Construction of High-Rate, Reliable Space-Time Codes

A Thesis
Submitted For the Degree of
Master of Science (Engineering)
in the Faculty of Engineering

by

Raj Kumar K.

Department of Electrical Communication Engineering
Indian Institute of Science
BANGALORE – 560 012
JUNE 2005
©Raj Kumar K.
JUNE 2005
All rights reserved
Acknowledgements

I would like to start out by thanking my parents for their unbounded support and affection, without which all this would not have been possible.

Working with Prof. P. Vijay Kumar over the past two years in IISc has been a most rewarding experience, both academically and otherwise. It has been both a privilege and a pleasure to have been associated with him.

During my stint in IISc, I have had the opportunity to collaborate (in person and through e-mail) with the following people, to whom I am grateful: Sameer Pawar (colleague, IISc), Petros Elia (visiting research scholar from the Univ. of Southern California, Los Angeles), Hsiao-feng Lu (Assistant Professor, National Chung-Cheng University, Taiwan) and B. A. Sethuraman (Professor, California State University, Northridge).

I have also been fortunate to have found several good acquaintances in IISc, a few of whom I will pause to mention and thank:

- Sameer, Petros and Anand - for all the good times spent both in the CSD lab and outside.
- Kaushik, my first semester roomie - I really enjoyed our stay in the quaint old H-12 room and the umpteen number of evenings spent in the IISc Gymkhana.
- The “A-Mess gang” - Yogesh, Srin, Govind, Raghavan, Prakash, Manikandan, Vishal and Rajesh - You guys made each and every meal enjoyable, irrespective of the menu!
- Azad and Virendra - for the excellent company provided in the ECE dept. and in the hostel.

K. Raj Kumar
Bangalore, June 2005.
Publications based on this Thesis

• Journal Submissions


• Conference Publications (In Chronological Order)


3. K. Raj Kumar, Sameer A. Pawar, Petros Elia, P. Vijay Kumar and Hsiao-feng Lu, “Two Generalizations of the Rank-Distance Space-Time Code”, Proc. of
the National Conference on Communications (NCC-2005), Kharagpur, India, Jan. 2005.


Abstract

The paradigm of “space-time coding”, introduced in 1998, involves the use of multiple antennas at both the transmitter and receiver for communication across a wireless channel. The advantages of using space-time codes over the wireless channel are two-fold: with intelligent signal design, better reliability and higher rates can be achieved, although there is a tradeoff between the two.

Two formulations which quantify the best possible reliability and rate that a space-time code can simultaneously achieve have been recently proposed in the literature: the rate-diversity tradeoff and the diversity-multiplexing gain (D-MG) tradeoff. The rate-diversity tradeoff applies in the context of a system with a fixed constellation (i.e., one that does not increase with signal to noise ratio (SNR)) and is based on pairwise error probability. The D-MG tradeoff however, is based on codeword error probability and employs an information theoretic formulation. The formulation assumes that the constellation grows with SNR. Further, the D-MG tradeoff is defined under the asymptotic of very high SNR.

In the initial part of this thesis, we will present an explicit construction of minimal-delay (square) space-time codes that achieve the D-MG tradeoff for arbitrary number of transmit and receive antennas. The tradeoff optimality of this construction is concluded from a recently proposed sufficient condition, which is part of a joint work that does not appear in this thesis. The sufficient condition, in conjunction with the constructions presented in this thesis, constitutes the first solution to the problem of providing explicit constructions of space-time codes that achieve the D-MG tradeoff for arbitrary number of transmit and receive antennas. The constructions presented in this thesis are based on
cyclic division algebras, which may be thought of as generalizations of the quaternions. Some work related to reducing the signalling complexity of these constructions is also presented. Following this, generalizations of the constructed square space-time codes to the rectangular case are dealt with.

The final part of our thesis will focus on providing a generalization to the rank-distance space-time (RDST) codes introduced recently by Lu and Kumar as a family of space-time codes that meet the rate-diversity tradeoff. The generalization proposed, called the variable-rank RDST code, is shown to have a better D-MG tradeoff than the RDST construction. The RDST codes and their generalization employ commonly used constellations such as QAM and PSK as their signalling alphabet, thus making them attractive for practical implementations.
Contents

Acknowledgements ii
Publications based on this Thesis iii
Abstract v

1 Introduction 1

2 MIMO systems and Space-Time Coding 5
  2.1 The Space-Time Channel ........................................... 6
  2.2 Pairwise Error Probability ........................................... 9
  2.3 Capacity of the Space-Time Channel ................................. 12
    2.3.1 Evaluation of the Capacity .................................... 14
    2.3.2 Non-Ergodic Channels ........................................... 15
  2.4 Space-Time Block Codes and Examples ............................... 16
    2.4.1 Orthogonal Design Based Space-Time Block Codes ............... 17
    2.4.2 V-BLAST ...................................................... 19

3 The Diversity-Multiplexing Gain Tradeoff 21
  3.1 Preliminaries ..................................................... 22
  3.2 Outage and Error Probability ..................................... 24
  3.3 The D-MG Tradeoff ................................................ 26
  3.4 D-MG Tradeoff of a Few Schemes from the Literature .............. 28

4 Achieving the D-MG Tradeoff 35
  4.1 Preliminaries ..................................................... 36
  4.2 The Sufficient Condition .......................................... 36
  4.3 The Non-Vanishing Determinant Property ........................... 37
  4.4 D-MG Optimality and Non-Vanishing Determinants .................. 38
    4.4.1 Signal Alphabet .............................................. 39
    4.4.2 Signal Energy ................................................ 40
    4.4.3 D-MG Optimality .............................................. 40
  4.5 D-MG Optimal Schemes from the Literature .......................... 42
    4.5.1 Tilted QAM codes of Yao and Wornell .......................... 42
    4.5.2 Dayal-Varanasi Construction .................................. 44
CONTENTS

4.5.3 Golden Code Construction ........................................... 44
4.5.4 LAST Codes ......................................................... 45
4.5.5 Cyclotomic Space-Time Codes of Wang and Xia .................... 46
4.5.6 Space-time Codes Derived from Division Algebras ................. 46

5 Review of Relevant Algebra and Number Theory 47
5.1 Brief Review of Groups, Rings and Fields .......................... 48
5.2 Basics of Algebraic Number Theory ................................ 50
5.2.1 Number Fields and Extensions ................................... 50
5.2.2 Algebraic Integers and Integral Bases .......................... 51
5.2.3 Ideals in an Algebraic Number Field ............................ 52
5.2.4 Properties of Ideals ............................................. 53
5.2.5 Factorization of Ideals ......................................... 54

6 Space-Time Codes from Division Algebras - An Overview 57
6.1 Division Algebras and Space-Time Codes .......................... 57
6.2 Structure of Division Algebras ..................................... 61
6.3 Space-Time Codes Arising from Cyclic Division Algebras ........... 62
6.4 Properties of ST codes Derived from CDA .......................... 64
6.5 The Non-vanishing Determinant Property ........................... 66
6.6 Summary of ST code construction from CDA with NVD .............. 68

7 Minimal Delay D-MG Optimal Space-Time Block Codes 69
7.1 CDA Based ST Codes with NVD are D-MG Optimal ................. 70
7.2 Previous Constructions of CDA Based ST Codes with NVD .......... 70
7.3 Cyclic Galois extensions ............................................ 71
7.3.1 Cyclic Galois extensions over \( \mathbb{Q} \) .......................... 71
7.3.2 Cyclic Galois Extensions over \( \mathbb{F} \) ......................... 74
7.4 Determining a Non-Norm Element .................................. 75
7.5 Reducing the Signalling Complexity ................................ 79
7.5.1 Cyclic Galois Extensions ....................................... 80
7.5.2 Non-Norm Elements ........................................... 82

8 Rectangular D-MG Optimal Space-Time Block Codes 85
8.1 Intuition Behind the Constructions ................................ 85
8.2 The Puncturing Construction ...................................... 86
8.3 The Stacking Construction ...................................... 88

9 Generalizations of the RDST Code ................................ 90
9.1 The Rate-Diversity Tradeoff .................................... 91
9.2 Rate-Diversity Tradeoff of Some Prior Constructions ............... 94
9.3 Comparing the Rate-Diversity and D-MG Tradeoffs ................. 97
9.4 A Unified Construction of RDST Codes ............................ 98
9.5 D-MG Tradeoff of the RDST code ................................. 102
## CONTENTS

9.6 The Variable-Rank RDST code ............................................. 104
9.7 D-MG Tradeoff of the Variable-Rank RDST Code .................... 104
  9.7.1 Upper Bound on Diversity Gain $d(r)$ .................................. 105
  9.7.2 Lower Bound on $d(r)$ ..................................................... 108
9.8 Rate-Diversity Tradeoff of the \(v\)-r RDST Code .................... 110
9.9 Other Generalizations of the RDST Code - A Review ................. 111
9.10 Generalized RDST Codes Via Local Power-Series Expansion ....... 113
  9.10.1 RDST Codes via Power-Series Expansion .......................... 113
  9.10.2 Case When $\beta$ is Not a Principal Ideal ......................... 115
  9.10.3 Examples ................................................................. 116

10 Simulation Results, Some Miscellany and Future Work ............... 118
  10.1 Decoding and Simulation of Square ST Codes Derived from CDA ... 119
  10.2 Review of Perfect Codes .................................................. 120
  10.3 Determining the Exact D-MG Tradeoff ................................. 121
  10.4 Approximate Universality ................................................. 122

A D-MG Optimality of the Puncturing Construction ....................... 124

B D-MG Optimality of the Stacking Construction .......................... 126

References ................................................................. 128
List of Tables

7.1 Non-Norm Elements ........................................ 79
7.2 Cyclic Galois Extensions with Reduced Signalling Complexity ........ 83
7.3 Values of $n$ for which Lower Signalling Complexity is Obtained than in [56] 84

9.1 Proportions $a_i$ for certain $r$ ($n = 4$) ........................ 109
List of Figures

2.1 The MIMO Channel ........................................ 6
2.2 The effect of diversity order on error probability ................. 11
2.3 Illustrating Linear Increase in Capacity: $n_t = n_r$. Plots are for SNR in range $0 \leq \text{SNR} \leq 35$ dB in 5 dB increments .......................... 15
2.4 The V-BLAST Architecture .................................. 19

3.1 The Diversity-Multiplexing Gain Tradeoff ($n_t = n_r = 4$, $T \geq 7$) .... 27
3.2 Bounds on the Diversity-Multiplexing Gain Tradeoff: $n_t = n_r = T = 4$ .......................... 28
3.3 D-MG tradeoff of the Alamouti Code: One Receive Antenna (left) and Two Receive Antennas (right) .......................... 30
3.4 D-MG Tradeoff of V-BLAST ($n_t = n_r = 4$) .................................. 31
3.5 The D-BLAST Architecture .................................. 32
3.6 D-MG Tradeoff of D-BLAST ($n_t = n_r = 4$) .................. 33
3.7 Bound on the D-MG Tradeoff of the Algebraic Space-Time Code of Damen et al. ($n_t = n_r = T = 2$) .......................... 34

4.1 D-MG Tradeoff of the Yao-Wornell Code ($n_t = n_r = T = 2$) .... 43

5.1 Ideals in Algebraic Number Fields .............................. 55

6.1 Space-Time Codes from Division Algebras ........................ 59
6.2 Structure of a Division Algebra ............................... 62
6.3 The Square and Hexagonal Lattices ................................ 67

7.1 Construction of Cyclic Extensions of $\mathbb{Q}(i)$ of Arbitrary Degree ........................ 72
7.2 The Fundamental Theorem of Galois Theory ........................ 73
7.3 Inertial degree in the compositum field ............................ 78
7.4 Cyclic Galois Extensions with Reduced Signalling Complexity .......................... 80
7.5 Non-Norm Elements - Resolving the Tower in Figure 7.4 ............ 84

8.1 Division Algebra ........................................... 86
8.2 The Rectangular Construction Framework .......................... 87

9.1 The Rate-Diversity Tradeoff ($n_t = 4$) .......................... 92
9.2 Bounds on the D-MG Tradeoff of the RDST Code ($n_t = n_r = T = 4$) [53] 102
9.3 The ‘Inside Story’ of the Bounds on the D-MG Tradeoff of the RDST Code 103
LIST OF FIGURES

9.4  D-MG tradeoff of the RDST and the variable-rank RDST codes . . . . . . 110
9.5  Dual-radii AM-PSK Constellation . . . . . . . . . . . . . . . . . . . . . . 112

10.1 Simulation Results for the $3 \times 3$ CDA Based Space-Time Codes . . . . . . 120
Chapter 1

Introduction

The invention of the electric telegraph in the early 19th century can be thought of as a turning point in the history of modern Communications Engineering. The concept of wireless communication, however, emerged only several decades subsequent to the invention the electric telegraph. In the 1860s, James Clerk Maxwell, a Scottish physicist, predicted the existence of radio waves, and in 1886 Heinrich Rudolph Hertz, a German physicist, demonstrated that rapid variations of electric current could be projected into space in the form of radio waves similar to those of light and heat. In 1893, Nikola Tesla, while carrying out experiments with high frequency electric currents, is claimed to have been the first person to demonstrate wireless communication. Independently, Guglielmo Marconi, an Italian inventor, proved the feasibility of radio communication. He sent and received his first radio signal in Italy in 1895. By 1899 he flashed the first wireless signal across the English Channel and three years later received the letter “S”, telegraphed from England to Newfoundland. This, in 1902, was the first successful transatlantic radiotelegraph message. Wireless communication brought with it’s advent, several new and interesting problems that needed to be solved to ensure reliable transmission of information.

There has been no looking back ever since these early inventions, and we have now entered an age in which the dream of achieving ubiquitous, reliable and high speed communications is becoming a reality. Most modern developments in the area, however,
may be attributed to the laying down of a “mathematical theory of communication” by Claude E. Shannon in a landmark paper in 1948. This pioneering work of Shannon introduced the field of information theory, owing to which the fields of coding theory and digital communication have advanced with rapid strides. Infact, it is quite appropriate to state that Shannon’s paper formed the basis of the entire digital communications revolution, from cell phones to the Internet.

The wireline channel, which can be modelled as an additive white Gaussian noise (AWGN) channel, is relatively benign when compared to the wireless channel. The wireless channel suffers from the effect of fading, caused due to radio waves scattered from different physical objects in the propagation environment interfering with each other at the receiver. This effect degrades the error probability of the wireless system and suitable methods to circumvent this problem need to be employed in wireless communication systems.

One way of achieving reliable communication over wireless channels is by employing “space-time codes”, the concept of which was introduced in [16, 17] in 1998. Space-time coding involves using multiple antennas at both the transmitter and the receiver, and is a natural extension of the diversity reception techniques prevalent earlier. This channel is appropriately called a multiple-input multiple-output (MIMO) channel. By employing intelligent signal design at the transmitter, data is suitably encoded and transmitted simultaneously through all transmit antennas. The antenna array at the receiver receives a noisy superposition of these transmitted signals. Space-time coding is all about the design of transmission schemes for MIMO channels (which have been called space-time codes in the literature) that enable the receiver to efficiently decode the transmitted data. The term “space-time” arises from the fact that coding of data is performed over the transmit antennas (spatial dimension) and over a block of transmission slots (time dimension). It is known that significant advantages result from the use of space-time codes over MIMO channels. In particular, both the reliability of transmission and data rates are improved; we will quantify these in a review of space-time codes to follow in Chapter 2.
Chapter 1. Introduction

Intuitively, it seems that rate and reliability are conflicting requirements, with one being achieved at the expense of the other. Recently, there have been inroads made into quantifying this tradeoff between rate and reliability. Two works, the rate diversity tradeoff introduced by Tarokh et al. in [16] and subsequently worked on by Lu and Kumar in [64], and the diversity-multiplexing gain (D-MG) tradeoff introduced by Zheng and Tse in [47] have quantified the best possible reliability that a space-time system can achieve at any given data rate. Although the basic idea behind both these tradeoffs is identical, certain differences in formulation result in the two different works. The D-MG tradeoff is presented in Chapter 3 while the rate-diversity tradeoff appears in Chapter 9.

Our primary interest in this thesis will be on the construction of space-time codes that achieve these tradeoffs. While there exist quite a few space-time codes that achieve the rate-diversity tradeoff, prior to our work, explicit construction of space-time codes for arbitrary number of transmit antennas that achieved the D-MG tradeoff remained an open problem. While a sufficient condition for a space-time scheme to achieve the D-MG tradeoff was presented in [54], this thesis presents a complete solution to the problem of constructing explicit space-time codes that satisfy this sufficient condition and hence achieve the D-MG tradeoff. Some insights into what it takes for a space-time scheme to achieve the D-MG tradeoff and a review of the sufficient condition from [54] are presented in Chapter 4. A primary result of this thesis, the construction of space-time codes achieving the D-MG tradeoff for arbitrary number of transmit and receive antennas appears in Chapter 7. While this chapter deals with the case of square space-time codes, an extension to the rectangular case is dealt with in Chapter 8. Some work related to reducing the signalling complexity of the constructed square space-time codes also appears in Chapter 7. Since our constructions are derived from abstract algebraic objects known as cyclic division algebras (CDA), Chapter 5 presents a review of some results from abstract algebra and number theory that we will use in our constructions and Chapter 6 contains an exposition of the idea behind deriving space-time codes from CDA.
Chapter 1. Introduction

The second part of this thesis will focus on providing a generalization to the rank-distance space-time (RDST) codes introduced recently by Lu and Kumar [64, 69] as a family of space-time codes that meet the rate-diversity tradeoff. The generalization proposed in Chapter 9, called the variable-rank RDST code, will be shown to improve upon the D-MG tradeoff of the RDST construction. The RDST codes and their generalizations have simple constellations as their signalling alphabet, thus making them attractive for practical implementations.

Simulation results for the D-MG optimal square designs from CDA appear in Chapter 10. Some other miscellaneous topics of interest and avenues for future work in Chapter 10 wrap up this thesis.
Chapter 2

MIMO systems and Space-Time Coding

The initial focus of this chapter will be to provide an overview of the field of space-time coding over MIMO (multiple input, multiple output) channels and highlight the advantages that it has to offer. This chapter will also serve to lay down the notation that we will follow in this thesis.

Keeping in mind that reliable transmission of data at high rates is the goal of any digital communication system, we will analyse both the error probability and the capacity of the MIMO channel. We will demonstrate that the MIMO channel has an improved diversity order (leading to significantly lower error probabilities) as well as a higher capacity when compared with the single antenna fading channel (which we will henceforth refer to as a single-input single-output (SISO) channel).

The area of space-time (ST) code construction initially comprised of two schools of thought. The first one aimed at constructing space-time codes which promised full diversity advantage according to the rank criterion of [16, 17]. They however did not exploit the huge increase in capacity offered by the space-time channel. Examples include the space-time trellis codes in [16, 17] and the orthogonal space-time block codes in [23, 22]. The aim of the other school of thought was to construct space-time codes which have high throughput and hence exploit the large capacity of the space-time channel.
These space-time codes however compromised on diversity advantage. The Bell Labs space-time architecture (BLAST) [19] is a typical example.

Some other constructions have aimed at simultaneously achieving both the above mentioned objectives. Among the first such constructions is the rank-distance space-time (RDST) construction of Lu and Kumar [64, 69]. The family of RDST codes allow us to sacrifice rate of transmission for better probability of error and vice-versa. Some other classes of codes in this “modern school of thought” which strive to achieve good rates with high reliability include the rotation based ST constructions of Yao-Wornell [42] and Dayal-Varanasi [44]. Among our main interests in this thesis will be to present space-time code constructions that are optimal with regard to trading-off rate for reliability. Towards this end, and in order to give the reader a flavour for the directions in space-time code constructions, we will present in this chapter a review of a few typical representatives of space-time constructions from the two traditional schools of thought.

### 2.1 The Space-Time Channel

![Diagram of MIMO Channel](image)

Figure 2.1: The MIMO Channel

We will work with the MIMO channel consisting of \( n_t \) transmit and \( n_r \) receive antennas shown in Figure 2.1. All antennas transmit simultaneously and we will assume that the signals transmitted from each transmit antenna have the same transmission period.
The signal received at each receive antenna is a linear superposition of faded copies of the \( n_t \) transmitted signals perturbed by noise.

We will assume that the fading is frequency non-selective (flat fading), which is equivalent to the assumption that the transmission period is significantly larger than the delay spread of the channel. Under this assumption, we can associate a fade coefficient \( h_{ij} \) corresponding to the \( j^{th} \) transmit and \( i^{th} \) receive antenna. We will assume that the channel is quasi-static, i.e., the fade coefficients \( h_{ij} \) remain constant for a duration of \( T \) channel uses, after which they change at random. \( T \) is called the quasi-static interval. This is equivalent to the assumption that the channel coherence time is approximately \( T \) times the symbol duration. We will also assume the distribution of the \( h_{ij} \) to be identical and independently distributed (i.i.d.) circularly symmetric complex Gaussian with zero mean and unit variance, \( \mathbb{C}N(0, 1) \). This corresponds to the Rayleigh fading model, which, by the central limit theorem, is an accurate model for the rich-scattering wireless channel. The assumption of independent Rayleigh paths is to be thought of as an idealized version of the result that for antenna elements placed on a rectangular lattice with half wavelength (\( \lambda/2 \)) spacing, the path losses tend to roughly decorrelate. Note that, for example, with a 5 GHz carrier frequency, \( \lambda/2 \) is only about 3 cm. So at sufficiently high carrier frequencies there can be great opportunity for accommodating numerous antennas in the regions of space occupied by the communicating stations.

Let the signal transmitted by the \( i^{th} \) antenna at the \( j^{th} \) time instant be denoted by \( x_{i,j} \). We will perform coding over our transmission signals for a duration equal to the quasi-static interval of \( T \) channel uses as well as over the signals transmitted from the \( n_t \) antennas during each of these \( T \) transmissions. This two-dimensional coding over the time and spatial (antenna) dimensions leads to the notion of a space-time code. A space-time code \( \mathcal{X} \) is therefore represented by a set of \( (n_t \times T) \) matrices, where the \((i, j)^{th}\) entry of each codeword matrix \( X \in \mathcal{X} \) represents the complex baseband equivalent of the
signal transmitted from the $i^{th}$ antenna at the $j^{th}$ time instant,

$$X = \begin{bmatrix}
    x_{1,1} & x_{1,2} & \cdots & x_{1,T} \\
    x_{2,1} & x_{2,2} & \cdots & x_{2,T} \\
    \vdots & \vdots & \ddots & \vdots \\
    x_{n_t,1} & x_{n_t,2} & \cdots & x_{n_t,T}
\end{bmatrix}_{n_t \times T}, \quad x_{i,j} \in \mathbb{C} \forall i,j. \quad (2.1)$$

Throughout this thesis, we will assume that $T \geq n_t$.

Under this framework, when the code matrix $X \in \mathcal{X}$ is transmitted, the $(n_r \times T)$ received signal matrix (the $(i,j)^{th}$ entry of which represents the signal received by the $i^{th}$ receive antenna at the $j^{th}$ time instant) is given by

$$Y = HX + W, \quad (2.2)$$

where $H = \{h_{ij}\}$ is the $(n_r \times n_t)$ channel matrix and $W$ is the $(n_r \times T)$ noise matrix. As mentioned before, the entries of $H$ are assumed to be i.i.d., circularly symmetric complex Gaussian $\mathcal{CN}(0,1)$ random variables. We will assume that the entries of the additive noise matrix $W$ also follow an identical distribution.

In order to ensure fair comparison between systems employing different number of transmit antennas, we will impose the following average power constraint,

$$\mathbb{E}(\|X\|_F^2) \leq T \text{ SNR}, \quad (2.3)$$

where $\| \cdot \|_F$ denotes the Frobenius norm. This essentially means that on an average, the transmitted power across all transmit antennas at any time instant is atmost SNR (noise variance is assumed to be unity).

We will assume that the receiver knows the exact realization of the channel matrix (i.e., the case of coherent reception), but that the transmitter knows only the statistics of the channel (and not the exact realization).
Definition 1. The rate of a space-time code $\mathcal{X}$ in bits per channel use is defined to be

$$R = \frac{\log_2 |\mathcal{X}|}{T},$$

where $|\mathcal{X}|$ denotes the cardinality of the space-time code.

As mentioned previously, we will now proceed to demonstrate the twin benefits of MIMO systems - better reliability quantified by probability of error and higher data rates achievable, quantified by the capacity of the MIMO channel.

### 2.2 Pairwise Error Probability

Analysis of the pairwise error probability (PEP) of MIMO systems with space-time coding was carried out in [16] and [17]. We present here a brief review of their work. Although the codeword error probability is of more relevance in quantifying the reliability of the system, an analysis of the PEP provides several useful insights and also provides an upper bound on the codeword error probability through the union bound.

Let $X_1, \ldots, X_{|\mathcal{X}|}$ denote the codeword matrices that constitute an $(n_t \times T)$ space-time code $\mathcal{X}$ ($T \geq n_t$). Assume without loss of generality that $X_j \in \mathcal{X}$ was transmitted. We are interested in the pairwise error probability $P(X_j \rightarrow X_k)$ that the receiver incorrectly decodes in favour of some other codeword matrix $X_k \in \mathcal{X}$, $j \neq k$. Define the difference matrix $\Delta X_{jk} = X_j - X_k$ and let $\{\lambda_i\}_{i=1}^{\nu}$ be the eigenvalues of $\Delta X_{jk} \Delta X_{jk}^\dagger$ (here $\Delta X_{jk}^\dagger$ denotes the hermitian of $\Delta X_{jk}$). It is shown in [16, 17] that the pairwise error probability

$$P(X_j \rightarrow X_k) \leq \left[ \frac{1}{\prod_{i=1}^{n_t} \left( 1 + \frac{\lambda_i \text{SNR}}{4n_t} \right)} \right]^{n_r}.$$  

If the rank of $\Delta X_{jk} \Delta X_{jk}^\dagger$ is $\nu$, and $\lambda_1, \lambda_2, \ldots, \lambda_\nu$ are its non-zero eigenvalues, it follows that

$$P(X_j \rightarrow X_k) \leq \left( \prod_{i=1}^{\nu} \lambda_i \right)^{-n_r} \left( \frac{\text{SNR}}{4n_t} \right)^{-\nu n_r}. \tag{2.4}$$
Before analysing the above expression, we digress briefly to introduce a couple of definitions.

**Definition 2.** The diversity advantage is the exponent of SNR in the denominator of the expression for pairwise error probability.

**Definition 3.** The coding advantage is an approximate measure of the gain over an uncoded system operating with the same diversity advantage.

It is now clear from (2.4) that a diversity advantage of $\nu n_r$ and a coding advantage of $(\lambda_1 \lambda_2 \ldots \lambda_\nu)^\frac{1}{\nu}$ is achieved. From the above analysis, the following design criteria are proposed [16, 17].

- **The Rank Criterion:** In order to achieve the maximum diversity advantage $n_t n_r$, the matrix $\Delta X_{jk}$ has to be full rank for any two codewords $X_j, X_k$. If $\Delta X_{jk}$ has minimum rank $\nu$ over the set of two tuples of distinct code matrices, then a diversity advantage of $\nu n_r$ is achieved.

- **The Determinant Criterion:** Suppose that a diversity advantage of $\nu n_r$ is our target. The design target is to ensure that the product of the non-zero singular values of the difference matrix $\Delta X_{jk}$ be as large as possible for all $j \neq k$. If a diversity advantage of $n_t n_r$ is the design target, then the criterion reduces to ensuring that the minimum of the determinant of $\Delta X_{jk}$ taken over all pairs of distinct codewords $X_j, X_k$ is maximized.

Another commonly used terminology in the literature is *transmit diversity*, defined below.

**Definition 4.** A space-time code $\mathcal{X}$ is said to achieve transmit diversity $\nu$ if the difference $\Delta X$ of any two distinct code matrices from $\mathcal{X}$ has rank at least $\nu$. Such a space-time code is known as a rank-$\nu$ space-time code and achieves a diversity advantage of $n_t \nu$.

