed-base access (ex. the IEEE 802.16 working group), and mobile wireless access (ex. the IEEE 802.20 working group). The contributions are as follows:

- A limited feedback precoding methodology for orthogonal space-time block

[‡]Partial CSI is defined as any form of channel knowledge at the transmitter that is not perfect channel knowledge. This includes statistical knowledge, knowledge of a quantized channel, knowledge of a quantized beamforming vector, etc.

codes transmitting over narrowband, matrix Rayleigh fading channels.

1. Developing a system model where the precoder is chosen from a finite set, or codebook, of possible precoding matrices designed off-line and known to both the transmitter and receiver.
2. Developing a selection criterion for selecting the optimal precoder matrix from the codebook using a bound on the symbol error rate.
3. Deriving a codebook design criterion based on the selection criterion.

• A generalized method of limited feedback precoding for spatial multiplexing systems transmitting over narrowband, matrix Rayleigh fading channels (partially reported in [72], [73]).

1. Developing a system model where the precoder is chosen from a finite set, or codebook, of possible precoding matrices designed off-line and known to both the transmitter and receiver.
2. Developing criteria for selecting the optimal precoder matrix from the codebook.
3. Deriving criteria, based on the selection criterion chosen, for designing the precoding matrix codebook.

• Multi-mode precoding for narrowband, matrix Rayleigh fading channels (partially reported in [40]).

1. Developing a diversity-multiplexing tradeoff approach that generalizes the results in [35], [36], [39], [74], [75] to allow any number of substreams.
2. Deriving selection criteria for choosing the optimal number of substreams.

3. Designing multi-mode codebooks and feedback techniques based on the first two contributions.

1.5 Organization of Dissertation

Chapter 2 presents a detailed system level description of spatial multiplexing, space-time coding, precoding, and diversity-multiplexing tradeoff. A limited feedback framework for precoded orthogonal space-time block codes is presented in Chapter 3. Chapter 4 discusses a methodology for limited feedback precoded spatial multiplexing. A new adaptive transmission scheme called multi-mode precoding is discussed in Chapter 5. Practical effects on limited feedback are discussed in Chapter 6. Chapter 7 concludes the dissertation.

Chapter 2

Background

This chapter gives background on MIMO wireless systems for later chapters. Section 2.1 overviews the mathematical notation used throughout the dissertation. The MIMO system model is presented in Section 2.2. Beamforming and combining are reviewed in Section 2.3. Spatial multiplexing is discussed in Section 2.4. Section 2.5 overviews OSTBCs.

2.1 Notation

Much of the work in this dissertation deals with concepts from linear algebra. The (k, l) entry of a matrix $\mathbf{X}$ is denoted by $x_{k,l}$. $\mathbb{C}^m$ and $\mathbb{C}^{m\times n}$ are used to refer to the m-dimensional complex vector space and the set of $m \times n$ complex matrices, respectively. The matrix transposition, conjugate transposition, inverse, and pseudo-inverse operators are give by ^{T}, *, $^{-1}$, and †, respectively. The k^{th} largest singular value of a matrix $\mathbf{H}$ will be denoted as $\lambda_k\{\mathbf{H}\}$. This dissertation uses $\|\cdot\|_2$ for the matrix two-norm (i.e. $\|\mathbf{H}\|_2 = \lambda_1\{\mathbf{H}\}$), $\|\cdot\|_F$ for the matrix Frobenius norm (i.e. $\|\mathbf{H}\|_F^2 = \sum_k \sum_l |h_{k,l}|^2$), $\|\cdot\|_1$ for the matrix one-norm (i.e. $\|\mathbf{H}\|_1 = \max_l \sum_k |h_{k,l}|$), and $\|\cdot\|_\infty$ for the vector sup-norm. The trace is represented by $tr(\cdot)$ and the determinant by $det(\cdot)$. The set of $M_t \times M$ matrices with orthonormal columns is denoted by $\mathcal{U}(M_t, M)$. The set of $M_t \times M$ matrices with largest singular values less than one is denoted by $\mathcal{L}(M_t, M)$. Both $card(\cdot)$ and $|\cdot|$ will define functions that return the cardinality of a set. The notation $|\cdot|$ will only be used for cardinality when there is no confusion with absolute value.

$\mathcal{CN}(0, \sigma^2)$ represents the distribution of a complex random variable with independent real and imaginary parts each distributed according to the real normal distribution $\mathcal{N}(0, \sigma^2/2)$. $E_y[\cdot]$ is used to denote expectation with respect to y. The ceiling of a number is returned by $\lceil\cdot\rceil$. The function argmax (or argmin) is defined to return a single, global maximizer (or minimizer).

2.2 System Model

A MIMO wireless system with M_t transmit antennas and M_r receive antennas is shown in Fig. 2.1. An $M_t \times T$ matrix $\mathbf{X}(k)$ is generated by a coding and modulation block at transmission k. At time l of the k^{th} transmission, $x_{m,l}(k)$ is transmitted from the m^{th} transmit antenna.

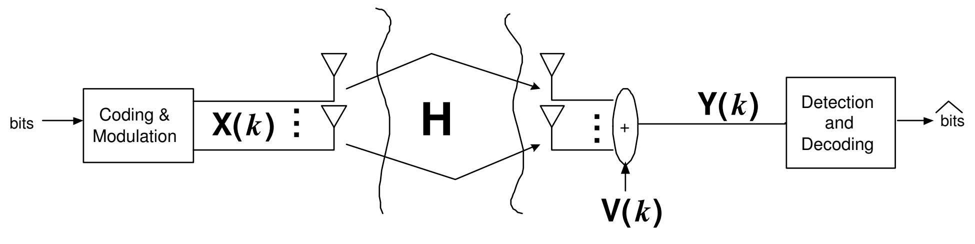

Figure 2.1: Block diagram of a MIMO system.

Assuming optimal pulse-shaping, match-filtering, and sampling, the received signal matrix for the k^{th} matrix transmission will be modeled as

$$\mathbf{Y}(k) = \mathbf{H}\mathbf{X}(k) + \mathbf{V}(k) \tag{2.1}$$

where $\mathbf{H}$ is an $M_r \times M_t$ channel matrix and $\mathbf{V}(k)$ is an $M_r \times T$ noise matrix. Because of the need to characterize the average performance of a wireless communication link, the channel matrix $\mathbf{H}$ will be modeled as a random matrix [16], [17], [76]. The channel will be modeled as a spatially uncorrelated, Rayleigh flat fading matrix. This corresponds to a random matrix with independent entries each distributed according to $\mathcal{CN}(0, 1)$ and assumes sufficient scattering and the absence of a line-of-sight component. This model assumes narrowband transmission and sufficiently spaced antennas. The channel is assumed to follow a block-fading model. With this model, the channel is constant for several blocks of T transmissions and then changes independently to a new realization. The noise matrix will be modeled as having independent entries each distributed according to $\mathcal{CN}(0, N_0)$ where N_0 is determined for the specific system. In general, the N_0 variance can be normalized out since only the signal-to-noise ratio (SNR) is of interest.

After reception, the matrix $\mathbf{Y}(k)$ is fed into a decoding and detection unit where the transmitted signal will be detected. The receiver is assumed to have perfect knowledge of $\mathbf{H}$. Since the analysis only depends on symbol-by-symbol detection, the burst index, k, will be omitted to simplify notation. The matrix $\mathbf{X}$ and the decoder to detect the matrix $\mathbf{X}$ can take a number of different forms . For this reason, Sections 2.3-2.6 will overview some common signaling methods and receiver architectures.

2.3 Beamforming

One of the earliest MIMO signaling methods was beamforming. In this method the matrix $\mathbf{X}$ can be written as

$$\mathbf{X} = \mathbf{w}s \tag{2.2}$$

where $\mathbf{w} \in \mathbb{C}^{M_t}$ (i.e. $T = 1$) and $s \in \mathbb{C}$. The vector $\mathbf{w}$ is called the beamforming vector while s is a symbol constructed from some constellation $\mathcal{S} \subset \mathbb{C}$ (ex. binary phase shift keying (BPSK), 4-quadrature amplitude modulation (4-QAM)). Power is controlled by requiring that $E_s\left[|s|^2\right] = \mathcal{E}_s$ and $\mathbf{w}^*\mathbf{w} = 1$. Beamforming converts the matrix channel $\mathbf{H}$ into an effective vector channel $\mathbf{Hw}$. The symbol s can then be detected by performing single-dimensional maximum likelihood (ML) detection on the symbol estimate

$$\mathbf{z}^*\mathbf{Y} = \mathbf{z}^*\mathbf{Hw}s + \mathbf{z}^*\mathbf{V} \tag{2.3}$$

where $\mathbf{z}$ is known as a combining vector. When $\mathbf{z} = \mathbf{Hw}/\|\mathbf{Hw}\|_2$ the combining-ML receiver corresponds to optimal detection of s.

Note that the vector $\mathbf{w}$ must be chosen as a function of the channel in order to provide any kind of performance benefit. For example, assuming an optimal combiner,

$$\mathbf{w} = \underset{\mathbf{w}^*\mathbf{w}=1}{\operatorname{argmax}} \|\mathbf{Hw}\|_2 \tag{2.4}$$

in order to minimize the probability of error and maximize the capacity [23], [26], [37]. This means that beamforming is a closed-loop MIMO signaling method. Limited feedback methods for beamforming were developed in [18], [19], [22], [27], [52]–[63].

2.4 Spatial Multiplexing

Spatial multiplexing is a capacity-achieving MIMO signaling technique. Traditional open-loop spatial multiplexing and closed-loop precoded spatial multiplexing will be discussed.

2.4.1 Open-Loop

In spatial multiplexing, a symbol stream $s_1, s_2, s_3, \ldots, s_{M_t}$ with $s_i \in \mathcal{S}$ for all i, where $\mathcal{S}$ is some constellation, is demultiplexed into a vector $\mathbf{X} = \mathbf{s} = [s_1\ s_2\ \ldots\ s_{M_t}]^T$ [3]. Thus once again, $T = 1$. The received vector is then

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{v}. \tag{2.5}$$

Power constraints require that $E_{\mathbf{s}}[\mathbf{s}\mathbf{s}^*] = (\mathcal{E}_s/M_t)\,\mathbf{I}_{M_t}$ so that the transmit power can be conserved. Note that lower-case letters have been adopted in this discussion to avoid confusion with the matrix space-time coding techniques presented later.

Unlike beamforming, spatial multiplexing sends multiple symbols or multiple substreams at each channel use. This property complicates detection since the simple linear combining and single-dimensional detection are no longer optimal. Optimal detection is actually multi-dimensional ML detection where vectors in $\mathcal{S}^{M_t}$ are detected rather than symbols in $\mathcal{S}$ as was the case for beamforming [16].

A number of different sub-optimal methods have been proposed for detection of $\mathbf{s}$. One easily implemented method is the use of a linear receiver. In this form, the receiver decodes to a vector

$$\hat{\mathbf{s}} = \mathbf{Q}(\mathbf{G}\mathbf{y}) \tag{2.6}$$

where $\mathbf{G}$ is an $M_t \times M_r$ matrix and $\mathbf{Q}$ is a function that performs entry-by-entry ML detection. Note that entry-by-entry ML detection grows linearly with M_t as opposed to multi-dimensional ML detection which grows exponentially with M_t. The matrix $\mathbf{G}$ can take many forms depending on system needs. Two methods for choosing $\mathbf{G}$ that will be used later are zero-forcing (ZF) decoding and minimum mean squared error (MMSE) decoding. If ZF decoding is employed $\mathbf{G} = \mathbf{H}^{\dagger}$ and if MMSE decoding is employed $\mathbf{G} = (MN_0/\mathcal{E}_s\mathbf{I}_{M_t} + \mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^*$ [3], [77].

2.4.2 Closed-Loop Precoding

Notice that spatial multiplexing only sends each symbol one time from one antenna. If one antenna is shadowed, for example, the receiver will lose the symbol regardless of the form of receive processing because of the lack of spatial or temporal redundancy. One method for overcoming this problem is precoding. The basic idea behind precoding is to provide improved diversity advantage by attempting to send the spatial multiplexing vector only on the eigen-directions that provide the best performance. When spatial multiplexing is precoded, $\mathbf{X}$ is of the form

$$\mathbf{X} = \mathbf{F}\mathbf{s} \tag{2.7}$$

where $\mathbf{F}$ is an $M_t \times M$ matrix with $M \leq M_t$. The matrix $\mathbf{F}$ can be thought of as converting the $M_t \times M_r$ antenna system into an effective $M \times M_r$ antenna system. The symbol vector $\mathbf{s}$ is then designed as for an M antenna transmission. This yields an input/output relation

$$\mathbf{y} = \mathbf{H}\mathbf{F}\mathbf{s} + \mathbf{v} \tag{2.8}$$

where $\mathbf{F}$ is chosen from some feasible set* $\mathcal{F}$. The received vector can then be decoded just as traditional spatial multiplexing by thinking of $\mathbf{HF}$ as an effective channel.

To date, there have been two different areas of emphasis in closed-loop spatial multiplexing: MMSE precoding and antenna selection. MMSE precoding was studied in [28], [29], [78]. The feasible sets considered were

$$\mathcal{F} = \{\mathbf{F} \in \mathbb{C}^{M_t \times M} \;\mid\; \lambda_{\max}\{\mathbf{F}\} \leq \mathcal{P}\} \tag{2.9}$$

or

$$\mathcal{F} = \{\mathbf{F} \in \mathbb{C}^{M_t \times M} \;\mid\; tr\left(\mathbf{F}\mathbf{F}^*\right) \leq \mathcal{P}\} \tag{2.10}$$

where $\lambda_{\max}\{\cdot\}$ denotes the maximum singular value and $\mathcal{P}$ is a power constraint. Assuming MMSE linear decoding, a selection criterion is to minimize the trace or determinant of the mean squared error matrix between the soft symbol vector estimate $\mathbf{Gy}$ and the transmitted vector $\mathbf{s}$ [28], [78]. This average mean squared error matrix is given by

$$\overline{\mathrm{MSE}}(\mathbf{F}) = \frac{\mathcal{E}_s}{M}\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0}\mathbf{F}^*\mathbf{H}^*\mathbf{H}\mathbf{F}\right)^{-1}. \tag{2.11}$$

The precoder can then be chosen according to

$$\mathbf{F} = \operatorname*{argmin}_{\mathbf{F}' \in \mathcal{F}} m\left(\overline{\mathrm{MSE}}(\mathbf{F}')\right) \tag{2.12}$$

where $m(\cdot)$ is either $tr(\cdot)$ [28], $det(\cdot)$ [28], [29], or a weighted version of the trace [78]. Note that MMSE precoding assumes an uncountable feasible set and perfect CSI. The perfect CSI assumption, in particular, makes MMSE precoding difficult from a practical perspective.

*A feasible set is the set that a cost function is optimized over.

A more practical precoder is antenna selection. In this case $\mathbf{F}$ chooses M out of the M_t antennas to transmit on and turns the other $M_t - M$ antennas off. Many different criteria have been proposed for antenna selection. To describe these criteria, assume that the set of possible subsets has been indexed from 1 to $\binom{M_t}{M}$. Maximum minimum distance antenna selection [79] works well for ML decoding by choosing the subset index

$$k = \underset{1 \le i \le \binom{M_t}{M}}{\operatorname{argmax}} \min_{\mathbf{s}_1 \neq \mathbf{s}_2} \|\mathbf{H}_{[i]}(\mathbf{s}_1 - \mathbf{s}_2)\|_2 \tag{2.13}$$

where $\mathbf{H}_{[k]}$ is the channel corresponding to signaling on subset k and $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{S}^M$. Maximizing the minimum singular value of $\mathbf{HF}$, $\lambda_{\min}\{\mathbf{HF}\}$, (see for example [71], [79]) has been used to approximately maximize the minimum distance between codewords for ML decoding and maximize the minimum SNR of over all substreams for linear receivers. This criterion chooses

$$k = \underset{1 \le i \le \binom{M_t}{M}}{\operatorname{argmax}} \lambda_{\min}\{\mathbf{H}_{[i]}\}. \tag{2.14}$$

Yet another criterion is to try to maximize the mutual information assuming an uncorrelated complex Gaussian source [68], [80]. The mutual information assuming an uncorrelated complex Gaussian source given subset choice $\mathbf{H}_{[k]}$ is

$$I(\mathbf{H}_{[k]}) = \log_2 det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{M N_0} \mathbf{H}_{[k]}^* \mathbf{H}_{[k]} \right). \tag{2.15}$$

Therefore a capacity inspired selection criterion is given by

$$k = \underset{1 \le i \le \binom{M_t}{M}}{\operatorname{argmax}} I(\mathbf{H}_{[i]}). \tag{2.16}$$

The only previously proposed closed-loop precoding methods that can be directly implemented in systems without transmit channel knowledge are the antenna selection precoders. Antenna selection is always a limited feedback system because there are only $\binom{M_t}{M}$ ways to choose M antenna subsets from M_t antennas so $\mathbf{F}$ can be conveyed from the receiver to the transmitter using $\lceil \log_2 \binom{M_t}{M} \rceil$ bits.

Unfortunately, antenna selection suffers in performance when compared with the unconstrained precoders (see for example the simulations in [72],[73]). One of the objectives of this dissertation is to extend the concepts and benefits of antenna selection to limited feedback, unconstrained precoders. This would allow the limited feedback system to use any number of bits of feedback rather than simply $\lceil \log_2 \binom{\mathrm{M}_t}{M} \rceil$ and benefit from the increased array gain available with more feedback.

Example: An example of the performance degradation on a 5×4 system using 4-QAM and four substreams is presented in Fig. 2.2. Antenna selection using the minimum singular value with a ZF decoder and MMSE precoding are plotted along with 4×4 spatial multiplexing using ZF and ML decoding. MMSE with the determinant cost function outperforms antenna selection by 1.3dB. As well MMSE with the trace cost function outperforms antenna selection by approximately 1.7dB.

2.5 Orthogonal Space-Time Block Coding

OTSBCs were some of the first proposed space-time block codes [13],[14]. The following discussion will overview OSTBCs for use in open-loop and closed-loop systems. The explanation of OSTBCs will follow from the linear dispersion formulation taken from [81].

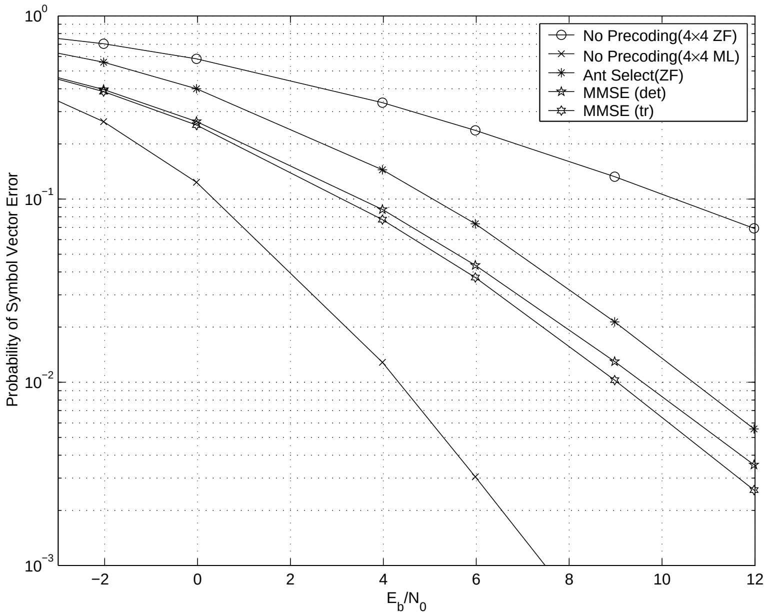

Figure 2.2: Probability of error comparison of antenna selection and MMSE precoding.

2.5.1 Open-Loop

An OSTBC codeword is created from a set of n_s symbols $s_1, s_2, \ldots, s_{n_s}$ all taken from the same constellation $\mathcal{S} \subset \mathbb{C}$. An $M_t \times T$ codeword is formulated as

$$\mathbf{X} = \sum_{i=1}^{n_s} (\mathbf{s}_i \mathbf{A}_i + \mathbf{s}_i^* \mathbf{B}_i) \tag{2.17}$$

where $\mathbf{A}_i$ and $\mathbf{B}_i$ for all i are matrices that satisfy several orthogonality properties [16]. Note that for OSTBCs, $T > 1$. Because of the power constraint properties,

$$\mathbf{X}\mathbf{X}^* = \left(\sum_{i=1}^{n_s} |s_i|^2 \right) \mathbf{I}_{M_t} \tag{2.18}$$

[16]. For power constraint reasons, it is assumed that $E_{s_i}\left[|s_i|^2\right] = \mathcal{E}_s/M_t$.

The orthogonality properties allow the received matrix $\mathbf{Y}$ to have a simple linear decoding structure [13], [14]. They also allow for a simple upper bound on the probability of error [12], [43], [45],

$$Pr(ERROR) \leq e^{-\gamma\|\mathbf{H}\|_F^2} \tag{2.19}$$

where γ is a constant scale factor that is a function of $\mathcal{S}$ and M_t. Because γ is constant, the performance of OSTBC is thus a decreasing function of the channel power $\|\mathbf{H}\|_F$.

2.5.2 Closed-Loop Precoding

Closed-loop OSTBC using precoding was studied in [42]–[49]. When precoding is employed,

$$\mathbf{X} = \mathbf{FC} \tag{2.20}$$

where $\mathbf{F}$ is an $M_t \times M$ matrix chosen from a feasible set $\mathcal{F}$ and $\mathbf{C}$ is an $M \times T$ OSTBC codeword. Thus the OSTBC will be designed as if it were being sent over M transmit antenna.

Current work in precoded OSTBCs has used a selection criterion for $\mathbf{F}$ that minimizes (2.19) using an effective channel $\mathbf{HF}$ [42]–[49]. Thus, minimizing (2.19) is equivalent to maximizing $\|\mathbf{HF}\|_F$, so $\mathbf{F}$ can be chosen by

$$\mathbf{F} = \underset{\mathbf{F}' \in \mathcal{F}}{\operatorname{argmax}} \|\mathbf{HF}'\|_F. \tag{2.21}$$

Various feasible sets for $\mathbf{F}$ have been considered. The restriction of $\mathbf{F}$ to diagonal matrices with a trace constraint was studied in [48], [49]. Note that this method is not a limited feedback method because $\mathcal{F}$ is uncountable. The authors in [42]–[44], [46] consider precoders where $M = M_t$ and

$$\mathcal{F} = \{\mathbf{F} \in \mathbb{C}^{M_t \times M} \ \mid \ \|\mathbf{F}\|_F^2 \leq M\}. \tag{2.22}$$

All of these precoding techniques lead to uncountable $\mathcal{F}$. Antenna selection uses a finite feasible set and chooses the antenna subset k such that

$$k = \underset{1 \le i \le \binom{M_t}{M}}{\operatorname{argmax}} \|\mathbf{H}_{[i]}\|_F. \tag{2.23}$$

Only two of these approaches work when the transmitter has no CSI. Once again, antenna selection is a limited feedback system because the subset chosen can be conveyed from receiver to transmitter using $\lceil \log_2 \binom{M_t}{M} \rceil$ bits. A limited feedback method was proposed in [44], [46] by requiring the receiver to convey a quantized version of the channel, $\mathbf{H}_{quant}$, using a limited number of bits to the transmitter. With this approach $\mathbf{F}$ is chosen from

$$\mathbf{F} = \underset{\mathbf{F}' \in \mathcal{F}}{\operatorname{argmax}} \|\mathbf{H}_{quant}\mathbf{F}'\|_F. \tag{2.24}$$

Unfortunately a design method for constructing a quantized feasible set (i.e. finite $card(\mathcal{F})$) other than antenna selection has not yet been derived. The limited feedback unconstrained precoders using a quantized channel are not easily implemented and require difficult optimizations [44], [46]. Therefore, one of the main objectives of the proposed research is to understand how quantized feasible sets for precoded OSTBCs should be designed. Solving this problem will convert the difficult optimizations in [44], [46] a simple brute force search over the codebook matrices.

2.6 Diversity-Multiplexing Tradeoff

A simple characterization of space-time systems is their diversity-multiplexing tradeoff [36]. The parameters of diversity and multiplexing have precise mathematical definitions. A system is said to have diversity gain (also called diversity order) of d if

$$d = -\lim_{\mathcal{E}_s/N_0 \to \infty} \frac{\log\left(Pr(ERROR)\right)}{\log\left(\mathcal{E}_s/N_0\right)}. \tag{2.25}$$

A system has a multiplexing gain of r if the supported data rate satisfies

$$R(\mathcal{E}_s/N_0) \approx r \log_2(\mathcal{E}_s/N_0) \tag{2.26}$$

where $R(\mathcal{E}_s/N_0)$ is the rate supported when the SNR is $\mathcal{E}_s/N_0$.

These two parameters can fundamentally characterize the performance of a MIMO signaling scheme [36]. Spatial multiplexing, for example, achieves maximum multiplexing gain [3], while OSTBC achieves maximum diversity gain [13], [14]. Space-time block codes have been designed that tradeoff the two gains (see for example codes designed from [81]), but a simple tradeoff can be made using closed-loop MIMO techniques [39].

The basic idea behind [39] is to fix the rate of the signaling scheme per channel use and then to choose spatial multiplexing when $d_{min,SM} > d_{min,OSTBC}$ and to choose OTSBC when $d_{min,SM} < d_{min,OSTBC}$ where $d_{min,SM}$ is the received minimum distance of spatial multiplexing and $d_{min,OSTBC}$ is the received minimum distance for an OSTBC. This method achieves the maximum diversity gain [39] and can be easily implemented using 1 bit of feedback. Unfortunately, OSTBCs suffer from a rate loss for $M_t > 2$ for all constellations except pulse amplitude modulation (PAM) and have design methods that do not easily generalize to arbitrary M_t [14]. Therefore, a generalized method of [39] that is not dependent on OSTBCs is of practical interest.

Chapter 3

Limited Feedback Precoding for Space-Time Codes

This chapter proposes a limited feedback framework for precoded orthogonal space-time block codes. The chapter is organized as follows. Section 3.1 provides a general overview of the system under consideration. A selection criterion for limited feedback OSTBC precoding is derived in Section 3.2. The codebook design problem is solved in Section 3.3. Section 3.4 shows that the designed codebooks provide full diversity order in Rayleigh fading channels. We present Monte Carlo simulation results in Section 3.5.

3.1 System Overview

An M_t transmit antenna and M_r receive antenna limited feedback precoded OSTBC system is illustrated in Fig. 3.1. A block of n_s symbols $s_1, s_2, \ldots, s_{n_s}$

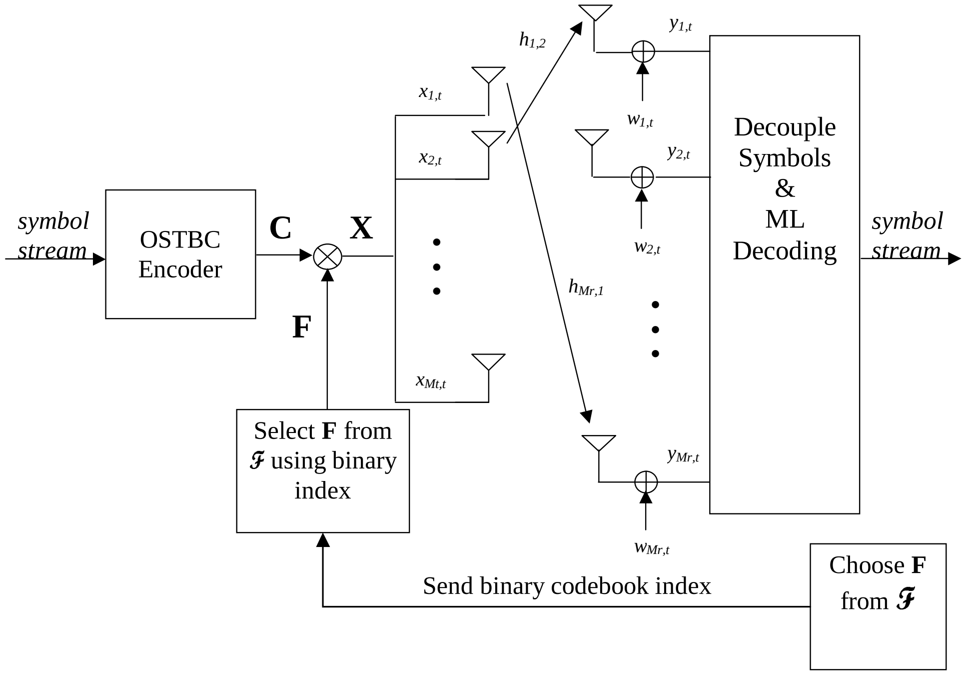

Figure 3.1: Block diagram of a limited feedback precoding system.

taken from a constellation $\mathcal{S}$ is fed into an OSTBC encoder producing an $M \times T$ (with $M < M_t$) space-time code matrix $\mathbf{C}(k) = [\mathbf{c}_1(k)\ \mathbf{c}_2(k)\ \ldots\ \mathbf{c}_T(k)]$ at the k^{th} transmission. Note that for any codeword

$$\mathbf{C}(k)\mathbf{C}(k)^* = \left(\sum_{l=1}^{n_s} |s_l|^2\right) \mathbf{I}_M \tag{3.1}$$

[16]. For power constraint reasons, we will assume that $E_{s_l}\left[|s_l|^2\right] = 1$.

Before transmission, the space-time codeword is premultiplied by an $M_t \times M$ matrix $\mathbf{F}$. The matrix $\mathbf{F}$ is chosen by a function $f : \mathbb{C}^{M_r \times M_t} \rightarrow \mathcal{F} = \{\mathbf{F}_1,\ \mathbf{F}_2, \ldots,\ \mathbf{F}_N\}$ evaluated using the current channel conditions and then conveyed to the transmitter using $\lceil \log_2 N \rceil$ bits of feedback. The set $\mathcal{F}$ can be thought of as a codebook and functions similarly to those found in the standard source coding literature (see for example [82]). This codebook is designed offline and known to both the transmitter and receiver. We will impose a peak power limit for all $\mathbf{F} \in \mathcal{F}$ by requiring that

$$\max_{\mathbf{s} \in \mathcal{S}^M} \frac{\|\mathbf{F}\mathbf{s}\|_2}{\|\mathbf{s}\|_2} \leq 1$$

for any constellation $\mathcal{S} \subseteq \mathbb{C}$. Because $\mathcal{S}$ is non-empty and $0 \notin \mathcal{S}$, this constraint corresponds to the assumption that $\lambda_1\{\mathbf{F}\} \leq 1$ for all $\mathbf{F} \in \mathcal{F}$. We will further restrict the form of the matrices in $\mathcal{F}$, but the justification will be presented in Section 3.2.

We assume that the transfer function between the transmitter and receiver can be modeled as a spatially uncorrelated, memoryless linear channel that is constant over several codeword transmissions before independently taking on a new value. Assuming optimal pulse shaping, match-filtering, and sampling, the received signal can thus be written as

$$\mathbf{Y} = \sqrt{\frac{\rho}{M}}\mathbf{HFC} + \mathbf{W} \tag{3.2}$$

where $\mathbf{H}$ is the $M_r \times M_t$ channel matrix with independent entries distributed as $\mathcal{CN}(0,1)$, $\mathbf{W}$ is an $M_r \times T$ noise matrix with independent entries distributed according to $\mathcal{CN}(0,1)$, and ρ is the signal-to-noise ratio (SNR). $\mathbf{F}$ corresponds to the evaluation of the mapping function for the current channel realization $\mathbf{H}$, i.e.

$$\mathbf{F} = f(\mathbf{H}).$$

Note that we have suppressed the temporal parameter k in $\mathbf{C}(k)$ because of our channel model and interest in the use of codeword-by-codeword detection. After reception, the receiver performs maximum likelihood (ML) decoding on the OSTBC using the received matrix $\mathbf{Y}$ and the effective channel $\mathbf{HF}$.

Problem Statement: Given a system illustrated by Fig. 3.1 and described by (3.2), two main problems arise. The first problem we address is the selection of an optimal precoder for given channel from a codebook of matrices $\mathcal{F}$. The second problem is how to design the set $\mathcal{F}$ given the chosen selection criterion. Solving these problems is the objective of this chapter.

3.2 Selection Criterion

The first issue with limited feedback unitary precoding we address is the design of the precoder selection mapping $f(\cdot)$. We assume an arbitrary set $\mathcal{F} \subset \mathcal{L}(M_t, M)$ and ML decoding.

Our goal in this chapter is to minimize the symbol error rate (SER) given $\mathbf{H}$, $Pr(ERROR \mid \mathbf{H})$. Closed-form expressions for the SER would be extremely difficult to obtain because they would be a function of the effective channel $\mathbf{HF}$ that is in general no longer matrix Rayleigh fading and the selection function $f(\cdot)$.

We will therefore take the approach of [14], [42], [43], [45] and will use a bound on the probability of error. Using the ML detection properties of OSTBCs, it can be shown that

$$Pr(ERROR \mid \mathbf{H}) \leq exp\left(-\gamma\|\mathbf{HF}\|_F^2\right) \tag{3.3}$$

where γ is a function that depends on M, ρ, and $\mathcal{S}$ [16]. Because γ is fixed, minimizing the bound in (3.3) requires the maximization of $\|\mathbf{HF}\|_F$. Note that this maximization is equivalent to choosing the matrix $\mathbf{F} \in \mathcal{F}$ that maximizes the receive minimum distance

$$d_{min} = \min_{\mathbf{C}_k \neq \mathbf{C}_l} \|\mathbf{HF}\left(\mathbf{C}_k - \mathbf{C}_l\right)\|_F \tag{3.4}$$

because of the orthogonality of the OSTBC codewords in (3.1). Therefore, we can state the following selection criterion:

Selection Criterion: Choose the limited feedback precoder according to

$$\mathbf{F} = f(\mathbf{H}) = \underset{\mathbf{F}' \in \mathcal{F}}{\operatorname{argmax}} \|\mathbf{H}\mathbf{F}'\|_F. \tag{3.5}$$

This selection criterion can be implemented by computing a matrix multiplication and Frobenius norm for each of the N codebook matrices. Ties in distance between codeword matrices are broken arbitrarily by selecting the precoding matrix with the lowest index because they occur with zero probability.

