University of California
Los Angeles

Universal Channel Codes and Trellis State-Diagram Reduction

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Electrical Engineering

by

Jun Shi

2005

The dissertation of Jun Shi is approved.

Michael P. Fitz

Kung Yao

Kirby A. Baker

Richard D. Wesel, Committee Chair

University of California, Los Angeles

2005

To my parents

Table of Contents

List of Figures

List of Tables

Acknowledgments

When I arrived at UCLA, I was hesitating what I should pursue for my Ph.D. study. My advisor Rick Wesel provided me with a position in his research group. His support and encouragement have been the driving force behind my success. I am extremely grateful for his help with my academic and personal life.

I would like to thank Prof. Yao and Prof. Fitz for giving me their invaluable advice on my research. I also want to thank Prof. Baker from the Mathematics Department for serving on my committee. Also from the Mathematics Department, Prof. Biskup taught me probability theory, random walks and stochastic processes. He trained me with the rigorousness of thinking, which shaped my way of conducting research. His profound knowledge in physics and mathematics helped me to understand my research from a totally different perspective.

It was my pleasure to be a member of Wesel's research group. Working with my teammates was both inspiring and exciting. I particularly enjoyed conversations with Adina, Cenk, Wenyen, Aditya and Chris during mid-afternoon coffee breaks. Adina's cubicle was just next to me, she told me funny stories about JPL. Cenk had talent in reading fortune from coffee. Wenyen was my tennis partner, who instructed me on forehand and backhand strokes. It was very joyful to discuss with Aditya about network coding problems. Our efforts resulted in an information theory paper. Chris introduced the FPGA implementation of LDPC codes to me. I also enjoyed working with junior members of the group Andres, Esteban and Miguel. My thanks extend to the alumni: Wei, Tom, Xueting, Christina and Mark. They helped me in various stages of my study.

Thanks to Tracey Ho at MIT with whom I collaborated on network coding papers. I appreciate the friendship with James Liu, who was always warm-hearted.

I successfully convinced Xiaolong to move into the UCLA apartment at 3777 Mentone Ave with me. He was a wonderful friend and roommate who always taught me new things.

Engling was my mentor at ST Microelectronics. He educated me on VLSI design which has been part of my career ever since. Stephan was my boss at Realtek. He had an incredible intuition in channel coding. My accomplishment at Realtek was mainly due to his support.

I try to make this list as complete as possible. I apologize for omitting any of you who deserve to be here.

Most of all, I thank my parents and my brother Feng. You are the source of constant love and support. You always have my love. To my parents, I dedicate this dissertation.

VITA

1975	Born, Shaoxing, Zhejiang, China
1992–1997	B.S., Electrical Engineering Tsinghua University Beijing, China
1997–2000	M.S., Electrical Engineering Tsinghua University Beijing, China
2001–2005	Ph.D., Electrical Engineering University of California, Los Angeles Los Angeles, California.

PUBLICATIONS

A. Ramamoorthy, J. Shi, and R. D. Wesel, "On the Capacity of Network Coding for Random Networks," IEEE Trans. on Info. Th. Vol. 51, No. 8, Aug. 2005, pp. 2878-2885.

J. Shi and R. D. Wesel, "Channel-Eigenvector Invariant Space-Time Constellations," IEEE Global Telecommunications Conference, 2004.

J. Shi and R. D. Wesel, "Rotationally Invariant Space-Time Constellations," IEEE International Symposium on Information Theory, 2004.

J. Shi and R. D. Wesel, "Effective Computation of Trellis Code Generating Functions," IEEE Trans. on Comm. Vol. 52, No. 2, Feb. 2004, pp. 219-227.

A. Ramamoorthy, J. Shi, and R. D. Wesel, "On the Capacity of Network Coding for Random Networks," the 41st Allerton Conference on Communication,Control and Computing, October 2003.

T. Ho, M. Medard, J. Shi, M. Effros, and D. Karger, "On Randomized Network Coding," the 41st Allerton Conference on Communication,Control and Computing, October 2003.

J. Shi and R. D. Wesel, "Further Error Event Diagram Reduction Using Algorithmic Techniques," International Conference on Communications 2003, Anchorage, Alaska.

A. Matache, R. D. Wesel, and J. Shi, "Trellis Coding for Diagonally Layered Space-Time Systems," International Conference on Communications 2002, New York.

Abstract of the Dissertation

Universal Channel Codes and Trellis State-Diagram Reduction

by

Jun Shi

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2005

Professor Richard D. Wesel, Chair

The compound channel coding theorem was originally studied in the 1960's. It states that for a compound channel consisting of a family of channels, there exists a single code that achieves the capacity of the compound channel with probability of error converging to zero as the block-length goes to infinity. This theorem lays the theoretical foundation of universal codes that perform uniformly well under all the channels in the family. During the design phase, one can take all the possible scenarios into consideration and form a compound channel. Codes designed for the compound channel possess robustness.

The compound-channel coding theorem is asymptotic in the block-length. It does not address the practical situation when the block-lengths of the codes are finite. This dissertation evaluates the performance of universal codes as the block-lengths vary. The sphere packing bound and random coding bound are used in the evaluation. It is found that the performance can vary significantly when the block-length is short. However, universal behavior is a practical goal when the block-length is as long as 5000. Furthermore low density parity check (LDPC) codes fit naturally as candidates for universal codes.

The concept of universal codes is further extended to multiple-antenna systems. We limit ourselves to space-time constellations. A class of constellations that possess the channel-eigenvector invariance (CEI) property is studied. The necessary and sufficient condition for these constellations is found. It is shown that they are extensions of the orthogonal designs. These constellations, however, can not be a building block of a universal system due to the inherent mutual information loss. Constellation and coding have to be designed simultaneously to achieve universality.

A state-diagram reduction algorithm for trellis coded modulation is proposed in the dissertation. The algorithm is based on the finite-state-machine (FSM) theory. We show that an arbitrary trellis code needs at most half of the all possible product-states to describe the error events. Our algorithm is able to reduce the number of states further by applying the FSM minimization algorithm in the forward and backward directions. The reduction is achieved with low complexity.

CHAPTER 1

Introduction

1.1 Universal Codes

More than half a century has passed since Shannon published his seminal paper *The Mathematical Theory of Communication*. Numerous channel coding schemes have been proposed and implemented in communication systems. The advances in channel coding have impact on many areas including wireless communications, sensor networks, and magnetic storage. It has been pointed out by McEliece at the ISIT 2004 Shannon Lecture that channel coding accounts for 37% of the orders of magnitude improvement in image bit rate for deep space communications [1].

A simple communication system is illustrated in Fig. 1.1. The input symbol is x, which belongs to the alphabet $\mathcal{X}$. The channel is basically a probabilistic mapper which maps input x to output $y \in \mathcal{Y}$ with some conditional probability determined by the characteristics of the channel. The capacity of this channel is the maximum mutual information between x and y over all the possible input distributions, i.e. $C = \max_{x}\{I(x;y)\}$. Shannon proved that any data rate less than C is achievable with probability of error converging to zero as the blocklength grows to infinity. The goal of channel coding is to design input sequences that are able to achieve the capacity of the channel. Coding theorists have taken distinct approaches to meet the goal, including block codes, convolutional codes, trellis coded modulation, and more recently, iterative decoding based turbo codes and LDPC codes. A prominent achievement is the invention and application of iterative decoding, which approaches the Shannon capacity of additive white Gaussian noise (AWGN) channels within 0.0045dB with a block-length of 10^7 [2].

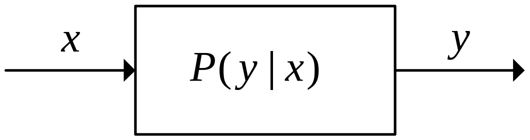


Figure 1.1: A simple channel

The channel characteristics are usually assumed to be known *a priori* when designing error-correcting codes. For example, AWGN channels or binary symmetric channels (BSCs) are often the targeted channels. In practice due to channel uncertainty, it is desirable to have robust codes that perform well not only on a specific channel, but universally well across a family of channels. These codes are called universal codes and are the main focus of this dissertation.

The theorem of universal codes is also known as the compound channel coding theorem in the literature [3]. A compound channel is illustrated in Fig. 1.2. The channel is described by the condition probability $P(y|x,s)$ indexed by s. We assume that the transmitter knows the channel belongs to a family of channels but does not know exactly what the channel is. We also assume the coherence time of the channel is larger than the codeword length so that the compound channel assumption is valid. The universal coding theorem states that the capacity of a compound channel can be achieved by a single code. This theorem is the theoretical foundation of universal codes.

The prominent advantage of the universal codes is robustness. If the designer of the system is able to predict the variation of the channel, he can form a com-

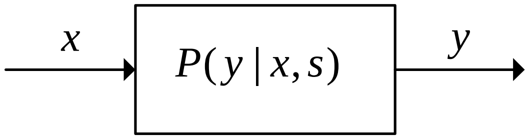


Figure 1.2: A compound channel

pound channel consisting all the possible variations. A universal code designed for this compound channel will work for all individual ones. Universal codes lend themselves to easy implementation. Systems equipped with universal codes are relatively simpler than those with adaptive schemes. This is extremely important for those applications where power consumption and silicon area are major concerns.

The theorem of universal codes was first studied by Blackwell et. al. [4], Dobrushin [5] and Wolfowitz [6] in the 1960's. The results were extended to Gaussian channels by Root and Variaya [7]. The design of universal codes nevertheless is not a simple task. Universal code designs based on the maximin criterion were proposed and studied for periodic channels and multiple-antenna systems by Wesel et. al. [8, 9]. Constellation designs for both single antenna and multiple antenna systems were studied by Tse and Viswanath [10], based on the optimality of diversity and multiplexing gain. In this dissertation, we will study the finite block-length behavior of universal codes and explore the applications to space-time constellations.

1.2 Trellis State-Diagram Reduction

Combining coding and modulation, trellis coded modulation (TCM) is an efficient method to increase the spectral efficiency and power efficiency. It was first proposed by Ungerboeck [11] and found immediate applications in high-data-rate communications over band-limited channels. For example, TCM is used in International Telecommunication Union (ITU) V.32, V.32bis and V.34 modem systems.

There are at least two distinct methods to design TCMs. The first method is set-partitioning, proposed by Ungerboeck in [11]. By partitioning the constellation points into different subsets that have large intra-subset distances, this method guarantees large Euclidean distance between codewords despite the fact that some bits are not explicitly coded. The second method is a full computer search, in which all bits are explicitly coded. The performance of a TCM can be obtained by computer simulation or evaluated by error bounds such as the union bound. Due to the nonlinearity of the constellation mapper, a product-state diagram is needed for performance evaluation. As a result, the number of states increases from N to N^2, where N is the number of states of the underlying convolutional code. Complexity reduction of the state diagram for TCMs has been extensively studied in the literature. Most studies were focused on particular structures of the codes that guarantee a simple state diagram. In this dissertation, we will follow an FSM-based approach originally proposed in [12] and present an iterative reduction method that can be applied to generic TCMs.

1.3 Outline

The rest of the dissertation is organized as follows. Chapter 2 introduces the concept of universal codes and studies the finite block-length properties of universal codes. Chapter 3 applies the universal codes to MIMO systems and investigates the channel-eigenvector-invariant space-time constellations. Chapter 4 presents an FSM-based trellis code state diagram reduction algorithm. Examples are used to illustrate the efficiency of the algorithm. Chapter 5 concludes the dissertation and outlines future research directions.

CHAPTER 2

Universal Codes with Finite Block-Lengths

2.1 Background

Traditional code design is targeted at a specific channel. The design relies on two key assumptions: first the channel can be identified accurately, second the characteristics of the channel do not change dramatically from those for which the code is designed. In reality, there exists a variety of situations for which these assumptions can not be satisfied. When this happens, the performance of such designed optimal codes might deteriorate significantly. For example, an optimal AWGN code might not perform well under periodic erasure channels [8] or partial band jamming channels [13]. A space-time code optimal for Rayleigh fading channels may not work well under some special scenarios, including singular channels that result from a single line-of-sight path, or the “keyhole” channel [14].

One approach to solve this problem is to design individual optimal codes for each channel condition. However this scheme requires storage for all the possible codes at the transmitter and the receiver and the ability of both sides to intelligently identify and adapt to the environment. An alternative approach is to design a code that works reasonably well under most, if not all, possible scenarios. In this chapter, we study both the theoretic and practical aspects of the latter approach.

2.2 Compound Channel Coding Theorem

A compound channel arises when users communicate under some channel uncertainty [3], i.e. users know the channel belongs to a family of channels but they do not know exactly what the channel is. Throughout this chapter, we will restrict our discussion to discrete memoryless channels (DMCs).

Definition 1 *A compound channel is a family of channels indexed by $i \in \mathcal{I}$ denoted by*

$$\{P(y|x,i), x \in \mathcal{X}, y \in \mathcal{Y}, i \in \mathcal{I}\}, \tag{2.1}$$

where $\mathcal{X}$ and $\mathcal{Y}$ are the input and output alphabet, respectively. $\mathcal{I}$ is the index set which can be finite, countably infinite or uncountably infinite. $P(y|x,i)$ is the conditional probability governing the channel with index i.

It is assumed that the channel index remains unchanged during the course of the transmission. Or at least the time that the channel index stays the same is longer than the codeword block-length. If the index varies arbitrarily from symbol to symbol, then such a channel is referred as an arbitrarily varying channel [3], which is not the focus of this chapter.

The capacity of a compound channel is defined as

$$C(\mathcal{I}) = \sup_{Q(x)} \inf_{i \in \mathcal{I}} I_i(x;y), \tag{2.2}$$

where $I(x;y)$ is the mutual information between the input and output random variables. Define the infimum of the capacities of the individual channels as

$$C_{\text{inf}}(\mathcal{I}) = \inf_{i \in \mathcal{I}} \sup_{Q(x)} I_i(x;y). \tag{2.3}$$

Note that $C(\mathcal{I}) \leq C_{\text{inf}}(\mathcal{I})$. However it can be shown $C(\mathcal{I}) = 0$ if and only if $C_{\text{inf}}(\mathcal{I}) = 0$ [6]. So any set of positive-capacity channels will have a positive compound channel capacity.

Blackwell, Breiman and Thomasian [4] proved that the capacity of a compound channel with discrete alphabet can be achieved by a single sequence of codes. A similar result also appeared in [6]. This result was extended by Root and Varaiya [7] to $p \times p$ (square) linear Gaussian compound channels where the alphabet is continuous. The slight generalization of their theorem to the $p \times q$ (rectangular) MIMO channels is stated below.

Theorem 1 *A family of Gaussian multiple-input multiple-output (MIMO) channels is denoted as $\{\boldsymbol{H}_i, \boldsymbol{K}_i, i \in \mathcal{I}\}$, where the index set $\mathcal{I}$ is an arbitrary set. The channel input output is governed by $\boldsymbol{y} = \boldsymbol{H}_i\boldsymbol{x} + \boldsymbol{z}_i$, where $\boldsymbol{H}_i$ is a $p \times q$ matrix and $\boldsymbol{z}_i$, the noise, is a Gaussian random vector of dimension $p \times 1$ with zero mean and the covariance matrix $\boldsymbol{K}_i$. Assume there exist real numbers a, α_1, α_2 such that for each $i \in \mathcal{I}$,*

1. *$\|\boldsymbol{H}_i\| \leq a$, where $\|\cdot\|$ is the spectral norm of a matrix, i.e. the square root of the largest eigenvalue of $\boldsymbol{H}_i^*\boldsymbol{H}_i$.*
2. *$\alpha_1 \leq \frac{\boldsymbol{x}^*\boldsymbol{K}_i\boldsymbol{x}}{\|\boldsymbol{x}_i\|^2} \leq \alpha_2$, for any $p \times 1$ vector $\boldsymbol{x}$.*

Then any rate $R < C(\mathcal{I})$ defined in (2.2) is achievable. i.e. there exists a sequence of $(2^{nR}, n)$ codes such that the probability of error under any channel in the family approaches zero as the block-length approaches infinity.

Proof: The theorem for square channel matrices first appeared in [7]. The original proof is however mathematically involved. In the Appendix, we give an alternative proof based on random codes and typical set decoding. ■

The surprising fact about the compound channel coding theorem is that the probability of error goes to zero uniformly as long as the code rate is less than the compound channel capacity no matter what channel the sequence of codes is actually encountering. The uniform convergence does not mean that the error probability of each channel in the compound channel goes to zero at exactly the same speed, but the speed is at least lower bounded. The difference is negligible at large block-length, but significant for codes with relatively short block-length. This short-block-length behavior is the focus of this chapter.

2.3 Codes with Finite Block-Lengths

Before we analyze the performance of an error-correcting code under various channels, a fair and convenient figure of merit is needed. This figure should automatically take channel conditions into consideration and act consistently across all channels.

2.3.1 Figure of Merit

The code performance in an AWGN channel is usually gauged by the signal-to-noise-ratio (SNR) required to achieve a certain target bit error rate (BER) or frame error rate (FER). SNR can be used to evaluate different codes under the same channel condition and the same rate. However it is not suitable for codes with distinct rates or under various channel conditions. A few metrics have been proposed in the literature. For example, Forney proposes the normalized SNR in the high-SNR regime [15], which is defined as

$$\mathrm{SNR}_{\mathrm{norm}} = \frac{\mathrm{SNR}}{2^{2R}-1}, \tag{2.4}$$

where R is the code rate. $10\log_{10}(\mathrm{SNR}_{\mathrm{norm}})$ is often called the SNR gap of a code. The value of the SNR gap indicates how far a system is operating from the Shannon limit. In the low-SNR regime and the wide-band regime, Verdú suggests that the figure of merit is not SNR, but rather the normalized energy per information bit, E_b/N_0 [16].

We propose excess mutual information (EMI) as the figure of merit for the purpose of universal code design. It is defined as:

$$\mathrm{EMI}(\mathrm{SNR}, R) = I(\mathrm{SNR}) - R, \tag{2.5}$$

where R is the rate of the code and SNR is the signal-to-noise-ratio at which the code achieves a certain target probability of error, $I(\mathrm{SNR})$ is the mutual information of the channel at that SNR. For the AWGN channel with Gaussian input, $I(\mathrm{SNR})$ coincides with the capacity of the channel. EMI indicates how much penalty is paid due to the imperfectness of the code.

For an AWGN channel, EMI and SNR gap only differ by some constant factor in the high-SNR and high-rate regime. Assuming a real AWGN channel with Gaussian input, SNR gap can be written as

$$\begin{aligned}\text{SNR gap} &= 10\log_{10}(\mathrm{SNR}) - 10\log_{10}(2^{2R}-1)\\ &\approx 10\log_{10}(\mathrm{SNR}) - 20R\log_{10}(2).\end{aligned} \tag{2.6}$$

EMI can be written as

$$\begin{aligned}\mathrm{EMI} &= \frac{1}{2}\log_2(1+\mathrm{SNR}) - R\\ &\approx \frac{1}{2}\log_2(\mathrm{SNR}) - R.\end{aligned} \tag{2.7}$$

The approximations in (2.6) and (2.7) are valid when R and SNR are large respectively. Under these assumptions, the EMI and the SNR gap differ by a factor of $20\log_{10} 2$. In the low SNR and low rate regime, this linear relationship is no longer valid. But there still exists a monotonic bijection between them.

It is interesting to note that EMI appears in the probability of error under typical set decoding. Recall that in the proof of the capacity of a Gaussian channel, the error probability of a Gaussian code book of block-length n is upper bounded by [17, p. 245]

$$\begin{aligned} P_e^{(n)} &\leq 2\epsilon + 2^{3n\epsilon}2^{-n(I(X;Y)-R)} \\ &= 2\epsilon + 2^{3n\epsilon}2^{-n\mathrm{EMI}}. \end{aligned} \tag{2.8}$$

The EMI is not the error exponent which is defined as

$$E_t = \frac{\ln(P_e^{(n)})}{n}. \tag{2.9}$$

Even for typical set decoding, the ϵ in (2.8) hides too much information about the error exponent. However, EMI is indeed related to the probability of error.

2.3.2 Design of Universal Codes

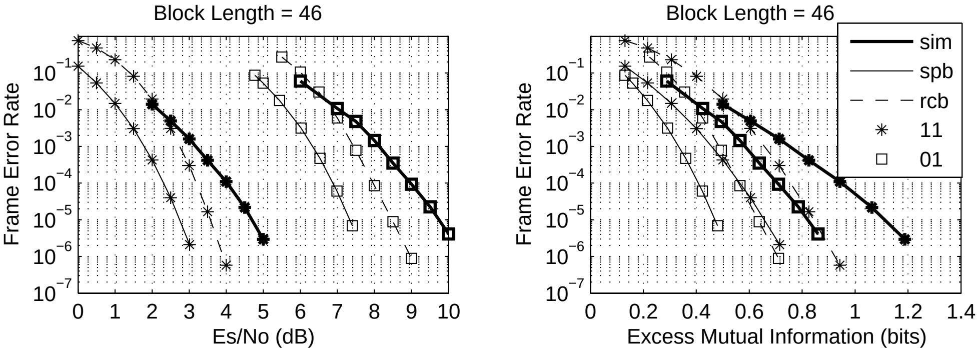
Universal code design is inherently a multicriterion optimization problem. One approach could be to design a code which optimizes the sum of some metric over all the channels in the family. Another possible criterion is to optimize the worst case performance. We will show examples to illustrate these criteria.

2.3.2.1 Trellis codes for Periodic Erasure Channels

The performance of a trellis code can be upper bounded by the transfer function bound. At high SNR, the bound can be approximated by its first term as listed in (2.10).

$$N_b Q\left(\sqrt{\frac{\mathrm{SNR} d_{\min}}{2}}\right), \tag{2.10}$$

where N_b is the number of bits in error associated with events having the minimum Euclidean distance $d_{\min}$. A search can be conducted to find the code that minimizes $\sum_i \mathrm{SNR}_i$ or $\sum_i \mathrm{EMI}_i$ at a certain target probability of error, where the sum is over all the erasure patterns in the compound channel family. The trellis code whose performance is shown in Fig. 2.1 is one of those codes optimized for EMI under the period-2 erasure [8]. It exhibits consistently good performance under both channels. A detailed explanation of this figure will be rendered later in this chapter.



(a) FER vs SNR

(b) FER vs EMI

Figure 2.1: Sphere packing bound (spb), random coding bound (rcb) and computer simulation results (sim) for AWGN channel with block-length 46. 11–non-erasure channel, 01–erasure channel.

2.3.2.2 Universal Space Time Trellis Codes

Root and Varaiya's compound channel coding theorem was proven for MIMO channels although it is only recently that MIMO channels have drawn considerable attention. So a natural application of the compound channel coding theorem is space-time code design.

However, caution has to be taken before applying the theorem directly. A MIMO channel is usually described as $\boldsymbol{y} = \boldsymbol{Hx} + \boldsymbol{z}$, where $\boldsymbol{H}$ obeys Rayleigh or Rician fading law. It can be modelled as a compound channel if the fading process is slow relative to the code block-length. The so-called quasi-static fading or block fading model fits in this category. Under this model, the outage capacity is a more relevant quantity than the ergodic capacity. Furthermore, the two conditions in the theorem on the norm of $\boldsymbol{H}$ and the norm of the covariance matrix of $\boldsymbol{z}$ need to be satisfied. It is commonly assumed that the noise is i.i.d. and has unit variance. So the condition on the noise covariance matrix is satisfied. The condition on the channel matrix is not always satisfied because no matter whether $\boldsymbol{H}$ follows Rayleigh or Rician fading, its norm can be arbitrarily large with some positive probability. Nevertheless we need only to consider the channels whose norm is bounded from above number because the propagation environments are passive.