**Remark 1.** The diversity advantage is not to be confused with the term “diversity gain” to be introduced later.
The above analysis demonstrates clearly the advantage that space-time coding offers over the SISO fading channel in terms of error probability. The SISO fading channel is known to have unit diversity order, i.e., the error probability falls inverse linearly with SNR. However, through intelligent signal design, we can achieve a diversity advantage of \( n_t n_r \) by employing space-time codes. Thus, the inverse linear fall of error probability with SNR is converted into an inverse polynomial fall, with the order of decay equal to the diversity advantage of the system. The maximum order of decay of \( n_t n_r \) intuitively corresponds to averaging over the \( n_t n_r \) independent fade coefficients between each pair of transmit and receive antennas. Figure 2.2 illustrates the effect of increase in diversity order on the fall of error probability with SNR.

![Figure 2.2: The effect of diversity order on error probability](image)

We shall see in the following subsection that diversity advantage is only one side of the coin when employing space-time coding; a significant increase in channel capacity is also obtained.
2.3 Capacity of the Space-Time Channel

The previous section projected MIMO systems as a method to combat fading. We now present the view that fading can in fact be beneficial, by increasing the number of degrees of freedom available for communication. When the path gains between transmit and receive antennas fade independently, the channel matrix is full rank with probability one, in which case multiple spatial channels are created. This enables parallel streams of independent data to be transmitted across the MIMO channel, thereby increasing the data rate. This effect is known as *spatial multiplexing* [47].

We will now present a quantitative analysis of the capacity of the MIMO channel along the lines of Telatar [18]. The results in this section do not require the quasi-static channel assumption. We will consider the following channel model in which the received vector $y \in \mathbb{C}^{n_r}$ and the transmitted vector $x \in \mathbb{C}^{n_t}$ are related as

$$y = Hx + w,$$

where the channel matrix $H \in \mathbb{C}^{n_r \times n_t}$ and $w$ is a vector of i.i.d. $\mathcal{CN}(0,1)$ entries. The entries of $H$ also constitute a collection of i.i.d. $\mathcal{CN}(0,1)$ random variables. The entries of $H$ change independently for every channel use. The power constraint analogous to that in (2.3) is imposed, i.e.,

$$\mathbb{E}(x^\dagger x) \leq \text{SNR}.$$  

If $Q$ represents the covariance matrix of the random vector $x$, this is equivalent to,

$$\text{tr}[\mathbb{E}(xx^\dagger)] = \text{tr}(Q) \leq \text{SNR}. \quad (2.6)$$

The reader will notice, that apart from the quasi-static assumption, the above model is the exact vector equivalent of the general MIMO model presented in (2.2).

In the analysis to follow, we will assume as before that the receiver knows the exact realization of the channel while the transmitter knows only the distribution of the channel.

The mutual information between the input and output given a particular realization
of the channel $H$ can be written as

$$I(x; y|H) = \mathcal{H}(y|H) - \mathcal{H}(y|x, H)$$

$$= \mathcal{H}(y|H) - \mathcal{H}(w), \quad (2.7)$$

where $\mathcal{H}(\cdot)$ denotes the differential entropy.

We will make use of a couple of standard results in information theory [13]. The first result is that the differential entropy of a complex Gaussian random vector with covariance matrix $Q$ is given by $\log \det(\pi e Q)$. The other result is that given a covariance matrix $Q$, the complex random vector that maximizes the differential entropy is a circularly symmetric complex Gaussian random vector with covariance matrix $Q$.

$(2.7)$ now becomes

$$I(x; y|H) = \mathcal{H}(y|H) - \log \det(\pi e I)$$

From the power constraint on the input $x$ in $(2.6)$ and by restricting our attention to zero-mean $x$, we have that $E(yy^\dagger) = I + HQH^\dagger$. Now using the entropy maximizing property of the circularly symmetric Gaussian random vector,

$$I(x; y|H) \leq \log \det[\pi e(I + HQH^\dagger)] - \log \det(\pi e I)$$

$$= \log \det(I + HQH^\dagger)$$

Equality in the above relation holds when the input $x$ is circularly symmetric complex Gaussian. The **ergodic capacity** is obtained by averaging the above quantity with respect to the channel and maximizing over the choice of non-negative definite $Q$ subject to the power constraint in $(2.6)$,

$$C_{er\bar{g}}(\text{SNR}) = \max_{\text{tr}(Q) \leq \text{SNR}} E[\log \det(I + HQH^\dagger)].$$

Further, it can be shown that the optimal $Q$ must be a scaled identity matrix $\alpha I$ [18]. It is
now clear that the maximum is achieved when $\alpha$ is the largest possible, viz. $\alpha = \text{SNR}/n_t$. We thus arrive at the following theorem.

**Theorem 1.** The ergodic capacity of the $n_t$ transmit, $n_r$ receive antenna MIMO channel is given by

$$C_{\text{erg}}(\text{SNR}) = \mathbb{E}\left[ \log \det \left( I + \frac{\text{SNR}}{n_t} HH^\dagger \right) \right].$$

The capacity is achieved when the input is zero-mean circularly symmetric complex Gaussian with covariance matrix $(\text{SNR}/n_t)I$.

Note that the above theorem says that the optimal power allocation strategy when the transmitter has no knowledge of the channel is to allocate equal power among all transmit antennas. It can be shown that in the case when the transmitter knows the channel, the optimal $Q$ is diagonal and the entries can be found through standard “water-filling” arguments [18]. However, in this thesis, we will only be concerned with the case when the transmitter has no knowledge of the channel.

### 2.3.1 Evaluation of the Capacity

The evaluation of the ergodic capacity given in Theorem 1 by averaging over the channel is not straightforward. It was shown by Foschini [19] that at high SNR, the ergodic capacity for the $n_t$ transmit, $n_r$ receive antenna MIMO channel with i.i.d Rayleigh faded gains between the antenna pairs is given by

$$C_{\text{erg}}(\text{SNR}) \approx \min\{n_t, n_r\} \log \text{SNR}. \quad (2.8)$$

The huge increase in channel capacity of the space-time channel is now evident. At high SNR, the capacity of an AWGN channel is approximately $\log \text{SNR}$. The MIMO channel can therefore be thought of as equivalent to $\min\{n_t, n_r\}$ parallel AWGN channels. Notice also that an increase in capacity over the AWGN channel is observed only when antenna diversity is employed at both the transmitter and the receiver. A plot showing the linear increase in capacity with increase in the number of transmit and receive antennas from
[18] is shown in Figure 2.3.

![Figure 2.3: Illustrating Linear Increase in Capacity: $n_t = n_r$. Plots are for SNR in range $0 \leq \text{SNR} \leq 35 \text{ dB}$ in 5 dB increments](image)

2.3.2 Non-Ergodic Channels

The investigation of non-ergodic channels will serve to introduce a very important concept, that of outage probability.

The results of the previous section were based on the fact that the maximum mutual information of the channel is the capacity. This is true only if the process that generates the channel $H$ is ergodic. To recall, an ergodic random process is one for which the time-average equals the ensemble average.

In this subsection, we will investigate the case when $H$ is chosen randomly at the beginning of time but is held fixed for all uses of the channel. Notice that in this case, the process is non-ergodic. The Shannon capacity in this scenario is zero, since there is a non-zero probability that the channel realization picked will be incapable of supporting the target data rate (irrespective of the code length). As a performance measure, we will
define the outage probability $P_{\text{out}}(R, \text{SNR})$ as the probability that a data rate of $R$ is not supported by the channel under a power constraint of SNR,

$$P_{\text{out}}(R, \text{SNR}) = \inf_{Q : Q \geq 0, \text{tr}(Q) \leq \text{SNR}} \text{Pr}[\log \det(I + HQH^\dagger) < R].$$  \hspace{1cm} (2.9)

We will see in the next chapter that the outage probability is a fundamental property of the channel. It will be shown later that the error probability for transmission over a channel is lower bounded by the outage probability for that channel.

\section*{2.4 Space-Time Block Codes and Examples}

Construction of coding schemes to exploit the above mentioned desirable properties of the space-time channel has been an active area of research over the past decade. The paradigm of space-time block codes was introduced in the seminal work by Tarokh et al. \cite{Tarokh}. We will first review some definitions pertaining to space-time block codes.

\textbf{Definition 5.} An $n_t \times T$ ($T \geq n_t$) space-time block code $\mathcal{X}$ is a finite number of $n_t \times T$ matrices with complex entries.

\textbf{Definition 6.} An $n_t \times T$ space-time design is an $n_t \times T$ matrix whose entries are complex linear combinations of a set of indeterminates $\{x_i\}_{i=1}^k$. A space-time block code over the base alphabet $\mathcal{A}$ is obtained by restricting the indeterminates to take values from a signal constellation $\mathcal{A}$. The rate of such a space-time block code is $k/T$ symbols from $\mathcal{A}$ per channel use.

As mentioned at the beginning of this chapter, space-time block code design initially followed one of two objectives: achieving full diversity or transmitting at rates close to capacity. More recent works have focussed on trading off one for the other. A brief exposition follows through examples from each case.
2.4.1 Orthogonal Design Based Space-Time Block Codes

The first example of an orthogonal space-time block code (OSTBC) was the Alamouti code [22]. The concept of orthogonal space-time block codes was introduced by Tarokh et al. [23] in order to provide generalizations of the Alamouti code. The orthogonal space-time block codes are full rank space-time codes which have an additional desirable property that they permit single-symbol decodability due to their orthogonality property.

**Definition 7.** A generalized linear processing real (complex) orthogonal design in variables $x_1, x_2, \ldots, x_k$ is an $n_t \times T$ ($T \geq n_t$) matrix $E$ such that

- the entries of $E$ are real (complex) linear combinations of the variables $x_1, x_2, \ldots, x_k$
- $EE^H = D$, where $D$ is a diagonal matrix with $(i, i)^{th}$ diagonal element of the form $l_i(|x_1|^2 + |x_2|^2 + \ldots + |x_k|^2)$, where the $l_i$ are strictly positive numbers.

The rate of such a design is defined to be $k/T$.

The Alamouti Code [22] was the first orthogonal design constructed. The Alamouti code is a $2 \times 2$ design given by

$$X = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix}. \quad (2.10)$$

We will use the Alamouti code to illustrate the characteristic properties of an orthogonal design. The orthogonality condition of the Alamouti code is demonstrated by the following equation.

$$XX^\dagger = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix} \begin{bmatrix} x_1^* & -x_2 \\ x_2^* & x_1 \end{bmatrix} = (|x_1|^2 + |x_2|^2)I_{2 \times 2}$$

The single symbol decodability property of the Alamouti code is a consequence of it’s orthogonality. By considering $x_1 = u_1 + u_2$ and $x_2 = u_3 + u_4$, where $i = \sqrt{-1}$, we
may rewrite the Alamouti design as

$$X = u_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + u_2 \begin{bmatrix} \iota & 0 \\ 0 & -\iota \end{bmatrix} + u_3 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + u_4 \begin{bmatrix} 0 & \iota \\ \iota & 0 \end{bmatrix} = \sum_{i=1}^{4} u_i A_i \text{ (say)}$$

Define the inner product on $\mathbb{C}^{2\times 2}$, the vector space of all $2 \times 2$ matrices over $\mathbb{C}$, as,

$$< A, B > = \text{Re}[\text{tr}(AB^\dagger)] \quad \forall \ A, B \in \mathbb{C}^{2\times 2}. \quad (2.11)$$

The space-time channel model (2.2) becomes

$$Y = H \sum_{i=1}^{4} u_i A_i + W.$$ 

Let $h^2 \coloneqq ||H||^2_F$. Using the orthogonality property of the Alamouti code, it can be easily shown that the matrices $\{\varphi_i := \frac{HA_i}{h}\}_{i=1}^{4}$ are a set of pairwise orthonormal matrices (Note that the inner product under consideration for orthogonality is the one defined in (2.11)).

The channel model now reduces to

$$Y = h \sum_{i=1}^{4} u_i \varphi_i + W.$$

At this point, it is clear how the Alamouti code permits single-symbol decoding. The decoding metric for each $u_i$ is obtained at the receiver by computing the inner product of the received matrix $Y$ with $\varphi_i$. This operation reduces the space-time channel to a bank of parallel channels,

$$y_i = hu_i + w_i, \quad w_i \sim \mathcal{N}\left(0, \frac{1}{2}\right) \text{ for } i = 1, 2, 3, 4.$$ 

Thus decisions can be made on each of the constituent $u_i$, resulting in symbol by symbol decoding.

The above properties carry over to other orthogonal designs also. The existence of full-rate real and complex orthogonal designs is answered by the following theorem.
Theorem 2. • \( n \times n \) (square) real orthogonal designs with rate = 1 exist if and only if \( n = 2, 4 \) or 8 \([23]\).

• \( p \times n \) (rectangular) real orthogonal designs with rate = 1 exist for all \( n \) for some value of \( p > n \). For details regarding the value of \( p \), see \([23]\).

• \( p \times n \) (rectangular) complex orthogonal designs with rate = 1 do not exist for \( n > 2 \). The Alamouti code is the only full rate complex orthogonal design \([25]\).

To summarize, orthogonal designs have the following salient features to their credit: full diversity and minimal decoding complexity. They however, fail to exploit the full capacity of the space-time channel, as will be seen in subsequent chapters, and are also limited by their existence only for certain sporadic values of number of transmit antennas.

2.4.2 V-BLAST

The vertical Bell-labs layered space-time architecture (V-BLAST) was introduced in \([20]\). In V-BLAST, the incoming data stream is de-multiplexed into \( n_t \) substreams and each substream is encoded into symbols (typically QAM) and fed to the respective transmitter. Each transmitter transmits through a separate antenna and all transmitters work in parallel. Unlike conventional TDMA, FDMA or CDMA, in V-BLAST, all transmitters address the entire bandwidth for the entire duration of transmission, and no orthogonality in code-domain is introduced during transmission. This is what enables
V-BLAST to achieve much higher spectral efficiencies than the multiple access schemes. The propagation environment itself, which is assumed to exhibit significant multipath, is exploited to achieve the signal decorrelation necessary to separate the co-channel signals.

The receiver is assumed to employ $n_r$ receive antennas, each feeding a conventional QAM receiver. It is assumed that $n_r \geq n_t$. These receivers also operate co-channel, each receiving the signals radiated from all $n_t$ transmit antennas. The decoding algorithm proposed in [20] employs “nulling and cancellation”, which is a low-complexity sub-optimal decoding algorithm that uses intelligent signal processing. We refer the reader to [20] for the details.

It is evident from inspection that the minimum rank of the difference of any two distinct code matrices drawn from the V-BLAST scheme is 1. This is because the code matrices consist of uncoded entries, which results in the existence of a pair of codeword matrices which differ in just one entry. V-BLAST therefore fails to achieve full transmit diversity.

It will be shown through the diversity-multiplexing tradeoff in Chapter 3 that V-BLAST with the nulling and cancelling detector is capable of operating close to the capacity of the MIMO channel. Using a laboratory prototype, enormous spectral efficiencies of $20 - 40 \text{ bps/Hz}$ in an indoor propagation environment at realistic SNRs and error rates have been demonstrated [20].
Chapter 3

The Diversity-Multiplexing Gain Tradeoff

The reader would have noticed that some discussions regarding the rate and reliability of the space-time code constructions presented in the previous chapter were qualitative in nature. Rate and reliability seem to be conflicting requirements, although the exact relationship between these quantities remains hazy. Further, it is not immediately clear as to how to compare two given space-time constructions. For example, one is not sure on what scale to compare say the Alamouti code (designed to provide full diversity) with V-BLAST employing two transmit antennas (designed to provide high rates). In this chapter, we provide a simultaneous solution to both these problems in the form of the diversity-multiplexing gain (D-MG) tradeoff introduced by Zheng and Tse in [47].

The D-MG tradeoff quantifies the best possible reliability that any space-time scheme can achieve for a given value of rate of transmission, under the assumption of high SNR. In this chapter, we review the formulation and details of the D-MG tradeoff. Following this, we will present the D-MG tradeoff of some space-time constructions from the literature.
3.1 Preliminaries

In the previous chapter, we saw that most of the initial space-time research focussed on developing schemes which extracted maximal diversity advantage (e.g., orthogonal STBCs) or maximal rates (e.g., V-BLAST). We also saw that each design goal compromised on the other. The crux of the D-MG tradeoff is that both high rates and good reliability can be simultaneously achieved, but there is a fundamental tradeoff between these two quantities. Better values of one quantity come at the expense of sacrificing the other.

We will first establish the basic notation and framework under which the D-MG tradeoff is defined.

**Definition 8.** The symbol \( \overset{\sim}{=} \) denotes exponential equality, i.e., we write \( f(SNR) \overset{\sim}{=} SNR^b \) to denote

\[
\lim_{SNR \to \infty} \frac{\log f(SNR)}{\log SNR} = b.
\]

The symbols \( >, \, <, \, \geq \) and \( \leq \) are similarly defined.

The space-time channel model is as defined in (2.2). The D-MG tradeoff is defined under the asymptotics of very high SNR.

**Definition 9.** A space-time scheme \( \{ X(SNR) \} \) is a family of space-time codes of block length \( T \), one at each SNR level.

The concept of a space-time scheme assumes importance since the ergodic capacity of the space-time channel increases linearly with \( \log SNR \), see (2.8). In order to achieve a non-vanishing fraction of the capacity, we consider schemes whose data rate also scales similarly with SNR. Typically, this is achieved by varying with SNR, the size of the base constellation over which the space-time code is constructed. Let \( R(SNR) \) denote the rate of the space-time scheme \( X(SNR) \) in bits/channel use, i.e.,

\[
R(SNR) = \frac{1}{T} \log_2 |X(SNR)|.
\]
We say that the scheme \( \mathcal{X}(\text{SNR}) \) achieves a spatial multiplexing gain of \( r \) if the data rate
\[
R(\text{SNR}) \approx r \log_2 \text{SNR}.
\] (3.2)

According to this formulation, any space-time scheme which has fixed rate (i.e., a rate that does not vary with SNR) has a multiplexing gain of zero. Also, from (3.1) and (3.2), we conclude that for a space-time scheme \( \mathcal{X}(\text{SNR}) \) to have a spatial multiplexing gain of \( r \), it’s cardinality must scale as
\[
|\mathcal{X}(\text{SNR})| = \text{SNR}^r T.
\] (3.3)

Wherever it is clear from the context, we will write just \( \mathcal{X} \) in place of the scheme \( \mathcal{X}(\text{SNR}) \).

The reliability of a space-time scheme will be quantified in terms of diversity gain. We say that a space-time code achieves a diversity gain of \( d \) if the codeword error probability of the space-time code decays with an SNR exponent of \( d \). To make the above concepts precise, we have the following definition [47].

**Definition 10.** A space-time scheme \( \{\mathcal{X}(\text{SNR})\} \) is said to achieve spatial multiplexing gain \( r \) and diversity gain \( d \) if the data rate
\[
\lim_{\text{SNR} \to \infty} \frac{R(\text{SNR})}{\log \text{SNR}} = r
\]
and the average error probability
\[
\lim_{\text{SNR} \to \infty} \frac{\log P_e(\text{SNR})}{\log \text{SNR}} = -d \quad (i.e., \ P_e = \text{SNR}^{-d}).
\]

The diversity-multiplexing gain tradeoff quantifies the best possible diversity gain \( d^*(r) \) that a space-time scheme can achieve at every value of the spatial multiplexing gain \( r \).
3.2 Outage and Error Probability

The concept of channel outage and the connection with non-ergodic channels was presented in Section 2.3.2. The outage probability will prove to be a fundamental quantity in determining the diversity-multiplexing gain tradeoff of a channel.

For the convenience of the reader, the expression for the outage probability (2.9) is repeated below,

\[
P_{\text{out}}(R, \text{SNR}) = \inf_{Q : Q \geq 0, \text{tr}(Q) \leq \text{SNR}} \Pr[\log \det(I + HQH^\dagger) < R].
\]

It can be shown [47] that as far as the exponent of SNR is concerned, we may without loss of generality assume the input (Gaussian) distribution to have covariance matrix \( Q = \text{SNR} I \). The outage probability

\[
P_{\text{out}}(R) \doteq \Pr[\log \det(I + \text{SNR}HH^\dagger) < R] = \Pr[\log \det(I + \text{SNR}H^\dagger H) < R].
\]

The outage probability is thus a characteristic of the fading distribution of the channel. We will assume the standard i.i.d. rayleigh fading model for further computation and will relegate a discussion of other channel models to Chapter 10. Considering the above equation with \( R = r \log \text{SNR} \) and denoting \( \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n'} \) as the non-zero eigenvalues of \( H^\dagger H \) where \( n' := \min\{n_t, n_r\} \), we obtain

\[
P_{\text{out}}(R) \doteq \Pr\left[\prod_{i=1}^{n'} (1 + \text{SNR}\lambda_i) < \text{SNR}^r\right].
\]

Let \( \lambda_i = \text{SNR}^{-\alpha_i} \). At high SNR, the above equation reduces to

\[
P_{\text{out}}(R) \doteq \Pr\left[\sum_{i=1}^{n'} (1 - \alpha_i)^+ < r\right]
\]
where \((x)^+\) denotes \(\max\{0, x\}\). The joint probability density function (PDF) of the ordered non-zero eigenvalues \(\lambda_i\) of \(H^\dagger H\) is given by [12] to be

\[
p(\alpha) = K [\log(SNR)]^{n'} \prod_{i=1}^{n'} (SNR)^{-\max(|n_r-n_t|+1, \alpha_i)} \prod_{i<j} (SNR^{-\alpha_i} - SNR^{-\alpha_j})^2 \exp\left( - \sum_{i=1}^{n'} SNR^{-\alpha_i} \right)
\]  

for some constant \(K\). It is shown in [47] that computation of the outage probability using the above PDF leads to the following theorem.

**Theorem 3.** The outage probability for the \(n_t\) transmit, \(n_r\) receive antenna i.i.d. rayleigh fading MIMO channel is given by

\[
P_{\text{out}}(r \log SNR) = SNR^{-d_{\text{out}}(r)}
\]

where \(d_{\text{out}}(r)\) is a piecewise linear function whose value at integral values of \(r = 0, 1, \ldots, \min\{n_t, n_r\}\) are given by

\[
d_{\text{out}}(r) = (n_t - r)(n_r - r).
\]

Following up on an explicit characterization of the outage probability for the i.i.d. Rayleigh fading channel, the following discussion links the outage with the error probability of the system. Intuitively, it seems that the outage probability should provide a lower bound to the codeword error probability; for whenever the channel is in outage, reliable communication is not possible (from Shannon’s channel capacity theorem) and an error is bound to happen (note here that reliable communication is not possible even under the assumption of coding over infinite length). Zheng and Tse [47] prove this fact rigourously using Fano’s inequality to arrive at the following theorem.

**Theorem 4.** (Outage bound) For the space-time channel in (2.2), let the data rate scale as \(R = r \log SNR\) (b/s/Hz). For any coding scheme, the probability of a detection error is lower-bounded by

\[
P_e(SNR) \geq SNR^{-d_{\text{out}}(r)}
\]

where \(d_{\text{out}}(r)\) is defined in Theorem 3.
3.3 The D-MG Tradeoff

We will denote $d^*(r)$ as the best possible diversity gain achievable by any scheme. Theorem 4 gives us an upper bound on the best possible diversity-multiplexing gain tradeoff, 

$$d^*(r) \leq d_{out}(r).$$  \hspace{0.5cm} (3.6)

To determine the exact D-MG tradeoff, the authors in [47] demonstrate the achievability of (3.6) by evaluating the performance of a particular code. Once this is done, we can conclude that $d^*(r)$ is exactly equal to $d_{out}(r)$. The probability of error for any code at a data rate $R = r \log \text{SNR}$ bits/channel use is given as

$$P_e(\text{SNR}) = P_{out}(R) P_r(\text{error} \mid \text{outage}) + P_r(\text{error, no outage}) \leq P_{out}(R) + P_r(\text{error, no outage}).$$  \hspace{0.5cm} (3.7)

To show the achievability of (3.6), we need to produce a scheme for which the error probability in the no-outage region is exponentially no worse than the outage probability. In [47], the authors show that this is the case with the ensemble of random Gaussian codes for the case when $T \geq n_r + n_t - 1$. The code matrices in the random Gaussian code are generated by choosing every entry of the $(n_t \times T)$ code matrices independently from a Gaussian distribution such that the power constraint in (2.3) is satisfied. An $(n_t \times T)$ code transmitting at a rate of $R$ bits per channel use consists of $\lceil 2^{RT} \rceil$ code matrices. It is shown in [47] that for coding length $T \geq n_t + n_r - 1$, the error probability averaged over the ensemble of random Gaussian codes is upper bounded as

$$P(\text{error, no outage}) \leq \text{SNR}^{-d_{out}(r)}.$$

From this, we can conclude that there exists one specific code in this ensemble whose performance is at least as good as the average performance. The outage bound (3.5) along with the above result proves the following theorem, which is the main result of
Chapter 3. The Diversity-Multiplexing Gain Tradeoff

Theorem 5. Assume $T \geq n_t + n_r - 1$. The optimal tradeoff curve $d^*(r)$ is given by the piecewise-linear function connecting the points $(k, d^*(k))$, $k = 0, 1, \ldots, \min\{n_t, n_r\}$, where

$$d^*(k) = (n_t - k)(n_r - k).$$

Corollary 6. Define $d^*_{\max} := d^*(0)$ and $r^*_{\max} := \sup\{r : d^*(r) > 0\}$. We then have that $d^*_{\max} = n_t n_r$ and $r^*_{\max} = \min\{n_t, n_r\}$.

The maximum value of diversity and multiplexing gains from the above corollary are in perfect agreement with our intuition. At high SNR, the maximum value of multiplexing gain $r^*_{\max}$ corresponds to rate approaching the ergodic capacity (2.8), while the maximum value of the diversity gain $d^*_{\max}$ corresponds to the maximum number of fade coefficients that a scheme can average over. An example plot for the D-MG tradeoff for $n_t = n_r = 4$, $T \geq 7$ is shown in Figure 3.1.

![Figure 3.1: The Diversity-Multiplexing Gain Tradeoff (n_t = n_r = 4, T \geq 7)](image)

For the case when $T < n_t + n_r - 1$, only bounds on the exact tradeoff were provided in [47]. Figure 3.2 shows the bounds on the optimal D-MG tradeoff for the case when $n_t = n_r = T = 4$. We refer the interested reader to [47] for the details regarding these bounds.
We will see in subsequent chapters that our results will extend the case for which the exact D-MG tradeoff is known to all \( T \geq n_t \).

### 3.4 D-MG Tradeoff of a Few Schemes from the Literature

Now that the framework of the D-MG tradeoff has been introduced, we examine the performance of some schemes which existed prior to this thesis on the D-MG tradeoff. In cases where obtaining the exact D-MG tradeoff of a space-time scheme is not tractable, we will present bounds on the D-MG tradeoff from [53, 52].

1. **Orthogonal Designs**

   Constructional details of the Orthogonal designs introduced by Tarokh et al. in [23] can be found in Section 2.4.1. They were examined in greater detail by Liang in [24], by Lu, Kumar and Chung in [26] and by Liang and Xia in [25].

   We will first analyse the D-MG tradeoff of the Alamouti code, which is the only rate one, complex orthogonal design that exists [25]. To recall, each code matrix
in the Alamouti code is of the form (2.10),

\[ X = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix}, \]  

(3.8)

where the \( x_i, i = 1, 2 \) are drawn from any fixed constellation. In [47], the outage probability of the Alamouti code is shown to be

\[ P_{out}(R) = \text{SNR}^{-n_t n_r (1-r)^+}. \]  

(3.9)

Following this, it is also shown that the scheme derived using a QAM constellation for each of the \( x_i \) achieves this outage probability. This essentially shows that the D-MG tradeoff of the Alamouti code is

\[ d_{alamouti}(r) = n_t n_r (1 - r)^+. \]

It is seen from the above that the Alamouti code meets the optimal D-MG tradeoff for the case when \( n_r = 1, T \geq 2 \), the corresponding plot is shown in Figure 3.3. However, for the case when \( n_r \geq 2 \), the Alamouti code is strictly suboptimal on the D-MG tradeoff. For example, the plot for the case when \( n_t = n_r = T = 2 \) is shown in Figure 3.3. In this case, the Alamouti code achieves the maximum possible diversity gain but fails to achieve the maximum spatial multiplexing gain. Heuristically, this can be explained by observing that while the received signal with \( n_t = n_r = T = 2 \) potentially lies in 4-dimensional complex space, the Alamouti code confines itself to a 2-dimensional subspace.