To bound the performance with quantized precoding, it is of interest to characterize an optimal unquantized precoder over the set $\mathcal{L}(M_t, M)$. Note that this matrix $\mathbf{F}_{opt}$ will not be unique over $\mathcal{L}(M_t, M)$ because for any $\mathbf{U} \in \mathcal{U}(M, M)$, $\|\mathbf{H}\mathbf{F}_{opt}\|_F = \|\mathbf{H}\mathbf{F}_{opt}\mathbf{U}\|_F$. Let the singular value decomposition of $\mathbf{H}$ be given by

$$\mathbf{H} = \mathbf{V}_L \boldsymbol{\Sigma} \mathbf{V}_R^* \tag{3.6}$$

where $\mathbf{V}_L \in \mathcal{U}(M_r, M_r)$, $\mathbf{V}_R \in \mathcal{U}(M_t, M_t)$, and $\boldsymbol{\Sigma}$ is an $M_r \times M_t$ diagonal matrix with $\lambda_i\{\mathbf{H}\}$ at entry (i, i). Let $\overline{\mathbf{V}}_R$ be the matrix formed from the first M columns of $\mathbf{V}_R$. The following lemma shows an optimal matrix in this unquantized scenario.

Lemma 1 *An optimal unquantized precoder* $\mathbf{F}_{opt}$ *over* $\mathcal{L}(M_t, M)$ *that maximizes* $\|\mathbf{H}\mathbf{F}_{opt}\|_F$ *is given by* $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R$.

Proof Let $\mathbf{F}$ have singular value decomposition $\mathbf{F} = \mathbf{U}_L \boldsymbol{\Gamma} \mathbf{U}_R^*$ where $\mathbf{U}_L \in \mathcal{U}(M_t, M_t)$, $\mathbf{U}_R \in \mathcal{U}(M, M)$, and $\boldsymbol{\Gamma}$ is an $M_t \times M$ diagonal matrix with $\lambda_i\{\mathbf{F}\}$

at entry (i, i). It then follows that

$$\begin{aligned}\|\mathbf{HF}\|_F &= \|\mathbf{\Sigma V}_R^* \mathbf{U}_L \mathbf{\Gamma}\|_F &(3.7)\\ &\leq \left\|\mathbf{\Sigma V}_R^* \mathbf{U}_L \left[\mathbf{I}_M \ \mathbf{0}\right]^T\right\|_F \\ &\leq \sqrt{\sum_{i=1}^{M} \lambda_i\{\mathbf{U}_L^* \mathbf{V}_R \mathbf{\Sigma}^T \mathbf{\Sigma V}_R^* \mathbf{U}_L\}} \\ &= \left\|\overline{\mathbf{\Sigma}}\right\|_F . &(3.8)\end{aligned}$$

by using the invariance of the Frobenius norm to unitary transformation [83] in (3.7) and the singular values bounds of $\mathbf{F}$ where $\mathbf{0}$ is $M \times (M_t - M)$ matrix of zeros and $\overline{\mathbf{\Sigma}}$ is the matrix consisting of the first M columns of $\mathbf{\Sigma}$. Equality in (3.8) is obtained when $\mathbf{F} = \overline{\mathbf{V}}_R$. ∎

Lemma 1 is interesting because it says that not only is $\lambda_1\{\mathbf{F}_{opt}\} = 1$ but $\lambda_i\{\mathbf{F}_{opt}\} = 1$ for $1 \leq i \leq M$. Intuitively, this means that we should *always* transmit at full power on the precoder effective channel modes when performing optimal precoding. This result would be expected given the mean squared error results for precoded spatial multiplexing in [28]. We are, however, interested in sub-optimal precoders constructed using a limited feedback path. The following lemma shows that *all* precoding matrices should have full power in the singular values.

Lemma 2 *If* $\mathbf{F} \in \mathcal{L}(M_t, M)$ *has a singular value decomposition* $\mathbf{F} = \mathbf{U}_L \mathbf{\Gamma} \mathbf{U}_R^*$ *and* $\lambda_M\{\mathbf{F}\} < 1$, *then* $\tilde{\mathbf{F}} = \mathbf{U}_L[\mathbf{I}_M \ \mathbf{0}]^T$ *satisfies* $\|\mathbf{HF}\|_F < \|\mathbf{H}\tilde{\mathbf{F}}\|_F$.

Proof Assume $\mathbf{F}$ has a singular value decomposition as shown in the first lemma. Then

$$\|\mathbf{HF}\|_F = \|\mathbf{HU}_L \mathbf{\Gamma}\|_F < \|\mathbf{HU}_L[\mathbf{I}_M \ \mathbf{0}]^T\|_F = \|\mathbf{H}\tilde{\mathbf{F}}\|_F$$

using the invariance of the Frobenius norm to unitary transformation [83], the singular value bound for $\mathbf{F}$, and $\tilde{\mathbf{F}} = \mathbf{U}_L[\mathbf{I}_M \ \mathbf{0}]^T$. ∎

Lemma 2 demonstrates the significant effect that the power restriction has on optimal precoders. The sum power constraint results in [28], [47], [78] all yield optimal precoders with unequal power allocation among the precoder singular values. In contrast, unequal power pouring when using a maximum singular value constraint only reduces the *total transmit power.*

Lemma 2 reveals that the precoder should *always* be designed to have singular values that are as large as possible. Because of the maximum singular value restriction, the precoder should be chosen from the set $\mathcal{U}(M_t, M)$ rather than $\mathcal{L}(M_t, M)$. For this reason we will restrict $\mathcal{F} \subset \mathcal{U}(M_t, M)$ and thus *only* design our precoders over $\mathcal{U}(M_t, M)$.

3.3 Chordal Distance Precoding: Motivation and Codebook Design

In this section, we derive a codebook design criterion that follows directly from the Section 3.2 codeword selection result. This criterion is based on a distortion function that we define in order to minimize the average SER. The criterion turns out to relate to the famous applied mathematics problem of Grassmannian subspace packing.

3.3.1 Distortion Function

To design a codebook we will propose a distortion measure that is a function of the channel and then find a codebook that minimizes the average distortion.

This distortion function must differ, however, from the distortion functions commonly used in vector quantization such as mean squared error (see for example [82]) because *we are interested in improving system performance rather than improving the quality of the estimated precoder at the transmitter.*

Essentially, we would like the quantized equivalent channel $\mathbf{HF}$ to provide SER performance close in some sense to that provided by the optimal precoded equivalent channel $\mathbf{HF}_{opt}$ where $\mathbf{F} \in \mathcal{F}$ is the "best" precoding matrix in the codebook. Along these lines, consider the total effective power $\|\mathbf{HF}\|_F^2$, which according to (3.3) relates to the SER. The loss in received channel power is expressed as

$$\min_{\mathbf{F}' \in \mathcal{F}} \left(\|\mathbf{HF}_{opt}\|_F^2 - \|\mathbf{HF}'\|_F^2 \right). \tag{3.9}$$

We propose to design the codebook by considering, as a measure of distortion, the average of the loss in received channel power given by

$$E_{\mathbf{H}} \left[\min_{\mathbf{F}' \in \mathcal{F}} \left(\|\mathbf{HF}_{opt}\|_F^2 - \|\mathbf{HF}'\|_F^2 \right) \right]. \tag{3.10}$$

Recall that $\|\mathbf{HF}_{opt}\|_F \geq \|\mathbf{HF}\|_F$ for all $\mathbf{F} \in \mathcal{U}(M_t, M)$ so the function will always be nonnegative. Notice also that minimizing this distortion function relates directly to minimizing the bound on the SER in (3.3). Thus the distortion has *physical meaning* in terms of the SER, as opposed to a mean squared error distance function.

The distortion function in (3.10) yields little insight into the codebook design problem directly, thus we will derive a tight upper bound on the distortion. Using Lemma 1 and the definition of $\overline{\mathbf{\Sigma}}$, the distortion for this arbitrary matrix is written as

$$
\begin{aligned}
&\min_{\mathbf{F}'\in\mathcal{F}} \left(\|\mathbf{H}\mathbf{F}_{opt}\|_F^2 - \|\mathbf{H}\mathbf{F}'\|_F^2\right) \\
&\quad = \min_{\mathbf{F}'\in\mathcal{F}} \left(tr\left(\overline{\mathbf{\Sigma}}\overline{\mathbf{\Sigma}}^T\right) - tr\left(\mathbf{\Sigma}\mathbf{V}_R^*\mathbf{F}'\mathbf{F}'^*\mathbf{V}_R\mathbf{\Sigma}^T\right)\right) \\
&\quad \leq \min_{\mathbf{F}'\in\mathcal{F}} \left(tr\left(\overline{\mathbf{\Sigma}}\overline{\mathbf{\Sigma}}^T\right) - tr\left(\overline{\mathbf{\Sigma}}\overline{\mathbf{V}}_R^*\mathbf{F}'\mathbf{F}'^*\overline{\mathbf{V}}_R\overline{\mathbf{\Sigma}}^T\right)\right) && (3.11) \\
&\quad = \min_{\mathbf{F}'\in\mathcal{F}} tr\left(\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\left(\mathbf{I}_M - \overline{\mathbf{V}}_R^*\mathbf{F}'\mathbf{F}'^*\overline{\mathbf{V}}_R\right)\right) && (3.12) \\
&\quad \leq \lambda_1^2\{\mathbf{H}\} \min_{\mathbf{F}'\in\mathcal{F}} tr\left(\mathbf{I}_M - \overline{\mathbf{V}}_R^*\mathbf{F}'\mathbf{F}'^*\overline{\mathbf{V}}_R\right) && (3.13) \\
&\quad = \lambda_1^2\{\mathbf{H}\} \min_{\mathbf{F}'\in\mathcal{F}} \frac{1}{2}\left\|\overline{\mathbf{V}}_R\overline{\mathbf{V}}_R^* - \mathbf{F}'\mathbf{F}'^*\right\|_F^2 && (3.14)
\end{aligned}
$$

where (3.11) follows from zeroing the least $M_t - M$ singular values of $\mathbf{H}$, (3.13) follows by substituting $\lambda_1\{\mathbf{H}\}$ for the other non-zero singular values in (3.12), and (3.14) uses the alternative representation for subspace distance [84].

To find a good codebook for many channel realizations, we are interested in the *average* distortion of our codebook given by

$$
E_{\mathbf{H}}\left[\min_{\mathbf{F}'\in\mathcal{F}} \left(\|\mathbf{H}\mathbf{F}_{opt}\|_F^2 - \|\mathbf{H}\mathbf{F}'\|_F^2\right)\right]. \tag{3.15}
$$

Using (3.14) and the independence of $\mathbf{\Sigma}$ and $\mathbf{V}_R$ [85],

$$
\begin{aligned}
&E_{\mathbf{H}}\left[\min_{\mathbf{F}'\in\mathcal{F}} \left(\|\mathbf{H}\mathbf{F}_{opt}\|_F^2 - \|\mathbf{H}\mathbf{F}'\|_F^2\right)\right] \\
&\quad \leq E_{\mathbf{H}}\left[\lambda_1^2\{\mathbf{H}\}\right] E_{\mathbf{H}}\left[\min_{\mathbf{F}'\in\mathcal{F}} \frac{1}{2}\left\|\overline{\mathbf{V}}_R\overline{\mathbf{V}}_R^* - \mathbf{F}'\mathbf{F}'^*\right\|_F^2\right]. && (3.16)
\end{aligned}
$$

Thus by bounding the distortion function we can think of the limited feedback performance as being characterized by two different terms, one related to the distribution of the maximum channel singular value and another representing the "quality" of the codebook $\mathcal{F}$.

3.3.2 Codebook Design Criterion

Minimizing (3.16) requires finding the distribution of $\overline{\mathbf{V}}_R$. Recall that a matrix $\mathbf{V}$ is *isotropically distributed* on $\mathcal{U}(M_t, M)$ if for any $\mathbf{\Theta} \in \mathcal{U}(M_t, M_t)$, $\mathbf{\Theta}^*\mathbf{V} \stackrel{d}{=} \mathbf{V}$ with $\stackrel{d}{=}$ denoting equivalence in distribution [86]. The following lemma gives this distribution.

Lemma 3 *The optimal precoding matrix $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R$ for a memoryless, i.i.d. Rayleigh channel $\mathbf{H}$ is isotropically distributed on $\mathcal{U}(M_t, M)$.*

Proof First note that $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R = \mathbf{V}_R[\mathbf{I}_M\ \mathbf{0}]^T$. Since $\mathbf{V}_R$ is isotropically distributed [85], it is easily seen that $\mathbf{\Theta}^*\mathbf{F}_{opt} \stackrel{d}{=} \mathbf{F}_{opt}$. ∎

To propose a design for $\mathcal{F}$, let us review some common properties of finite subsets of $\mathcal{U}(M_t, M)$. These properties are found in the Grassmannian subspace packing literature (for example [84], [87]–[90]). The set $\mathcal{U}(M_t, M)$ defines the Stiefel manifold [90]. Each matrix in $\mathcal{U}(M_t, M)$ generates an M-dimensional subspace of the complex M_t-dimensional vector space $\mathbb{C}^{M_t}$. We will adopt Grassmannian packing notation and define the set of all column spaces of the matrices in $\mathcal{U}(M_t, M)$ to be the complex Grassmann manifold $\mathcal{G}(M_t, M)$. Thus if $\mathbf{F}_1, \mathbf{F}_2 \in \mathcal{U}(M_t, M)$ then the column spaces of $\mathbf{F}_1$ and $\mathbf{F}_2$, $\mathcal{P}_{\mathbf{F}_1}$ and $\mathcal{P}_{\mathbf{F}_2}$ respectively, are contained in $\mathcal{G}(M_t, M)$. A normalized invariant measure μ is induced on $\mathcal{G}(M_t, M)$ by the Haar measure on $\mathcal{U}(M_t, M)$. This measure allows the computation of volumes within $\mathcal{G}(M_t, M)$. Subspaces within the Grassmann manifold can be related by their distance from each other. The *chordal distance* between the two subspaces $\mathcal{P}_{\mathbf{F}_1}$ and $\mathcal{P}_{\mathbf{F}_2}$ is given by

$$d(\mathbf{F}_1, \mathbf{F}_2) = \frac{1}{\sqrt{2}}\|\mathbf{F}_1\mathbf{F}_1^* - \mathbf{F}_2\mathbf{F}_2^*\|_F. \tag{3.17}$$

Let $\mathcal{V} = \{\mathcal{P}_{\mathbf{F}_1}, \mathcal{P}_{\mathbf{F}_2}, \ldots, \mathcal{P}_{\mathbf{F}_N}\}$ be the set of column spaces where $\mathcal{P}_{\mathbf{F}_k}$ is the column space of $\mathbf{F}_k$. This set $\mathcal{V} \subset \mathcal{G}(M_t, M)$ is a *packing* of subspaces in $\mathcal{G}(M_t, M)$. A packing can be described by its minimum distance

$$\delta = \min_{1 \leq k < l \leq N} d(\mathbf{F}_k, \mathbf{F}_l). \tag{3.18}$$

The Grassmannian subspace packing problem is the problem of finding the set of N subspaces in $\mathcal{G}(M_t, M)$ such that δ is as large as possible.

Consider the open ball in $\mathcal{G}(M_t, M)$ of radius $\delta/2$ defined as

$$\mathcal{B}_{\mathbf{F}_k}(\delta/2) = \{\mathcal{P}_{\mathbf{U}} \in \mathcal{G}(M_t, M) \;\mid\; d(\mathbf{U}, \mathbf{F}_k) < \delta/2\}. \tag{3.19}$$

Notice that the balls are disjoint by the triangle inequality of metrics [84]. This observation allows the *density* of a chordal subspace packing to be defined as

$$\Delta(\mathcal{F}) = \mu\left(\bigcup_{k=1}^{N} \mathcal{B}_{\mathbf{F}_k}(\delta/2)\right) = \sum_{k=1}^{N} \mu\left(\mathcal{B}_{\mathbf{F}_k}(\delta/2)\right). \tag{3.20}$$

Using the density, we find the probability of the isotropically distributed $\overline{\mathbf{V}}_R$ falling in one of the sets $\mathcal{B}_{\mathbf{F}_k}(\delta/2)$ can be expressed as

$$Pr\left(\overline{\mathbf{V}}_R \in \bigcup_{k=1}^{N} \mathcal{B}_{\mathbf{F}_k}(\delta/2)\right) = \Delta(\mathcal{F}). \tag{3.21}$$

For large M_t, it has been shown in [84] that

$$\Delta(\mathcal{F}) \approx N\left(\frac{\delta}{2\sqrt{M}}\right)^{2M_tM + o(M_t)}. \tag{3.22}$$

We can thus bound the "codebook quality" term in (3.16) as

$$\begin{aligned} &E_{\mathbf{H}}\left[\min_{\mathbf{F}' \in \mathcal{F}} \frac{1}{2}\left\|\overline{\mathbf{V}}_R\overline{\mathbf{V}}_R^* - \mathbf{F}'\mathbf{F}'^*\right\|_F^2\right] \\ &\quad \leq \frac{1}{4}\delta^2\Delta(\mathcal{F}) + M(1 - \Delta(\mathcal{F})) \\ &\quad \approx M + N\left(\frac{\delta}{2\sqrt{M}}\right)^{2M_tM + o(M_t)}\left(\frac{1}{4}\delta^2 - M\right). \end{aligned} \tag{3.23}$$

Differentiating (3.23) and noting $\delta < \sqrt{M}$ easily shows that the bound is a decreasing function of δ when $2M_tM + o(M_t) > 2/3$. We know from the probabilistic analysis in [59] that this is always satisfied for any M_t when $M = 1$ because $o(M_t) = -2$. As well, we also know that this assumption is also satisfied for large M_t because of the definition of the little-o Landau symbol. We have experimentally verified this assumption for many other M and conjecture that it is always true. Thus practically (3.23) is minimized by maximizing δ. We have now established that designing low distortion codebooks is equivalent to packing subspaces in the Grassmann manifold using the chordal distance metric.

Therefore we now understand how to design codebooks for limited feedback precoding. Maximizing the minimum subspace distance between any pair of codebook column spaces approximately minimizes our distortion bound. Thus, we state the following design method for creating limited feedback precoding codebooks.

Codebook Design Summary: The codebook $\mathcal{F} = \{\mathbf{F}_1,\ \mathbf{F}_2, \ldots,\ \mathbf{F}_N\}$ should be designed such that $\delta = \min_{1 \leq k < l \leq N} d(\mathbf{F}_k, \mathbf{F}_l)$ is as large as possible.

3.3.3 Practical Codebook Designs

Finding good precoder codebooks from Grassmannian packings for arbitrary $M_t,\ M$, and N is actually quite difficult [87]–[89]. For instance, in the simplest case of $M = 1$ where the Rankin lower bound on line packing correlation [89] can be employed, packings that achieve equality with the lower bound are often impossible to design. A comprehensive tabulation of real packings can be found on [91]. The most practical method for generating these packings is to use codebooks designed from the non-coherent space-time modulation designs in [87] and [92].

The search algorithm in [92] can be very easily implemented and yields codebooks with large minimum distances. The algorithm works by considering codebooks of the form

$$\mathcal{F} = \{\mathbf{F}_{DFT},\ \mathbf{\Theta}\mathbf{F}_{DFT}, \ldots,\ \mathbf{\Theta}^{N-1}\mathbf{F}_{DFT}\}$$

where $\mathbf{F}_{DFT}$ is an $M_t \times M$ matrix with $\frac{1}{\sqrt{M_t}} e^{j\frac{2\pi}{M_t}kl}$ at entry (k,l) and $\mathbf{\Theta}$ is a diagonal matrix given by

$$\mathbf{\Theta} = \begin{bmatrix} e^{j\frac{2\pi}{N}u_1} & 0 & \cdots & 0 \\ 0 & e^{j\frac{2\pi}{N}u_2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{j\frac{2\pi}{N}u_{M_t}} \end{bmatrix}$$

where

$$0 \leq u_1, \ldots, u_{M_t} \leq N-1.$$

The values for $u_1, u_2, \ldots, u_{M_t}$ are chosen according to the entries of the vector $\mathbf{u} = [u_1\ u_2\ \cdots\ u_{M_t}]^T$ from the set $\mathcal{Z} = \{\mathbf{u} \in \mathbb{Z}^{M_t} \mid \forall k,\ 0 \leq u_k \leq N-1\}$ given by

$$\mathbf{u} = \underset{\mathcal{Z}}{\operatorname{argmax}} \min_{1 \leq l \leq N-1} d(\mathbf{F}_{DFT}, \mathbf{\Theta}^l \mathbf{F}_{DFT}).$$

Thus, there are N^{M_t} different possibilities for $\mathbf{\Theta}$ that must be checked. For small transmit antenna arrays and/or low feedback rates, it is possible to search over all possible values of $\mathbf{u}$ in $\mathcal{Z}$. In general, however, random search methods must be employed to design $\mathcal{F}$. These methods optimize the cost function by randomly testing values of $\mathbf{u}$ using a uniform distribution on $\mathcal{Z}$.

Codebooks designed from [92] also have the added benefit of easy memory storage. The codebook $\mathcal{F}$ can be stored at the transmitter/receiver using $\lceil M_t \log_2 N \rceil$ bits because only the numbers $u_1, u_2, \ldots, u_{M_t}$ need to be stored. The chosen codeword matrix $\mathbf{F}$ can then be easily calculated by computing $\mathbf{\Theta}^{l-1}\mathbf{F}_{DFT}$ for the chosen binary codeword index l. Note that the matrix $\mathbf{F}_{DFT}$ can be either stored or computed at each codebook update.

3.4 Performance Analysis

Because of the difficulty in deriving closed-form SER expressions, we will characterize the diversity of our limited feedback precoders. A signaling scheme is said to obtain diversity order d if [36]

$$d = -\lim_{\rho \to \infty} \frac{\log Pr(ERROR)}{\log \rho}.$$

We can bound the asymptotic performance of limited feedback precoding given a channel using the SER bound in (3.3). Thus in order to understand the diversity performance of limited feedback precoding, we will bound $\max_{\mathbf{F}' \in \mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_F$.

It follows from the Poincaré separation theorem [83], pp. 190 that

$$\max_{\mathbf{F}' \in \mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_F \leq \|\mathbf{H}\|_F.$$

Note that $\|\mathbf{H}\|_F$ is the post-processing channel gain of an M_t antenna OSTBC that is known to obtain a diversity order of $M_t M_r$ [14]. Therefore, the diversity order of our limited feedback precoders is less than or equal to $M_t M_r$. Thus all that is needed is a lower bound on diversity order.

Let $\mathbf{f}_{k,l}$ denote the l^{th} column of codebook matrix $\mathbf{F}_k$. The following lemma will prove useful in lower bounding $\max_{\mathbf{F}' \in \mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_F$.

Lemma 4 *Let $\tilde{\mathcal{F}} = \{\tilde{\mathbf{F}}_1,\ \tilde{\mathbf{F}}_2, \ldots,\ \tilde{\mathbf{F}}_N\}$ be a codebook with packing minimum distance $\tilde{\delta}$ and $N \geq M_t/M$. If there exists a vector $\mathbf{v} \in \mathbb{C}^{M_t}$ such that $\tilde{\mathbf{F}}_k^* \mathbf{v} = \mathbf{0}$ for all $1 \leq k \leq N$, then*
1.) There exists (k', l') such that $\tilde{\mathbf{f}}_{k',l'} = \sum_{i=1}^{m} \alpha_i \tilde{\mathbf{f}}_{k_i,l_i}$ where k_i and l_i are indexing sequences with $(k_i, l_i) \neq (k', l')$, $0 < |\alpha_i| < 1$, and $1 \leq m < M_t$,
2.) A new codebook $\mathcal{F}$ with

$$\mathbf{f}_{k,l} = \begin{cases} \tilde{\mathbf{f}}_{k,l} & \text{if } (k,l) \neq (k',l'); \\ \mathbf{v} & \text{if } (k,l) = (k',l') \end{cases}$$

has minimum distance $\delta \geq \tilde{\delta}$.

Proof Part 1 follows from the fact that a basis for the column space of the matrix $\tilde{\mathbf{E}} = \left[\tilde{\mathbf{F}}_1\ \tilde{\mathbf{F}}_2\ \cdots\ \tilde{\mathbf{F}}_N\right]$ can be formed from $m < M_t$ columns of the matrix $\tilde{\mathbf{E}}$. Part 2 is a result of the fact that for $k \neq k'$, $\|\tilde{\mathbf{F}}_k^* \tilde{\mathbf{F}}_{k'}\|_F = \|\mathbf{F}_k^* \tilde{\mathbf{F}}_{k'}\|_F \geq \|\mathbf{F}_k^* \mathbf{F}_{k'}\|_F$ where

$$\mathbf{F}_{k'} = [\tilde{\mathbf{f}}_{k',1}\ \cdots\ \tilde{\mathbf{f}}_{k',l'-1}\ \mathbf{v}\ \tilde{\mathbf{f}}_{k',l'+1}\ \cdots\ \tilde{\mathbf{f}}_{k',M}].$$

∎

By applying Lemma 4 repeatedly, any $N \geq M_t/M$ matrix codebook $\tilde{\mathcal{F}}$ with minimum distance $\tilde{\delta}$ and columns $\{\tilde{\mathbf{f}}_{k,l}\}$ that do not span $\mathbb{C}^{M_t}$ can be trivially modified to a codebook $\mathcal{F}$ with columns $\{\mathbf{f}_{k,l}\}$ that span $\mathbb{C}^{M_t}$ and minimum distance $\delta \geq \tilde{\delta}$. Therefore we can state the following theorem.

Theorem 1 *If $N \geq M_t/M$ and the columns $\{\mathbf{f}_{k,l}\}$ span $\mathbb{C}^{M_t}$, then $\mathcal{F} = \{\mathbf{F}_1,\ \mathbf{F}_2, \ldots,\ \mathbf{F}_N\}$ provides full diversity order.*

Proof First note that the channel power is given by $\max_{\mathbf{F}'\in\mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_F^2$. Notice that

$$\begin{aligned}\max_{\mathbf{F}'\in\mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_F^2 &\geq \max_{k,l} \|\mathbf{H}\mathbf{f}'_{k,l}\|_2^2 \\ &\geq \frac{1}{M_r} \max_{k,l} \|\mathbf{H}\mathbf{f}'_{k,l}\|_1^2 \\ &= \frac{1}{M_r}\|\mathbf{H}\mathbf{E}\|_1^2\end{aligned}$$

where $\mathbf{E} = [\mathbf{F}_1\ \mathbf{F}_2\ \cdots\ \mathbf{F}_N]$. Because the columns $\{\mathbf{f}_{k,l}\}$ span $\mathbb{C}^{M_t}$, we can find $\mathbf{U}_{E,L} \in \mathcal{U}(M_t, M_t)$, $\mathbf{U}_{E,R} \in \mathcal{U}(NM, NM)$, and $\mathbf{\Phi}$ with diagonal entries $\phi_1 \geq \phi_2 \geq \ldots \geq \phi_{M_t} > 0$ such that $\mathbf{E} = \mathbf{U}_{E,L}\mathbf{\Phi}\mathbf{U}_{E,R}^*$. Then using the invariance of complex normal matrices to unitary transformation [85] and the bounds in [59],

$$\begin{aligned}\max_{\mathbf{F}'\in\mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_F^2 &\stackrel{d}{=} \max_{\mathbf{F}'\in\mathcal{F}} \|\mathbf{H}\mathbf{U}_{E,L}^*\mathbf{F}'\|_F^2 \\ &\geq \frac{1}{NMM_r}\|\mathbf{H}\mathbf{\Phi}\|_1^2 \\ &\geq \frac{\phi_{M_t}^2}{NMM_r}\|\mathbf{H}\|_1^2 \\ &\geq \frac{\phi_{M_t}^2}{NMM_t^2M_r}\|\mathbf{H}\|_F^2\end{aligned}$$

The lower bound $\frac{\phi_{M_t}^2}{NMM_t^2M_r}\|\mathbf{H}\|_F^2$ is the channel gain of an M_t antenna OSTBC with an array gain shift $\frac{\phi_{M_t}^2}{NMM_t^2M_r}$. Systems with effective channel gains of this form are known to obtain a diversity order of M_tM_r [14]. We can conclude now that the limited feedback precoders designed using the chordal distance obtain full diversity order. ∎

Because limited feedback precoding maintains full diversity performance, this allows the generalization of OSTBCs to transmission over any transmit antenna configuration with full diversity. This lemma generalizes the result in [45] that proves antenna selection achieves full diversity order.

3.5 Simulations

Precoded OSTBCs using chordal distance limited feedback precoders have been simulated in the following experiments. The precoder codebooks were designed using the criterion proposed in Section 3.3 and implemented with packings designed from [92]. For comparison, OSTBCs with antenna selection have also been simulated [45].

Note that by our power scaling definitions, ρ is the ratio of the total transmitted energy to noise power for each transmission. For example, if a two antenna Alamouti code were transmitted then

$$\mathbf{X} = \sqrt{\frac{\rho}{2}}\mathbf{F}\begin{bmatrix} s_1 & s_2^* \\ s_2 & -s_1^* \end{bmatrix}$$

with $E_{s_1}\left[|s_1|^2\right] = E_{s_2}\left[|s_2|^2\right] = 1$.

Experiment 1: The first experiment compares antenna selection and eight bit chordal distance precoding for a precoded Alamouti code on an 8×1 wireless system using 16-quadrature amplitude modulation (QAM). The SER performance for a 2×1 Alamouti code is shown for comparison. The results are shown in Fig. 3.2. Note that antenna selection adds approximately a 7.5dB gain over the unprecoded system at an error rate of 10^{-2}. Using chordal distance precoding provides approximately a 1.4dB gain over antenna selection.

Experiment 2: Fig. 3.3 considers a 4×2 system using a precoded Alamouti code with 4-QAM. Antenna selection provides a 3dB array gain compared

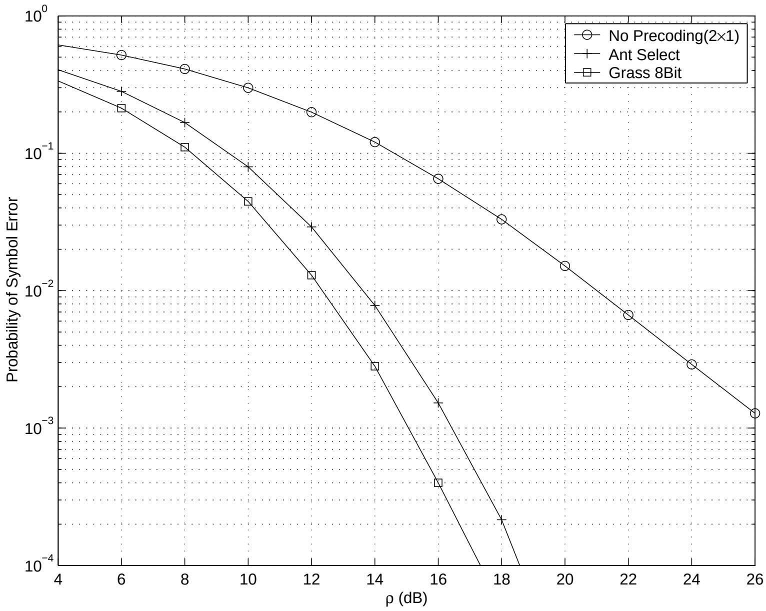

Figure 3.2: SER comparison of various OSTBC precoding schemes for a two substream 8×1 system using 16-QAM.

with an unprecoded OSTBC. Chordal distance precoding with a three bit codebook provides over a 0.3dB array gain and with a six bit codebook provides over a 0.7dB array gain over antenna selection. Interestingly, antenna selection requires $\lceil \log_2 \binom{4}{2} \rceil = 3$ bits of feedback. Thus, by simply lifting the restriction of $\mathbf{F}$ having columns of $\mathbf{I}_{M_t}$ adds a 0.3dB gain.

Experiment 3: The third experiment addresses the performance of a 5×4 wireless system using the OSTBC

$$\mathbf{C} = \begin{bmatrix} s_1 & 0 & s_2 & -s_3 \\ 0 & s_1 & s_3^* & s_2^* \\ -s_2^* & -s_3 & s_1^* & 0 \end{bmatrix} \tag{3.24}$$

obtained from [16] with 16-QAM. The results are shown in Fig. 3.4. Using antenna selection adds around a 1.5dB improvement over a 3×4 OSTBC. An

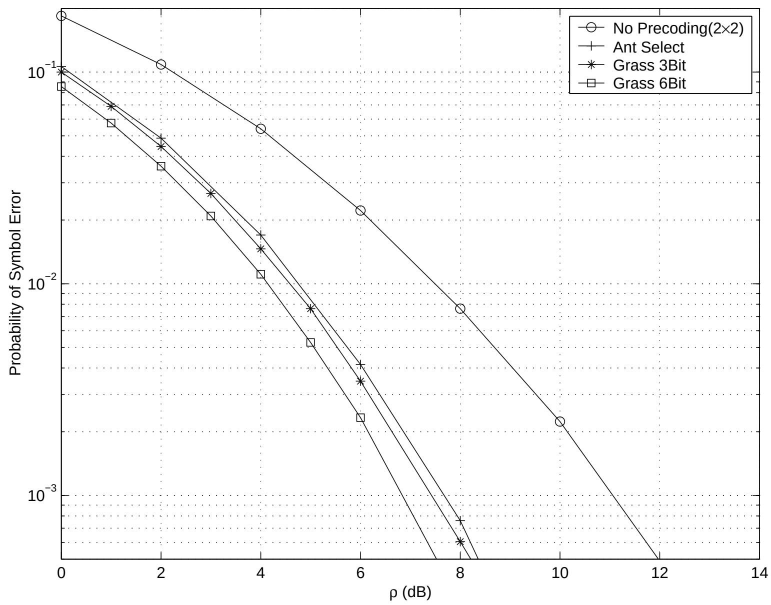

Figure 3.3: SER comparison of various OSTBC precoding schemes for a two substream 4×2 system using 4-QAM.

array gain of approximately 0.4dB is added by using six bit chordal distance precoding instead of antenna selection. Chordal distance precoding with an eight bit codebook outperforms antenna selection by more than 0.5dB.

Experiment 4: The fourth experiment, shown in Fig. 3.5, addresses the performance of a precoded OSTBC using the code in (3.24) over a 6×3 wireless with 16-QAM. Six and eight bit chordal distance precoding have array gains of approximately 2.1dB and 2.3dB, respectively, compared to unprecoded 3×3 OSTBCs. Note that six bits provides most of the array gain available with eight bits, thus a careful simulation analysis must be done in any practical system before choosing a given feedback rate.