Having justified the applicability, we are in the position to formulate the space time code design problem in the context of the compound channel theorem. First we need to identify the family of channels. Any channel whose mutual information is less than data rate R should not be included because according to Shannon's coding theorem, reliable communication is not possible through this channel. The critical set of channels are those whose mutual information is exactly equal to R. So an interesting compound channel is defined as all the $\boldsymbol{H}$'s that satisfy:

$$\log\det\left(\boldsymbol{I}+\boldsymbol{H}\boldsymbol{H}^{\dagger}\right)=R \tag{2.11}$$

The compound channel capacity of this set is equal to the minimum mutual information of all the individual channels with a uniform input. For an error event $\boldsymbol{X}\rightarrow\hat{\boldsymbol{X}}$ and a particular channel $\boldsymbol{H}$, the associated Euclidean distance is

$$d(\boldsymbol{X}\rightarrow\hat{\boldsymbol{X}})=\mathrm{Tr}\left(\Delta\boldsymbol{X}^{\dagger}\boldsymbol{H}^{\dagger}\boldsymbol{H}\Delta\boldsymbol{X}\right), \tag{2.12}$$

where $\Delta\boldsymbol{X}=\boldsymbol{X}-\hat{\boldsymbol{X}}$, and $\mathrm{Tr}(\cdot)$ is the trace of a matrix. Using (2.10), we can compute the SNR required to achieve a certain probability of error. Since there are infinitely many channels satisfying (2.11), proper sampling is needed. The objective is to search for a code that minimizes the maximum EMI over all the sample channels. Such designed trellis codes for 2 × 2 systems were shown in [9]. However this approach becomes difficult as the system becomes large.

2.4 Periodic Erasure Channels

A simple way to generate a family of channels is to erase transmitted symbols periodically with different patterns. Each erasure pattern generates a channel in the family. We will focus on periodic erasure channels due to its simplicity.

Consider a binary symmetric periodic erasure channel with input symbols $x_i \in \{-1, 1\}$. $f(x_i, p)$ is the channel function, i.e. with probability of $1 - p$, $f(x_i, p) = x_i$, while with probability of p, $f(x_i, p) = -x_i$. $a_i \in \{0, 1\}$ are the erasing coefficients with period T, i.e. $a_i = a_{i+T}$. Similarly an AWGN periodic erasure channel can be formulated as $y_i = a_i x_i + n_i$, when n_i is the Gaussian noise.

The periodic erasure channel can be regarded as a simplified model for frequency-hopped systems where partial band interference arises due to frequency dependent disturbance [13]. It can also model communications in the presence of jamming.

2.5 Error Bounds and Error Exponents

A finite block-length code is often compared to the Shannon capacity to measure its imperfectness. However this comparison is not completely fair because in most cases, the Shannon capacity can only be achieved as the block-length goes to infinity. Among available finite block-length analysis tools are the sphere packing bound (SPB) [18] and the random coding bound (RCB) [19]. The probability of error for codes with finite block-length is lower bounded by the SPB.

The RCB, characterizing the average performance of a randomly selected codes, serves as an upper bound on the probability of error for an optimal code. However in reality it might be the case that even the RCB can not be achieved by a carefully designed code due to the increasing decoding complexity. A fair assessment of a finite block-length code can be made by measuring its EMI against the EMI of the SPB or the RCB. We state the SPB and the RCB for the BSC and the AWGN channel, respectively, then extend them to periodic erasure channels.

2.5.1 Sphere Packing Bound (BSC)

The derivation of the SPB for the BSC is combinatorial. For an $(n,\, k)$ binary code, the bound can be written as [19, 20]

$$\begin{aligned} P_w(n,k,p) &\geq p^{r+1}(1-p)^{n-r-1}\left(\sum_{i=0}^{r+1}\binom{n}{i} - 2^{n-k}\right) \\ &\quad + \sum_{i=r+2}^{n}\binom{n}{i}p^i(1-p)^{n-i} \\ &= \sum_{i=r+1}^{n}\binom{n}{i}p^i(1-p)^{n-i} \\ &\quad - p^{r+1}(1-p)^{n-r-1}\left(2^{n-k} - \sum_{i=0}^{r}\binom{n}{i}\right), \end{aligned} \tag{2.13}$$

where p is the crossover probability of the BSC and r is the maximum integer such that $\sum_{i=0}^{r}\binom{n}{i} \leq 2^{n-k}$.

2.5.2 Sphere Packing Bound (AWGN)

The derivation of Shannon's SPB for the AWGN channel is essentially geometric. The codewords of block-length n are regarded as points on an $(n-1)$-dimensional sphere with radius $\sqrt{nE_s}$. The error probability is lower bounded by the probability that an n-dimensional Gaussian random variable falls out a cone which corresponds to the Vonoroi region of the transmitted codeword. The error probability is given as [18, 21]:

$$P_w \geq Q_n(\theta_s, A), \tag{2.14}$$

where $A = \sqrt{E_s/N_o}$, θ_s is the half angle of a cone whose solid angle is a fraction $1/2^k$ of the total solid angle of the sphere. k is the information bit length. θ_s satisfies

$$\Omega_n(\theta_s) = \int_0^{\theta_s} \frac{(n-1)\Gamma(\frac{n}{2}+1)(\sin\phi)^{(n-2)}}{n\Gamma(\frac{n+1}{2})\sqrt{\pi}} d\phi = \frac{1}{2^k}. \tag{2.15}$$

$Q_n(\theta_s, A)$ is given as

$$\begin{aligned} Q_n(\theta_s, A) = &\int_{\theta_s}^{\pi} \frac{(n-1)(\sin\phi)^{(n-2)}}{2^{n/2}\sqrt{\pi}\Gamma(\frac{n+1}{2})} \\ &\int_0^{\infty} s^{(n-1)} e^{-(s^2+nA^2-2s\sqrt{n}A\cos\phi)/2} ds d\phi. \end{aligned} \tag{2.16}$$

The computation in (2.15) and (2.16) becomes numerically unstable when n becomes large. The following asymptotic approximations should be used for large n:

$$\Omega_n(\theta_s) \approx \frac{\Gamma(\frac{n}{2}+1)(\sin\theta_s)^{(n-1)}}{n\Gamma(\frac{n+1}{2})\sqrt(\pi)\cos\theta_s} \approx \frac{(\sin\theta_s)^{(n-1)}}{\sqrt{2\pi n}\cos\theta_s}, \tag{2.17}$$

$$\begin{aligned} Q_n(\theta_s, A) \approx &\frac{1}{\sqrt{n\pi}\sqrt{1+G^2(\theta_s, A)}\sin\theta_s} \times \\ &\frac{[G(\theta_s, A)\sin\theta_s e^{-(A^2-AG(\theta_s,A)\cos\theta_s)/2}]^n}{AG(\theta_s, A)\sin^2\theta_s - \cos\theta_s}, \end{aligned} \tag{2.18}$$

where $G(\theta_s, A) = (1/2)[A\cos\theta_s + \sqrt{A^2\cos^2\theta_s + 4}]$.

2.5.3 Random Coding Bound and Error Exponent (BSC)

The RCB for the BSC is computed as [19, p. 146]

$$P_w^n \le e^{-nE_r(R)}, \tag{2.19}$$

where $E_r(R)$ is the random coding exponent. However, even for such a simple channel, $E_r(R)$ does not have a simple explicit form. We need first to compute an intermediate parameter δ whose relation with the code rate R (in nats) is given as:

$$R = \ln 2 - H(\delta), \tag{2.20}$$

where $H(\delta)$ is the binary entropy function (in nats).

For δ in the range

$$p \leq \delta \leq \frac{\sqrt{p}}{\sqrt{p} + \sqrt{1-p}}, \tag{2.21}$$

where p is the crossover probability, the random coding exponent of the BSC is

$$E_r(R) = -\delta \ln p - (1-\delta)\ln(1-p) - H(\delta). \tag{2.22}$$

For

$$R < \ln 2 - H(\frac{\sqrt{p}}{\sqrt{p} + \sqrt{1-p}}), \tag{2.23}$$

the exponent becomes

$$E_r(R) = \ln 2 - 2\ln(\sqrt{p} + \sqrt{1-p}) - R. \tag{2.24}$$

2.5.4 Random Coding Bound and Error Exponent (AWGN)

Assume Gaussian input, the RCB for the AWGN channel has an explicit form [19, p. 340]. As in the BSC case, we need to compute the error exponent. For the rate R (in nat) in the range:

$$\frac{1}{2}\ln\left(\frac{1}{2} + \frac{A}{4} + \frac{1}{2}\sqrt{1 + \frac{A^2}{4}}\right) \leq R \leq \frac{1}{2}\ln(1+A), \tag{2.25}$$

where $A = E_s/N_o$. The error exponent is

$$\begin{aligned} E_r(R) = {} & \frac{A}{4\beta}\left[\beta + 1 - (\beta - 1)\sqrt{1 + \frac{4\beta}{A(\beta-1)}}\right] \\ & + \frac{1}{2}\ln\left[\beta - \frac{A(\beta-1)}{2}\left(\sqrt{1 + \frac{4\beta}{A(\beta-1)}} - 1\right)\right], \end{aligned} \tag{2.26}$$

where $\beta = e^{2R}$. When R is less than the left hand side of (2.25), the error exponent becomes

$$E_r(R) = 1 - \beta + \frac{A}{2} + \frac{1}{2}\ln\left(\beta - \frac{A}{2}\right) + \frac{1}{2}\ln\beta - R, \tag{2.27}$$

where

$$\beta = \frac{1}{2}\left(1 + \frac{A}{2} + \sqrt{1 + \frac{A^2}{4}}\right). \tag{2.28}$$

2.5.5 Extension to Periodic Erasure Channels

As far as the effect of block-length is concerned, sending a codeword through a periodic channel is equivalent to puncturing the codewords first then sending through a standard channel. The previously listed bounds and error exponents can be extended to erasure channels by this equivalence. For example, the random coding bound of a rate-1/4 length-N code with erasure pattern "10" is the same as that of a rate-1/2 length-$N/2$ code without erasures.

2.6 Numerical Analysis and Simulation Results

2.6.1 Random Coding Error Exponents

The random coding error exponent indicates the rate at which the error probability of the ensemble codes approaches zero as block-length grows. We consider a BSC and an AWGN compound channel with erasure patterns "1111" and "0111". For a fair comparison, the BSC's have a crossover probability 0.11 for the first erasure pattern and 0.0615 for the second pattern such that both channels have capacity 0.5 bit. In the AWGN case, we set the SNR to be 3 for the first pattern and 5.3496 for the second such that both channels have capacity 1.0 bit. The random coding error exponents are plotted in Fig. 2.2 for various code rates. It is clear that the error exponents are distinct in both the AWGN and the BSC compound channels.

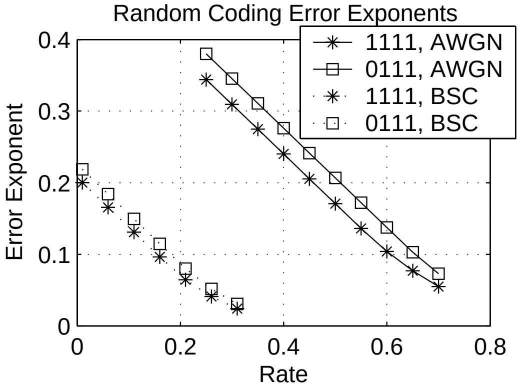


Figure 2.2: Random coding error exponents for a BSC with capacity 0.5 bit and an AWGN periodic erasure channel with capacity 1.0 bit. 1111–non-erasure channel, 0111–erasure channel.

2.6.2 Periodic Erasure BSC

Suppose CH1 is the standard BSC with crossover probability p and CH2 is a BSC with erasure pattern "0111". This example is different from above in that both channels have the same p. The capacities of two channels are:

$$C_1(p) = 1 - H(p), \qquad C_2(p) = \frac{3}{4}(1 - H(p)), \tag{2.29}$$

where $H(p) = -(1-p)\log_2(1-p) - p\log_2(p)$ is the binary entropy function. In both cases, the code rate is 1/4. According to (2.5), the EMI for CH1 is calculated by $\text{EMI}_1 = C_1(p) - 0.25$ and the EMI for CH2 is $\text{EMI}_2 = C_2(p) - 0.25$. Bounds on the frame error rate for block-length 52 are plotted against EMI in Fig. 2.3.

It is observed that the erasure channel is a more efficient channel in the sense

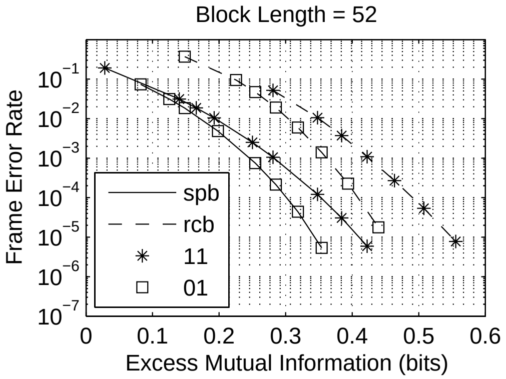


Figure 2.3: Sphere packing bound (SPB) and random coding bound (RCB) for BSC with block-length 52. 11–non-erasure channel, 01–erasure channel.

that it uses less EMI according to the SPB and the RCB. So one would expect a short block-length universal code to have better performance in an erasure channel than in a standard channel in terms of EMI.

2.6.3 Periodic Erasure AWGN Channel

A more interesting case is AWGN channel. In what follows we compare the SPB and the RCB performance to that of three different codes, a trellis code, an LDPC code and a turbo code.

Example 1 *A trellis code*

The rate 1/3 trellis code [171 46 133] (in octal) prosposed in [8] was simulated. Gray labeled 8PSK constellation was employed. The block-length was 46 symbols. Again, we denote the standard AWGN channel by CH1 and the one with erasure pattern "01" by CH2. The EMI for CH1 is (per complex symbol)

$$\mathrm{EMI}_1 = \log(1 + \mathrm{SNR}) - R. \tag{2.30}$$

And the EMI for CH2 is

$$\mathrm{EMI}_2 = 1/2 \log(1 + \mathrm{SNR}) - R. \tag{2.31}$$

The bounds on frame error rate are plotted against SNR in Fig. 2.1(a) and against EMI in Fig.2.1(b), together with the simulation results. If only looking at the SNR plot and ignoring the SPB and RCB, one might assume that the code performance is much worse in the erasure channel because it needs larger SNR. However this is not correct. The erasure channel inherently requires a larger SNR than the AWGN channel at the same capacity. Like the BSC case, the erasure channel requires less EMI according to the bounds. The simulation results agree with the bounds. This phenomenon can be justified by the following computation.

Suppose a code operates at SNR_1 and SNR_2 in CH1 and CH2 (as defined above) respectively in order to achieve a certain target FER. If the following is satisfied,

$$\mathrm{SNR}_1 \leq \mathrm{SNR}_2 \leq \mathrm{SNR}_1^2 + 2\mathrm{SNR}_1, \tag{2.32}$$

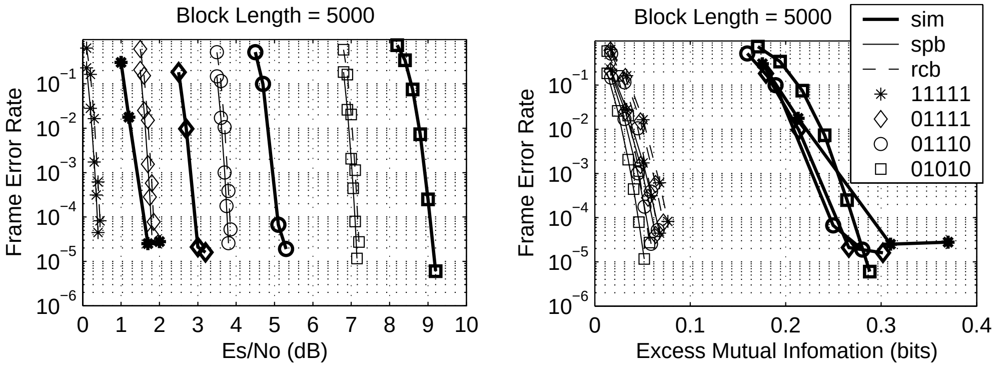
then according to (2.30) and (2.31), we have $\mathrm{EMI}_1 \leq \mathrm{EMI}_2$. The SNR's for this trellis code happened to fall in the region defined in (2.32).

Example 2 *An LDPC code*

Although Fig. 2.1 shows that the bounds for the two channels differ considerably at short block-length, meaning that constant EMI is not possible across both channels, this gap becomes much smaller at longer block-lengths. As shown in Fig. 2.4, the difference among the bounds becomes negligible as the block-length becomes large. Also shown in the Fig. 2.4 are simulation results of a rate 1/4 block-length 20000 binary LDPC code mapped to 5000 16QAM symbols. This code was optimized for AWGN channel. Its parity check matrix was generated according to the degree distribution in (2.33). The graph-conditioning method in [22], [23] and [24] was used to lower the error floor. Systematic design of LDPC codes for periodic erasure channel through density evolution was presented in [25].

Four different channels were considered, including the standard AWGN channel and three erasure channels. This code performs uniformly well under all four channels in terms of EMI.

$$\begin{aligned}\lambda(x) &= 0.356x + 0.219x^2 + 0.175x^5 \\ &\quad + 0.057x^6 + 0.1x^{15} + 0.0926x^{16} \\ \rho(x) &= 0.5x^3 + 0.5x^4.\end{aligned} \tag{2.33}$$



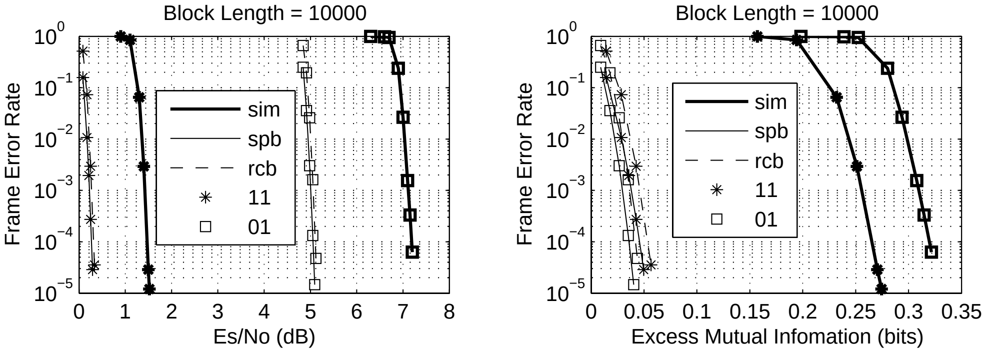
(a) FER vs SNR

(b) FER vs EMI

Figure 2.4: Sphere packing bound (spb), random coding bound (rcb) and computer simulation results (sim) for AWGN channel with block-length 5000. 11111–non-erasure channel, 01110, 01110, 01010–three erasure channels.

Example 3 *A turbo code*

The rate 1/3 turbo code SC-5 proposed in [26] is shown in Fig. 2.5. The block-length was 10000 8PSK symbols. The bounds still suggest that the erasure channel uses less mutual information, but the simulation result shows the opposite. This is because the 8PSK symbol in the erasure channel carries 2 bits of information, close to the saturation regime of the 8PSK-constrained capacity. The same phenomenon is observed in the previous LDPC example, where the code performs worse in terms of EMI in the most erased channel.



(a) FER vs SNR

(b) FER vs EMI

Figure 2.5: Sphere packing bound (spb), random coding bound (rcb) and computer simulation results (sim) for AWGN channel with block-length 10000. 11–non-erasure channel, 10–erasure channel.

2.7 Discussion and Summary

EMI is a useful metric to fairly compare a single code operating at a single rate. If codes of different rates are to be compared, a better metric is a relative quantity, e.g. R/C, $(C - R)/C$ or $(C - R)/R$, where C is the capacity of the channel. For example, consider an AWGN erasure channel which erases all the odd symbols, the code rate is R and channel capacity is $1/4 \log_2(1 + \text{SNR})$.

Thus EMI is $1/4 \log_2(1 + \mathrm{SNR}) - R$. Equivalently, we can puncture all the odd symbols then transmit it through a standard AWGN channel. In this case, the code rate is $2R$ and channel capacity is $1/2 \log_2(1 + \mathrm{SNR})$. Thus EMI becomes $1/2 \log_2(1 + \mathrm{SNR}) - 2R$. In this example, EMI changed due to our different perspectives. The relative quantities listed above behave consistently.

The compound channel coding theorem is reviewed in this chapter. A new proof of the theorem is presented. Like most coding theorems, only the asymptotic behavior of codes for a compound channel is stated in the theorem. We investigate the performance limits of universal codes with finite block-lengths by using the random coding bound and the sphere packing bound. It is shown that although probability of error approaches zero uniformly under all the channels in the family, the difference could be significant at finite block-lengths. Once the input block-length is on the order of 5000 bits, the bounds indicate that uniform behavior over all channels is a practical goal. Furthermore, LDPC codes seem well-suited to provide this behavior. It is interesting to notice that in all three examples, the codes perform approximately 0.2 bit EMI away from the SPC. So the short block-length trellis code in the example is as universal as the long block-length codes.

CHAPTER 3

Channel Eigenvector Invariance Space Time Constellations

3.1 Introduction

Recent years have seen intensive research on multiple antenna systems. Multiple-input-multiple-output (MIMO) systems make high-data-throughput, low-error-rate communications possible. A commonly used model for MIMO systems is described as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}, \tag{3.1}$$

where $\boldsymbol{x} \in \mathbb{C}^{n_t}$ is the transmitted signal from n_t antennas, $\boldsymbol{y} \in \mathbb{C}^{n_r}$ is the received signal at n_r antennas. $\boldsymbol{H}$ is an $n_r \times n_t$ complex matrix characterizing the fading channel. $\boldsymbol{n}$ is a complex vector whose elements are i.i.d. zero-mean, complex Gaussian random variables with variance-0.5 real and imaginary parts. If the channel state infomation is not available at the transmitter, a uniform input distribution is optimal [27, 28]. With an i.i.d. Gaussian input, the mutual information $\mathcal{I}$ between the transmitted and the received signals in (3.1) given a realization of $\boldsymbol{H}$, is shown in (3.2).

$$\begin{aligned}\mathcal{I}(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{H}) &= \log\det\left(\boldsymbol{I} + \frac{P}{n_t}\boldsymbol{H}\boldsymbol{H}^\dagger\right)\\ &= \log\prod_i\left(1 + \frac{P}{n_t}\lambda_i(\boldsymbol{H}\boldsymbol{H}^\dagger)\right),\end{aligned} \tag{3.2}$$

where $\dagger$ is the complex conjugate transpose. P is the average power of the transmitted vector $\boldsymbol{x}$. $\lambda_i(\cdot)$ stands for the ith eigenvalue of a matrix.

The statistics of $\boldsymbol{H}$ vary from model to model. When the fading is slow, the so-called block fading model is usually used, where $\boldsymbol{H}$ stays fixed during one block and then new entries are drawn independently according to the complex Gaussian distribution at the next block. In the block fading model, the transmitted signal $\boldsymbol{x}$ in (3.1) extends to an $n_t \times T$ matrix $\boldsymbol{X}$ and $\boldsymbol{y}$ becomes an $n_r \times T$ matrix $\boldsymbol{Y}$, where T is the block-length.

Various space-time code design criteria, such as the rank criterion and the determinant criterion [29, 30], have been derived to achieve large diversity and coding gains. The underlying idea is to minimize the average pairwise error probability. These designs depend highly on channel distributions. Codes designed for a Rayleigh fading channel do not necessarily work well in certain singular or unitary channels.

Alternatively, codes designed from the compound channel theory enjoy more universality. A block fading channel is an example of a compound channel [3], where the channel state information $\boldsymbol{H}$ is unknown at the transmitter except the knowledge that it belongs to a certain set. Compound channel theory states that one can find a code to achieve rates arbitrarily close to the capacity of the compound channel [7, 6]. In the block-fading scenario, the capacity of the compound channel happens to be the minimum mutual information among all the channels with uniform inputs [9, 28]. So by using a single code, we can achieve arbitrarily small probability of error for all channels with mutual information no less than the rate. This code achieves the outage capacity of the fading channel under every fading distribution as the block-length goes to infinity. Such a code is the best one can expect in a block fading channel.

The existence result in [7, 6] sheds little light on the construction of such universal codes. The code design is inherently a multicriterion optimization problem [8]. Some success with trellis codes [9] and LDPC codes [25] has been achieved. It is observed that the channel matrix $\boldsymbol{H}$ contributes to the mutual information in (3.2) only through its squared singular values (i.e. the eigenvalues of $\boldsymbol{H}\boldsymbol{H}^{\dagger}$). A family of channels with the same set of singular values has identical mutual information. So a universal channel code should perform identically under these channels. A code that achieves the same performance irrespective of the eigenvectors is called channel-eigenvector invariant. In this chapter we investigate the feasibility of constructing channel-eigenvector invariant space-time constellations. Any code equipped with these constellations is automatically channel-eigenvector invariant.