It can similarly be shown that all other orthogonal designs fail to achieve the optimal D-MG tradeoff. This fact is also justified by heuristic arguments similar to the one presented above.

2. V-BLAST

The vertical Bell-labs space-time architecture (V-BLAST), introduced in Section
2.4.2, was designed to maximize throughput. As mentioned previously, V-BLAST has unit transmit diversity irrespective of the number of transmit antennas.

The performance of V-BLAST under the nulling and cancelling scheme mentioned in Section 2.4.2 is limited by the error probability in decoding the first substream. The D-MG tradeoff of this scheme (which we shall call V-BLAST(1)) for the case when \( n_t = n_r = n \) (say) has been shown to be [47]

\[
d(r) = 1 - \frac{r}{n},
\]

and is plotted in Figure 3.4. Hence V-BLAST achieves the maximum multiplexing gain of \( \min\{n_t, n_r\} = n \). The maximum diversity gain achieved is in agreement with the transmit diversity times the number of receive antennas, viz., \( n \).

Among the methods to increase the reliability of V-BLAST are to decode at each stage that substream which corresponds to the maximum SNR at the output of the decorrelator (V-BLAST(2)) and to assign different rates to different substreams based on SNR (V-BLAST(3)). The D-MG tradeoff of these schemes are also plotted in Figure 3.4 [47].
3. **D-BLAST**

Notice that the D-MG tradeoff performance of V-BLAST was limited by the fact that no coding was employed over antennas, i.e., each codeword was transmitted through a single antenna and hence experienced only $n_r$ fade coefficients. Diagonal BLAST (D-BLAST) [19], with coding over the signals transmitted on different antennas, promises a higher diversity gain.

In D-BLAST, the input data stream is divided into substreams, each of which is transmitted on different antennas and time slots in a diagonal fashion, see Figure 3.5.

For example, in a $2 \times 2$ system, the transmitted signal in matrix form is

$$
\begin{bmatrix}
0 & x_1^{(1)} & x_1^{(2)} & \cdots \\
 x_2^{(1)} & x_2^{(2)} & x_2^{(3)} & \cdots 
\end{bmatrix},
$$

where $x_i^{(k)}$ denotes the symbols transmitted on the $i$th antenna for substream $k$.

The receiver uses a successive nulling and cancelling process [19], which demands that an overhead is required to start the detection process, corresponding to the 0 symbol in the above example. The D-MG tradeoff of D-BLAST with the nulling
receiver is computed in [47] and is plotted in Figure 3.6.

Since it has been observed that the nulling receiver causes the degradation of diversity, it is natural to replace the nulling step with a linear minimum mean-square (MMSE) receiver. It turns out that with the MMSE receiver, D-BLAST achieves the optimal D-MG tradeoff, provided one ignores the overhead that is required to start the D-BLAST processing [47], see Figure 3.6. With the overhead, the actual achieved data rate is decreased and neither the nulling nor the MMSE D-BLAST achieve the optimal D-MG tradeoff.

In a recent development, the authors in [79] show that using what are known as permutation codes [78] for the layers in D-BLAST, the optimal D-MG tradeoff can be achieved (while not ignoring the overhead) for the case when the number of receive antennas is either one or two. Their analysis makes use of joint ML decoding at the receiver.

4. **Algebraic Space-Time Code of Damen et al.**

In [28], the authors construct a $2 \times 2$ full-rank space-time code. Let $n_t = T = 2$ and let $\mathbb{Z}[i] = \{a + ib : a, b \in \mathbb{Z}\}$ denote the Gaussian integers. Let $\mathcal{B} \subseteq \mathbb{Z}[i]$. Then
the $(2 \times 2)$ space-time code constructed by Damen et al. [28] is given by

\[
X = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix}
    b_1 + \phi b_2 & \theta (b_3 + \phi b_4) \\
    \theta (b_3 - \phi b_4) & b_1 - \phi b_2
\end{bmatrix} : b_i \in B \right\}.
\]

where $\theta = \sqrt{\phi}, \phi = \exp(i \lambda), i = \sqrt{-1}$ and $\lambda \in \mathbb{R}$ is chosen so as to make the element $\exp(i \lambda)$ transcendental over $\mathbb{Z}(i)$.

Lower bounds on the D-MG tradeoff of the above construction are evaluated using [53] and are plotted in Figure 3.7. We conclude from the analysis using [53] that among the factors that lead to this construction having a poor performance on the D-MG tradeoff is that the minimum determinant of the code vanishes as SNR increases.
Figure 3.7: Bound on the D-MG Tradeoff of the Algebraic Space-Time Code of Damen et al. \((n_t = n_r = T = 2)\)
Chapter 4

Achieving the D-MG Tradeoff

Armed with the intuition gained from a review of the D-MG tradeoff of some space-time constructions, the focus of this chapter will be to “dissect a D-MG optimal space-time code and see what makes it tick”! Until recently, it was not clear among the research community as to what properties would ensure that a given space-time code achieves the D-MG tradeoff. The first part of this chapter will present a review of a sufficient condition for a space-time scheme to achieve the D-MG tradeoff given in [54]. Following this, we will demonstrate the optimality of a class of minimal-delay (square) space-time codes endowed with what is known as the non-vanishing determinant property. We will use this sufficient condition in subsequent chapters to identify and construct space-time block codes which achieve the diversity-multiplexing gain tradeoff for arbitrary number of transmit and receive antennas. Before actually moving on to do so, we will present in this chapter a review of some D-MG optimal space-time schemes which existed prior to our work.
4.1 Preliminaries

The formulation of the D-MG tradeoff demands that the data rate $R$ of the space-time scheme $\mathcal{X}$ scale with SNR as

$$R = r \log \text{SNR} \text{ bits/channel use},$$

where $r$ is defined as the multiplexing gain of $\mathcal{X}$. Recollect that the cardinality of the space-time scheme $\mathcal{X}$ must satisfy (3.3),

$$|\mathcal{X}| = \text{SNR}^r T.$$

To achieve the maximum diversity point of $d^*(0) = n_t n_r$, it is clear that we must employ a rank-$n_t$ code $\mathcal{X}$ (assume $T \geq n_t$). If $\mathcal{X}$ is a maximal rank-$n_t$ code, considering the entries of $\mathcal{X}$ to be drawn from a constellation $\mathcal{A}$, we obtain from the Singleton bound that

$$|\mathcal{X}| = |\mathcal{A}|^T = \text{SNR}^{Tr} \Rightarrow |\mathcal{A}| = \text{SNR}^r. \quad (4.1)$$

4.2 The Sufficient Condition

The following theorem, taken from [54, 55], presents a sufficient condition that ensures that a space-time scheme meets the optimal D-MG tradeoff.

**Theorem 7.** Consider a maximal rank $(n_t \times T)$ space-time code $\mathcal{X}$ ($T \geq n_t$). Let $\Delta X$ denote the difference of any two codeword matrices drawn from $\mathcal{X}$. Define

$$\min_{\forall \Delta X} \det(\Delta X \Delta X^\dagger) := \text{SNR}^\delta.$$

If the parameter $\delta$ of $\mathcal{X}$ satisfies

$$\delta = n_t - r,$$
then $\mathbf{X}$ is optimal with respect to the D-MG tradeoff for any number of receive antennas.

Proof. See [54, 55].

We see that the sufficient condition connects the SNR exponent of the minimum determinant of the space-time code to the rate of transmission (represented by the multiplexing gain). This calls for the design of energy efficient constellations and space-time code construction techniques over these constellations that result in a large value of the minimum determinant (also known as coding gain in conventional space-time literature). At high SNR, the sufficient condition also represents the best possible minimum determinant that a code can possess (upto SNR exponent) for a given rate and energy and vice versa. The energy is not immediately visible in the sufficient condition because of the normalization assumed in (2.3). Another heartening point to be noted is that the sufficient condition is analogous to the rate-distance criteria in conventional error-correcting code theory. The metric of hamming distance in error-correction code theory is supplanted by the rank and the minimum determinant in space-time coding. The intimate connection between space-time and error-correcting codes does seem natural in retrospect, considering the fact that when a space-time code over a constellation $\mathcal{A}$ is collapsed into a single column whose components are drawn from a ‘super-alphabet’ of size $|\mathcal{A}|^T$, the resulting column vector can be thought of as a conventional error-correcting code.

4.3 The Non-Vanishing Determinant Property

The non-vanishing determinant property was introduced in [38, 39, 40], as a method to preserve the spectral efficiency when we want to employ some outer block coded modulation scheme, which usually entails a signal set expansion. This scenario is very much relevant to the formulation of the D-MG tradeoff, where the data rate is assumed to scale with log SNR.

Definition 11. A space-time code $\mathcal{X}$ is said to have non-vanishing determinant (NVD) if the determinant $\det(\Delta \mathbf{X})$ of the difference $\Delta \mathbf{X}$ of any pair of distinct code matrices in
$X$, is bounded away from 0 even in the limit as the SNR and hence the size of the signal constellation, tends to infinity.

The importance of the non-vanishing determinant property will be realized in the subsequent section, where it will be shown that space-time codes built over energy-efficient alphabets endowed with the NVD property achieve the optimal D-MG tradeoff.

### 4.4 D-MG Optimality and Non-Vanishing Determinants

We will consider a class of “energy-efficient” square space-time codes that are endowed with the non-vanishing determinant property and show that these hypothetical codes satisfy the sufficient condition given in Theorem 7. Towards this goal, we will first define some parameters relating to the energy and signalling alphabets of space-time codes.

For clarity of exposition, we will assume the following space-time channel model with $n_t$ transmit and $n_r$ receive antennas in the sequel:

$$Y = \theta HX + W.$$ (4.2)

$X$ is drawn from a space-time code $\mathcal{X}$ and the entries of $H$ and $W$ are assumed to be i.i.d. circularly symmetric $\mathbb{C}N(0, 1)$ as before. $\theta$ is chosen to ensure

$$\mathbb{E}(\|\theta X\|^2_2) \leq T.SNR.$$ (4.3)

The above channel model is seen to be in agreement with that in (2.2), except for the change in notation.
4.4.1 Signal Alphabet

**Definition 12.** An alphabet $A$ is said to be scalably dense if $A$ can be scaled with SNR (we write $A(SNR)$ to identify its value for a specific SNR) in such a way that

$$a \in A(SNR) \Rightarrow |a|^2 \leq |A(SNR)|$$

It can be verified that the following constellations are scalably dense.

**Example 1.** QAM Constellation

$$A_{\text{QAM}}(SNR) = \{a + ib \mid -M + 1 \leq a, b \leq M - 1, a, b \text{ odd}\}$$

with $M^2$ elements.

**Example 2.** Rotated-QAM Constellation

$$A_{r-\text{QAM}}(SNR) = \left\{a + \omega b \mid -M + 1 \leq a, b \leq M - 1, a, b \text{ odd, } \omega = \exp \left(\frac{i2\pi}{3}\right)\right\}$$

with $M^2$ elements.

**Example 3.** PAM constellation

$$A_{\text{PAM}}(SNR) = \{a \mid -M^2 + 1 \leq a \leq M^2 - 1, a \text{ odd}\}$$

with $M^2$ elements.

**Example 4.** Quadratic Extension Constellation

Let $d = 3 \pmod{4}$. Then

$$A_{\text{QE}}(SNR) = \left\{a + \frac{1 + \sqrt{-d}}{2}b \mid -M + 1 \leq a, b \leq M - 1, a, b \text{ odd}\right\}$$

with $M^2$ elements.
Definition 13. We will call a ST code $X$ $A$-linear, if every entry $X_{ij}$ of each code matrix $X$ is of the form $X_{ij} = \sum_{k=1}^{m} c_{ijk} \alpha_{ijk}$, $c_{ijk} \in A$, for some $m$, and for some fixed set of complex numbers $\{\alpha_{ijk}\}$ which do not vary with SNR.

The space-time codes that we will consider are assumed to be $A$-linear over a scalably dense alphabet $A$. Further, we will assume our space-time code to be full rate over this scalably dense alphabet, where the full-rate property is as defined below.

Definition 14. An $A$-linear space-time code $X$ is said to be full-rate if for every SNR, the size of the code satisfies

$$|X| = |A|^{ntT}.$$  

4.4.2 Signal Energy

Let $A$ be scalably dense. For an $A$-linear square $(nt \times nt)$ ST code $X$, we have for every $X \in X$,

$$||X||^2_F \leq \sum_{i,j=1}^{nt} \sum_{k=1}^{m} |c_{ijk}|^2 |\alpha_{ijk}|^2 \leq M^2.$$  

since $\alpha_{ijk}$ are fixed and independent of SNR and $|c_{ijk}|^2 \leq M^2$ (for $A$ is scalably dense).

From (4.3), it follows that a valid choice of $\theta^2$ (upto SNR exponent) is

$$\theta^2 = \frac{\text{SNR}}{M^2}.$$  

(4.4)

4.4.3 D-MG Optimality

We are now ready to state and prove the D-MG optimality of a particular class of minimal-delay ($T = nt$) space-time constructions that have the NVD property.

Theorem 8. Let $A$ be a scalably dense alphabet and $X$ be an $nt \times nt$ (square) space-time code which
• is $\mathcal{A}$-linear,
• is full rate over $\mathcal{A}$, and
• has the non-vanishing determinant property.

Then $\mathcal{X}$ achieves the diversity-multiplexing gain tradeoff for any number of receive antennas.

Proof. Let the cardinality of the base alphabet

$$|\mathcal{A}| = M^2.$$ 

In accordance with the formulation of the D-MG tradeoff, the transmission rate of the space-time code $\mathcal{X}$ is $R = r \log \text{SNR}$ bits/channel use. Using the fact that $\mathcal{X}$ is full-rate over $\mathcal{A}$, we obtain,

$$|\mathcal{X}| = \text{SNR}^r n_t = (M^2)^{n_t}$$

$$\Rightarrow M^2 = \text{SNR}^{\frac{r}{n_t}}.$$ 

(4.4) gives us that

$$\theta^2 = \text{SNR}^{1 - \frac{r}{n_t}}.$$ 

Let $\Delta X$ denote the difference of any two codeword matrices from $\mathcal{X}$. The non-vanishing determinant property of $\mathcal{X}$ is equivalent to

$$\min_{\Delta X} \det(\Delta X \Delta X^\dagger) \geq \text{SNR}^0.$$ 

The parameter $\delta$ (defined in Theorem 7) for $\mathcal{X}$ can therefore be evaluated as

$$\text{SNR}^\delta = \min_{\Delta X} \det[(\theta \Delta X)(\theta \Delta X)^\dagger]$$

$$= (\theta^2)^{n_t} \min_{\Delta X} \det(\Delta X \Delta X^\dagger)$$

$$\Rightarrow \delta = n_t - r.$$
Using Theorem 7, we have the desired result.

\section*{4.5 D-MG Optimal Schemes from the Literature}

Having reviewed a sufficient condition that ensures optimality on the D-MG tradeoff, the natural direction to pursue is to see whether any existing schemes satisfy this condition. One scheme which has already been mentioned to be optimal on the D-MG tradeoff is the Alamouti scheme for $n_r = 1$. Among the schemes to be presented in this section, the D-MG tradeoff optimality of a few schemes were proven through independent approaches earlier while the optimality of the rest follow from the sufficient condition presented in [54] (reproduced in Theorem 7 of this thesis). This section will also reveal that space-time schemes which satisfy Theorem 7 existed previously in the literature only for a few sporadic values of number of transmit antennas.

\subsection*{4.5.1 Tilted QAM codes of Yao and Wornell}

Yao and Wornell were the first to present a $(2 \times 2)$ space-time code construction that achieves the D-MG tradeoff for the case of two receive antennas [43, 42]. The constructional details are as follows.

Each codeword in the Yao-Wornell space-time code is of the form

\[
\begin{bmatrix}
  x_1 & y_2 \\
  y_1 & x_2
\end{bmatrix},
\]

where

\[
\begin{bmatrix}
  x_1 \\
  x_2
\end{bmatrix} = M_{\theta_1} u,
\]

\[
\begin{bmatrix}
  y_1 \\
  y_2
\end{bmatrix} = M_{\theta_2} v
\]

with $u, v \in \mathcal{A}_{QAM}^2$ and

\[
M_{\theta_i} = \begin{bmatrix}
  \cos(\theta_i) & -\sin(\theta_i) \\
  \sin(\theta_i) & \cos(\theta_i)
\end{bmatrix}.
\]
It is shown that setting

$$\theta_1 = \frac{1}{2} \arctan \left( \frac{1}{2} \right), \quad \theta_2 = \frac{1}{2} \arctan(2)$$

maximizes the minimum determinant of the difference of two space-time code matrices. The minimum determinant is shown to exceed $\frac{1}{2\sqrt{5}}$. Thus, the Yao-Wornell construction is endowed with the non-vanishing determinant property. Yao and Wornell have the distinction of being the first to use the non-vanishing determinant to establish the fact that their code achieves the upper bound on the D-MG tradeoff, see Figure 4.1 (for details, see [43, 42]). By achieving the upper bound on D-MG tradeoff for $T = 2$ ($< nt + nr - 1 = 3$), the authors also establish that the upper bound on D-MG tradeoff provided by Zheng and Tse is exact when $T = nt = nr = 2$.

Figure 4.1: D-MG Tradeoff of the Yao-Wornell Code ($nt = nr = T = 2$)
4.5.2 Dayal-Varanasi Construction

In [44], Dayal and Varanasi consider a construction of a space-time code $\mathcal{X}$ with the same parameters $n_t = T = 2$ given by

$$\mathcal{X}(\theta, \phi) = \begin{cases} \begin{bmatrix} x_1 & \phi \frac{1}{2} y_1 \\ \phi \frac{1}{2} y_2 & x_2 \end{bmatrix} \end{cases},$$

where $|\phi| = 1,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = M_{\theta} u, \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = M_{\theta} v,$$

with $u, v \in \mathcal{A}_{\text{QAM}}^2$ and

$$M_{\theta} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}.$$

This construction draws from earlier constructions in [28, 32] and special instances of this construction yield the earlier constructions by Damen et al. [28] and El Gamal and Damen [32] (when specialized to the case of two antennas). The authors of [44] show that choosing $\phi = -i$ and $\theta = \frac{1}{2} \arctan(2)$ optimizes the coding gain. As in the case of the Yao-Wornell construction, the optimized code of Dayal and Varanasi is endowed with the NVD property. Therefore, the D-MG optimality of the Dayal-Varanasi construction follows along the same lines as that of the Yao-Wornell code [43, 42]. The coding gain of the Dayal-Varanasi code, relevant for low values of SNR, is shown to improve upon that of the Yao-Wornell code.

4.5.3 Golden Code Construction

A third construction of $n_t = n_r = T = 2$ space-time codes that achieve the D-MG tradeoff$^1$ is the “Golden code” construction of Belfiore et al. [39], so called because of the appearance of the Golden number $\zeta = \frac{1+\sqrt{5}}{2}$ in the construction. The Golden code

$^1$Optimality with respect to the D-MG gain tradeoff is not discussed in either [44] or [39]. Optimality of the Golden code construction is pointed out in [53] as a consequence of the sufficient condition given in Theorem 7.
consists of all code matrices $X$ of the form
\[
X = \frac{1}{\sqrt{5}} \begin{bmatrix}
\alpha(a + \zeta b) & \alpha(c + \zeta d) \\
\gamma \alpha(c + \zeta d) & \alpha(a + \zeta b)
\end{bmatrix}
\]
where $\gamma = i$, $a, b, c, d \in A_{\text{QAM}}$ and $\alpha = 1 + i(1 - \zeta)$, $\overline{\alpha} = 1 + i(1 - \overline{\zeta})$ denotes the complex conjugates of $\alpha$ and $\overline{\zeta} = 1 - \zeta$. The minimum determinant of the Golden code is also non-vanishing.

### 4.5.4 LAST Codes

In [48], El Gamal, Caire and Damen consider a lattice-based construction of space-time block codes and call these codes LAST codes. In this construction, the columns of a space-time code matrix are stacked vertically to result in a column vector with $n_t T$ components. Thus a code matrix $X$ in the space-time code $\mathcal{C}$ is identified with a $(n_t T \times 1)$ vector $\underline{x}$. The construction calls for a lattice $\Lambda_c$ and a sublattice $\Lambda_s$. The message symbols are mapped in $1-1$ fashion onto coset representatives $\{\underline{c}\}$ of the subgroup $\Lambda_s$ of $\Lambda_c$ that lie within the fundamental region $V_s$ of the sublattice $\Lambda_s$. Thus the fundamental region of the sublattice serves as a shaping region for the lattice itself. The transmitted vector $\underline{x}$ is then given by
\[
\underline{x} = \underline{c} - \underline{u} \pmod{\Lambda_s},
\]
where $\underline{u}$ is a pseudorandom “dither” vector chosen with uniform probability from $V_s$. The dither is assumed to also be known to the receiver. The nested lattice $\Lambda_s \subseteq \Lambda_c$ is drawn from an ensemble of lattices having good “covering” properties, an example of which can be found in a paper by Loeliger [58]. It is shown that this ensemble of lattices contains a lattice such that the resultant space-time code, when suitably decoded using generalized minimum Euclidean distance lattice decoding, achieves the D-MG tradeoff for all $T \geq n_t + n_r - 1$. In actual code construction, a lattice drawn at random from the ensemble of lattices is used. A principal advantage of LAST codes is that in comparison to random Gaussian codes, the decoding is much simpler and does not require searching.
over the entire codebook.

4.5.5 Cyclotomic Space-Time Codes of Wang and Xia

In [45], the authors present space-time code designs based on cyclotomic lattices. An optimality theorem for single layer (diagonal) cyclotomic space-time codes is presented first, where the optimality is in the sense that for fixed mean transmission power, the minimum determinant of the code is maximized, or equivalently, for a fixed minimum determinant, the mean transmission power is minimized. These single-layer space-time codes however have only unit symbol rate.

The authors are able to construct full-rate (multi-layer) space-time codes with full-diversity only for the case when the number of transmit antennas is either two or three. Both these designs are endowed with the non-vanishing determinant property. Although the authors in [45] do not discuss the D-MG tradeoff of these constructions, the optimality of the two and three transmit antenna constructions follows from Theorems 7 and 8.

4.5.6 Space-time Codes Derived from Division Algebras

The algebraic framework behind the Golden code in Section 4.5.3 is what is known as a division algebra. The D-MG optimality of the Golden code leads us to suspect that space-time codes with NVD for larger number of transmit antennas derived from “suitable” division algebras might also be endowed with “good” D-MG tradeoff, if not optimal. We will relegate the discussion regarding the existence of such space-time codes for arbitrary number of transmit antennas and their D-MG tradeoff to the chapters to come.
Chapter 5

Review of Relevant Algebra and Number Theory

In the chapters to come, concepts from abstract algebra and algebraic number theory will play a critical role in our design of space-time block-codes. We will assume the reader to be familiar with some basic concepts from algebra and Galois theory (Good reference texts for algebra and Galois theory include the ones by Herstein [14] and by Dummit and Foote [7]).

The purpose of this chapter is to first, briefly refresh the reader’s memory with some definitions from abstract algebra and then, to provide an overview of the basics of algebraic number theory. Our coverage of algebraic number theory will be along the lines of the excellent introductory text by Pollard and Diamond [8]. The final section on ideal factorization is based on the exposition by Ribenboin [4]. The material presented here is tailored to provide an understanding of those concepts which we will make use of in our space-time code constructions and is by no means exhaustive. We recommend [4, 9] for further reading in algebraic number theory.
5.1 Brief Review of Groups, Rings and Fields

The group is a fundamental building block of abstract algebra, and is the first object to be introduced in most texts on the subject.

**Definition 15.** A non-empty set of elements $G$ is said to form a group if in $G$ there is defined a binary operation, called the product and denoted by $\cdot$, such that

1. $a, b \in G$ implies that $a \cdot b \in G$ (closed).

2. $a, b, c \in G$ implies that $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (associative law).

3. There exists an element $e \in G$ such that $a \cdot e = e \cdot a = a$ for all $a \in G$ (the existence of an identity element in $G$).

4. For every $a \in G$ there exists an element $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$ (the existence of inverses in $G$).

Following is a highly special, but very important class of groups.

**Definition 16.** A group $G$ is said to be abelian (or commutative) if for every $a, b \in G$, $a \cdot b = b \cdot a$.

The next level of abstraction leads us to a ring, which is a set of elements with two binary operations defined on them.

**Definition 17.** A non-empty set $R$ is said to be an (associative) ring if in $R$ there are defined two operations, denoted by $+$ and $\cdot$ respectively, such that for all $a, b, c \in R$:

1. $R$ is an abelian group under the operation $+$.

2. $a \cdot b$ is in $R$.

3. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

4. $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(b + c) \cdot a = b \cdot a + c \cdot a$ (the two distributive laws).
We will not have any occasion to consider the class of non-associative rings. In addition, if there is an element 1 in $R$ such that $a \cdot 1 = 1 \cdot a = a$ for every $a \in R$, we shall describe $R$ as a ring with unit element. If the operation $\cdot$ is commutative in $R$, i.e., $a \cdot b = b \cdot a$ for every $a, b \in R$, then we call $R$ a commutative ring. For simplicity of notation, we shall henceforth drop the dot in $a \cdot b$ and merely write this product as $ab$.

The following abstract structure, a special type of a ring, will be of prime importance to us.

**Definition 18.** A ring is said to be a division ring if its non-zero elements form a group under multiplication.

Notice that commutativity is not part of the definition of a division ring. We define the centre of a ring as follows.

**Definition 19.** The centre of a ring $R$ is the set of all elements which commute with every element in $R$, i.e., \{ $z \in R \mid zr = rz \ \forall \ r \in R$ \}.

Adding the property of commutativity to a division ring leads us to a higher level of abstraction, defined below.

**Definition 20.** A field is a commutative division ring.

The following remark is a consequence of the above definitions.

**Remark 2.** The centre of a division ring is a field.

The next algebraic object which is of importance to us is an algebra.

**Definition 21.** An associative ring $A$ is called an algebra over a field $F$ if $A$ is a vector space over $F$ such that for all $a, b \in A$ and $\alpha \in F$, $\alpha(ab) = (\alpha a)b = a(\alpha b)$.

All background required to define our object of interest is now in place. We wrap up this section with the definition of a division algebra.

**Definition 22.** A division algebra is simply a ring in which every non-zero element has a multiplicative inverse. The word “algebra” refers to the fact that such a ring is naturally a vector space over its centre.
5.2 Basics of Algebraic Number Theory

5.2.1 Number Fields and Extensions

By a number field $\mathbb{F}$, we will mean a field that contains the set of all rational numbers $\mathbb{Q}$. Note that the set of all rational numbers is a field in its own right as well. This is an example of what is known as a field extension. In general, any field $\mathbb{K}$ containing a field $\mathbb{F}$ is called an extension of $\mathbb{F}$. In the sections to follow, usage of the term “field” will always refer to a “number field”.

We say a number $\theta$ is algebraic over a field $\mathbb{F}$ if it satisfies a non-trivial polynomial equation

$$a_nx_n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

with coefficients in $\mathbb{F}$. $\theta$ need not belong to $\mathbb{F}$. For example, $\sqrt{2}$ satisfies $x^2 - 2 = 0$ over $\mathbb{R}$, the field of real numbers, but $\sqrt{2}$ is not in $\mathbb{R}$.

Suppose now that $\theta$ is algebraic over $\mathbb{F}$ and consider all polynomials over $\mathbb{F}$ of which $\theta$ is a root. Denote the polynomial of lowest degree having the leading coefficient to be 1 as $p(x)$. This polynomial $p(x)$ is called the minimal polynomial of $\theta$. $p(x)$ is clearly irreducible (if not, the minimum degree assumption will be violated) and it can be shown that $p(x)$ is also unique. If the degree of $p(x)$ is $n$, then $\theta$ is said to be of degree $n$ over $\mathbb{F}$.

Let $\theta_1, \theta_2, \ldots, \theta_n$ be the roots of $p(x)$, where $\theta_1 = \theta$. That these $n$ roots are distinct can be shown from the fact that $p(x)$ is irreducible. We call them the conjugates of $\theta$ over $\mathbb{F}$. When $\mathbb{F} = \mathbb{Q}$, we omit reference to the field and simply say conjugates.