Experiment 5: The effect of channel estimation error during precoder selection for an 8×1 Alamouti system is addressed in Fig. 3.6. Antenna

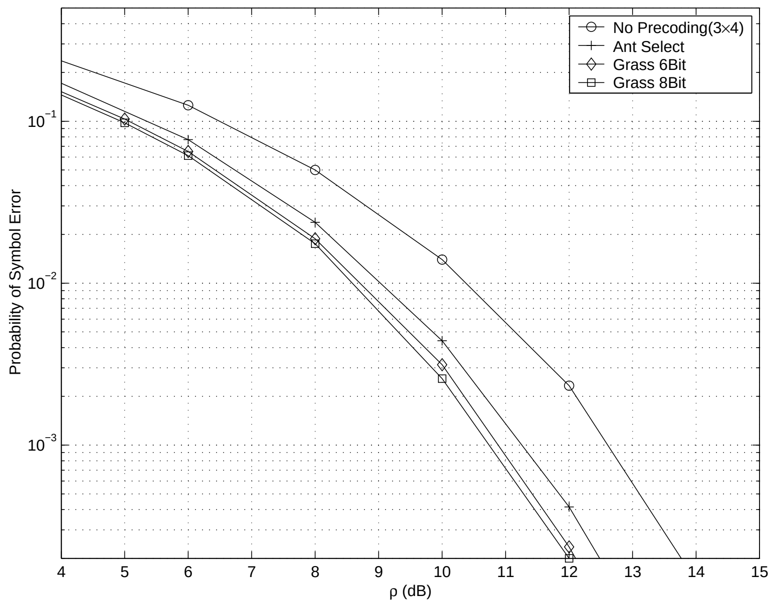

Figure 3.4: SER comparison of various OSTBC precoding schemes for a three substream 5×4 system using 16-QAM.

selection and eight bit chordal distance precoding are simulated for various channel correlations. We simulate estimation error by assuming that the receiver chooses $\mathbf{F} = f(\mathbf{H}_{est})$ using knowledge of a matrix $\mathbf{H}_{est}$ where

$$\mathbf{H}_{est} = \alpha\mathbf{H} + \sqrt{1-\alpha^2}\mathbf{H}_{error} \tag{3.25}$$

where the entries of $\mathbf{H}_{error}$ are independent and distributed according to $\mathcal{CN}(0,1)$. Notice that by our definition for α this means $E[vec(\mathbf{H}_{est})vec(\mathbf{H})^*] = \alpha\mathbf{I}_{M_rM_t}$. Note that $\mathbf{H}_{est}$ is a *biased estimate* of $\mathbf{H}$. This is to emphasize that the channel estimate is not only a *noisy estimate* but also a *dated estimate.* Thus the precoder was designed based on an outdated channel estimate.

It is assumed that the ML decoder has perfect knowledge of $\mathbf{HF}$ in order to isolate the effect of precoder mismatch. Interestingly, eight bit chordal distance precoding still outperforms perfect channel knowledge antenna subset

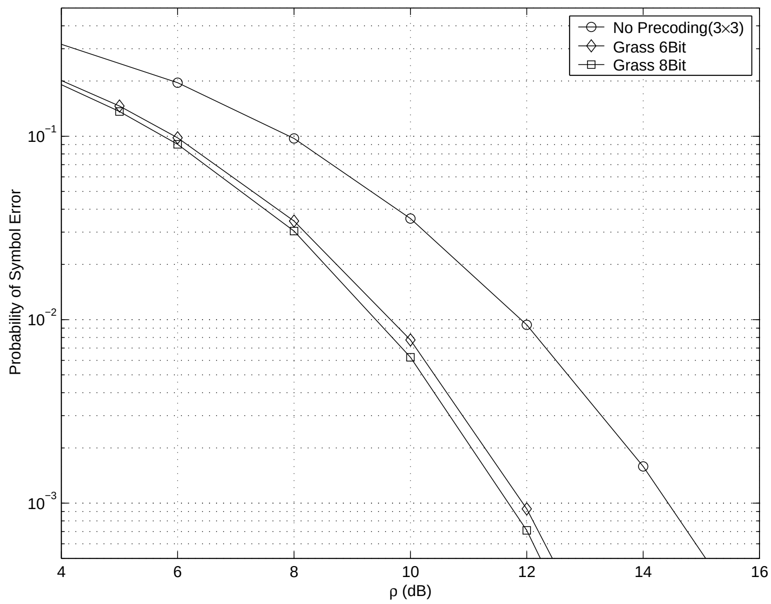

Figure 3.5: SER comparison of various OSTBC precoding schemes for a three substream 6×3 system using 16-QAM.

selection for $\alpha^2 = 0.9$. When $\alpha^2 = 0.75$, chordal distance precoding performs approximately the same as $\alpha^2 = 0.9$ antenna selection. This result demonstrates that chordal distance precoding allows a tradeoff between the quality of the channel estimation and the rate of the feedback path. At the expense of requiring more feedback, limited feedback chordal distance precoding can obtain a lower probability of error than antenna selection even in the presence of channel estimation error.

Experiment 6: The final experiment validates that the codebooks achieve the entropy lower bound on codeword bit length [93], pp. 86. In this experiment, an 8×8 wireless system transmitting a precoded Alamouti code was simulated in order to estimate the probability that each codeword precoder is selected. A five bit codebook was simulated. The results are shown in Fig.

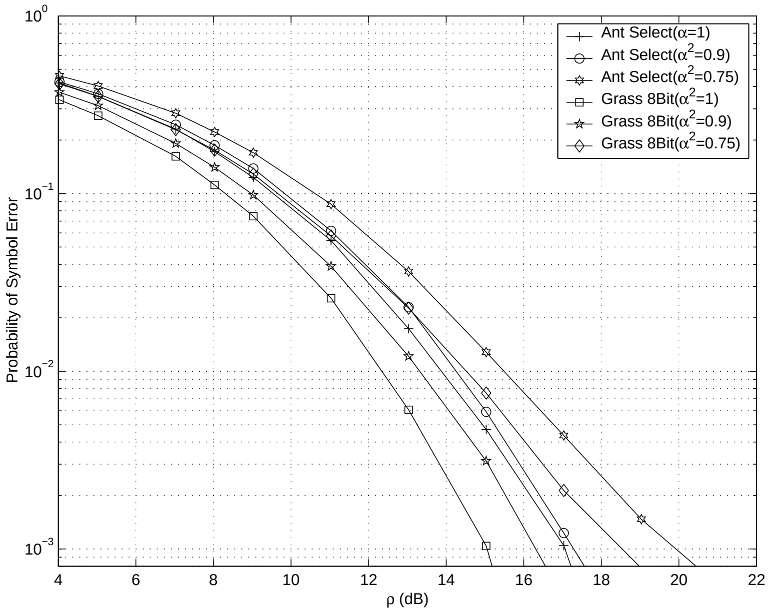

Figure 3.6: SER comparison of various OSTBC precoding schemes with channel estimation error for a two substream 8×1 system using 16-QAM.

3.7. Note that a uniform distribution across the codebook would mean that each matrix is selected with probability $1/2^5 = 0.03125$. All estimated selection probabilities are within 2.5×10^{-4} of the uniform selection probability. Thus the precoder matrix to be encoded, $\mathbf{F} = f(\mathbf{H})$, has an entropy of

$$\sum_{j=1}^{2^5} 2^{-5} \log_2(2^5) = 5 \text{ bits.}$$

An optimal data compression algorithm to encode the precoder matrix $\mathbf{F}$ can *at best* achieve an average codeword bit length of five bits. Because we transmit each precoder codeword with $\log_2 N = 5$ bits, our codebook has been experimentally shown to obtain the optimal codeword bit length lower bound [93].

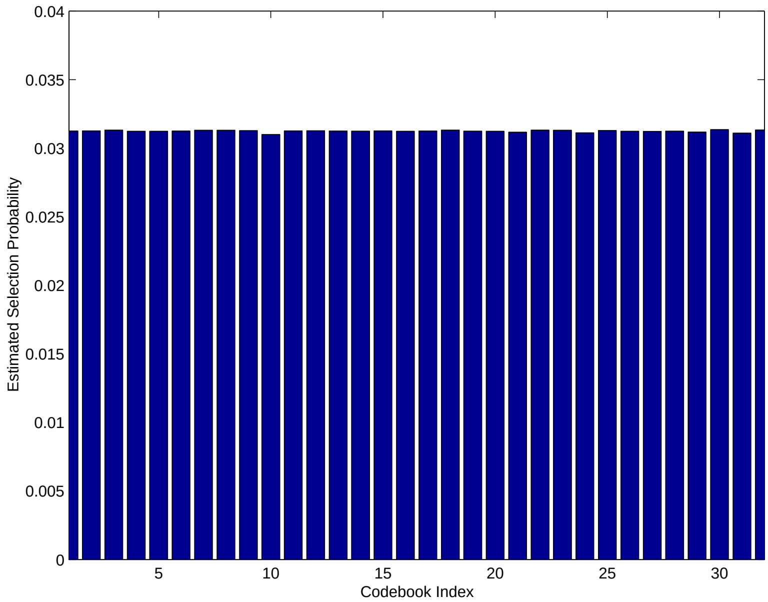

Figure 3.7: Probability of selection for each possible codebook precoding matrix for a precoded Alamouti code transmitted on an 8×8 wireless system.

Chapter 4

Limited Feedback Precoding for Spatial Multiplexing

In this chapter, a limited feedback framework for precoded spatial multiplexing is discussed. This chapter is organized as follows. Section 4.1 reviews the precoded spatial multiplexing system model. Criteria for choosing the optimal matrix from the codebook is presented in Section 4.2. Design criteria for creation of the precoder codebook are derived in Section 4.3. Section 4.4 illustrates the performance improvements over no precoding, unquantized precoding, and antenna subset selection using Monte Carlo simulations of the symbol error rate.

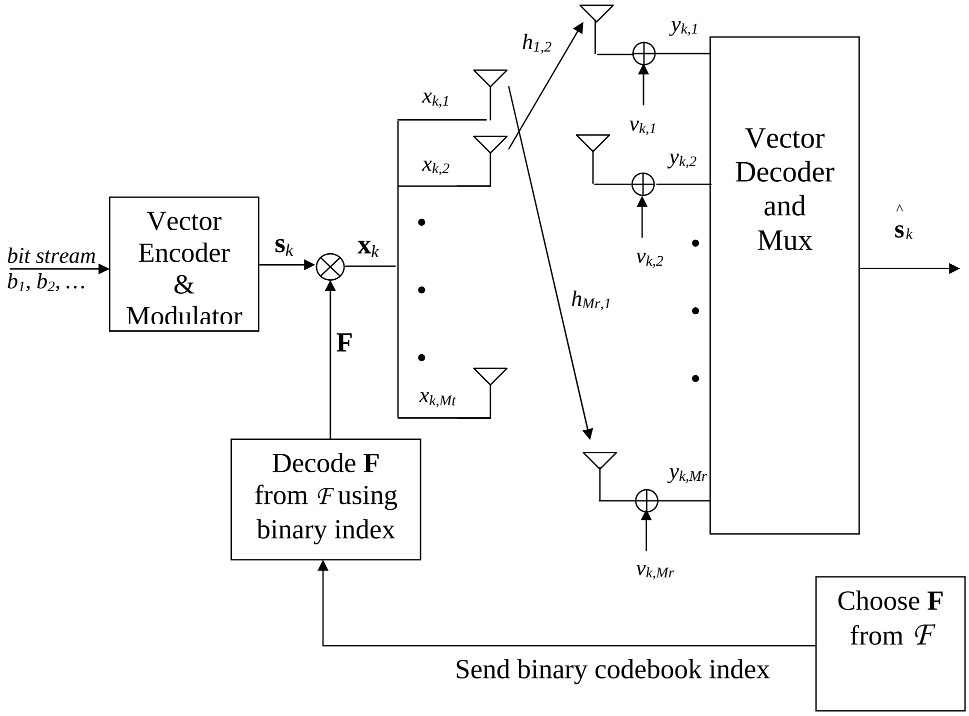

Figure 4.1: Block diagram of a limited feedback precoding MIMO system.

4.1 System Overview

The proposed system is illustrated in Fig. 4.1. A bit stream is sent into a vector encoder and modulator block where it is demultiplexed into M different bit streams. Each of the M bit streams is then modulated independently using the same constellation $\mathcal{W}$. This yields a symbol vector at time k of $\mathbf{s}_k = [s_{k,1}\ s_{k,2}\ \ldots\ s_{k,M}]^T$. For convenience, we will assume that $E_{\mathbf{s}_k}[\mathbf{s}_k\mathbf{s}_k^*] = \mathbf{I}_M$.

The symbol vector $\mathbf{s}_k$ is then multiplied by an $M_t \times M$ precoding matrix $\mathbf{F}$ (which is chosen as a function of the channel using criteria to be described) producing a length M_t vector $\mathbf{x}_k = \sqrt{\frac{\mathcal{E}_s}{M}}\mathbf{F}\mathbf{s}_k$ where $\mathcal{E}_s$ is the total transmit energy, M_t is the number of transmit antennas, and $M_t > M$. Assuming perfect timing, synchronization, sampling, and a memoryless linear matrix channel, this formulation allows the baseband, discrete-time equivalent received signal to be written as

$$\mathbf{y}_k = \sqrt{\frac{\mathcal{E}_s}{M}}\mathbf{H}\mathbf{F}\mathbf{s}_k + \mathbf{v}_k \tag{4.1}$$

where $\mathbf{H}$ is the channel matrix and $\mathbf{v}_k$ is the noise vector. We assume that the entries of $\mathbf{H}$ are independent and identically distributed (i.i.d.) according to $\mathcal{CN}(0,1)$ and the entries of $\mathbf{v}_k$ are independent and distributed according to $\mathcal{CN}(0,N_0)$. The received vector is then decoded by a vector decoder, assuming perfect knowledge of $\mathbf{HF}$, that produces a hard decoded symbol vector $\widehat{\mathbf{s}}_k$.

The receiver chooses a precoding matrix $\mathbf{F}$ from a finite set of possible precoding matrices $\mathcal{F} = \{\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_N\}$ and conveys the index of the chosen precoding matrix back to the transmitter over a limited capacity, zero-delay feedback link. We assume that each $\mathbf{F} \in \mathcal{F}$ has unit column vectors that are orthogonal. This assumption is not especially restrictive since it follows from the form of the optimal, full channel knowledge precoders derived in [28] assuming a maximum singular value constraint on $\mathbf{F}$. Thus the proposed codebook will satisfy $\mathcal{F} \subseteq \mathcal{U}(M_t, M)$. To simplify implementation, we will typically assume that B bits of feedback are available; thus the codebook consists of $N = 2^B$ matrices in $\mathcal{U}(M_t, M)$. The fact that the set $\mathcal{F}$ is discrete allows the receiver to solve for $\mathbf{F}$ by computing the selection metric of interest for each of the $N = 2^B$ codebook entries. The limitation of the codebook to 2^B matrices allows the system designer to constrain the precoding overhead and to take full advantage of the limited feedback channel.

To illustrate the concept, consider a codebook that corresponds to antenna subset selection [71]. Such a codebook would consist of the $\binom{M_t}{M}$ matrices consisting of M columns of $\mathbf{I}_{M_t}$. Notice that each set of M columns of $\mathbf{I}_{M_t}$ is an element of $\mathcal{U}(M_t, M)$. Naturally, antenna selection precoding can be directly implemented in a limited feedback system because a total feedback of only $\lceil \log_2 \binom{M_t}{M} \rceil$ bits is required. Unfortunately, the performance is highly limited because (i) the columns of $\mathbf{F}$ are restricted to being M columns of $\mathbf{I}_{M_t}$ and (ii) the size of the codebook is limited by M and M_t. It is of interest to remove any restrictions about the nature of the elements of the codebook as well as the number of elements in an effort to come closer to the gains of approximately-optimal precoding.

Problem Statement. The objective of this work is to solve the two key problems that are needed to effectively design and implement a limited feedback precoding system as proposed in Fig. 4.1. The first is to develop algorithms for selecting the optimal $\mathbf{F}$ from $\mathcal{F}$ as a function of the error probability or mutual information. This is the codeword or precoder selection problem. The second is to determine how to select a good codebook $\mathcal{F}$, based on a distortion measure that accounts for the fact that the channel is uncorrelated Rayleigh fading. This is the codebook design problem.

4.2 Precoding Criteria

In this section we discuss the criteria used for choosing the optimal precoding matrix from a given codebook. We outline criteria based on minimizing the error rate for the maximum likelihood or linear decoder and on maximizing the mutual information. When illustrative, we derive the optimal matrix over $\mathcal{U}(M_t, M)$.

4.2.1 Maximum Likelihood Receiver

The ML receiver solves the optimization problem

$$\widehat{\mathbf{s}}_k = \operatorname*{argmin}_{\mathbf{s}\in\mathcal{W}^M} \left\| \mathbf{y}_k - \sqrt{\frac{\mathcal{E}_s}{M}}\mathbf{HFs} \right\|_2^2 . \tag{4.2}$$

A closed-form expression for the probability of symbol vector error is difficult to derive. It has been shown that the probability of symbol vector error can be computed numerically [94], but this is cumbersome to implement in a real-time system. One approach is to observe that the probability of symbol vector error can be upper bounded for high signal-to-noise ratios (SNR) using the vector Union Bound [39]. This approach is motivated by the fact that the Union Bound provides an adequately tight prediction of the probability of error for large SNR. Since we assume $\mathcal{E}_s/N_0$ to be fixed, the Union Bound is solely a function of the receive minimum distance $d_{\min,R}$ of the multidimensional constellation $\mathcal{W}^M$ [79], which is given by

$$d_{\min,R} = \min_{\mathbf{s}_1,\mathbf{s}_2\in\mathcal{W}^M:\mathbf{s}_1\neq\mathbf{s}_2} \sqrt{\frac{\mathcal{E}_s}{M}} \|\mathbf{HF}\left(\mathbf{s}_1-\mathbf{s}_2\right)\|_2. \tag{4.3}$$

The computation of $d_{\min,R}$ requires a search over $\binom{card(\mathcal{W}^M)}{2}$ vectors.

Using (4.3), the minimum Euclidean distance criterion is to pick $\mathbf{F}$ from the codebook $\mathcal{F}$ for a given $\mathbf{H}$ assuming that $\mathcal{W}$ and $\mathcal{E}_s/N_0$ are fixed according to the following criterion.

SC-ML: ML Selection Criterion: Pick $\mathbf{F}$ such that

$$\mathbf{F} = \operatorname*{argmax}_{\mathbf{F}_i\in\mathcal{F}} d_{\min,R}. \tag{4.4}$$

Deriving a closed-form solution to *SC-ML* is difficult since the minimum distance depends on the constellation as well as the channel realization.

4.2.2 Linear Receiver

Linear receivers apply an $M \times M_r$ matrix $\mathbf{G}$, chosen according to some criterion, to produce $\widehat{\mathbf{s}}_k = \mathbf{Q}(\mathbf{G}\mathbf{y}_k)$ where $\mathbf{Q}(\cdot)$ is a function that performs single-dimension ML decoding for each entry of a vector. Criteria will be presented for two different forms of $\mathbf{G}$: zero-forcing and minimum mean square error. For a zero-forcing (ZF) linear decoder, $\mathbf{G} = (\mathbf{HF})^{\dagger}$. When a minimum mean square error (MMSE) linear decoder is used $\mathbf{G} = [\mathbf{F}^*\mathbf{H}^*\mathbf{HF} + (N_0 M/\mathcal{E}_s)\mathbf{I}_M]^{-1}\mathbf{F}^*\mathbf{H}^*$.

Minimum Singular Value

We will characterize the average probability of symbol vector error performance using the substream with the minimum SNR following the results given in [71]. It was shown in [71] that the SNR of the k^{th} substream is given by

$$SNR_k^{(ZF)} = \frac{\mathcal{E}_s}{MN_0[\mathbf{F}^*\mathbf{H}^*\mathbf{HF}]_{k,k}^{-1}} \tag{4.5}$$

for the ZF decoder and

$$SNR_k^{(MMSE)} = \frac{\mathcal{E}_s}{MN_0[\mathbf{F}^*\mathbf{H}^*\mathbf{HF} + MN_0/\mathcal{E}_s\mathbf{I}_M]_{k,k}^{-1}} - 1 \tag{4.6}$$

for the MMSE decoder, where $\mathbf{A}_{k,k}^{-1}$ is entry (k,k) of $\mathbf{A}^{-1}$. In [71] it is shown that in order to minimize a bound on the average probability of a symbol vector error, the minimum substream SNR must be maximized.

Using a selection criterion based on the minimum SNR requires the computation of the SNR of each of the M substreams and the estimation of the $\mathcal{E}_s/N_0$. The computational complexity combined with the possibility of estimation error makes the minimum cumbersome to implement. For this reason, [71] shows that the minimum SNR for ZF can be bounded using

$$SNR_{\min}^{(ZF)} = \min_{1\le k\le M} SNR_k^{(ZF)} \tag{4.7}$$

$$\ge \lambda_{\min}^2\{\mathbf{HF}\}\frac{\mathcal{E}_s}{MN_0} \tag{4.8}$$

where $\lambda_{\min}\{\mathbf{HF}\}$ is the minimum singular value of $\mathbf{HF}$.

We use (4.8) to obtain a requirement for choosing $\mathbf{F}$ from $\mathcal{F}$ for a given $\mathbf{H}$. We have assumed that $\mathcal{F}$ and $\mathcal{E}_s/N_0$ are fixed.

SC-MSV: Minimum Singular Value Criterion: Pick $\mathbf{F}$ such that

$$\mathbf{F} = \operatorname*{argmax}_{\mathbf{F}_i\in\mathcal{F}} \lambda_{\min}\{\mathbf{HF}_i\}. \tag{4.9}$$

This criterion provides a close approximation to maximizing the minimum SNR for dense constellations. The reason for this is that as $card(\mathcal{W})$ grows large, the probability of an error vector lying collinear to the minimum singular value direction goes to one.

Optimal Precoder over $\mathcal{U}(M_t, M)$

For comparison purposes, we also derive $\mathbf{F}_{opt} \in \mathcal{U}(M_t, M)$ that maximizes $\lambda_{\min}\{\mathbf{HF}_{opt}\}$. Note that when the feasible set* is $\mathcal{U}(M_t, M)$, $\mathbf{F}_{opt}$ is not unique. For example, if $\mathbf{F}_{opt}$ maximizes $\lambda_{\min}\{\mathbf{HF}_{opt}\}$ then so does $\mathbf{F}_{opt}\mathbf{U}$ for any $\mathcal{U}(M, M)$.

Let the singular value decomposition of $\mathbf{H}$ be given by

$$\mathbf{H} = \mathbf{V}_L\mathbf{\Sigma}\mathbf{V}_R^* \tag{4.10}$$

where $\mathbf{V}_L \in \mathcal{U}(M_r, M_r)$, $\mathbf{V}_R \in \mathcal{U}(M_t, M_t)$, and $\mathbf{\Sigma}$ is an $M_r \times M_t$ diagonal matrix with $\lambda_k\{\mathbf{H}\}$ denoting the k^{th} largest singular value of $\mathbf{H}$, at entry (k, k).

*The feasible set of an optimization is the domain that the cost function is optimized over.

Lemma 5 *An optimal precoder over $\mathcal{U}(M_t, M)$ for SC-MSV is $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R$ where $\overline{\mathbf{V}}_R$ is a matrix constructed from the first M columns of $\mathbf{V}_R$.*

Proof Let $\tilde{\mathbf{F}} = [\mathbf{F}_{opt}\ \tilde{\mathbf{f}}_1\ \ldots \tilde{\mathbf{f}}_{M_t-M}]$ where $\tilde{\mathbf{F}}^*\tilde{\mathbf{F}} = \mathbf{I}_{M_t}$. It is clear that the matrix $\tilde{\mathbf{F}}^*\mathbf{H}^*\mathbf{H}\tilde{\mathbf{F}}$ is a Hermitian matrix and $\mathbf{F}_{opt}^*\mathbf{H}^*\mathbf{H}\mathbf{F}_{opt}$ is obtained from $\tilde{\mathbf{F}}^*\mathbf{H}^*\mathbf{H}\tilde{\mathbf{F}}$ by simply taking the principle submatrix corresponding to the first M rows. By the *Inclusion Principle* [83],

$$\lambda_{\min}\{\mathbf{H}\} = \lambda_{M_t}\{\mathbf{H}\} \leq \lambda_{\min}\{\mathbf{H}\mathbf{F}_{opt}\} = \lambda_M\{\mathbf{H}\mathbf{F}_{opt}\} \leq \lambda_M\{\mathbf{H}\}. \tag{4.11}$$

This upper-bound can thus be achieved if $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R$. ∎

Minimum Mean Squared Error

Previous work [28] has considered improving the overall system performance by minimizing some function of the mean squared error (MSE) matrix

$$\overline{\mathrm{MSE}}(\mathbf{F}, \mathbf{G}) = E\left[\left(\mathbf{G}\mathbf{y}_k - \sqrt{\frac{\mathcal{E}_s}{M}}\mathbf{s}_k\right)\left(\mathbf{G}\mathbf{y}_k - \sqrt{\frac{\mathcal{E}_s}{M}}\mathbf{s}_k\right)^*\right]$$

where the expectation is taken over $\mathbf{s}_k$ and $\mathbf{v}_k$. When MMSE linear decoding is used, we express the MSE as

$$\overline{\mathrm{MSE}}(\mathbf{F}) = \frac{\mathcal{E}_s}{M}\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0}\mathbf{F}^*\mathbf{H}^*\mathbf{H}\mathbf{F}\right)^{-1}. \tag{4.12}$$

Using (4.12) we derive a selection criterion for choosing $\mathbf{F}$ from $\mathcal{F}$.
SC-MSE: Mean Squared Error Criterion: Pick $\mathbf{F}$ such that

$$\mathbf{F} = \underset{\mathbf{F}_i \in \mathcal{F}}{\operatorname{argmin}}\, m\left(\overline{\mathrm{MSE}}(\mathbf{F}_i)\right) \tag{4.13}$$

where $m(\cdot)$ is either $tr(\cdot)$ or $det(\cdot)$.

Note that minimizing the MSE does not specifically mean a reduction in the probability of error. In general, if the goal is to minimize the probability of error either *SC-MSV* should be chosen.

Optimal Precoding over $\mathcal{U}(M_t, M)$

Again, we present the optimal precoder over the unquantized set $\mathcal{U}(M_t, M)$ for subsequent comparisons. In [28], various constraints on $\mathbf{F}_{opt}$ were considered along with various mean squared error cost functions based on $\overline{\text{MSE}}(\mathbf{F}_{opt})$. Since we restrict our search to $\mathbf{F}_{opt} \in \mathcal{U}(M_t, M)$ we will consider the constraint in [28] where the maximum eigenvalue of $\mathbf{F}_{opt}\mathbf{F}_{opt}^*$ is unity. Note that all matrices in $\mathcal{U}(M_t, M)$ satisfy this constraint, but belonging to $\mathcal{U}(M_t, M)$ is not a necessary condition for this constraint.

It was shown in [28] that $\mathbf{F}_{opt}$ that minimizes $tr(\overline{\text{MSE}}(\mathbf{F}_{opt}))$ or $det(\overline{\text{MSE}}(\mathbf{F}_{opt}))$ under this maximum eigenvalue constraint is $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R$. Therefore we can state the following lemma as a consequence.

Lemma 6 *(Scaglione et al [28]) A matrix* $\mathbf{F}_{opt} \in \mathcal{U}(M_t, M)$ *that minimizes either of the two cost functions* $tr(\overline{MSE}(\mathbf{F}_{opt}))$ *and* $det(\overline{MSE}(\mathbf{F}_{opt}))$ *is* $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R$.

Once again $\mathbf{F}_{opt}$ is not unique because $tr(\overline{\text{MSE}}(\mathbf{F}_{opt})) = tr(\overline{\text{MSE}}(\mathbf{F}_{opt}\mathbf{U}))$ and $det(\overline{\text{MSE}}(\mathbf{F}_{opt})) = det(\overline{\text{MSE}}(\mathbf{F}_{opt}\mathbf{U}))$ for any $\mathbf{U} \in \mathcal{U}(M, M)$.

4.2.3 Capacity

In the context of antenna subset selection for spatial multiplexing systems, the mutual information (or capacity) has been used to formulate a precoder selection criterion [68], [80]. When the transmitter precodes with $\mathbf{F}$ before transmission, the equivalent channel is $\mathbf{HF}$. Thus the mutual information assuming an uncorrelated complex Gaussian source given $\mathbf{H}$ and a fixed $\mathbf{F}$ is

$$I(\mathbf{F}) = \log_2 det\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0}\mathbf{F}^*\mathbf{H}^*\mathbf{HF}\right). \tag{4.14}$$

Therefore we can state a capacity inspired selection criterion as follows.

SC-Capacity: Capacity Selection Criterion: Pick $\mathbf{F}$ such that

$$\mathbf{F} = \operatorname*{argmax}_{\mathbf{F}_i \in \mathcal{F}} I(\mathbf{F}_i). \tag{4.15}$$

Note that we call this selection criterion "*SC-Capacity*" for consistency with previous works [68], [71], [80].

Optimal Precoding over $\mathcal{U}(M_t, M)$

It is possible to find the optimal unquantized precoder $\mathbf{F}_{opt} \in \mathcal{U}(M_t, M)$ for the *SC-Capacity* criterion.

Lemma 7 *A precoder matrix $\mathbf{F}_{opt} \in \mathcal{U}(M_t, M)$ that maximizes $I(\mathbf{F}_{opt})$ is given by $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R$.*

Proof Note that maximizing $\log_2 det\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0}\mathbf{F}^*_{opt}\mathbf{H}^*\mathbf{HF}_{opt}\right)$ is equivalent to maximizing $det\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0}\mathbf{F}^*_{opt}\mathbf{H}^*\mathbf{HF}_{opt}\right)$ and thus minimizing $det\left(\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0}\mathbf{F}^*_{opt}\mathbf{H}^*\mathbf{HF}_{opt}\right)^{-1}\right)$. The latter expression differs from $det\left(\overline{\mathrm{MSE}}(\mathbf{F})\right)$ by a constant scale factor. It therefore follows from Lemma 6 that $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R$ maximizes $I(\mathbf{F}_{opt})$. ∎

Because of the relationship to *SC-MSE*, it is easily seen that $I(\mathbf{F}_{opt}) = I(\mathbf{F}_{opt}\mathbf{U})$ for any $\mathbf{U} \in \mathcal{U}(M, M)$.

4.3 Limited Feedback Precoding: Motivation and Codebook Design

In the previous section we derived criteria for selecting the optimal precoding matrix. It is important that codebook $\mathcal{F}$ is designed specifically for the chosen criterion. To understand the codebook design problem, we first perform a probabilistic characterization of the optimal precoding matrix. We then use this characterization to derive codebooks that maximize average bounds on each of the performance criteria.

4.3.1 Probabilistic Characterization of Optimal Precoding Matrix

Let the eigenvalue decomposition of $\mathbf{H}^*\mathbf{H}$ be given by

$$\mathbf{H}^*\mathbf{H} = \mathbf{V}_R \mathbf{\Sigma}^2 \mathbf{V}_R^* \tag{4.16}$$

with $\mathbf{V}_R$ and $\mathbf{\Sigma}$ defined as in (4.10). In [85], it is shown that for a MIMO Rayleigh fading channel $\mathbf{V}_R$, the right singular vector matrix, is isotropically distributed on $\mathcal{U}(M_t, M_t)$, the group of unitary matrices. An *isotropically distributed* $M_t \times M$ matrix $\mathbf{V}$ is a matrix where $\mathbf{\Theta}^*\mathbf{V} \overset{d}{=} \mathbf{V}$ for all $\mathbf{\Theta} \in \mathcal{U}(M_t, M_t)$ with $\overset{d}{=}$ denoting equivalence in distribution [86]. As stated in Section 4.2, the optimal precoder for *SC-MSV, SC-MSE,* and *SC-Capacity* is constructed by simply taking the first M columns of $\mathbf{V}_R$. Using the isotropic distribution of $\mathbf{V}_R$, it is possible to derive the distribution of $\mathbf{F}_{opt}$.

Lemma 8 *For a memoryless, i.i.d. Rayleigh fading channel* $\mathbf{H}$, $\mathbf{F}_{opt} = \overline{\mathbf{V}}_R$ *is isotropically distributed on* $\mathcal{U}(M_t, M)$.

Proof First note that

$$\mathbf{F}_{opt} = \overline{\mathbf{V}}_R = \mathbf{V}_R \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{(M_t-M)\times M} \end{bmatrix} \tag{4.17}$$

where $\mathbf{0}_{(M_t-M)\times M}$ is an $(M_t - M)\times M$ matrix of zeros. Since $\mathbf{V}_R$ is isotropically distributed,

$$\mathbf{\Theta}^*\mathbf{F}_{opt} = \mathbf{\Theta}^*\mathbf{V}_R \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{(M_t-M)\times M} \end{bmatrix} \stackrel{d}{=} \mathbf{V}_R \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{(M_t-M)\times M} \end{bmatrix} = \mathbf{F}_{opt}. \tag{4.18}$$

∎

Lemma 8 will allow the effect of the codebook on average distortion to be studied, with distortion to be defined later.

4.3.2 Grassmannian Subspace Packing

Before stating design criteria for each of the precoding matrix selection criteria, we present some relevant background about finite sets of matrices in $\mathcal{U}(M_t, M)$. The set $\mathcal{U}(M_t, M)$ defines the complex *Stiefel manifold* [90] of real dimension $2M_tM - M^2$. Each matrix in $\mathcal{U}(M_t, M)$ represents an M-dimensional subspace of $\mathbb{C}^{M_t}$. The set of all M-dimensional subspaces spanned by matrices in $\mathcal{U}(M_t, M)$ is the *complex Grassmann manifold,* denoted as $\mathcal{G}(M_t, M)$. Thus if $\mathbf{F}_1, \mathbf{F}_2 \in \mathcal{U}(M_t, M)$ then the column spaces of $\mathbf{F}_1$ and $\mathbf{F}_2$, $\mathcal{P}_{\mathbf{F}_1}$ and $\mathcal{P}_{\mathbf{F}_2}$ respectively, are contained in $\mathcal{G}(M_t, M)$. Note that the Grassmann manifold can be analyzed using a real or complex Stiefel manifold [90], however, we will only make use of complex subspaces. Our codebook $\mathcal{F}$, which consists of a finite number of matrices chosen from $\mathcal{U}(M_t, M)$, thus represents a set, or packing, of subspaces in the Grassmann manifold. Designing sets of N matrices that maximize the minimum subspace distance (where distance can be chosen in a number of different ways [84]) is known as *Grassmannian subspace packing.* We will use the interpretation of the precoding codebook $\mathcal{F}$ as a packing of subspaces to simplify notation and analysis.