3.2 Preliminaries

Since the design of channel-eigenvector invariant constellations will rely heavily on the Hurwitz-Radon matrix equations, we present a short overview of this topic. The interested readers are referred to [31, 32] for details.

We use T to denote the transpose of a matrix and $\dagger$ the complex conjugate transpose. $\boldsymbol{I}$ and $\boldsymbol{0}$ are the identity and all-zero matrices respectively. Their dimensions are implicitly decided in the context. $\mathbb{R}$ and $\mathbb{C}$ denote the field of the real numbers and the complex numbers respectively. $*$ denotes the complex conjugate. $\mathsf{i} = \sqrt{-1}$ is the pure imaginary unit.

3.2.1 Hurwitz-Radon Matrix Equations

The Hurwitz-Radon matrix equations were independently examined around 1920 by Hurwitz and Radon to study the composition of quadratic forms. These equations have been used in topology, algebra, and analysis beyond their inventors' original purpose [33].

Definition 2 *The following system is called the Hurwitz-Radon matrix equations over field $\mathbb{F}$,*

$$\begin{aligned} \boldsymbol{A}_i \boldsymbol{A}_i^{\intercal} &= \boldsymbol{I} \quad i = 1, 2, \ldots k, \\ \boldsymbol{A}_i \boldsymbol{A}_j^{\intercal} + \boldsymbol{A}_j \boldsymbol{A}_i^{\intercal} &= \boldsymbol{0} \quad 1 \leq i < j \leq k, \end{aligned} \tag{3.3}$$

where $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k \in \mathbb{F}^{n \times p}$. The set of matrices $\{\boldsymbol{A}_i, i = 1, \ldots, k\}$ is called a solution to the matrix equations. k is called the cardinality of the solution. ■

The Hurwitz-Radon matrix equations can be simply denoted by $[k, n, p]$ over the field $\mathbb{F}$ since these are the only quantities that matter. A major concern of the equations is to decide the relationship among k, n and p and to explicitly construct the $\boldsymbol{A}_i$'s. Although the problem when n and p are arbitrary integers still remains open, the case $n = p$ is totally settled for any field $\mathbb{F}$ whose characteristic is not 2. The results are listed below.

Theorem 2 *If $p = n = 2^a(2b+1)$ and 2 is not the characteristic of the field $\mathbb{F}$, then the maximum k such that the set of matrix equations (3.3) has a solution is* *equal to the Hurwitz-Radon function $\rho_1(n)$.*

$$\begin{aligned}\rho_1(n) &= \rho_1(2^a(2b+1)) \\ &= \begin{cases} 2a+1 & \mathit{if}\ a \equiv 0 \quad (\ \mathit{mod}\ 4), \\ 2a & \mathit{if}\ a \equiv 1 \quad (\ \mathit{mod}\ 4), \\ 2a & \mathit{if}\ a \equiv 2 \quad (\ \mathit{mod}\ 4), \\ 2a+2 & \mathit{if}\ a \equiv 3 \quad (\ \mathit{mod}\ 4). \end{cases}\end{aligned} \tag{3.4}$$

■

Although the cardinality of a solution to the Hurwitz-Radon matrix equations in the square case is limited by $\rho_1(n)$, the solution is not unique. A very important one among all the solutions is the integer solution where all the entries of $\boldsymbol{A}_i$ are in the set $\{0, \pm 1\}$. The existence of a solution to the $[\rho_1(n), n, n]$ Hurwitz-Radon matrix equations implies the existence of an integer solution. Hence restriction on the entries to the set $\{0, \pm 1\}$ does not reduce the cardinality of a solution.

3.2.2 Hermitian Hurwitz-Radon Matrix Equations

If the field $\mathbb{F} = \mathbb{C}$, (3.3) has the following variant which is very useful in the space-time code design.

Definition 3 *The following system is called the Hermitian Hurwitz-Radon matrix equations over the complex field $\mathbb{C}$,*

$$\begin{aligned}\boldsymbol{A}_i\boldsymbol{A}_i^\dagger &= \boldsymbol{I} \quad i = 1, 2, \ldots 2k, \\ \boldsymbol{A}_i\boldsymbol{A}_j^\dagger + \boldsymbol{A}_j\boldsymbol{A}_i^\dagger &= \boldsymbol{0} \quad 1 \le i < j \le 2k,\end{aligned} \tag{3.5}$$

where $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_{2k} \in \mathbb{C}^{n\times p}$. $2k$ is called the cardinality of the solution. ■

We denote the equations in (3.5) by $[k, n, p]$ over $\mathbb{C}$. When $n = p$, the maximum k is equal to

$$\rho_2(n) = \rho_2(2^a(2b+1)) = a + 1. \tag{3.6}$$

The solution to the $[\rho_2(n), n, n]$ Hermitian Hurwitz-Radon equations is not unique either. A solution is called an integer solution if the entries of the matrices $\{\boldsymbol{A}_i, i = 1, \ldots, \rho_2(n)\}$ are in $\{0, \pm 1\}$ and the entries of the remaining matrices $\{\boldsymbol{A}_i, i = \rho_2(n) + 1, \ldots, 2\rho_2(n)\}$ are in $\{0, \pm \mathrm{i}\}$. The existence of a solution to the $[\rho_2(n), n, n]$ Hermitian Hurwitz-Radon equations implies the existence of an integer solution.

We implicitly assume the solutions to (3.3) and (3.5) have the maximum k for given n, p and $\mathbb{F}$. That is, whenever $n = p$, we have $k = \rho_1(n)$ or $k = \rho_2(n)$ for (3.3) and (3.5), respectively. The following property of linear independence will be utilized shortly.

Lemma 1 *Let $\{\boldsymbol{A}_i, i = 1, \ldots, k\}$ be a solution to (3.3) (for $\mathbb{F} = \mathbb{R}$) or (3.5) (for $\mathbb{F} = \mathbb{C}$), then $\{\boldsymbol{A}_i\}$ are linearly independent with respect to real coefficients. To be specific, let $\{a_i\}$ be a set of real numbers.*

$$\sum_{i=1}^{k} a_i \boldsymbol{A}_i = \mathbf{0}_{n\times p}, \tag{3.7}$$

if and only if

$$a_i = 0, \quad i = 1, \ldots k. \tag{3.8}$$

■

Proof: Assume that (3.7) is true, then

$$\begin{aligned}\mathbf{0}_{n\times n} &= \mathbf{0}_{n\times p}\mathbf{0}^{\dagger}_{n\times p} \\ &= \left(\sum_{i=1}^{k} a_i \boldsymbol{A}_i\right)\left(\sum_{i=1}^{k} a_i \boldsymbol{A}_i\right)^{\dagger} \\ &= \left(\sum_{i=1}^{k} {a_i}^2\right)\boldsymbol{I}_{n\times n}.\end{aligned} \tag{3.9}$$

Since a_i's are real numbers, $a_i = 0$ for $i = 1, \ldots, k$. ■

Note that in the above lemma, $\{a_i\}$ are required to be real numbers while $\{\boldsymbol{A}_i\}$ can be complex matrices.

3.2.3 Orthogonal Designs

Alamouti [34] first proposed the 2×2 orthogonal design. This idea was generalized to square and rectangular matrices of arbitrary size [35]. We follow the definitions in [36].

Definition 4 *A real orthogonal design $\mathcal{O}_x$ is defined as an $n \times p$ matrix whose nonzero entries are the real indeterminates $\pm x_1, \pm x_2, \ldots, \pm x_k$ such that,*

$$\mathcal{O}_x \mathcal{O}_x^{\intercal} = \left(\sum_{i=1}^{k} x_i^2\right)\boldsymbol{I}. \tag{3.10}$$

The matrix $\mathcal{O}_x$ is also called a $[k, n, p]$ real orthogonal design. ■

The definition of its complex counterpart goes in parallel.

Definition 5 *A complex orthogonal design $\mathcal{O}_z$ is defined as an $n \times p$ matrix whose nonzero entries are the complex indeterminates $\pm z_1, \pm z_1^*, \pm z_2, \pm z_2^*, \ldots, \pm z_k, \pm z_k^*$,* *such that,*

$$\mathcal{O}_z \mathcal{O}_z^{\dagger} = \left(\sum_{i=1}^{k} |z_i|^2 \right) \boldsymbol{I}, \tag{3.11}$$

The matrix $\mathcal{O}_z$ is also called a $[k, n, p]$ complex orthogonal design. ■

Note the indeterminates x_i in (3.10) can be any real number and z_i in (3.11) can be any complex number. Such a design is also referred as the standard orthogonal design in this chapter. When x_i and z_i are confined to a finite set, the orthogonal design generates a set of matrices which is referred as the O-STC. It is different from the restricted orthogonal design in [37, 38]. The distinction will be elaborated later on.

The benefit of space-time codes from the orthogonal designs is their very low maximum-likelihood decoding complexity. Each symbol in an orthogonal design can be decoded individually, in contrast to the general case where all the symbols from one transmission have to be decoded jointly to achieve maximum-likelihood performance. The drawback of the orthogonal design is the significant rate and mutual information loss for MIMO systems with more than two antennas.

A real orthogonal design $\mathcal{O}_x$ can be expressed as a linear combination of constant real matrices as follows.

$$\mathcal{O}_x = \sum_{i=1}^{k} x_i \boldsymbol{A}_i, \tag{3.12}$$

where $\boldsymbol{A}_i \in \{0, \pm 1\}^{n \times p}$. Similarly, a complex orthogonal design $\mathcal{O}_z$ can be expressed as the following linear combination.

$$\mathcal{O}_z = \sum_{i=1}^{k} \left(\Re(z_i) \boldsymbol{A}_i + \Im(z_i) \boldsymbol{B}_i \right), \tag{3.13}$$

where $\Re(\cdot)$ and $\Im(\cdot)$ are the real part and the imaginary part respectively. $\boldsymbol{A}_i \in \{0, \pm 1\}^{n \times p}$, $\boldsymbol{B}_i \in \{0, \pm \mathrm{i}\}^{n \times p}$ and such that $\mathcal{O}_z$'s entries are in $\{\pm z_k, \pm z_k^*\}$. It can be shown [39] that $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ satisfy the following equalities.

$$\begin{gathered}
\boldsymbol{A}_i\boldsymbol{A}_i^\dagger = \boldsymbol{I}, \quad \boldsymbol{B}_i\boldsymbol{B}_i^\dagger = \boldsymbol{I}, \quad 1 \le i \le k \\
\boldsymbol{A}_i\boldsymbol{A}_j^\dagger = -\boldsymbol{A}_j\boldsymbol{A}_i^\dagger, \quad \boldsymbol{B}_i\boldsymbol{B}_j^\dagger = -\boldsymbol{B}_j\boldsymbol{B}_i^\dagger, \quad 1 \le i \ne j \le k \\
\boldsymbol{A}_i\boldsymbol{B}_j^\dagger = -\boldsymbol{B}_j\boldsymbol{A}_i^\dagger, \quad 1 \le i,j \le k
\end{gathered} \tag{3.14}$$

The orthogonal design and the Hurwitz-Radon equations are closely related in the following sense.

Theorem 3 *The existence of a $[k,n,p]$ real orthogonal design is equivalent to the existence of an integer solution to the $[k,n,p]$ Hurwitz-Radon matrix equations (3.3). The existence of a $[k,n,p]$ complex orthogonal design is equivalent to the existence of an integer solution to the $[k,n,p]$ Hermitian Hurwitz-Radon matrix equations (3.5).* ■

The real part of the theorem can be shown by modifying the proof in [36, Proposition 1] and the complex part by modifying the proof in [39, Appendix] to accommodate the integer constraint.

The integer constraint in Theorem 3 can be removed either in the square case $n = p$ where the constraint does not reduce the cardinality of the solution or in the linear processing orthogonal design case where the entries of $\mathcal{O}_x$ or $\mathcal{O}_z$ are allowed to be linear combinations of $\{\pm x_i\}$ or $\{\pm z_i, \pm z_i^*\}$, respectively.

3.3 Channel-Eigenvector Invariance

In a multi-antenna system, a constellation can be a set of matrices spanning both space and time. These matrices have dimension $n_t \times T$, where n_t is the number of transmit antennas and T is the time span. We call the set of matrices a space-time constellation. Our definition of the space-time constellation includes the well-known schemes such as the Alamouti code [34], the space-time block codes from orthogonal designs [35], and the linear dispersion codes (LDCs)[40].

Throughout the chapter we consider the block fading scenario where the channel coherence time is much larger than the matrices time span T. Thus all of the columns in a space-time matrix are affected by the same channel.

A space-time constellation is called channel-eigenvector invariant if its pairwise error probability does not depend on the eigenvectors of $\boldsymbol{H}^\dagger \boldsymbol{H}$. We write

$$\boldsymbol{H} \sim \boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_n) \tag{3.15}$$

if $\boldsymbol{H}^\dagger \boldsymbol{H}$ has eigenvalues $(\lambda_1, \cdots, \lambda_n)$. Channel-eigenvector invariance is rigorously defined as follows:

Definition 6 *Let $\mathcal{H}(\boldsymbol{\lambda})$ be the set of all $n_r \times n_t$ complex matrices $\boldsymbol{H} \sim \boldsymbol{\lambda}$ for some fixed $\boldsymbol{\lambda}$. A space-time constellation is* ***channel-eigenvector invariant*** *if for any constellation matrix pair $(\boldsymbol{X}_1, \boldsymbol{X}_2)$, both having dimension $n_t \times T$, the squared Euclidean distance between $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ is independent of the eigenvectors of $\boldsymbol{H}^\dagger \boldsymbol{H}$.*

$$\begin{aligned} d^2(\boldsymbol{X}_1 \to \boldsymbol{X}_2 | \boldsymbol{H}) &= Tr\left((\boldsymbol{X}_1 - \boldsymbol{X}_2)^\dagger \boldsymbol{H}^\dagger \boldsymbol{H} (\boldsymbol{X}_1 - \boldsymbol{X}_2)\right) \\ &= Tr\left(\boldsymbol{H}^\dagger \boldsymbol{H} (\boldsymbol{X}_1 - \boldsymbol{X}_2)(\boldsymbol{X}_1 - \boldsymbol{X}_2)^\dagger\right) \\ &= f(\Delta \boldsymbol{X}, \boldsymbol{\lambda}), \end{aligned} \tag{3.16}$$

for any $\boldsymbol{H} \in \mathcal{H}(\boldsymbol{\lambda})$, where $\Delta \boldsymbol{X} = \boldsymbol{X}_1 - \boldsymbol{X}_2$, $Tr(\cdot)$ is the trace of a matrix and $f(\cdot, \cdot)$ is a certain function of matrices. ■

The absence of the eigenvectors of $\boldsymbol{H}^\dagger \boldsymbol{H}$ in (3.16) justifies the name "channel-eigenvector invariant". It has been mentioned earlier that channel-eigenvector invariant constellations are practically useful for the design of compound channel codes. They are also conceptually pleasing. According to (3.2), the mutual information between the channel input and output (and ergodic capacity as well) depends only on the non-zero eigenvalues of $\boldsymbol{H}\boldsymbol{H}^\dagger$ (the same as those of $\boldsymbol{H}^\dagger \boldsymbol{H}$), one would expect that a good constellation also has this property.

A necessary and sufficient condition for a constellation to be channel-eigenvector invariant is based on the following lemma.

Lemma 2 *[9] If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two $n \times n$ positive semidefinite matrices with non-negative eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots, \geq \lambda_n$ and $\gamma_1 \geq \gamma_2 \geq \cdots, \geq \gamma_n$, respectively, then*

$$\sum_{i=1}^{n} \lambda_i \gamma_{n+1-i} \leq Tr(\boldsymbol{AB}) \leq \sum_{i=1}^{n} \lambda_i \gamma_i, \tag{3.17}$$

■

Proof: Suppose that $\boldsymbol{A}$ and $\boldsymbol{B}$ have the following eigenvalue decompositions.

$$\boldsymbol{A} = \boldsymbol{U}^\dagger \boldsymbol{\Lambda} \boldsymbol{U}, \quad \boldsymbol{B} = \boldsymbol{V}^\dagger \boldsymbol{\Gamma} \boldsymbol{V}, \tag{3.18}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices, $\boldsymbol{\Lambda}$ and $\boldsymbol{\Gamma}$ are diagonal matrices. Let $\boldsymbol{Q} = \boldsymbol{U}\boldsymbol{V}^\dagger$, then $\boldsymbol{Q}$ is also a unitary matrix.

$$\begin{aligned} \mathrm{Tr}(\boldsymbol{AB}) &= \mathrm{Tr}(\boldsymbol{U}^\dagger \boldsymbol{\Lambda} \boldsymbol{U} \boldsymbol{V}^\dagger \boldsymbol{\Gamma} \boldsymbol{V}) \\ &= \mathrm{Tr}(\boldsymbol{\Lambda} \boldsymbol{Q} \boldsymbol{\Gamma} \boldsymbol{Q}^\dagger) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} |Q_{ij}|^2 \lambda_i \gamma_j \end{aligned} \tag{3.19}$$

where Q_{ij} is the (i,j) element of $\boldsymbol{Q}$. The lower bound can be obtained by solving the following linear programming problem.

$$\begin{aligned} \min & \sum_{i=1}^{n}\sum_{j=1}^{n} |Q_{ij}|^2 \lambda_i \gamma_j \\ \text{s.t.} & \sum_{i=1}^{n} |Q_{ij}|^2 = 1, \quad i = 1, \cdots, n \\ & \sum_{j=1}^{n} |Q_{ij}|^2 = 1. \quad j = 1, \cdots, n \end{aligned} \tag{3.20}$$

Since both the objective function and the constraints are linear in $|Q_{ij}|^2$, the optimal value is achieved at vertices, which corresponds to $\{|Q_{ij}|^2\}$ being entries of permutation matrices. It is easy to verify the optimal solution is the lower bound in (3.17). The upper bound can be obtained in a similar fashion. ■

Note that the lower bound and the upper bound are equal if and only if at least one of the matrices has identical eigenvalues. Replacing $\boldsymbol{A}$ and $\boldsymbol{B}$ with $\boldsymbol{H}^\dagger\boldsymbol{H}$ and $\Delta\boldsymbol{X}\Delta\boldsymbol{X}^\dagger$ respectively, we have:

Theorem 4 *A space-time constellation is channel-eigenvector invariant if and only if every "squared" difference matrix $\Delta\boldsymbol{X}\Delta\boldsymbol{X}^\dagger$ has identical eigenvalues. i.e.*

$$\lambda = \lambda_1\left(\Delta\boldsymbol{X}\Delta\boldsymbol{X}^\dagger\right) = \cdots = \lambda_n\left(\Delta\boldsymbol{X}\Delta\boldsymbol{X}^\dagger\right), \tag{3.21}$$

where $\Delta\boldsymbol{X} = \boldsymbol{X}_1 - \boldsymbol{X}_2$. $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are two arbitrary matrices in the space-time constellation. i.e.

$$(\boldsymbol{X}_1 - \boldsymbol{X}_2)(\boldsymbol{X}_1 - \boldsymbol{X}_2)^\dagger = \Delta\boldsymbol{X}\Delta\boldsymbol{X}^\dagger = \lambda\boldsymbol{I}, \tag{3.22}$$

■

3.4 Channel-Eigenvector Invariant STCs

We will first study the structure of channel-eigenvector invariant space-time constellations under the linear constraint.

3.4.1 Linear CEI-STCs

Definition 7 *A real linear space-time constellation $\mathcal{L}_x$ is a set of real matrices of dimension $n_t \times T$ such that*

$$\mathcal{L}_x = \{\boldsymbol{X} \mid \boldsymbol{X} = \sum_{i=1}^{k} x_i \boldsymbol{A}_i\}, \tag{3.23}$$

where $\{\boldsymbol{A}_i \in \mathbb{R}^{n_t \times T},\ i = 1, \ldots, k\}$ is called the basis of $\mathcal{L}_x$. Each x_i is a real number. Similarly, a complex linear space-time constellation $\mathcal{L}_z$ is a set of complex matrices of dimension $n_t \times T$ such that

$$\mathcal{L}_z = \{\boldsymbol{X} \mid \boldsymbol{X} = \sum_{i=1}^{k} (\Re(z_i)\boldsymbol{A}_i + \Im(z_i)\boldsymbol{B}_i)\}. \tag{3.24}$$

where $\{\boldsymbol{A}_i,\ \boldsymbol{B}_i \in \mathbb{C}^{n_t \times T},\ i = 1, \ldots, k\}$ is called the basis of $\mathcal{L}_z$. Each z_i is a complex number. ■

Note that equations (3.23) and (3.24) are similar to (3.12) and (3.13) respectively. This is not surprising because orthogonal designs are linear space-time constellations. The following lemma establishes the conditions for a linear space-time constellation to be channel-eigenvector invariant.

Lemma 3 *A real linear space-time constellation defined in (3.23) is channel-eigenvector invariant if and only if*

$$\boldsymbol{A}_i \boldsymbol{A}_i^{\intercal} = \xi_i \boldsymbol{I}, \quad i = 1, \cdots, k, \tag{3.25}$$

$$\boldsymbol{A}_i\boldsymbol{A}_j^{\intercal} + \boldsymbol{A}_j\boldsymbol{A}_i^{\intercal} = \xi_{ij}\boldsymbol{I}, \quad 1 \le i < j = k, \tag{3.26}$$

where each ξ_i is a positive real number and each ξ_{ij} is a real number.

Similarly a complex linear space-time constellation defined in (3.24) is channel-eigenvector invariant if and only if

$$\boldsymbol{A}_i\boldsymbol{A}_i^{\dagger} = \xi_{A_i}\boldsymbol{I}, \quad \boldsymbol{B}_i\boldsymbol{B}_i^{\dagger} = \xi_{B_i}\boldsymbol{I}, \quad 1 \le i \le k \tag{3.27}$$

$$\boldsymbol{A}_i\boldsymbol{A}_j^{\dagger} + \boldsymbol{A}_j\boldsymbol{A}_i^{\dagger} = \xi_{A_{ij}}\boldsymbol{I}, \quad 1 \le i \ne j \le k \tag{3.28}$$

$$\boldsymbol{B}_i\boldsymbol{B}_j^{\dagger} + \boldsymbol{B}_j\boldsymbol{B}_i^{\dagger} = \xi_{B_{ij}}\boldsymbol{I}, \quad 1 \le i \ne j \le k \tag{3.29}$$

$$\boldsymbol{A}_i\boldsymbol{B}_j^{\dagger} + \boldsymbol{B}_j\boldsymbol{A}_i^{\dagger} = \xi_{A_iB_j}\boldsymbol{I}, \quad 1 \le i, j \le k, \tag{3.30}$$

where ξ_{A_i} and ξ_{B_i} are positive real numbers. $\xi_{A_{ij}}$, $\xi_{B_{ij}}$ and $\xi_{A_iB_j}$ are real numbers.
■

Proof: We first prove the real case. Suppose $\mathcal{L}_x$ is a channel-eigenvector invariant constellation. Let $\boldsymbol{X}_1$, $\boldsymbol{X}_2 \in \mathcal{L}_x$ be expressed as follows.

$$\boldsymbol{X}_1 = \sum_{i=1}^{k} x_{1i}\boldsymbol{A}_i, \quad \boldsymbol{X}_2 = \sum_{i=1}^{k} x_{2i}\boldsymbol{A}_i. \tag{3.31}$$

$\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are said to differ at i if $x_{1i} \ne x_{2i}$, but $x_{1j} = x_{2j}$ for all $j \ne i$. They are said to differ at (i, j) if $x_{1k} \ne x_{2k}, k = i, j$ but $x_{1k} = x_{2k}$ for all $k \ne i, j$. According to Theorem 4, we have the following equality:

$$\begin{aligned}\lambda \boldsymbol{I} &= (\boldsymbol{X}_1 - \boldsymbol{X}_2)(\boldsymbol{X}_1 - \boldsymbol{X}_2)^{\mathsf{T}} \\ &= \sum_{i=1}^{k} (x_{1i} - x_{2i})^2 \boldsymbol{A}_i \boldsymbol{A}_i^{\mathsf{T}} \\ &+ \sum_{i=1}^{k} \sum_{j>i}^{k} (x_{1i} - x_{2i})(x_{1j} - x_{2j})(\boldsymbol{A}_i \boldsymbol{A}_j^{\mathsf{T}} + \boldsymbol{A}_j \boldsymbol{A}_i^{\mathsf{T}})\end{aligned}, \tag{3.32}$$

where λ is a non-negative real number dependent on the matrix pair. By letting $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ differ at i, we have (3.25). And by letting $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ differ at (i, j), we obtain (3.26).