If $\theta$ is algebraic over $\mathbb{F}$, then $\mathbb{K} = \mathbb{F}(\theta)$ is defined to be the smallest field containing both $\mathbb{F}$ and $\theta$. $\mathbb{K}$ is called a simple algebraic extension of $\mathbb{F}$. Analogously, if $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ are numbers algebraic over $\mathbb{F}$, then the smallest field $\mathbb{K} = \mathbb{F}(\alpha_1, \ldots, \alpha_n)$ containing $\mathbb{F}$ and the $\alpha_i$ is called a multiple algebraic extension of $\mathbb{F}$. Further, it can be shown that every multiple algebraic extension of $\mathbb{F}$ is a simple algebraic extension.

A number $\theta$ is said to be an algebraic number if it is algebraic over the field $\mathbb{Q}$ of rational numbers. Numbers which are not algebraic are transcendental.
A set of numbers \( \{ \beta_1, \beta_2, \ldots, \beta_n \} \) in \( K \) is said to form a basis for \( K \) over \( F \) if for each element \( \beta \) in \( K \), there exists a unique set of numbers \( d_1, d_2, \ldots, d_n \) in \( F \) such that

\[
\beta = d_1 \beta_1 + d_2 \beta_2 + \cdots + d_n \beta_n.
\]

Observe that the \( \beta_i \) are linearly independent. It turns out that although there may exist several bases for \( K \) over \( F \), the number of elements \( n \) in each of these bases is the same. \( n \) is called the degree of \( K \) over \( F \), and \( K \) is called a finite extension of degree \( n \) over \( F \). We write \( n = (K|F) \).

It can be shown that \( 1, \theta, \ldots, \theta^{n-1} \) is a basis for \( F(\theta) \) over \( F \), where \( n \) is the degree of \( \theta \) over \( F \).

There exist intimate connections between finite and algebraic extensions. If \( K \) is a finite extension over \( F \), then every element \( \alpha \) of \( K \) is algebraic over \( F \). This suggests a method for making an extension which is not finite. Let \( F = \mathbb{Q} \), the field of rational numbers and let \( \xi \) be a transcendental number. Then the field \( K = F(\xi) \) is not a finite extension. Further, it can also be shown that an extension \( K \) of \( F \) is finite if and only if it is a simple algebraic extension.

The above discussion leads us to what is known as an algebraic number field. An algebraic number field is any finite (hence simple) extension of \( \mathbb{Q} \).

### 5.2.2 Algebraic Integers and Integral Bases

Henceforth, for clarity of presentation, we will call the set of integers in \( \mathbb{Q} \) as rational integers. Our motivation in this section is to investigate the analogues of the rational integers in an arbitrary number field \( F \). For example, a natural analogous definition of “integers” in the field \( \mathbb{Q}(i) = \{a + ib \mid a, b \in \mathbb{Q}, i = \sqrt{-1}\} \) are the Gaussian integers \( \mathbb{Z}(i) = \{a + ib \mid a, b \in \mathbb{Z}\} \). In arbitrary number fields, the following definition holds. An algebraic number is an algebraic integer if it’s minimal polynomial has only rational integers as coefficients.

The set of algebraic integers have an interesting algebraic structure. If \( K = \mathbb{Q}(\theta) \) is
an algebraic number field, then the set of integers in it form a ring, which we will call the ring of integers of \( \mathbb{K} \), denoted as \( \mathcal{O}_K \).

**Example 5.** Consider the quadratic field \( \mathbb{Q}(\sqrt{D}) \), where \( D \) is a rational integer free of square factors. The algebraic integers of \( \mathbb{Q}(\sqrt{D}) \) consist of these classes:

1. all numbers of the form \( l + m\sqrt{D} \), where \( l \) and \( m \) are rational integers, and
2. if \( D \equiv 1 \pmod{4} \), but not otherwise, all numbers of the form \( (l + m\sqrt{D})/2 \), where \( l \) and \( m \) are odd.

If \( \alpha \) is in \( \mathcal{O}_K \) and \( \mathbb{K} \) is of degree \( n \) over \( \mathbb{Q} \), then \( \alpha \) has \( n \) conjugates \( \alpha_1, \ldots, \alpha_n \) for \( \mathbb{K} \).

We define the norm of \( \alpha \), written as \( N(\alpha) \) or \( N\alpha \), by

\[
N\alpha = \alpha_1\alpha_2\ldots\alpha_n.
\]

Note that \( N\alpha \) depends on the field \( \mathbb{K} \). For example, \( N2 = 2 \) in \( \mathbb{Q} \), but \( N2 = 4 \) in \( \mathbb{Q}(i) \).

The next important concept is that of an integral basis. A set of integers \( \alpha_1, \ldots, \alpha_s \) is called an integral basis of \( \mathbb{K} \) if every integer \( \alpha \) in \( \mathbb{K} \) can be written uniquely in the form

\[
\alpha = b_1\alpha_1 + \cdots + b_s\alpha_s,
\]

where the \( b_i \) are rational integers. It can be shown that an integral basis is a basis, from which it follows that \( s = n \), the degree of extension of the field. Also, every algebraic number field has at least one integral basis.

**Example 6.** This follows up on Example 5. An integral basis for \( \mathbb{Q}(\sqrt{D}) \) is \( \{1, \sqrt{D}\} \) if \( D \not\equiv 1 \pmod{4} \) and \( \{1, (1 + \sqrt{D})/2\} \) if \( D \equiv 1 \pmod{4} \).

### 5.2.3 Ideals in an Algebraic Number Field

Let \( \mathbb{K} \) be an algebraic number field. A set \( A \) of integers in \( \mathcal{O}_K \) is an ideal in \( \mathcal{O}_K \) if, together with any pair of integers \( \alpha \) and \( \beta \) in \( A \), the set also contains \( \lambda\alpha + \mu\beta \) for any integers \( \lambda \) and \( \mu \) in \( \mathcal{O}_K \). The ideal \( A \) “swallows up” multiplication by arbitrary integers.
in $\mathcal{O}_K$. A set of integers $\omega_1, \ldots, \omega_n$ in $A$ is said to form a basis for $A$ if every element $\alpha$ of $A$ can be uniquely represented in the form

$$\alpha = c_1\omega_1 + \cdots + c_n\omega_n,$$

where the $c_i$ are rational integers and $n$ denotes the degree of extension of $\mathbb{K}$.

An ideal $A$ is said to be generated by $\alpha_1, \ldots, \alpha_t$, written $A = (\alpha_1, \ldots, \alpha_t)$, if $A$ consists of all sums $\sum_{i=1}^{t} \lambda_i \alpha_i$, where the $\lambda_i$ are integers, not necessarily rational, in $\mathcal{O}_K$. Obviously, if $\omega_1, \ldots, \omega_n$ is a basis for $A$, then $A = (\omega_1, \ldots, \omega_n)$; but if $A = (\alpha_1, \ldots, \alpha_t)$, the $\alpha_i$ do not necessarily form a basis for $A$.

An ideal $A$ is principal if it is generated by a single integer, i.e., $A = (\alpha)$. We use the notation $\alpha\mathcal{O}_K$ and $(\alpha)$ interchangeably to mean the principal ideal generated by $\alpha \in \mathcal{O}_K$.

### 5.2.4 Properties of Ideals

By the product $AB$ of ideals of the ideals $A = (\alpha_1, \ldots, \alpha_s)$ and $B = (\beta_1, \ldots, \beta_t)$ in $\mathcal{O}_K$ we mean the ideal

$$AB = (\alpha_1\beta_1, \ldots, \alpha_i\beta_j, \ldots, \alpha_s\beta_t)$$

in $\mathcal{O}_K$ generated by all products $\alpha_i\beta_j$.

We will say that $A$ divides $B$, written as $A|B$, if an ideal $C$ exists so that $B = AC$, alternatively, if $B \subseteq A$. $A$ is then called a factor of $B$.

The norm of an ideal $A$ of the ring of integers $\mathcal{O}_K$ of an algebraic number field $\mathbb{K}$ is defined to be the cardinality of the additive quotient group $\mathcal{O}_K/A$. We will denote the norm of the ideal $A$ as $\|A\| = |\mathcal{O}_K/A|$. If $A = (a)$ for some $a \in \mathcal{O}_K$, then $\|A\| = |N_{\mathbb{K}/\mathbb{Q}}(a)|$.

We will now move towards establishing a theory of unique factorization for ideals similar to that which is known for the rational integers. The ideals which take over the function of the prime rational integers are naturally those ideals $P$ which have no factors except $P$ and $(1)$ (the ideal generated by 1 is the entire ring of integers of the number field). To be precise, an ideal $P$ is said to be irreducible if it has no factors except $P$ and $(1)$.
We will now introduce two other kinds of ideals. An ideal \( A \) is \textit{maximal} if it is included in no larger ideal except (1), the entire ring of integers. An ideal \( P \) different from (0) or (1) is \textit{prime} if it has the following property: whenever a product of integers \( \gamma \delta \) is in \( P \), so is either \( \gamma \) or \( \delta \). When considering ideals in the ring of integers of number fields (which is the only case with which we will work), it turns out that an ideal \( P \) different from (0) or (1) is maximal if and only if it is prime.

\subsection{5.2.5 Factorization of Ideals}

Following is the \textit{fundamental theorem of ideal theory}.

\textbf{Theorem 9.} Every ideal not (0) or (1) can be factored into the product of irreducible ideals. This factorization is unique except for the order of the factors.

It turns out that the ring of algebraic integers of an arbitrary number field is an example of a Dedekind domain. In such a scenario, prime and irreducible ideals are one and the same, which is shown by the following theorem [4].

\textbf{Theorem 10.} Let \( A \) be a Dedekind domain. An ideal of \( A \) is irreducible if and only if it is prime.

Therefore, in our cases of interest, we can conclude from the fundamental theorem of ideal theory that any non-trivial ideal can be factored uniquely (up to order) into a product of prime ideals.

We will now embark upon the study of extension of ideals. Our exposition will closely follow Figure 5.1.

Let \( \mathbb{K} \) be an algebraic number field, \( \mathbb{L}/\mathbb{K} \) be an extension of finite degree \( n \), and let \( \mathcal{O}_\mathbb{K} \) (respectively, \( \mathcal{O}_\mathbb{L} \)) be the rings of algebraic integers of \( \mathbb{K} \) (respectively, \( \mathbb{L} \)). Let \( P \) be any non-zero ideal of \( \mathcal{O}_\mathbb{K} \). Our aim is to relate the decomposition of \( P \) into prime ideals of \( \mathcal{O}_\mathbb{K} \), with the decomposition into prime ideals of \( \mathcal{O}_\mathbb{L} \), of the ideal of \( \mathcal{O}_\mathbb{L} \) generated by \( P \). From the fundamental theorem of ideal theory, we may, without loss of generality, restrict our attention to the case when \( P \) is a prime ideal in \( \mathcal{O}_\mathbb{K} \).
Let \( P \) be any non-zero prime ideal of \( \mathcal{O}_K \). Let \( Q = P\mathcal{O}_L \) denote the ideal in \( \mathcal{O}_L \) generated by \( P \); it consists of all sums \( \sum_{i=1}^{m} l_i x_i \) with \( m \geq 1 \), \( l_i \in \mathcal{O}_L \), \( x_i \in P \) for all \( i = 1, \ldots, m \). From the fundamental theorem of ideal theory, \( Q \) can be written in a unique way as a product of powers of prime ideals of \( \mathcal{O}_L \):

\[
Q = P\mathcal{O}_L = \prod_{i=1}^{g} Q_i^{e_i} \tag{5.1}
\]

Following [37], we will say that each prime ideal \( Q_i \) lies over the prime ideal \( P \), or \( P \) lies under \( Q_i \). Notice that \( Q_i \cap \mathcal{O}_K = P \). This means that \( Q_i \) cannot be the prime factor of any other ideal \( P_i\mathcal{O}_L \), where \( P_i \) is a prime ideal in \( \mathcal{O}_K \) different from \( P \).

With reference to (5.1), we introduce the following terminology. \( g \) is called the decomposition number of \( P \) in the extension \( L|K \). If \( Q \) is a prime ideal, then we say that \( P \) is inert in \( L|K \). For every \( i = 1, \ldots, g \), \( e_i \) is called the ramification index of \( Q_i \) in \( L|K \). We will also use the notation \( e(Q_i|P) \) for the ramification index \( e_i \). If \( e_i = 1 \), we say that \( Q_i \) is unramified in \( L|K \). It turns out that \( \mathcal{O}_L/Q_i \) is a vector space over the field \( \mathcal{O}_K/P \) of dimension \( [\mathcal{O}_L/Q_i : \mathcal{O}_K/P] \leq [L : K] \). The dimension \( f_i \) of \( \mathcal{O}_L/Q_i \) over \( \mathcal{O}_K/P \) is called
the inertial degree or residual degree of $Q_i$ in $L|K$. We use the notation $f_i = f(Q_i|P)$.

The norms of the two ideals $Q_i$ and $P$ are related as $\|Q_i\| = \|P\|^f_i$.

A notable simplification arises in the important case when $L|K$ is a Galois extension. The ramification indices $e_1 = \cdots = e_g := e$ and the inertial degrees $f_1 = \cdots = f_g := f$.

The following fundamental relation between the above parameters holds:

$$n = \sum_{i=1}^{g} e_i f_i. \quad (5.2)$$

We note explicitly the particular case when $K|F$ is Galois:

$$n = ef g. \quad (5.3)$$

Let $p$ be the unique prime in $\mathbb{Z}$ that lies under $P$. The transitivity of the ramification index and inertial degrees are as follows,

$$e(Q_i|p) = e(Q_i|P) \cdot e(P|p), \quad (5.3)$$

$$f(Q_i|p) = f(Q_i|P) \cdot f(P|p). \quad (5.4)$$

We call the primes $p, P, Q_i$ that lie one above the other as a prime triplet $(p, P, Q_i)$ [37].
Chapter 6

Space-Time Codes from Division Algebras - An Overview

In this chapter, we will present an overview of space-time block code construction from division algebras. We will first motivate this approach by elucidating the natural connection between division algebras and full-rank space-time codes and will then proceed to give the algebraic details of the construction. An excellent overview of this area can be found in [36]. In this chapter, we will confine ourselves to the minimum-delay case of $T = n_t$, i.e., each space-time codeword is a square matrix. It will be shown in the beginning of the chapter to follow that square space-time codes derived from “suitable” division algebras endowed with the NVD property satisfy Theorem 8 and hence meet the optimal D-MG tradeoff. Construction of such division algebras will be the focus of Chapter 7 while generalization to rectangular designs will follow in Chapter 8.

6.1 Division Algebras and Space-Time Codes

As we saw in Chapter 2, it was shown in [16, 17] that the minimum rank of the difference of any two codeword matrices from a space-time code $\mathcal{X}$ is an indication of how the pairwise error probability decays with the SNR. If we denote $\Delta X = X_1 - X_2$ as the difference between any two distinct codeword matrices $X_1, X_2 \in \mathcal{X}$, and $\text{rank}(\Delta X) = \nu$,
then it was shown in [16, 17] that the pairwise error probability

\[ PEP(X_1 \rightarrow X_2) \propto \frac{1}{\text{SNR}^{\nu \cdot r}}. \]

Also, recall from Chapter 4 that in order that we achieve the maximum diversity point of \( d^*(0) = n_t n_r \) on the D-MG tradeoff, it is imperative that we employ a full-rank space-time code (atleast for the segment \( 0 \leq r < 1 \)). Further, the sufficient condition from [54], reviewed in Chapter 4 assumes that the space-time scheme is a full-rank design. Our interest will therefore be in constructing a space-time code \( \mathcal{X} \) with full minimum difference rank \( \nu = \min\{n_t, T\} \). Furthermore, as mentioned before, we will confine our interest in this chapter to the case when \( n_t = T \), i.e., the minimum delay case.

An \((n \times n)\) full-rank space-time code is nothing but a set of matrices which have the property that the difference of any two matrices is of full-rank. Although it is not necessary that these matrices themselves be of full-rank, we will find it convenient to construct our space-time code out of such a set. Consider the set of all \((n \times n)\) invertible (i.e., full-rank) matrices over the field of complex numbers \( \mathbb{C} \). Let us first examine the algebraic structure of such a set. It is immediately evident that this infinite set is not closed under matrix addition. However, this set forms a non-abelian group under the operation of matrix multiplication.

A subset of the set of all \((n \times n)\) matrices that have the desired property of full difference-rank can be derived from a division algebra as follows [36]. Recall from Chapter 5 that a division algebra is a ring in which each non-zero element has a multiplicative inverse. What distinguishes a division algebra from a field is that the multiplication in a division algebra need not be commutative (the astute reader will notice that this is in accordance with the non-commutativity of matrix multiplication). The properties of the division algebra that we will exploit are those of invertibility and closure under addition. This guarantees that the difference of two elements in a division algebra is also an element of the division algebra, and is hence invertible. We would like to carry over this property to a finite set of \((n \times n)\) invertible matrices. The principle behind the
Chapter 6. Space-Time Codes from Division Algebras - An Overview

The construction of space-time codes from division algebras is given in the following theorem [36] and is illustrated in Figure 6.1.

**Theorem 11.** Let $f : D \to M_n(\mathbb{F})$ be a ring homomorphism from a division algebra $D$ to the set of all $n \times n$ matrices over some field $\mathbb{F}$. If $E$ is any finite subset of the image of $D$ under this map, then $E$ will have the property that the difference of any two elements in it will be of full rank.

![Figure 6.1: Space-Time Codes from Division Algebras](image)

The theorem essentially means the following: since the homomorphism $f$ is a map that preserves both the addition and multiplication operations, the difference of any two matrices is guaranteed to have a multiplicative inverse (and hence has full rank) since it is the image of some element in a division algebra (where multiplicative inverses exist for all elements).

The space-time code construction makes use of the representation as square matrices, of elements in a division algebra. Such a representation arises when one considers an element $d$ in a division algebra as a linear transformation corresponding to multiplication.
of elements in the division algebra by \( d \). We illustrate this concept by considering the field of complex numbers as a (commutative) division algebra over the reals.

**Example 7.** *(Matrix Representation of a Complex Number)* Consider the linear operator \( A : \mathbb{C} \to \mathbb{C} \), given by

\[
A(u + iv) = (a + ib)(u + iv). \tag{6.1}
\]

By regarding \( \mathbb{C} \) as a vector space over \( \mathbb{R} \), we arrive at the matrix representation

\[
A = \begin{bmatrix}
a & -b \\
b & a
\end{bmatrix}. \tag{6.2}
\]

Since the mapping:

\[
(a + ib) \to A = \begin{bmatrix}
a & -b \\
b & a
\end{bmatrix} \tag{6.3}
\]

is a ring homomorphism, as long as \( (a + ib) \neq 0 \), \( A \) is guaranteed to have an inverse.

Our next example will involve a non-commutative division algebra, the Hamilton’s Quaternions. We will consider the representation of quaternions as matrices over the complex numbers.

**Example 8.** *(Matrix representation of Quaternions)* The canonical example of a division algebra is the ring \( D = \mathbb{R}(e, i, j, k) \) of quaternions over the real numbers \( \mathbb{R} \) where \( e \) is the identity element with \( e^2 = 1 \) and \( i, j, k \) are elements satisfying

\[
i^2 = j^2 = k^2 = -1, \quad ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \tag{6.4}
\]

Since \( j^2 = -1 \), we can identify \( \mathbb{R}(j) \) with the complex numbers \( \mathbb{C} \), i.e., \( \mathbb{C} = \mathbb{R}(j) \) and \( \mathbb{R}(e, i, j, k) = \mathbb{C}(i) \).
Consider analogously, the linear operator \( A: \mathbb{D} \rightarrow \mathbb{D} \), given by

\[
A([u_1 + ju_2] + i[v_1 + jv_2]) = ([a_1 + ja_2] + i[b_1 + jb_2])([u_1 + ju_2] + i[v_1 + jv_2])
\]

\[
= [a_1 + ja_2][u_1 + ju_2] + i[b_1 + jb_2][u_1 + ju_2]
\]

\[
+ [a_1 + ja_2][v_1 + jv_2] + i[b_1 + jb_2][v_1 + jv_2]
\]

\[
= [a_1 + ja_2][u_1 + ju_2] + (-1)[a_1 - ja_2][v_1 + jv_2]
\]

\[
+ i\{[b_1 + jb_2][u_1 + ju_2] + [a_1 - ja_2][v_1 + jv_2]\}.
\]

By regarding \( D \) as a two-dimensional (right) vector space over \( \mathbb{C} \), \( D = \mathbb{C} + i\mathbb{C} \), we arrive at the matrix representation

\[
A = \begin{bmatrix}
a_1 + ja_2 & -(b_1 - jb_2) \\
b_1 + jb_2 & a_1 - ja_2
\end{bmatrix} = \begin{bmatrix}
a & -b^* \\
b & a^*
\end{bmatrix}.
\]

(6.5)

Here again, as long as \((a_1 + ja_2) + i(b_1 + jb_2) \neq 0\), \( A \) is guaranteed to have an inverse. Notice that the matrix representation has led us to the Alamouti code!

A generalization of the above idea will yield us our space-time code. Before proceeding to give this generalization, we present some structural properties of division algebras in the following section.

### 6.2 Structure of Division Algebras

The centre \( Z(D) \) of a division algebra \( D \) is the set of all elements in \( D \) which commute, i.e.,

\[
\{x \in D \mid xd = dx \ \forall \ d \in D\}
\]

It can be verified that \( Z(D) \) is a field. \( D \) can be considered as a vector space over \( Z(D) \).

If \( \mathbb{F} \) is the centre of a division algebra \( D \), it is known that the dimension \([D : \mathbb{F}]\) is a perfect square \((= n^2, \text{say})\).

A subfield of a division algebra is a field \( \mathbb{L} \) such that \( \mathbb{F} \subset \mathbb{L} \subset D \). \( \mathbb{L} \) can be considered
Figure 6.2: Structure of a Division Algebra

\[
\begin{array}{c|c}
D & \text{Division Algebra} \\
\hline
n & \\
\hline
\mathbb{L} & \text{Maximal Subfield} \\
\hline
n & \\
\hline
\mathbb{F} & \text{Centre} \\
\end{array}
\]

as a subspace of the $\mathbb{F}$-vector space $D$, which gives us that $[\mathbb{L} : \mathbb{F}]$ divides $[D : \mathbb{F}] = n^2$.

It is known that the maximum possible value of $[\mathbb{L} : \mathbb{F}]$ is $n$, and such a subfield is called a maximal subfield of $D$. Figure 6.2 illustrates the above discussion.

### 6.3 Space-Time Codes Arising from Cyclic Division Algebras

We will choose to derive our space-time codes from a specific class of division algebras known as cyclic division algebras (CDA). Cyclic division algebras have a particularly simple structure and a general technique for the construction of a CDA can be found in [1], Proposition 11 of [36], or Theorem 1 of [38]. We review this construction procedure and the method of obtaining space-time codes from CDA.

Let $\mathbb{F}, \mathbb{L}$ be number fields, with $\mathbb{L}$ a finite, cyclic Galois extension of $\mathbb{F}$ of degree $n$. Let $\sigma$ denote the generator of the Galois group $\text{Gal}(\mathbb{L}/\mathbb{F})$. Let $z$ be some symbol that satisfies the relations

\[
\ell z = z \sigma(\ell) \quad \forall \ \ell \in \mathbb{L} \quad \text{and} \quad z^n = \gamma, \quad (6.6)
\]
for some $\gamma \in \mathbb{F}^*$ having the property that the smallest integer $t$ for which $\gamma^t$ is the relative norm $N_{L/F}(\ell)$ of some element $\ell$ in $\mathbb{L}^*$, is $n$, i.e.,

$$\gamma^t \notin N_{L/F}(\mathbb{L}) \text{ for } 1 \leq i < n \text{ and } \gamma^n \in N_{L/F}(\mathbb{L}). \quad (6.7)$$

In a slight abuse of terminology, we shall refer to an element $\gamma$ satisfying this property, as a ‘non-norm’ element. Then a cyclic division algebra $D(\mathbb{L}/\mathbb{F}, \sigma, \gamma)$ with center $\mathbb{F}$ and maximal subfield $\mathbb{L}$ can be constructed by setting

$$D = \mathbb{L} \oplus z\mathbb{L} \oplus \ldots \oplus z^{n-1}\mathbb{L}. \quad (6.8)$$

Notice that the non-commutativity of multiplication in the division algebra is captured by (6.6).

A space-time code $\mathcal{X}$ can be associated to $D$ by selecting the set of matrices corresponding to the matrix representation of elements of a finite subset of $D$. This so-called “left-regular representation” is along the lines in which the Alamouti code was derived as the matrix representation of elements of the quaternion algebra in Example 8. The matrix corresponding to an element $d \in D$ corresponds to the left multiplication by the element $d$ in the division algebra. Let $\lambda_d$ denote this operation, $\lambda_d : D \to D$, defined by

$$\lambda_d(e) = de, \quad \forall \ e \in D.$$  

It can be verified that $\lambda_d$ is a $\mathbb{L}$-linear transformation of $D$. It is evident from (6.8) that $D$ is a right $\mathbb{L}$ vector space of dimension $n$. A natural basis is given by the decomposition in (6.8), we choose the basis $\{1, z, z^2, \ldots, z^{n-1}\}$. A typical element in the division algebra $D$ is $d = \ell_0 + z\ell_1 + \cdots + z^{n-1}\ell_{n-1}$, where the $\ell_i \in \mathbb{L}$. The matrix representation of the
\( \mathbb{L} \)-linear transformation \( \lambda_d \) under this basis can be shown to be [36]

\[
\begin{bmatrix}
\ell_0 & \gamma \sigma(\ell_{n-1}) & \gamma \sigma^2(\ell_{n-2}) & \hdots & \gamma \sigma^{n-1}(\ell_1) \\
\ell_1 & \sigma(\ell_0) & \gamma \sigma^2(\ell_{n-1}) & \hdots & \gamma \sigma^{n-1}(\ell_2) \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
\ell_{n-1} & \sigma(\ell_{n-2}) & \sigma^2(\ell_{n-3}) & \hdots & \sigma^{n-1}(\ell_0)
\end{bmatrix}
\]  

(6.9)

A set of such matrices, obtained by choosing a finite subset of elements in \( D \) constitutes our space-time code \( \mathcal{X} \). Each codeword matrix \( X \in \mathcal{X} \) is of the form shown in (6.9) and is the left regular representation of an element in the division algebra \( D \).

### 6.4 Properties of ST codes Derived from CDA

1. *Determinant lies in the centre \( \mathbb{F} \)*

Space-time codes derived from CDA have the property that the determinant of any codeword matrix is an element of the centre \( \mathbb{F} \). A brief outline of a proof can be found in [38] for some specific values of \( n \). A simple proof valid for all \( n \) is contained in [53] but is very likely known much earlier\(^1\).

While this result appears to be known (see earlier footnote), for the sake of completeness, we reproduce below, the proof of this result for general \( n \).

The division algebra \( D \) is naturally a right vector space over it’s maximal subfield \( \mathbb{L} \) of degree equal to it’s index \( n \). So any element in \( D \) is of the form

\[
\ell(z) = \sum_{i=0}^{n-1} z^i \ell_i, \ \ell_i \in \mathbb{L}.
\]

Let \( \mathfrak{L} \) denote it’s left regular representation and \( \mathfrak{z} \) be that of \( z \in D \). Consider the

\(^1\)We were informed that a proof can also be found in [3].
product

$$\ell(z).z = \sum_{i=0}^{n-1} z^i \ell_i z$$

$$= z \sum_{i=0}^{n-1} z^i \sigma(\ell_i) \quad \{ \because \ell_i z = z \sigma(\ell_i) \}$$

It is evident that the left regular representation of the element represented by
$$\sum_{i=0}^{n-1} z^i \sigma(\ell_i)$$
is nothing but $$\sigma(\mathcal{L})$$. We therefore have that

$$\mathcal{L} 3 = 3 \sigma(\mathcal{L})$$

$$3^{-1} \mathcal{L} 3 = \sigma(\mathcal{L})$$

$$\Rightarrow |\mathcal{L}| = |\sigma(\mathcal{L})| = \sigma(|\mathcal{L}|)$$

Thus the determinant is invariant under $$\sigma$$ and therefore lies in $$\mathbb{F}$$. Note that this
also applies to the determinant of the difference $$\Delta X = X_i - X_j$$ of any two distinct
matrices $$X_i, X_j \in \mathcal{X}$$.