A normalized invariant measure μ is induced on $\mathcal{G}(M_t, M)$ by the Haar measure in $\mathcal{U}(M_t, M)$. This measure allows the computation of volumes within $\mathcal{G}(M_t, M)$. Subspaces within the Grassmann manifold can be related by their distance from each other [84], [88], [89]. A number of different distances can be defined [84], [95], but we will only make use of three. The *chordal distance* between the two subspaces $\mathcal{P}_{\mathbf{F}_1}$ and $\mathcal{P}_{\mathbf{F}_2}$ is

$$d_{chord}(\mathbf{F}_1, \mathbf{F}_2) = \frac{1}{\sqrt{2}} \|\mathbf{F}_1\mathbf{F}_1^* - \mathbf{F}_2\mathbf{F}_2^*\|_F = \sqrt{M - \sum_{i=1}^{M} \lambda_i^2\{\mathbf{F}_1^*\mathbf{F}_2\}}. \tag{4.19}$$

The *projection two-norm distance* between two subspaces $\mathcal{P}_{\mathbf{F}_1}$ and $\mathcal{P}_{\mathbf{F}_2}$ is

$$d_{proj}(\mathbf{F}_1, \mathbf{F}_2) = \|\mathbf{F}_1\mathbf{F}_1^* - \mathbf{F}_2\mathbf{F}_2^*\|_2 = \sqrt{1 - \lambda_{\min}^2\{\mathbf{F}_1^*\mathbf{F}_2\}}. \tag{4.20}$$

The *Fubini-Study distance* between two subspaces $\mathcal{P}_{\mathbf{F}_1}$ and $\mathcal{P}_{\mathbf{F}_2}$ is

$$d_{FS}(\mathbf{F}_1, \mathbf{F}_2) = \arccos \left| det\left(\mathbf{F}_1^*\mathbf{F}_2\right)\right|. \tag{4.21}$$

Each of these distances correspond to different ideas of distance between subspaces. The chordal distance generalizes the distance between points on the unit sphere through an isometric embedding from $\mathcal{G}(M_t, M)$ to the unit sphere [88]. Maximizing this distance corresponds to minimizing the sum of the eigenvalues of $\mathbf{F}_2^*\mathbf{F}_1\mathbf{F}_1^*\mathbf{F}_2$ or similarly $\|\mathbf{F}_1^*\mathbf{F}_2\|_F$. The projection two-norm distance is maximized by minimizing the smallest singular value of $\mathbf{F}_1^*\mathbf{F}_2$, while the Fubini-Study distance is maximized by minimizing the product of the singular values of $\mathbf{F}_1^*\mathbf{F}_2$. Note that

$$\|\mathbf{F}_1^*\mathbf{F}_2\|_F^2 \geq M\lambda_{\min}^2\{\mathbf{F}_1^*\mathbf{F}_2\} \geq M\left|det\left(\mathbf{F}_1^*\mathbf{F}_2\right)\right|^2, \tag{4.22}$$

thus

$$d_{chord}(\mathbf{F}_1,\mathbf{F}_2) \leq \sqrt{M}d_{proj}(\mathbf{F}_1,\mathbf{F}_2) \leq \sqrt{M}\sin\left(d_{FS}(\mathbf{F}_1,\mathbf{F}_2)\right). \tag{4.23}$$

Let $\mathcal{S} = \{\mathcal{P}_{\mathbf{F}_1}, \mathcal{P}_{\mathbf{F}_2}, \ldots, \mathcal{P}_{\mathbf{F}_N}\}$ be the packing of column spaces of the codebook matrices where $\mathcal{P}_{\mathbf{F}_i}$ is the column space of $\mathbf{F}_i$. Similarly to binary error correcting codes [84], a packing can be characterized by its minimum distance

$$\delta = \min_{1\leq i<j\leq N} d(\mathbf{F}_i,\mathbf{F}_j) \tag{4.24}$$

where $d(\cdot,\cdot)$ is a distance function on $\mathcal{G}(M_t,M)$.

Consider the open ball in $\mathcal{G}(M_t,M)$ of radius $\gamma/2$ defined as

$$\mathcal{B}_{\mathbf{F}_i}(\gamma/2) = \{\mathcal{P}_{\mathbf{U}} \in \mathcal{G}(M_t,M) \;\mid\; d(\mathbf{U},\mathbf{F}_i) < \gamma/2\}. \tag{4.25}$$

This metric ball can be defined with respect to any of the distance function on $\mathcal{G}(M_t,M)$. Notice that

$$\mathcal{B}_{\mathbf{F}_i}(\gamma/2)\bigcap\mathcal{B}_{\mathbf{F}_j}(\gamma/2) = \phi$$

if $i \neq j$ and $\gamma \leq \delta$ with ϕ denoting the empty set because if $\mathcal{P}_{\mathbf{U}} \in \mathcal{B}_{\mathbf{F}_i}(\gamma/2)$ then

$$d(\mathbf{U},\mathbf{F}_j) \geq d(\mathbf{F}_i,\mathbf{F}_j) - d(\mathbf{F}_i,\mathbf{U}) \geq \gamma - \frac{1}{2}\gamma = \frac{1}{2}\gamma.$$

Note that if $d_{chord}(\mathbf{F}_1,\mathbf{F}_2) < \sqrt{1-\rho^2}$, with $0 \leq \rho \leq 1$, then we are guaranteed that $d_{proj}(\mathbf{F}_1,\mathbf{F}_2) < \sqrt{1-\rho^2}$ and $d_{FS}(\mathbf{F}_1,\mathbf{F}_2) < \arccos(\rho^M)$. This follows by restricting the largest $M-1$ singular values of $\mathbf{F}_1^*\mathbf{F}_2$ to be unity in order to find a lower bound on the minimum singular value. This observation yields

$$\mathcal{B}^{chord}_{\mathbf{F}_i}(\delta_{proj}/2) \subseteq \mathcal{B}^{proj}_{\mathbf{F}_i}(\delta_{proj}/2) \tag{4.26}$$

and

$$\mathcal{B}^{chord}_{\mathbf{F}_i}\left(\sqrt{1-\cos^{2/M}(\delta_{FS}/2)}\right) \subseteq \mathcal{B}^{FS}_{\mathbf{F}_i}(\delta_{FS}/2). \tag{4.27}$$

The *density* of a subspace packing with respect to a distance γ ($\gamma \leq \delta$) is

$$\Delta(\gamma) = \mu\left(\bigcup_{i=1}^{N} \mathcal{B}_{\mathbf{F}_i}(\gamma/2)\right) = \sum_{i=1}^{N} \mu\left(\mathcal{B}_{\mathbf{F}_i}(\gamma/2)\right) \tag{4.28}$$

where $\mathcal{B}_{\mathbf{F}_i}(\gamma/2)$ can be defined with respect to any distance function on the Grassmann manifold. The density of a packing is a measure of how well the codebook matrices "cover" $\mathcal{G}(M_t, M)$. The density allows the probability of the isotropically distributed $\overline{\mathbf{V}}_R$ falling in one of the set $\mathcal{B}_{\mathbf{F}_i}(\gamma/2)$, with $\gamma \leq \delta$, to be expressed as

$$Pr\left(\overline{\mathbf{V}}_R \in \bigcup_{i=1}^{N} \mathcal{B}_{\mathbf{F}_i}(\gamma/2)\right) = \Delta(\gamma). \tag{4.29}$$

Furthermore, (4.26) and (4.27) yield

$$\Delta_{chord}(\delta_{proj}) \leq \Delta_{proj}(\delta_{proj}) \tag{4.30}$$

and

$$\Delta_{chord}\left(2\sqrt{1-\cos^{2/M}(\delta_{FS}/2)}\right) \leq \Delta_{FS}(\delta_{FS}) \tag{4.31}$$

where the subscript indicates the distance used. Notice that the factor of 2 in (4.31) follows from the fact that the Fubini-Study minimum distance is halved inside of the cosine function. For large M_t it has been shown in [84] that

$$\Delta_{chord}(\delta) \approx N\left(\frac{\delta}{2\sqrt{M}}\right)^{2M_tM+o(M_t)}. \tag{4.32}$$

4.3.3 Codebook Design Criteria

We now derive the codebook design criteria for each specific selection criterion using the distribution of the optimal unquantized precoding matrix derived in Section 4.3.1 and the Grassmannian subspace packing results in Section 4.3.2.

SC-ML, SC-MSV, & SC-MSE (with Trace Cost Function)

Using (4.3), we can bound

$$\begin{aligned} d_{\min,R} &\geq \sqrt{\frac{\mathcal{E}_s}{M}} \left(\min_{\mathbf{s}_1,\mathbf{s}_2 \in \mathcal{W}^M : \mathbf{s}_1 \neq \mathbf{s}_2} \|\mathbf{s}_1 - \mathbf{s}_2\|_2 \right) \left(\min_{\mathbf{s}_1,\mathbf{s}_2 \in \mathcal{W}^M : \mathbf{s}_1 \neq \mathbf{s}_2} \|\mathbf{HFe}_{\mathbf{s}_1 - \mathbf{s}_2}\|_2 \right) \\ &\geq \sqrt{\frac{\mathcal{E}_s}{M}} \left(\min_{\mathbf{s}_1,\mathbf{s}_2 \in \mathcal{W}^M : \mathbf{s}_1 \neq \mathbf{s}_2} \|\mathbf{s}_1 - \mathbf{s}_2\|_2 \right) \lambda_{\min}\{\mathbf{HF}\} \end{aligned} \quad (4.33)$$

where $\mathbf{e}_{\mathbf{s}_1 - \mathbf{s}_2} = \frac{\mathbf{s}_1 - \mathbf{s}_2}{\|\mathbf{s}_1 - \mathbf{s}_2\|_2}$. Thus maximizing the lower bound on $d_{\min,R}$ is equivalent to maximizing $\lambda_{min}\{\mathbf{HF}\}$. Thus this bound shows *SC-ML* requires maximizing $\lambda_{\min}\{\mathbf{HF}\}$.

SC-MSE using the trace cost function chooses $\mathbf{F} \in \mathcal{F}$ that maximizes

$$tr\left(\frac{\mathcal{E}_s}{M} \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}^*\mathbf{H}^*\mathbf{HF} \right)^{-1} \right).$$

For high SNR, this can be approximated by $N_0 tr\left((\mathbf{F}^*\mathbf{H}^*\mathbf{HF})^{-1} \right)$, and we can bound

$$\min_{\mathbf{F}_i \in \mathcal{F}} N_0 tr\left((\mathbf{F}_i^*\mathbf{H}^*\mathbf{HF}_i)^{-1} \right) \leq \min_{\mathbf{F}_i \in \mathcal{F}} \frac{MN_0}{\lambda^2_{\min}\{\mathbf{HF}_i\}}. \quad (4.34)$$

The bound in (4.34) uses the fact that the maximum eigenvalue of $(\mathbf{F}^*\mathbf{H}^*\mathbf{HF})^{-1}$ is the inverse of the minimum eigenvalue of $\mathbf{F}^*\mathbf{H}^*\mathbf{HF}$. Minimizing the bound approximately minimizes the trace of the MSE matrix. Therefore, maximizing the $\lambda_{\min}\{\mathbf{HF}\}$ is an approximate method for minimizing the trace of the MSE matrix.

Based on (4.33) and (4.34), we can relate the selection of the optimal codeword in a given codebook for the *SC-ML* and *SC-MSE* (using the trace cost function) cases to selection of the optimal codeword based on *SC-MSV*.

To define a notion of an optimal codebook, we need a distortion measure with which to measure the average distortion. To design codebooks for the *SC-ML*, *SC-MSE*, and *SC-MSV* case we will use the error difference

$$\lambda_{\min}^2\{\mathbf{HF}_{opt}\} - \lambda_{\min}^2\{\mathbf{HF}_i\}$$

which is nonnegative for any choice of $\mathbf{F}_i \in \mathcal{F}$. Thus we will choose our codebook to maximize the average distortion

$$E_{\mathbf{H}}\left[\lambda_{\min}^2\{\mathbf{HF}_{opt}\} - \max_{\mathbf{F}_i\in\mathcal{F}} \lambda_{\min}^2\{\mathbf{HF}_i\}\right] = E_{\mathbf{H}}\left[\lambda_M^2\{\mathbf{H}\} - \max_{\mathbf{F}_i\in\mathcal{F}} \lambda_{\min}^2\{\mathbf{HF}_i\}\right]. \tag{4.35}$$

Evaluating the expectation exactly in (4.35) is difficult therefore we will minimize an upper bound on the average distortion.

Using the singular value representation used in Section 4.2 and the properties of Grassmannian subspace packing,

$$\begin{aligned}
E_{\mathbf{H}}\left[\max_{\mathbf{F}_i\in\mathcal{F}} \lambda_{\min}^2\{\mathbf{HF}_i\}\right] &= E_{\mathbf{H}}\left[\max_{\mathbf{F}_i\in\mathcal{F}} \lambda_{\min}^2\{\mathbf{\Sigma V}_R^*\mathbf{F}_i\}\right] \\
&\geq E_{\mathbf{H}}\left[\max_{\mathbf{F}_i\in\mathcal{F}} \lambda_{\min}^2\{\overline{\mathbf{\Sigma V}}_R^*\mathbf{F}_i\}\right] && (4.36)\\
&\geq E_{\mathbf{H}}\left[\lambda_M^2\{\mathbf{H}\}\right] E_{\mathbf{H}}\left[\max_{\mathbf{F}_i\in\mathcal{F}} \lambda_{\min}^2\{\overline{\mathbf{V}}_R^*\mathbf{F}_i\}\right] && (4.37)
\end{aligned}$$

where $\overline{\mathbf{\Sigma}}$ is the matrix constructed from the first M columns of $\mathbf{\Sigma}$. The result in (4.37) follows from the fact that singular values and singular vectors of complex normal matrices are independent [85], [90]. Due to the results in (4.35) and

(4.37) we obtain

$$\begin{aligned}
&E_{\mathbf{H}}\left[\lambda^2_{\min}\{\mathbf{H}\mathbf{F}_{opt}\} - \max_{\mathbf{F}_i \in \mathcal{F}} \lambda^2_{\min}\{\mathbf{H}\mathbf{F}_i\}\right] \\
&\quad \leq E_{\mathbf{H}}\left[\lambda^2_M\{\mathbf{H}\}\right] E_{\mathbf{H}}\left[\left(1 - \max_{\mathbf{F}_i \in \mathcal{F}} \lambda^2_{\min}\{\overline{\mathbf{V}}^*_R \mathbf{F}_i\}\right)\right] \\
&\quad \leq E_{\mathbf{H}}\left[\lambda^2_M\{\mathbf{H}\}\right] \left(\frac{\delta^2_{proj}}{4}\Delta_{proj}(\delta_{proj}) + (1 - \Delta_{proj}(\delta_{proj}))\right) && (4.38) \\
&\quad \lessapprox E_{\mathbf{H}}\left[\lambda^2_M\{\mathbf{H}\}\right] \left(1 + N\left(\frac{\delta_{proj}}{2\sqrt{M}}\right)^{2M_tM+o(M_t)} \left(\frac{\delta^2_{proj}}{4} - 1\right)\right). && (4.39)
\end{aligned}$$

The bound in (4.38) is a result of partitioning the possible outcomes into two cases: i.) the subspace of $\overline{\mathbf{V}}_R$ falls within a codeword metric ball of radius δ_{proj} and ii.) the subspace of $\overline{\mathbf{V}}_R$ does not fall within a codeword metric ball. The codewords fall within a metric ball with probability $\Delta_{proj}(\delta_{proj})$ and must have distance less than $\delta_{proj}/2$ from some codeword when they fall within a metric ball. Substituting the density bound in (4.30) and the approximation in (4.32) results in (4.39). Differentiation and making the assumption that $2M_tM + o(M_t) > 2/3$, gives the following design criterion.

Codebook Design Criterion: A codebook $\mathcal{F}$ for a system using *SC-ML*, *SC-MSV*, or *SC-MSE* (using the trace cost function) to select $\mathbf{F}$ from $\mathcal{F}$ should be designed by maximizing the minimum projection two-norm distance between any pair of codeword matrix column spaces.

SC-MSE (with Determinant Cost Function) & SC-Capacity

Selecting $\mathbf{F} \in \mathcal{F}$ using *SC-MSE* with the determinant cost function requires solving for $\mathbf{F}$ that minimizes

$$det\left(\frac{\mathcal{E}_s}{M}\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\mathbf{F}^*\mathbf{H}^*\mathbf{H}\mathbf{F}\right)^{-1}\right)=$$
$$\left(\frac{\mathcal{E}_s}{M}\right)^M\left(det\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\mathbf{F}^*\mathbf{H}^*\mathbf{H}\mathbf{F}\right)\right)^{-1}. \tag{4.40}$$

This is equivalent to solving for the $\mathbf{F}$ that maximizes $I(\mathbf{F}) = det\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\mathbf{F}^*\mathbf{H}^*\mathbf{H}\mathbf{F}\right)$, the same expression maximized in *SC-Capacity*. Using the SVD representation of $\mathbf{H}$,

$$I(\mathbf{F}) = det\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\mathbf{F}^*\mathbf{V}_R\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{V}_R^*\mathbf{F}\right) \tag{4.41}$$
$$= det\left(\mathbf{F}^*\mathbf{V}_R\left(\mathbf{I}_{M_t}+\frac{\mathcal{E}_s}{MN_0}\mathbf{\Sigma}^T\mathbf{\Sigma}\right)\mathbf{V}_R^*\mathbf{F}\right) \tag{4.42}$$
$$\geq det\left(\mathbf{F}^*\overline{\mathbf{V}}_R\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\right)\overline{\mathbf{V}}_R^*\mathbf{F}\right) \tag{4.43}$$
$$= \left|det\left(\overline{\mathbf{V}}_R^*\mathbf{F}\right)\right|^2 det\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\right). \tag{4.44}$$

To define a notion of an optimal codebook, we need a distortion measure to measure the average performance loss in this case. Since $\mathbf{F}_{opt}$ that maximizes the mutual information over $\mathcal{U}(M_t, M)$ gives $I(\mathbf{F}_{opt}) = det\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\right)$, we will use the error difference

$$I(\mathbf{F}_{opt}) - \left|det\left(\overline{\mathbf{V}}_R^*\mathbf{F}\right)\right|^2 det\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\right)$$

which is nonnegative for any choice of $\mathbf{F}_i \in \mathcal{F}$. Thus we will choose our codebook to minimize the average distortion

$$E_{\mathbf{H}}\left[det\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\right)-\max_{\mathbf{F}_i\in\mathcal{F}}\left|det\left(\overline{\mathbf{V}}_R^*\mathbf{F}_i\right)\right|^2 det\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\right)\right]$$
$$= E_{\mathbf{H}}\left[det\left(\mathbf{I}_M+\frac{\mathcal{E}_s}{MN_0}\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\right)\right]\left(1-\max_{\mathbf{F}_i\in\mathcal{F}}E_{\mathbf{H}}\left[\left|det\left(\overline{\mathbf{V}}_R^*\mathbf{F}\right)\right|^2\right]\right). \tag{4.45}$$

where (4.45) follows from the independence of $\mathbf{\Sigma}$ and $\mathbf{V}_R$ [85], [90].

The distortion cost function can be bounded as

$$E_{\mathbf{H}}\left[det\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0}\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\right)\right]\left(1 - \max_{\mathbf{F}_i\in\mathcal{F}} E_{\mathbf{H}}\left[\left|det\left(\overline{\mathbf{V}}_R^*\mathbf{F}\right)\right|^2\right]\right)$$

$$\leq E_{\mathbf{H}}\left[det\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0}\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\right)\right]\left(1 - \cos^2\left(\delta_{FS}/2\right)\Delta_{FS}(\delta_{FS})\right) \tag{4.46}$$

$$\lessapprox E_{\mathbf{H}}\left[det\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0}\overline{\mathbf{\Sigma}}^T\overline{\mathbf{\Sigma}}\right)\right]\cdot \tag{4.47}$$

$$\left(1 - \cos^2\left(\delta_{FS}/2\right)N\left(\sqrt{\frac{1-\cos^{2/M}(\delta_{FS}/2)}{M}}\right)^{2M_tM+o(M_t)}\right). \tag{4.48}$$

The result in (4.46) follows from the facts that the subspace of $\overline{\mathbf{V}}_R$ lies within a codeword metric ball with probability $\Delta_{FS}(\delta_{FS})$ and that all subspaces within the metric balls have distance less than $\delta_{FS}/2$. Using the density bound in (4.31) and the approximation in (4.32) yields (4.48). Differentiating this bound, and assuming that $M_t + o(M_t)/(2M) \geq (2^{1/M} - 1)$, tells us that we want to maximize δ_{FS} in order to minimize the distortion cost function.We can now state the following.

Codebook Design Criterion: A codebook $\mathcal{F}$ for a system using *SC-MSE* with the determinant cost function or *SC-Capacity* to select $\mathbf{F}$ from $\mathcal{F}$ should be designed by maximizing the minimum Fubini-Study distance between any pair of codeword matrix column spaces.

Discussion

In summary, thinking of the codebook $\mathcal{F}$ as a packing of M-dimensional subspaces rather than a set of $M_t \times M$ matrices allows us to bound the distortion for each of the selection criteria proposed in Section 4.2. The distortion bound for *SC-ML, SC-MSV,* and *SC-MSE* with the trace cost function is minimized by maximizing the minimum projection two-norm distance between any pair of codebook subspaces. The *SC-Capacity* and *SC-MSE* with the determinant cost function bound is minimized by maximizing the minimum Fubini-Study distance between any pair of codebook subspaces. Thus the codebook design is equivalent to subspace packing in the Grassmann manifold.

Observe that both design criteria make assumptions on the relation between M_t, M, and the $o(M_t)$ term. Numerical experiments have shown that for most $M_t > 2$ the assumptions are satisfied. When $M = 1$, it is also known that the $o(M_t)$ term is always -2 and (4.32) is an exact expression [59].

Finding good packings in the Grassmann manifold for arbitrary M_t, M, and N, and thus finding good codebooks, is difficult (see for example [87]–[89]). The problem is exasperated by the use of the projection two-norm and Fubini-Study distances instead of the more common chordal distance [84], [95]. For instance, in the simplest case of $M = 1$ where the Rankin lower bound on line packing correlation [89] can be employed, packings that achieve equality with the lower bound are often impossible to design. One simple method for designing good packings with arbitrary distance functions is to use the non-coherent constellation designs from [92]. We have found that the algorithms for constellation design in [92] yield codebooks with large minimum distances and can be easily modified to work with any distance function on the Grassmann manifold.

4.4 Simulations

Monte Carlo simulations were performed to illustrate the performance of Grassmannian precoders. The codebooks were designed using the criteria proposed in Section 4.3.3. For each of the precoding systems using M_t transmit antennas and M substreams we also plotted the $M \times M_r$ spatial multiplexing results with both ZF and ML decoding. In addition we simulated the unquantized MMSE precoding using the trace cost function and both the sum power and maximum singular value constraints [28] and maximum minimum singular value antenna selection [71].

Experiment 1: In this simulation we compared precoding schemes using 16 quadrature amplitude modulation (QAM) and two substreams. Fig. 4.2 shows the probability of symbol vector error curves for a 4×2 limited feedback precoding with six bits of feedback using *SC-Capacity, SC-MSV,* and *SC-MSE* with both cost functions. All four of the selection criteria perform approximately the same, 1.5dB better than antenna selection. The four selection criteria are approximately 1dB away from unquantized MMSE precoding using the sum power constraint.

Experiment 2: This simulation used binary phase shift keying (BPSK) modulation and two substreams on a 4×2 wireless system. The results are shown in Fig. 4.3. We simulated six bit limited feedback precoding using *SC-Capacity, SC-MSV,* and *SC-ML*. ZF decoding and precoding using *SC-Capacity* provided more than a 4dB performance gain at a probability of symbol vector error of 2×10^{-3} over unprecoded decoding using ML decoding. Precoding using *SC-MSV* provided a 0.5dB gain over *SC-Capacity.* Unquantized MMSE precoding with the sum power constraint performs approximately 1.5dB better than limited feedback precoding using *SC-MSV*. As expected, ML decoding combined with *SC-ML* provided a large performance gain and outperformed unquantized MMSE precoding at a probability of symbol vector error of 10^{-3} by around 2.5dB.

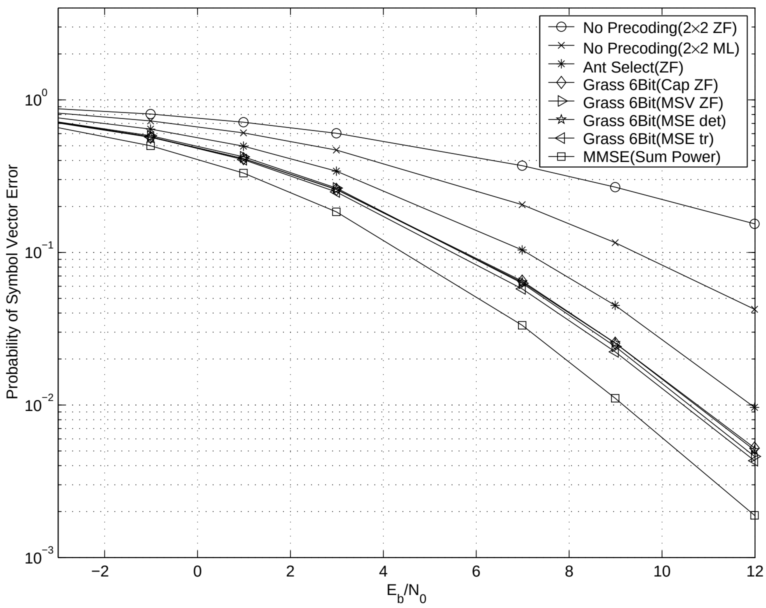

Figure 4.2: Probability of symbol vector error comparison of various precoding schemes for a 2 substream 4×2 system using 16-QAM.

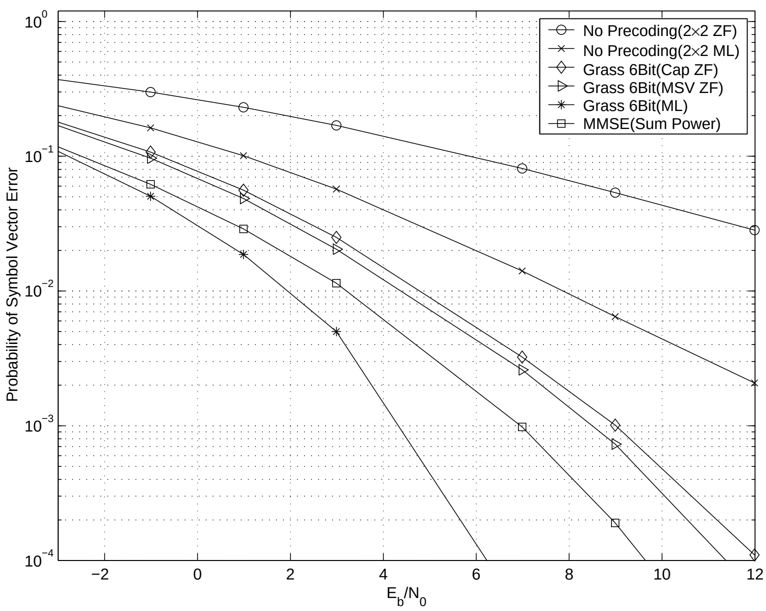

Figure 4.3: Probability of symbol vector error comparison of various precoding schemes for a 2 substream 4×2 system using BPSK.

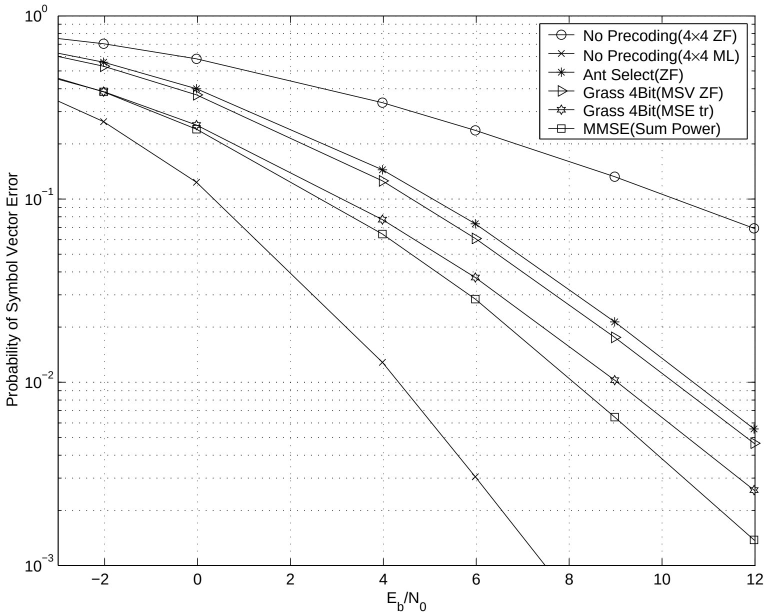

Figure 4.4: Probability of symbol vector error comparison of various precoding schemes for a 4 substream 5×4 system using 4-QAM.

Experiment 3: This simulation shows performance comparisons of precoding schemes using 4-QAM modulation and four substreams. The results are shown in Fig. 4.4. In this case, we simulated the probability of symbol vector error curves for a 5×4 limited feedback precoding with four bits of feedback using *SC-MSV* and *SC-MSE*. This simulation shows the performance differences between the different precoder selection criteria. Precoding using *SC-MSV* yields approximately a 0.4dB improvement over antenna selection. Using limited feedback precoding with *SC-MSE* and the trace cost function within 1.1dB of unquantized MMSE precoding with the sum power constraint.

Experiment 4: We simulated three substream precoding on a 6×3 wireless system in this experiment using 16-QAM with results shown in Fig. 4.5. We used limited feedback precoding with six bits of feedback. Limited

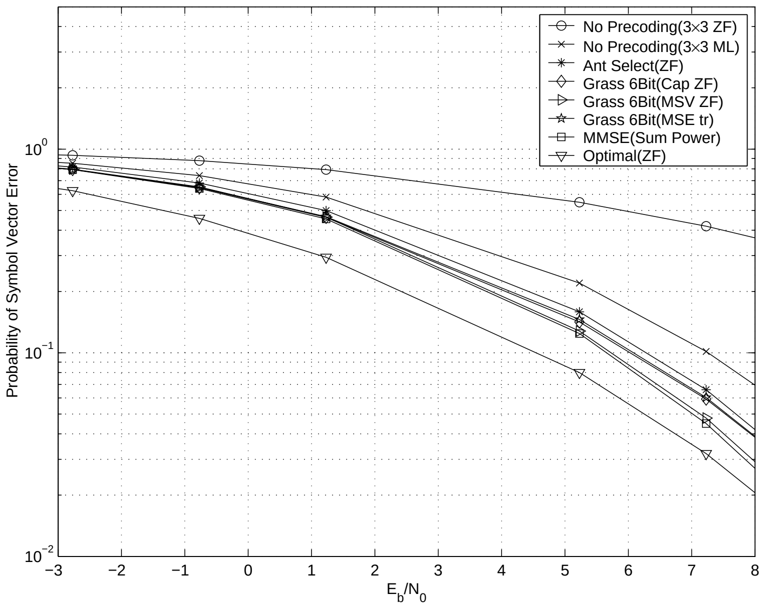

Figure 4.5: Probability of symbol vector error comparison of various precoding schemes for a 3 substream 6×3 system using 16-QAM.

feedback precoding with *SC-Capacity* performed approximately the 0.25dB better than antenna selection. Limited feedback precoding using *SC-MSV* and *SC-MSE* with the trace cost function both performed approximately the same. They both provide around a 0.25dB improvement over *SC-Capacity* and perform within 1dB of unquantized optimal precoding using ZF decoding. Note that all selection criteria outperformed unprecoded spatial multiplexing using an ML receiver. This shows though the power of precoding: near ML or better than ML performance with low complexity receivers at the expense of feedback.

Experiment 5: This experiment simulated three substream precoding on a 6×3 wireless system in this experiment using BPSK. This experiment's results are presented in Fig. 4.6. Once again, six bit limited feedback precod-

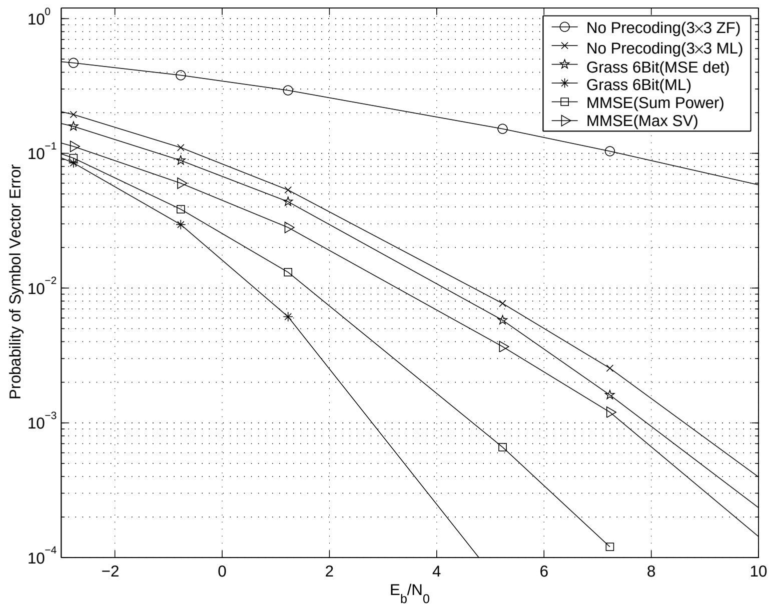

Figure 4.6: Probability of symbol vector error comparison of various precoding schemes for a 3 substream 6×3 system using BPSK.

ing was used. Precoding using *SC-MSE* with the determinant cost function provides approximately a 0.4 dB array gain compared to unprecoded spatial multiplexing using ML decoding. Unquantized MMSE precoding with the maximum singular value constraint performs only 0.8dB better than the six bit limited feedback precoder. Using quantized precoding with ZF decoding provides approximately a 1dB gain over quantized precoding using *SC-MSE* with the determinant cost function at a probability of symbol vector error of 10^{-3}. Quantized precoding using ML decoding and *SC-ML* outperforms unquantized MMSE precoding with the sum power constraint by more than 1.8dB.