The converse can be proved by simply verifying that (3.32), and hence (3.22) holds if (3.25) and (3.26) are true.

The complex case follows in the same fashion by letting $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ differ at i's real part or imaginary part. ■

Apparently, Lemma 3 implies that constellations from the orthogonal designs are channel-eigenvector invariant. Actually we can say more about this.

Theorem 5 *A real linear space-time constellation $\mathcal{L}_x$ is channel-eigenvector invariant if and only if each matrix $\boldsymbol{A}_i$ can be written as a linear combination of a solution to the Hurwitz-Radon matrix equations. To be specific, $\boldsymbol{A}_i$ can be written as*

$$\boldsymbol{A}_i = \sum_{j=1}^{m} a_{ij} \boldsymbol{O}_j, \quad i = 1, \cdots, k, \tag{3.33}$$

where $m \leq k$. Each a_{ij} is a real number. $\{\boldsymbol{O}_j, j = 1, \cdots, m\}$ is a solution to the $[m, n_t, T]$ Hurwitz-Radon matrix equations.

Similarly, a complex linear space-time constellation $\mathcal{L}_z$ is channel-eigenvector invariant if and only if matrices $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ can be written as a linear combination *of a solution to the Hermitian Hurwitz-Radon matrix equations. To be specific,* $\boldsymbol{A}_i$ *and* $\boldsymbol{B}_i$ *can be written as*

$$\begin{aligned}\boldsymbol{A}_i &= \sum_{j=1}^{m} a_{ij}\boldsymbol{O}_j + \sum_{j=1}^{m} b_{ij}\boldsymbol{Q}_j, \quad i = 1, \cdots, k \\ \boldsymbol{B}_i &= \sum_{j=1}^{m} c_{ij}\boldsymbol{O}_j + \sum_{j=1}^{m} d_{ij}\boldsymbol{Q}_j, \quad i = 1, \cdots, k\end{aligned} \tag{3.34}$$

where $m \leq k$*.* a_{ij}*,* b_{ij}*,* c_{ij} *and* d_{ij} *are real numbers.* $\{\boldsymbol{O}_j,\, \boldsymbol{Q}_j,\, j = 1, \cdots, m\}$ *is a solution to the* $[m, n_t, T]$ *Hermitian Hurwitz-Radon matrix equations.* ■

Proof: See Appendix. ■

By Theorem 5, we can rewrite any linear CEI-STC $\boldsymbol{X}$ as

$$\begin{aligned}\boldsymbol{X} &= \sum_{i=1}^{k} x_i \boldsymbol{A}_i \\ &= \sum_{i=1}^{k} x_i \sum_{j=1}^{m} a_{ij}\boldsymbol{O}_j \\ &= \sum_{j=1}^{m} \left(\sum_{i=1}^{k} x_i a_{ij} \boldsymbol{O}_j \right) \\ &= \sum_{j=1}^{m} x'_j \boldsymbol{O}_j,\end{aligned} \tag{3.35}$$

where $x'_j = \sum_{i=1}^{k} x_i a_{ij}$. This indicates that $\boldsymbol{X}$ is a generalization of a linear processing orthogonal design [35, Definition 3.4.1] since (3.35) means $\boldsymbol{X}\boldsymbol{X}^{\mathsf{T}} = \sum_{j=1}^{m} \left(\sum_{i=1}^{k} x_i a_{ij} \right)^2 \boldsymbol{I}$. It degrades to a linear processing orthogonal design if $m = k$ and $a_{ij} = 0$ for $i \neq j$. Although this generalization provides more freedom in the choice of the basis $\{\boldsymbol{A}_i\}$ ($\{\boldsymbol{A}_i, \boldsymbol{B}_i\}$ for the complex case), it sacrifices the low decoding complexity enjoyed by the orthogonal design because the symbols might be mixed together. Furthermore, it does not increase the maximum number of linearly independent matrices, which is limited by the solution to the Hurwitz-Radon matrix equations.

3.4.2 Restricted CEI-STCs

We have studied the structure of the linear CEI-STC. Recall that there is no requirement on the entries of the matrices defined in (3.23) and (3.24). However, the transmitted signals are usually confined to a finite set, say points from a QPSK or an 8-PSK constellation. This set is called the base set, denoted as $\mathcal{S}$. This section explores how the restricted base set changes the characterization of a CEI-STC.

The base set is real if it consists of real numbers, otherwise it is called complex. If $\mathcal{S}$ is real, we assume that it is symmetric about the origin, i.e. $x_i \in \mathcal{S}$ implies $-x_i \in \mathcal{S}$. If $\mathcal{S}$ is complex, we assume that it is symmetric relative to both the x-axis and the y-axis on the complex plane, i.e. $z_i \in \mathcal{S}$ implies $-z_i$ and $\pm z_i^* \in \mathcal{S}$. The symmetry of $\mathcal{S}$ is also used in O-STCs and LDCs where the negation and complex conjugate of a transmitted signal are allowed. Furthermore, 0 is assumed to be in $\mathcal{S}$ which allows an antenna to send nothing during a time slot.

Denote $\mathcal{N}$ to be a set of $n_t \times T$ matrices whose entries are drawn from $\mathcal{S}$. $\mathcal{N}$ is the set of all matrices that could possibly be transmitted. A subset $\mathcal{M} \subset \mathcal{N}$ is channel-eigenvector invariant if it satisfies (3.22) in Theorem 4. The subset $\mathcal{M}$ is called a restricted CEI-STC. To distinguish the restricted CEI-STC from the linear one, note that the linear constraint (3.23) or (3.24) is absent but there is a restriction of the base set. After proving Theorem 6, we will see that the restricted CEI-STC has an affine structure.

For the restricted CEI-STC, a major concern is the cardinality of $\mathcal{M}$. The cardinality does not make sense in the linear CEI-STC case because there is no restriction on the base set. However the cardinality of a restricted CEI-STC decides the raw data rate of the space-time constellation. For example, an Alamouti code over QPSK generates 16 matrices of dimension 2×2, meaning each matrix conveys 4 bits information. A fundamental problem is to determine the maximum cardinality of $\mathcal{M}$, given the dimension of the matrix $n_t \times T$ and the base set $\mathcal{S}$. In the following, we will mainly concentrate on the $n_t = T$ case and briefly discuss the case where $n_t < T$. It is clear that if $n_t > T$, then $|\mathcal{M}| = 0$.

Care should also be taken to distinguish between the O-STC and the restricted orthogonal design [37, 38]. The former is the set of matrices generated by the standard orthogonal design. The orthogonal design as defined in (3.10) or (3.11) is first constructed and then the limitation of the variables to a finite set is applied. As for the latter, the variables are first limited to a finite set and then the orthogonal design condition is met by properly choosing the basis matrices. It has been shown that the restricted orthogonal design can be different from the O-STC if $\mathcal{S}$ is sufficiently constrained. Take an example from [37], let $\boldsymbol{H}$ be an $n \times n$ Hadamard matrix, $\boldsymbol{\Lambda}_x$ be an $n \times n$ diagonal matrix with x_1, x_2, ..., x_n along the diagonal. Each x_i can only take values 1 or -1. $\boldsymbol{H}\boldsymbol{\Lambda}_x$ forms an restricted orthogonal design which has n variables rather than $\rho_1(n)$. Note that $n > \rho_1(n)$ for $n > 8$, for example, $\rho_1(10) = 2$. The set of matrices $\boldsymbol{H}\boldsymbol{\Lambda}_x$ with entries in $\{\pm 1\}$ has cardinality of 2^n. Later we will show that there does not exist an $n \times n$ restricted CEI-STC with the same cardinality over $\{\pm 1\}$.

A sufficient condition for a restricted orthogonal design to be equivalent to an standard orthogonal design was found in [41]. However for a given base set, the maximum number of matrices that form a restricted orthogonal design remains an open question. Furthermore, unlike the standard case, not every restricted orthogonal design is channel-eigenvector invariant. The difference of an arbitrary pair of matrices in the above Hadamard matrix example might not even be of full rank. Thus we will only compare the restricted CEI-STC to the O-STC rather than the restricted orthogonal design.

Our main results focus on the case where $n_t = T$. The matrices generated from the orthogonal design form a channel-eigenvector invariant set. However, it is not known whether it is a maximum set for the restricted CEI-STC. Our main result is presented in the following theorem.

Theorem 6 *Let $\mathcal{M}$ be a maximum-cardinality channel-eigenvector invariant subset of $\mathcal{N}$ where the matrices are of dimension $n \times n$. Denote the size of the base set by $s = |\mathcal{S}|$. Then, $|\mathcal{M}| = s^{\rho_1(n)}$ when $\mathcal{S}$ is real and*

$$s^{\rho_2(n)} \leq |\mathcal{M}| \leq (\lfloor \frac{s}{2} \rfloor + 1)^{2\rho_2(n)} \tag{3.36}$$

when $\mathcal{S}$ is complex. For the latter, both bounds are tight in the sense that each of them is the maximum cardinality under certain selection of n and $\mathcal{S}$. ■

To unify the proof, $|\mathcal{M}| = s^{\rho_1(n)}$ in the real case can be regarded as $s^{\rho_1(n)} \leq |\mathcal{M}| \leq s^{\rho_1(n)}$. First consider the lower bound for both cases. Since the set of matrices generated by the $[\rho_1(n), n, n]$ real orthogonal design or the $[\rho_2(n), n, n]$ complex orthogonal design over $\mathcal{S}$ satisfies (3.22), the size of this set serves as a lower bound. There are $s^{\rho_1(n)}$ ways to choose $\{x_i\}$ in (3.10) and $s^{\rho_2(n)}$ ways to choose $\{z_i\}$ in (3.11).

To consider the upper bound, we need to define the difference set.

Definition 8 *Let $\mathcal{S}$ be a set of numbers. The difference set of $\mathcal{S}$, denoted $\Delta_{\mathcal{S}}$ is defined by*

$$\Delta_{\mathcal{S}} = \{s_i - s_j | \forall s_i, s_j \in \mathcal{S}\} \tag{3.37}$$

■

Assume $\mathcal{M} = \{\boldsymbol{A}_i, i = 1, \ldots, k\}$ to be a maximum channel-eigenvector invariant constellation for a base set $\mathcal{S}$. Define

$$\boldsymbol{B}_i = \boldsymbol{A}_i - \boldsymbol{A}_1, \quad i = 2, \ldots, k. \tag{3.38}$$

The following lemma states that $\boldsymbol{B}_i$ is generated by a solution to the (Hermitian) Hurwitz-Radon matrix equations.

Lemma 4 *Let $\{\boldsymbol{A}_i\}$ and $\{\boldsymbol{B}_i\}$ be defined above. Let $|\Delta_{\mathcal{S}}| = \delta_s$. If $\Delta_{\mathcal{S}}$ is real, then there exists a solution to the $[\rho_1(n), n, n]$ Hurwitz-Radon matrix equations, denoted as $\{\boldsymbol{O}_j, j = 1, \ldots, \rho_1(n)\}$, such that*

$$\boldsymbol{B}_i = \sum_{j=1}^{\rho_1(n)} a_{ij}\boldsymbol{O}_j, \quad i = 2, \ldots, k, \tag{3.39}$$

where each a_{ij} is a real number. Similarly, if $\Delta_{\mathcal{S}}$ is complex, then there exists a solution to the $[\rho_2(n), n, n]$ Hermitian Hurwitz-Radon matrix equations, denoted as $\{\boldsymbol{O}_j, \boldsymbol{Q}_j, j = 1, \ldots, \rho_2(n)\}$, such that

$$\boldsymbol{B}_i = \sum_{j=1}^{\rho_2(n)} c_{ij}\boldsymbol{O}_j + \sum_{j=1}^{\rho_2(n)} d_{ij}\boldsymbol{Q}_j, \quad i = 2, \cdots, k, \tag{3.40}$$

where c_{ij}, d_{ij} are real numbers. ■

Proof: Consider the complex case. According to the definition, $\{\boldsymbol{A}_i\}$ and $\{\boldsymbol{B}_i\}$ have the following relationship.

$$\boldsymbol{B}_i - \boldsymbol{B}_j = \boldsymbol{A}_i - \boldsymbol{A}_j \quad 2 \leq i, j \leq k \tag{3.41}$$

Since $\{\boldsymbol{A}_i\}$ is channel-eigenvector invariant, $\boldsymbol{B}_i$ satisfies

$$\begin{aligned} \boldsymbol{B}_i\boldsymbol{B}_i^\dagger &= \xi_i \boldsymbol{I} \quad i = 2, \dots k \\ (\boldsymbol{B}_i - \boldsymbol{B}_j)(\boldsymbol{B}_i - \boldsymbol{B}_j)^\dagger &= \xi_{ij}\boldsymbol{I} \quad 2 \leq i < j \leq k, \end{aligned} \tag{3.42}$$

where ξ_i and ξ_{ij} are positive real numbers. Now expanding the second equation, we have

$$\boldsymbol{B}_i\boldsymbol{B}_j^\dagger + \boldsymbol{B}_j\boldsymbol{B}_i^\dagger = \xi'_{ij}\boldsymbol{I} \quad 2 \leq i < j \leq k, \tag{3.43}$$

where ξ'_{ij} is real. Using the Gram-Schmidt technique from Lemma 11 in the Appendix, we conclude that $\boldsymbol{B}_i$ can be written as

$$\boldsymbol{B}_i = \sum_{j=1}^{\rho_2(n)} c_{ij}\boldsymbol{O}_j + \sum_{j=1}^{\rho_2(n)} d_{ij}\boldsymbol{Q}_j, \quad i = 2, \cdots, k, \tag{3.44}$$

where $\{\boldsymbol{O}_i, \boldsymbol{Q}_i,\ i = 1, \dots, \rho_2(n)\}$ is a solution to the the $[\rho_2(n), n, n]$ Hermitian Hurwitz-Radon matrix equations. c_{ij} and d_{ij} are real numbers. The real case follows in the same manner. ■

For any channel-eigenvector invariant set $\{\boldsymbol{A}_i\}$, $\{\boldsymbol{B}_i\}$ have the form in (3.39) or (3.40) with all entries in $\Delta_{\mathcal{S}}$. Because of the structure, we can first bound the maximum size of $\{\boldsymbol{B}_i\}$ then obtain $\{\boldsymbol{A}_i\}$ by properly selecting $\boldsymbol{A}_1$. This idea is used to find the upper bound.

Proof of Theorem 6 (the upper bound): Assume $\mathcal{S}$ is real. We will show that $\boldsymbol{A}_i$ is completely determined by the values of its $\rho_1(n)$ entries. Let $\{\boldsymbol{B}_i\}$ and $\{\boldsymbol{O}_i\}$ be defined as in Lemma 4. Since all $\boldsymbol{O}_i$'s are linearly independent, (3.39) says that only $\rho_1(n)$ entries of $\boldsymbol{B}_i$ can assume values freely in $\Delta_{\mathcal{S}}$. These entries determine a_{ij}'s in (3.39) completely, which in turn fix the remaining entries of $\boldsymbol{B}_i$. Because $\boldsymbol{A}_i = \boldsymbol{B}_i + \boldsymbol{A}_1$, $\boldsymbol{A}_i$ is also determined by its corresponding $\rho_1(n)$ entries for any given $\boldsymbol{A}_1$. There are only $s^{\rho_1(n)}$ possible combinations for these $\rho_1(n)$ entries of $\boldsymbol{A}_i$ because all of them must be in $\mathcal{S}$.

When $\mathcal{S}$ is complex, Lemma 1 only ensures that $\{\boldsymbol{O}_i, \boldsymbol{Q}_i, i = 1, \ldots, \rho_2(n)\}$ are linearly independent with real coefficients. We have to consider the real and imaginary parts of (3.40) separately. Let $\boldsymbol{B}_i = \hat{\boldsymbol{B}}_i + \mathrm{i}\tilde{\boldsymbol{B}}_i$, $\boldsymbol{O}_i = \hat{\boldsymbol{O}}_i + \mathrm{i}\tilde{\boldsymbol{O}}_i$ and $\boldsymbol{Q}_i = \hat{\boldsymbol{Q}}_i + \mathrm{i}\tilde{\boldsymbol{Q}}_i$. The variables with $\hat{}$ are the real parts and the ones with $\tilde{}$ are the imaginary parts. (3.40) is equivalent to the following:

$$\hat{\boldsymbol{B}}_i = \sum_{j=1}^{\rho_2(n)} c_{ij}\hat{\boldsymbol{O}}_j + \sum_{j=1}^{\rho_2(n)} d_{ij}\hat{\boldsymbol{Q}}_j, \quad i = 2, \cdots, k, \tag{3.45}$$

$$\tilde{\boldsymbol{B}}_i = \sum_{j=1}^{\rho_2(n)} c_{ij}\tilde{\boldsymbol{O}}_j + \sum_{j=1}^{\rho_2(n)} d_{ij}\tilde{\boldsymbol{Q}}_j, \quad i = 2, \cdots, k. \tag{3.46}$$

According to Lemma 1, $\{\boldsymbol{O}_i, \boldsymbol{Q}_i, i = 1, \ldots, \rho_2(n)\}$ are linearly independent with respect to real coefficients, which implies either $\{\hat{\boldsymbol{O}}_i, \hat{\boldsymbol{Q}}_i, i = 1, \ldots, \rho_2(n)\}$ or $\{\tilde{\boldsymbol{O}}_i, \tilde{\boldsymbol{Q}}_i, i = 1, \ldots, \rho_2(n)\}$ or both are linearly independent. Without loss of generality, we assume the first case is true. Then c_{ij} and d_{ij} can be solved by specifying $2\rho_2(n)$ entries of $\hat{\boldsymbol{B}}_i$. Fix $\boldsymbol{A}_1$, let $\boldsymbol{A}_i = \boldsymbol{A}_1 + \boldsymbol{B}_i$. Then $\boldsymbol{A}_i$ is also determined by the real part of the corresponding $2\rho_2(n)$ entries. Since each entry of $\boldsymbol{A}_i$ has to be in $\mathcal{S}$, the real part can assume at most $(\lfloor \frac{s}{2} \rfloor + 1)$ values due to the symmetry of $\mathcal{S}$. Thus there are at most $(\lfloor \frac{s}{2} \rfloor + 1)^{2\rho_2(n)}$ possible values for $\boldsymbol{A}_i$.

This gives the upper bound. Cases where each of the bounds is the maximum cardinality are shown in the example below. ■

The proof reveals the affine structure of the restricted CEI-STC which is formed by a linear CEI-STC $\{\boldsymbol{B}_i\}$ shifted by $\boldsymbol{A}_1$. Recall we did not explicitly impose any linear constraint like (3.23) or (3.24), the linearity comes into play automatically. If the all-zero matrix is in the restricted CEI-STC, then it is also a linear CEI-STC.

Example 1 (achievability of the bounds): We first show that the lower bound is the maximum cardinality when $n = 2$ and $\mathcal{S} = \{0, \pm 1 \pm \mathsf{i}\}$, which is the standard QPSK with the origin. Following the notation in Lemma 4, notice that if a 2×2 complex matrix $\boldsymbol{B}_i$ satisfies

$$\boldsymbol{B}_i \boldsymbol{B}_i^{\dagger} = \xi_i \boldsymbol{I}, \tag{3.47}$$

where ξ_i is a positive number, then it has the following form.

$$\begin{bmatrix} a & b \\ -b^* e^{\mathsf{i}(\delta+2\theta)} & a e^{\mathsf{i}\delta} \end{bmatrix} \tag{3.48}$$

where a, b are arbitrary complex numbers. $a = |a|e^{\mathsf{i}\theta}$, δ is an arbitrary angle. Since all the entries of (3.48) are in $\Delta_{\mathcal{S}}$ which has 13 points, $\delta + 2\theta$ has to be one of the following values $\{0,\ \pi/2,\ \pi,\ 3\pi/2\}$. Thus $\boldsymbol{B}_i$ takes one of the following four forms.

$$\begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}, \quad \begin{bmatrix} a & b \\ b^* & -a^* \end{bmatrix}, \quad \begin{bmatrix} a & b \\ b^*\mathsf{i} & -a^*\mathsf{i} \end{bmatrix}, \quad \begin{bmatrix} a & b \\ -b^*\mathsf{i} & a^*\mathsf{i} \end{bmatrix}. \tag{3.49}$$

Suppose that each $\boldsymbol{B}_i$ takes the first form, then (3.43) is automatically satisfied. It is left to choose a proper $\boldsymbol{A}_1$. Define $\mathcal{M}$ to be a subset of $\{\boldsymbol{A}_1, \boldsymbol{A}_1 + \boldsymbol{B}_i, i =$

$2, \ldots m\}$ such that $\forall\ \boldsymbol{X} \in \mathcal{M}$, all of $\boldsymbol{X}$'s entries are in $\mathcal{S}$. An exhaustive computer search shows that $|\mathcal{M}|$ is maximized if and only if $\boldsymbol{A}_1$ takes the same form as $\boldsymbol{B}_i$. This leads to an orthogonal design and the lower bound of (3.36) is the best that can be achieved. A similar argument works if $\boldsymbol{B}_i$ takes the other forms.

Next we show that the upper bound is the maximum cardinality when $n = 8$ and $\mathcal{S} = \{0, \pm a_i, \pm \mathrm{i} b_i |\ a_i, b_i \in \mathbb{R},\ i = 1, \ldots, m\}$, i.e., the base set consists of real and pure imaginary numbers. Let

$$\boldsymbol{Z} = \begin{bmatrix} x_1 & \mathrm{i}y_1 & x_2 & \mathrm{i}y_2 & x_3 & \mathrm{i}y_2 & x_4 & \mathrm{i}y_4 \\ \mathrm{i}y_1 & x_1 & -\mathrm{i}y_2 & -x_2 & -\mathrm{i}y_3 & -x_3 & \mathrm{i}y_4 & x_4 \\ -x_2 & -\mathrm{i}y_2 & x_1 & \mathrm{i}y_1 & x_4 & \mathrm{i}y_4 & -x_3 & -\mathrm{i}y_3 \\ \mathrm{i}y_2 & x_2 & \mathrm{i}y_1 & x_1 & -\mathrm{i}y_4 & -x_4 & -\mathrm{i}y_3 & -x_3 \\ -x_3 & -\mathrm{i}y_3 & -x_4 & -\mathrm{i}y_4 & x_1 & \mathrm{i}y_1 & x_2 & \mathrm{i}y_2 \\ \mathrm{i}y_3 & x_3 & \mathrm{i}y_4 & x_4 & \mathrm{i}y_1 & x_1 & \mathrm{i}y_2 & x_2 \\ -x_4 & \mathrm{i}y_4 & x_3 & -\mathrm{i}y_3 & -x_2 & \mathrm{i}y_2 & x_1 & -\mathrm{i}y_1 \\ \mathrm{i}y_4 & -x_4 & -\mathrm{i}y_3 & x_3 & \mathrm{i}y_2 & -x_2 & -\mathrm{i}y_1 & x_1 \end{bmatrix}.$$

$\boldsymbol{Z}$ generates a restricted CEI-STC. Since each x_i can assume any value in $\{0, a_i, i = 1, \ldots, m\}$ and each y_i can assume any value in $\{0, b_i, i = 1, \ldots, m\}$, $\boldsymbol{Z}$ can be any of the $(m+1)^8$ distinct matrices. Recall that $\rho_2(8) = 4$ and $s = 2m+1$. The upper bound of (3.36) is achieved.

