Typeset with NDTheSiS version 2.14 (2000/09/08)
on April 14, 2004

for

Guangyue Han

entitled

SPACE TIME CODING WITH MULTIPLE ANTENNA SYSTEMS

This class conforms to the University of Notre Dame style guidelines established Fall 2000. However it is still possible to generate a non-conformant document if the published instructions are not followed! Be sure to refer to the published Graduate School guidelines as well.

This summary page can be disabled by specifying the nosummary *option to the class invocation.*(i.e., \documentclass[nosummary]{ndthesis})

THIS PAGE IS *NOT* PART OF THE THESIS, BUT SHOULD BE TURNED IN TO THE PROOFREADER!

NDTheSiS documentation can be found at these locations:

`http://www.nd.edu/~afsunix/faq/tetexdoc/latex/ndthesis/`
`http://www.gsu.nd.edu/Committees/ITC/ndthesis.pdf`
`http://www.gsu.nd.edu/Committees/ITC/sample_ndthesis.tar.gz`

General LaTeX documentation and info:

On-line docs:	
ND installation	`http://www.nd.edu/~afsunix/faq/tetexdoc/`
TeX User's Group	`http://www.tug.org/`
Books:	
A Guide...for Beg. & Adv. Users	by Kopka/Daly
LaTeX User's Guide ...	by Lamport
The LaTeX Companion	by Goossens/Mittelbach/Samarin
Packages: (check on-line docs)	
rotating	sideways tables and figures
longtable	multi-page tables
graphicx	using Postscript and other figures

SPACE TIME CODING WITH MULTIPLE ANTENNA SYSTEMS

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Guangyue Han, B.S., M.S.

Joachim Rosenthal, Director

Graduate Programs in Mathematics
Notre Dame, Indiana
April 2004

SPACE TIME CODING WITH MULTIPLE ANTENNA SYSTEMS

Abstract

by

Guangyue Han

Recently there has been considerable interest in wireless communication systems using multiple antennas. Theoretical analysis shows that such systems can provide spacial and temporal diversity gain if appropriate signal sets are employed. This dissertation analyzes and designs suitable signal constellations for such systems.

Diversity product and diversity sum will be derived as two important parameters to design constellations. A basic question is what the maximal diversity product or diversity sum could be. In this dissertation we derive general upper bounds on the diversity sum and the diversity product for unitary constellations of any dimension M and any size L using packing techniques on the compact Lie group $U(M)$.

We generalize one complex dimensional phase shift keying (PSK) signal and introduce space time constellations from generalized phase shift keying (GPSK) signals based on the complex and real orthogonal designs. The resulting space time constellations reallocate the energy among transmitting antennas and feature good diversity, consequently their performances are better than some of the existing compatible codes. Moreover, since the decoding of our proposed codes can be decomposed into one dimensional PSK signal demodulation, maximum likelihood (ML) decoding of our codes can be implemented in a very efficient way.

Group structure and other algebraic approaches have been considered to construct fully diverse unitary constellations in the literature. Our observation, how-

ever, indicates that full diversity can be easily obtained with h.d. random constellations. In this dissertation we also propose constellations with suitable structure which allow one to construct codes of any dimension M and any size L with excellent diversity using geometrical symmetry and numerical methods. We demonstrate how these structured constellations out-perform currently existing constellations and explain why the proposed constellation structure admit simple decoding algorithms such as sphere decoding. The presented design methods apply to any dimensional constellation of any size. Moreover, codes based on the proposed structure are very flexible and can be optimized for any signal to noise ratio.

To my parents

and to Jun,

who always light up for me,

with all my love.

CONTENTS

TABLES

FIGURES

ACKNOWLEDGMENTS

I would like to thank all the people who encouraged and supported my work towards this dissertation. First I want to cordially thank my advisor, Joachim Rosenthal, who lead me into this exciting research area and has encouraged and supported me constantly along the way. The dissertation wouldn't be possible without him. I am also grateful to my nice friends, Adrish, Ali, Arvind, Carmelo, Ching, Deepak, Marcin, Michael and XiaoWei, for their friendship and many things I learned from them.

I would like to thank the members of my defense committee, Andrew Sommese, Karen Chandler and Nicholas Laneman, for their time, effort and observations. I would like to thank all of the faculty, staff and graduate students of the Mathematics Department of University of Notre Dame for providing an extraordinary environment.

I would also like to acknowledge all of the institutions which have supported my work: the Department of Mathematics and the Center for Applied Mathematics at Notre Dame for their support and fellowships, also the National Science Foundation, for its grant support.