Experiment 6: The purpose of this experiment is to demonstrate the problems associated with directly quantizing the matrix channel $\mathbf{H}$. The re-

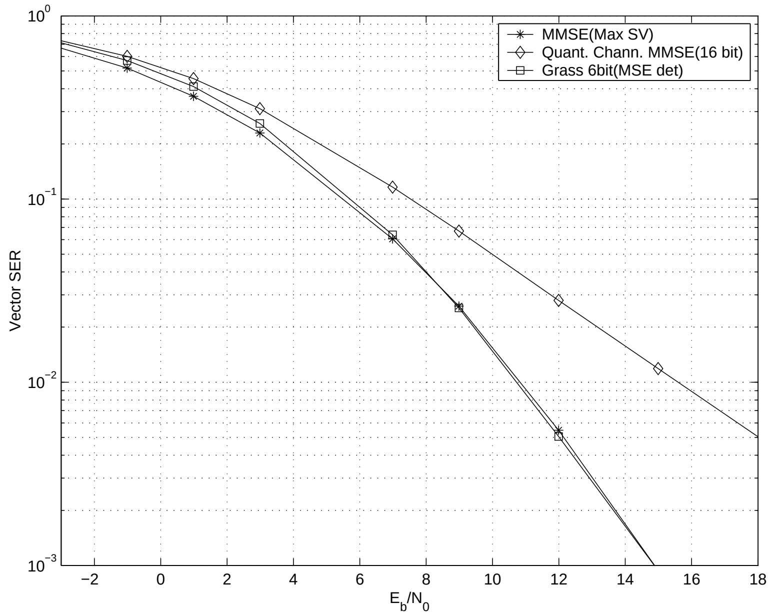

Figure 4.7: Probability of symbol vector error comparison for direct channel quantization, unquantized MMSE, and limited feedback precoding on a 2 substream 4×2 system using 16-QAM.

sults are presented in Fig. 4.7. This experiment considered a 4×2 wireless system using two substrems and 16-QAM. Directly quantizing the channel with sixteen bits of feedback performs approximately 4.7dB worse than a six bit limited feedback precoder at a probability of symbol vector error of 10^{-2}. The limited feedback precoder obtains performance approximately identical to that of the unquantized MMSE precoder with the maximum singular value power constraint.

Chapter 5

Multi-Mode Precoding

This chapter overviews a new adaptive modulation for MIMO systems called multi-mode precoding. Section 5.1 introduces the multi-mode precoded spatial multiplexing system model. Criteria for choosing the optimal matrix from the codebook is presented in Section 5.2. We discuss the implementation of multi-mode precoding in systems with perfect CSI at the transmitter in Section 5.3. The case of no CSI at the transmitter is considered in Section 5.4. Section 5.5 characterizes the diversity and multiplexing gains of multi-mode precoding. The relationship between limited feedback multi-mode precoding and covariance quantization is explored in Section 5.6. Section 5.7 illustrates the performance improvements over previously proposed techniques using Monte Carlo simulations of the symbol error rate and mutual information.

5.1 System Overview

The M_t transmit and M_r receive antenna MIMO wireless system studied in this chapter is shown in Fig. 5.1. For each channel use, R bits are demultiplexed into M different bit streams. Each bit stream is modulated using the same constellation $\mathcal{S}$, producing a vector $\mathbf{s}_k$ at the k^{th} channel use. This means that *each substream carries R/M bits of information.* The spatial multiplexing symbol vector $\mathbf{s}_k = [s_{k,1}\ s_{k,2}\ \ldots\ s_{k,M}]^T$ is assumed to have power constraints so that $E_{\mathbf{s}_k}[\mathbf{s}_k\mathbf{s}_k^*] = \frac{\mathcal{E}_s}{M}\mathbf{I}_M$. Note that this means the *average of the total transmitted power* at any channel use is *independent of the number of substreams M*.

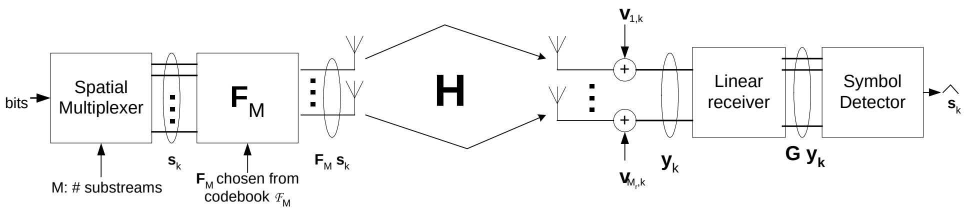

Figure 5.1: Block diagram of a limited feedback precoding MIMO system.

An $M_t \times M$ linear precoding matrix $\mathbf{F}_M$ maps $\mathbf{s}_k$ to an M_t-dimensional spatio-temporal signal that is transmitted on M_t transmit antennas. The transmitted signal vector encounters an $M_r \times M_t$ matrix channel $\mathbf{H}$ before being added with an M_r-dimensional white Gaussian noise vector $\mathbf{v}_k$. Assuming perfect pulse-shaping, sampling, and timing, this formulation yields an input-output relationship

$$\mathbf{y} = \mathbf{H}\mathbf{F}_M\mathbf{s} + \mathbf{v} \tag{5.1}$$

where the channel use index k has been suppressed because we are interested in vector-by-vector detection of $\mathbf{s}_k$. We assume that $\mathbf{H}$ has i.i.d. entries with each distributed according to $\mathcal{CN}(0,1)$. We employ a block-fading model where the channel is constant over multiple frames before independently taking a new realization. As well, the noise vector $\mathbf{v}$ is assumed to have i.i.d. entries distributed according to $\mathcal{CN}(0, N_0)$.

The matrix $\mathbf{HF}_M$ can be thought of as an *effective channel*. The receiver decodes $\mathbf{y}$ using this effective channel and a linear receiver. It is assumed that the receiver has perfect knowledge of $\mathbf{H}$ and $\mathbf{F}_M$. The receiver applies an $M \times M_r$ matrix transformation $\mathbf{G}$ to $\mathbf{y}$ and then independently detects each entry of $\mathbf{Gy}$. If a zero-forcing (ZF) decoder is used, $\mathbf{G} = (\mathbf{HF}_M)^{\dagger}$. Minimum mean squared error (MMSE) decoding uses $\mathbf{G} = [\mathbf{F}_M^*\mathbf{H}^*\mathbf{HF}_M + (N_0 M/\mathcal{E}_s)\mathbf{I}_M]^{-1}\mathbf{F}_M^*\mathbf{H}^*$.

Note that the total instantaneous transmitted power for this system is given by $\mathbf{s}^*\mathbf{F}_M^*\mathbf{F}_M\mathbf{s}$. The precoder matrix $\mathbf{F}_M$ must therefore be constrained in order to limit the transmitted power. We will restrict $\lambda_1^2\{\mathbf{F}_M\} \leq 1$ in order to limit the peak-to-average ratio. This means that $E_{\mathbf{s}}\left[\mathbf{s}^*\mathbf{F}_M^*\mathbf{F}_M\mathbf{s}\right] \leq \mathcal{E}_s$ regardless of the modulation scheme or the value of M. It was shown in [28], [72], [73] that matrices of this form that optimize MSE, capacity, and total channel power are all members of the set $\mathcal{U}(M_t, M) = \{\mathbf{U} \in \mathbb{C}^{M_t,M} \mid \mathbf{U}^*\mathbf{U} = \mathbf{I}_M\}$. For this reason, we will further restrict that $\mathbf{F}_M \in \mathcal{U}(M_t, M)$ for any chosen value of M.

The *key difference* between multi-mode precoding and previously proposed linear precoders is that M *is adapted using current channel conditions* by allowing M to vary between 1 and M_t. We refer to the value of M as the *mode* of the precoder. Usually, only a subset of the M_t possible modes can be chosen. Examples of why only subsets of $\{1, \ldots, M_t\}$ might be chosen include that i.) R/M is an integer for only a few of the modes between 1 and M_t

and ii.) system architecture only supports a subset. We will denote the set of supported mode values as $\mathcal{M}$. For example, if $R = 8$ bits and $M_t = 4$ then only modes in the set $\mathcal{M} = \{1, 2, 4\}$ can be supported. Another example is a *dual-mode precoding* where $\mathcal{M} = \{1,\ M_t\}$.

We assume that a selection function $g : \mathbb{C}^{M_r \times M_t} \rightarrow \mathcal{M}$ picks the "best" mode according to some criterion, and we set $M = g(\mathbf{H})$. After M is chosen, the precoder $\mathbf{F}_M$ is chosen from a set $\mathcal{F}_M \subseteq \mathcal{U}(M_t, M)$ using a selection function $f_M : \mathbb{C}^{M_r \times M_t} \rightarrow \mathcal{F}_M$. Therefore, $\mathbf{F}_M = f_M(\mathbf{H})$. This means that there are $|\mathcal{M}|$ different precoder selection functions. Note that we have made no assumptions about the number of matrices in $\mathcal{F}_M$.

We consider two different transmitter CSI scenarios: full transmitter CSI and zero transmitter CSI. These are the two extremes of transmitter channel knowledge. The full transmitter CSI assumption is approximately valid for a time division duplexing (TDD) system transmitting over a channel that changes slowly in comparison to the duplexing period. The assumption of zero channel knowledge approximates a frequency division duplexing (FDD) system where the forward and reverse frequency bands are separated by a frequency bandwidth much larger than the channel coherence bandwidth.

5.2 Multi-Mode Precoder Selection

The selection of the mode and precoder matrix will determine the performance of the entire system. Because we are interested in constructing a high-rate signaling scheme with low error rates, we will present bounds on the probability of symbol vector error (i.e. the probability that the linear decoder returns at least one symbol in error). As stated earlier, we consider receivers using ZF and MMSE linear receivers. We will also review the capacity results for MIMO systems both with and without transmitter CSI.

5.2.1 Performance Discussion

The selection criterion used to choose M and $\mathbf{F}_M$ must tie directly to the resulting performance of the precoded spatial multiplexing system. We will address selection details based on two different performance measures: probability of error and capacity.

Probability of Error

It was shown in [71] that the effective SNR of the k^{th} substream after linear processing is given by

$$SNR_k^{(ZF)}(\mathbf{F}_M) = \frac{\mathcal{E}_s}{MN_0[\mathbf{F}_M^*\mathbf{H}^*\mathbf{H}\mathbf{F}_M]_{k,k}^{-1}} \tag{5.2}$$

for ZF decoding and

$$SNR_k^{(MMSE)}(\mathbf{F}_M) = \frac{\mathcal{E}_s}{MN_0[\mathbf{F}_M^*\mathbf{H}^*\mathbf{H}\mathbf{F}_M \ + \ (MN_0/\mathcal{E}_s)\mathbf{I}_M]_{k,k}^{-1}} - 1 \tag{5.3}$$

for MMSE decoding where $\mathbf{A}_{k,k}^{-1}$ is entry (k,k) of $\mathbf{A}^{-1}$. The minimum substream SNR, given by

$$SNR_{min}(\mathbf{F}_M) = \min_{1\le k\le M} SNR_k(\mathbf{F}_M), \tag{5.4}$$

is an important parameter that will be used to characterize performance. For ZF and MMSE decoding, $SNR_{min}(\mathbf{F}_M)$ can be bounded by [71]

$$SNR_{min}(\mathbf{F}_M) \ge \lambda_M^2\{\mathbf{H}\mathbf{F}_M\}\frac{\mathcal{E}_s}{MN_0}. \tag{5.5}$$

Therefore, *the minimum singular value of the effective channel is often an important parameter in linear precoded MIMO systems.* Because the linear precoding matrix $\mathbf{F}_M$ is restricted to lie in the set $\mathcal{U}(M_t, M)$,

$$\lambda_M\{\mathbf{H}\mathbf{F}_M\} \leq \lambda_M\{\mathbf{H}\} \tag{5.6}$$

by the Poincaré separation theorem [83], p. 190.

We are more interested, however, in tight bounds on the probability of symbol vector error. Let $d_{min}(M, R)$ denote the minimum distance for the constellation $\mathcal{S}$ used when R/M bits are modulated per substream and $N_e(M, R)$ denote the average number of nearest neighbors of $\mathcal{S}$. Given a matrix channel $\tilde{\mathbf{H}}$, the conditional probability of symbol vector error can be bounded by the Nearest Neighbor Union Bound (NNUB) as [96]

$$P_e(\tilde{\mathbf{H}}) \leq 1 - \left(1 - N_e(M, R) Q\left(\sqrt{SNR_{min}\frac{d_{min}^2(M, R)}{2}}\right)\right)^M . \tag{5.7}$$

where SNR_{min} is computed for the given linear receiver. In stochastic channels, this average NNUB on the probability of symbol vector error is given by taking the expectation of (5.7),

$$\begin{aligned} P_e &= E_{\tilde{\mathbf{H}}}[P_e(\tilde{\mathbf{H}})] \\ &\leq 1 - E_{\mathbf{H}}\left[\left(1 - N_e(M, R) Q\left(\sqrt{SNR_{min}\frac{d_{min}^2(M, R)}{2}}\right)\right)^M\right]. \end{aligned} \tag{5.8}$$

The NNUB is asymptotically tight [96], so the bound can be used to determine the *diversity order* of a communication system. A MIMO wireless system is said to have diversity order d [36] if

$$d = -\lim_{\mathcal{E}_s/N_0 \to \infty} \frac{\log(P_e)}{\log(\mathcal{E}_s/N_0)}. \tag{5.9}$$

Diversity order, or diversity gain, is one of the fundamental parameters for MIMO systems. The diversity order is always bounded above by the product of the number of transmit and receive antennas, $M_t M_r$.

Capacity

The mutual information of the channel $\mathbf{HF}_M$ assuming uncorrelated Gaussian signaling on each antenna, denoted C_{UT}, is well-known to be

$$C_{UT}(\mathbf{F}_M) = \log_2 det\left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0}\mathbf{F}_M^*\mathbf{H}^*\mathbf{HF}_M\right). \tag{5.10}$$

The notation C_{UT} is used because this is commonly called the uninformed transmitter (UT) capacity (i.e. no transmitter CSI) [16]. Note that this is not really a "capacity" expression in the sense of distribution maximization because we assume a fixed distribution [93]. We will, however, refer to (5.10) as the capacity of the effective channel $\mathbf{HF}_M$ when there is no form of CSI at the transmitter in order to follow existing terminology in the MIMO literature.

When transmitter and receiver both have perfect knowledge of $\mathbf{H}$ and $\mathbf{F}_M$ the capacity is found by waterfilling [1], [16], [17]. Let $r = rank(\mathbf{HF}_M)$. The capacity of the informed transmitter is given by

$$C_{IT}(\mathbf{F}_M) = \sum_{j=1}^{r}\log_2\left(1 + \frac{\mathcal{E}_s}{MN_0}\gamma_j\lambda_j^2\{\mathbf{HF}_M\}\right) \tag{5.11}$$

where

$$\textstyle\sum_{j=1}^{r}\gamma_j = M \qquad \text{and} \qquad \gamma_j = \max\left(0, \mu - \frac{MN_0}{\mathcal{E}_s\lambda_j^2\{\mathbf{HF}_M\}}\right).$$

Just as diversity order is often used to characterize probability of error performance, *multiplexing gain* can be used to characterize spectral efficiency. Let $R(\mathcal{E}_S/N_0)$ denote the supported data rate as a function of SNR. A MIMO wireless system is said to have a multiplexing gain of c [36] if

$$c = \lim_{\mathcal{E}_s/N_0\to\infty}\frac{R(\mathcal{E}_s/N_0)}{\log(\mathcal{E}_s/N_0)}. \tag{5.12}$$

Multiplexing gain is a fundamental property and, in our system, will be bounded from above by $\min(M_t, M_r)$.

5.2.2 Selection Criteria

We will present probability of error and capacity selection. The probability of error selection is based on the previous work in [40], [71], [97], [98] while the capacity selection is similar to work presented in [65], [71], [99].

Probability of Error Selection

Assuming a probability of error selection criterion, the optimal selection criterion would obviously be to choose the mode and precoder that provide the lowest probability of symbol vector error. Selection using this criterion, however, is unrealistic because closed-form expressions for the probability of vector symbol error conditioned on a channel realization are not available to the authors' knowledge. The NNUB can be successfully employed in place of this bound for asymptotically tight selection. Using the NNUB result in Section 5.2.1, the following selection criterion is obtained.

NNUB Selection Criterion: Choose M and $\mathbf{F}_M$ such that

$$g^{(UB)}(\mathbf{H}) = \underset{m\in\mathcal{M}}{\operatorname{argmin}} \left(1 - \left(1 - N_e(m,R) Q\left(\sqrt{SNR_{min}(f_m^{(UB)}(\mathbf{H}))\frac{d_{min}^2(m,R)}{2}}\right)\right)^m\right) \tag{5.13}$$

$$f_M^{(UB)}(\mathbf{H}) = \underset{\mathbf{F}'\in\mathcal{F}_M}{\operatorname{argmin}} 1 - \left(1 - N_e(M,R) Q\left(\sqrt{SNR_{min}(\mathbf{F}')\frac{d_{min}^2(M,R)}{2}}\right)\right)^M. \tag{5.14}$$

The value of $N_e(M,R)$ and $d_{min}(M,R)$ can be computed offline and stored in the receiver's memory. The cost function can be implemented using a brute force search when there are only a finite number of matrices per codebook. Unfortunately, the Q-function is often problematic to implement.

This issue can be overcome by further simplifying the cost function evaluated in (5.14). The Q-function in (5.14) numerically dominates the constant terms because it decreases exponentially. For this reason, it is often beneficial to use the following approximate selection criterion.

SNR Selection Criterion: Choose M and $\mathbf{F}_M$ such that

$$g^{(SNR)}(\mathbf{H}) = \underset{m \in \mathcal{M}}{\operatorname{argmax}}\, SNR_{min}(f_m^{(SNR)}(\mathbf{H}))d_{min}^2(m, R)$$
$$f_M^{(SNR)}(\mathbf{H}) = \underset{\mathbf{F}' \in \mathcal{F}_M}{\operatorname{argmax}}\, SNR_{min}(\mathbf{F}'). \quad (5.15)$$

The SNR criterion represents a large complexity cost savings compared to the NNUB criterion because of the removal of the Q-function. There are two problems, however, with the SNR selection criterion. First, the functions $f_M^{(SNR)}(\mathbf{H})$ and $g^{(SNR)}(\mathbf{H})$ are *coupled*. This means that the functions can not be implemented serially because $f_m^{(SNR)}(\mathbf{H})$ must be computed for all $m \in \mathcal{M}$ before $g^{(SNR)}(\mathbf{H})$ is evaluated. The second problem is that *the NNUB and SNR selection criteria provide no intuition about the effect of a channel's structure on mode selection.* Intuitively, the singular value structure of the channel should have a *natural mode* of communication.

Note that $\mathbf{HF}_M$ is at most a rank M matrix. Thus for each value of M

$$\frac{\mathcal{E}_s}{MN_0} \max_{\mathbf{F}' \in \mathcal{F}_M} \lambda_{M_t}^2\{\mathbf{HF}'\} \leq \max_{\mathbf{F}' \in \mathcal{F}_M} SNR_{min}(\mathbf{F}') \leq \frac{\mathcal{E}_s}{MN_0} \max_{\mathbf{F}' \in \mathcal{F}_M} \lambda_M^2\{\mathbf{HF}'\} \quad (5.16)$$

by the Poincare Separation Theorem [83], pp. 190. A selection criterion based on selecting a mode only if *the best achievable SNR approximately minimizes the NNUB.* This corresponds to the following selection criterion.

Singular Value (SV) Selection Criterion: Choose M and $\mathbf{F}_M$ such that

$$g^{(SV)}(\mathbf{H}) = \underset{m \in \mathcal{M}}{\operatorname{argmax}} \frac{\lambda_m^2\{\mathbf{H}\}}{m} d_{min}^2(m, R) \tag{5.17}$$

$$f_M^{(SV)}(\mathbf{H}) = \underset{\mathbf{F}' \in \mathcal{F}_M}{\operatorname{argmax}} \lambda_M^2\{\mathbf{HF}'\}. \tag{5.18}$$

Despite the fact that this bound is approximate, the probability of error performance was shown in [40] to be quite close to both the SNR selection and NNUB selection. The important advantage of this criterion is that the selection of M and $\mathbf{F}_M$ are decoupled. In particular, the computation of the smallest singular value of $\mathbf{HF}'$ for all codebook matrices in $\mathcal{F}_M$ must only be completed for one value of M rather than for every value in $\mathcal{M}$. The criterion also gives a clear picture of the relationship between a given channel realization and mode selection. *The channel rank and conditioning dictate the dimensionality of the transmitted symbol vector.*

Capacity Selection

While capacity selection is not optimal from a probability of error point-of-view, it can provide insight into the attainable spectral efficiencies given the multi-mode precoding system model. Because the uninformed transmitter capacity is evaluated in closed-form given a matrix channel, the following criterion can be succinctly stated.

Capacity Selection Criterion: Choose M and $\mathbf{F}_M$ such that

$$g^{(Cap)}(\mathbf{H}) = \underset{m \in \mathcal{M}}{\operatorname{argmax}} C_{UT}(f_m(\mathbf{H}))$$

$$f_M^{(Cap)}(\mathbf{H}) = \underset{\mathbf{F}' \in \mathcal{F}_M}{\operatorname{argmax}} C_{UT}(\mathbf{F}_M). \tag{5.19}$$

5.2.3 Minimum Distance Calculations

The minimum distance expressions for $d_{min}(M, R)$ often have closed-form expressions for common calculations. Consider the minimum distance of the beamforming constellation (i.e. $M = 1$) $d_{min}(1, R)$. The squared minimum beamforming constellation distance for QAM constellations is given by

$$d^2_{min}(1, R) = \frac{6\mathcal{E}_s}{2^R - 1}$$

and for PSK constellations by

$$d^2_{min}(1, R) = 4\mathcal{E}_s \sin^2\left(\frac{\pi}{2^R}\right).$$

When $M > 1$,

$$d^2_{min}(M, R) = d^2_{min}(1, R/M).$$

The SNR and SV selection criteria both require the use of the ratios of $d^2_{min}(m_1, R)/d^2_{min}(m_2, R)$ for $m_1, m_r \in \mathcal{M}$. This simplifies to

$$\frac{d^2_{min}(m_1, R)}{d^2_{min}(m_2, R)} = \frac{2^{R/m_2} - 1}{2^{R/m_1} - 1}$$

for QAM constellations and

$$\frac{d^2_{min}(m_1, R)}{d^2_{min}(m_2, R)} = \frac{\sin^2\left(\frac{\pi}{2^{R/m_1}}\right)}{\sin^2\left(\frac{\pi}{2^{R/m_2}}\right)}$$

for PSK constellations.

Digital implementation of the cosine function is often challenging. For high-rates, the small angle approximation can be used for PSK constellations yielding

$$d^2_{min}(1, R) \approx 4\mathcal{E}_s \left(\frac{\pi}{2^R}\right)^2. \quad (5.20)$$

The QAM distance ratios also have a simplified form for large R. For large R and QAM constellations,

$$\frac{d_{min}^2(m_1, R)}{d_{min}^2(m_2, R)} = 2^{R(1/m_2 - 1/m_1)}. \tag{5.21}$$

5.3 Multi-Mode Precoding: Perfect Transmitter Channel Knowledge

We will first address the application of multi-mode precoding when the transmitter has perfect CSI. In this case, the transmitter knows $\mathbf{H}$ *perfectly.* An example of this scenario is when the forwardlink channel matrix $\mathbf{H}$ can be exactly estimated from the reverselink channel.

In this situation, the set $\mathcal{F}_m$ for each $m \in \mathcal{M}$ will be defined as

$$\mathcal{F}_m = \mathcal{U}(M_t, m). \tag{5.22}$$

The optimization of the different selection criteria is now over an *uncountable set.* Thus the discrete search techniques employed in antenna selection are no longer applicable. To our knowledge, the NNUB and SNR selection criteria are not easily solved in the infinite feedback case. The SV and capacity selection criteria, however, are easily solved.

Let $\mathbf{H}$ have singular value decomposition (SVD) given by

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^*$$

where $\mathbf{U} \in \mathcal{U}(M_r, M_r)$, $\mathbf{V} \in \mathcal{U}(M_t, M_t)$, and $\mathbf{\Lambda}$ is a diagonal matrix with $\lambda_j\{\mathbf{H}\}$ in position (j, j). Define

$$\overline{\mathbf{V}}_M = \mathbf{V} \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0} \end{bmatrix}.$$

Note that the bound in (5.6) is achieved when

$$\mathbf{F}_M = \overline{\mathbf{V}}_M. \tag{5.23}$$

Because $\overline{\mathbf{V}}_M$ achieves the global upper bound over $\mathcal{U}(M_t, M)$, we can conclude that for the SV selection criterion

$$f_M^{(SV)}(\mathbf{H}) = \overline{\mathbf{V}}_M. \tag{5.24}$$

It was observed in [72], [73] that $\mathbf{F}_M$ is a maximizer of C_{UT} if and only if it is a minimizer of the determinant of the mean squared error (MSE) matrix considered in [28]. In [28], it was shown that $\overline{\mathbf{V}}_M$ minimizes the determinant of the MSE matrix. Therefore, the capacity selection criterion matrix selection function is defined as

$$f_M^{(Cap)}(\mathbf{H}) = \overline{\mathbf{V}}_M.$$

The form of the optimal precoding matrices illuminates an important benefit of the proposed precoder. Multi-mode precoding is of practical interest in TDD systems because it is actually *reduced complexity waterfilling*. Instead of performing power pouring, only the precoder rank is adapted. This complexity reduction is dramatic from both a power view and from a complexity view. While waterfilling requires the matrix design techniques outlined in Section 5.2.1, multi-mode precoding only requires knowledge of the M most dominant singular vectors.

5.4 Limited Feedback Multi-Mode Precoding: Zero Transmitter Channel Knowledge

We now consider the implementation of multi-mode precoding when the transmitter has *no form of channel knowledge.* This design makes multi-mode precoding practical even in systems that do not meet the assumption of full transmitter CSI.

5.4.1 Codebook Model

The design of an adaptive modulator in a system without transmitter CSI is daunting because we must find a simple method that allows the transmitter to adapt to *current* channel conditions. We will overcome the lack of transmitter CSI by using a *low-rate feedback channel* that can carry a limited number of information bits, denoted by B, from the receiver to the transmitter.

Because the feedback channel can only support a limited number of bits, the uncountable matrix sets employed in Section 5.3 are unusable. We will therefore be required to take a different approach to the precoder design. In this limited feedback scenario, the precoder $\mathbf{F}_M$ is chosen from a finite set, or *codebook*, of N_M different $M_t \times M$ precoder matrices $\mathcal{F}_M = \{\mathbf{F}_{M,1}, \mathbf{F}_{M,2}, \ldots, \mathbf{F}_{M,N_M}\}$. *Thus, we assume that there is a codebook for each supported mode value.* Because there are a total of

$$N_{total} = \sum_{m \in \mathcal{M}} N_m$$

codeword matrices, a total feedback of

$$B = \lceil \log_2(N_{total}) \rceil = \left\lceil \log_2 \left(\sum_{m \in \mathcal{M}} N_m \right) \right\rceil$$

bits is required for feedback. Feedback can thus be kept to a reasonable amount by varying the size of N_M for each mode.

There are two main problems associated with this codebook-based limited feedback system. First, we must determine how to distribute the N_{total} codewords among the modes in $\mathcal{M}$. The second problem is how to design $\mathbf{F}_M$ given M and N_M. We present solutions for both of these problems in Sections 5.4.2 and 5.4.3.

5.4.2 Codeword Distribution

The feedback amount B is often specified offline by general system design constraints. For example, only B bits of control overhead might be available in the reverselink frames. For this reason, we will assume that B *is a fixed system parameter.* Thus, we wish to understand how to distribute $N_{tot} = 2^B$ codeword matrices among the $|\mathcal{M}|$ modes.

The first step in assigning codewords is the determination of a distortion function. The distortion function must be specific to the selection function used in order to maximize performance. We will design the distortion function by attempting to force the quantized multi-mode precoder to perform identically to the unquantized (or perfect CSI) multi-mode precoder.

Because the NNUB, SNR, and SV selection all relate directly or indirectly to maximizing $\lambda_M^2\{\mathbf{HF}_M\}$, we will define the distortion conditioned on $\mathbf{H}$ and $\mathbf{M}$ to be

$$D^{(Pe)}(\mathbf{F}_M, M) = \left|\lambda_M^2\{\mathbf{H}\overline{\mathbf{V}}_M\} - \lambda_M^2\{\mathbf{HF}_M\}\right|^2 = \left|\lambda_M^2\{\mathbf{H}\} - \lambda_M^2\{\mathbf{HF}_M\}\right|^2.$$

The capacity selection will use the conditional distortion function given by

$$D^{(Cap)}(\mathbf{F}_M, M) = \left| \log_2 det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \overline{\mathbf{V}}_M^* \mathbf{H}^* \mathbf{H} \overline{\mathbf{V}}_M \right) \right.$$
$$\left. - \log_2 det \left(\mathbf{I}_M + \frac{\mathcal{E}_s}{MN_0} \mathbf{F}_M^* \mathbf{H}^* \mathbf{H} \mathbf{F}_M \right) \right|^2 . \qquad (5.25)$$

A *key difference* between two distortions is that the probability of error distortion is *independent of the SNR* while the capacity distortion is *dependent on the SNR*. This means that the capacity codebook would have to be modified each time the SNR changes. This same SNR dependence also exists in the capacity codebooks designed in [100].

Let $\mathcal{M} = \{m_1, m_2, ..., m_{|\mathcal{M}|}\}$. We wish to minimize the *average distortion*

$$\overline{D}\left(\mathcal{F}_{m_1}, \ldots, \mathcal{F}_{m_{|\mathcal{M}|}}\right) = E_M \left[E_{\mathbf{H}} \left[D(f_M(\mathbf{H}), M) \;\mid\; M \right] \right] \qquad (5.26)$$

using the appropriate distortion. We will approximate $E_{\mathbf{H}} \left[D(f_M(\mathbf{H}), M) \;\mid\; M \right]$ using a *uniform, high-resolution approximation* [82], p. 163. This assumption yields

$$E_{\mathbf{H}} \left[D(f_M(\mathbf{H}), M) \;\mid\; M \right] \approx \frac{K}{N_M^2} \qquad (5.27)$$

where K is a positive real constant that depends on the chosen selection function but is constant for all mode numbers. Without this assumption the problem is virtually unsolvable because i.) the effective channel will no longer be a Gaussian matrix and ii.) the distortion depends directly on the codebook. This assumption also allows the codeword matrices to be allocated for the different selection functions using the same optimization methods. Thus, the optimization will not directly deal with $D^{(Pe)}(\cdot, \cdot)$ or $D^{(Cap)}(\cdot, \cdot)$. While this approximation is very rough, it is sufficient enough for us to approximately allocate the codewords among the $|\mathcal{M}|$ supported levels.

Plugging (5.27) into (5.26),

$$\overline{D}\left(\mathcal{F}_{m_1}, \ldots, \mathcal{F}_{m_{|\mathcal{M}|}}\right) \approx E_M\left[\frac{K}{N_M^2}\right] = \sum_{m\in\mathcal{M}} p_m \frac{K}{N_m^2} \tag{5.28}$$

with $p_m = Prob(M = m)$ meaning the probability that the mode random variable chooses m streams. Specifically, $p_m = Prob\left(g^{(SV)}(\mathbf{H}) = m\right)$ for probability of error selection and $p_m = Prob\left(g^{(Cap)}(\mathbf{H}) = m\right)$ for capacity selection. We will minimize (5.28) given that

$$\sum_{m\in\mathcal{M}} N_m = N_{tot} \tag{5.29}$$

and $0 < N_m < N_{tot}$ for all $m \in \mathcal{M}$. Let $\mathbf{N} = \left[N_{m_1}\ N_{m_2}\ \cdots\ N_{m_{|\mathcal{M}|-1}}\right]^T$. We rewrite (5.28) using the function

$$D_{approx}(\mathbf{N}) = \sum_{l=1}^{|\mathcal{M}|-1} p_{m_l}\frac{K}{N_{m_l}^2} + p_{m_{|\mathcal{M}|}}\frac{K}{\left(N_{tot} - \sum_{k=1}^{|\mathcal{M}|-1} N_{m_k}\right)^2}.$$

Thus we will solve for the matrix allocation values $N_{m_1}, N_{m_2}, \ldots, N_{m_{|\mathcal{M}|-1}}$ that minimize D_{approx} and then set

$$N_{m_{|\mathcal{M}|}} = N_{tot} - \sum_{l=1}^{|\mathcal{M}|-1} N_{m_l}.$$

Because the Hessian of D_{approx} is a positive and diagonally dominant, and thus positive-definite, matrix if $0 < N_m < N_{tot}$ for all $m \in \mathcal{M}$, the function D_{approx} is convex. Thus the minimizer of D_{approx} can be found by solving for $\mathbf{N}$ that sets the gradient of D_{approx} equal to zero. Evaluating the partial derivative of D_{approx} with respect to N_m for $m \in \mathcal{M}$ gives

$$\frac{\partial D_{approx}}{\partial N_m}(\mathbf{N}) = -p_m\frac{2K}{N_m^3} + p_{m_{|\mathcal{M}|}}\frac{2K}{\left(N_{tot} - \sum_{k=1}^{|\mathcal{M}|-1} N_{m_k}\right)^3}. \tag{5.30}$$

The partial derivative $\frac{\partial D_{approx}}{\partial N_m}(\mathbf{N})$ will be equal to zero when

$$\sqrt[3]{p_m}\left(N_{tot} - \sum_{k=1}^{|\mathcal{M}|-1} N_{m_k}\right) - \sqrt[3]{p_{m_{|\mathcal{M}|}}}N_m = 0.$$

Solving the set of $|\mathcal{M}| - 1$ linear equations to find the values of $\mathbf{N}$ that sets the gradient equal to zero corresponds to the solution

$$\mathbf{N} = N_{tot}\mathbf{\Phi}^{-1}\mathbf{p} \tag{5.31}$$

where

$$\mathbf{\Phi} = \sqrt[3]{p_{m_{|\mathcal{M}|}}}\ \mathbf{I}_{|\mathcal{M}|-1} + \begin{bmatrix} \sqrt[3]{p_{m_1}} & \sqrt[3]{p_{m_1}} & \cdots & \sqrt[3]{p_{m_1}} \\ \sqrt[3]{p_{m_2}} & \sqrt[3]{p_{m_2}} & \cdots & \sqrt[3]{p_{m_2}} \\ \vdots & & \ddots & \vdots \\ \sqrt[3]{p_{m_{|\mathcal{M}|-1}}} & \cdots & \sqrt[3]{p_{m_{|\mathcal{M}|-1}}} & \sqrt[3]{p_{m_{|\mathcal{M}|-1}}} \end{bmatrix}$$

and

$$\mathbf{p} = \left[\sqrt[3]{p_{m_1}}\ \sqrt[3]{p_{m_2}}\ \cdots\ \sqrt[3]{p_{m_{|\mathcal{M}|-1}}}\right]^T. \tag{5.32}$$

Eq. (5.31) will return values of N_m that are real numbers rather than integers. This point can be corrected by simply rounding the numbers and reenforcing (5.29). Care must also be taken when $M_t \in \mathcal{M}$. The codebook $\mathcal{F}_{M_t}$ corresponds to choosing traditional spatial multiplexing. Thus $\mathbf{I}_{M_t} \in \mathcal{F}_{M_t}$. The identity matrix, however, is *the only matrix in* $\mathcal{F}_{M_t}$. The reason for this is that if there existed $\mathbf{F} \in \mathcal{F}_{M_t}$ with $\mathbf{F} \neq \mathbf{I}_{M_t}$ then *the performance of precoded spatial multiplexing using* $\mathbf{F}$ *over uncorrelated Rayleigh fading would exactly match the performance of traditional spatial multiplexing.* This follows easily from the fact that $\mathbf{HF}$ has the same distribution as $\mathbf{H}$ when $\mathbf{H}$ is an uncorrelated, Rayleigh fading matrix [85]. Therefore either $\mathcal{F}_{M_t} = \{\mathbf{I}_{M_t}\}$ or $\mathcal{F}_{M_t}$ is empty. Thus when (5.31) gives $N_{M_t} > 1$, (5.31) can be used to recompute N_m for all

$m \in \underline{\mathcal{M}} = \mathcal{M}\backslash\{M_t\}$. This requires replacing the p_m probabilities in (5.31) with $p_{m|\underline{\mathcal{M}}}$, the probability of mode $m \in \underline{\mathcal{M}}$ given that the chosen mode is in $\underline{\mathcal{M}}$ and the constraint

$$\sum_{m \in \underline{\mathcal{M}}} N_m = N_{tot} - 1 \tag{5.33}$$

is satisfied.