$\boldsymbol{Z}$ is neither a real nor a complex orthogonal design because its entries can not be arbitrary real or complex numbers. In contrast, the 8×8 complex orthogonal design with entries restricted to $\mathcal{S}$ generates $(2m+1)^4$ matrices, which is equal to the lower bound of (3.36). So in this example, the restricted CEI-STC is larger than the complex O-STC. The equality $\boldsymbol{Z}\boldsymbol{Z}^\dagger = \sum_{i=1}^{4}(x_i^2 + y_i^2)$ shows that $\boldsymbol{Z}$ itself is a restricted orthogonal design. It is not clear nonetheless whether all restricted CEI-STCs are also restricted orthogonal designs. As a by-product, this example verifies that a restricted orthogonal design can be different than the standard one.

We now briefly comments on the case $n_t < T$. Although the basic idea should be the same, there exist two difficulties. First, we have extensively used the solution to the $[k, n, n]$ Hurwitz-Radon matrix equations to obtain Theorem 6. However, the solution to the $[k, n_t, T]$ Hurwitz-Radon matrix equations when $n_t \neq T$ remains open. Second, it has been shown the $[k, n, n]$ Hurwitz-Radon matrix equations always have integer solutions as long as there exists an arbitrary solution. Whether this holds when $n_t \neq T$ has not been proved, although it is conjectured to be true [31, p. 314].

Efforts have been made to obtain the maximum rate of k/T when the solution is integral. Techniques presented in [36] can be used to construct orthogonal designs with arbitrary n_t, however there is no control over the value of T. Once we obtain a $[k, n_t, T]$ orthogonal design, then the number of the matrices generated by the orthogonal design always serves as a lower bound on the cardinality of a CEI-STC.

3.5 Mutual Information Loss of CEI-STCs

Due to the stringent constraint in (3.21), a CEI-STC in general does not have enough degrees of freedom to maintain the mutual information promised by a multiple-antenna system. To analyze the mutual information loss, we shall rewrite (3.23) and (3.24) as (3.50) and (3.51) respectively.

$$\boldsymbol{X} = \begin{bmatrix} \hat{\boldsymbol{A}}_1\boldsymbol{x} & \hat{\boldsymbol{A}}_2\boldsymbol{x} & \cdots & \hat{\boldsymbol{A}}_T\boldsymbol{x} \end{bmatrix} \tag{3.50}$$

$$\boldsymbol{X} = \begin{bmatrix} \hat{\boldsymbol{A}}_1\hat{\boldsymbol{z}} + \hat{\boldsymbol{B}}_1\tilde{\boldsymbol{z}} & \hat{\boldsymbol{A}}_2\hat{\boldsymbol{z}} + \hat{\boldsymbol{B}}_2\tilde{\boldsymbol{z}} & \cdots & \hat{\boldsymbol{A}}_T\hat{\boldsymbol{z}} + \hat{\boldsymbol{B}}_T\tilde{\boldsymbol{z}} \end{bmatrix} \tag{3.51}$$

where

$$\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_k \end{bmatrix}^{\mathsf{T}} \tag{3.52}$$

$$\hat{\boldsymbol{z}} = \begin{bmatrix} \Re(z_1) & \Re(z_2) & \cdots & \Re(z_k) \end{bmatrix}^{\mathsf{T}} \tag{3.53}$$

$$\tilde{\boldsymbol{z}} = \begin{bmatrix} \Im(z_1) & \Im(z_2) & \cdots & \Im(z_k) \end{bmatrix}^{\mathsf{T}}. \tag{3.54}$$

Each $\hat{\boldsymbol{A}}_i$ consists of the ith columns of $\{\boldsymbol{A}_j,\ j = 1, \dots k\}$ and each $\hat{\boldsymbol{B}}_i$ consists of the ith columns of $\{\boldsymbol{B}_j,\ j = 1, \dots k\}$.

The space-time constellation constrained channel mutual information can be computed by forming a large MIMO channel including the matrices $\{\hat{\boldsymbol{A}}_i\}$ and $\{\hat{\boldsymbol{B}}_i\}$ [40]. We need to distinguish the space-time constellation constrained channel from the conventional constellation-constrained channel. In the former case, the constraint comes from the the space and time formation of the input signals which are still Gaussian distributed, while in the latter case, the inputs are limited to a finite set.

A space-time constellation is **mutual-information lossy** if the mutual information of the constellation constrained channel is less than that of the unconstrained one for i.i.d. Gaussian input. For a constellation defined in (3.50) or (3.51) to be mutual information preserving, the following condition must be met.

Theorem 7 *Consider the linear constellation defined in (3.23) and (3.24). Assume the basis matrices satisfying the following power constraints:*

$$Tr\left(\hat{\boldsymbol{A}}_i\hat{\boldsymbol{A}}_i^{\mathsf{T}}\right) = n_t, \quad i = 1, \dots, k \tag{3.55}$$

$$Tr\left(\hat{\boldsymbol{A}}_i\hat{\boldsymbol{A}}_i^{\dagger} + \hat{\boldsymbol{B}}_i\hat{\boldsymbol{B}}_i^{\dagger}\right) = 2n_t, \quad i = 1, \dots, k \tag{3.56}$$

where (3.55) is for the real case and (3.56) is for the complex case. The constellation is mutual-information lossy on average under Rayleigh fading channels with i.i.d Gaussian input if $k < \min(n_t, n_r) \times T$. ■

Proof: See Appendix. ■

We do not claim that Theorem 7 is new. Actually, it is a *de facto* standard in the LDC design. For example, the authors in [40] assert $k = \min(n_t, n_r) \times T$ "since this tends to maximize the mutual information ". However we are not able to find a rigorous proof. The caveat is that the mutual information loss in Theorem 7 is on average. There exists a particular realization of the channel under which the space-time constellation constrained mutual information is actually larger than the unconstrained one. This is because the space-time constellation effectively performs some "water-filling". We use mutual information instead of capacity because the capacity of a constrained channel is not achieved by an i.i.d. Gaussian input given the transmitter knows the matrices $\{\hat{\boldsymbol{A}}_i\}$ and $\{\hat{\boldsymbol{B}}_i\}$. The space-time constellation constrained channel can be viewed as a fading channel with correlation. The capacity of such channels was derived in [42].

An immediate implication of Theorem 7 is:

Corollary 1 *If a square CEI-STC can be written as (3.50) or (3.51), which satisfies the power constraints in (3.55) or (3.56), then the CEI-STC is mutual-information lossy under Rayleigh fading channels with i.i.d. Gaussian input except the following cases:* $n_t = T = 2$ *or* 4 *or* 8, $n_r = 1$ *if the constellation is real and* $n_t = T = 2$, $n_r = 1$ *if the constellation is complex.* ■

Proof: We need only compare k and $\min(n_t, n_r) \times T$ according to Theorem 7. For the real linear CEI-STC, (3.35) says that $k \leq \rho_1(n)$. So $k \leq \min(n_t, n_r) \times T$ except the cases listed in the corollary. A similar result holds for the complex case. For the restricted CEI-STC whose entries are limited to a finite set, the Gaussian input mutual information should be interpreted as follows. We find a set of independent variables whose values determine the entire matrix in a restricted CEI-STC. These variables are regarded as the input of the channel and assumed to be i.i.d. Gaussian distributed. From the proof of Theorem 7, the number of independent variables is at most $\rho_1(n)$ for the real case and at most $2\rho_2(n)$ real parts for the complex case. The result follows by comparing these numbers with $\min(n_t, n_r) \times T$.

3.6 Summary

It is interesting to compare the orthogonal design with the CEI-STC. The two concepts originate from different perspectives. The former focuses on the matrices themselves, mainly pursued for its low decoding complexity, while the latter places requirements on their difference, mainly pursued for the invariance of probability of error under different channel eigenvectors. Depending on the base set, a complex restricted CEI-STC can have larger cardinality than the complex O-STC. If the entries are real numbers, then the restricted CEI-STC has the same cardinality of the O-STC.

No matter whether the entries are restricted in a finite set or whether the linear constraints (3.23) or (3.24) are imposed, the CEI-STC are limited by the solution to the (Hermitian) Hurwitz-Radon matrix equations, which is the direct cause of mutual information loss. It has been shown by examples that the approach that first restricts the base set then orthogonalizes the matrices may result in a large cardinality than the approach that first orthogonalizes the matrices then restricts the base set. So we should take the former approach whenever the base set is known *a priori*.

The maximum cardinality of a restricted CEI-STC is specified in Theorem 6, but the restricted CEI-STC of the maximum cardinality is not unique. This is due to the nonuniqueness of the solution to the Hurwitz-Radon matrix equations. As shown in Example 1, the 2×2 restricted CEI-STC can be of any forms in (3.49). Each of them generates a restricted CEI-STCs of the same cardinality.

Just like the conventional TCM, a space-time constellation can be set-partitioned to achieve coded performance. Some work in this direction has been seen in [43, 44] where the O-STC was expanded before set partitioning. However the expanded set is no longer channel-eigenvector invariant. This is confirmed by computer simulation in [9].

Our original motivation was to investigate the structure of the universal codes which perform equally well whenever the channels provide enough mutual information. However except for a few cases, the CEI-STC is mutual-information lossy. Thus the divide-and-conquer method, first obtaining eigenvector invariance then obtaining eigenvalue spread invariance, does not seem to be promising. There is definitely a tradeoff between channel-eigenvector invariance and eigenvalue spread invariance. It might be wise to consider them jointly. Future work can be focused on trading eigenvector invariance for less mutual information loss.

CHAPTER 4

Trellis State-Diagram Reduction

4.1 Introduction

Viterbi's seminal work [45] introduced generating function techniques for computing a union bound on the bit error rate of convolutional codes. For a convolutional code with N states, an N-state error event diagram is sufficient to produce the correct generating function. However, general trellis codes may require up to N^2 states for the error event diagram because the error event distance depends on the transmitted sequence. The computation of the generating function of a trellis code may be separated into two stages. The first stage reduces the number of states as much as possible using low-complexity approaches. The second stage produces the generating function from the reduced-state diagram through some form of matrix inversion, which has a relatively high complexity.

Zehavi & Wolf [46], Rouanne & Costello [47], Liu, Oka & Biglieri [48] and Biglieri & McLane [49] derived some sufficient conditions (e.g. quasi-regularity) for the error event diagram to be independent of the transmitted sequence, which means only N states are necessary. But not all error diagrams can be independent of the transmitted sequence. Techniques working on generic trellis codes are particularly useful in code searches. By taking advantage of the symmetry of the constellations, Wesel [50] proposed a method that is able to reduce the number of states to some quantity between N and N^2 without resorting to quasi-regularity. But symmetry of constellation does not cover all the reducible cases. More generic approaches based on the FSM (finite-state-machine) theory were proposed by Schlegel [12], Kucukyavuz & Fitz [51] and Ryan [52]. However their FSM-theory-based reduction algorithms preserve not only the generating function but also the input-output relationship which is an extra constraint preventing further reduction. In this chapter we propose an improved reduction technique by eliminating this constraint.

4.2 Preliminaries

4.2.1 Product-State Diagrams

For a trellis code with N states, we define a product-state as the ordered pair (e_i, s_i), where s_i is the correct state of the encoder and e_i is the bitwise exclusive-or of the correct state and a (possibly wrong) receiver state. There are N^2 product-states in a product-state diagram.

Let $g(s_1, s_2)$ or $g(e_1, e_2)$ denote the binary label in response to the transition $s_1 \rightarrow s_2$ or $e_1 \rightarrow e_2$ respectively, and $f[g(s_1, s_2)]$ denote the corresponding constellation point. In a product-state diagram, an edge from product-state (e_1, s_1) to (e_2, s_2) exists if and only if both transitions $e_1 \rightarrow e_2$ and $s_1 \rightarrow s_2$ exist in the encoder state diagram. The edge $(e_1, s_1) \rightarrow (e_2, s_2)$ has the label

$$p(s_1 \rightarrow s_2 | s_1) \sum_n p_n I^{d_H(i_{12})} W^{d_E^2(f[g(s_1,s_2)],f[g(s_1,s_2) \oplus g(e_1,e_2)])}, \tag{4.1}$$

where $p(s_1 \rightarrow s_2 | s_1)$ is the conditional probability of transition $s_1 \rightarrow s_2$ given state s_1. The summation is over all parallel transitions. p_n is the probability of having the n^{th} transition among all possible parallel transitions. For trellis codes with no parallel transitions, the summation becomes a single nonzero term. The exponent of I is the Hamming weight of the binary input i_{12} that causes the transition $e_1 \rightarrow e_2$, i.e., the number of input bits in error. The exponent of W is the squared Euclidean distance between the transmitted and received constellation points.

4.2.2 The Generating Function

The generating function can be computed by adapting Biglieri's approach [53]. For simplicity, we define the index of the product-state (e_1, s_1) to be $e_1 \times N + s_1$. We call those product-states indexed from 0 to $N-1$ "correct states" and the remaining $(N^2 - N)$ product-states "erroneous states". Consider an $N^2 \times N^2$ transition matrix $\boldsymbol{G}$ whose entry G_{ij} is either the label of the transition from product-states j to i as shown in (4.1) or zero if no such transition exists. The matrix $\boldsymbol{G}$ has the form

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{d} & \boldsymbol{c} \\ \boldsymbol{b} & \boldsymbol{A} \end{bmatrix}, \tag{4.2}$$

where $\boldsymbol{d}$ is an $N \times N$ matrix for transitions among correct states. $\boldsymbol{c}$ is an $N \times (N^2 - N)$ matrix for transitions from an erroneous state to a correct state. $\boldsymbol{b}$ is an $(N^2 - N) \times N$ matrix for transitions from a correct state to an erroneous state. $\boldsymbol{A}$ is an $(N^2 - N) \times (N^2 - N)$ matrix for transitions between two erroneous states. The generating function is calculated as

$$T(W, I) = \boldsymbol{1dp} + \boldsymbol{1c(I - A)}^{-1}\boldsymbol{bp}, \tag{4.3}$$

where $\boldsymbol{1}$ is the $1 \times N$ row vector of ones, $\boldsymbol{p}$ is the $N \times 1$ column vector whose elements are all $1/N$, $\boldsymbol{I}$ is the $(N^2 - N) \times (N^2 - N)$ identity matrix. The commonly used upper bound on bit error probability is then given as [54]

$$P_e \leq \frac{1}{k} e^{d_{free}^2/4N_0} Q\left(\sqrt{d_{free}^2/2N_0}\right) \frac{\partial T(W, I)}{\partial I}|_{I=1, W=e^{-1/4N_0}}, \tag{4.4}$$

where

$$\begin{aligned}\frac{\partial T(W,I)}{\partial I} &= \mathbf{1}\boldsymbol{d}'\boldsymbol{p} + \mathbf{1}\boldsymbol{c}(\boldsymbol{I}-\boldsymbol{A})^{-\mathbf{1}}\boldsymbol{b}'\boldsymbol{p} \\ &+ \mathbf{1}\boldsymbol{c}'(\boldsymbol{I}-\boldsymbol{A})^{-\mathbf{1}}\boldsymbol{b}\boldsymbol{p} \\ &+ \mathbf{1}\boldsymbol{c}(\boldsymbol{I}-\boldsymbol{A})^{-\mathbf{1}}\boldsymbol{A}'(\boldsymbol{I}-\boldsymbol{A})^{-\mathbf{1}}\boldsymbol{b}\boldsymbol{p}.\end{aligned}$$

4.3 Finite State Machine Minimization

A product-state diagram can be regarded as the state diagram of an FSM which has correct states as both its sources and sinks. In the standard FSM theory, two FSM's are equivalent if they have the same input-output relationship [55]. In contrast we consider two product-state diagrams to be equivalent if they produce the same generating function. According to this definition, we have the following lemma.

Lemma 5 *An unreduced N^2-state diagram is equivalent to an (N^2-N+1)-state diagram by combining all correct states into one state.*

Proof: We construct a transition matrix with lower dimension and show that the new matrix has the same generating function as the original one. The unreduced transition matrix $\boldsymbol{G}$ is defined in (4.2). Define an $(N^2-N+1)\times(N^2-N+1)$ matrix $\hat{\boldsymbol{G}}$ of the form

$$\hat{\boldsymbol{G}} = \begin{bmatrix}\hat{\boldsymbol{d}} & \hat{\boldsymbol{c}} \\ \hat{\boldsymbol{b}} & \boldsymbol{A}\end{bmatrix},$$

where $\hat{d} = \mathbf{1}\boldsymbol{d}\boldsymbol{p}$, $\hat{\boldsymbol{c}} = \mathbf{1}\boldsymbol{c}$, $\hat{\boldsymbol{b}} = \boldsymbol{b}\boldsymbol{p}$. $\hat{\boldsymbol{G}}$ is the transition matrix of a reduced product-state diagram which has the same erroneous states as the original one but has only one correct state. Since the generating function of the reduced diagram

$$\begin{aligned}\hat{T}(W,I) &= \hat{d} + \hat{\boldsymbol{c}}(\boldsymbol{I}-\boldsymbol{A})^{-\mathbf{1}}\hat{\boldsymbol{b}} \\ &= \mathbf{1}\boldsymbol{dp} + \mathbf{1}\boldsymbol{c}(\boldsymbol{I}-\boldsymbol{A})^{-\mathbf{1}}\boldsymbol{bp} \\ &= T(W,I),\end{aligned}$$

which is the generating function of the original state diagram, the two diagrams are equivalent. ■

Note that $\hat{\boldsymbol{G}}$ is formed by first replacing leftmost N columns of $\boldsymbol{G}$ with their sum multiplied by $1/N$ and then replacing topmost N rows of $\boldsymbol{G}$ with their sum. The first column and row of $\hat{\boldsymbol{G}}$ provide the labels of all the edges connected with the new correct state.

Lemma 5 states that all correct states can always reduce to one state. The next lemma shows that further reductions are possible under certain conditions.

Lemma 6 *If there are two identical rows or columns in the transition matrix* $\boldsymbol{G}$*, the corresponding two states can reduce to one state.*

Similar statement for convolutional codes also appeared in [56]. A matrix with identical rows or columns was called row-uniform or column-uniform respectively in [48] and [52]. But row or column-uniform is defined on the error weight matrices which are sub-matrices of $\boldsymbol{G}$.

Proof: Since Lemma 5 states that the original N^2-state diagram is always equivalent to an reduced $(N^2 - N + 1)$-state diagram with one correct state, we only need to show Lemma 2 is valid for the reduced diagram whose transition matrix has the same form as shown in (4.2) except that $\boldsymbol{d}$ is now a scalar and $\boldsymbol{b}$ and $\boldsymbol{c}$ are vectors.

Let us first consider the identical column case. It is easy to verify that switching a pair of rows and the corresponding columns of $\boldsymbol{G}$ does not change the generating function. So without loss of generality, we assume the last two columns of $\boldsymbol{G}$ are identical. Define two operational matrices,

$$\boldsymbol{E}_1 = \begin{bmatrix} \boldsymbol{I}_{m-2} & & \\ & 1/2 & 0 \\ & 1/2 & 0 \end{bmatrix}, \boldsymbol{E}_2 = \begin{bmatrix} \boldsymbol{I}_{m-2} & & \\ & 1 & 1 \\ & 0 & 0 \end{bmatrix}$$

where $m = N^2 - N + 1$, $\boldsymbol{I}_{m-2}$ is the $(m-2) \times (m-2)$ identity matrix. Define

$$\begin{aligned} \hat{\boldsymbol{c}} = \boldsymbol{c}\boldsymbol{E}_1 = \begin{bmatrix} \tilde{\boldsymbol{c}} & 0 \end{bmatrix} \quad & \hat{\boldsymbol{b}} = \boldsymbol{E}_2\boldsymbol{b} = \begin{bmatrix} \tilde{\boldsymbol{b}} \\ 0 \end{bmatrix} \\ \hat{\boldsymbol{A}} = \boldsymbol{E}_2\boldsymbol{A}\boldsymbol{E}_1 = \begin{bmatrix} \tilde{\boldsymbol{A}} & \boldsymbol{0} \\ \boldsymbol{0} & 0 \end{bmatrix} \quad & \tilde{\boldsymbol{G}} = \begin{bmatrix} d & \tilde{\boldsymbol{c}} \\ \tilde{\boldsymbol{b}} & \tilde{\boldsymbol{A}} \end{bmatrix} \end{aligned} \tag{4.5}$$

where $\tilde{\boldsymbol{c}}$ and $\tilde{\boldsymbol{b}}$ are $1 \times (N^2 - N - 1)$ and $(N^2 - N - 1) \times 1$ vectors. $\tilde{\boldsymbol{A}}$ is a $(N^2 - N - 1) \times (N^2 - N - 1)$ matrix. $\tilde{\boldsymbol{G}}$ is the new transition matrix which has the same states as $\boldsymbol{G}$ except that the two states corresponding to the two identical columns reduce to one state. Now we need to show that $\tilde{\boldsymbol{G}}$ and $\boldsymbol{G}$ are equivalent, i.e., their generating functions are identical.

$$\begin{aligned} \tilde{T}(W, I) &= d + \tilde{\boldsymbol{c}}(\boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{b}} \\ &= d + \sum_{n=1}^{\infty} \tilde{\boldsymbol{c}}\tilde{\boldsymbol{A}}^n\tilde{\boldsymbol{b}} \\ &= d + \sum_{n=1}^{\infty} \boldsymbol{c}\boldsymbol{A}^n\boldsymbol{b} \\ &= d + \boldsymbol{c}(\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} \\ &= T(W, I) \end{aligned} \tag{4.6}$$

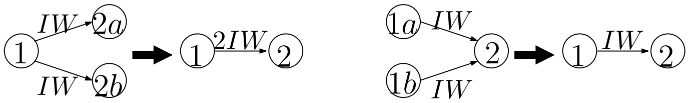
Eq.(4.6) is true because for any positive integer n, we have,

$$\begin{aligned}\tilde{\boldsymbol{c}}\tilde{\boldsymbol{A}}^{\boldsymbol{n}}\tilde{\boldsymbol{b}} &= \hat{\boldsymbol{c}}\hat{\boldsymbol{A}}^{\boldsymbol{n}}\hat{\boldsymbol{b}} \\ &= \boldsymbol{cE_1E_2AE_1E_2AE_1\ldots E_2AE_1E_2b} \\ &= \boldsymbol{c}\boldsymbol{A}^n\boldsymbol{b}\end{aligned} \tag{4.7}$$

where the last equality follows from the fact that $\boldsymbol{A}\boldsymbol{E}_1\boldsymbol{E}_2 = \boldsymbol{A}$ and $\boldsymbol{c}\boldsymbol{E}_1\boldsymbol{E}_2 = \boldsymbol{c}$ since the last two columns of $\boldsymbol{G}$ are identical.

Note that the generating functions for the transition matrix $\boldsymbol{G}$ and its transpose $\boldsymbol{G}^T$ are equal. So identical rows of $\boldsymbol{G}$ can be regarded as identical columns of $\boldsymbol{G}^T$. The argument follows. ■

As shown in the proof, the reduced transition matrix $\tilde{\boldsymbol{G}}$ is actually obtained by summing two rows and halving the sum of two columns in the identical column case. If we let $\boldsymbol{E}_1 = \boldsymbol{E}_2^T$ and $\boldsymbol{E}_2 = \boldsymbol{E}_1^T$, we will get a new matrix with the same generating function by summing two identical columns and halving the sum of the corresponding rows. So we have two ways to reduce the diagram. This is also true for the identical row case. In this chapter we adopt the convention illustrated in Fig. 4.1 where the incoming and outgoing edge labels correspond to the rows and columns of $\boldsymbol{G}$ respectively.



(a) Identical rows (b) Identical columns

Figure 4.1: Illustration of Lemma 6

Lemma 6 requires identical rows or columns for state reduction, which is a very strict constraint. However we are able to extend it by introducing the concept of equivalent states.

Definition 9 *State s_1 and s_2 are said to be forward (backward) equivalent if there exists a one-to-one mapping on their outgoing (incoming) edges such that*

1. *The corresponding edges have identical labels.*
2. *The succeeding (preceding) states associated with corresponding edges are forward (backward) equivalent.*

According to the definition, state $1a$ and $1b$ are forward equivalent in Fig. 4.1(b), and state $2a$ and $2b$ are backward equivalent in Fig. 4.1(a). The equivalence relation is usually defined as forward equivalence in the FSM theory where the input-output relationship needs to be preserved. However in our particular application since only the generating function is concerned, the backward equivalence is also a valid equivalence relation. Note that the definition above is recursive because the word "equivalent" also appears in the second condition. The recursive definition can be converted to an iterative one by introducing the concept "k-equivalent". We will use forward equivalence in our iterative definition of equivalence. There is an analogous iterative definition of backward equivalence.