2. Full-rate property

It is evident from an inspection of (6.9) that a ST code matrix transmits $$n$$ symbols
from $$\mathbb{L}$$, which is equivalent to a rate of 1 symbol from $$\mathbb{L}$$ per channel use. Further,
since the field $$\mathbb{L}$$ is an $$n$$ degree extension over the centre $$\mathbb{F}$$, we can write each
element $$l_i \in \mathbb{L}$$ as

$$l_i = \sum_{k=1}^{n} e_{i,k} \beta_k,$$

where $$e_{i,k} \in \mathbb{F}$$ and $$\{\beta_k\}_{k=1}^{n}$$ forms a basis for $$\mathbb{L}/\mathbb{F}$$. Therefore, the space-time code $$\mathcal{X}$$ transmits $$n$$
symbols from $$\mathbb{F}$$ per channel use and is therefore full-rate over $$\mathbb{F}$$.
6.5 The Non-vanishing Determinant Property

The non-vanishing determinant property was introduced in Section 4.3 and the connection between the NVD property and achieving the D-MG tradeoff was given in Section 4.4. In this section, our aim will be to show how space-time codes derived from CDA may be endowed with the NVD property.

One of the steps in constructing a cyclic division algebra involves the identification of a suitable non-norm element $\gamma$ satisfying (6.7), which is in general, not an easy task. In [36], the element $\gamma$ is chosen to be transcendental over $\mathbb{L}$, which ensures that $(\mathbb{L}(\gamma)/\mathbb{F}(\gamma), \sigma, \gamma)$ is a division algebra. However, as was pointed out in [38, 39, 40], this has the disadvantage that the minimum determinant of the space-time code tends to zero as the size of the signal constellation keeps increasing.

The space-time code derived from CDA can be endowed with the NVD property as follows [38, 39, 40]. Let $\mathcal{O}_F$ denote the ring of integers of the field $\mathbb{F}$. Our choice of $\mathbb{F}$ will be such that the desired base alphabet $\mathcal{A} \subset \mathcal{O}_F$. Also, let $\mathcal{O}_L$ be the integral closure of $\mathcal{O}_F$ in $\mathbb{L}$. Let $\beta_i, i = 1, 2, \cdots, n$ be an integral basis for $\mathcal{O}_L/\mathcal{O}_F$ where $n = [\mathbb{L} : \mathbb{F}]$. The desired space-time code is then obtained by restricting the elements $\ell_i$ appearing in (6.9) to be of the form

$$\ell_i = \sum_{k=1}^{n} e_{i,k} \beta_k, \quad e_{i,k} \in \mathcal{A} \subset \mathcal{O}_F. \quad (6.10)$$

Note that the space-time code is now full-rate with respect to the alphabet $\mathcal{A}$. Further, we will choose our non-norm element $\gamma$ to belong to the ring of integers $\mathcal{O}_F$. The above two restrictions ensure that the determinant is an algebraic integer in $\mathbb{L}$, since all entries of the codeword matrix themselves now belong to $\mathcal{O}_L$. We have already shown that the determinant of any codeword matrix is an element of the centre (Section 6.4). Hence, the determinant of any code matrix (and that of the difference of any two codeword matrices) belongs to $\mathbb{F} \cap \mathcal{O}_L = \mathcal{O}_F$.

One final trick remains to ensure that the space-time code is bestowed with the NVD property. If we were to choose our centre $\mathbb{F}$ such that it’s ring of integers $\mathcal{O}_F$ constitutes a discrete lattice, we can then lower bound the determinant, an element of $\mathcal{O}_F$, by
the minimum norm of the lattice. This will ensure that the minimum determinant is bounded away from zero by a constant independent of SNR and is hence non-vanishing. For example, consider the case when $\mathbb{F} = \mathbb{Q}(i)$ or $\mathbb{F} = \mathbb{Q}(j)$, where $i := \sqrt{-1}$ and $j := \exp(i \frac{2\pi}{3})$. The ring of integers are $\mathcal{O}_\mathbb{F} = \mathbb{Z}(i)$ (square lattice) and $\mathcal{O}_\mathbb{F} = \mathbb{Z}(j)$ (hexagonal lattice) respectively. The square and hexagonal lattices are plotted in Figure 6.5.

![Square and Hexagonal Lattices](image)

Figure 6.3: The Square and Hexagonal Lattices

The minimum distance of both the square and hexagonal lattice is 1. Therefore, if $\mathcal{X}$ is a space-time code derived from a division algebra having either $\mathbb{Q}(i)$ or $\mathbb{Q}(j)$ as its centre, we can claim that the minimum determinant

$$\min_{\Delta \mathcal{X}} \det(\Delta \mathcal{X}) \geq 1 \div \text{SNR}^0,$$

where $\Delta \mathcal{X}$ denotes the difference of any two distinct matrices from $\mathcal{X}$. Thus $\mathcal{X}$ is endowed with the NVD property.
6.6 Summary of ST code construction from CDA with NVD

To summarize, in order to construct a space-time code from CDA which is bestowed with the NVD property, we need to accomplish the following tasks:

1. Choose a field $F$ which has the following properties:
   - The desired transmission alphabet is a subset of the ring of integers $\mathcal{O}_F$ of $F$. In most cases of practical interest, $F$ is typically a number field.
   - The ring of integers $\mathcal{O}_F$ constitutes a discrete lattice.

2. Identify a field $L$, which is an $n$ degree cyclic Galois extension of $F$.

3. Identify a non-norm element $\gamma$ (see (6.7)) that belongs to $\mathcal{O}_F$.

4. Construct a cyclic division algebra $D(L/F, \sigma, \gamma)$ with centre $F$ and maximal subfield $L$.

5. Obtain the space-time codeword matrices using the left-regular representation of a finite set of elements of the division algebra $D$ as in (6.9). The finite subset of $D$ is identified in accordance with restricting the elements $\ell_i$ in (6.9) to be of the form specified in (6.10).
Chapter 7

Minimal Delay D-MG Optimal Space-Time Block Codes

The previous chapter presented the theory and intuition behind the construction of space-time codes from division algebras. This chapter will first show that minimal delay space-time codes derived from suitable cyclic division algebras (CDA) having the non-vanishing determinant (NVD) property achieve the diversity-multiplexing gain tradeoff for any number of receive antennas. The optimality is proven by demonstrating that these space-time codes satisfy Theorem 8 presented in Chapter 4. A literature survey reveals that prior to our work, such space-time codes existed only for a few sporadic values of number of transmit antennas. We then present a construction of square space-time codes from cyclic division algebras for arbitrary number of transmit antennas which are endowed with the non-vanishing determinant property, hence providing an explicit construction of space-time block codes that achieve the D-MG tradeoff for every pair of transmit and receive antennas. This is achieved by clearing the two impediments that previously hampered the construction of CDA for all values of the index $n$: construction of cyclic Galois extensions of arbitrary degree and techniques for the identification of non-norm elements. Some work related to reducing the signalling complexity of the constructed space-time codes wraps up this chapter.
7.1 CDA Based ST Codes with NVD are D-MG Optimal

We will choose to construct our space-time codes derived from CDA over the scalably dense $\mathcal{A}_{QAM}$ base alphabet. Observe that $\mathcal{A}_{QAM} \subset \mathbb{Z}(i)$, and hence we will choose the centre of the division algebra $\mathbb{F} = \mathbb{Q}(i)$.

It is clear from the exposition in Chapter 6 that the space-time codes derived from CDA with the NVD property, constructed over scalably dense alphabets satisfy all the conditions of Theorem 8 and hence achieve the D-MG tradeoff. However, as we will see in the ensuing section, such space-time codes were available in the literature prior to our work only for certain sporadic number of transmit antennas due to the twin requirements of construction of cyclic Galois extensions over $\mathbb{Q}(i)$ and identification of non-norm elements $\in \mathbb{Z}(i)$ not being achieved in the general case. We present a solution to both these tasks in Sections 7.3 and 7.4 respectively.

7.2 Previous Constructions of CDA Based ST Codes with NVD

Now that the optimality of CDA based space-time codes with NVD has been established, the natural question that follows is whether such codes already exist in the literature, and if so, for what values of number of transmit antennas. This section presents a review of such codes which existed prior to our work.

As mentioned in Section 4.3, the non-vanishing determinant property was first introduced in [38] as a tool to ensure a good coding gain even when the spectral efficiency increases. However, the authors in [38] were only able to construct codes derived from CDA with NVD for the number of transmit antennas equal to 2, 3 or 4. Non-norm elements were determined through a computer search. The Golden code, introduced in [39], is a $2 \times 2$ CDA based code with NVD in which it is ensured that there is no shaping loss in the transmission alphabet. This translates into a good coding gain, relevant for finite
values of SNR. This concept was generalized to the $3 \times 3$, $4 \times 4$ and $6 \times 6$ cases in [40]. In [41], an additional constraint of uniform average transmitted energy per antenna is imposed and square space-time codes are constructed for the number of transmit antennas being 2, 3, 4 or 6 which are termed as “perfect codes”. This requirement translates into ensuring that the non-norm element $\gamma$ has unit magnitude, which corresponds to a reduction in the average transmission energy.

In a recent advancement, the authors in [37] identify a general principle for determining non-norm elements and construct space-time codes derived from CDA with the NVD property for the number of transmit antennas of the form $2^k$ or $3.2^k$ or $q^k(q-1)/2$, where $q$ is a prime of the form $4k + 3$ and $k$ is an arbitrary integer.

To summarize, the construction of appropriate cyclic Galois extensions and techniques for determining non-norm elements were only available for some sporadic values of $n_t$ prior to our work. In the following section, we overcome these restrictions and present constructions of space-time codes derived from CDA which are endowed with the NVD property for arbitrary number of transmit antennas.

### 7.3 Cyclic Galois extensions

In this section, we provide a systematic means of constructing cyclic Galois extensions of $\mathbb{Q}(\iota)$, of arbitrary degree $n = 2^m \prod_{i=2}^{r} p_i^{n_i}$. $p_i$ an odd prime. Figure 7.1 provides an overview of our construction. As a first step, we will construct cyclic Galois extensions over $\mathbb{Q}$.

#### 7.3.1 Cyclic Galois extensions over $\mathbb{Q}$

We will use $\omega_m$ to denote a complex, primitive $m^{th}$ root of unity. We will construct cyclic Galois extensions over $\mathbb{Q}$ starting from cyclotomic extensions $\mathbb{Q}(\omega_m)$. The following lemma establishes the conditions under which the cyclotomic extension $\mathbb{Q}(\omega_m)$ is cyclic over $\mathbb{Q}$.

**Lemma 12.** [4] The Galois group $\text{Gal}(\mathbb{Q}(\omega_m)/\mathbb{Q}) \cong \mathbb{Z}_m^*$ is a cyclic group iff $m =$
Figure 7.1: Construction of Cyclic Extensions of $\mathbb{Q}(i)$ of Arbitrary Degree

$2, 4, p^e, 2p^e$ where $e \geq 1$ and $p$ is an odd prime.

The following theorem, illustrated in Figure 7.2, will prove to be very useful in identifying cyclic subfields of cyclotomic extensions.

**Theorem 13.** (Fundamental Theorem of Galois Theory) Let $\mathbb{K}$ be a finite Galois extension of $\mathbb{F}$. Then there is a one-one correspondence between the subfields $\mathbb{K}_1$ of $\mathbb{K}$ containing $\mathbb{F}$ and subgroups $H$ of $G = \text{Gal}(\mathbb{K}/\mathbb{F})$. The correspondence maps a subgroup $H$ of $G$ to the largest subfield $\mathbb{K}_1$ of $\mathbb{K}$ fixed by the subgroup and subfields $\mathbb{K}_1$ of $\mathbb{K}$ containing $\mathbb{F}$ to the largest subgroup of $G$ fixing the subfield. Under this correspondence, $[\mathbb{K}_1 : \mathbb{F}] = |G/H|$.

For reasons that will become clearer as we proceed, we set $\mathbb{F}_1 = \mathbb{Q}(\omega_{p_1^{m_1+1}}) = \mathbb{Q}(i)$ (i.e., $p_1 = 2$ and $m_1 = 1$). The Galois group $\text{Gal}(\mathbb{F}_1/\mathbb{Q})$ is cyclic of degree 2. Now consider $\mathbb{R}_i = \mathbb{Q}(\omega_{p_i^{m_i+1}})$, $p_i$ are distinct odd primes, $i = 2, \cdots, r$ (See Figure 7.1). Then each $\mathbb{R}_i/\mathbb{Q}$ is an extension of degree $\phi(p_i^{m_i+1}) = p_i^{m_i}(p_i - 1)$, with Galois group $G_i$.
Figure 7.2: The Fundamental Theorem of Galois Theory

isomorphic to $\mathbb{Z}_{p_i^{m_i+1}}^*$. From Lemma 12, $G_i$ is cyclic over $\mathbb{Q}$. It is trivial to see that $G_i$ has a unique cyclic subgroup $H_i$ of order $(p_i - 1)$. From Theorem 13, there is a unique subfield $\mathbb{F}_i$ of $\mathbb{R}_i$ which is fixed by the subgroup $H_i$ such that $\mathbb{Q} \subseteq \mathbb{F}_i \subseteq \mathbb{R}_i$. Also, it follows from Theorem 13, that $[\mathbb{R}_i : \mathbb{F}_i] = |H_i| = (p_i - 1)$. Next, we show that $\mathbb{F}_i$ is a cyclic Galois extension of $\mathbb{Q}$ of degree $p_i^{m_i}$. Since $\text{Gal}(\mathbb{F}_i/\mathbb{Q}) \cong G_i/H_i$, it is sufficient to show that $G_i/H_i$ is cyclic.

**Theorem 14.** Let $G = \mathbb{Z}_{p_i^{m_i+1}}^*$, and $H$ be the unique subgroup of $G$ of size $(p - 1)$ consisting of elements of order dividing $(p - 1)$. Then $G/H$ is cyclic of order $p^m$.

**Proof.** Let $\psi$ be the homomorphism from $G$ onto $\psi(G)$ defined by

$$\psi(g) = g^{p-1}, \ g \in G.$$ 

Clearly, $\psi(G)$ is a cyclic subgroup of $G$ of order $p^m$. The kernel $K_\psi$ of $\psi$ is the set of all elements of $G$ whose order divides $(p - 1)$, i.e.,

$$K_\psi = H.$$
It follows that $G/H \cong \psi(G)$ and is therefore cyclic.

7.3.2 Cyclic Galois Extensions over $\mathbb{F}$

In the previous subsection, we constructed cyclic extensions $\mathbb{F}_i$ of $\mathbb{Q}$. Our main interest is however in using these $\mathbb{F}_i$, which are cyclic Galois over $\mathbb{Q}$, to construct cyclic Galois extensions over $\mathbb{F} = \mathbb{Q}(\iota)$ of arbitrary degree $n$. We will now accomplish this task.

**Theorem 15.** Let $S = S_1S_2\cdots S_r$ be a field which is the compositum of the fields $S_1, S_2, \ldots, S_r$. If each $S_i$ is a cyclic Galois extension over $\mathbb{F}$ of degree $n_i$ (where the $n_i$ are pairwise relatively prime), then $S$ is a cyclic Galois extension over $\mathbb{F}$ of degree $\prod_{i=1}^r n_i$.

**Proof.** Consider the compositum $S = S_1S_2\cdots S_r$. Each $S_i/\mathbb{F}$ is a cyclic Galois extension of order $n_i$ such that $(n_i, n_j) = 1$ $\forall$ $i \neq j$. Thus,

$$S_i \cap S_j = \mathbb{F} \forall i \neq j$$

$$\Rightarrow Gal(S/\mathbb{F}) = Gal(S_1/\mathbb{F}) \times \cdots \times Gal(S_r/\mathbb{F})$$

It follows that $S/\mathbb{F}$ is cyclic of degree $n$. □

The composite field $\mathbb{K} = F_1F_2\ldots F_r$, with $F_i$ as defined above and as shown in Figure 7.1, has $\mathbb{Q}(\iota)$ as a sub-field, since $F_1 = \mathbb{Q}(\iota)$. From the above theorem, $\mathbb{K}$ is a cyclic extension of $\mathbb{Q}$. $Gal(\mathbb{K}/\mathbb{Q}(\iota))$ is a subgroup of the cyclic group $Gal(\mathbb{K}/\mathbb{Q})$ and is hence cyclic of order $\prod_{i=2}^r p_i^{m_i}$. Now if we are able to generate a cyclic Galois extension $E$ of degree $2^m$ for any $m$ over $\mathbb{Q}(\iota)$, then by using the compositum $\mathbb{L}$ of $\mathbb{K}$ and $E$, we will obtain a cyclic Galois extension $\mathbb{L}$ of arbitrary degree $n = 2^m \prod_{i=2}^r p_i^{m_i}$ over $\mathbb{Q}(\iota)$. The field $E = \mathbb{Q}(\omega_{2^{m+2}})$ shown in Figure 7.1 has the desired properties.

**Theorem 16.** $\mathbb{Q}(\omega_{2^{m+2}})/\mathbb{Q}(\iota)$ is a cyclic Galois extension of degree $2^m$.

**Proof.** Note that in $\mathbb{Z}_{2^{m+2}}^*$, the element 5 has maximal possible order $2^m$. This follows
Chapter 7. Minimal Delay D-MG Optimal Space-Time Block Codes

since $5^{2^{m-1}} \equiv (1 + 2^2)^{2^{m-1}} \equiv 1 + 2^{m-1} \cdot 2^2 \not\equiv 1 \pmod{2^{m+2}}$. Next, consider the automorphisms $\sigma_k, 0 \leq k \leq 2^m - 1$, of $\mathbb{Q}(\omega_{2^m+2})/\mathbb{Q}$ given by

$$\sigma_k(\omega_{2^m+2}) = \omega_{2^m+2}^{5^k}, \quad 0 \leq k \leq 2^m - 1.$$ 

These form a cyclic group of order $2^m$ and the fixed field of this group is $\mathbb{Q}(i)$ since $r^{5^k} = i$, all $k$. It follows that $\mathbb{Q}(\omega_{2^m+2})/\mathbb{Q}(i)$ is cyclic of degree $2^m$. □

Finally, we observe that $\mathbb{L}$ is a cyclic Galois extension of arbitrary degree $n = 2^m \prod_{i=2}^r p_i^{m_i}$ over $\mathbb{Q}(i)$, as desired.

### 7.4 Determining a Non-Norm Element

In this section, we prove the existence and give a procedure to determine a non-norm element, by which we mean, as stated earlier, an element $\gamma \in \mathbb{Q}(i)$ such that the smallest $i$ for which $\gamma^i \in N_{L/\mathbb{Q}(i)}(L)$ is $n$. Here $L$ is a degree $n$ cyclic Galois extension of $\mathbb{F} = \mathbb{Q}(i)$ (constructed using the techniques given in the previous section). We will build on the work in [37]. Following is the main theorem of [37].

**Theorem 17.** Let $\mathbb{K}$ be a degree $n$ Galois extension of a number field $\mathbb{F}$ and let $\mathfrak{p}$ be a prime ideal in $\mathcal{O}_\mathbb{F}$, the ring of algebraic integers in $\mathbb{F}$, that is below the prime ideal $\mathfrak{P} \subset \mathcal{O}_\mathbb{K}$ whose norm is given by $\|\mathfrak{P}\| = \|\mathfrak{p}\|^f$, where $f$ the inertial degree of $\mathfrak{P}$ over $\mathfrak{p}$. If $\gamma$ is any element of $\mathfrak{p} \setminus \mathfrak{p}^2$, then $\gamma^i \notin N_{\mathbb{K}/\mathbb{F}}(\mathbb{K})$ for any $i = 1, 2, \ldots, f - 1$.

In particular, if $\text{Gal}(\mathbb{K}/\mathbb{F}) = \langle \sigma \rangle$ with $[\mathbb{K} : \mathbb{F}] = n$, then the cyclic algebra $(\mathbb{K}/\mathbb{F}, \sigma, \gamma)$ is a division algebra if $\gamma \in \mathfrak{p} \setminus \mathfrak{p}^2$ for some prime triplet $(p; \mathfrak{p}; \mathfrak{P})$ with $f(\mathfrak{P}|p) = n$, where $p$ is the characteristic of the field $\mathcal{O}_\mathbb{F}/\mathfrak{p}$.

**Proof.** See [37]. □

Restricting our attention to the case when $\mathbb{F} = \mathbb{Q}(i)$, in order to find a non-norm element $\gamma$ in $\mathbb{Z}(i)$, it is therefore sufficient that we find a prime ideal in the ring of integers
\( O_\mathfrak{p} = \mathbb{Z}[i] \) such that it’s inertial degree \( f \) in \( \mathbb{L}/\mathbb{Q}(i) \) is \( f = [\mathbb{L} : \mathbb{Q}(i)] = n \). Such an ideal is said to be inert in \( \mathbb{L}/\mathbb{Q}(i) \).

Our approach towards determining \( \gamma \) will be as follows. First, we will show that there exist prime ideals (possibly different) in \( \mathbb{Z} \) that are inert in each branch \( \mathbb{F}_i/\mathbb{Q} \) (see Figure 7.1). We will then use these ideals to find a prime ideal in \( \mathbb{Z} \), which is inert in the compositum \( \mathbb{K}/\mathbb{Q} \). Obviously, the (prime) ideal that lies above it in \( \mathbb{Z}(i) \) is inert in \( \mathbb{K}/\mathbb{Q}(i) \). Once this is accomplished, we can use similar arguments to show that if an inert ideal exists in the branch \( \mathbb{Q}(\omega_{2m+2})/\mathbb{Q}(i) \), then we can find an inert ideal in \( \mathbb{L}/\mathbb{Q}(i) \).

We now proceed to find an inert ideal in \( \mathbb{F}_1/\mathbb{Q} \) for \( i = 2, \cdots, r \). For this purpose, we need the following lemma.

**Lemma 18.** [4] Let \( p \) be any odd prime. Then for any integer \( k \), the group \( \mathbb{Z}_{p^k}^* \) is cyclic of order \( \phi(p^k) \). For any integer \( f \) dividing \( \phi(p^k) \) there exists an integer \( a \in \mathbb{Z}_{p^k}^* \) such that \( a \) has order \( f \) in \( \mathbb{Z}_{p^k}^* \).

It will turn out that our interest is in finding a prime ‘\( a \)’ that has multiplicative order \( f = \phi(p^k) \), under modulo-\( p^k \) arithmetic. The above lemma guarantees the existence of some integer \( a \) with this property, where \( a \) is not necessarily prime. Dirichlet’s theorem proves the existence of an ‘\( a \)’ that is prime.

**Theorem 19.** (Dirichlet’s theorem)

Let \( a, m \) be integers such that \( 1 \leq a \leq m \), \( \gcd(a, m) = 1 \). Then the arithmetic progression

\[
\{a, a + m, a + 2m, \ldots, a + km, \ldots\}
\]

contains infinitely many prime numbers.

Setting the value of \( m \) in the above theorem to \( m = p^k \) validates our claim regarding the existence of a prime number \( a \) with multiplicative order \( f = \phi(p^k) \mod p^k \).

Let us now go back to Figure 7.1. We have \( \mathbb{R}_i = \mathbb{Q}(\omega_{p_i^{m+1}}) \), where \( p_i \) is an odd prime. Let \( A_i \) be the ring of integers of \( \mathbb{R}_i \). Then, we have the following lemma from [4].
Lemma 20. [4] Let $q_i$ be any prime number distinct from $p_i$, let $f_i \geq 1$ be the smallest integer such that $q_i^{f_i} \equiv 1 \pmod{p_i^{m_i+1}}$ and let $g_i = \phi(p_i^{m_i+1})/f_i$. Then $q_iA_i = \beta_1 \cdots \beta_{g_i}$ where $\beta_1, \ldots, \beta_{g_i}$ are distinct prime ideals of $A_i$ and $N(\beta_j) = q_i^{f_i}$, all $j$.

From Lemma 18 and Dirichlet’s theorem, we can find a prime $q_i$ with primitive order $f_i = \phi(p_i^{m_i+1})$ for $i = 2, \ldots, r$. Then, using Lemma 20, we conclude that the ideal $q_iA_i$ is prime in $A_i$. Hence $q_i$ is inert in $\mathbb{R}_i/\mathbb{Q}$, which also implies that it is inert in $F_i/\mathbb{Q}$. We have now identified primes $q_i$ which have full inertial degree in $F_i/\mathbb{Q}$ i.e., $f_i = [F_i : \mathbb{Q}] = p_i^{m_i}$ for $i = 2, \ldots, r$. For $i = 1$, i.e., $F = F_1 = \mathbb{Q}(\iota)$, we will always use the prime $q_1 = 5$, whose inertial degree in $\mathbb{Q}(\omega_{2^{m+2}})/\mathbb{Q}$ is $2^m$ for all $m$ [2, 37]. Note that $5 = (2 + i)(2 - i)$ in $\mathbb{Q}(\iota)$, which implies that the inertial degree of $q_1$ in $F_1/\mathbb{Q}$ is $f_1 = 1$. This in turn implies that the inertial degree in $\mathbb{Q}(\omega_{2^{m+2}})/\mathbb{Q}(\iota)$ of the ideal in $\mathbb{Z}(\iota)$ above $q_1$ is $2^m$ (this follows from the transitivity property of inertial degrees, see (5.4)). We now use these $q_i$ to obtain an inert ideal in $\mathbb{K}/\mathbb{Q}(\iota)$ and then in $\mathbb{L}/\mathbb{Q}(\iota)$. For this purpose, we will use the following theorem.

Theorem 21. If there exist primes $q_i \in \mathbb{Z}$ having primitive multiplicative order $f_i$ in $\mathbb{Z}/p_i^{m_i+1}$ for $i = 1, 2, \ldots, r$ with all $p_i$ being distinct primes, then there exists a prime $q \in \mathbb{Z}$ which has primitive multiplicative order $f_i$ in $\mathbb{Z}/p_i^{m_i+1}$ for $i = 1, 2, \ldots, r$.

Proof. Using the Chinese remainder theorem, there exists an integer $a \in \mathbb{Z}$ such that,

$$a \equiv q_i \pmod{p_i^{m_i+1}}, \forall i.$$

Hence, the multiplicative order of $a$ is equal to the multiplicative order of $q_i \pmod{p_i^{m_i+1}}$ for all $i = 1, 2, \ldots, r$.

Let $n' = \prod_{i=2}^{r} p_i^{m_i}$. Then $\{a, a + n', \ldots, a + kn', \ldots\}$ forms an arithmetic progression. Note that $(a, n') = 1$. Using Dirichlet’s theorem, we can find a prime $q$ satisfying,

$$q \equiv a \pmod{p_i^{m_i+1}}, \forall i = 1, 2, \ldots, r.$$
Using the above theorem, we can find a prime $q$ which has inertial degree $f_i$ in $\mathbb{F}_i/\mathbb{Q}$ for all $i = 1, 2, \ldots, r$. Now consider,

**Theorem 22.** Let $\mathbb{K}_1$ and $\mathbb{K}_2$ be Galois extensions of a field $\mathbb{F}$ of degree $p_1^{n_1}$ and $p_2^{n_2}$ respectively, for distinct primes $p_1$ and $p_2$. If an ideal $J \in \mathcal{O}_{\mathbb{F}}$ has inertial degree $f_1 = p_1^{n_1}$ and $f_2 = p_2^{n_2}$ in $\mathbb{K}_1/\mathbb{F}$ and $\mathbb{K}_2/\mathbb{F}$ respectively, then it’s inertial degree in the compositum field $\mathbb{K}_1\mathbb{K}_2$ is $f = f_1 f_2 = p_1^{n_1} p_2^{n_2}$.

**Proof.** Consider the field extension diagram in Fig.7.3. Let $f$ be the total inertial degree of $J$ in the compositum $\mathbb{K}_1\mathbb{K}_2$. Then we have $f = f_1 f'_1 \Rightarrow f_1 | f$, and also $f = f_2 f'_2 \Rightarrow f_2 | f$.

But we know that $(f_1, f_2) = 1$ and $f \leq [\mathbb{K}_1\mathbb{K}_2 : \mathbb{F}]$. This implies that $f = f_1 f_2 = p_1^{n_1} p_2^{n_2}$.