5.4.3 Codebook Criterion Given the Number of Substreams

Now that we have determined an algorithm that gives an approximate allocation of the possible codebook matrices among the modes, it is now imperative to present the design of $\mathcal{F}_M$ for each mode. There are two special cases of the multi-mode precoder codebooks. The column vectors in $\mathcal{F}_1$ correspond to beamforming vectors [59]. The design of limited feedback beamforming was explored in [22], [52], [53], [59], [63]. In particular, it was shown in [59], [63] that the set of vectors should be designed by thinking of the vectors as representing lines in $\mathbb{C}^{M_t}$. The lines can then be optimally spaced by maximizing the minimum angular separation between any two lines. The set $\mathcal{F}_{M_t}$ is trivially designed because we will require that $\mathcal{F}_{M_t} = \{\mathbf{I}_{M_t}\}$. This precoder matrix corresponds to sending a standard spatial multiplexing vector.

Codebook design for limited feedback precoding assuming a fixed number of substreams was studied in [72], [73], [101], [102] for a variety of selection functions. The NNUB and SNR selection criteria both depend on $SNR_{min}(\mathbf{F}_M)$. As stated earlier, the minimum SNR can be approximately maximized by maximizing $\lambda_M\{\mathbf{HF}_M\}$. It was shown in [102] that $\mathcal{F}_M$ should be designed by maximizing the minimum *projection two-norm subspace dis-*

tance between any two matrices in $\mathcal{F}_M$ given by

$$d_{proj}(\mathbf{F}_{M,1}, \mathbf{F}_{M,2}) = \|\mathbf{F}_{M,1}\mathbf{F}^*_{M,1} - \mathbf{F}_{M,2}\mathbf{F}^*_{M,2}\|_2. \tag{5.34}$$

The capacity selection criterion, however, motivates the use of the *Fubini-Study subspace distance.* Thus for the capacity selection criterion we would try to maximize the minimum pairwise distance given by

$$d_{FS}(\mathbf{F}_{M,1}, \mathbf{F}_{M,2}) = \arccos \left| det \left(\mathbf{F}^*_{M,1}\mathbf{F}_{M,2} \right) \right|. \tag{5.35}$$

These codebooks can be easily designed using the matrix codebooks designed for non-coherent constellations in [92]. The only modification needed is to change the distance function to the projection two-norm subspace distance. For more information on these designs consult [102].

5.5 Diversity Order & Multiplexing Gain

MIMO wireless systems have fundamental limits on the maximum attainable diversity order and multiplexing gain. Both parameters are fundamental in MIMO systems [36]. We will derive conditions for multi-mode precoding to have maximum diversity order and multiplexing gain.

The following theorem addresses the conditions sufficient to obtain full diversity order.

Theorem 2 *If $1 \in \mathcal{M}$ and $N_1 \geq M_t$ then multi-mode precoding provides full diversity order when using the NNUB selection criterion.*

Proof It is shown in [38], that a limited feedback beamformer with a vector codebook that spans $\mathbb{C}^{M_t}$ obtains full diversity order. It was shown in [59] that this condition is satisfied when $N_1 \geq M_t$ and the codebook design techniques outlined in Section 5.4.3 are used. Because the NNUB presents an upper-bound on the probability of error, a selection function based on this result will only *decrease* the probability of error asymptotically when compared with a beamforming-only system. ∎

The achievability of full diversity order is a substantial benefit. Spatial multiplexing has limited diversity order performance (ex. achieves diversity order M_r for overly complex maximum likelihood decoding), so a large diversity increase such as this adds considerable error rate improvements.

The next theorem addresses multiplexing gain.

Theorem 3 *If $M_t \in \mathcal{M}$ and $N_{M_t} > 0$ then multi-mode precoding provides full multiplexing gain.*

Proof It is shown in [36] that spatial multiplexing (i.e. using mode $M = M_t$) provides full multiplexing gain. Because the transmission rate is independent of the mode chosen, the rate growth will be proportional to $\min(M_t, M_r)$. ∎

5.6 Relation to Covariance Quantization

The capacity analysis for MIMO systems with transmitter CSI relies on *optimizing the transmit covariance matrix.* A MIMO system has a general input-output relationship

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$$

with $\mathbf{H}$ and $\mathbf{v}$ defined as in (5.1). The capacity maximizes the covariance matrix

$$\mathbf{Q} = E_{\mathbf{x}}\left[\mathbf{x}\mathbf{x}^*\right]. \tag{5.36}$$

Covariance quantization, proposed in [100], [103], chooses $\mathbf{Q}$ from a codebook $\mathcal{Q} = \{\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_N\}$. Assuming that $\mathbf{s}$ in (5.1) consists of independent entries distributed according to $\mathcal{CN}(0, \mathcal{E}_s/M)$, the covariance matrix will be $(\mathcal{E}_s/M)\mathbf{F}_M\mathbf{F}_M^*$. *Thus multi-mode precoding quantizes the set of covariance matrices assuming a rank constraint.* Let

$$\mathcal{Q}_M = \{(\mathcal{E}_s/M)\mathbf{F}\mathbf{F}^* \ \mid \ \forall\ \mathbf{F} \in \mathcal{F}_M\}.$$

This allows multi-mode precoding to be reformulated as covariance quantization with a codebook

$$\mathcal{Q} = \bigcup_{m \in \mathcal{M}} \mathcal{Q}_m.$$

Multi-mode precoding is a *rank constrained* covariance quantization. While the codebook matrices in [100] attempt to quantize a waterfilling solution, chooses a covariance rank and then allocates equal power among each mode. This avoids the power allocation problems associated with waterfilling.

5.7 Simulations

Limited feedback multi-mode precoding was simulated to exhibit the available decrease in probability of error and the increase in capacity. The capacity results are compared with the results in [100].

Experiment 1: This experiment addresses the probability of symbol vector error of 4×4 dual-mode precoding with a ZF receiver. The results are shown in Fig. 5.2. The rate is fixed at $R = 8$ bits per channel use with QAM constellations. Because the system is dual-mode, the set of supported modes is $\mathcal{M} = \{1, 4\}$. Four bits of feedback is assumed to be available. Limited feedback beamforming using four bits (see [59], [63]) and spatial multiplexing are simulated for comparison. Multi-mode precoding with the SV selection criterion provides approximately a 0.25 dB performance improvement over limited feedback beamforming. The SNR and NNUB selection criteria perform approximately the same, 0.55 dB better than the SV selection criterion. These gains are modest because of the restriction to dual-mode precoding.

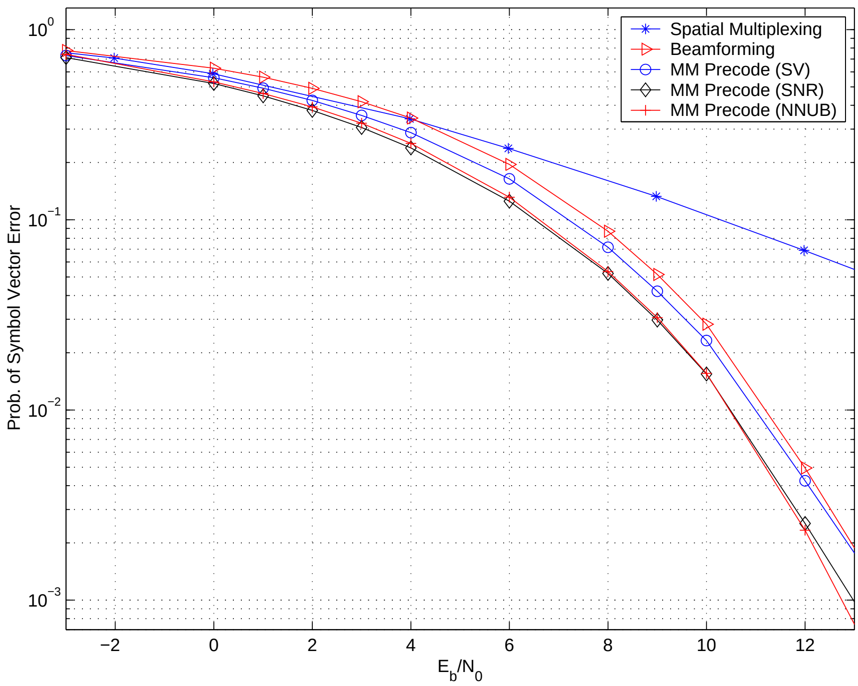

Figure 5.2: Probability of symbol vector error performance for limited feedback dual-mode precoding, limited feedback beamforming, and spatial multiplexing.

Experiment 2: In contrast to the first experiment, this experiment considers a 4×4 MIMO system transmitting $R = 8$ bits per channel use with $\mathcal{M} = \{1, 2, 4\}$. Codebooks were designed using $B = 5$ bits of feedback. Constellations were restricted to be QAM. Fig. 5.3 presents the simulation results. Spatial multiplexing, unquantized beamforming (i.e. perfect CSI at the transmitter), and unquantized MMSE precoding are shown for comparison. The

MMSE precoding was implemented as in [28] with the sum power constraint and the trace cost function. Note that all of the selection criteria provide approximately the same probability of symbol vector error performance. Five bit multi-mode precoding provides approximately a 5dB gain over full CSI beamforming. There is more than a 8.5dB gain over spatial multiplexing at an error rate of 10^{-1}. Interestingly, MMSE precoding with full transmit channel knowledge, a less restrictive power constraint, and a superior receiver provides only a 1.2dB gain over limited feedback multi-mode precoding.

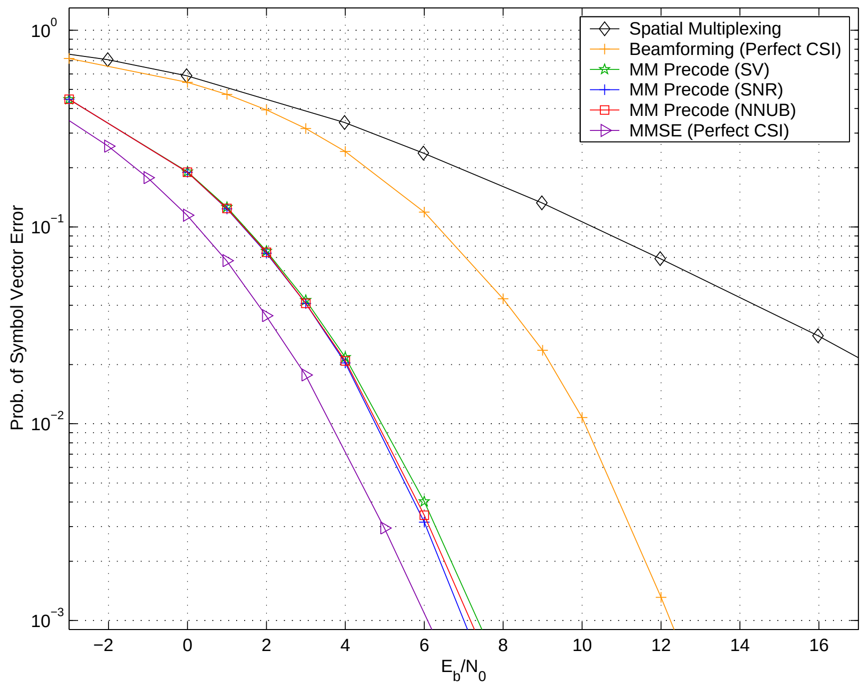

Figure 5.3: Probability of symbol vector error performance for limited feedback multi-mode precoding, beamforming, and spatial multiplexing.

Experiment 3: The third experiment compares limited feedback and perfect CSI multi-mode precoding. Again, we considered a 4×4 MIMO system transmitting $R = 8$ bits per channel use. We used a zero-forcing receiver, QAM constellations, and $B = 5$ bits of feedback for the zero CSI case. The NNUB selection criterion was employed for the limited feedback multi-mode precoder and the SV selection criterion for the perfect CSI multi-mode precoder. At low SNR, there is only a 1dB loss when employing limited feedback precoding. The loss decreases as the SNR increases. Perfect CSI beamforming is shown to demonstrate the multi-mode precoding gain.

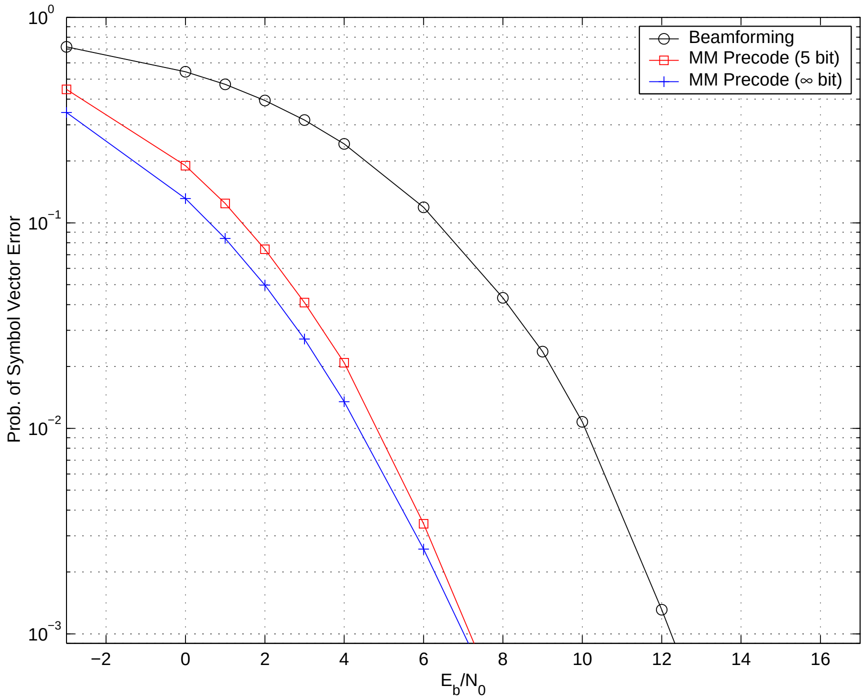

Figure 5.4: Probability of symbol vector error performance comparison for limited feedback and perfect CSI multi-mode precoding.

Experiment 4: The capacity gains available with the capacity selection criterion are illustrated in Fig. 5.5 for a 2×2 MIMO system. The plot shows the ratio of the computed mutual information with the capacity of a transmitter with perfect CSI. The capacity of an uninformed transmitter (UT) and the limited feedback covariance optimization mutual information results published in [100] are shown for comparison. Note that limited feedback precoding outperforms limited feedback covariance optimization for both two and three bits of feedback. This result is striking because, unlike covariance optimization, multi-mode precoding *does not require any form of waterfilling.* Thus our scheme, on average, always transmits with the same power on each symbol substream. The high-rate feedback performance difference between limited feedback covariance optimization and multi-mode precoding can be most likely attributed to this power-pouring.

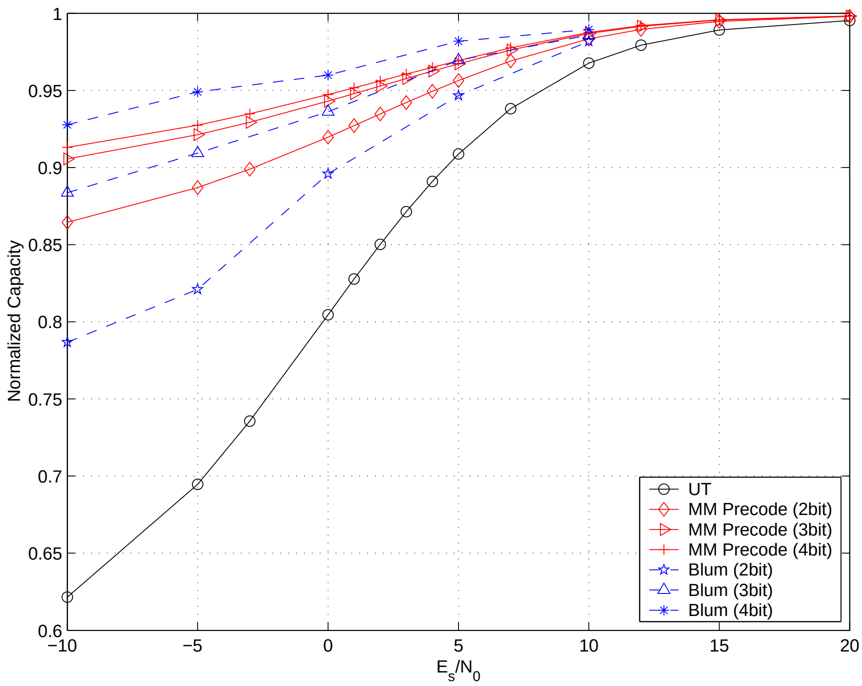

Figure 5.5: Capacity comparison of multi-mode precoding, limited feedback covariance optimization [100], and the uninformed transmitter.

Experiment 5: The fifth experiment, shown in Fig. 5.6, compares the simulated mutual information for multi-mode precoding with the capacity of the UT and beamforming (i.e. unit rank covariance constraint). We assumed perfect transmitter CSI for the multi-mode precoding and a supported mode set $\mathcal{M} = \{1, 2, 3, 4\}$. The mutual informations were normalized at each SNR value by the capacity of a MIMO system with perfect channel state information. Multi-mode precoding provides dramatic gains over other the other systems by always obtaining greater than 97.5% of the total system capacity. Thus the equal power assumption and rank tradeoff provide a close approximation to optimal waterfilling. This is an amazing achievement considering that rank adaptation is only rudimentary waterfilling.

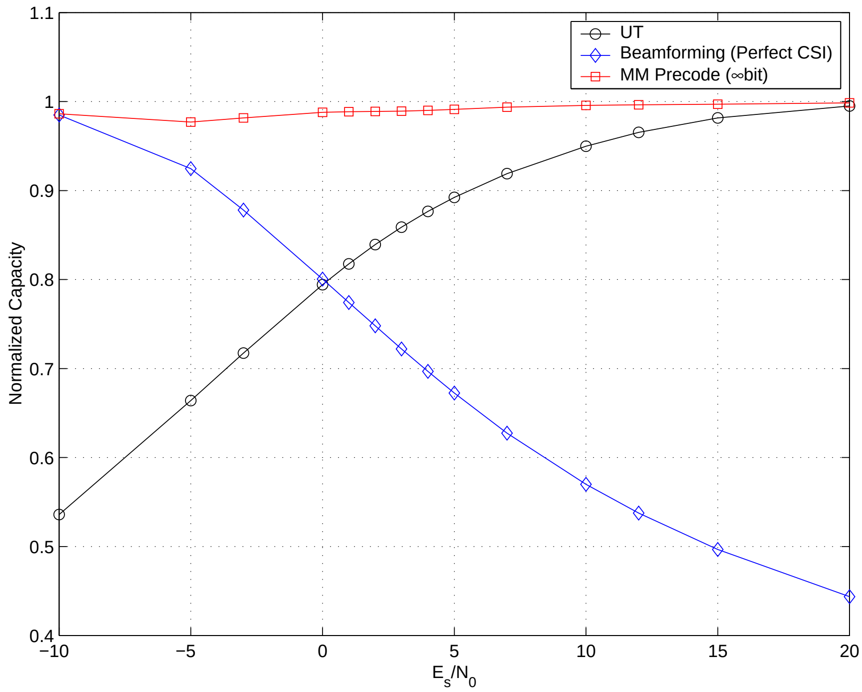

Figure 5.6: Capacity comparison of multi-mode precoding with an infinite amount of feedback, beamforming, and an uninformed transmitter.

Experiment 6: The final experiment compares infinite resolution multi-mode precoding, limited feedback multi-mode precoding, and the UT capacity for a 3×3 MIMO system. The results are shown in Fig. 5.7. We considered supported modes of $\mathcal{M} = \{1, 2, 3\}$ and four feedback bits for the limited feedback case. We designed the limited feedback codebook using techniques from Section 5.4. The matrix codeword allocations for the limited feedback case are given in Table 1. Note that the codebook differs when the SNR changes. Particularly, the number of codeword matrices shifts from the smaller to higher modes as the SNR increases. Once again, the infinite resolution multi-mode precoder obtains within 98.5% of the system capacity when the channel is known to both the transmitter and receiver. The limited feedback case obtains more than 86% of the perfect transmitter CSI system capacity. This comes with the benefit of only requiring four bits of feedback, a small number in relation to the total number of transmit antennas.

SNR	Mode 1	Mode 2	Mode 3
-10db	13	3	0
-5db	10	6	0
-3db	9	7	0
0db	6	9	1
1db	6	9	1
2db	5	10	1
3db	5	10	1
4db	4	11	1
5db	3	12	1
7db	2	13	1
10db	1	14	1
12db	1	14	1
15db	0	15	1
20db	0	15	1

Table 5.1: Codeword allocation generated for a 3x3 system with four bits of feedback.

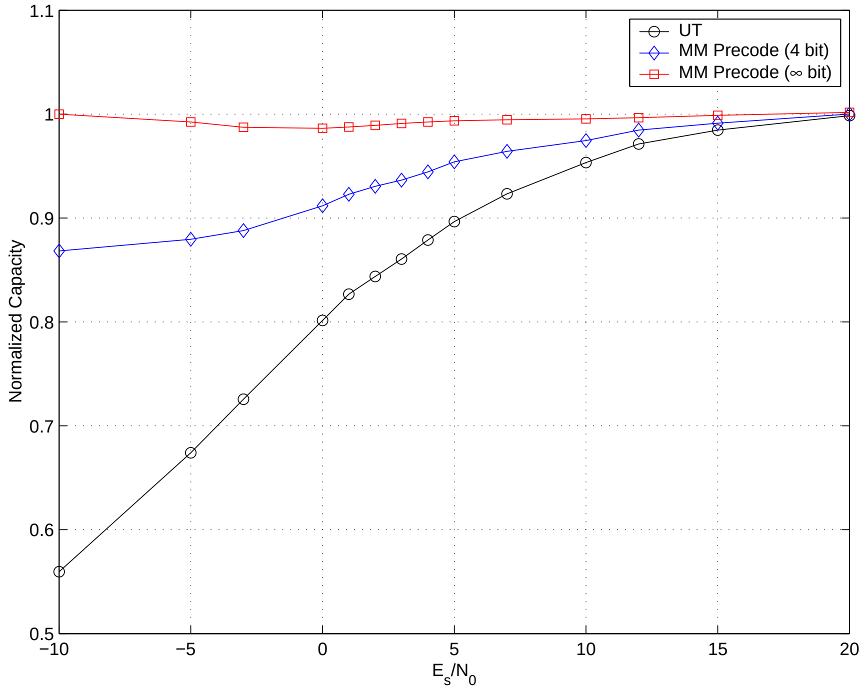

Figure 5.7: Capacity comparison of multi-mode precoding with an infinite amount of feedback, multi-mode precoding with four feedback bits, and an uninformed transmitter.

Chapter 6

Practical Aspects

The analysis presented in the preceding chapters made several assumptions that would be questionable in real systems. This section will address four of the major issues: spatial correlation, feedback delay, channel estimation error, and feedback error. The use of limited feedback schemes in broadband communications is also addressed.

6.1 Effect of Spatial Correlation

The effect of spatial correlation is often incorporated into the channel model. While the matrix Rayleigh fading channel is commonly used for system design [16], [17], several modifications have been proposed to the channel model to more realistically simulate systems. We will consider these systems along with measurement data.

In any wireless system, the "optimal" solution to test system performance is to actually build the system. While this is impractical in nearly every setting, simulations using actual channel measurements can provide a flexible environment for making algorithmic tradeoffs. Prof. Murat Torlak at The University of Texas at Dallas (UTD) has spearheaded the development of an eight transmit antenna by one receive antenna testbed. The system uses a carrier frequency of 2415 MHz and transmits a single tone at 10 dBm.

Figure 6.1: The UTD antenna array is an eight antenna circular array.

Experiment 1: Precoded OSTBCing was simulated on the channel data collected by Torlak. The date was collected in a corridor of the Erik Jonsson building with the antenna array transmitting from a lab approximately 50 feet away. The transmission was non-line-of-sight with the transmitter and receiver being separated by three walls. This experiment tested antenna subset selection, six bit precoding, and an unprecoded two-antenna Alamouti code on an 8×1 system transmitting 16-QAM. The results are shown in Fig. 6.2. The limited feedback precoding system still provides around a 1dB gain compared to antenna subset selection. This clearly presents the benefits of limited feedback in "real-world channels."

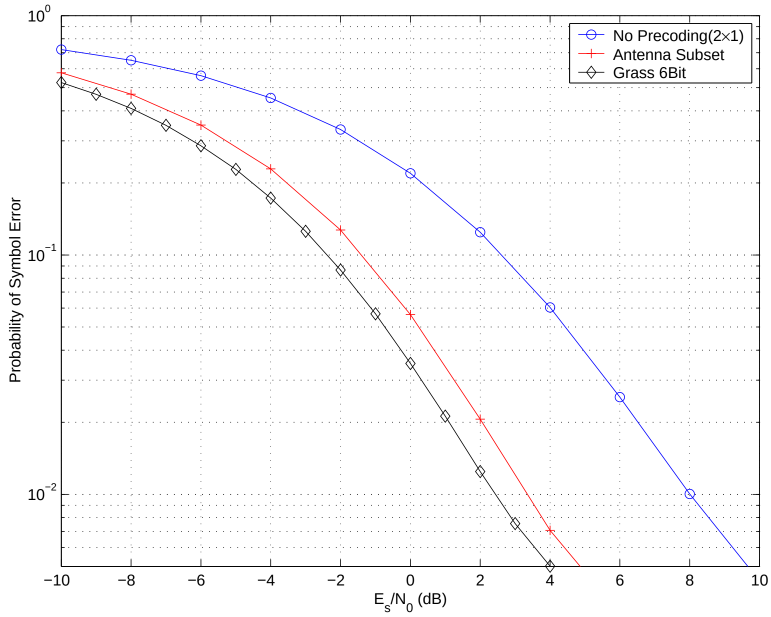

Figure 6.2: Precoded OSTBC performance tested on UTD channel measurements.

The first MIMO channel models proposed by the measurement community were ray based models. Ray based channel models are popular because they mathematically simulate (in some sense) the effect of actual scatters. The most well-known of the MIMO ray based models is the 3GPP/3GPP2 channel model. The 3GPP/3GPP2 MIMO channel model, overviewed in [104], generates each entry of the channel matrix $\mathbf{H}$ as a function of the number of specular wave components received K. The model uses numerous parameters in order to generate a realization. This model is primarily a function of the angle of arrival (AoA) and angle of departure (AoD) of each of the multipaths. The model allows for simulation in both line-of-sight (LOS) and non-line-of-sight (NLOS).

The channel model can use three different simulation environments: urban macrocell, suburban macrocell, and urban microcell. Further descriptions of the assumptions for each of these simulation environments is included in [104]. The channel is a function of 13 different parameters including AOA, AOD, wave number, transmit antenna gain, receive antenna gain, mobile velocity, mobile velocity vector angle, etc.

Despite the large number of variables, the channel is essentially a standard ray based channel model for transmit/receive uniform linear arrays (ULAs). Let D_r (D_t) denote the receive (transmit) antenna spacing, θ_r (θ_t) denote the AoA (AoD), and λ denote the carrier wavelength. A ray based channel is simulated as [105]–[107]

$$\mathbf{H} = \sum_{l=1}^{K} \alpha_l \cdot \mathbf{a}_r(\theta_{r_l})^* \mathbf{a}_t(\theta_{t_l}) \tag{6.1}$$

where

$$\mathbf{a}_r(\theta_r) = \left[1 \; exp\left(\frac{2\pi}{\lambda} j D_r \sin(\theta_r)\right) \; \cdots \; exp\left(\frac{2\pi}{\lambda} j (M_r - 1) D_r \sin(\theta_r)\right)\right],$$

$$\mathbf{a}_t(\theta_t) = \left[1 \; exp\left(\frac{2\pi}{\lambda} j D_t \sin(\theta_t)\right) \; \cdots \; exp\left(\frac{2\pi}{\lambda} j (M_t - 1) D_t \sin(\theta_t)\right)\right],$$

θ_{r_l} (θ_{t_l}) is the AoA (AoD) for the l^{th} specular wave component, and α_l is the per wave gain.

Channel generation using (6.1) requires a distribution for θ_r and θ_t. These two angles are usually generated independently using the same distribution. Popular distributions include the Laplacian, Gaussian, and truncated uniform. This model can also be generalized to fit into the stochastic local area channel (SLAC) model discussed in [105] by changing the form of the array response vectors. The SLAC model sums over all specular components assuming known path delays, wavevectors, and receiver translation vectors.

The Cost 259 model is a popular MIMO ray-based channel model [108], [109]. The model uses a fixed number of clusters and generates a certain number of rays per cluster. Each cluster has a mean AoA and AoD associated with it. Each ray has a different Rayleigh distributed path gain.

Experiment 2: This experiment simulated a 4×4 precoded spatial multiplexing system transmitting 4-QAM on a six cluster Cost 259 channel model. Two substreams were transmitted. To model insufficient local scattering, the number of rays per cluster was set to one. The results are shown in Fig. 6.3. Note that a diversity difference is still evident between the precoded and unprecoded systems. In particular, there is a gain of approximately 4dB at an error rate of 10^{-2} when using precoding.

Since the introduction of MIMO, researchers in the areas of wireless channel measurement and electromagnetics recognized the need to develop new stochastic channel models that closely matched reality. This specifically motivated the Information Society Technologies (IST) Multi-Element Transmit and Receive Antennas (METRA) Project [110] funded by the European Union and a consortium of governments and companies. The goal was to design a channel model for MIMO performance evaluation that would match measurements. In particular, their model generalizes the ray based models for moderate to large numbers of scatterers. Practically, this number has been found to be approximately $K > 10$ [106].

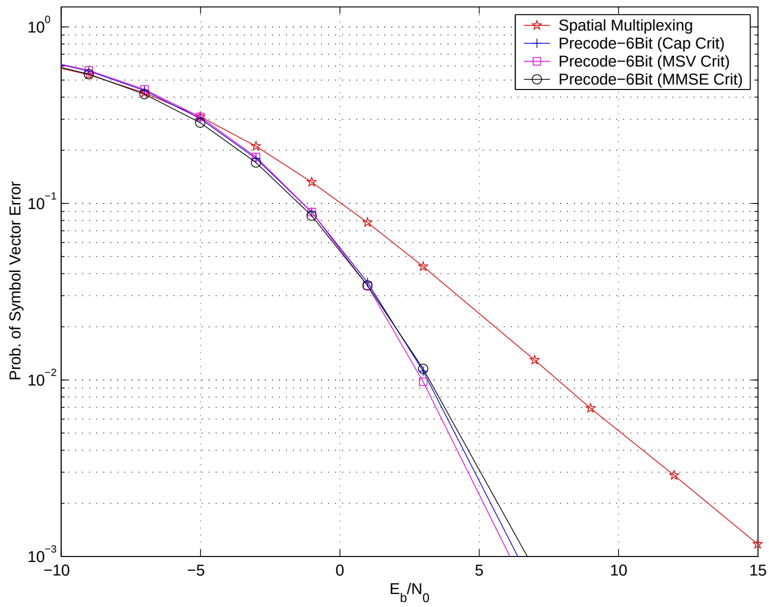

Figure 6.3: Precoded spatial multiplexing tested on the Cost 259 channel model.

The basic model is to generate the channel as [111]

$$\mathbf{H} = \mathbf{R}_R^{1/2}\mathbf{G}\mathbf{R}_T^{1/2} \tag{6.2}$$

where $\mathbf{G}$ is a random matrix with i.i.d. $\mathcal{CN}(0,1)$ entries, $\mathbf{R}_R$ is the receive antenna correlation matrix, and $\mathbf{R}_T$ is the transmit antenna correlation matrix. These correlation matrices can be computed or measured. Realistic computational methods for obtaining the correlations were studied in [107], [111], [112], while experimentally measured correlation matrices were presented in [113].

The IST METRA model was experimentally verified in [113]–[118]. The model has been verified in a number of criteria including error rate, capacity, and singular value distribution. This model has been adopted as the MIMO channel model by the IEEE 802.11N wireless LAN high throughput study group [111] and has been studied for use in IEEE 802.20 working group on mobile broadband wireless access [119]. The model is thus expected to be the cornerstone by which all future MIMO algorithms are tested before implementation.

Unlike ray based models, analytical error rate results can be derived for the IST METRA model. Specifically, it is of interest to understand the requirements that $\mathbf{R}_R$ and $\mathbf{R}_T$ must maintain to guarantee full diversity order. We will first characterize the conditions needed for a beamformer to achieve full diversity order.