Definition 10 *State s_1 and s_2 are said to be k-equivalent (where k is an integer) if there exists a one-to-one mapping on their outgoing edges such that*

1. *The corresponding edges have identical labels.*
2. *The succeeding states associated with corresponding edges are (k-1)-equivalent.*

The partition induced by k-equivalence is denoted as P_k. If the initial condition is given as "all states are 0-equivalent", we can find the partition P_k for any positive integer k by following Definition 10 iteratively. If two states are k-equivalent, they are also (k-1)-equivalent. So a subset in P_k is always contained in a subset in P_{k-1}, which means P_k is a finer partition than P_{k-1}. Consequently, the number of subsets in P_k is monotonically non-decreasing with regard to k. Since the number of subsets is upper bounded by the number of states, there exists a finite integer n which satisfies the equality $P_{n-1} = P_n$, the stopping condition. Once the stopping condition is met, two states in a subset of P_n have corresponding succeeding states that are both (n-1)-equivalent and n-equivalent. So n-equivalence satisfies the recursive condition in Definition 9. The above description implies the following algorithm.

Forward (Backward) Partitioning Algorithm

1. $k = 0$. The initial partition denoted as P_0 is a single set containing all the states.

2. $k = k + 1$. Refine the partition. Two states in a subset of P_{k-1} are in a subset of P_k if there exists a one-to-one mapping on their outgoing (incoming) edges such that the corresponding edges have the same labels and the succeeding (preceding) states associated with corresponding edges are in the same subset of P_{k-1}.

3. Go to step (2) until $P_k = P_{k-1}$. P_k is the final forward (backward) partition.

Since an equivalence class in the final partition can be represented by any of its members, reduction is possible whenever there is an equivalence class consisting of more than one state. The following theorem gives the upper bound on the number of product-states required for linear trellis codes, i.e., trelllis codes consisting of a linear convolutinal encoder and a (usually nonlinear) signal mapper.

Theorem 8 *For a linear trellis code, its product-state diagram requires at most* $\frac{N^2-N}{2} + 1$ *states.*

Proof: Group the N^2 product-states according to their error states such that those sharing the same error state are in the same group. There are N groups with N product-states each. Then partition each group by pairing up (e, s) with $(e, s \oplus e)$. We will show that the paired product-states are equivalent.

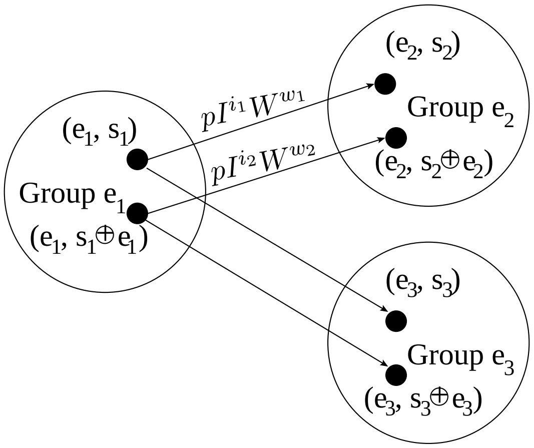


Figure 4.2: Paired product-states are equivalent

As shown in Fig. 4.2, consider arbitrary pair (e_1, s_1) and $(e_1, s_1 \oplus e_1)$ which belongs to group e_1. Assume state (e_1, s_1) has an outgoing edge arriving at state (e_2, s_2) with the label $pI^{i_1}W^{w_1}$, where p is a constant independent of the states. If the transition $e_1 \rightarrow e_2$ is caused by the input i_e and the transition $s_1 \rightarrow s_2$ is caused by the input i_s, then feeding $i_e \oplus i_s$ to the state $e_1 \oplus s_1$ causes the transition $e_1 \oplus s_1 \rightarrow s_2 \oplus e_2$. So there is always an edge from $(e_1, s_1 \oplus e_1)$ to $(e_2, s_2 \oplus e_2)$ with the label $pI^{i_2}W^{w_2}$. Since both edges start at error state e_1 and end at error state e_2, we have $i_1 = i_2$. As for w_1 and w_2, we have

$$\begin{aligned}
w_1 =& d_E^2(f[g(s_1, s_2)], f[g(s_1, s_2) \oplus g(e_1, e_2)]), \\
w_2 =& d_E^2(f[g(s_1 \oplus e_1, s_2 \oplus e_2)], \\
& f[g(s_1 \oplus e_1, s_2 \oplus e_2) \oplus g(e_1, e_2)]).
\end{aligned}$$

Because of the linearity of the constituent convolutional code we get

$$\begin{aligned} & g(s_1, s_2) \oplus g(e_1, e_2) = g(s_1 \oplus e_1, s_2 \oplus e_2), \\ & g(s_1 \oplus e_1, s_2 \oplus e_2) \oplus g(e_1, e_2) \\ & = g(s_1, s_2) \oplus g(e_1, e_2) \oplus g(e_1, e_2) = g(s_1, s_2). \end{aligned}$$

So $w_1 = w_2$. Condition 1 of Definition 1 holds now. Notice that the ending states (e_2, s_2) and $(e_2, s_2 \oplus e_2)$ are paired too. The second condition is also satisfied. The same is true for other outgoing edges. Hence the paired states are equivalent. There are $N^2/2$ such pairs in the product-state diagram. Combining this result with Lemma 5, the product-state diagram requires at most $\frac{N^2-N}{2} + 1$ states. ■

Now we are ready for the iterative FSM minimization algorithm.

Iterative FSM Minimization Algorithm

1. Form the product-state diagram and denote it by SD.

2. Apply forward partitioning algorithm to SD. Reduce the diagram by combining states in the same equivalence classes in the final partition into one state. The new label is obtained by summing over the labels of the edges sharing the same starting and ending equivalence classes and then dividing the sum by the number of members in the starting equivalence class. Denote the reduced diagram as SD_1.

3. Apply backward partitioning algorithm to SD_1. Reduce the diagram in the same way as in step (2). Denote the reduced diagram as SD_2.

4. $SD = SD_2$. Go to (2) until neither step (2) nor (3) is able to reduce the diagram. SD is the final reduced diagram.

Questions may arise like: Since according to the FSM theory the forward partitioning alone gives the minimal equivalent diagram, why could the backward equivalence provide further reduction? The answer lies in the definition of equivalence. In the FSM theory, two states are equivalent if they generate the same output sequence for every possible input sequence. That is, there is no difference between them by only looking at the input and output sequences. However, when it comes to state diagram reduction, two states are equivalent if they contribute identically to the generating function. Having same input-output relationship certainly implies contributing equally to the generating function, but not vice versa.

For instance, consider two states in a diagram. Assume one of them generates an output sequence $(a_1,\ a_2,\ a_3,\ a_4)$ then ends at the sink node, while the other generates sequence $(a_4,\ a_3,\ a_2,\ a_1)$ responding to the same input then ends at the same sink node. For this input sequence, both of them contribute $\prod_{i=1}^{4} a_i$ to the generating function. These states could be equivalent if that is their entire contribution to the generating function.

Furthermore, two states can reduce to one state even though they generate totally different output symbols. This is possible because two states could be backward equivalent. We will show an example from a real code later in this chapter.

4.4 Computing the Generating Function

The reduction in the previous section does not typically produce a one-state diagram. Usually, we have to compute $\boldsymbol{c}(\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}$ as shown in (4.3) (but for a reduced diagram) to obtain the generating function. We only consider the case where $\boldsymbol{b}$ and $\boldsymbol{c}$ are vectors because all correct states can be combined into one state. This computation involves matrix inversion which is generally an $O(n^3)$ process. [1] In the rest of this section, we will study matrix inversion from two perspectives.

4.4.1 Gaussian Elimination

Matrix inverse in (4.3) can be computed by standard Gaussian elimination [58]. To be more specific, we first form the matrix

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{I} - \boldsymbol{A} & \boldsymbol{b} \\ -\boldsymbol{c} & \boldsymbol{0} \end{bmatrix}. \tag{4.8}$$

Then apply Gaussian elimination to $\boldsymbol{M}$ until $\boldsymbol{c}$ becomes a zero vector. Now we have

$$\boldsymbol{M'} = \begin{bmatrix} \boldsymbol{A'} & \boldsymbol{b'} \\ \boldsymbol{0} & \boldsymbol{d'} \end{bmatrix}. \tag{4.9}$$

Note that this is possible only if $\boldsymbol{I} - \boldsymbol{A}$ is non-singular. $\boldsymbol{d'}$ is exactly what we want.

$$\boldsymbol{d'} = \boldsymbol{c}(\boldsymbol{I} - \boldsymbol{A})^{-\mathbf{1}}\boldsymbol{b}. \tag{4.10}$$

4.4.2 Node Elimination for the Signal-flow Graph

The product-state diagram is also a signal-flow graph [59], where the label of an edge is called transmittance and the matrix $\boldsymbol{G}$ in (4.2) is called connection matrix of the graph. Computing $\boldsymbol{c}(\boldsymbol{I} - \boldsymbol{A})^{-\mathbf{1}}\boldsymbol{b}$ amounts to "solving" the graph, i.e., successively eliminating the nodes of the graph until only source nodes and

[1] Theoretically there exist $O(n^{2.496})$ matrix inversion algorithms [57], however they are quite involved and do not provide insight into this problem. Our purpose is to show that matrix inversion is more complex than the FSM minimization. This is also valid for those $O(n^{2.496})$ algorithms.

sink nodes remain. In the case of computing the generating functions, we need to eliminate all the erroneous states until the graph reduces to one state with a self loop whose transmittance is the generating function.

Elimination of a node must be performed in a manner such that the transmittances between remaining nodes do not change. The transmittances lost due to node elimination need to be transferred to other edges. Node elimination is equivalent to proper operations on the connection matrix $\boldsymbol{G}$. Suppose the connection matrix before the elimination has the following form.

$$\boldsymbol{G} = \begin{bmatrix} g_{0,0} & g_{0,1} & \cdots & g_{0,n-2} & g_{0,n-1} \\ g_{1,0} & g_{1,1} & \cdots & g_{1,n-2} & g_{1,n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ g_{n-2,0} & g_{n-2,1} & \cdots & g_{n-2,n-2} & g_{n-2,n-1} \\ g_{n-1,0} & g_{n-1,1} & \cdots & g_{n-1,n-2} & g_{n-1,n-1} \end{bmatrix}. \tag{4.11}$$

Now we eliminate the node corresponding to the last row and the last column of $\boldsymbol{G}$. Assume this node does not have any self loop, i.e. $g_{n-1,n-1} = 0$. Node elimination is achieved by disposing of the last column and row of $\boldsymbol{G}$ and modifying the remaining entries. The reduced connection matrix $\boldsymbol{G}_1$ is shown in (4.12).

$$\boldsymbol{G}_1 = \begin{bmatrix} g_{0,0} + g_{n-1,0}g_{0,n-1} & g_{0,1} + g_{n-1,1}g_{0,n-1} & \cdots & g_{0,n-2} + g_{n-1,n-2}g_{0,n-1} \\ g_{1,0} + g_{n-1,0}g_{1,n-1} & g_{1,1} + g_{n-1,1}g_{1,n-1} & \cdots & g_{1,n-2} + g_{n-1,n-2}g_{n-2,n-1} \\ \vdots & \vdots & & \vdots \\ g_{n-2,0} + g_{n-1,0}g_{n-2,n-1} & g_{n-2,1} + g_{n-1,1}g_{n-2,n-1} & \cdots & g_{n-2,n-2} + g_{n-1,n-2}g_{n-2,n-1} \end{bmatrix} \tag{4.12}$$

If the node has a self loop, we should divide the last column of $\boldsymbol{G}$ by $1 - g_{n-1,n-1}$ before applying the elimination described above. Continue the elimination until only the sink and source nodes (which happen to be the same node) remain. At this point $\boldsymbol{G}$ becomes a scalar which is equal to the generating function we get from Gaussian elimination. This approach is detailed in [60].

4.5 Complexity Analysis

Once the first stage has reduced the state diagram as much as possible, the complexity of computing $\boldsymbol{c}(\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}$ is clear. We will discuss its complexity first. Then we will give an upper bound on the complexity of the first stage.

4.5.1 Gaussian Elimination

The complexity analysis of Gaussian elimination is rather straightforward. We count the number of multiplications and additions in computing $\boldsymbol{c}(\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}$. Assume the matrix $\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{c}$ in (4.8) have the dimension $(n-1)\times(n-1)$, $(n-1)\times 1$, $1\times(n-1)$, respectively. When eliminating the i^{th} column, we need $(n-i+1)(n-i)$ multiplications and $(n-i)^2$ additions. Let C_{Mm}, C_{Ma}be the total number of multiplications and additions respectively. They are determined by the following equations.

$$C_{Mm}=\sum_{i=1}^{n-1} i^2+\sum_{i=1}^{n-1} i=\frac{n(n-1)(2n-1)}{6}+\frac{n(n-1)}{2}. \tag{4.13}$$

$$C_{Ma}=\sum_{i=1}^{n-1} i^2=\frac{n(n-1)(2n-1)}{6}. \tag{4.14}$$

4.5.2 Signal-Flow Graph Node Elimination

As described in the previous section, the complexity varies upon whether the node to be eliminated has a self loop or not. If it does, extra multiplications are needed. When we eliminate the i^{th} node, $(i-1)^2+M_i$ multiplications and $(i-1)^2$ additions are needed, where M_i is zero if the node does not have a self loop or $n-i$ if it has a self loop. Denote C_{Gm} and C_{Ga} as the total number of multiplications and additions respectively. we have

$$C_{Gm} = \sum_{i=1}^{n-1} i^2 + \sum_{i=1}^{n-1} M_i = \frac{n(n-1)(2n-1)}{6} + \sum_{i=1}^{n-1} M_i. \tag{4.15}$$

$$C_{Ga} = \sum_{i=1}^{n-1} i^2 = \frac{n(n-1)(2n-1)}{6}. \tag{4.16}$$

The second term in (4.15) could be $\frac{n(n-1)}{2}$ in the worst case where all nodes have self loops or zero in the best case where none of the nodes have a self loop. In reality, the term takes some value between these extreme cases. Note that C_{Ma} always equals to C_{Ga}. However C_{Gm} is upper bounded by C_{Mm}. Although both of them are $O(n^3)$ processes, the signal-flow graph node elimination is slightly less complex than Gaussian elimination.

4.5.3 The FSM Minimization

Unlike Gaussian elimination and the signal-flow graph node elimination, the FSM minimization can not often reduce a diagram to a single state. The number of states resulting from the FSM minimization is actually unknown until the minimization process completes. So its complexity is unpredictable and varies from case to case. However, we are still able to find an upper bound.

Consider a minimization process starting at n states and ending at m states, where $n \geq m$. According to the description in Section III, the reduction consists of at most $n - m$ partitionings. The worst case is that each partitioning creates one singleton set. That is, after the first partitioning, there are two sets, one singleton set and one set with $n - 1$ elements. This partitioning needs at most $2n - 3$ comparisons. The second partitioning results in two singleton sets and a set with $n - 2$ elements. At most $2n - 5$ comparisons are needed at this step. Continue in this fashion until in the end there are $m - 1$ singleton sets and one set with $n-m+1$ elements. Note comparisons are merely additions. Denote the total number of additions as $C_{Fa}(m)$, which is a function of m, the final size of the diagram. $C_{Fa}(m)$ is given as

$$C_{Fa}(m) = 2^{2k}\left[\sum_{i=n-m}^{n-1} 2i-(n-1)\right] = 2^{2k}[-m^2+(2n-1)m-(n-1)], m = 1, 2, \dots n. \tag{4.17}$$

where 2^{2k} is the number of outgoing edges from each states. Now we compare it with the signal-flow graph node elimination. Define $C_{Ga}(m)$ to be the number of additions needed for the same reduction if the node elimination is used. The analysis in the previous section shows

$$C_{Ga}(m) = \sum_{i=m}^{n-1} i^2, m = 1, 2, \dots n. \tag{4.18}$$

We examine two extreme cases where $m = 1$ and $m = n$. Evaluate $C_{Fa}(m)$ and $C_{Ga}(m)$ at these values, we have

$$C_{Fa}(1) = 2^{2k}(n-1) \qquad C_{Ga}(1) = \frac{n(n-1)(2n-1)}{6}, \tag{4.19}$$

$$C_{Fa}(n) = 2^{2k}(n-1)^2 \qquad C_{Ga}(n) = 0. \tag{4.20}$$

On one hand, $C_{Ga}(1) > C_{Fa}(1)$ if $2^{2k} < \frac{n(2n-1)}{6(n-1)}$. That is, if both techniques can reduce the diagram to one state, the FSM minimization has fewer additions for small k. On the other hand, $C_{Ga}(n) < C_{Fa}(n)$. If the diagram is so asymmetric that going through the FSM technique does not reduce the diagram at all, the node elimination technique is better because nothing is gained by the FSM computation. We will show that there exists a critical point t between 1 and n where $C_{Ga}(t) = C_{Fa}(t)$. On one side of this point, the FSM technique has less complexity. On the other side of the point, the opposite is true.

Take the first derivative of $C_{Fa}(m)$ and $C_{Ga}(m)$ with regard to m, we get

$$C'_{Fa}(m) = 2^{2k}(2n-2m-1), \tag{4.21}$$

$$C'_{Ga}(m) = -m^2 + m + 1/6. \tag{4.22}$$

For $1 \leq m < n$, $C'_{Fa}(m) > 0$ but $C'_{Ga}(m) < 0$. So $C_{Fa}(m)$ is monotonically increasing and $C_{Ga}(m)$ is monotonically decreasing. Given the boundary conditions at $m = 1$ and $m = n$, there is one and only one point where two curves meet. Fig. 4.3 shows $C_{Fa}(m)$ and $C_{Ga}(m)$ for $n = 256$ and different k's. Because commonly used constellations have at least one axis of symmetry [50], the FSM minimization usually brings the size of the product-state diagram below the critical point. In addition, the FSM minimization does not involve any multiplications. Thus even our loose bound of FSM complexity indicates that the overall complexity of the FSM technique is lower than that of the matrix inversion techniques to provide the same degree of graph simplification.

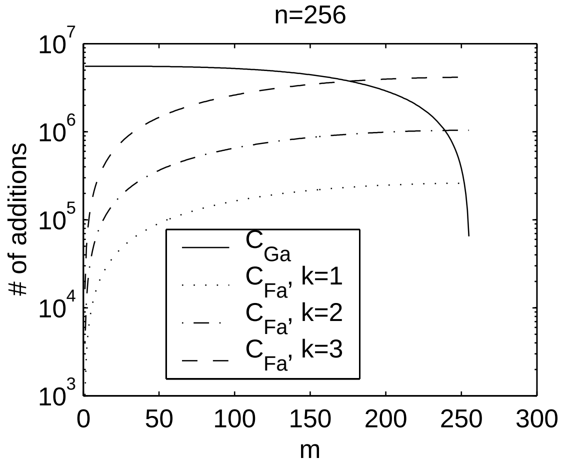


Figure 4.3: Complexity comparisons

As for the iterative FSM reduction, it is hard to have a bound because the number of iterations also varies from code to code. But we observe that iteration does not increase computational burden dramatically for two reasons. First, the number of iterations is usually quite small as shown in Table 4.1 where we list the results from exhaustive search over the rate 1/3 codes of the form

$\begin{bmatrix} g_0(D) & g_1(D) & g_2(D) \end{bmatrix}$ with up to 4 memory elements ($N = 16$). Second, the number of comparisons at each iteration is diminishing because the size of the diagram is monotonically decreasing.

Table 4.1: Number of iteration for rate 1/3 codes

Number of memory elements		Maximum number of iterations	Average number of iterations
2	8PSK	4	2.57
	8PAM	4	2.73
3	8PSK	4	2.73
	8PAM	4	3.37
4	8PSK	4	2.83
	8PAM	6	3.70

4.6 The Algorithm

Having analyzed the complexity of alternatives for each of the two stages, we are in a position to propose an efficient algorithm for computing the generating function. Due to the monotonicity of $C_{Fa}(m)$, the FSM minimization is upper bounded by $O(2^{2k}n^2)$.[2] It has lower complexity than Gaussian elimination or the signal-flow graph node elimination. We select it as the first step in reducing a product-state diagram. As for the second step, both the signal-flow graph node elimination and Gaussian elimination have complexity of $O(n^3)$. We choose

[2] The existence of an $O(2^{2k}nlogn)$ minimization algorithm [61] confirms our claim that the FSM minimization should be employed in the first place.

the former not only because it has slightly less complexity but also because it embodies the idea of state diagram reduction.

Our algorithm goes as follows. For a trellis code with N encoder states, if it has a linear convolutional encoder, form the $\frac{N^2-N}{2}+1$ product-state diagram according to Theorem 1. Otherwise form the N^2-N+1 product-state diagram according to Lemma 5. Then employ the iterative FSM minimization technique to reduce the diagram. Finally use the signal-flow graph node elimination to further reduce the diagram to one state. The label of the self loop in the one-state diagram is the generating function.

4.7 Examples and Discussion

4.7.1 Quasi-regular Trellis Codes

A trellis code is quasi-regular [47] if the code is composed of a linear binary encoder and a signal mapper, and the the quantity

$$W(s,e_o)=\sum_{x_s} p(x_s|s)W^{d_E^2(x_s,e_o)} \tag{4.23}$$

is constant for all encoder state s for a fixed binary error output e_o, where x_s is any possible output from state s. Since quasi-regularity implies that all the product state (e,s) sharing the same error state e are forward equivalent, the FSM minimization reduces the product-state diagram of a quasi-regular trellis code to at most N states.

4.7.2 Degraded Trellis Codes

We define degraded trellis codes to be trellis codes whose product-state diagram can be reduced to less than N states by the iterative FSM minimization.

Although any product-state diagram can be reduced to less than N states by signal flow graph elimination or Gaussian elimination, a degraded trellis code permits this reduction with a low-complexity stage-one algorithm. The example we present here is a rate 1/3 trellis code $\begin{bmatrix} D^2+1 & D & 1 \end{bmatrix}$ with natural labeling 8-PSK. As shown in Fig. 4.4, the reduced diagram has only 3 states. Similar observations were mentioned in [51] and [52]. Note that this is not achievable with quasi-regularity [46, 47, 48], Wesel's technique [50], or by the row and column uniform properties [49].

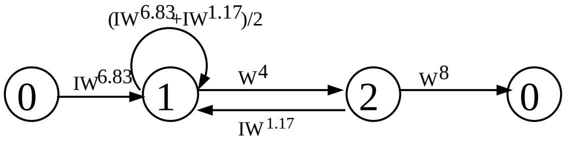
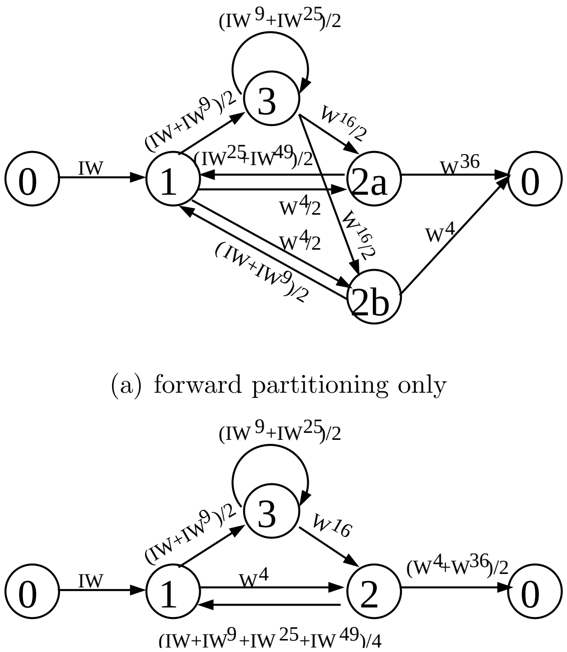


Figure 4.4: A degraded trellis code

4.7.3 Backward Equivalence

Section III shows that backward equivalence is theoretically able to further reduce the diagram. Here we give an example from a real code. Consider another rate 1/3 trellis code $\begin{bmatrix} D^2 & D^2+D & 1 \end{bmatrix}$ with natural labeling 8-PAM. Note that this trellis code is not row-uniform, column-uniform, or quasi-regular. The existing FSM theory based techniques presented in [51] or [52] are able to reduce it to 5 states as shown in Fig. 4.5(a). Close examination shows that although state $2a$ and $2b$ have different outgoing labels they are actually backward equivalent. Backward partitioning reduces them to one state as shown in Fig. 4.5(b).