CHAPTER 1

INTRODUCTION

Digital communication systems are used to transmit information in digital forms from sources to destinations. In this introduction, the basic principles and methodologies to characterize, analyze and design digital communication systems will be presented. The characteristics of the physical channels, through which the information is transmitted, greatly affect the analysis and design of the communication systems. Several simple sample physical channels will be discussed and their mathematical models will be analyzed in a information theoretical way. This chapter is intended for mathematicians who are not very familiar with information and communication theory. We refer the readers to [49, 10] for more detailed discussion about a digital communication systems and information theory.

1.1 Elements of Digital Communication Systems

A generic digital communication system can be pictured as in Figure 1.1. The source may send either an analog signal or a digital one, however in a digital communication system, the source output is always converted into a sequence of binary digits. People are seeking an efficient transmission scheme for every particular communication channel, namely they are trying to represent the source output with as few digits as possible. This process of converting the analog or digital messages into a sequence of binary digits and removing redundancy in the messages is called source

encoding or data compression. For example, an MP3 file is a compressed digital representation of a song.

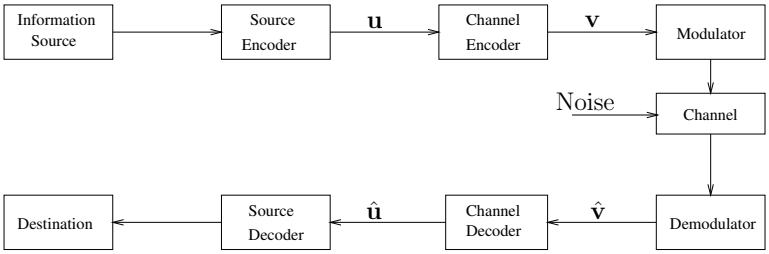


Figure 1.1. Block diagram of a typical digital communication system

The purpose of the channel encoder, however, is to add some redundancy into the sequence of binary digits from the source encoder. This process, which is often called channel encoding or error correction coding, is done in a controlled way such that the receiver can take advantage of the added redundant information to reliably recover the corrupted information sequence (the sequence of binary digits from the source encoder) by the noise, fading or interference in the transmission of the signals through the channel. The simplest error correction code is the repetition code: instead of sending a binary digit through the channel once, m copies of the digit will be sent. More complicated codes, such as linear block codes [33] and convolutional codes [33] have been studied in the literature. Recently the very powerful codes, turbo codes [6] and low density parity check (LDPC) codes [15] featuring efficient iterative decoding have been presented, and they are being studied actively. Every coding scheme adds the redundant information to overcome the corruption of the noise, fading or interference in its own way, which may require the receiver to take different measures to recover the originally sent information. The amount of added

redundancy is, however, universally characterized by the ratio of the number of digits of the sent information sequence to those of the encoded sequence, which is called the rate of the code.

The digital modulator will take over the binary sequence from the channel encoder and map the binary information sequence into signal waveforms, which most of the communication systems are capable of transmitting. For example the binary digital modulator operating in additive white Gaussian noise (AWGN) channel can map digit 0 to $\sqrt{E}\sin(t)$ and map digit 1 to $\sqrt{E}\cos(t)$, where $0 \leq t \leq T$. In this manner, each bit from the channel encoder is transmitted separately, namely the transmission rate of this scheme is 1. Alternatively the modulator can map b binary digit sequence to a constellation of size M. In this case the transmission rate will be $\log_2 M/b$. In this dissertation we will be concerned with unitary space time modulation (USTM), which maps a sequence of binary digits to a "matrix" of signal waveforms. We shall analyze a wireless communication system with multiple antennas, present the design criterion for signal constellations (mathematically this can be viewed as a finite set of special matrices), and explicitly construct good-performing constellations.

As indicated in the preceding discussion, the communication channel provides the connection between the transmitter and the receiver. From the transmitter the signal waveforms have to go through the physical medium to reach the receiver. There are many kinds of channels featuring different kinds of transmission media, such as the atmosphere in wireless transmission, or wired lines in computer networks. Regardless of the channel physical medium, the corruption of the signal waveforms is involved with uncertainty. Mathematically all the channel noises will be characterized by a random variable or stochastic process. A fundamental parameter of a channel is the capacity, which indicates how fast the data can be reliably

transmitted through the channel.

The task of the digital demodulator is to detect and estimate the channel-corrupted signal waveforms and to map the estimated waveforms to the binary digit or real number sequence for soft decoding (for instance the iterative decoding of a turbo code or LDPC code). Then the channel decoder will take over the binary or real number sequence, remove the redundancy added by the channel encoder and recover the originally sent information sequence produced by the source encoder.

The bit error rate (BER) is defined to be the ratio of the number of the incorrectly decoded information digits to the number of binary digits sent from the source encoder. It is a very important indicator of how well the receiver is performing. Many parameters of the communication system will affect BER, such as the characteristics of the channel, the modulation and coding method used during the transmission, the nature of the noise, fading and interference and the signal-to-noise ratio (SNR) of the channel.