Theorem 4 *A wireless system employing limited feedback beamforming and combining over memoryless, correlated Rayleigh fading channels provides full diversity order if and only if the vectors in the beamforming feasible set span $\mathbb{C}^{M_t}$, the vectors in the combining feasible set span $\mathbb{C}^{M_r}$, and the matrices $\mathbf{R}_T$ and $\mathbf{R}_R$ are invertible.*

Proof We will first prove the sufficiency of the conditions. Suppose that the beamforming and combining vectors span $\mathbb{C}^{M_t}$ and $\mathbb{C}^{M_r}$, respectively. Note that the diversity order will always be less than or equal to M_tM_r because there are only M_tM_r independently fading parameters. Let $\mathcal{W}$ denote the beamforming feasible set and $\mathcal{Z}$ denote the combining feasible set. We can construct an invertible matrix $\mathbf{B} = \mathbf{R}_T^{1/2}\mathbf{W} = \mathbf{R}_T^{1/2}[\mathbf{w}_1 \ \mathbf{w}_2 \ \cdots \ \mathbf{w}_{M_t}]$ where $\mathbf{w}_i \in \mathcal{W}$ for all i. We can similarly construct an invertible matrix $\mathbf{C} = \mathbf{R}_R^{1/2^*}\mathbf{Z} = \mathbf{R}_R^{1/2^*}[\mathbf{z}_1 \ \mathbf{z}_2 \ \cdots \ \mathbf{z}_{M_r}]$ where $\mathbf{z}_i \in \mathcal{Z}$ for all i. Since the matrices are invertible, we can define a singular value decomposition of each matrix

$$\mathbf{B} = \mathbf{V}_L \mathbf{\Lambda} \mathbf{V}_R^* \quad \text{and} \quad \mathbf{C} = \mathbf{U}_L \mathbf{\Phi} \mathbf{U}_R^* \tag{6.3}$$

where $\mathbf{V}_L$ and $\mathbf{V}_R$ are $M_t \times M_t$ unitary matrices, $\mathbf{\Lambda}$ is a diagonal matrix with diagonal entries $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{M_t} > 0$, $\mathbf{U}_L$ and $\mathbf{U}_R$ are $M_r \times M_r$ unitary matrices, and $\mathbf{\Phi}$ is a diagonal matrix with diagonal entries $\phi_1 \geq \phi_2 \geq \ldots \geq \phi_{M_r} > 0$.

For this system,

$$\begin{aligned}
\mathbf{\Gamma}_r &= \max_{\mathbf{w}\in\mathcal{W}} \max_{\mathbf{z}\in\mathcal{Z}} \left|\mathbf{z}^*\mathbf{H}\mathbf{w}\right|^2 \\
&\geq \max_{1\leq p\leq M_r} \max_{1\leq q\leq M_t} \left|\mathbf{z}_p^*\mathbf{H}\mathbf{w}_q\right|^2 \\
&= \max_{1\leq p\leq M_r} \max_{1\leq q\leq M_t} \left|\left(\mathbf{U}_R\right)_p^* \mathbf{\Phi}\mathbf{U}_L^*\mathbf{G}\mathbf{V}_L\mathbf{\Lambda}\left(\mathbf{V}_R\right)_q\right|^2 \\
&\overset{d}{=} \max_{1\leq p\leq M_r} \max_{1\leq q\leq M_t} \left|\left(\mathbf{U}_R^*\mathbf{\Phi}\mathbf{G}\mathbf{\Lambda}\mathbf{V}_R\right)_{(p,q)}\right|^2 \qquad (6.4) \\
&\geq \frac{1}{M_t M_r}\left\|\mathbf{\Phi}\mathbf{G}\mathbf{\Lambda}\right\|_2^2 \\
&\geq \frac{1}{M_t M_r}\phi_{M_r}^2\lambda_{M_t}^2 \max_{1\leq p\leq M_r} \max_{1\leq q\leq M_t} \left|g_{p,q}\right|^2 \qquad (6.5)
\end{aligned}$$

where $\left(\mathbf{A}\right)_p$ represents the p^{th} column of a matrix $\mathbf{A}$, $\left(\mathbf{A}\right)_{(p,q)}$ is the (p,q) entry of a matrix $\mathbf{A}$, and $\overset{d}{=}$ denotes equivalence in distribution. Eq. (6.4) used the invariance of complex normal matrices to unitary transformation [85], and (6.5) follows from the matrix norm bounds described in [120].

Noting that the maximum over the entries $\left|g_{p,q}\right|^2$ is the effective channel gain for SD systems, we thus have that

$$P_s(Error) \leq E_{\mathbf{G}}\left[P_s\left(Error \;\middle|\; \frac{1}{M_t M_r}\phi_{M_r}^2\lambda_{M_t}^2\overline{\gamma}_{r_{SD}}\right)\right] \tag{6.6}$$

with

$$\overline{\gamma}_{r_{SD}} = \max_{1\leq p\leq M_r} \max_{1\leq q\leq M_t} \left|g_{p,q}\right|^2 \frac{\mathcal{E}}{N_0}.$$

This is the probability of symbol error for an uncorrelated Rayleigh fading SD system with array gain shift of $\frac{1}{M_t M_r}\phi_{M_r}^2 \lambda_{M_t}^2$ which provides a diversity of order $M_t M_r$. We have thus proven that the system achieves a diversity order of $M_t M_r$.

Now we prove necessity. Let the subspace spanned by the vectors in $\mathcal{W}$ after multiplication by $\mathbf{R}_T^{1/2}$, denoted by $\mathcal{S}_W$, be of dimension M_w and the subspace spanned by the vectors in $\mathcal{Z}$ after multiplication by $\mathbf{R}_R^{1/2^*}$, denoted by $\mathcal{S}_Z$, be of dimension M_z. Suppose that $M_w M_z < M_t M_r$. We can then find an $M_t \times M_w$ matrix $\mathbf{V}$ that spans $\mathcal{S}_W$ and an $M_r \times M_z$ matrix $\mathbf{U}$ that spans $\mathcal{S}_Z$ with $\mathbf{V}^*\mathbf{V} = \mathbf{I}_{M_w}$ and $\mathbf{U}^*\mathbf{U} = \mathbf{I}_{M_z}$. For both matrices, we can construct square unitary matrices $\overline{\mathbf{V}}$ and $\overline{\mathbf{U}}$ by concatenating $\mathrm{M}_t - M_w$ and $M_r - M_z$ orthonormal vectors to $\mathbf{V}$ and $\mathbf{U}$ respectively.

Therefore,

$$\begin{aligned}
\mathbf{\Gamma}_r &= \max_{\mathbf{w}\in\mathcal{W}} \max_{\mathbf{z}\in\mathcal{Z}} |\mathbf{z}^*\mathbf{H}\mathbf{w}|^2 \\
&\leq \max_{\mathbf{a}:\mathbf{a}^*\mathbf{a}=1} \max_{\mathbf{d}:\mathbf{d}^*\mathbf{d}=1} \beta\, |\mathbf{d}^*\mathbf{U}^*\mathbf{G}\mathbf{V}\mathbf{a}|^2 \\
&= \max_{\mathbf{a}:\mathbf{a}^*\mathbf{a}=1} \max_{\mathbf{d}:\mathbf{d}^*\mathbf{d}=1} \beta \left| \begin{pmatrix} \mathbf{d} \\ \mathbf{0} \end{pmatrix}^* \overline{\mathbf{U}}^* \mathbf{G} \overline{\mathbf{V}} \begin{pmatrix} \mathbf{a} \\ \widetilde{\mathbf{0}} \end{pmatrix} \right|^2 \\
&\stackrel{d}{=} \max_{\mathbf{a}:\mathbf{a}^*\mathbf{a}=1} \max_{\mathbf{d}:\mathbf{d}^*\mathbf{d}=1} \beta \left| \begin{pmatrix} \mathbf{d} \\ \mathbf{0} \end{pmatrix}^* \mathbf{G} \begin{pmatrix} \mathbf{a} \\ \widetilde{\mathbf{0}} \end{pmatrix} \right|^2 \\
&= \beta \left\| \mathbf{G}_{(1:M_z,1:M_w)} \right\|_2^2
\end{aligned} \tag{6.7}$$

where $\beta = \|\mathbf{R}_T\|_2 \|\mathbf{R}_R\|_2$, $\mathbf{0}$ and $\widetilde{\mathbf{0}}$ are zero vectors, and $\mathbf{G}_{(1:M_z,1:M_w)}$ represents the matrix formed from the first M_z rows and M_w columns of $\mathbf{G}$. The expression in (6.7) is the effective channel gain for an $M_w \times M_z$ MRT/MRC system with an array gain shift. Therefore,

$$P_s(Error) \geq E_{\mathbf{G}}\left[P_s\left(Error \,\middle|\, \|\mathbf{R}_T\|_2\,\|\mathbf{R}_R\|_2\overline{\gamma}_{r_{MRC}}\right)\right] \tag{6.8}$$

with

$$\overline{\gamma}_{r_{MRC}} = \|\mathbf{G}_{(1:M_z,1:M_w)}\|_2^2\frac{\mathcal{E}}{N_0}.$$

Note that an $M_w \times M_z$ MRT/MRC system provides a diversity order of M_wM_z. Thus, the bounded system provides a diversity order less than or equal to M_wM_z and thus less than M_tM_r. Since this bound is true for arbitrary $\frac{\mathcal{E}}{N_0}$, we can conclude that the system does not achieve full diversity. ∎

Theorem 4 provides the intuition that the set of possible beamformers *must always* form a basis for the vector space of possible dominant singular value directions. The correlation matrix plays an important role in the performance because it *shapes* the set of beamforming vectors. Interestingly a similar result can be derived for precoded OSTBCs.

Theorem 5 *A precoded OSTBCing system with full rank correlation matrices obtains a diversity of order M_tM_r if and only if the matrix $\mathbf{E} = [\mathbf{F}_1\ \mathbf{F}_2\ \cdots\ \mathbf{F}_N]$ is full rank.*

Proof We will first prove the "if" part of the statement using the contrapositive. Suppose that $\mathbf{E}$ *is not* full rank.

Note that

$$\max_{\mathbf{F}'\in\mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_F^2 \leq M \max_{\mathbf{F}'\in\mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_2^2 \leq M\|\mathbf{H}\mathbf{E}\|_2^2. \tag{6.9}$$

Let $\mathbf{R}_T^{1/2}\mathbf{E} = \mathbf{U}_L \mathbf{\Sigma} \mathbf{U}_R^*$ where $\mathbf{U}_L \in \mathcal{U}(M_t, M_t)$, $\mathbf{\Sigma}$ an $M_t \times NM$ diagonal matrix with entry $\lambda_i\{\mathbf{R}_T^{1/2}\mathbf{E}\}$ at entry (i,i), and $\mathbf{U}_R \in \mathcal{U}(NM, NM)$. Because of the rank deficiency in $\mathbf{E}$, $\mathbf{\Sigma}$ will be of the form $\mathbf{\Sigma} = [\overline{\mathbf{\Sigma}}^T \ \mathbf{0}]^T$ where $\overline{\mathbf{\Sigma}}$ consists of the first $rank(\mathbf{E})$ rows of $\mathbf{\Sigma}$ and $\mathbf{0}$ is a zero matrix. Substitution of the singular value decomposition of $\mathbf{R}_T^{1/2}\mathbf{E}$ in to (6.9) gives

$$\begin{aligned}
\max_{\mathbf{F}' \in \mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_F^2 &\leq M \|\mathbf{R}_R\|_2 \|\mathbf{G}\mathbf{U}_L \mathbf{\Sigma} \mathbf{U}_R^*\|_2^2 \\
&= M \|\mathbf{R}_R\|_2 \|\mathbf{G}\mathbf{U}_L [\overline{\mathbf{\Sigma}}^T \ \mathbf{0}]^T\|_2^2 \\
&= M \|\mathbf{R}_R\|_2 \|\overline{\mathbf{G}\mathbf{U}_L} \ \overline{\mathbf{\Sigma}}\|_2^2 \\
&\leq M \|\mathbf{R}_R\|_2 \|\mathbf{R}_T^{1/2}\mathbf{E}\|_2^2 \|\overline{\mathbf{G}\mathbf{U}_L}\|_F^2 \qquad (6.10)
\end{aligned}$$

where $\overline{\mathbf{G}\mathbf{U}_L}$ is the first $rank(\mathbf{E})$ columns of $\mathbf{G}\mathbf{U}_L$ and (6.10) follows from using the submultiplicative property of a matrix norm and of noting that $\|\overline{\mathbf{\Sigma}}\|_2^2 = \|\mathbf{R}_T^{1/2}\mathbf{E}\|_2^2$.

Note that the matrix $\mathbf{G}\mathbf{U}_L$ is equivalent in distribution to $\mathbf{G}$ [85]. Therefore, we have upper-bounded the effective channel gain by the effective channel gain of an $rank(\mathbf{E}) \times M_r$ multi-antenna system experiencing matrix Rayleigh fading. A transmit diversity system with the effective channel gain given by this upper bound would obtain a diversity order of $rank(\mathbf{E}) \cdot M_r < M_t M_r$ [16]. Therefore, we can conclude that if $\mathbf{E}$ is not full rank then the precoded OSTBC *does not* obtain full diversity order.

Now we will prove the converse. Suppose $\mathbf{E}$ is full rank. Then

$$\max_{\mathbf{F}' \in \mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_F^2 \geq \max_{1 \leq i \leq NM} \|\mathbf{H}\mathbf{e}_i\|_2^2 \geq \frac{1}{M_r} \|\mathbf{H}\mathbf{E}\|_1^2.$$

Once again let $\mathbf{R}_T^{1/2}\mathbf{E}$ have a singular value decomposition of $\mathbf{R}_T^{1/2}\mathbf{E} = \mathbf{U}_L \mathbf{\Sigma} \mathbf{U}_R^*$.

By substitution,

$$\begin{aligned}
\max_{\mathbf{F}'\in\mathcal{F}} \|\mathbf{H}\mathbf{F}'\|_F^2 &\geq \frac{1}{M_r}\|\mathbf{H}\mathbf{E}\|_1^2 \\
&= \frac{1}{M_r}\|\mathbf{R}_R^{1/2}\mathbf{G}\mathbf{R}_T^{1/2}\mathbf{E}\|_1^2 \\
&\geq \frac{1}{NMM_r}\|\mathbf{R}_R^{1/2}\mathbf{G}\mathbf{U}_L\boldsymbol{\Sigma}\|_1^2 \qquad (6.11) \\
&\geq \frac{\lambda_{M_t}^2\{\mathbf{R}_T^{1/2}\mathbf{E}\}}{NMM_r}\|\mathbf{R}_R^{1/2}\mathbf{G}\mathbf{U}_L\|_1^2 \\
&\geq \frac{\lambda_{M_t}^2\{\mathbf{R}_T^{1/2}\mathbf{E}\}}{NMM_rM_t^2}\|\mathbf{R}_R^{1/2}\mathbf{G}\mathbf{U}_L\|_F^2 \qquad (6.12)
\end{aligned}$$

where (6.11) uses the fact that $\|\mathbf{AU}\|_1^2 \geq \frac{1}{n}\|\mathbf{A}\|_1^2$ when $\mathbf{A} \in \mathbb{C}^{m\times n}$ and $\mathbf{U} \in \mathcal{U}(n,n)$ [59].

Therefore, we have lower-bounded the effective channel gain by the effective channel gain of an $M_t \times M_r$ transmit diversity system. $\mathbf{GU}_L$ is equivalent in distribution to an uncorrelated matrix Rayleigh fading channel [85]. A system with a effective channel gain given by (6.12) has M_tM_r diversity order. We can conclude that the precoded OSTBC obtains full diversity order [16]. ∎

We also performed numerous Monte Carlo simulations to verify the system performance. These simulations used covariance matrices taken from actual measurement data and experimentally validated in [113]. The correlation matrices were assumed fixed during the simulation.

Experiment 3: This experiment dealt with the simulation of precoded 16-QAM Alamouti codes on a 4×4 system. The channel was modeled using the "pico decorrelated" measurements METRA model from [113]. Fig. 6.4 shows the performance of six bit limited feedback precoding, antenna subset selection, and a 2×4 Alamouti code. Note that limited feedback precoding outperforms antenna subset selection by approximately 1dB and unprecoded Alamouti by approximately 4dB. This is a relatively uncorrelated channel model so there is little difference in the probability of error results.

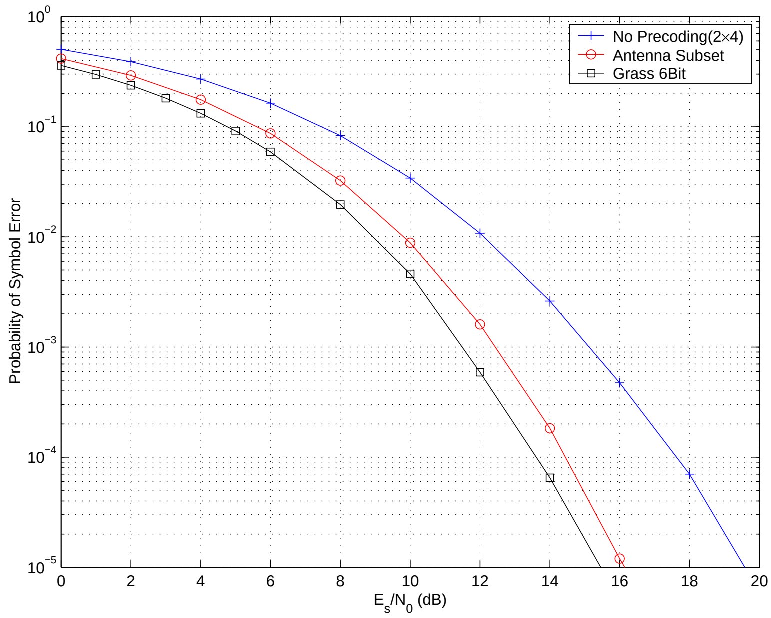

Figure 6.4: Precoded OSTBC tested on the "pico decorrelated" model from [113].

Experiment 4: The channel in Experiment 3 was modified to use the "micro decorrelated" model in [113]. The results are shown in Fig. 6.5. Note that the performance gap between antenna subset selection and limited feedback precoding has widened to around 2dB. The benefits of limited feedback, however, have diminished with the correlation. Limited feedback precoding now only has a 3dB difference compared to an unprecoded Alamouti system. Interestingly, the plots in Fig. 6.4 and 6.5 do not show much of the limited feedback diversity benefits. This can be explained by the fact that the diversity benefits only show up at extremely high SNRs.

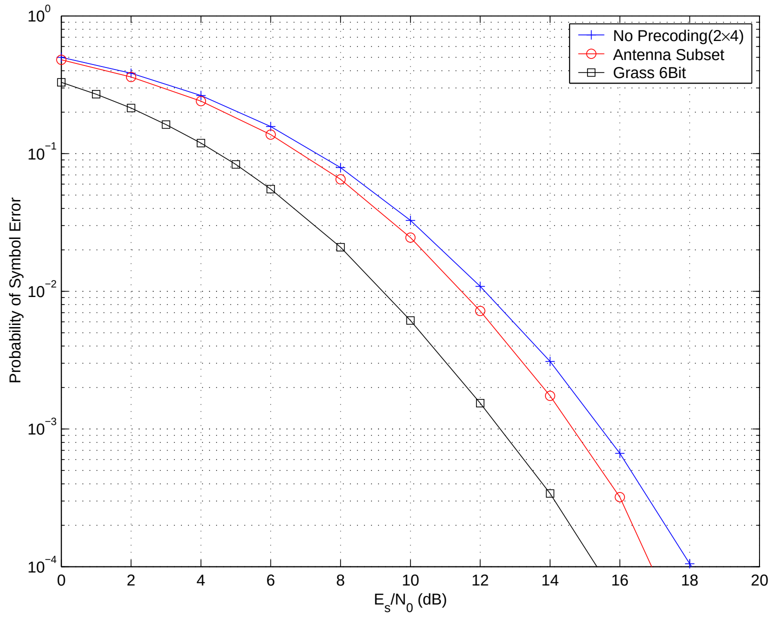

Figure 6.5: Precoded OSTBC tested on the "micro correlated" model from [113].

Experiment 5: The performance of limited feedback precoded spatial multiplexing on the "micro correlated" model is demonstrated in Fig. 6.6. A two substream 4×4 system using six bits of feedback was simulated using a 4-QAM constellation. Note in the plot that the criteria begin to vary greatly in performance. There is now more than a 3.5dB difference between the capacity and minimum singular value criteria. The effects of precoding are also startling. Limited feedback precoding using the singular value criterion outperforms unprecoded spatial multiplexing by more than 6.5dB.

Experiment 6: The 4×4 "pico decorrelated" model from [113] was simulated with multi-mode precoding using eight bits per channel use. Fig. 6.7 shows the results. Both eight bit limited feedback and perfect channel knowledge multi-mode precoding are plotted. The perfect channel knowledge

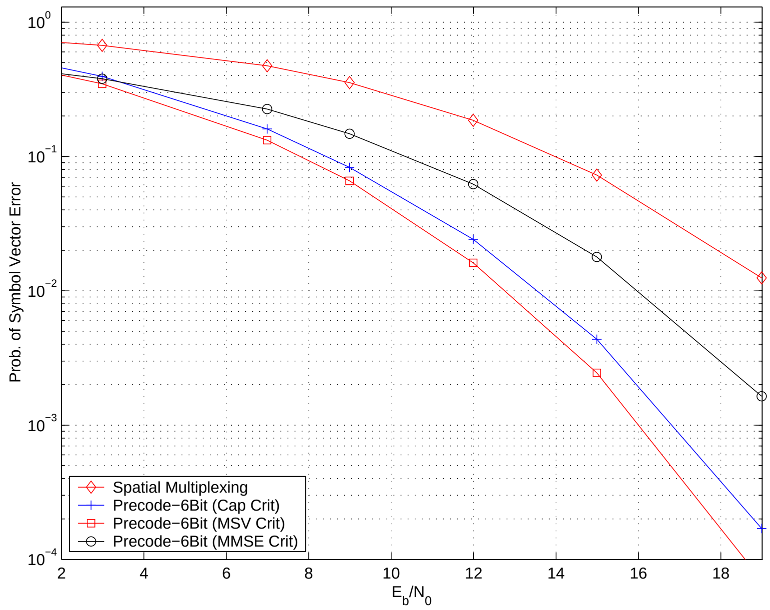

Figure 6.6: Precoded spatial multiplexing tested on the "micro correlated" model in [113].

precoder used the singular value selection criterion. At an error rate of 10^{-1} limited feedback multi-mode precoding outperforms 4×4 spatial multiplexing by approximately 12.5dB. This is a dramatic gain. Note also that the selection criteria plays an important role in performance.

Experiment 7: The performance of the same system used in Experiment 6 transmitting over the "micro correlated" model from [113] is shown in Fig. 6.8. Note the poor performance of spatial multiplexing in a highly correlated channel. This would be expected because spatial multiplexing is highly dependent on the channel singular value structure. In this simulation, the effect of limited feedback on performance is also large. The perfect channel knowledge case outperforms limited feedback by approximately 5dB.

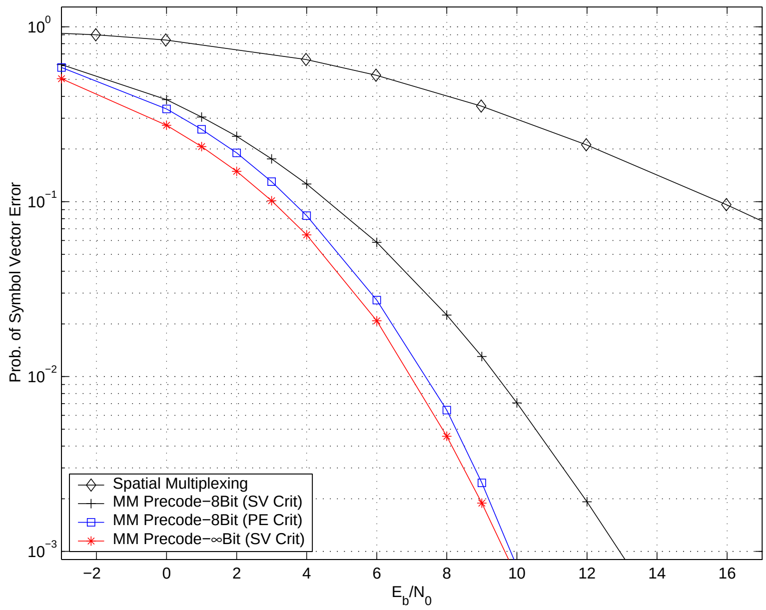

Figure 6.7: Multi-mode precoding tested on the "pico decorrelated" model in [113].

6.2 Effect of Feedback Delay

In the previous chapters, it was assumed that there was no delay in the feedback channel. This assumption is quite erroneous in any real environment where the receiver channel estimation period and low rate of the feedback channel combine to produce an out-of-date precoder at the transmitter.

In order to quantify this performance loss, we simulated using the Clark-Gans time correlation model from [76]. In this model, Gaussian random variables are multiplied in the frequency domain by a spectral mask before being taken into the time domain via an inverse fast Fourier transform (IFFT). This approach can be easily generalized to allow for different doppler spreads, sampling periods, etc. We assumed that the system was using a carrier frequency of 2400 MHz.

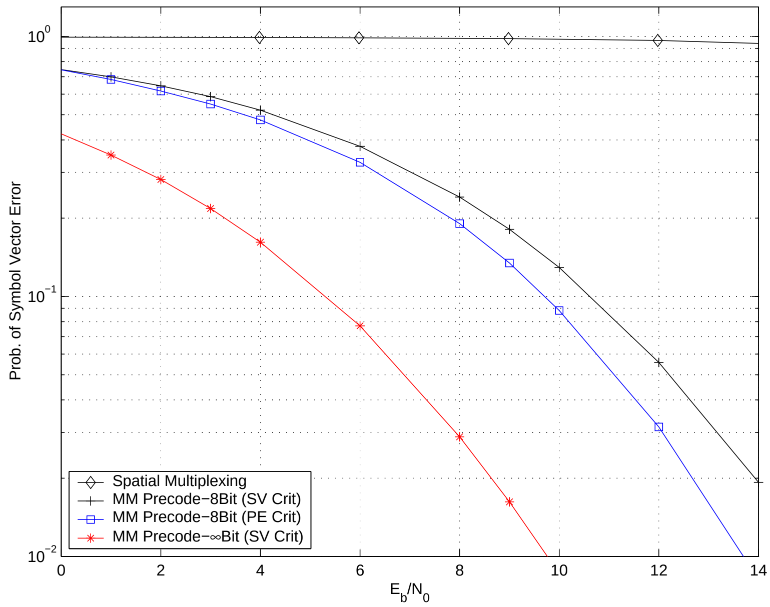

Figure 6.8: Multi-mode precoding tested on the "micro correlated" model in [113].

Experiment 1: The first experiment tests the performance of limited feedback precoded OSTBCs. A six bit precoded Alamouti code was tested on a 4×2 MIMO system transmitting 16-QAM. It was assumed that the environment had a maximum speed of 1m/s. Several different sampling rates were simulated and are shown in Fig. 6.9. The limited feedback precoding system still provides around a 1dB gain compared to antenna subset selection. This clearly presents the benefits of limited feedback in "real-world channels." The plot shows that an update period of less than 0.1ms is required. Using an update period of 0.2ms causes a 3.5dB performance loss.

Experiment 2: Six bit limited feedback precoded spatial multiplexing was simulated at a speed of 0.5m/s on a 4×4 system transmitting 4-QAM. The zero-delay case was simulated for comparison. The results are shown in

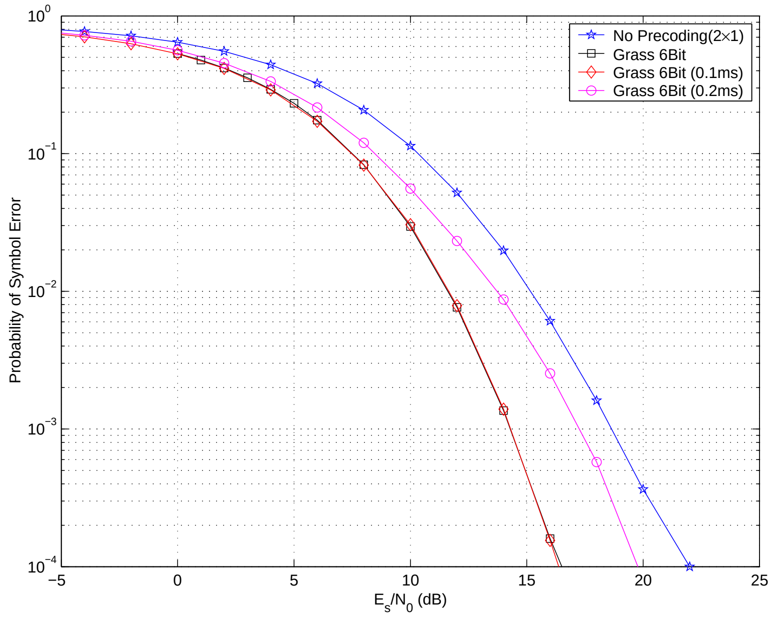

Figure 6.9: Precoded OSTBC travelling at 1m/s with various sampling rates.

Fig. 6.10. At this reasonable indoor speed, the update period can be as high as 10ms without causing any performance loss. Assuming a feedback channel is present that can carry data at 1.5kb/s, we can support up to 15 bits of feedback in this scenario.

Experiment 3: Fig. 6.11 shows the simulated performance of 4×4 multi-mode precoding with five feedback bits at a speed of 0.5m/s. It was assumed that eight bits were transmitted at each channel use. The zero-delay case was also simulated. Update periods of 6ms and 10ms are considered. At these update periods, the performance loss only seems to show up at high SNRs. This 10ms performance loss is approximately 0.5dB at an error rate of 10^{-3}.

Experiment 4: The system in Experiment 2 was simulated assuming a transmit SNR of 4dB. It was assumed that the mobile was traveling at

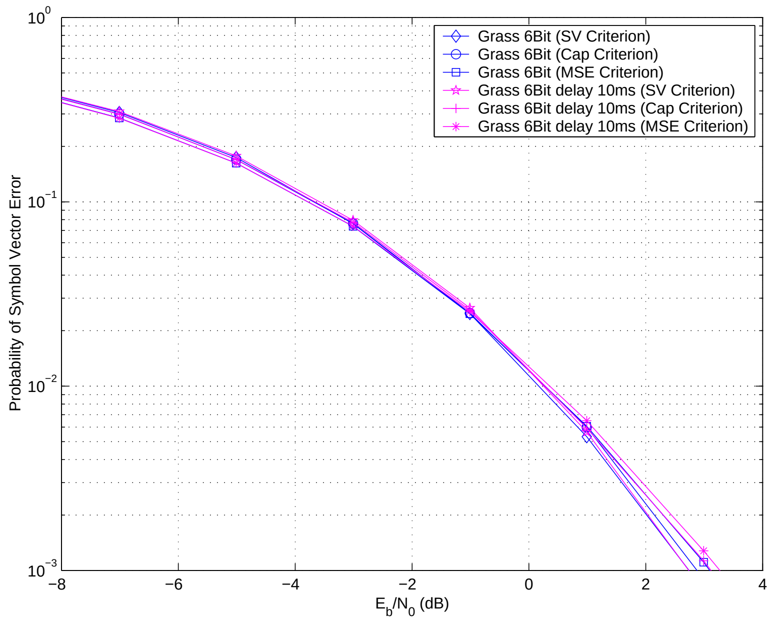

Figure 6.10: Precoded spatial multiplexing travelling at 0.5m/s with various sampling rates.

1m/s. Fig. 6.12 presents the results. Note that all three criteria start to perform worse than spatial multiplexing when the update period is larger than the coherence time. Intuitively, this makes sense because the coherence time of the channel is an approximate measure of the length of time before channel realizations become independent. The plot also indicates that in most scenarios an update period that is less than half the coherence time is needed.

6.3 Effect of Channel Estimation Error

An important consideration in MIMO wireless links is the estimation of the matrix channel $\mathbf{H}$. The limited feedback techniques developed in this dissertation require that the receiver have knowledge of the current channel conditions.

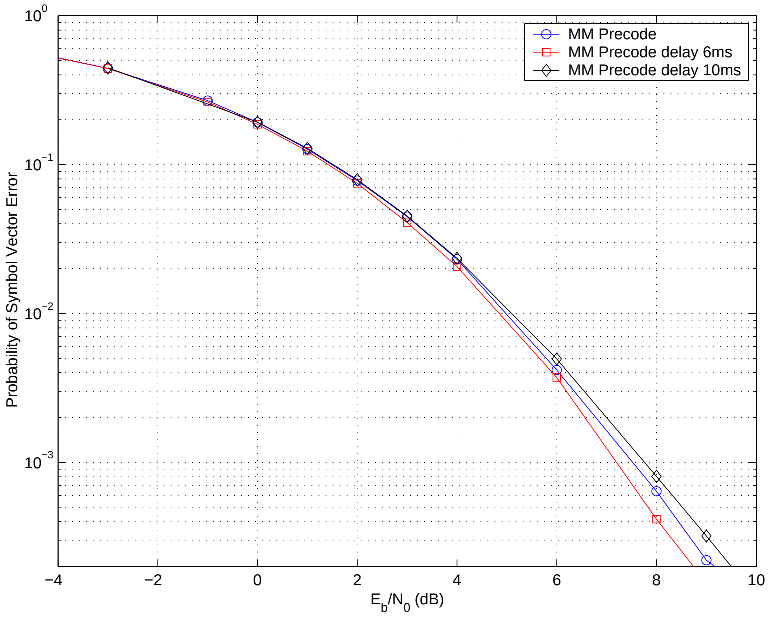

Figure 6.11: Multi-mode precoding travelling at 0.5m/s with various sampling rates.

This knowledge can be obtained by *training,* sending a signal known to both the transmitter and receiver and using the known transmission to estimate the channel.