(a) forward partitioning only

(b) both forward and backward partitioning

Figure 4.5: Backward equivalence in a real code

4.7.4 The Union Bound

We can compute the union bound using (4.4). The process is similar to computing the generating function except that we will encounter a slightly complicated term $\boldsymbol{c}(\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{A}'$, where $\boldsymbol{A}'$ is a matrix instead of a vector. Extension to the node elimination algorithm is needed. Assume $\boldsymbol{A}$ and $\boldsymbol{A}'$ are both $(n-1)\times(n-1)$ matrix, $\boldsymbol{c}$ is a $1\times(n-1)$ vector, Define an $n\times 2(n-1)$ matrix as follows.

$$\boldsymbol{G}=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{c}\\ \boldsymbol{A}' & \boldsymbol{A}\end{bmatrix}. \tag{4.24}$$

$\boldsymbol{G}$ is the connection matrix of a signal-flow graph with $n-1$ source nodes and one sink node. The technique described in Section IV is used to eliminate the intermediate nodes until $\boldsymbol{G}$ becomes a $1\times(n-1)$ vector which is equal to $\boldsymbol{c}(\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{A}'$.

Consider the rate 1/4 trellis code [117 155 145 137] with Gray labeling 16-QAM. This code is given as an irreducible example in [50]. The numbers of reduced states obtained with different techniques are listed in Table 4.2. The iterative FSM minimization reduces the number of states significantly.

Table 4.2: Reduction in the first stage

Variables or Techniques	Number of states
N	64
N^2	4096
$(N^2-N)/2+1$	2017
Wesel [50]	4096
Kucukyavuz & Fitz [51]	1240
The standard FSM	653
The iterative FSM	527

The number of comparisons for the standard FSM minimization is approximately 3.0×10^5. For the iterative FSM minimization approximately 3.4×10^5 comparisons are needed. The upper bound from (4.17) for each case is 9.8×10^6 and 8.0×10^6 respectively. This verifies that the iterative process only increases complexity slightly. It also shows that the upper bound is very loose.

The CPU time required to compute the union bound at a fixed SNR using a 1.2 GHz Pentium 4 computer is reported in Table 4.3, where 'FSM', 'IFSM', 'GE', 'SFG' stand for the FSM minimization, the iterative FSM minimization, Gaussian elimination, the signal-flow graph node elimination respectively. According to the table the iterative FSM minimization followed by node elimination is the most efficient algorithm. The node elimination is slightly faster than Gaussian elimination as we predicted in the complexity analysis.

Table 4.3: Comparison of different techniques

Techniques	CPU time (seconds)
Direct GE	3510
Direct SFG	3261
FSM+GE	15.91
FSM+SFG	14.48
IFSM+GE	8.2
IFSM+SFG	7.54

4.8 Summary

We have proposed an efficient approach to compute the generating function for trellis codes. The proposed approach consists of two stages. A low-complexity iterative FSM minimization is performed in the first stage and a signal-flow graph node elimination algorithm is employed in the second stage. We have also presented detailed comparisons between alternatives for each of the two stages to show the efficiency of our techniques.

For trellis codes comprised of a linear convolutional encoder and a signal mapper, we have derived an upper bound on the size of the product-state diagram to be $\frac{N^2-N}{2}+1$. This bound halves the size of an N^2-product-state diagram without introducing any computational cost.

CHAPTER 5

Conclusion

5.1 Summary

Chapter 2 reviewed the compound channel coding theorem and formally introduced the universal codes. The performance of universal codes was evaluated against the sphere packing bound and the random coding bound as the block-lengths varied. The universality is an asymptotic behavior, which is a practical goal for long codes such as turbo or LDPC codes.

Chapter 3 applied the universal codes to MIMO systems. A class of space-time constellations with the channel-eigenvector-invariance property were shown to be an extension of the orthogonal designs. This class of constellations were however not universal due to varying mutual information loss under different MIMO channels. Both constellation and channel coding should be used to achieve universality.

Chapter 4 presented a trellis state diagram reduction algorithm based on the FSM theory. Compared to existing methods, the proposed method was applicable to more generic codes and had lower complexity. The complexity of matrix inversion was dominant in the union bound computation. The proposed method should be used as the first step to reduce the size of the state transition matrix.

5.2 Future Work

The channel-eigenvector-invariant space-time constellations found in Chapter 3 has the drawback of mutual information loss. The trade-off between eigenvector invariance and eigenvalue invariance has been observed. It would be interesting to examine the exact trade-off mathematically. Also, how to achieve the trade-off by designing proper channel codes and constellations can be a subject for future study.

CHAPTER 6

Appendix

6.1 Chapter 2

6.1.1 Proof of Theorem 1

Theorem 1 *A family of Gaussian multiple-input multiple-output (MIMO) channels is denoted as $\{\boldsymbol{H}_i, \boldsymbol{K}_i, i \in \mathcal{I}\}$, where the index set $\mathcal{I}$ is an arbitrary set. The channel input output is governed by $\boldsymbol{y} = \boldsymbol{H}_i\boldsymbol{x} + \boldsymbol{z}_i$, where $\boldsymbol{H}_i$ is a $p \times q$ matrix and $\boldsymbol{z}_i$, the noise, is a Gaussian random vector of dimension $p \times 1$ with zero mean and the covariance matrix $\boldsymbol{K}_i$. Assume there exist real numbers a, α_1, α_2 such that for each $i \in \mathcal{I}$,*

1. *$\|\boldsymbol{H}_i\|_2 \leq a$, where $\| \cdot \|_2$ is the spectral norm of a matrix, i.e. the square root of the largest eigenvalue of $\boldsymbol{H}_i^*\boldsymbol{H}_i$.*

2. *$\alpha_1 \leq \frac{\boldsymbol{x}^*\boldsymbol{K}_i\boldsymbol{x}}{\|\boldsymbol{x}_i\|^2} \leq \alpha_2$, for any $p \times 1$ vector $\boldsymbol{x}$.*

Then any rate $R < C(\mathcal{I})$ defined in (2.2) is achievable. i.e. there exists a sequence of $(2^{nR}, n)$ codes such that the probability of error under any channel in the family approaches zero as the block-length approaches infinity. ■

Proof: We start with the case $p = q$. The result will be generalized to $p \neq q$ afterwards. Following [7], we divide the proof into two steps. We will show the theorem is true when the set of the channels is finite, then extend the result to infinitely many channels by bounding the error probability of an arbitrary channel.

Definition 11 *[17] The set $A_\epsilon^{(n)}$ of jointly typical sequence $\boldsymbol{x}^{(n)}, \boldsymbol{y}^{(n)}$ with respect to the distribution $p(x, y)$ is the set of n-sequences with empirical entropies ϵ-close to the true entropies, i.e.,*

$$\begin{aligned} A_\epsilon^{(n)} = \{ & \, (\boldsymbol{x}^{(n)}, \boldsymbol{y}^{(n)}) \in \mathcal{X}^{(n)} \times \mathcal{Y}^{(n)} : \\ & \left| -\frac{1}{n} \log p(\boldsymbol{x}^{(n)}) - H(\boldsymbol{X}) \right| < \epsilon, \\ & \left| -\frac{1}{n} \log p(\boldsymbol{y}^{(n)}) - H(\boldsymbol{Y}) \right| < \epsilon, \\ & \left| -\frac{1}{n} \log p(\boldsymbol{x}^{(n)}, \boldsymbol{y}^{(n)}) - H(\boldsymbol{X}, \boldsymbol{Y}) \right| < \epsilon, \} \end{aligned} \tag{6.1}$$

where

$$p(\boldsymbol{x}^{(n)}, \boldsymbol{y}^{(n)}) = \prod_{i=1}^{n} p(\boldsymbol{x}_i, \boldsymbol{y}_j). \tag{6.2}$$

For the MIMO gaussian channel in Theorem 1, the probability of the typical set can be bounded as follows.

Lemma 7 *Let $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{z}$, where $\boldsymbol{H}$ is a deterministic matrix of dimension $p \times p$ and $\boldsymbol{x}$ and $\boldsymbol{z}$ are independent Gaussian random vectors of dimension $p \times 1$ with zero mean and covariance matrices $\boldsymbol{S}$ and $\boldsymbol{K}$, respectively. Consider i.i.d. drawn length-n sequence $(\boldsymbol{x}^{(n)}, \boldsymbol{y}^{(n)})$ of dimension $2pn \times 1$,*

$$Pr\big\{(\boldsymbol{x}^{(n)}, \boldsymbol{y}^{(n)}) \notin A_\epsilon^{(n)}\big\} < 6e^{-\frac{n\epsilon^2}{4p}}. \tag{6.3}$$

Proof: We bound the probability when the first inequality in (6.1) is violated.

$$\begin{aligned}
&\Pr\left\{-\log p(\boldsymbol{x}^{(n)}) < nH(\boldsymbol{X}) - n\epsilon\right\} \\
&\overset{(a)}{=} \Pr\left\{\frac{np}{2} - n\epsilon - \boldsymbol{x}^{(n)\mathsf{T}}\boldsymbol{K}^{(n)^{-1}}\boldsymbol{x}^{(n)} \geq 0\right\} \\
&\overset{(b)}{\leq} \left(e^{t(\frac{p}{2}-\epsilon)}\mathsf{E}e^{-\frac{t}{2}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{K}^{-1}\boldsymbol{x}}\right)^{n} \\
&= \left(e^{\frac{pt}{2}-\epsilon t-\frac{p}{2}\log(1+t)}\right)^{n} \\
&\overset{(c)}{\leq} \left(e^{\frac{pt^2}{4}-\epsilon t}\right)^{n} \\
&\overset{(d)}{=} \left(e^{-\frac{\epsilon^2}{p}}\right)^{n},
\end{aligned} \tag{6.4}$$

where $\boldsymbol{K}^{(n)}$ in (a) is a block diagonal matrix of dimension $np \times np$ with $\boldsymbol{K}$ along the diagonal. (b) is due to the Chernoff bound and the fact $\boldsymbol{x}^{(n)}$ is i.i.d. drawn. t is an arbitrary positive number and E stands for expectation. (c) is because $t - 1/2\log(1+t) \leq t^2/2$ for $0 \leq t < 1$. We substitue $t = 2\epsilon/p < 1$ in (c) and arrive at (d).

The other direction goes similarly.

$$\begin{aligned}
&\Pr\left\{-\log p(\boldsymbol{x}^{(n)}) > nH(\boldsymbol{X}) + n\epsilon\right\} \\
&= \mathrm{P}\left\{-\frac{np}{2} - n\epsilon + \boldsymbol{x}^{(n)\mathsf{T}}\boldsymbol{K}^{(n)^{-1}}\boldsymbol{x}^{(n)} \geq 0\right\} \\
&\leq \left(e^{-t(\frac{p}{2}+\epsilon)}\mathsf{E}e^{\frac{t}{2}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{K}^{-1}\boldsymbol{x}}\right)^{n} \\
&= \left(e^{-\frac{pt}{2}-\epsilon t-\frac{p}{2}\log(1-t)}\right)^{n} \\
&\overset{(a)}{\leq} \left(e^{\frac{pt^2}{2}-\epsilon t}\right)^{n} \\
&\overset{(b)}{=} \left(e^{-\frac{\epsilon^2}{2p}}\right)^{n},
\end{aligned} \tag{6.5}$$

where (a) is because $t + \log(1-t) \geq -t^2$ for $0 \leq t \leq 0.5$. We obtain (d) by substituting $t = \epsilon/p$ in (a). Combine (6.4) and (6.5) we get

$$\Pr\left\{\left|-\frac{1}{n}\log p(\boldsymbol{x}^{(n)}) - H(\boldsymbol{X})\right| > \epsilon\right\} < 2e^{-\frac{n\epsilon^2}{2p}}. \tag{6.6}$$

Similarly, we can prove

$$\Pr\left\{\left|-\frac{1}{n}\log p(\boldsymbol{y}^{(n)}) - H(\boldsymbol{Y})\right| > \epsilon\right\} < 2e^{-\frac{n\epsilon^2}{2p}}, \tag{6.7}$$

$$\Pr\left\{\left|-\frac{1}{n}\log p(\boldsymbol{x}^{(n)}, \boldsymbol{y}^{(n)}) - H(\boldsymbol{X}, \boldsymbol{Y})\right| > \epsilon\right\} < 2e^{-\frac{n\epsilon^2}{4p}}. \tag{6.8}$$

Finally, we arrive at (6.3) by the union bound. ■

The following lemma gives bounds on the power of input and output of the channel.

Lemma 8 *With the same setup in Lemma 7, denote the total power of vector $\boldsymbol{x}^{(n)}$ by $P_x = Tr(\boldsymbol{S})$. Then for any $\epsilon > 0$,*

$$Pr\left\{\| \boldsymbol{x}^{(n)} \|^2 \geq n(P_x + \epsilon)\right\} \leq e^{-c_1 n}, \tag{6.9}$$

where $c_1 = \frac{1}{2}[\frac{\epsilon}{2P} - \log(1 + \frac{\epsilon}{P})]$. Furthermore, if the channel satisfies the conditions in Theorem 1, for any input complying with the power constraint $\| \boldsymbol{x}^{(n)} \|^2 \leq nP_x$, the output satisfies,

$$Pr\left\{\| \boldsymbol{y}^{(n)} \|^2 \geq nP_y \mid \boldsymbol{x}^{(n)}\right\} \leq e^{-c_2 n}, \tag{6.10}$$

where $P_y = 2a^2 P_x + 2p\alpha_1 + 2$ and $c_2 = \frac{1}{p\alpha_1} - \log(1 + \frac{1}{p\alpha_0})$.

Proof: This is a direct result of the Chernoff bound. See Lemma 5,8 of [7]. ■

In the following lemma, we prove the existence of universal codes for a finite set of channels.

Lemma 9 *With the same setup up as Theorem 1, denote $L = |\mathcal{I}|$, the cardinality of the index set $\mathcal{I}$. Assume L to be finite, any rate $R < C(\mathcal{I})$ is achievable.*

Proof: We will use the same ideas in [17], namely random codes and joint typical decoding. We generate i.i.d. codewords according to the distribution of $\boldsymbol{x}$. The codewords are denoted by $\boldsymbol{x}^{(n)}(w), w = 1, 2, ..., 2^{nR}$. The receiver looks for codewords that are jointly typical with the received vector. If a single codeword is found, it is declared to be the transmitted codeword. Otherwise an error is declared. The receiver also declares an error if the chosen codeword does not satisfy the power constraint. Without loss of generality, assume that codeword 1 is sent.

Define the following events:

$$\mathrm{E}_0 = \left\{\| \boldsymbol{x}^{(n)}(1) \|^2 > P_x\right\}, \tag{6.11}$$

$$\mathrm{E}_i = \left\{(\boldsymbol{x}^{(n)}(i), \boldsymbol{y}^{(n)}) \in A_\epsilon^{(n)})\right\}. \tag{6.12}$$

Let P_s^n be the average of the sum of error probabilities under individual channels when codeword 1 is sent.

$$\begin{aligned}
P_s^n &= \sum_{i=1}^{L} P_i \left\{\mathcal{E}|W = 1\right\} \\
&\overset{(a)}{\leq} \sum_{i=1}^{L} \left(P_i(\mathrm{E}_0) + P_i(\mathrm{E}_1^c) + \sum_{j=2}^{2^{nR}} P_i(\mathrm{E}_j) \right) \\
&\overset{(b)}{\leq} \sum_{i=1}^{L} \left(e^{-c_1 n} + 6e^{-\frac{n\epsilon^2}{4p}} + (\log 2) e^{-n(I_i(x;y)-R)} e^{6n\epsilon} \right) \\
&\overset{(c)}{\leq} L \left(e^{-c_1 n} + 6e^{-\frac{n\epsilon^2}{4p}} + (\log 2) e^{-n\epsilon} \right) \longrightarrow 0, \text{as } n \longrightarrow \infty.
\end{aligned} \tag{6.13}$$

where (a) is due to the union bound, $P_i(\cdot)$ is the probability of an event under the ith channel. (b) is due to Lemma 7 and 8. (c) follows from the fact $R < C(\mathcal{I}) \leq I_i(x, y), \forall i \in \mathcal{I}$ and n sufficiently large. Thus, for individual channel, the error probability also approaches zero. By deleting the worst half of the codewords we obtain a code with low maximal probability of error. ■

To extend the result to arbitrary set $\mathcal{I}$, we need to find a dense finite subset and establish the relationship of the error probabilities between an arbitrary channel and its neighbor in the subset. The following lemma reveals the relationship.

Lemma 10 *Let ($\boldsymbol{H}_1$,$\boldsymbol{K}_1$) and ($\boldsymbol{H}_2$,$\boldsymbol{K}_2$) be two channels satisfying the constraints in Theorem 1. Denote $\boldsymbol{x}^{(n)}$ and $\boldsymbol{y}^{(n)}$ to be the input and output n-sequence of p-dimensional vectors, respectively. Let $P_{\boldsymbol{H}_1,\boldsymbol{K}_1}\{\boldsymbol{y}^{(n)}|\boldsymbol{x}^{(n)}\}$ and $P_{\boldsymbol{H}_2,\boldsymbol{K}_2}\{\boldsymbol{y}^{(n)}|\boldsymbol{x}^{(n)}\}$ are the np-variate probability density for the output signal sequence $\boldsymbol{y}^{(n)}$ given $\boldsymbol{x}^{(n)}$, for the n-extension of the two channel ($\boldsymbol{H}_1$,$\boldsymbol{K}_1$) and ($\boldsymbol{H}_2$,$\boldsymbol{K}_2$), respectively. Then for those $\boldsymbol{x}^{(n)}$ satisfying $\|\boldsymbol{x}^{(n)}\|^2 \leq nP_x$ and $\boldsymbol{y}^{(n)}$ satisfying $\|\boldsymbol{y}^{(n)}\|^2 \leq nP_y$,*

$$\frac{P_{\boldsymbol{H}_1,\boldsymbol{K}_1}\{\boldsymbol{y}^{(n)}, \boldsymbol{x}^{(n)}\}}{P_{\boldsymbol{H}_2,\boldsymbol{K}_2}\{\boldsymbol{y}^{(n)}, \boldsymbol{x}^{(n)}\}} \leq e^{n(c_3(\delta,\eta)+c_4)}, \tag{6.14}$$

where

$$\begin{aligned} c_3(\delta,\eta) = & \frac{1}{2\alpha_0^2}(P_y + a^2P_x + a\sqrt{P_xP_y})\delta \\ & + \frac{1}{\alpha_0^2}(aP_x + \sqrt{P_xP_y})\eta, \end{aligned} \tag{6.15}$$

$$c_4 = \log\det(\boldsymbol{K}_2) - \log\det(\boldsymbol{K}_1), \tag{6.16}$$

The variables $\delta = \|\boldsymbol{K}_1 - \boldsymbol{K}_2\|$, $\eta = \|\boldsymbol{H}_1 - \boldsymbol{H}_2\|$, the numbers a and α_0 are defined in Theorem 1.

Proof: See Lemma 7 of [7]. ■

Now we are ready to prove the theorem. Define the δ-neighborhood of the channel $(\boldsymbol{H}, \boldsymbol{K}) \in \mathcal{I}$ to be the all the channels $(\boldsymbol{H}', \boldsymbol{K}') \in \mathcal{I}$ satisfying $\|\boldsymbol{K} - \boldsymbol{K}'\| \leq \delta$ and $\|\boldsymbol{H} - \boldsymbol{H}'\| \leq \delta$. The conditions in Theorem 1 guarantee that the channel space is compact. We can select a finite subset $\mathcal{I}' \subset \mathcal{I}$ such that for an arbitrary channel in $\mathcal{I}$, in its neighborhood there exists at least one channel belonging to $\mathcal{I}'$. We denote $|\mathcal{I}'|$ to be L_δ to emphasize its relationship with δ.

For any $R < C(\mathcal{I}) \leq C(\mathcal{I}'))$, by Lemma 9, we can find a sequence of codes whose probability of error over $\mathcal{I}'$ vanishes as the block-length grows. The code can be applied to the whole channel space in the following manner. If the channel is in $\mathcal{I}'$, then the receiver uses its own typical set decoder described in Lemma 9, otherwise the receiver borrows the typical set decoder from its neighbor that is in $\mathcal{I}'$. The probability of error when the receiver uses its neighbor's decoder can be bounded. To be specific, let $(\boldsymbol{H}, \boldsymbol{K}) \in \mathcal{I}$ and $(\boldsymbol{H}', \boldsymbol{K}') \in \mathcal{I}'$ satisfy $\|\boldsymbol{K} - \boldsymbol{K}'\| \leq \delta$ and $\|\boldsymbol{H} - \boldsymbol{H}'\| \leq \delta$, by Lemma 10 we get,

$$P_{\boldsymbol{H},\boldsymbol{K}}\left\{\mathcal{E} \cap \mathrm{E}_0^c \cap \mathrm{E}_y^c\right\} \leq e^{n(c_3(\delta,\delta)+c_4)} P_{\boldsymbol{H}',\boldsymbol{K}'}\left\{\mathcal{E} \cap \mathrm{E}_0^c \cap \mathrm{E}_y^c\right\}, \tag{6.17}$$

where $\mathcal{E}$ is the event that the receiver makes an error, E_0 is defined in (6.11) and E_y is the event that $\boldsymbol{y}^{(n)}$ violates the power constraints.

Using the Taylor expansion of $\det(\cdot)$ at $\boldsymbol{K}'$ we can show that $\det(\boldsymbol{K})/\det(\boldsymbol{K}') \leq 1 + P(\delta)/\alpha_0^p$, where $P(\delta)$ is a polynomial with $P(0) = 0$. Then, (6.17) becomes

$$\begin{aligned}
&P_{\boldsymbol{H},\boldsymbol{K}}\left\{\mathcal{E} \cap \mathrm{E}_0^c \cap \mathrm{E}_y^c\right\} \\
&\leq \exp\left\{n\left[c_3(\delta,\delta) + \frac{1}{2}\log\left(1 + \frac{P(\delta)}{\alpha_0^p}\right)\right]\right\} P_{\boldsymbol{H}',\boldsymbol{K}'}\left\{\mathcal{E} \cap \mathrm{E}_0^c \cap \mathrm{E}_y^c\right\} \\
&\leq \exp\left\{n\left[c_3(\delta,\delta) + \frac{1}{2}\log\left(1 + \frac{P(\delta)}{\alpha_0^p}\right)\right]\right\} P_{\boldsymbol{H}',\boldsymbol{K}'}\left\{\mathcal{E}\right\} \\
&\leq \exp\left\{n\left[c_3(\delta,\delta) + \frac{1}{2}\log\left(1 + \frac{P(\delta)}{\alpha_0^p}\right)\right]\right\} \cdot \\
&\quad L_\delta\left(\exp\left\{-c_1 n\right\} + 6\exp\left\{-\frac{n\epsilon^2}{4p}\right\} + (\log 2)\exp\left\{-n\epsilon\right\}\right)
\end{aligned} \tag{6.18}$$

The last inequality is by Lemma 9. Since $c_3(\delta, \delta)$ and $P(\delta)$ approach zero when δ goes zero, we can select sufficiently small δ to ensure the overall exponent in the last expression to be negative. Then as n goes zero, $P_{\boldsymbol{H},\boldsymbol{K}}\left\{\mathcal{E} \cap \mathrm{E}_0^c \cap \mathrm{E}_y^c\right\}$ vanishes. Now we use the union bound,

$$P_{\boldsymbol{H},\boldsymbol{K}}\{\mathcal{E}\} \leq P_{\boldsymbol{H},\boldsymbol{K}}\{\mathrm{E}_0\} + P_{\boldsymbol{H},\boldsymbol{K}}\{\mathrm{E}_y\} + P_{\boldsymbol{H},\boldsymbol{K}}\left\{\mathcal{E} \cap \mathrm{E}_0^c \cap \mathrm{E}_y^c\right\}. \tag{6.19}$$

According to Lemma 8, the first two terms vanish as the block-length approaches infinity, thus the code works for any channel in $\mathcal{I}$. The converse is due to the fact that there exists a channel in $\mathcal{I}$ whose mutual information is less than $C + \epsilon$. So any rate great than $C + \epsilon$ will not be achievable for that channel, where ϵ is an arbitrarily small number.