From Theorem 22 and Figure 7.1, the total inertial degree of $q$ in $\mathbb{K}/\mathbb{Q}$ is $\prod_{i=2}^{r} p_i^{m_i} = [\mathbb{K} : \mathbb{Q}]/2$. Since the inertial degree of $q$ in $\mathbb{Q}(i)/\mathbb{Q}$ is $f_1 = 1$, from (5.4), we conclude that the inertial degree in $\mathbb{K}/\mathbb{Q}(i)$ of the ideal in $\mathbb{Z}(i)$ above $q$ is equal to $\prod_{i=2}^{r} p_i^{m_i}$.

The above arguments have essentially achieved the following: by identifying inert ideals in the individual branches $\mathbb{F}_i/\mathbb{Q}$, $i = 2, \ldots, r$, we have “moved up the tower” (Figure 7.1) and produced an inert ideal in the compositum $\mathbb{K}/\mathbb{Q}(i)$. In the sequel, we will repeat the above procedure to move further up the tower from $\mathbb{K}$ to $\mathbb{L}$.
As mentioned previously, $q_1 = 5$ has inertial degree $2^m$ in $\mathbb{Q}(\omega_{2^{m+2}})/\mathbb{Q}(i)$ [2, 37]. Through an application of Theorems 21 and 22, we can find a prime $q_0$ which has inertial degree in $L/Q(i)$ equal to $f = [L : Q(i)] = 2^m \prod_{i=2}^r p_i^{m_i} = n$. It is clear that $q_0$ and $q$ have the same inertial degree of 1 in $Q(i)/Q$. Therefore, from Lemma 20, $q_0Z(i) = B_1B_2$, where $B_1$ and $B_2$ are distinct prime ideals in $\mathbb{Z}[i]$. Choose $\gamma \in B_1 \setminus B_1^2$ as the required non-norm element (since $\mathbb{Z}[i]$ is a principal ideal domain, the generator of the principal ideal $B_1$ is an immediate candidate for $\gamma$). Having proven the existence of $\gamma$ for any number of transmit antennas, we explicitly calculate $\gamma$ for a few cases, shown in table 7.1.

### Table 7.1: Non-Norm Elements

<table>
<thead>
<tr>
<th>No. of Antennas</th>
<th>prime 'q'</th>
<th>Non-norm 'γ'</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5</td>
<td>$(2 + i)$</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>$(2 + i)$</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>$(2 + i)$</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>$(3 + 2i)$</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>$(2 + i)$</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>$(2 + i)$</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>$(2 + i)$</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>$(2 + i)$</td>
</tr>
<tr>
<td>10</td>
<td>13</td>
<td>$(3 + 2i)$</td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>$(3 + 2i)$</td>
</tr>
</tbody>
</table>

7.5 Reducing the Signalling Complexity

The field $L$ obtained in Section 7.3 is a subfield of the product of cyclotomic extensions $\mathbb{Q}(\omega_{2^{m+2}}) \prod_{i=2}^r R_i = \mathbb{Q}(\omega_v)$ (say). The entries of the space-time code matrices constructed from a division algebra having $L$ as its maximal subfield belong to $L$. As a result, the signalling alphabet of the space-time code is an AM$-v$-PSK constellation [56], where the parameter $v$ is determined by the field $L$. In this section, our interest will be twofold:

- to construct cyclic Galois extensions $L$ which have minimal signalling complexity, i.e., to minimize the value of $v$, and,
• to present methods of determining non-norm elements in this case.

7.5.1 Cyclic Galois Extensions

Consider the following theorem.

Theorem 23. (Kronecker-Weber) [4] If $L \mid \mathbb{Q}$ is an Abelian extension (that is, a Galois extension with Abelian Galois group) of finite degree, then there exists a root of unity $\xi$, such that $L \subseteq \mathbb{Q}(\xi)$.

Since all cyclic extensions are Abelian, we can conclude from the above theorem that it is sufficient that we restrict our hunt for cyclic Galois extensions of degree $n$ to hunting among subfields of cyclotomic extensions. Moreover, we will aim to reduce the degree of $\mathbb{Q}(\xi)$ in accordance with our desire to minimize the signalling complexity. We will show that this problem can be recast in our space-time code construction framework as follows:

“Find the minimum positive integer $v$ such that $\mathbb{Z}_v^*$ has a cyclic subgroup of order $n$ and such that $\mathbb{Q}(i) \subset \mathbb{Q}(\omega_v)$.”

\[
\begin{array}{c}
\mathbb{Q}(\omega_v) \\
\phi(v)/(2n) \\
L \\
n \\
\mathbb{Q}(i) \\
2 \\
\mathbb{Q}
\end{array}
\]

Figure 7.4: Cyclic Galois Extensions with Reduced Signalling Complexity

The method of determining the smallest such $v$ is through a bounded search: we have an upper bound on $v$ from the previous subsection to be $2^{m+2} \prod_{i=2}^{r} p_i^{m_i+1}$. A similar
approach was used in [56], with the additional restriction that $Q(\omega_v)$ be cyclic. The framework here includes the case when $Q(\omega_v)$ is not cyclic too, and is hence more general than [56]. The situation in this case is as depicted in Figure 7.4. Since $Q(i) \subset Q(\omega_v)$, we have that $4|v$. Before delving into the details of the search for $v$, we will present three theorems regarding residue classes from [4].

**Theorem 24.** Let $n = \prod_{i=1}^{r} p_i^{e_i}$ be the decomposition of $n > 1$ into prime powers (with $e_i > 0$ for $i = 1, 2, \ldots, r$). Then there exists a ring-isomorphism

$$\varphi : \mathbb{Z}_n \rightarrow \prod_{i=1}^{r} \mathbb{Z}_{p_i^{e_i}}.$$ 

**Proof.** Let $\varphi : \mathbb{Z} \rightarrow \prod_{i=1}^{r} \mathbb{Z}_{p_i^{e_i}}$ be the map defined as $\varphi(n) = (\nu_1, \nu_2, \ldots, \nu_r)$ where $\nu_i = n \pmod{p_i^{e_i}}$ for $i = 1, 2, \ldots, r$. It is clear that $\varphi$ is a ring-homomorphism. The kernel of this homomorphism is the set of all elements divisible by $p_i^{e_i}$ for $i = 1, 2, \ldots, r$, viz., the ideal $(n)$ consisting of all multiples of $n$. Thus, we have that the map

$$\varphi : \mathbb{Z}/(n) = \mathbb{Z}_n \rightarrow \prod_{i=1}^{r} \mathbb{Z}_{p_i^{e_i}}$$

is a one-one ring homomorphism.

Since the cardinality of $\prod_{i=1}^{r} \mathbb{Z}_{p_i^{e_i}}$ is $\prod_{i=1}^{r} p_i^{e_i} = n$, we conclude that $\varphi$ is onto and is hence a ring isomorphism. \[ \square \]

**Theorem 25.** If $n = \prod_{i=1}^{r} p_i^{e_i}$ is the prime power decomposition of $n$, then

$$\mathbb{Z}_n^* \cong \prod_{i=1}^{r} \mathbb{Z}_{p_i^{e_i}}^*.$$ 

**Proof.** Consider the ring-isomorphism $\varphi : \mathbb{Z}_n \rightarrow \prod_{i=1}^{r} \mathbb{Z}_{p_i^{e_i}}$ from Theorem 24. Any element $\overline{a} \in \mathbb{Z}_n$ is invertible iff it’s image under $\varphi$ is invertible, i.e. the $i^{th}$ component of $\varphi(\overline{a})$ is invertible in $\mathbb{Z}_{p_i^{e_i}}$ for $i = 1, 2, \ldots, r$. This gives us that $\varphi(\mathbb{Z}_n^*) = \prod_{i=1}^{r} \mathbb{Z}_{p_i^{e_i}}^*$, which proves the claim. \[ \square \]
Theorem 26. If $p \neq 2$, $e \geq 1$, then $\mathbb{Z}_{p^e}^*$ is a cyclic group and

$$\mathbb{Z}_{p^e}^* \cong \mathbb{Z}_{p^{e-1}} \times \mathbb{Z}_{p-1}$$

Proof. See [4].

From the above theorems, we have for any given value $v = \prod_{i=1}^{s} q_i^{s_i}$ that

$$\mathbb{Z}_v^* \cong \prod_{i=1}^{s} \mathbb{Z}_{q_i^{s_i}}^*$$

$$\cong \prod_{i=1}^{s} \mathbb{Z}_{q_i^{s_i-1}} \times \mathbb{Z}_{q_i-1}$$

(7.1)

Now consider again Figure 7.4. We know that $\text{Gal}(\mathbb{Q}(i)/\mathbb{Q}) \cong \mathbb{Z}_2$ and we require that $\text{Gal}(L/\mathbb{Q}(i)) \cong \mathbb{Z}_n$. Also, using (7.1) and Theorem 24, we may write $\text{Gal}(\mathbb{Q}(\omega_v)/\mathbb{Q}) \cong \prod_{i=1}^{s} \mathbb{Z}_{n_i}$ for some integers $s$ and $n_i$, $i = 1, 2, \ldots, s$, where each $n_i$ is a prime power. Note that this Galois group is not necessarily cyclic. Using Galois theory, it is evident that $\text{Gal}(L/\mathbb{Q}(i))$ is isomorphic to a subgroup of $\text{Gal}(\mathbb{Q}(\omega_v)/\mathbb{Q})$. It is now clear that we must search for the smallest value of $v$ such that $\text{Gal}(\mathbb{Q}(\omega_v)/\mathbb{Q})$ contains the two cyclic subgroups of interest, $\mathbb{Z}_2$ (corresponding to $\text{Gal}(\mathbb{Q}(i)/\mathbb{Q})$) and $\mathbb{Z}_n$ (corresponding to $\text{Gal}(L/\mathbb{Q}(i))$). This is equivalent to searching for the smallest $v$ such that

$$\mathbb{Z}_v^* \cong \mathbb{Z}_2 \times \mathbb{Z}_n \times \prod_{i=1}^{k} \mathbb{Z}_{k_i}$$

for some integer values of $k, k_i \forall i$. Values of $v$ determined using this procedure are tabulated in table 7.2 for few $n$. Further, table 7.3 shows some values of $v$ for those $n$ which improve upon the values reported in [56].

7.5.2 Non-Norm Elements

Determining non-norm elements in this case turns out to be similar to the procedure given in Section 7.4. The reduced signalling complexity construction of cyclic Galois
Table 7.2: Cyclic Galois Extensions with Reduced Signalling Complexity

<table>
<thead>
<tr>
<th>n</th>
<th>v</th>
<th>n</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8</td>
<td>14</td>
<td>116</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>15</td>
<td>124</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>16</td>
<td>64</td>
</tr>
<tr>
<td>5</td>
<td>44</td>
<td>17</td>
<td>412</td>
</tr>
<tr>
<td>6</td>
<td>28</td>
<td>18</td>
<td>76</td>
</tr>
<tr>
<td>7</td>
<td>116</td>
<td>19</td>
<td>764</td>
</tr>
<tr>
<td>8</td>
<td>32</td>
<td>20</td>
<td>88</td>
</tr>
<tr>
<td>9</td>
<td>76</td>
<td>21</td>
<td>172</td>
</tr>
<tr>
<td>10</td>
<td>44</td>
<td>22</td>
<td>92</td>
</tr>
<tr>
<td>11</td>
<td>92</td>
<td>23</td>
<td>188</td>
</tr>
<tr>
<td>12</td>
<td>52</td>
<td>24</td>
<td>112</td>
</tr>
<tr>
<td>13</td>
<td>212</td>
<td>25</td>
<td>404</td>
</tr>
</tbody>
</table>

extensions in Section 7.5.1 makes use of the chain of fields shown in Figure 7.4. Let us assume the following prime factorization for $v$,

$$v = 2^\alpha_1 \prod_{i=2}^{s} q_i^{\alpha_i}.$$  

The single tower in Figure 7.4 can then be resolved into the equivalent structure shown in Figure 7.5.

It is now evident that the procedure for determining a non-norm element $\gamma$ using the tower in Figure 7.5 is exactly similar to the one in Section 7.4, owing to the similarity in structure of the towers in Figures 7.1 and 7.5.
Table 7.3: Values of $n$ for which Lower Signalling Complexity is Obtained than in [56]

<table>
<thead>
<tr>
<th>$n$</th>
<th>$v$ from [56]</th>
<th>$v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>100</td>
<td>88</td>
</tr>
<tr>
<td>24</td>
<td>224</td>
<td>112</td>
</tr>
<tr>
<td>44</td>
<td>356</td>
<td>184</td>
</tr>
<tr>
<td>48</td>
<td>388</td>
<td>224</td>
</tr>
<tr>
<td>80</td>
<td>704</td>
<td>352</td>
</tr>
<tr>
<td>84</td>
<td>688</td>
<td>344</td>
</tr>
<tr>
<td>92</td>
<td>752</td>
<td>376</td>
</tr>
<tr>
<td>104</td>
<td>1252</td>
<td>848</td>
</tr>
<tr>
<td>116</td>
<td>932</td>
<td>472</td>
</tr>
<tr>
<td>120</td>
<td>964</td>
<td>476</td>
</tr>
<tr>
<td>123</td>
<td>2956</td>
<td>2324</td>
</tr>
<tr>
<td>132</td>
<td>1072</td>
<td>536</td>
</tr>
<tr>
<td>140</td>
<td>1124</td>
<td>568</td>
</tr>
<tr>
<td>144</td>
<td>1126</td>
<td>608</td>
</tr>
<tr>
<td>145</td>
<td>5804</td>
<td>2596</td>
</tr>
<tr>
<td>159</td>
<td>12724</td>
<td>2996</td>
</tr>
<tr>
<td>160</td>
<td>1408</td>
<td>704</td>
</tr>
<tr>
<td>164</td>
<td>1328</td>
<td>664</td>
</tr>
<tr>
<td>168</td>
<td>1348</td>
<td>688</td>
</tr>
<tr>
<td>176</td>
<td>1412</td>
<td>736</td>
</tr>
<tr>
<td>184</td>
<td>1504</td>
<td>752</td>
</tr>
<tr>
<td>188</td>
<td>3764</td>
<td>2264</td>
</tr>
<tr>
<td>195</td>
<td>4684</td>
<td>3476</td>
</tr>
</tbody>
</table>

Figure 7.5: Non-Norm Elements - Resolving the Tower in Figure 7.4
Chapter 8

Rectangular D-MG Optimal Space-Time Block Codes

The objective of this chapter is to generalize the construction of square ($T = n_t$) D-MG optimal space-time codes derived from division algebras to the rectangular case ($T > n_t$). We will provide two constructions, the puncturing and stacking constructions, and show that these rectangular constructions achieve the upper bound on the optimal D-MG tradeoff using the sufficient condition reviewed in Chapter 4.

8.1 Intuition Behind the Constructions

Square space-time codes from division algebras were derived through the left regular representation of elements in a division algebra. Any division algebra $D$ is an $n$-dimensional vector space over its maximal subfield $L$ for some $n$, see Figure 8.1. The left regular representation $\lambda_d : e \mapsto de$, $\forall e \in D$ is an $L$–linear transformation from $D \rightarrow D$ (see Chapter 6). The matrix representation of this ‘dimension-preserving’ linear transformation is therefore square in dimension. The most natural generalization to the rectangular case is to concoct linear transformations which are ‘rank-decreasing’ and obtain our space-time codes as the matrix representations of these transformations. In particular, to obtain an $(n_t \times T)$ space-time code (assume $T \geq n_t$), we will produce linear transformations from a
division algebra $D$ of index $T$ (which is a $T$-dimensional vector-space over the maximal subfield), to an $n_t$-dimensional subspace $D'$. This setup is illustrated in Figure 8.2 and is the idea behind the puncturing rectangular construction in Section 8.2.

The stacking construction presented in Section 8.3 is obtained by horizontally stacking a series of puncturing constructions, and is aimed at minimizing the signalling complexity.

### 8.2 The Puncturing Construction

The starting point to construct an $(n_t \times T)$ punctured space-time code $\mathcal{X}$ is a $(T \times T)$ space-time code $\mathcal{X}'$ obtained using the construction given in Chapter 7. $\mathcal{X}'$ satisfies the sufficient condition in Theorem 7, i.e.,

$$\delta_{\mathcal{X}'} = T - r.$$ 

We construct the $(n_t \times T)$ ST code $\mathcal{X}$ from $\mathcal{X}'$ by deleting (puncturing) the last $(T - n_t)$ rows of each matrix $X' \in \mathcal{X}'$, i.e., every $X \in \mathcal{X}$ is

$$X_{n_t \times T} = \begin{bmatrix} I_{n_t \times n_t} & 0_{n_t \times (T-n_t)} \end{bmatrix} X' \quad \text{for some } X' \in \mathcal{X}'. \quad (8.1)$$
We can also write this as

\[ X_{T \times T}' = \begin{bmatrix} X_{n_t \times T} \\ Z_{(T-n_t) \times T} \end{bmatrix}, \text{ for some particular } Z \in \mathbb{C}^{(T-n_t) \times T}. \]

From (6.9), each codeword \( X \in \mathcal{X} \) is of the form

\[ X = \begin{bmatrix} \ell_0 & \gamma \sigma(\ell_{T-1}) & \cdots & \gamma \sigma^{T-1}(\ell_1) \\ \ell_1 & \sigma(\ell_0) & \cdots & \gamma \sigma^{T-1}(\ell_2) \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{n_t-1} & \sigma(\ell_{n_t-2}) & \cdots & \gamma \sigma^{T-1}(\ell_{n_t}) \end{bmatrix}_{n_t \times T}. \tag{8.2} \]

We pause to note that in principle, however, any set of \( (T-n_t) \) rows may be deleted instead of the last \( (T-n_t) \). In line with the discussion in Section 8.1, each matrix \( X \) in (8.2) corresponds to the matrix representation of the composition of two linear transformations,

1. the left regular representation of elements in the division algebra \( D \)
2. projection onto \( D' \), which is an \( n_t \)-dimensional subspace of \( D \) (regarded as a vector
space over the maximal subfield).

The D-MG optimality of the puncturing construction is provided in the following theorem.

**Theorem 27.** (D-MG optimality of the puncturing construction) Let \( \mathcal{X}' \) be a \((T \times T)\) D-MG optimal space-time code that satisfies the sufficiency condition in Theorem 7. The \((n_t \times T)\) space-time code \( \mathcal{X} \) (assume \( T > n_t \)) derived from \( \mathcal{X}' \) by retaining any set of \( n_t \) rows from each codeword matrix in \( \mathcal{X}' \) achieves the optimal diversity-multiplexing gain tradeoff.

**Proof.** See Appendix A.

\[ \square \]

### 8.3 The Stacking Construction

The \((n_t \times T)\) stacking construction \((T \geq n_t)\) is obtained by horizontally stacking several punctured codeword matrices of lesser size. The motivation behind this approach is the reduction in signalling complexity achieved.

Let \( t \) be the smallest factor of \( T \) which is greater than or equal to \( n_t \). Such a factor is always guaranteed to exist because \( T \) is trivially a factor of itself and \( T \geq n_t \) by assumption. The \((n_t \times T)\) stacking space-time code construction \( \mathcal{X} \) will be obtained by juxtaposing \( k = \frac{T}{t} \) number of punctured rectangular \((n_t \times t)\) ST code constructions \( \mathcal{X}_i, \ i = 1, 2, \ldots, k \) as shown below.

\[
\mathcal{X} = \left[ \begin{bmatrix} \mathcal{X}_1 \end{bmatrix}_{n_t \times t} \begin{bmatrix} \mathcal{X}_2 \end{bmatrix}_{n_t \times t} \cdots \begin{bmatrix} \mathcal{X}_k \end{bmatrix}_{n_t \times t} \right]_{n_t \times T}
\]

Each \( \mathcal{X}_i \) carries independent information. The entries of the stacked space-time codeword matrices lie in a field \( \mathbb{L} \) which is a degree \( t \) cyclic Galois extension of \( \mathbb{Q}(i) \). However, the entries of the corresponding \((n_t \times T)\) punctured codeword matrices will lie in a degree \( T \) extension of \( \mathbb{Q}(i) \). The reduction in signalling complexity of the stacking construction is evident from this fact (see also Section 7.5).
Remark 3. When \( T \) is prime, or if the smallest factor \( t \geq n_t \) is equal to \( T \), the puncturing and stacking constructions coincide.

The following theorem claims the D-MG optimality of the stacking construction.

**Theorem 28.** The \((n_t \times T)\) stacking rectangular space-time code construction \( X \) achieves the optimal D-MG tradeoff.

*Proof. See Appendix B.*
Chapter 9

Generalizations of the RDST Code

We begin this chapter with a review of the rate-diversity tradeoff and by evaluating the rate-diversity tradeoff of some common space-time codes from the literature. The rank-distance space-time (RDST) codes were introduced by Lu and Kumar in [69, 64] as a unified construction of space-time codes over a wide variety of constellations that achieve the rate-diversity tradeoff. A brief overview of the RDST construction appears in this chapter.

Among the salient points of the RDST code is that it includes as its transmission alphabet several commonly used constellations such as PAM, QAM and PSK. Such a property is very desirable in practical systems, where simplicity and ease of implementation are of paramount importance. In comparison, the constellations of most other space-time codes, including those derived from division algebras in Chapters 7 and 8, do not have a simple representation.

Recent work in [52, 53] have provided bounds on the D-MG tradeoff of the RDST scheme specialized to the QAM alphabet. They have shown the QAM-RDST code to have a reasonably good D-MG tradeoff. We present a generalization to the RDST scheme known as the variable-rank RDST (v-r RDST) code and show that the scheme specialized to the QAM-alphabet has a significantly better D-MG tradeoff than the original QAM-RDST scheme. The rate-diversity of the v-r RDST code is also determined. It is hoped that the v-r RDST code will prove to be an attractive option for practical systems, as
it offers the advantages of having a simple constellation as it’s transmission alphabet as well as a good D-MG tradeoff.

Some other recent generalizations in the literature (see Section 9.9 for a review) have extended the RDST construction over constellations such as AM-PSK. In Section 9.10, we provide a construction for space-time codes based on power-series expansion with respect to a prime ideal in a number field. An advantage of this approach is that it relates to a broad class of constellations and special instances of this construction yield the RDST code and some of it’s subsequent generalizations.

9.1 The Rate-Diversity Tradeoff

The rate-diversity tradeoff was introduced by Tarokh et al. in [16] and was subsequently worked on by Lu and Kumar in [64]. Let \( \mathcal{X} \) be an \( (n_t \times T) \) space-time code whose entries are drawn from an alphabet \( \mathcal{A} \). Define the symbol rate \( R' \) of \( \mathcal{X} \) [16] as

\[
R' := \frac{1}{T} \log_{|\mathcal{A}|} |\mathcal{X}|.
\] (9.1)

Under this definition, a rate-one code corresponds to an ST code of size \( |\mathcal{A}|^T \), i.e., a code which transmits on an average, one symbol from the signal constellation \( \mathcal{A} \) per channel use. Recall from Section 2.2 that a space-time code \( \mathcal{X} \) is said to achieve a diversity advantage of \( n_r \nu \) if \( [16, 17, 21] \)

\[
n_r \nu = - \lim_{\text{SNR} \to \infty} \frac{\log(\text{PEP})}{\log(\text{SNR})},
\]

where \( \nu \) denotes the transmit diversity. A space-time code achieves transmit diversity \( \nu \) if and only if for every \( X_1, X_2 \in \mathcal{X} \), the difference matrix \( \Delta X = X_1 - X_2 \) has rank at least \( \nu \) over the field of complex numbers \( \mathbb{C} \). Under the assumption that \( T \geq n_t \), the maximum transmit diversity achievable is therefore \( n_t \).

For a fixed, finite signal constellation \( \mathcal{A} \), the tradeoff between the symbol-rate \( R' \) and the transmit diversity \( \nu \) of a space-time code \( \mathcal{X} \) (\( T \geq n_t \)) is known as the rate-diversity
tradeoff [64, 16] and is given by

\[ R' \leq n_t - \nu + 1. \]  \tag{9.2} \]

The proof of the rate-diversity tradeoff can be found in [64]. The tradeoff for the case when \( n_t = 4 \) is illustrated in Figure 9.1. In particular, a space-time code that achieves maximal transmit diversity = \( n_t \) must have symbol rate \( R' \leq 1 \).

![Figure 9.1: The Rate-Diversity Tradeoff (\( n_t = 4 \))](image)

A space-time block code having transmit diversity equal to \( \nu \), \( 1 \leq \nu \leq n_t \), and rate \( R' = n_t - \nu + 1 \) will be said to achieve the rate-diversity tradeoff and called a rate-diversity optimal code. A rate-diversity optimal code will be called a minimal-delay rate-diversity optimal code if in addition to optimality, the block length \( T \) is the minimum possible, i.e., \( T = n_t \).

A few remarks are in order [64, 69].

**Remark 4.** For the case when \( T < n_t \), the tradeoff takes on the form

\[ R' \leq n_t - \frac{n_t(\nu - 1)}{T}. \]  \tag{9.3} \]

**Remark 5.** The tradeoffs given in (9.2) and (9.3) apply even for the case when each
symbol $X_{ij}$ of the space-time matrix $X$ is drawn from a different signal constellation, provided that the constellations are all of the same size.

**Remark 6.** The rate diversity tradeoff is independent of the number of receive antennas $n_r$. However, borrowing terminology from [47], we set $d^*(R')$ to be the maximum achievable diversity advantage for a given symbol rate $R'$. Since the diversity advantage must be an integer, we have from (9.2) that

$$d^*(R') \leq n_r [n_t - R' + 1],$$

where $[\cdot]$ denotes the floor function.

**Remark 7.** The rate-diversity tradeoff issue is first discussed by Tarokh et al. in Theorem 3.3.1 and its corollary in [16] where the bound on the rate of the space-time code is phrased in terms of the maximal size $A_q(n_t, \nu)$ of an error-correcting code over an alphabet of size $q = |A|^T$. The Singleton bound applied to such error-correcting codes states that

$$A_q(n_t, \nu) \leq q^{n_t - \nu + 1}$$

and can be used to give an alternate proof of the bound (9.2).

At times in the literature on space-time codes, one comes across papers with transmit diversity $\nu$ and rates $R'$ that seem to be in conflict with what is permitted under (9.2). This is often due to the adoption of a different notion of symbol rate. For example, in some papers, the transmission symbol alphabet is built up from a smaller base (or building) alphabet $B$ and the symbol rate is measured with respect to the size of this base alphabet, i.e., according to

$$R'' := \frac{1}{T} \log_{|B|} |X|.$$
For example, if the symbols in $\mathcal{A}$ are linear combinations in the form of

$$
\sum_{i=1}^{k} \gamma_i b_i
$$

where the coefficients $\{\gamma_i\}$ are drawn from a fixed set of $k$ complex numbers and where $\{b_i\}$ are taken from a base alphabet $\mathcal{B}$, then potentially, we could have

$$|\mathcal{A}| = |\mathcal{B}|^k.$$

In this situation, the bound on the rate $R''$ (for $T \geq n_t$) takes on the form

$$R'' \leq k[n_t - \nu + 1],$$

which shows that even with $\nu = n_t$, rates $R'' > 1$ may be possible\(^1\).

### 9.2 Rate-Diversity Tradeoff of Some Prior Constructions

While we have presented below a review of the performance of some common space-time codes on the rate-diversity tradeoff, a more exhaustive list can be found in [69].

1. **Orthogonal Designs**

   The constructional details of the orthogonal designs, introduced in [22, 23] can be found in Section 2.4.1. Orthogonal designs achieve maximal diversity advantage and have rate $R' \leq 1$.

   \begin{itemize}
   \item Let $n_t = 2^{4c+d}(2b+1)$, $b$, $c \geq 0$, $0 \leq d \leq 3$. Then the maximum possible symbol rate of an $(n_t \times T)$ real, square (i.e., $T = n_t$) orthogonal design is
   \end{itemize}

\(^1\)A similar observation is made in [36], where a distinction is made between space-time codes over an alphabet $\mathcal{A}$ and space-time codes completely over an alphabet $\mathcal{A}$.
Chapter 9. Generalizations of the RDST Code

given by

\[ R' = \frac{8c + 2^d}{n_t}. \]

Thus, real square orthogonal designs with maximum possible rate \( R' = 1 \) exist only for \( n_t = 2, 4 \) or 8.