For example, supposed the $M_t \times T$ matrix $\mathbf{S}$ was used for training. This would yield an input-output relationship of

$$\mathbf{Y} = \sqrt{\frac{\mathcal{E}_s}{M_t}}\mathbf{HS} + \mathbf{N} \tag{6.13}$$

where $tr\left(\mathbf{S}^*\mathbf{S}\right) \leq M_t T$ and the entries of $\mathbf{N}$ are i.i.d. $\mathcal{CN}(0, N_0)$. The ML channel estimate would be given by [121]

$$\widehat{\mathbf{H}}_{ML} = \sqrt{\frac{M_t}{\mathcal{E}_s}}\mathbf{YS}^*\left(\mathbf{SS}^*\right)^{-1}, \tag{6.14}$$

and the MMSE estimate is given by

$$\widehat{\mathbf{H}}_{MMSE} = \sqrt{\frac{M_t}{\mathcal{E}_s}}\mathbf{YS}^*\left(\mathbf{SS}^* + \frac{M_t N_0}{\mathcal{E}_s}\mathbf{I}_{M_t}\right)^{-1}. \tag{6.15}$$

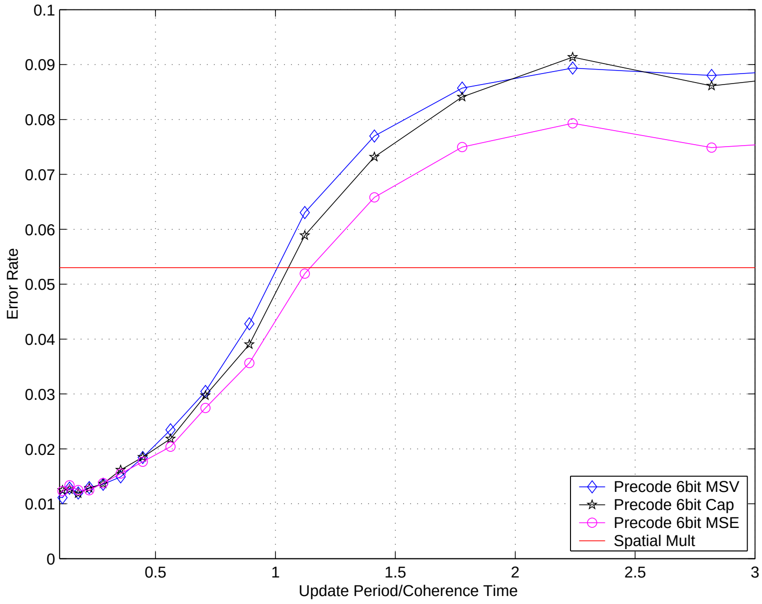

Figure 6.12: Precoded spatial multiplexing travelling at 1m/s with various sampling rates and a fixed SNR.

The expression in (6.14) and (6.15) tell us that $\mathbf{S}$ must have a rank greater than M_t. This means that $T \geq M_t$. A simple training signal would be to use T/M_t concatenated identity matrices,

$$\mathbf{S} = \left[\mathbf{I}_{M_t} \;\; \cdots \;\; \mathbf{I}_{M_t}\right]. \tag{6.16}$$

This would simplify (6.14) to

$$\widehat{\mathbf{H}}_{ML} = \frac{M_t}{T}\sqrt{\frac{M_t}{\mathcal{E}_s}}\sum_{k=1}^{T/M_t}\widehat{\mathbf{H}}_k \tag{6.17}$$

where $\widehat{\mathbf{H}}_k$ is the received signal obtained from transmission of the k^{th} identity matrix.

Experiment 1: Multi-mode precoding with five bits of feedback and eight bits per transmission was simulated using ML channel estimation. The identity matrix training signal in (6.16) was used with $T/M_t = 1,\ 4$, and 8. The receiver used the estimated channel for both precoder design and decoding. Note that estimation error can significantly affect the receiver performance. Even an eight step training incurs a 2dB difference. Notice that this plot does not present an error floor. This result is obtained because at high SNR we are getting a *better* channel estimate. The channel estimate power in (6.17) was scaled to agree with the system operating SNR.

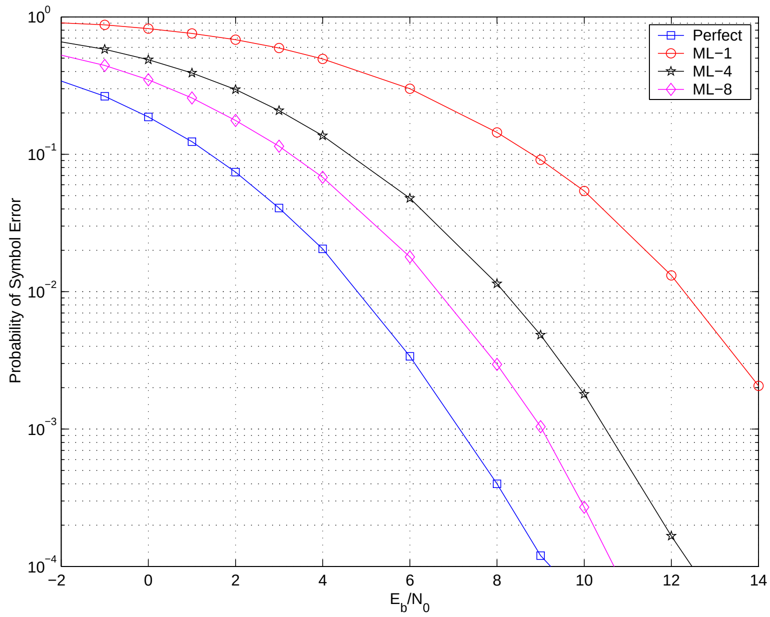

Figure 6.13: Multi-mode precoding performance with various estimation methods.

6.4 Feedback Error

Another issue in a real system is the absence of an error free feedback channel. Real systems will have feedback channels with non-negligible error rates. This follows because a full duplex wireless system will use a feedback channel that is transmitted over a possibly fading wireless link. The following experiment addresses this problem

Experiment 1: A 4×2 precoded Alamouti system transmitting 4-QAM was simulated with various feedback SNRs. It was assumed that the feedback channel transmitted with BPSK signaling. Note that a loss of approximately 3dB occurs when the feedback channel has an SNR of 0dB, and feedback at 5dB encounters around a 1dB loss. In contrast, 10dB feedback causes approximately no performance loss.

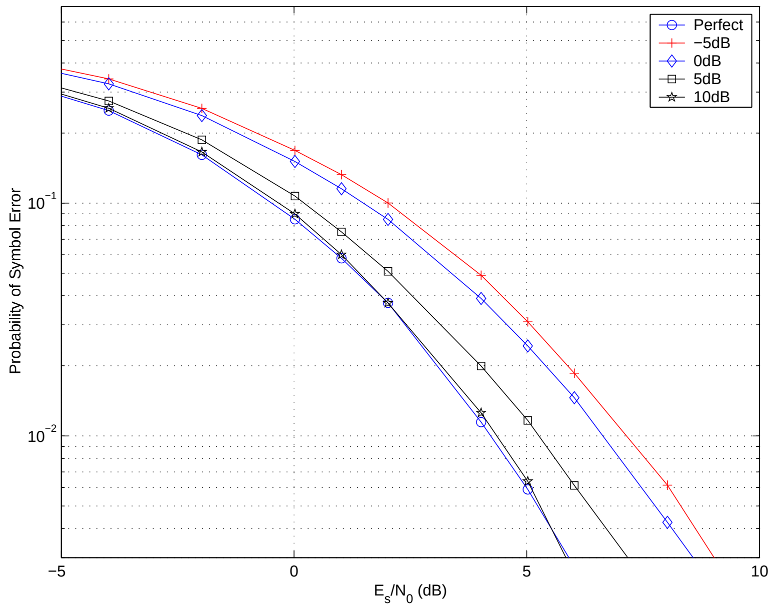

Figure 6.14: Precoded OSTBCing performance with various feedback SNRs.

6.5 Application to Broadband Communications

While the majority of the analysis in this dissertation has been for narrowband communications, the results are directly applicable to MIMO-OFDM systems. OFDM works by dividing a large bandwidth into small narrowbands. This is visually described in Fig. 6.15.

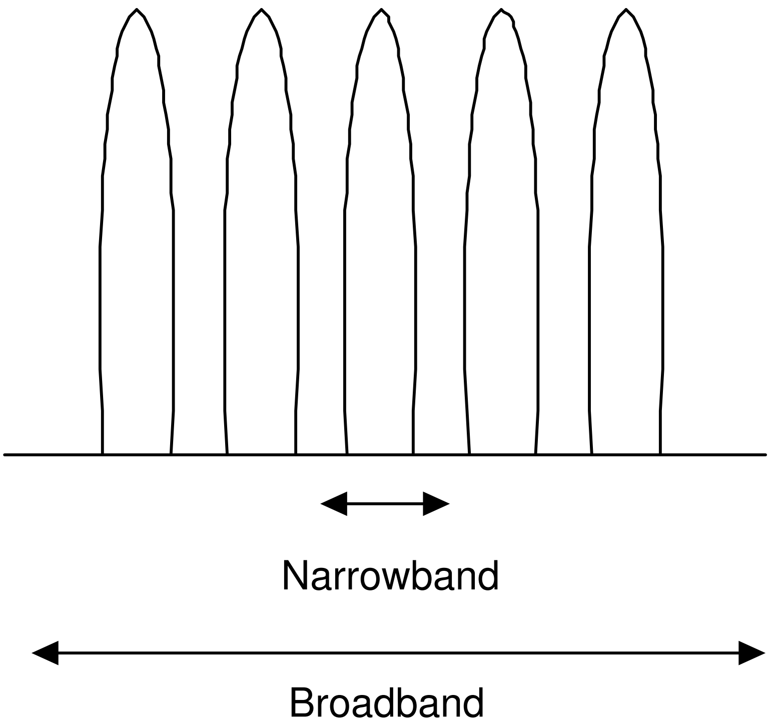

Figure 6.15: Illustration of OFDM frequency division.

Broadband channels are challenging to transmit over because they have multi-tap channels. This means that symbols will interfere with each other because of multipath in the wireless channel. OFDM systems are designed in frequency and transmitted in time using an IFFT. The receiver uses a fast Fourier transform (FFT) to convert the OFDM time symbol back to frequency. The key to OFDM is the use of a cyclic prefix that is appended to the OFDM time symbol before transmission and removed at the receiver before the FFT. This cyclic prefix converts the multi-tap channel from a traditional linear convolution to cyclic convolution. Thus the transmit and receive orthogonal transformations convert the multi-tap channel into tone-by-tone parallel narrowband channels in the frequency domain.

Limited feedback techniques can be done tone-by-tone over these parallel channels. The feedback can either be done for each tone or be transmitted back for pilot tones*. When the feedback is transmitted on pilots only, special interpolation techniques must be used to recover the adaptive signaling technique for each tone.

*Pilot tones are tones interspersed in an OFDM symbol that carry a known symbol instead of actual data. These tones can be used for channel estimation.

Chapter 7

Conclusions

This dissertation derived limited feedback closed-loop methods for MIMO wireless systems. The current state of research in the area of MIMO open-loop and closed-loop wireless systems is reviewed. Closed-loop methods require transmit channel knowledge, which is not feasible in many wireless systems. Limited feedback precoding remedies this deficiency by allowing for a limited number of bits to be conveyed from the receiver to transmitter over a feedback channel.

7.1 Summary

After a review of space-time signaling in Chapter 2, limited feedback MIMO signaling techniques are proposed in Chapters 3-5. Limited feedback precoding for spatial multiplexing and OSTBCs requires that the receiver choose a precoding matrix from a codebook of possible precoding matrices known to both the transmitter and receiver. The receiver can then convey this matrix to the transmitter using a limited number of bits. The codebook is assumed to be designed offline and fixed for all time.

Chapter 3 discussed limited feedback precoding for OSTBCs. It is found that the codebook should be thought of as a subspace code using the chordal distance function on the Grassmann manifold. Limited feedback precoding for spatial multiplexing is proposed in Chapter 4. Distortion functions are derived as a function of the precoder selection function. It is shown that the codebook should once again be thought of as a subspace code using either the projection two-norm distance or the Fubini-Study distance.

While precoding assumes a fixed number of substreams, both rate and reliability can be further improved by allowing the number of substreams to be varied based on current channel conditions. Chapter 5 proposed dual-mode and multi-mode precoding systems as a precoding approach that can make tradeoffs between diversity gain and multiplexing gain. Multi-mode precoding can be understood as precoded spatial multiplexing with a variable number of substream.

The effect of practical considerations such as spatial correlation, feedback delay, estimation error, and feedback error are discussed. In particular, it is shown that limited feedback techniques are very resilient to these problems. Necessary and sufficient conditions for full diversity order in a spatially correlated system are derived for the amount of feedback that must be allotted.

7.2 Future Work

This dissertation represents only an initial foray into the uses of limited feedback in MIMO systems. The following areas of research still need to be addressed.

Multiuser MIMO: The systems considered in this dissertation were point-to-point or single-user. Multiuser techniques using feedback have recently been discussed in [122], [123]. It is of both practical and theoretical interest to find methods to integrate limited feedback into these algorithms.

Limited Feedback in MIMO-OFDM: Orthogonal frequency division multiplexing (OFDM) is a popular technique for signaling on wideband channels. OFDM works by dividing a large bandwidth into many different narrowband channels. MIMO systems using OFDM, called MIMO-OFDM, are expected to the prevalent architecture in next generation wireless. Employing limited feedback in MIMO-OFDM is challenging because the system can be viewed as multiple narrowband limited feedback MIMO systems operating in parallel. This creates many different problems in feedback design.

Information Theoretic Effects: Questions such as, "What is the capacity when B bits of feedback are allowed?" still have yet to be answered. These kinds of issues must be addressed for researchers to understand the fundamental limits of feedback.

These topics represent a very small sample of the vast open feedback issues. Limited feedback MIMO is still an open and deep area of research. This dissertation answers many of the fundamental issues in precoded MIMO with limited feedback.

Bibliography

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm.*, vol. 10, no. 6, pp. 585–595, 1999.

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

[3] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, pp. 41–59, Autumn 1996.

[4] W. C. Jakes, *Microwave Mobile Communications.* New York: John Wiley and Sons, 1974.

[5] D. G. Brennan, "Linear diversity combining techniques," *Proc. IRE*, vol. 47, pp. 1075–1102, June 1959.

[6] A. Wittneben, "Analysis and comparison of optimal predictive transmitter selection and combining diversity for DECT," in *Proc. IEEE Glob. Telecom. Conf.*, vol. 2, pp. 1527–1531, Nov. 1995.

[7] A. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast

systems using distributed transmission/directional reception (DTDR)." U. S. patent #5345599, Sept. 1994.

[8] I. E. Telatar, "Capacity of multi-antenna Gaussian channels." tech. rep., Lucent Technologies Bell Laboratories Internal Technical Memorandum, Oct. 1995.

[9] IEEE-SA Standards Board, "IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems," IEEE Std 802.16.-2001, Dec. 2001.

[10] H. Holma and A. Toskala, eds., *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications, Revised Edition*. John Wiley & Sons, 1990.

[11] J. Heiskala and J. Terry, *OFDM Wireless LANs: A Theoretical and Practical Guide*. Indianapolis: SAMS, 2002.

[12] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Info. Th.*, vol. 44, pp. 744–765, March 1998.

[13] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Jour. Select. Areas in Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[14] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Info. Th.*, vol. 45, pp. 1456–1467, July 1999.

[15] G. Bauch and J. Hagenauer, "Smart versus dumb antennas-capacities and FEC performance," *IEEE Commun. Lett.*, vol. 6, pp. 55–57, Feb. 2002.

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

[17] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications.* New York: Cambridge University Press, 2003.

[18] N. R. Sollenberger, "Diversity and automatic link transfer for a TDMA wireless access link," in *Proc. IEEE Glob. Telecom. Conf.*, vol. 1, pp. 532–536, Nov.-Dec. 1993.

[19] S. Thoen, L. Van der Perre, B. Gyselinckx, and M. Engels, "Performance analysis of combined transmit-SC/receive-MRC," *IEEE Trans. Commun.*, vol. 49, pp. 5–8, Jan. 2001.

[20] C.-H. Tse, K.-W. Yip, and T.-S. Ng, "Performance tradeoffs between maximum ratio transmission and switched-transmit diversity," in *Proc. 11th IEEE Int. Symp. on Personal, Indoor and Mobile Radio Comm.*, vol. 2, pp. 1485–1489, Sept. 2000.

[21] T. K. Y. Lo, "Maximum ratio transmission," *IEEE Trans. Commun.*, vol. 47, pp. 1458–1461, Oct. 1999.

[22] D. J. Love and R. W. Heath, Jr., "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. Commun.*, vol. 51, pp. 1102–1110, July 2003.

[23] P. A. Dighe, R. K. Mallik, and S. S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *IEEE Trans. Commun.*, vol. 51, pp. 694–703, April 2003.

[24] M. Kang and M.-S. Alouini, "Largest eigenvalue of complex wishart matrices and performance analysis of MIMO MRC systems," *IEEE Jour. Select. Areas in Commun.*, vol. 21, pp. 418–426, April 2003.

[25] M. Kang and M.-S. Alouini, "Performance analysis of MIMO MRC systems over Rician fading channels," in *Proc. IEEE Veh. Technol. Conf.*, vol. 2, pp. 869–873, Sept. 2002.

[26] J. Bach Anderson, "Antenna arrays in mobile communications: gain, diversity, and channel capacity," *IEEE Antennas and Propagation Magazine*, vol. 42, pp. 12–16, April 2000.

[27] D. J. Love and R. W. Heath, Jr., "Equal gain transmission in multiple-input multiple-output wireless systems," in *Proc. IEEE Glob. Telecom. Conf.*, vol. 2, pp. 1124–1128, Nov. 2002.

[28] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannaks, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Sig. Proc.*, vol. 50, pp. 1051–1064, May 2002.

[29] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Commun. Lett.*, vol. 6, pp. 239–241, June 2002.

[30] E. Visotsky and U. Madhow, "Space-time transmit precoding with im-

perfect feedback," *IEEE Trans. Info. Th.*, vol. 47, pp. 2632–2639, Sept. 2001.

[31] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Jour. Select. Areas in Commun.*, vol. 21, pp. 684–702, June 2003.

[32] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Sig. Proc.*, vol. 50, pp. 2599–2613, Oct. 2002.

[33] Z. Hong, K. Liu, R. W. Heath, Jr., and A. M. Sayeed, "Spatial multiplexing in correlated fading via the virtual channel representation," *IEEE Jour. Select. Areas in Commun.*, vol. 21, pp. 856–866, June 2003.

[34] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE Int. Conf. on Commun.*, vol. 7, pp. 11–14, June 2001.

[35] R. U. Nabar, H. Bolcskei, V. Erceg, D. Gesbert, and A. J. Paulraj, "Performance of multiantenna signaling techniques in the presence of polarization diversity," *IEEE Trans. Sig. Proc.*, vol. 50, pp. 2553–2562, Oct. 2002.

[36] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Info. Th.*, vol. 49, pp. 1073–1096, May 2003.

[37] M. K. Simon and M. S. Alouini, *Digital Communications over Fading Channels*. New York: John Wiley and Sons, 2000.

[38] D. J. Love and R. W. Heath, Jr., "Necessary and sufficient conditions for full diversity order in correlated Rayleigh fading beamforming and combining systems," *IEEE Trans. Wireless Comm.*, Sept. 2003. Accepted.

[39] R. W. Heath Jr. and A. Paulraj, "Switching between spatial multiplexing and transmit diversity based on constellation distance," in *Proc. of Allerton Conf. on Comm. Cont. and Comp.*, Oct. 2000.

[40] R. W. Heath, Jr. and D. J. Love, "Multi-mode antenna selection for spatial multiplexing systems with linear receivers." submitted to *IEEE Trans. Sig. Proc.*, Nov. 2003.

[41] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, Mar. 1998.

[42] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Info. Th.*, vol. 48, pp. 611–627, March 2002.

[43] E. G. Larsson, G. Ganesan, P. Stoica, and W.-H. Wong, "On the performance of orthogonal space-time block coding with quantized feedback," *IEEE Commun. Lett.*, vol. 6, pp. 487–489, Nov. 2002.

[44] G. Jöngren and M. Skoglund, "Improving orthogonal space-time block codes by utilizing quantized feedback information," in *Proc. IEEE Int. Symp. Info. Th.*, p. 220, June 2001.

[45] D. A. Gore and A. J. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Trans. Sig. Proc.*, vol. 50, pp. 2580–2588, Oct. 2002.

[46] G. Jöngren, M. Skoglund, and B. Ottersten, "Utilizing partial channel information in the design of space-time block codes," in *Proc. Int. Symp. Wireless Pres. Mult. Commun.*, vol. 2, pp. 681–685, Oct. 2002.

[47] W.-H. Wong and E. G. Larsson, "Orthogonal space-time block coding with antenna selection and power allocation," *Elect. Lett.*, vol. 39, pp. 379–381, Feb. 2003.

[48] G. Ganesan, P. Stoica, and E. G. Larsson, "Diagonally weighted orthogonal space-time block codes," in *Proc. 36th Asilomar Conf. on Sig., Sys., and Comp.*, vol. 2, pp. 1147–1151, Nov. 2002.

[49] T. Xiaofeng, H. Harald, Y. Zhuizhuan, Q. Haiyan, and Z. Ping;, "Closed loop space-time block code," in *Proc. IEEE Veh. Technol. Conf.*, vol. 2, pp. 1093–1096, Oct. 2001.

[50] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Info. Th.*, vol. 49, pp. 1673–1690, July 2003.

[51] G. Jöngren, M. Skoglund, and B. Ottersten, "Design of channel estimate dependent space-time block codes." submitted to *IEEE Trans. Commun.*, June 2001.

[52] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in mulitiple-antenna data transmission over fading

channels," *IEEE Jour. Select. Areas in Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.

[53] R. W. Heath Jr. and A. Paulraj, "A simple scheme for transmit diversity using partial channel feedback," in *Proc. of the 32nd Annual Asil. Conf. on Sig. Sys. and Comp.*, vol. 2, pp. 1073–1078, Nov. 1998.

[54] K. K. Mukkavilli, A. Sabharwal, M. Orchard, and B. Aazhang, "Space-time coding with feedback," in *Proc. Int. Symp. Telecom.*, Sept. 2001.

[55] K. K. Mukkavilli, A. Sabharwal, and B. Aazhang, "Design of multiple antenna coding schemes with channel feedback," in *Proc. of the 35th Annual Asil. Conf. on Sig. Sys. and Comp.*, pp. 1009–1013, Nov. 2001.

[56] D. J. Love, R. W. Heath, Jr., and T. Strohmer, "Quantized antenna weighting for multiple-input multiple-output wireless systems," in *Proc. of Allerton Conf. on Comm. Cont. and Comp.*, Oct. 2002.

[57] D. J. Love, R. W. Heath, Jr., and T. Strohmer, "Quantized maximum ratio transmission codebook design for multiple-input multiple-output wireless systems," in *Proc. 36th Asilomar Conf. on Sig., Sys., and Comp.*, vol. 1, pp. 531–535, Nov. 2002.

[58] D. J. Love, R. W. Heath, Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," in *Proc. IEEE Int. Conf. on Commun.*, vol. 4, pp. 2618–2622, May 2003.

[59] D. J. Love, R. W. Heath Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Info. Th.*, vol. 49, pp. 2735–2747, Oct. 2003.

[60] K. K. Mukkavilli, A. Sabharwal, and B. Aazhang, "Finite rate feedback design for multiple transmit antennas," in *Proc. of Allerton Conf. on Comm. Cont. and Comp.*, Oct. 2002.

[61] K. K. Mukkavilli, A. Sabharwal, B. Aazhang, and E. Erkip, "Performance limits on beamforming with finite rate feedback for multiple antenna systems," in *Proc. of the 36th Annual Asil. Conf. on Sig. Sys. and Comp.*, vol. 1, pp. 536–540, Nov. 2002.

[62] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "Beamformer design with feedback rate constraints: criteria and constructions," in *Proc. IEEE Int. Symp. Info. Th.*, July 2003.

[63] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple antenna systems," *IEEE Trans. Info. Th.*, vol. 49, pp. 2562–2579, Oct. 2003.

[64] J. H. Wilkinson, *The Algebraic Eigenvalue Problem.* Oxford: Clarendon Press, 1988.

[65] R. S. Blum and J. H. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, pp. 322–324, Aug. 2002.

[66] M. Gharavi-Alkhansari and A. B. Gershman, "Fast antenna subset selection in wireless MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 5, pp. 57–60, April 2003.

[67] D. A. Gore, R. W. Heath, Jr., and A. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, pp. 491–493, Nov. 2002.

[68] D. Gore, R. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 5, pp. 2785–2788, June 2000.

[69] A. Gorokhov, "Antenna selection algorithms for MEA transmission systems," in *Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 3, pp. 2857–2860, May 2002.

[70] A. F. Molisch, M. Z. Win, and J. H. Winters, "Capacity of MIMO systems with antenna selection," in *Proc. IEEE Int. Conf. on Commun.*, vol. 2, pp. 570–574, June 2001.

[71] R. W. Heath Jr., S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, pp. 142–144, April 2001.

[72] D. J. Love and R. W. Heath, Jr., "Limited feedback precoding for spatial multiplexing systems using linear receivers," in *Proc. IEEE Mil. Comm. Conf.*, Oct. 2003.

[73] D. J. Love and R. W. Heath, Jr., "Limited feedback precoding for spatial multiplexing systems," in *Proc. IEEE Glob. Telecom. Conf.*, vol. 4, pp. 1857–1861, Dec. 2003.

[74] B. Varadarajan and J. R. Barry, "The rate-diversity trade-off for linear space-time codes," in *Proc. IEEE Veh. Technol. Conf.*, vol. 1, pp. 67 – 71, 2002.

[75] M. Godavarti and A. O. Hero, "Diversity and degrees of freedom in

wireless communications," in *Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 3, pp. 2861–2854, May 2002.

[76] T. S. Rappaport, *Wireless Communications.* Upper Saddle River, NJ: Prentice Hall, second ed., 2002.

[77] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *Elec. Lett.*, vol. 35, pp. 14–15, Jan. 1999.

[78] H. Sampath, P. Stoica, and A. Paulraj, "Generalized precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, pp. 2198–2206, Dec. 2001.

[79] R. W. Heath Jr. and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. IEEE Int. Conf. on Commun.*, vol. 7, pp. 2276–2280, June 2001.

[80] S. Sandhu, R. U. Nabar, D. A. Gore, and A. Paulraj, "Near-optimal selection of transmit antennas for a MIMO channel based on Shannon capacity," in *Proc. 34th Asilomar Conf. on Sig., Sys., and Comp.*, vol. 1, Nov. 2000.

[81] B. Hassibi and B. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Info. Th.*, vol. 48, pp. 1804–1824, July 2002.

[82] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston: Kluwer, 1992.

[83] R. A. Horn and C. R. Johnson, *Matrix Analysis.* New York: Cambridge University Press, March 2001.

[84] A. Barg and D. Y. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Info. Th.*, vol. 48, pp. 2450–2454, Sept. 2002.

[85] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Annals of Mathematical Statistics*, vol. 35, pp. 475–501, June 1964.

[86] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Info. Th.*, vol. 45, pp. 139–157, Jan. 1999.

[87] D. Agrawal, T. J. Richardson, and R. L. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Info. Th.*, vol. 47, pp. 2618–2626, Sept. 2001.

[88] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Mathematics*, vol. 5, pp. 139–159, 1996.

[89] T. Strohmer and R. W. Heath, Jr., "Grassmannian frames with applications to coding and communications," *Appl. Comput. Harmonic. Analysis*, vol. 14, pp. 257–275, May 2003.

[90] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Info. Th.*, vol. 48, pp. 359–383, Feb. 2002.

[91] N. J. A. Sloane, "Tables of real Grassmannian packings." available at *http : //www.research.att.com/ njas/*, 2003.

[92] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Info. Th.*, vol. 46, pp. 1962–1973, Sept. 2000.

[93] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: John Wiley & Sons, 1992.

[94] S. Talwar and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array," *IEEE Trans. Sig. Proc.*, vol. 45, pp. 706–718, March 1997.

[95] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthonormality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.

[96] J. Cioffi, "Digital transmission: Volume I." Stanford University EE 379A: www.stanford.edu/class/ee379a/.

[97] R. W. Heath Jr. and D. J. Love, "Multi-mode antenna selection for spatial multiplexing," in *Proc. of Allerton Conf. on Comm. Cont. and Comp.*, Oct. 2003.

[98] R. W. Heath Jr. and D. J. Love, "Dual-mode antenna selection for spatial multiplexing systems with linear receivers," in *Proc. 37th Asilomar Conf. on Sig., Sys., and Comp.*, Nov. 2003.

[99] R. Nabar, D. Gore, and A. Paulraj, "Selection and use of optimal transmit antennas in wireless systems," in *Proc. ICT*, May 2000.

[100] R. S. Blum, "MIMO with limited feedback of channel state information," in *Proc. IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 4, pp. 89–92, April 2003.

[101] D. J. Love and R. W. Heath Jr., "Grassmannian precoding for spatial multiplexing systems," in *Proc. of Allerton Conf. on Comm. Cont. and Comp.*, Oct. 2003.

[102] D. J. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for spatial multiplexing systems." submitted to *IEEE Trans. Info. Th.*, July 2003.

[103] W. Santipach, Y. Sun, and M. L. Honig, "Benefits of limited feedback for wireless channels," in *Proc. of Allerton Conf. on Comm. Cont. and Comp.*, Oct. 2003.

[104] 3GPP standard contribution, "Spatial Channel Model Text Description," SCM-134, April 22, 2003.

[105] G. D. Durgin, *Space-Time Wireless Channels.* Upper Saddle River, NJ: Prentice Hall, 2003.

[106] A. M. Rao and D. L. Jones, "Efficient detection with arrays in the presence of angular spreading," *IEEE Trans. Sig. Proc.*, vol. 51, pp. 301–312, Feb. 2003.

[107] L. Schumacher, K. I. Pedersen, and P. E. Mogensen, "From antenna spacings to theoretical capacities - guidelines for simulating MIMO systems," in *Proc. IEEE Int. Symp. Personal, Indoor, and Mobile Radio Commun.*, vol. 2, pp. 587–592, Sept. 2002.

[108] M. Stege, J. Jelitto, M. Bronzel, and G. Fettweis, "A multiple input-multiple output channel model for simulation of TX-and RX-diversity wireless systems," in *Proc. IEEE Veh. Technol. Conf.*, vol. 2, pp. 833–839, Sept. 2000.

[109] M. Stege, M. Bronzel, and G. Fettweis, "MIMO-capacities for COST 259 scenarios," in *Proc. of the International Zurich Seminar*, pp. 29/1 – 29/6, Feb. 2002.

[110] J. R. Fonollosa, R. Gaspa, X. Mestre, A. Pages, M. Heikkila, J. P. Kermoal, L. Schumacher, A. Pollard, and J. Ylitalo, "The IST METRA project," *IEEE Comm. Mag.*, vol. 40, pp. 78–86, July 2002.

[111] V. Erceg, L. Schumacher, P. Kyritsi, A. Molisch, D. S. Baum, A. Y. Gorokhov, C. Oestges, V. J. Rhodes, J. Medbo, D. Michelson, M. Webster, E. Jacobsen, D. Cheung, Q. Li, C. Prettie, M. Ho, K. Yu, L. Jengx, A. Jagannatham, N. Tal, and C. Lanzl, "Indoor MIMO WLAN channel models." IEEE 802.11-03/161r2, Sept. 2003.

[112] A. Forenza, D. J. Love, and R. W. Heath, Jr., "A low complexity algorithm to simulate the spatial covariance matrix for clustered MIMO channel models," in *Proc. IEEE Veh. Technol. Conf.*, May. 2004.

[113] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Jour. Select. Areas in Commun.*, vol. 20, pp. 1211–1226, Aug. 2002.

[114] L. Schumacher, J. P. Kermoal, F. Frederiksen, K. I. Pedersen, A. Algans,

and P. E. Mogensen, "MIMO channel characterisation." tech. rept., European IST-1999-11729 Project METRA, Feb. 2001.

[115] Y. Tan, M. Pereira, M. Mewburn, and M. Faulkner, "Investigation of singular value distributions of MIMO channels in indoor environment." Available at *http://www.telecommunications.crc.org.au/content/ConfPapers/Tan.pdf.*

[116] J. W. Wallace and M. A. Jensen, "Modeling the indoor MIMO wireless channel," *IEEE Trans. Antennas Propagat.*, vol. 50, pp. 591–599, May 2002.

[117] J. P. Kermoal, L. Schumacher, F. Frederiksen, and P. E. Mogensen, "Polarization diversity in MIMO radio channels: experimental validation of a stochastic model and performance assessment," in *Proc. IEEE Veh. Technol. Conf.*, vol. 1, pp. 22–26, Fall 2001.

[118] J. P. Kermoal, L. Schumacher, P. E. Mogensen, and K. I. Pedersen, "Experimental investigation of correlation properties of MIMO radio channels for indoor picocell scenarios," in *Proc. IEEE Veh. Technol. Conf.*, vol. 1, pp. 14–21, Sept. 2000.

[119] I. Sohn, G. D. Golden, H. Lee, and J. W. Ahn, "Overview of METRA model for MBWA MIMO channel." IEEE C802.20-03-50, May. 2003.

[120] G. H. Golub and C. F. Van Loan, *Matrix Computations.* Baltimore: Johns Hopkins University Press, third ed., 1996.

[121] B. Hassibi and B. M. Hochwald, "How much training is needed in

multiple-antenna wireless links?," *IEEE Trans. Info. Th.*, vol. 49, pp. 951–963, April 2003.

[122] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels." Submitted to *IEEE Trans. Info. Th.* in Aug. 2002.

[123] S. Vishwanath, N. Jindal, and A. Goldsmith, "On the capacity of multiple-input multiple-output broadcast channels," in *Proc. IEEE Int. Conf. on Commun.*, vol. 3, pp. 1444–1450, April-May 2002.

Vita

David James Love was born in Garland, Texas. He received the B.S. (highest honors), M.S.E., and Ph.D. degrees in electrical engineering from The University of Texas at Austin in 2000, 2002, and 2004, respectively.

During the summers of 2000 and 2002, he was affiliated with the Texas Instruments DSPS R&D Center in Dallas, Texas. At Texas Instruments, he performed research on physical layer system design for next generation wireless systems employing multiple antennas. In the fall of 2004, he will join the School of Electrical and Computer Engineering at Purdue University, West Lafayette, IN, as an Assistant Professor. His current research interests are in the design and analysis of wireless systems and the theory of codes based on subspace packings.

Mr. Love is a member of the IEEE, SIAM, Tau Beta Pi and Eta Kappa Nu. He was the recipient of the 2003 IEEE Vehicular Technology Society Daniel E. Noble Fellowship. He is married to Elaina Love.

Permanent Address: 718 Royal Birkdale
Garland, Texas 75044

This dissertation was typeset with $\LaTeX\,2_{\varepsilon}$[*] by the author.

[*] $\LaTeX\,2_{\varepsilon}$ is an extension of $\LaTeX$. $\LaTeX$ is a collection of macros for $\TeX$. $\TeX$ is a trademark of the American Mathematical Society. The macros used in formatting this dissertation were written by Dinesh Das, Department of Computer Sciences, The University of Texas at Austin, and extended by Bert Kay and James A. Bednar.