To extend it to $p \neq q$, let $m = \max(p, q)$ we can expand $\boldsymbol{H}$ to be of dimension $m \times m$ by padding zero columns or rows. Simultaneously we expand $\boldsymbol{y}$ by appending zeros when $p < q$ or expand $\boldsymbol{x}$ and $\boldsymbol{z}$ when $p > q$. The code works for the expanded compound channel also works for the original compound channel.

Remark: The above proof simplifies the error probability computation by using typical set decoding. This requires channel side information at the receiver. The original proof in [7] is stronger in the sense that it does not have this assumption. One could however argue that the rate lose due to channel estimation is negligible when the block-length goes to infinity.

6.2 Chapter 3

6.2.1 Orthogonalization Lemma

Lemma 11 *Let $\{\boldsymbol{A}_i, i = 1, \ldots, m\}$ be a set of $n \times n$ matrices satisfying*

$$\begin{aligned} \boldsymbol{A}_i\boldsymbol{A}_i^\dagger &= \xi_i\boldsymbol{I} \quad 1 \le i \le m, \\ \boldsymbol{A}_i\boldsymbol{A}_j^\dagger + \boldsymbol{A}_j\boldsymbol{A}_i^\dagger &= \xi_{ij}\boldsymbol{I} \quad 1 \le i < j \le m, \end{aligned} \tag{6.20}$$

where $\{\xi_i\}$ are positive real numbers and $\{\xi_{ij}\}$ are some real numbers. Each $\boldsymbol{A}_i$ can be written as a linear combination of matrices from a solution to the $[\rho_2(n), n, n]$ Hermitian Hurwitz-Radon matrix equations. i.e.,

$$\boldsymbol{A}_i = \sum_{j=1}^{2\rho_2(n)} a_{ij}\boldsymbol{C}_j, \quad 1 \le i \le m, \tag{6.21}$$

where $\{\boldsymbol{C}_i\}$ is a solution to the Hermitian Hurwitz-Radon matrix equations and $\{a_{ij}\}$ are real numbers. ■

Proof: Define the norm $\|\boldsymbol{A}_i\| = \sqrt{\xi_i}$ if $\boldsymbol{A}_i\boldsymbol{A}_i^\dagger = \xi_i\boldsymbol{I}$, and the inner product $(\boldsymbol{A}_i, \boldsymbol{A}_j) = \xi_{ij}$ if $\boldsymbol{A}_i\boldsymbol{A}_j^\dagger + \boldsymbol{A}_j\boldsymbol{A}_i^\dagger = \xi_{ij}\boldsymbol{I}$. Select a maximum linearly independent (with respect to real coefficients) subset of $\{\boldsymbol{A}_i\}$. Without loss of generality, assume they are the first p matrices. In the next step, we make $\boldsymbol{A}_1$ perpendicular to all of the remaining matrices. A new set of matrices $\{\boldsymbol{B}_{1i}\}$ are defined as follows.

$$\begin{aligned} \boldsymbol{B}_{11} &= \frac{\boldsymbol{A}_1}{\|\boldsymbol{A}_1\|}, \\ \boldsymbol{B}_{1i} &= \boldsymbol{A}_i - \boldsymbol{A}_1\frac{(\boldsymbol{A}_1, \boldsymbol{A}_i)}{2\|\boldsymbol{A}_1\|^2}, \quad i = 2, \cdots, p \end{aligned} \tag{6.22}$$

It is easy to verify the following properties.

$$\begin{aligned} \|\boldsymbol{B}_{11}\| &= 1, \\ (\boldsymbol{B}_{11}, \boldsymbol{B}_{1i}) &= 0, \quad i = 2, \cdots, p, \\ \|\boldsymbol{B}_{1i}\| &= c_{1i} > 0, \quad i = 2, \cdots, p, \\ (\boldsymbol{B}_{1i}, \boldsymbol{B}_{1j}) &= d_{ij}, \quad i, j = 2, \cdots, p, i \neq j, \end{aligned} \tag{6.23}$$

where c_{1i} and d_{ij} are real numbers. The last two equalities ensure that the new matrices still satisfy the eigenvector invariance conditions (6.20). The exact values of c_{1i} and d_{ij} are not of importance.

Then, we orthogonalize $\{\boldsymbol{B}_{12}\}$ by defining $\{\boldsymbol{B}_{2i}\}$.

$$\begin{aligned}\boldsymbol{B}_{22} &= \frac{\boldsymbol{B}_{12}}{\|\boldsymbol{B}_{12}\|}, \\ \boldsymbol{B}_{2i} &= \boldsymbol{B}_{1i} - \boldsymbol{B}_{12}\frac{(\boldsymbol{B}_{12}, \boldsymbol{B}_{1i})}{2\|\boldsymbol{B}_{12}\|^2}, \quad i = 3, \cdots, p\end{aligned} \tag{6.24}$$

The whole procedure resembles the Gram-Schmidt orthogonalization. We can proceed in this way until all of $\{\boldsymbol{A}_i, i = 1, \ldots, p\}$ are orthogonalized. This is possible since these matrices are initially linearly independent. Let $\boldsymbol{C}_i = \boldsymbol{B}_{ii}$, then $\{\boldsymbol{C}_i\}$ is a (possibly subset of) solution to the Hurwitz-Radon matrix equations. Each matrix $\boldsymbol{A}_i$ is a linear combination (with real coefficients) of the solution. ■

6.2.2 Proof of Theorem 5

Theorem 5 *A real linear space-time constellation $\mathcal{L}_x$ is channel-eigenvector invariant if and only if each matrix $\boldsymbol{A}_i$ can be written as a linear combination of a solution to the Hurwitz-Radon matrix equations. To be specific, $\boldsymbol{A}_i$ can be written as*

$$\boldsymbol{A}_i = \sum_{j=1}^{m} a_{ij}\boldsymbol{O}_j, \quad i = 1, \cdots, k, \tag{6.25}$$

where $m \leq k$. Each a_{ij} is a real number. $\{\boldsymbol{O}_j, j = 1, \cdots, m\}$ is a solution to the $[m, n_t, T]$ Hurwitz-Radon matrix equations.

Similarly, a complex linear space-time constellation $\mathcal{L}_z$ is channel-eigenvector invariant if and only if matrices $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ can be written as a linear combination of a solution to the Hermitian Hurwitz-Radon matrix equations. To be specific, $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ can be written as

$$\begin{aligned} \boldsymbol{A}_i &= \sum_{j=1}^{m} a_{ij}\boldsymbol{O}_j + \sum_{j=1}^{m} b_{ij}\boldsymbol{Q}_j, \quad i = 1, \cdots, k \\ \boldsymbol{B}_i &= \sum_{j=1}^{m} c_{ij}\boldsymbol{O}_j + \sum_{j=1}^{m} d_{ij}\boldsymbol{Q}_j, \quad i = 1, \cdots, k \end{aligned} \tag{6.26}$$

where $m \leq k$. a_{ij}, b_{ij}, c_{ij} and d_{ij} are real numbers. $\{\boldsymbol{O}_j, \boldsymbol{Q}_j, j = 1, \cdots, m\}$ is a solution to the $[m, n_t, T]$ Hermitian Hurwitz-Radon matrix equations. ■

Proof: If (6.25) holds, then we can show that (3.25) and (3.26) hold by straightforward algebra. We first verify (3.25).

$$\begin{aligned} \boldsymbol{A}_i\boldsymbol{A}_i^{\mathsf{T}} &= \left(\sum_{j=1}^{m} a_{ij}\boldsymbol{O}_j\right)\left(\sum_{j=1}^{m} a_{ij}\boldsymbol{O}_j\right)^{\mathsf{T}} \\ &= \sum_{j=1}^{m} a_{ij}^2\boldsymbol{O}_j\boldsymbol{O}_j^{\mathsf{T}} \\ &\quad + \sum_{j=1}^{m}\sum_{k>j}^{m} a_{ij}a_{ik}\left(\boldsymbol{O}_j\boldsymbol{O}_k^{\mathsf{T}} + \boldsymbol{O}_k\boldsymbol{O}_j^{\mathsf{T}}\right) \\ &= \left(\sum_{j=1}^{m} a_{ij}^2\right)\boldsymbol{I}. \end{aligned} \tag{6.27}$$

The last step is due to $\{\boldsymbol{O}_j, j = 1, \dots m\}$ from a solution to the Hurwitz-Radon matrix equations. Similarly, we also get (3.26).

$$\begin{aligned} &\boldsymbol{A}_i\boldsymbol{A}_j^{\mathsf{T}} + \boldsymbol{A}_j\boldsymbol{A}_i^{\mathsf{T}} \\ &= \left(\sum_{k=1}^{m} a_{ik}\boldsymbol{O}_k\right)\left(\sum_{l=1}^{m} a_{jl}\boldsymbol{O}_l\right)^{\mathsf{T}} \\ &\quad + \left(\sum_{l=1}^{m} a_{jl}\boldsymbol{O}_l\right)\left(\sum_{k=1}^{m} a_{ik}\boldsymbol{O}_l\right)^{\mathsf{T}} \\ &= \sum_{k=1}^{m}\sum_{l=1}^{m} a_{ik}a_{jl}\left(\boldsymbol{O}_k\boldsymbol{O}_l^{\mathsf{T}} + \boldsymbol{O}_l\boldsymbol{O}_k^{\mathsf{T}}\right) \\ &= \left(2\sum_{k=1}^{m} a_{ik}a_{jk}\right)\boldsymbol{I}. \end{aligned} \tag{6.28}$$

Conversely, if $\{\boldsymbol{A}_i, i = 1, \ldots, k\}$ is the basis of a linear CEI-STC, they satisfy the conditions in Lemma 3. The result is obtained by applying the orthogonalization lemma. The extension to the complex case is trivial. ■

6.2.3 Proof of Theorem 7

Theorem 7 *Consider the linear constellation defined in (3.23) and (3.24). Assume the basis matrices satisfying the following power constraints:*

$$Tr\left(\hat{\boldsymbol{A}}_i \hat{\boldsymbol{A}}_i^{\mathsf{T}}\right) = n_t, \quad i = 1, \ldots, k \tag{6.29}$$

$$Tr\left(\hat{\boldsymbol{A}}_i \hat{\boldsymbol{A}}_i^{\dagger} + \hat{\boldsymbol{B}}_i \hat{\boldsymbol{B}}_i^{\dagger}\right) = 2n_t, \quad i = 1, \ldots, k \tag{6.30}$$

where (3.55) is for the real case and (3.56) is for the complex case. The constellation is mutual-information lossy on average under Rayleigh fading channels with i.i.d. Gaussian input if $k < \min(n_t, n_r) \times T$. ■

Proof: We consider the real case. The MIMO system can be written as $\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X}$, where each entry of $\boldsymbol{H}$ is a Gaussian random variable $\mathcal{N}(0, 1)$. $\boldsymbol{X}$ is defined in (3.50). By stacking the columns of $\boldsymbol{Y}$ on top of each other, we have

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}\hat{\boldsymbol{A}}_1 \\ \vdots \\ \boldsymbol{H}\hat{\boldsymbol{A}}_T \end{bmatrix} \boldsymbol{x} \tag{6.31}$$

where $\boldsymbol{y}_i$ is the ith column of $\boldsymbol{Y}$, $\boldsymbol{x}$ is defined in (3.52). Let ρ be the signal to noise ratio. Denote $\mathcal{H} = \left[(\boldsymbol{H}\hat{\boldsymbol{A}}_1)^{\mathsf{T}} \ \cdots \ (\boldsymbol{H}\hat{\boldsymbol{A}}_T)^{\mathsf{T}}\right]^{\mathsf{T}}$ The mutual information of the channel (6.31) with i.i.d. Gaussian input is

$$
\begin{aligned}
\mathrm{MI} &= \frac{1}{T}\mathrm{E}\log\det\left(\boldsymbol{I}+\rho\mathcal{H}\mathcal{H}^{\mathsf{T}}\right) \\
&\overset{(a)}{\leq} \frac{1}{T}\sum_{i=1}^{T}\mathrm{E}\log\det\left(\boldsymbol{I}+\rho\boldsymbol{H}\hat{\boldsymbol{A}}_i\left(\boldsymbol{H}\hat{\boldsymbol{A}}_i\right)^{\mathsf{T}}\right) \\
&\overset{(b)}{=} \frac{1}{T}\sum_{i=1}^{T}\mathrm{E}\log\det\left(\boldsymbol{I}+\rho\boldsymbol{H}\boldsymbol{\Lambda}_i\boldsymbol{H}^{\mathsf{T}}\right) \\
&\overset{(c)}{\leq} \mathrm{E}\log\det\left(\boldsymbol{I}+\rho\boldsymbol{H}\boldsymbol{H}^{\mathsf{T}}\right)
\end{aligned}
\tag{6.32}
$$

In the above, (a) is due to Fischer's inequality [62, C.1.d. Theorem, p. 226]. (b) is because the distribution of $\boldsymbol{H}$ does not change when multiplied by an orthogonal matrix. $\boldsymbol{\Lambda}_i$ is a diagonal matrix with eigenvalues of $\hat{\boldsymbol{A}}_i\hat{\boldsymbol{A}}_i^{\mathsf{T}}$. (c) is due to the concavity of logdet$(\cdot)$ and the power constraint $\mathrm{Tr}\left(\hat{\boldsymbol{A}}_i\hat{\boldsymbol{A}}_i^{\mathsf{T}}\right) = n_t$. The same approach applies to the complex case. ■

References

[1] R. J. McEliece. Are there Turbo-Codes on Mars? *IEEE Information Theory Society Newsletter*, Dec. 2004.

[2] S. Chung, G. D. Forney, and T. J. Richardson. On the Design of Low-Density Parity-Check Codes within 0.0045dB of the Shannon Limit. *IEEE Comm. Let.*, 5(2), Feb 2001.

[3] A. Lapidoth and P. Narayan. Reliable Communication Under Channel Uncertainty. *IEEE Trans. on Info. Th.*, 44(6):2148–2177, Oct. 1998.

[4] D. Blackwell, L. Breiman, and A. J. Thomasian. The Capacity of a Class of Channels. *Ann. Math. Statist.*, 30:1229–1241, 1959.

[5] R. L. Dobrushin. Optimum Information Transmission Through a Channel with Unknown Parameters. *Radio Eng. Elec.*, 4(12), 1959.

[6] J. Wolfowitz. *Coding Theorems of Infomation Theory*, volume 3rd ed. Springer-Verlag, Berlin, Germany, 1978.

[7] W. L. Root and P. P. Varaiya. Capacity of Classes of Gaussian Channels. *SIAM J. Appl. Math.*, 16(6):1350–1393, Nov. 1968.

[8] R. D. Wesel, X. Liu, and W. Shi. Trellis Codes for Periodic Erasures. *IEEE Trans. on Comm.*, 28(1):938–947, Jun. 2000.

[9] C. Köse and R. D. Wesel. Universal Space-Time Trellis Codes. *IEEE Trans. on Info. Th.*, 49(10):2717–2727, Oct. 2003.

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

[11] Ungerboeck, G. Trellis-Coded Modulation with Redundant Signal Sets: Parts I and II. *IEEE Comm. Mag.*, 25(2):5–20, Feb. 1987.

[12] C. Schlegel. *Trellis Coding*. IEEE Press, 1997.

[13] M. Pursely and J. Skinner. Decoding strategies for turbo product codes in frequency-hop wireless communications. In *Proc. IEEE ICC–03*, May 2003.

[14] D. Chizhik, G. J. Foschini, M. J. Gans, and R. A. Valenzuela. Keyholes, Correlations, and Capacities of Multielement Transmit and Receive Anttenas. *IEEE Trans. on Wireless Comm*, 1(2):361–368, 2002.

[15] G. D. Forney. Modulation and Coding for Linear Gaussian Channels. *IEEE Trans. on Info. Th.*, 44(6):2384–2415, Oct. 1998.

[16] S. Verdu. Spectral Efficiency in the Wideband Regime. *IEEE Trans. on Info. Th.*, 48(6):1319–1343, June. 2002.

[17] T. M. Cover and J. A. Thomas. *Elements of Infomation Theory.* Wiley, New York, 1991.

[18] C. E. Shannon. Probability of Error for Optimal Codes in a Gaussian Channel. *Bell Sysst. Tech. J.*, 38:611–656, 1959.

[19] R. Gallager. *Information Theory and Reliabile Communication.* John Wiley & Sons, New York, 1968.

[20] S. J. MacMullan and O. M. Collins. A Comparison of Known Codes, Random Codes and the Best Codes. *IEEE Trans. on Info. Th.*, 44(7):3009–3022, Nov. 1998.

[21] S. Dolinar, D. Divsalar, and F. Pollara. Code Performance as a Function of Block Size. *TMO Progress Report 42-133*, May 1998.

[22] T. Tian, C. Jones, J. Villasenor, and R. Wesel. Construction of Irregular LDPC Codes with Low Error Floors. In *Proc. IEEE ICC-03*, Anchorage, May 2003.

[23] A. Ramamoorthy and R. Wesel. Construction of Short Block Length Irregular LDPC Codes. In *Proceedings of ICC*, Jun. 2004.

[24] T. Tian, C. Jones, J. D. Villasenor, and R. D. Wesel. Selective Avoidance of Cycles in Irregular LDPC Code Construction. *IEEE Trans. on Comm.*, To appear, Aug. 2004.

[25] C. Jones, A. Matache, T. Tian, J. Villasenor, and R. Wesel. The Univerality of LDPC Codes on Wireless Channels. In *Proceedings of Military Communications Conference*, Oct. 2003.

[26] W. Weng, C. Köse, and R. D. Wesel. Universal Serially Concatenated Trellis Coded Modulation for Perioidic Erasures and Periodic Fading. *IEEE Trans. on Comm.*, submitted 2004.

[27] I. E. Telatar. Capacity of Multi-antenna Gaussian Channels. *Europ. Trans. Telecommu.*, 10:585–595, Nov./Dec. 1999.

[28] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas. Uniform Power Allocation in MIMO Channels: a Game-Theoretic Approach. *IEEE Trans. on Info. Th.*, 49(7):1707–1727, Jul. 2003.

[29] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo. Signal Design for Transmitter Diversity Wireless Communication Systems over Rayleigh Fading Channels. In *Proc. IEEE VTC–96*, pages 136–140, 1996.

[30] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-Time Codes for High Data Rate Wirless Communication: Performance Criteria and Code Construction. *IEEE Trans. on Info. Th.*, 44(2):744–765, March 1998.

[31] D. Shapiro. *Compositions of Quadratic Forms.* W. de Gruyter, Berlin, Germany, 2000.

[32] A.V. Geramita and J. Seberry. *Orthogonal Designs: Quadratic Forms and Hadamard Matrices (Lecture Notes in Pure and Applied Mathematics)*, volume 43. Marcel Dekker, New York and Basel, 1979.

[33] B. Eckmann. Topology, Algebra, Analysis – Relations and Missing Links. *Notices of the AMS*, 46(5):520–527, May 1999.

[34] S. M. Alamouti. A Simple Transmit Diversity Tecnhique for Wireless Communications. *IEEE J. Select. Areas Commun.*, 16(8):1451–1458, Nov 1998.

[35] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-Time Block Codes from Orthogonal Designs. *IEEE Trans. on Info. Th.*, 45(5):1456–1467, July 1999.

[36] X. Liang. Orthogonal Designs With Maximal Rates. *IEEE Trans. on Info. Th.*, 49(10):2468–2503, Oct. 2003.

[37] L-F. Lu, P.V. Kumar, and H. Chung. On orthogonal designs and space-time codes. In *Proc. IEEE ISIT–02*, 2002.

[38] H. Jafarkhani and F. Taherkhani. Psuedo Orthognal Designs as Space-Time Block Codes. In *Proc. IEEE Int. Symp. Advances in Wireless Communications*, Sept. 2002.

[39] G. Ganesan and P. Stoica. Space-Time Block Codes: a Maximum SNR Approach. *IEEE Trans. on Info. Th.*, 47(4):1650–1656, Apr. 2001.

[40] B. Hassibi and B. M. Hochwald. High-Rate Codes That Are Linear in Space and Time. *IEEE Trans. on Info. Th.*, 48(7):1804–1824, Jul. 2002.

[41] H. Wang and X. -G. Xia. Upper Bounds of Rates of Complex Orthogonal Space-Time Block Codes. *IEEE Trans. on Info. Th.*, 49(10):2788–2796, Oct. 2003.

[42] S. Jafar, S. Vishwanath, and A. Goldsmith. Channel Capacity and Beamforming for Multiple Transmit and Receive Antennas with Covariance Feedback. In *Proc. IEEE ICC*, 2001.

[43] S. Siwamogsatham and M. P. Fitz. Improved High-Rate Space-Time Codes via Othogonality and Set Partitioning. In *Proc. IEEE WCNC-03*, volume 1, Mar. 2003.

[44] H. Jafarkhani and N. Seshadri. Super-Othogonal Space-Time Trellis Codes. *IEEE Trans. on Info. Th.*, 49(4):937–950, Apr. 2003.

[45] A. J. Viterbi and J. K. Omura. *Principles of Digital Communication and Coding*. McGraw-Hill, New York, 1979.

[46] E. Zehavi and J. K. Wolf. On the performance evaluation of trellis codes. *IEEE Trans. on Info. Th.*, 33(2):196–202, Mar. 1987.

[47] M. Rouanne and D. J. Costello, Jr. An algorithm for computing the distance spectrum of trellis codes. *IEEE J. Select. Areas Commun.*, 7(6):929–940, Aug. 1989.

[48] Y.-J. Liu, I. Oka, and E. Biglieri. Error probability for digital transmission over nonlinear channels with application to TCM. *IEEE Trans. on Info. Th.*, 36(5):1101–1110, Sept. 1990.

[49] E. Biglieri and P. J. McLance. Uniform distance and error probability properities of TCM schemes. 39(1):41–53, Jan. 1991.

[50] R. D. Wesel. Reduced-State Representations for Trellis Codes Using Constellation Symmetry. *IEEE Trans. on Comm.*, 52(8):1302–1310, Aug. 2004.

[51] D. Kucukyavuz and M. P. Fitz. New views of transfer function based performance analysis of coded modulation. In *Proc. 33rd Asilomaar Conf. on Signal, System and Computers*, pages 44–46, Pacific Grove, CA, Oct. 1999.

[52] W. Ryan and Z. Tang. Reduced-complexity error state diagrams in TCM and ISI channel performance evaluation. *IEEE Trans. on Comm.*, 52(12), Dec. 2004.

[53] E. Biglieri. High-level modulation and coding for nonlinear satellite channels. 32(5):616–626, May 1984.

[54] S. Benedetto, M. Mondin, and G. Montorsi. Performance evaluation of trellis coded modulation schemes. *Proc. IEEE*, 83(6):833–855, Jun. 1994.

[55] Z. Kohavi. *Switching and Finite Automata Theory*. McGraw-Hill, 1970.

[56] W. Zhang and C. Schlegel. Reduced state computation of distance spectrum for certain optimal convoluational codes. In *SICOM/ICIE 93*, Singapore, Sept. 1993.

[57] V. Pan. How Can We Speed Up Matrix Multiplication? *SIAM Review*, 26(3):393–415, Jul. 1984.

[58] F. R. Gantmaher. *The Theory of Matrices*. Chelsea Publishing Co., New York, 1959.

[59] N. Balabanian and T. A. Brickart. *Electrical Network Theory*. Robert E. Krieger Publishing Co., Malabar, Florida, 1983.

[60] C. Pimentel. On the computation of weight enumerators for convolutional codes. 51(3):313–317, Mar. 2003.

[61] J. Hopcroft. An nlogn Algorithm for Minimizing States in a Finite Automaton. *Theory of Machines and Computation, Kohavi ed.*, pages 189–196, 1971.

[62] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York, 1979.