The source decoder will do the "inverse" mapping of the processing procedure of the source encoder to recover the originally sent messages. For lossless source encoding, the source decoder will go through the exact inverse process of a source encoder's encoding procedure. For a lossy source encoding, such as the digital image compression, the source decoder will try to combine the received data and the knowledge of the encoding method to reconstruct the message. Many times the estimated message is only an approximation of the originally sent message. In this regard the "difference" between the estimated message and the originally sent message will be used to measure how well the whole communication system is performing.

One common problem in signal transmission through any channel is additive noise. In general, additive noise is generated internally by components such as

resistors and solid-state devices used to implement the communication system. This is sometimes called thermal noise. Other sources of noise and interference may arise externally to the system, such as interference from other users of the channel. When such noise and interference occupy the same frequency band as the desired signal, their effect can be minimized by the proper design of the transmitted signal and its demodulator at the receiver. Other types of signal degradations that may be encountered in transmission over the channel are signal attenuation, amplitude and phase distortion and multi-path distortion.

The effects of noise may be minimized by increasing the power in the transmitted signal. However, equipment and other practical constraints limit the power level in the transmitted signal. Another basic limitation is the available channel bandwidth. Physically a bandwidth constraint is usually due to the physical limitations of the medium and the electronic components used to implement the transmitter and the receiver. These two limitations constrain the amount of data that can be transmitted reliably over any communication channel.

We find it convenient to construct mathematical models that reflect the most important characteristics of the transmission medium, then the mathematical model for the channel is used in the design of the channel encoder and modulator at the transmitter and the demodulator and channel decoders at the receiver. Below, we provide a brief description of additive-noise channels that are frequently used to characterize many of the physical channels in practice.

An additive-noise channel is one of the simplest and most common models for a communication channel. The transmitted signal is corrupted by an additive noise, which is a very common stochastic process. Physically additive noise is caused internally by the electronic components which are used to implement the communication system, or from the interference encountered in the transmission, such as

the interference caused by other users in the communication channel. If the noise is introduced primarily by electronic components and amplifier at the receiver, it may be characterized as thermal noise. This type of noise is characterized statistically as a white Gaussian noise process. Hence, the resulting mathematical model for the channel is usually called the additive white Gaussian noise (AWGN) channel. Because this channel model can be applied to a broad class of physical communication channels and because of its mathematical tactility, this is the predominant channel model studied in the communication system analysis and design. We will analyze the AWGN channel from an information theoretic point of view.

1.2 Introduction to Information Theory

Information theory, or the mathematical theory of communication, originated from Shannon's groundbreaking paper [39] in 1948. Based on the mathematical models for practical channels, information theory analyzes the communication channels using the tools from mathematics, such as analysis, probability and statistics.

As a measure of uncertainty of a random variable, entropy is a very basic concept in information theory. In the following let X denote a discrete random variable in alphabet $\mathcal{X}$ with probability mass function $p_X(x) = \Pr(X = x)$, $x \in \mathcal{X}$. For simplicity, the probability mass function will be denoted by $p(x)$ rather than $p_X(x)$. In this way $p(x)$ and $p(y)$ refer to two different random variables, and are in fact different probability mass functions, $p_X(x)$ and $p_Y(y)$, respectively.

Definition 1.1. The entropy $H(X)$ of a discrete random variable X is defined to be

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log p(x).$$

The joint entropy $H(X_1, X_2, \cdots, X_n)$ of a number of discrete random variables

$(X_1, X_2, \cdots, X_n)$ with a joint distribution $p(x_1, x_2, \cdots, x_n)$ is defined to be

$$H(X_1, X_2, \cdots, X_n) = -\sum_{x_1 \in \mathcal{X}_1} \sum_{x_2 \in \mathcal{X}_2} \cdots \sum_{x_n \in \mathcal{X}_n} p(x_1, x_2, \cdots, x_n) \log p(x_1, x_2, \cdots, x_n).$$

If $(X, Y) \sim p(x, y)$, then the conditional entropy $H(Y|X)$ is defined to be

$$H(Y|X) = \sum_{x \in \mathcal{X}} p(x) H(Y|X = x) = -\sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y|x) \log p(y|x)$$

$$= -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(y|x) = -E_{p(x,y)} \log p(Y|X).$$

Another interpretation of the entropy $H(X)$ is that it can viewed as the expected value of $\log \frac{1}{p(X)}$, where $p(x)$ is the probability mass function for X. Therefore

$$H(X) = E_p \log \frac{1}{p(X)}.$$

The same interpretation can be applied to joint entropy $H(X_1, X_2, \cdots, X_n)$. So we have

$$H(X_1, X_2, \cdots, X_n) = -E \log p(X_1, X_2, \cdots, X_n).$$

In the following one will see the conditional entropy is closely related to joint entropy. Let $X_1, X_2, \cdots, X_n$ be drawn according to $p(x_1, x_2, \cdots, x_n)$. Then

$$H(X_1, X_2, \cdots, X_n) = \sum_{i=1}^{n} H(X_i | X_{i-1}, \cdots, X_1).$$

Mutual information between two random variables X and Y, as defined below, is a measure of to what extent X is determined by Y, or vice versa.