- The smallest possible value of \( T \) for which a real, orthogonal design of rate \( R' = 1 \) with \( n_t \) transmit antennas exists is the smallest value of \( T = 2^{4c+d}(2b+1) \), \( 0 \leq d \leq 3 \), \( b, c \geq 0 \), such that the associated integers \( c, d \) satisfy

\[ 8c + 2^d \geq n_t. \]

Thus, rate \( R' = 1 \) real orthogonal designs are rarely of minimal delay.

- Let \( n_t = 2^a(2b+1) \), \( a, b \geq 0 \). Then the maximum possible rate of an \((n_t \times T)\) complex, square orthogonal design is given by

\[ R' = \frac{a + 1}{n_t}. \]

Thus, square, complex, orthogonal designs with maximum possible rate \( R' = 1 \) exist only for \( n_t = 2 \).

- For \( T \geq n_t \geq 3 \), the maximum possible rate of an \((n_t \times T)\) complex orthogonal design is given by

\[ R' = \frac{T - 1}{T}. \]

This proves the non-existence of rate-one STBC from generalized linear processing complex orthogonal designs for more than two antennas.

In summary, orthogonal designs provide maximum the possible diversity advantage of \( n_t n_r \). For this diversity advantage, the maximum possible symbol rate is \( R' = 1 \). With the exception of the Alamouti code, complex orthogonal designs cannot achieve symbol rate \( R' = 1 \). Real orthogonal designs that are of rate \( R' = 1 \) can be found. However, the corresponding space-time block codes are restricted to real
alphabets and are of minimum possible delay $T = n_t$ only in the cases $n_t = 2, 4,$ or $8.$

2. **BPSK and QPSK Constructions of Hammons & El-Gamal**

In [63], Hammons and El Gamal introduced binary design criteria that facilitated the algebraic design of space-time codes over BPSK and QPSK constellations, that achieve maximal spatial diversity. It is shown how, starting with a collection of maximal-rank $(n_t \times T)$ binary $\{0, 1\}$ matrices closed under modulo-2 addition, one can construct a family of real, $\{\pm 1\}$ matrices suitable for BPSK modulation whose pairwise difference is always of full-rank. A similar construction for sets of $\{\pm 1, \pm i\}$ matrices, corresponding to QPSK modulation, with pairwise difference having maximal rank is given. The starting point in the complex case is sets of quaternary matrices over the integers modulo 4, that are closed under addition, having the additional property that certain binary projections of these matrices have full-rank. Methods of constructing the requisite binary and quaternary matrices are presented. While the paper mainly focuses on rank, some example codes of rate 1 including codes for $n_t = 2$ and $n_t = 3$ transmit antennas are also provided.

3. **Miscellaneous Algebraic Constructions**

The algebraic space-time code of Damen et al. [28], presented in Section 3.4, is a full-rank code has symbol rate $R' = 1$, and hence achieves the rate-diversity tradeoff.

The diagonal algebraic space-time (DAST) code construction by Damen et al. [29] involve Hadamard matrices. Here the authors first rotate a given base alphabet to ensure that no two symbols in the alphabet share a coordinate in common. The codes achieve rate $R' = 1$ and also have maximal diversity. Hence, they are rate-diversity optimal.

In [30], the authors construct maximal diversity space-time codes with symbol rate $R' = 1$, whose alphabet lies in a cyclotomic extension of the rational numbers $\mathbb{Q}$ that contain $\mathbb{Q}(i)$. 


The concept of threaded algebraic space-time (TAST) codes is introduced in [31] and algebraic construction techniques are provided. A generalization of the “single-layer” DAST code described above to multiple layers is provided in [33]. These TAST codes also achieve symbol rate $R' = 1$ and maximal diversity gain.

In [36], Sethuraman et al. provide a detailed exposition of how full-rank (often with symbol rate $R' = 1$) space-time codes can be constructed by making use of the regular representation of elements in either an extension field of the rational numbers $\mathbb{Q}$, or else, elements in a division algebra whose centre contains the rational numbers. Constructional details of this space-time code were presented in Chapter 6.

### 9.3 Comparing the Rate-Diversity and D-MG Tradeoffs

A few clarifications regarding the difference in formulation of the D-MG tradeoff from that of the rate-diversity tradeoff are in order here.

First, the two tradeoffs use different definitions of diversity. For clarity of exposition and to emphasize the difference, we have used the term ‘diversity gain’ while considering the D-MG tradeoff and ‘transmit diversity’ for the rate-diversity tradeoff. Further, in this thesis, we have used the term ‘diversity advantage’ for the standard definition of diversity in space-time coding literature (see, for example, [16]), which is $n_r$ times the transmit diversity. Diversity advantage and diversity gain differ in the following aspects:

- The diversity gain involves the SNR exponent of the codeword error probability, as against the pairwise error probability used in defining the diversity advantage and transmit diversity.

- The diversity advantage and transmit diversity are asymptotic performance measures of one fixed code. The diversity gain is however the performance metric of a scheme, which is a sequence of codes. The diversity gain is the speed with
which the error probability (of a maximum likelihood (ML) detector) decays as SNR increases.

Further, the definitions of rate used in the rate-diversity and D-MG tradeoffs are also different. While the rate diversity tradeoff uses the symbol rate $R'$ (9.1), the D-MG tradeoff formulation uses the rate $R$ in bits per channel use (3.1). The above mentioned differences in formulation therefore leads to two different tradeoffs.

Both tradeoffs are useful performance measures for a space-time code. For space-time codes that use a finite fixed constellation, the rate-diversity tradeoff assumes significance. Most space-time systems implemented in practice use simple constellations such as QAM and PSK as their signalling alphabet and hence fall into this category. The D-MG tradeoff is a good performance measure in cases where an outer code with variable rate is used in conjunction with the space-time code.

Another desirable property of the rate-diversity tradeoff is that its computation is simple and tractable in most cases. In general, computation of the exact D-MG tradeoff of a given scheme is however a difficult task. However, bounds on the D-MG tradeoff presented in [53] can be easily applied for a given space-time scheme and are reasonably tight in most cases.

### 9.4 A Unified Construction of RDST Codes

The RDST codes are derived from what are known as maximal-rank codes. We will first review a few concepts and definitions from [69].

**Definition 23. (Rank-$\nu$ Codes)** Let $\mathbb{F}$ be a field and $\mathcal{Q} \subseteq \mathbb{F}$. Let $C$ be a (matrix) code over $\mathcal{Q}$, where the components of each codeword are arranged in the form of an $(m \times n)$ matrix. The code $C$ will be called an $(m \times n)$ rank-$\nu$ code over $\mathcal{Q}$ if the difference $\Delta C = C_1 - C_2$ of any two distinct matrices $C_1, C_2 \in C$ has rank $\geq \nu$.

The following result was first proven by Gabidulin [66], by making a connection with the Singleton bound of coding theory.
Theorem 29. (Singleton Bound) Let \( n \geq m \) and \( C \) be a rank-\( \nu \) \((m \times n)\) code over \( \mathbb{Q} \). Then
\[
| C | \leq | \mathbb{Q} | ^{n(m-\nu+1)}.
\] (9.4)

Proof. See [69]. \( \square \)

A rank-\( \nu \) code whose size meets the Singleton bound will be referred to as a maximal rank-\( \nu \) code. When \( \nu = \min\{m, n\} \), we will refer to the maximal code as a maximal, full-rank \((m \times n)\) code.

Methods to construct maximal rank-\( \nu \) binary codes (i.e., codes over \( \mathbb{Q} = \mathbb{F}_2 = \{0, 1\} \), the finite field of size two) are provided in [64, 69, 66]. They will form a basic ingredient for constructing the RDST codes.

The following theorem, the main result of [69], is a general construction of space-time codes that achieve the rate-diversity tradeoff given in (9.2). The construction is called a unified construction as by specializing the construction, it is possible to construct space-time block codes over a large class of signal constellations.

Theorem 30. (Unified Construction) Let \( K \) and \( s \) be positive integers. Let \( \theta \) be a complex, primitive \( 2^k \)th root of unity. Let \( \{C_{u,k} : 0 \leq u \leq s-1, 0 \leq k \leq K-1\} \) be a collection of \( sK \) maximal, rank-\( \nu \) \((n_t \times T)\) binary codes.

Let \( 0 \neq \eta \in 2\mathbb{Z}[\theta] \), where \( 2\mathbb{Z}[\theta] \) is the ideal generated by 2 in \( \mathbb{Z}[\theta] \), and \( \mathbb{Z}[\theta] \) is the ring of algebraic integers in the cyclotomic number field \( \mathbb{Q}(\theta) \).

Let
\[
\mu : C_{0,0} \times C_{0,1} \times \cdots \times C_{s-1,K-1} \to X \subset \mathbb{C}^{n_t \times T}
\]
be a map, termed “unified mapper”, defined by
\[
(C_{0,0}, C_{0,1}, \ldots, C_{s-1,K-1}) \mapsto \kappa \sum_{u=0}^{s-1} \eta^u \theta^{u \sum_{k=0}^{K-1} 2^k} C_{u,k}
\] (9.5)
where $\kappa$ is a non-zero complex number.

Then the resultant rank-distance space-time (RDST) code

$$
X = \left\{ \kappa \sum_{u=0}^{s-1} \eta^u \theta^{K-1} 2^k C_{u,k} : C_{u,k} \in C_{u,K} \right\}
$$

achieves transmit diversity $\nu$ and has symbol rate $R' = n_t - \nu + 1$. $X$ is, therefore optimal with respect to the rate-diversity tradeoff.

Proof. See [69].

It will be found useful to think of each value of the index $u$ in (9.5) as constituting a layer of the RDST code. We will also refer to the set of binary codes $\{C_{u,k}\}$ as the component binary codes of $X$.

Remark 8. The unified mapper $\mu$ in (9.5) should be interpreted on a component-by-component basis, i.e., the right-hand side of this map is an $(n_t \times T)$ matrix whose $(n,t)$th component equals

$$
\kappa \sum_{u=0}^{s-1} \eta^u \theta^{K-1} 2^k [C_{u,k}]_{n,t}
$$

where $[C_{u,k}]_{n,t}$ is the $(n,t)$th component of the binary code matrix $C_{u,k}$.

Remark 9. Note that Theorem 30 does not require that the maximal rank-$\nu$ codes $C_{u,k}$ to be distinct. Indeed, in practice it is expected that the codes will, in fact, be chosen to be identical, for the sake of simplicity in hardware/software implementation.

We will now show how RDST codes over the PAM, QAM and PSK constellations can be obtained as special cases of the unified mapper.

Example 9. (PAM Constellation) A PAM constellation results when we set $K = 1$, $\eta = 2$, $\theta = -1$ and $\kappa = 1$. The resultant space-time code $X$ now takes on the form

$$
X = \left\{ \sum_{u=0}^{s-1} 2^u (-1)^{C_{u,0}} : C_{u,0} \in C_{u,0} \right\}
$$
and has symbol rate \( R' = n_t - \nu + 1 \) if the ingredient codes \( C_{u,0} \) are \((n_t \times T)\) maximal, rank-\( \nu \), binary codes, \( 1 \leq \nu \leq n_t \). The signal constellation is easily verified to be the \( 2^s\)-PAM constellation given by

\[
\mathcal{A} = \{(2^s - 1 - 2u) : 0 \leq u \leq 2^s - 1\}.
\]

**Example 10. (QAM Constellation)** A QAM constellation results when we set \( K = 2, \eta = 2, \theta = i \) and \( \kappa = (1 + i) \) with \( i := \sqrt{-1} \). The resultant space-time code \( \mathcal{X} \) now takes on the form

\[
\mathcal{X} = \left\{ \sum_{u=0}^{s-1} 2^u C_{u,0} + 2 C_{u,1} : C_{u,0} \in C_{u,0}, \ C_{u,1} \in C_{u,1} \right\}
\]

and has symbol rate \( R' = n_t - \nu + 1 \) if the codes \( C_{u,0} \) and \( C_{u,1} \) are \((n_t \times T)\) maximal, rank-\( \nu \) binary codes, \( 1 \leq \nu \leq n_t \). Further specializing to the case when the maximal, rank-\( \nu \) codes \( C_{u,0} = C_{u,1} := C(\nu) \), we can rewrite the above space-time code as [69]

\[
\mathcal{X} = \left\{ \sum_{k=0}^{T-1} \left( -1 \right)^{A_k} + i \left( -1 \right)^{B_k} : A_k, B_k \in C(\nu) \right\} . \quad (9.6)
\]

The resulting constellation can now be seen to be the \( A_{\text{QAM}} \) constellation given by

\[
A_{\text{QAM}} = \{ a + ib \mid (-M + 1) \leq a, b \leq (M - 1), \ a, b \text{ odd} \}
\]

where \( M \) is even and \( M^2 = 4^s \).

**Example 11. (PSK Constellation)** A PSK constellation results when we set \( U = 1, \eta = 2, \theta \) to be a complex, primitive \( 2^K \)th root of unity and \( \kappa = 1 \). The resultant space-time code \( \mathcal{X} \) now takes on the form

\[
\mathcal{X} = \left\{ \sum_{k=0}^{K-1} 2^k C_{0,k} : C_{0,k} \in C_{0,k} \right\}
\]

and has symbol rate \( R' = n_t - \nu + 1 \) if the codes \( C_{0,k} \) are \((n_t \times T)\) maximal, rank-\( \nu \) codes,
The resulting constellation is clearly the $2^K$-ary PSK constellation

$$A = \{ \theta^k : 0 \leq k \leq 2^K - 1 \}.$$ 

In the sections to come, for clarity of exposition, we will restrict our attention to the RDST code over the $4^s$-ary QAM constellation given in (9.6). In the sequel, unless otherwise specified, the term ‘RDST code’ will refer to the QAM-RDST code of (9.6).

### 9.5 D-MG Tradeoff of the RDST code

In [52, 53], bounds on the D-MG tradeoff of the QAM-RDST construction are provided. The bounds from [53] for the $(4 \times 4)$ QAM-RDST code assuming $n_r = 4$ is given in Figure 9.2. This plot is actually a composite plot of the D-MG tradeoff of four schemes, the rank-$\nu$ RDST code for $\nu = 1, \ldots, 4$, with the plotted diversity gain at each $r$ being the maximum over the four schemes. Figure 9.3, which shows the D-MG plots of each constituent rank-$\nu$ RDST code separately leads us to some interesting observations.

Figure 9.2: Bounds on the D-MG Tradeoff of the RDST Code ($n_t = n_r = T = 4$) [53]
Figure 9.3: The ‘Inside Story’ of the Bounds on the D-MG Tradeoff of the RDST Code

We see that the full rank RDST design performs better in comparison to designs of lesser rank only in the low $r$ region. A different picture is seen close to the maximum multiplexing-gain, where the uncoded (rank $\nu = 1$) RDST performs best. The middle regions of $r$ are dominated in performance by the designs of intermediary ranks.

These observations suggest constructing RDST codes where the ranks of the component codes in the various layers are chosen individually based on the value of the multiplexing-gain parameter $r$. This leads us to the generalization of the RDST code presented in the following section.
9.6 The Variable-Rank RDST code

An \((n \times n)\) variable-rank RDST (v-r RDST) code \(X\) over the \(M^2\)-point alphabet \(A_{QAM}\) is described as follows:

\[
X = \left\{ \sum_{k=0}^{s-1} 2^k \left[ (-1)^{A_k} + i(-1)^{B_k} \right] : \{A_k, B_k\}_{k=0}^{s_n-1} \in \mathcal{C}(n), \right. \\
\left. \{A_k, B_k\}_{k=s_n}^{s_n+s_n-1} \in \mathcal{C}(n-1), \ldots, \{A_k, B_k\}_{k=s_n+\ldots+s_2}^{s-1} \in \mathcal{C}(1) \right\} \tag{9.7}
\]

where \(M^2 = 4^s\) and \(\mathcal{C}(\nu)\) denotes an \((n \times n)\) maximal rank-\(\nu\) binary code. Note that in this construction, to form a single v-r RDST codeword matrix, \(2^{s_k}\) codewords (not necessarily distinct) drawn from the code \(\mathcal{C}(k)\) are used for \(1 \leq k \leq n\), where \(s = \sum_{k=1}^{n} s_k\). By setting \(s_k = s\) for some \(k\), we recover the original fixed rank-\(k\) RDST code of (9.6). We proceed to show the gains in the D-MG tradeoff accrued over that of the RDST code.

9.7 D-MG Tradeoff of the Variable-Rank RDST Code

We consider the channel model in (4.2),

\[
Y = \theta H X + W,
\]

with the parameter \(\theta\) chosen to ensure that

\[
\mathbb{E}(\| \theta X \|_F^2) \leq T \text{ SNR}.
\]

The exponential inequality is sufficient in the above energy constraint since our interest is confined to computing the D-MG tradeoff of the space-time scheme.

Cardinality of the v-r RDST code: Equations (9.7) and (3.3) along with the maximal rank property of the constituent binary codes give us the size of the \((n \times n)\) v-r
RDST code in terms of its spatial multiplexing gain $r$ as

$$|\mathcal{X}| = (4^n)^s (4^{2n})^{s_{n-1}} \cdots (4^{n^2})^{s_1} = \text{SNR}^{nr}. \quad (9.8)$$

Using the fact that $M^2 = 4^s$ and setting $s_i = a_i s$, $0 \leq a_i \leq 1$ with $\sum_i a_i = 1$, results in

$$\text{SNR}^{nr} = (4^n)^r \left( \sum_{i=1}^{s} s_{i} a_{n-i+1} \right)^{-1} \Rightarrow \text{SNR} \left( \sum_{i=1}^{s} i a_{n-i+1} \right)^{-1} = M^2 := \text{SNR}^{\delta(r)} \text{ (say).} \quad (9.9)$$

**Computation of the average transmission energy:** The transmission alphabet of the v-r RDST code $\mathcal{X}$ is $\mathcal{A}_{\text{QAM}}$. Since the $M^2$-size constellation $\mathcal{A}_{\text{QAM}}$ is scalably dense (see Definition 12),

$$a \in \mathcal{A}_{\text{QAM}} \Rightarrow |a|^2 \leq M^2.$$

It can also be verified that if $a$ denotes an element in $\mathcal{A}_{\text{QAM}}$, then,

$$\mathbb{E}(|a|^2) \leq M^2.$$

Using the fact that for any code matrix $X \in \mathcal{X}$, we have $||X||_F^2 \leq 2n^2 M^2 \leq M^2$, a valid choice for $\theta$ is

$$\theta^2 \leq \frac{\text{SNR}}{M^2} = \text{SNR}^{1-\delta(r)}.$$

### 9.7.1 Upper Bound on Diversity Gain $d(r)$

Let $X$ be any codeword drawn from the v-r RDST code $\mathcal{X}$,

$$X = \sum_{k=0}^{s-1} 2^k \left[ (-1)^{A_k} \; + \; i(-1)^{B_k} \right].$$
For any $X \in \mathcal{X}$, there always exists a second codeword matrix $X'$ such that the difference matrix is

$$
\Delta X_0 = X - X' = (-1)^A_0 - (-1)^{A'_0}
$$

with all entries of $\Delta X_0$ belonging to the set $\{\pm 2, 0\}$. Let the rank of $\Delta X_0$ be $\nu$ and $l_1 \geq l_2 \geq \cdots \geq l_\nu$ be its $\nu$ non-zero ordered eigenvalues. We have,

$$
Tr(\Delta X_0 \Delta X_0^\dagger) = SNR^0 = \prod_{j=1}^{\nu} l_j.
$$

The corresponding pairwise error probability $PEP(X \rightarrow X')$ then serves as a lower bound to the codeword error probability $P_e$, i.e.,

$$
P_e \geq PEP(X \rightarrow X') \quad (9.10)
$$

Similarly, it follows that there exist for each $X \in \mathcal{X}$, codewords $X'$ such that the difference

$$
\Delta X_i = X - X' = 2^i[(-1)^A_i - (-1)^{A'_i}], \quad i = 1, 2, \ldots, s-1.
$$

The product of the $\nu$ non-zero eigenvalues of $\Delta X_i \Delta X_i^\dagger$ is given by

$$
\prod_{j=1}^{\nu} l_j = 2^i \text{SNR}^0. \quad (9.11)
$$

The pairwise error probability between the two codewords whose difference is $\Delta X_i$ (assuming $n$ receive antennas) is given in [16, 17, 21] to be

$$
PEP \leq \frac{1}{(\theta^2)^\nu \left(\prod_{i=1}^{\nu} l_i\right)^n}. \quad (9.12)
$$

When the fractions $a_i$ are fixed, the PEP depends on the rank $\nu$ of the difference matrix and it’s eigenvalues $l_i$. Contrary to the original RDST code, in the variable-rank RDST code, the PEP does not monotonically decrease with increasing rank $\nu$. To see this, consider two codeword matrices $X$ and $X'$ constructed using $\{A_k, B_k\}_{k=0}^{s-1}$ and
Chapter 9. Generalizations of the RDST Code

\{A_k', B_k'\}_{k=0}^{s-1}$ respectively,

\[
X = \sum_{k=0}^{s-1} 2^k \left[ (-1)^{A_k} \pm i(-1)^{B_k} \right]
\]

\[
X' = \sum_{k=0}^{s-1} 2^k \left[ (-1)^{A_k'} \pm i(-1)^{B_k'} \right].
\]

Suppose that the smallest value of $k$ for which either $A_k \neq A_k'$ or $B_k \neq B_k'$ is $k = k_{\text{min}}$. Then, the difference matrix

\[
\Delta X = X - X' = \sum_{k=k_{\text{min}}}^{s-1} 2^k \left[ (-1)^{A_k} - (-1)^{A_k'} + i(-1)^{B_k} - i(-1)^{B_k'} \right].
\]

The rank of $\Delta X$ depends on this value $k_{\text{min}}$. In fact, it can be shown using techniques similar to those used in [69] to prove Theorem 30, that,

\[
\text{rank}(\Delta X) = \begin{cases} 
  n, & \text{if } 0 \leq k_{\text{min}} \leq s_n - 1 \\
  \nu, & \text{if } s_{\nu+1} + s_{\nu+2} + \cdots + s_n \leq k_{\text{min}} \leq s_{\nu} + \cdots + s_n - 1, \ 1 \leq \nu \leq n - 1
\end{cases}
\]

(9.13)

As the rank $\nu$ of $\Delta X$ falls, the scaling factor $2^{k_{\text{min}}}$ which is common to all remaining component matrices of $\Delta X$ increases. From (9.11), this scaling factor increases the lower bound on $\prod_i l_i$ which in turn decreases the PEP. Our aim is thus to minimize the worst case PEP by optimizing the fractions $a_i$. This amounts to determining

\[
\text{arg max}_{a_i} \left\{ \min \left[ \left\{ (\theta^2)^{\nu} \left( \prod_{i=1}^{\nu} l_i \right)^n \right\}^{\nu=1} \right] \right\}.
\]

(9.14)

In above equation we determine the minimum denominator of (9.12) which in turn determines the worst case PEP among all ranks. This minimum denominator is maximized by optimizing over all proportions $\{a_i\}$ to minimize worst case PEP.
9.7.2 Lower Bound on $d(r)$

Let $X_0$ be the transmitted codeword. The error probability when $X_0$ is transmitted is

$$P_e(X_0) \leq \sum_{i=1}^{|X|-1} \text{PEP}(X_0 \rightarrow X_i) \leq \sum_{\nu=1}^n N_{\nu} \text{PEP}_{\nu},$$  \hspace{1cm} (9.15)

where $N_{\nu}$ is the number of $X_i$ for which $\Delta X = X_0 - X_i$ has rank $\nu$ and $\text{PEP}_{\nu}$ is the worst case PEP among all such $X_i$. Given a code with particular values of $\{s_i\}_{i=1}^n$, we have for each $\nu$ that,

$$N_{\nu} \leq \prod_{j=1}^{\nu} (4^{(n-j+1)n})^{s_j}. \hspace{1cm} (9.16)$$

Define $p_{\nu} = \sum_{j=1}^{n-\nu+1} (n-j+1)a_j$. Let $(\prod_{i=1}^n l_i)_{\min}$ be the lower bound on the product of eigenvalues of $\Delta X \Delta X^\dagger$, among all $\Delta X$ with rank $\nu$. Using this lower bound in (9.12) gives $\text{PEP}_{\nu}$.

Notice, that since both $\text{PEP}_{\nu}$ and $N_{\nu}$ are independent of the transmitted codeword $X_0$, we can write the codeword error probability as

$$P_e \leq \sum_{\nu=1}^n N_{\nu} \text{PEP}_{\nu}. \hspace{1cm} (9.17)$$

We noted earlier that as the rank of $\Delta X$ falls, the scalar multiple $2^{k_{\min}}$ of codeword matrices increases. It is evident that after factoring out this $2^{k_{\min}}$, what remains is an integer matrix with rank at least $\nu$, for which the product of the $\nu$ non-zero eigenvalues is lower bounded by $\text{SNR}^0$. This fact and (9.13) lead to

$$\prod_{j=1}^{\nu} l_j \geq 4^\nu \sum_{j=\nu+1}^{n} s_j \text{SNR}^0. \hspace{1cm} (9.17)$$
Using (9.12), (9.16) and (9.17) in (9.15),

\[ P_e(r) \leq \sum_{\nu=1}^{n} \left( \frac{S_{\mathrm{NR}} \delta(r)^{\nu p_{n-\nu+1}}}{S_{\mathrm{NR}}^{1-\delta(r)} \nu \sum_{j=\nu+1}^{n} a_j} \right) \]

By definition, for a given \( r \), the negative exponent of SNR in the above quantity is equal to a lower bound on \( d(r) \). We maximize this lower bound by optimizing with respect to the \( a_i \).

\[ d(r) \geq \max_{a_i \forall i} \left[ \min \left\{ n\nu (1 - \delta(r)) + \nu \delta(r) \sum_{j=\nu+1}^{n} a_j - n\delta(r)p_{n-\nu+1} \right\} \right] \]

(9.18)

In Figure 9.4, bounds on \( d(r) \) for \( n = n_r = 4 \) are plotted for a variable-rank RDST code (obtained using (9.18), (9.14) and the \( \lambda_{\text{min}} \) sphere bound [53]) in comparison to those for the fixed-rank RSDT code (bounds from [53]). Significant improvement in both lower and upper bounds is observed. Table 9.1 shows the proportions \( \{a_i\}_{i=1}^{4} \) for certain \( r \) which yield the best lower bound on \( d(r) \). These values are in agreement with our intuition that the proportion of lesser rank codes should increase with multiplexing gain.

<table>
<thead>
<tr>
<th>( r )</th>
<th>( a_4 )</th>
<th>( a_3 )</th>
<th>( a_2 )</th>
<th>( a_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.4</td>
<td>0.62</td>
<td>0.38</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.8</td>
<td>0.34</td>
<td>0.43</td>
<td>0.23</td>
<td>0</td>
</tr>
<tr>
<td>1.2</td>
<td>0.24</td>
<td>0.29</td>
<td>0.37</td>
<td>0.10</td>
</tr>
<tr>
<td>1.6</td>
<td>0.18</td>
<td>0.22</td>
<td>0.28</td>
<td>0.32</td>
</tr>
<tr>
<td>1.7-4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 9.1: Proportions \( a_i \) for certain \( r \) (\( n = 4 \))
Chapter 9. Generalizations of the RDST Code

9.8 Rate-Diversity Tradeoff of the v-r RDST Code

Evaluating the rate-diversity tradeoff of the v-r RDST code leads to some interesting observations. From (9.8), we evaluate the symbol rate

\[
R' = \log_4 \left( \frac{4^n \sum_{i=1}^{n} s_i (n-i+1)}{n} \right)
\]

For any given v-r RDST code \( \mathcal{X} \), we will assume that the least \( i \) for which \( s_i \) is non-zero is \( i = \nu \). This essentially means that the transmit diversity of \( \mathcal{X} \) is \( \nu \).

\[
R' = \log_4 \left( 4^{\sum_{i=\nu}^{n} s_i (n-i+1)} \right) = \sum_{i=\nu}^{n} a_i (n - i + 1)
\]

(9.19)

**Theorem 31.** The class of variable-rank RDST codes excluding the family of fixed-rank RDST codes (which is a special case of the v-r RDST construction) is strictly suboptimal on the rate-diversity tradeoff. In other words, the only codes from the family of all
variable-rank RDST codes that achieve the rate-diversity tradeoff are the family of RDST codes.

Proof. We wish to investigate the symbol rate of the v-r RDST code in (9.19) and show that the maximum possible value of \( R' = n - \nu + 1 \) is achieved by only the fixed-rank RDST. Setting \( a'_i = a_{n-i+1} \), we can rewrite (9.19) as

\[
R' = \sum_{i=1}^{n-\nu+1} a'_i i \quad \text{where} \quad \sum_{i=1}^{n-\nu+1} a'_i = 1.
\]

This is just a weighted sum of the first \( n - \nu + 1 \) integers. A simple analysis along the lines of the proof of Theorem 2 in [54] (the “mismatched eigenvalue bound”) leads us to conclude that the maximum value of \( R' = n - \nu + 1 \) is achieved only when \( a'_{n-\nu+1} = 1 \) and \( a'_i = 0 \) for \( 1 \leq i \leq n - \nu \). This corresponds to \( a_\nu = 1 \) and \( a_i = 0 \) for \( i \neq \nu \). Our claim is proved.

We note that the v-r RDST code in fact sacrifices performance on the rate-diversity tradeoff to attain a better D-MG tradeoff than the fixed-rank RDST code!

9.9 Other Generalizations of the RDST Code - A Review

The unified mapper in (9.5) provides a systematic method, termed unified construction [69] of constructing space-time codes over a wide variety of signal constellations \( \mathcal{A} \) whose size is a power of 2. These include the \( 2^m \)-PSK and the \( 4^m \)-QAM constellations. The authors in [69] subsequently generalize this unified construction to the cases when the size of the constellation \( \mathcal{A} \) is a power of some prime \( \varphi \), \( \varphi > 2 \). This generalized unified mapper makes use of a class of maximal, rank-\( \nu \), \( \varphi \)-codes over \( \mathbb{F}_\varphi \).

Motivated by the potential advantages of dual radii AM-PSK constellations over conventional PSK constellations, Hammons develops new space-time codes for AM-PSK constellations in [70, 72] that also achieve the rate-diversity tradeoff. Space-time code
constructions for both dual-radii (see Figure 9.5) and multi-radii AM-PSK constellations are provided. Further, a “super-unified” space-time code construction that incorporates both the Lu-Kumar unified RDST construction and the new multi-radii AM-PSK codes is presented.

In [71], Lu generalizes Hammons’ dual-radii construction to the cases when the size of the constellations to the cases when the size of the dual-radii constellation is a power of some prime $\wp$, $\wp \geq 2$. The resulting space-time code is also rate-diversity optimal and has an AM-PSK constellation with signal alphabets distributed over $\wp$-concentric circles in the complex plane, i.e., there are $\wp$ radii. The generalized construction is thus termed “generalized $\wp$-radii construction”. Also, several rich classes of subset-subcodes of the newly constructed space-time codes are identified and are shown to be rate-diversity optimal. Finally, a “generalized super-unified construction” is presented by extending the super-unified construction to the $\wp$-adic case.
9.10 Generalized RDST Codes Via Local Power-Series Expansion

In this section, we further extend the class of RDST codes to include those having a description in terms of power-series expansion with respect to a prime ideal of a number field.

Let $\mathbb{Q}(\theta)$ be a number field with $\mathbb{Z}[\theta]$ as the ring of integers. Let $\beta$ be a generator of the maximal ideal $I$. Initially we will assume that $I = \langle \beta \rangle$ but we will subsequently relax this assumption. Let $\mathbb{F}_q$ denote the finite field $\mathbb{F}_q = \mathbb{Z}[\theta]/I$ of size $q = p^f$ where $p$ is prime and $q$ is the norm of the ideal $I$.

Let $A \subseteq \mathbb{C}$ be the signal constellation given by

$$A = \left\{ s - 1 \sum_{i=0}^{s-1} a_i \beta^i \mid a_i \in \mathbb{F}_q \right\}.$$

By $a_i \in \mathbb{F}_q$, we mean that $a_i$ is a suitably chosen pre-image in $\mathbb{Z}[\theta]$ associated to the isomorphism $\mathbb{Z}[\theta]/I \cong \mathbb{F}_q$. Let $C(\nu)$ be an $(n \times n)$ maximal rank code over $\mathbb{F}_q$. This means that the code is of maximal size $(q^\nu)^{n-\nu+1}$ given the rank requirement. While constructions for such codes have previously been provided for $q$ prime, these constructions are easily extendable to the case of any finite field.

9.10.1 RDST Codes via Power-Series Expansion

We now define our space-time code $X$ to be given by:

$$X \in X \iff X = \sum_{i=0}^{s-1} A_i \beta^i, A_i \in C(\nu), \forall i.$$

**Proposition 32.** Let $X_1, X_2 \in X$, $X_1 \neq X_2$. Then as a complex matrix, $\Delta X = X_1 - X_2$ has rank $\geq \nu$.

**Proof.** (The proof is identical to that in [69].) We focus on the $\nu = n$ case. The general case follows, by restricting attention to a submatrix of $X$ of size $\nu \times \nu$. Assume that the
rank is $< n$. Let $\mathbf{u} \in \mathbb{Q}(\theta)$ be such that $\mathbf{u}^T \Delta \mathbf{X} = \mathbf{0}^T$. Without loss of generality, we can assume that $\mathbf{u} \in \mathbb{Z}[\theta]$. Thus if

$$X_1 = \sum_{i=0}^{s-1} A_i \beta^i, \quad X_2 = \sum_{i=0}^{s-1} B_i \beta^i$$

$$\Rightarrow \mathbf{u}^T \mathbf{0}^T \Leftrightarrow \mathbf{u}^T \sum_{i=0}^{s-1} \beta^i [(A_i - B_i)] = \mathbf{0}^T$$

$$\Rightarrow \mathbf{u}^T [A_0 - B_0] = \mathbf{0}^T \pmod{I}.$$  

If $\mathbf{u} \neq \mathbf{0} \pmod{I}$ we are done. If $\mathbf{u} = \mathbf{0} \pmod{I}$ we have $\beta | \mathbf{u}$ and can set $\mathbf{u} = \beta \mathbf{v}$. Clearly $\mathbf{u}^T (A_0 - B_0) = \mathbf{0}^T \pmod{I}$ with $\mathbf{v} \neq \mathbf{0} \pmod{I}$ which is a contradiction. If $A_0 = B_0$ we can repeat a similar argument with $\Delta \mathbf{X}$

Note that the space-time code $\mathcal{X}$ constructed above is of size $q^{n(n-\nu+1)s}$ and is hence optimal with respect to the rate-diversity tradeoff [64].

**Variation Using a Subset of the Constellation** Consider a variation of the above construction in which $\theta, \beta, I, q$ all remain as above in Section 9.10, but where this time, each code matrix $\mathbf{X}$ in the space-time code has the modified expansion

$$\mathbf{X} \in \mathcal{X} \Leftrightarrow \mathbf{X} = A_0 + \sum_{i=1}^{e_1-1} J_i(A_0) \beta^i + A_1 \beta^{e_1} + \sum_{i=e_1+1}^{e_2-1} J_i(A_0, A_1) \beta^i + A_2 \beta^{e_2} + \cdots$$

in which the integers $e_i$ satisfy $1 \leq e_1 < e_2 < \cdots < e_r < s$ for some integer $r$ and where each $J_i$ is a matrix over $\mathbb{F}_q = \mathbb{Z}[\theta]/I$ which is an arbitrary componentwise function of its matrix arguments, i.e., each function $J_i$ operates on a component-by-component basis and the same arbitrary function applies to all components. It can be verified that $A_i \in \mathcal{C}(\nu)$ guarantees that the space-time code $\mathcal{X}$ has all difference matrices of rank $\nu$ or larger and that the resulting space-time code achieves the rate-diversity tradeoff over the corresponding constellation.
This generalization permits considerable variation in signal constellations. For instance, the space-time codes of [70] having the AM-PSK constellation shown in Figure 9.5 can be so derived.

9.10.2 Case When $\beta$ is Not a Principal Ideal

Let $K = \mathbb{Q}(\theta)$ be a number field such that $R = \mathbb{Z}[\theta]$ is the corresponding ring of integers. Let $\beta$ be a prime ideal in $R$ and $\gamma \in \beta$. The focus here is on the case when $\beta$ is not a principal ideal. Let $\mathbb{F}_q = R/\beta$ where $q$ denotes the size of the finite field. Let the $(n \times n)$ matrix $X$ over $R$ have the decomposition

$$X = C_0 + \gamma C_1 + \cdots + \gamma^{s-1} C_{s-1}$$

where the matrices $C_i$ are either $[0]$ or else such that their reductions $D_i = C_i \pmod{\beta}$ are of full rank over $\mathbb{F}_q$. We assume that at least one of the $C_i$ is nonzero.

**Theorem 33.** $X$ has full rank over $K$.

**Proof.** Suppose not. We first assume that $C_0 \neq [0]$. Then without loss of generality, we can assume that

$$u^T X = 0^T$$

for some $u \in \mathbb{Z}[\theta]^n$. Suppose $u = 0 \pmod{\beta^j}$ but $u \neq 0 \pmod{\beta^{j+1}}$. Reducing both sides modulo $\beta^{j+1}$ yields

$$u^T X = 0^T \pmod{\beta^{j+1}}. \quad (9.20)$$

Let $R_\beta$ denote the local ring obtained by localization with respect to $\beta$. Noting that for any $j \geq 1$, $R/\beta^j R \cong R_\beta/\beta^j R_\beta$, we have that (9.20) has the local ring version

$$u^T X = 0^T \pmod{\beta^{j+1} R_\beta}. \quad (9.21)$$

Let $\pi$ be the local generator of $\beta R_\beta$, i.e., $\beta R_\beta = < \pi >$. Since $R_\beta$ is a PID and $u \in \beta^j R_\beta^n$, 

we can set \( u = \pi^l v \), for some \( v \in R^0_\beta \). With this, we can rewrite (9.21) as

\[
v^T C_0 = 0^T, \quad (\text{mod } \beta R_\beta).
\]

Next, if

\[
v_i = \frac{a_i}{b_i}, \quad a_i \in R, \quad b_i \in R \setminus \beta,
\]

we note that \( b = \prod_{i=1}^n b_i \) does not belong to \( \beta \). By replacing \( v \) by \( b v \), we obtain

\[
(bv)^T C_0 = 0^T, \quad (\text{mod } \beta R_\beta), \quad bv \in R^n,
\]

or equivalently,

\[
(bv)^T C_0 = 0^T, \quad (\text{mod } \beta), \quad bv \in R^n.
\]

This however is a contradiction since \( C_0 \) (mod \( \beta \)) has full rank over \( \mathbb{F}_q \).

Next consider the case when \( C_0 = [0] \). Then

\[
\frac{X}{\gamma} = C_1 + \gamma C_2 + \cdots + \gamma^{s-1} R_{s-1}
\]

and we are in a position to repeat the previous argument with \( C_1 \) in place of \( C_0 \) etc. \( \square \)

### 9.10.3 Examples

**Example 12.** Let \( p \) be prime, \( \theta = \exp \left( \frac{2\pi i}{p} \right) \). Then we can choose \( \beta = (1 - \theta) \), i.e., choose \( I = < 1 - \theta > \). From number theory it follows that the norm \( N(I) = p \). As a special case of this example, if we set \( p = 2, \theta = -1 \), then \( \beta = 2 \) and \( N(I) = 2 \), we recover the PAM- RDST code in [69].

**Example 13.** Let \( \theta = \exp \left( \frac{2\pi i}{m} \right) \). Let \( p \) be a prime such that the minimal (cyclotomic) polynomial \( f(x) \) of \( \theta \) over \( \mathbb{Z} \) is irreducible (mod \( p \)). Let

\[
m = p^a m_1, \quad p \nmid m_1,
\]

and
and let the order of $m_1 \pmod{p} = f = \phi(m)$. Then in $\mathbb{Z}[[\theta]]$, $p$ remains prime and

$$\mathbb{Z}[[\theta]]/p\mathbb{Z}[[\theta]] \cong GF(p^f)$$

We now choose $\beta = p$ and proceed as before.

For instance, we could take $m = 9, p = 2$. 
Chapter 10

Simulation Results, Some Miscellany and Future Work

This chapter presents simulation results for the square D-MG optimal CDA based space-time codes presented in Chapter 7. The performance of our construction is compared with other schemes from the literature.

The D-MG tradeoff is an excellent performance metric of space-time schemes under the assumption of very high SNR. At low values of SNR, the performance of space-time schemes are in addition dictated by factors such as constellation shaping to ensure energy efficiency. This thesis does not address the issue of performance at finite SNR. However, we review a recent work in this direction, the “perfect codes” construction in [41].

By providing an explicit construction of space-time codes for $T \geq n_r$, our work contributes to the theory behind the D-MG tradeoff by establishing the exact optimal tradeoff for all $T \geq n_t$. This constitutes an improvement to the result of Zheng and Tse [47], who were able to establish the same only for $T \geq n_t + n_r - 1$. This is made precise in Section 10.3.

Another topic of recent interest is that of “approximate universality” [15, 77, 78, 80]. Using this concept, it has been concluded from the analysis in [80] that the space-time code constructions in Chapters 7 and 8 of this thesis not only achieve the D-MG tradeoff of the i.i.d. Rayleigh fading channel but achieve the D-MG tradeoff of every
fading channel. This result has key implications in considering our space-time codes for implementation in practical systems, where the actual channel characteristics might not be known apriori.

## 10.1 Decoding and Simulation of Square ST Codes Derived from CDA

We first introduce the concept of a linear dispersion code from [76].

**Definition 24.** A linear dispersion code $\mathcal{X}$ consists of $(n_t \times T)$ matrices of the form

$$X = \sum_{q=1}^{Q} s_q A_q + s_q^* B_q$$

where $s_q$ belongs to a signal set $\mathcal{A}$ and $A_q, B_q$ where $q = 1, 2, \ldots, Q$ are $2Q$ number of $(n_t \times T)$ complex matrices that define the code.

It is seen that the CDA based space-time codes presented in this thesis are linear dispersion codes. It has been shown in [76] that linear dispersion codes can be decoded using the sphere decoding algorithm [73, 74, 75]. The sphere decoder implements a low complexity ML decoding algorithm.

As an illustrative example, we present simulation results for the $(3 \times 3)$ space-time code constructed in Chapter 7 assuming 3 receive antennas in Figure 10.1. The base alphabet assumed is the 4-point $\mathcal{A}_{QAM}$ constellation. Shown alongside for comparison are the $3 \times 3$ codes of Sethuraman et al. [36] which use non-norm elements that are transcendental (with same base constellation), and the $3 \times 3$ “perfect codes” of Oggier et al. [41], which uses a 4-point HEX base constellation [41].

Note that the codes in [36] are not endowed with the non-vanishing determinant property. It is not clear what the exact D-MG tradeoff of these codes are. It has been conjectured through simulations in [51] that these codes achieve the D-MG tradeoff, but this is yet to be confirmed analytically.
The perfect codes presented in [41] are designed to ensure good performance at finite SNRs. A review follows in the subsequent section.

10.2 Review of Perfect Codes

“Perfect codes” are square designs derived from CDA presented in [41] for number of transmit antennas $n_t = 2, 3, 4, 6$. The basic structure of the perfect codes is the same as those presented in Chapter 6. Non-vanishing determinant and constellation shaping are the two key properties of perfect codes.

**Definition 25.** [41] A square $(n_t \times n_t)$ space-time block code $X$ is called a perfect code if and only if

- it is a full rate linear dispersion code using $n_t^2$ information symbols drawn from either QAM or HEX.
- $X$ is a full-rank code.
Chapter 10. Simulation Results, Some Miscellany and Future Work

• the $2n_t^2$-dimensional real lattice generated by the vectorized codewords is either $\mathbb{Z}^{2n_t^2}$ or $A_2^{n_t^2}$, where $A_2$ denotes the hexagonal lattice [41].

• it induces uniform average transmitted energy per antenna in all $T$ time slots, i.e., all the coded symbols in the code matrix have the same average energy.

Notice that the restriction of uniform average transmitted demands that the non-norm element $\gamma$ have unit magnitude, which incidentally is the least possible value that $|\gamma|$ can assume.

The idea behind the constellation shaping is as follows. The transmission alphabet of space-time codes derived from CDA is a finite subset of the maximal subfield $\mathbb{L}$. In our approach (Chapter 6), we choose a finite subset of the centre $\mathbb{F} = \mathbb{Q}(i)$ (the base constellation $A_{QAM}$), which chooses a finite subset of $\mathbb{L}$ indirectly through an integral basis. In [41], the reverse approach is pursued. A suitable subset of $\mathbb{L}$ is first chosen, which dictates the base constellation. In choosing this transmission alphabet, energy efficiency is ensured as follows.

The transmission alphabet is chosen to be a finite subset of a suitably chosen ideal $I$ of $\mathcal{O}_\mathbb{L}$. Each ideal may be associated to a complex algebraic lattice $\Lambda(I)$ through what is known as the canonical embedding (see [61, 41] for details). The basic idea is to choose $I$ such that $\Lambda(I)$ is a rotated version of either the $\mathbb{Z}^{2n_t}$ or $A_2^{n_t}$ lattices (according to whether QAM or HEX is the base alphabet).

The above properties ensure that the perfect codes have good performance at finite SNRs. That the perfect codes are D-MG optimal is a consequence of the sufficient condition in [54], reviewed in this thesis in Chapter 4. However, such perfect codes are known to exist only for the number of transmit antennas being $n_t = 2, 3, 4$ or 6.

10.3 Determining the Exact D-MG Tradeoff

In [47], the authors provide an exact characterization of the optimal D-MG tradeoff for the case when $T \geq n_t + n_r - 1$. This characterization is obtained by demonstrating the achievability of the outage curve of the i.i.d. Rayleigh fading channel by random
Gaussian codes with $T \geq n_t + n_r - 1$ (See Section 3.3). However, for $T < n_t + n_r - 1$, Zheng and Tse were only able to provide upper and lower bounds on the optimal D-MG tradeoff.

By providing an explicit construction of D-MG optimal space-time codes for all $T \geq n_t$ (Chapters 7 and 8), our results contribute to the theory of the D-MG tradeoff by showing that the Zheng-Tse upper bound on the optimal D-MG tradeoff is in fact achievable for all $T \geq n_t$.

**Corollary 34.** The optimal D-MG tradeoff for the $n_t$ transmit $n_r$ receive antenna i.i.d. Rayleigh fading channel $d^*(r)$ for all values of $T \geq n_t$ is given by the piecewise-linear function connecting the points $(k, d^*(k))$, $k = 0, 1, \ldots, \min\{n_t, n_r\}$, where

$$d^*(k) = (n_t - k)(n_r - k).$$

### 10.4 Approximate Universality

A space-time code is said to be *approximately universal* if it optimally trades off the diversity and multiplexing gains for every statistical characterization of the fading channel [15, 80, 77, 78]. Thus, an approximately universal code achieves the outage curve of every slow fading channel. Approximately universal codes are especially significant for practical systems, where it might not be possible to model the channel accurately.

In [80, 78, 77], the authors present a necessary and sufficient condition that ensures that a space-time scheme is approximately universal. The approximate universality condition constrains the smallest $\min\{n_t, n_r\}$ singular values of the normalized codeword difference matrix.

**Theorem 35.** A space-time code communicating at rate $R$ bits/symbol is approximately universal over the MIMO channel if and only if, for every pair of codewords,

$$\lambda_1^2 \lambda_2^2 \ldots \lambda_{\min\{n_t, n_r\}}^2 > \frac{\text{SNR}^{\min\{n_t, n_r\}}}{2^R + o(\log \text{SNR})},$$
where $\lambda_1, \lambda_2, \ldots, \lambda_{\min\{n_t, n_r\}}$ are the smallest $\min\{n_t, n_r\}$ singular values of the normalized codeword difference matrix.

Proof. See [80].

The above condition on the singular values of the codeword matrices essentially ensures that the probability of error for the code when the channel is not in outage decays exponentially with SNR. From (3.7), it is evident that this ensures the desired result of the space-time code achieving the outage probability.

The square and rectangular space-time codes constructed Chapters 7 and 8 are D-MG optimal over the i.i.d. Rayleigh fading channel, as a consequence of satisfying the sufficient condition for D-MG optimality (Theorem 7) reviewed in Chapter 4. Further, it is immediate from Theorem 35 that these codes also satisfy the approximate universality condition and hence achieve the outage curve of any slow fading channel.
Appendix A

D-MG Optimality of the Puncturing Construction

Before proving the optimality of the puncturing construction on the D-MG tradeoff, we will make note of the following useful lemma [10] 1. In the sequel, the notation $\ell_k(A)$ will denote the $k^{th}$ ordered eigenvalue of the matrix $A \in \mathbb{C}^{n \times n}$, assumed to be arranged in increasing order,

$$\ell_1(A) \leq \ell_2(A) \leq \cdots \leq \ell_n(A).$$

Lemma 36. Let $A \in \mathbb{C}^{T \times T}$ be a Hermitian matrix, $n_t$ be an integer such that $1 \leq n_t \leq T$, and $A_{n_t}$ denote any $n_t$-by-$n_t$ principal submatrix of $A$ (obtained by deleting $T - n_t$ rows and the corresponding columns from $A$). For each integer $k$ such that $1 \leq k \leq n_t$, we have,

$$\ell_k(A) \leq \ell_k(A_{n_t}) \leq \ell_{k+T-n_t}(A).$$

Proof. (Theorem 27) The channel model assumed is the one given in (4.2). $\mathcal{X}'$ is a ($T \times T$) D-MG optimal space-time code with non-vanishing determinant, i.e.,

$$\min_{\Delta X'} \det[\Delta X'\Delta X'^{\dagger}] \geq \text{SNR}^0,$$

1I would like to thank Prof. B. A. Sethuraman for having pointed out this lemma.
where $\Delta X'$ denotes the difference of any two distinct matrices from $X'$. From (4.5), the energy normalizing factor

$$\theta^2 = \text{SNR}^{1-\hat{r}}.$$  

If $X'$ is any codeword matrix drawn from $X'$, the energy constraint in (4.3) gives us that

$$\|\theta X'\|_F^2 \leq T \text{SNR} \Rightarrow \|X'\|_F^2 \leq \text{SNR} \hat{r}.$$  

Consider the $(n_t \times T)$ space-time code $X$ derived from $X'$ by retaining the first $n_t$ rows and let $\Delta X$ denote the difference of any two distinct matrices drawn from it. Then,

$$\Delta X' \Delta X' = \begin{bmatrix} \Delta X \\ \Delta Z \end{bmatrix} \begin{bmatrix} \Delta X' \Delta Z' \\ \Delta Z' \Delta X' \end{bmatrix}$$

for some $\Delta Z \in \mathbb{C}^{(T-n_t) \times T}$. In order to derive an expression for the parameter $\delta_X$, we use Lemma 36. Let $\ell_{\text{max}}$ be the maximum eigenvalue of $\Delta X' \Delta X'^\dagger$. Then direct application of Lemma 36 along with $\ell_{\text{max}} \leq \|\Delta X'\|_F^2 \leq \text{SNR} \hat{r}$ gives

$$\min_{\Delta X} \det(\Delta X' \Delta X'^\dagger) \geq \frac{\min_{\Delta X'} \det(\Delta X' \Delta X'^\dagger)}{(\ell_{\text{max}})^{T-n_t}} \geq \text{SNR}^{-\frac{(T-n_t)}{\hat{r}}}.$$  

The parameter $\delta_X$ for $X$ is evaluated as

$$\text{SNR}^\delta = \min_{\Delta X} \det[(\theta \Delta X)(\theta \Delta X)^\dagger] = (\theta^2)^{n_t} \min_{\Delta X} \det[\Delta X' \Delta X'^\dagger] \Rightarrow \delta_X = n_t - r.$$  

$X$ therefore satisfies Theorem 7 and hence achieves the optimal D-MG tradeoff.
Appendix B

D-MG Optimality of the Stacking Construction

The following lemma from [10] will be useful in proving the D-MG optimality of the stacking construction.

Lemma 37. Let $A, B \in \mathbb{C}^{n \times n}$ be Hermitian. Assume that $B$ is positive semidefinite and that the eigenvalues of $A$ and $A + B$ are arranged in increasing order. Then

$$\ell_k(A) \leq \ell_k(A + B) \text{ for all } k = 1, 2, \ldots, n.$$ 

Proof. (Theorem 28) We will assume a normalized rate of transmission $R = r \log \text{SNR}$ bits/channel use. From (3.3), the cardinality of the space-time code is given as $|\mathcal{X}| = \text{SNR}^{rT}$. Using the fact that in the rectangular design, there are $tk$ independent variables that assume values from the maximal subfield $\mathbb{L}$, an expression for $M^2$, the size of the base QAM constellation can be derived as

$$|\mathcal{X}| = \left[(M^2)^t\right]^{tk} = \text{SNR}^{Tr}$$

$$\Rightarrow M^2 = \text{SNR}^\frac{r}{t}.$$
This also implies that the energy normalizing factor

\[ \theta^2 = \frac{\text{SNR}}{M^2} = \text{SNR}^{1-\epsilon}. \]

The difference between any two distinct matrices in \( \mathcal{X} \), \( \Delta X \), is of the form

\[ \Delta X = [\Delta X_1 \ldots \Delta X_k], \]

where \( \Delta X_i \) denotes the corresponding difference matrix in \( \mathcal{X}_i \) \( \forall i \). Choosing some value \( j \) such that \( \Delta X_j \neq 0 \), we can write

\[ \Delta X \Delta X^\dagger = \Delta X_j \Delta X_j^\dagger + \sum_{l \neq j} \Delta X_l \Delta X_l^\dagger. \]

Let \( \{ \ell_i \}_{i=1}^{n_t} \) and \( \{ \ell'_i \}_{i=1}^{n_t} \) denote the eigenvalues of \( \Delta X \Delta X^\dagger \) and \( \Delta X_j \Delta X_j^\dagger \) respectively. From Lemma 37, we have,

\[ \ell'_i \leq \ell_i \ \forall \ i = 1, 2, \ldots, n_t \]

\[ \Rightarrow \min_{\Delta X} \det(\Delta X \Delta X^\dagger) = \min_{\Delta X} \prod_{i=1}^{n_t} \ell_i \]

\[ \geq \min_{\Delta X_j} \prod_{i=1}^{n_t} \ell'_i \]

\[ \geq \text{SNR}^{-\frac{n_t-r}{r}} \quad \text{(from (A.1))}. \]

In the framework of Theorem 7, this essentially means that for \( \mathcal{X} \), we have

\[ \text{SNR}^{\delta_X} = \min_{\Delta X} \det[(\theta \Delta X)(\theta \Delta X)^\dagger] \]

\[ \Rightarrow \delta_X = n_t - r. \]

Hence from Theorem 7, \( \mathcal{X} \) achieves the optimal D-MG tradeoff. \( \square \)
References


REFERENCES


[49] L. Grokop and D. N. C. Tse, “Diversity Multiplexing Tradeoff in ISI channels,”


[51] V. Shashidhar, B. Sundar Rajan and P. Vijay Kumar, “STBCs with optimal
diversity-multiplexing tradeoff for 2,3 and 4 transmit antennas,” Proc. IEEE Int.

[52] P. Vijay Kumar, Petros Elia, Hsiao-feng (Francis) Lu, Sameer Pawar and K. Raj
Kumar, “On the Decoding and Diversity-Multiplexing Gain Tradeoff of a Recent

[53] Petros Elia, P. Vijay Kumar, Sameer Pawar, K. Raj Kumar, B. Sundar Rajan and
Hsiao-feng (Francis) Lu, “Diversity-Multiplexing Tradeoff Analysis of a few Algebraic

[54] Petros Elia, K. Raj Kumar, Sameer A. Pawar, P. Vijay Kumar and Hsiao-feng Lu,
“Explicit, Minimum-Delay Space-Time Codes Achieving The Diversity-Multiplexing

[55] Petros Elia, K. Raj Kumar, Sameer A. Pawar, P. Vijay Kumar and Hsiao-feng Lu,
“Explicit, Minimum-Delay Space-Time Codes Achieving The Diversity-Multiplexing
Gain Tradeoff,” Accepted for presentation in the 2005 IEEE Intl. Symp. Info. Theory
(ISIT-05), Sept. 2005.

[56] Hsiao-feng Lu, Petros Elia, K. Raj Kumar, Sameer A. Pawar and P. Vijay Ku-
mar, “Space-time codes meeting the diversity-multiplexing gain tradeoff with low


REFERENCES