Definition 1.2. Let X and Y be two random variables with a joint probability mass function $p(x, y)$ and marginal probability mass function $p(x)$ and $p(y)$. The mutual information $I(X; Y)$ is defined to be

$$I(X; Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} = E_{p(x,y)} \log \frac{p(X, Y)}{p(X)p(Y)}.$$

The following theorem indicates a way to calculate the mutual information.

Theorem 1.3.

$$I(Y;X) = I(X;Y) = H(X)+H(Y)-H(X,Y) = H(X)-H(X|Y) = H(Y)-H(Y|X)$$

Particularly we have

$$I(X;X) = H(X).$$

A discrete channel is a system consisting of an input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ and a probability transition matrix $p(y|x)$ that expresses the probability of observing the output symbol y given that the symbol x is sent. The channel is said to be memoryless if the probability distribution of the output depends only on the input at that time and is conditionally independent of the previous channel inputs or outputs.

Definition 1.4. The information channel capacity of a discrete memoryless channel is defined to be

$$C = \max_{p(x)} I(X;Y),$$

where the maximum is taken over all possible input distributions $p(x)$.

The following theorem has been proved by Shannon. It is the fundamental result in channel coding literature. Basically it is indicated that as long as the transmission rate is lower than the capacity, one can always make the transmission through the channel reliable by using long codes.

Theorem 1.5. *(The channel coding theorem): All rates below capacity C are achievable. Specifically, for every rate $R < C$, there exists a sequence of $(2^{nR}, n)$ codes with maximum probability of error $\lambda^{(n)} \to 0$. Conversely, any sequence of $(2^{nR}, n)$ codes with $\lambda^{(n)} \to 0$ must have $R \leq C$.*

Similar to the entropy of a discrete random variable, the concept of differential entropy of a continuous random variable can be defined. Differential entropy is similar in many ways to the entropy of a discrete random variable. But there are some important differences, and there is need for some care in using the concept.

Let X be a random variable with cumulative distribution function $F(x) = \Pr(X \leq x)$. If $F(x)$ is continuous, the random variable is said to be continuous. Let $f(x) = F'(x)$ when the derivative is defined. If $\int_{-\infty}^{\infty} f(x) = 1$, then $f(x)$ is called the probability density function for X. The set where $f(x) > 0$ is called the support set of X. Similar to the discrete case, we can extend the definition of differential entropy of a single random variable to several random variables.

Definition 1.6. The differential entropy $h(X)$ of a continuous random variable X with a probability density $f(x)$ is defined to be

$$h(X) = -\int_S f(x) \log f(x) dx,$$

where S is the support set of the random variable. The differential entropy of a set $X_1, X_2, \cdots, X_n$ of random variables with probability density $f(x_1, x_2, \cdots, x_n)$ is defined to be

$$h(X_1, X_2, \cdots, X_n) = -\int f(x_1, x_2, \cdots, x_n) \log f(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n.$$

If X, Y have a joint density function $f(x, y)$, we can define the conditional differential entropy $h(X|Y)$ to be

$$h(X|Y) = -\int f(x, y) \log f(x|y) dx dy.$$

Definition 1.7. The mutual information $I(X;Y)$ between two random variables with joint probability density $f(x, y)$ is defined to be

$$I(X;Y) = \int f(x, y) \log \frac{f(x, y)}{f(x) f(y)} dx dy.$$

1.3 Analysis of Several Channel Models

1.3.1 Additive White Gaussian Noise Channel

Figure 1.2 shows the most important continuous channel: the Gaussian channel. This is a time discrete channel with input X_i and output Y_i, which is the sum of

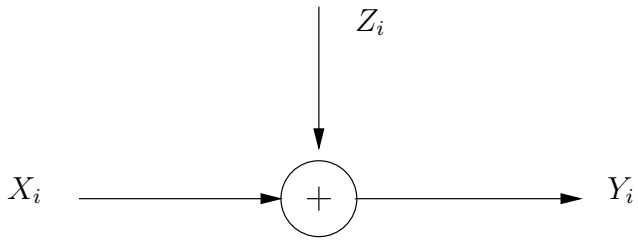


Figure 1.2. Discrete time AWGN channel

the input and the noise Z_i. The noise Z_i is a i.i.d. random variable with a Gaussian distribution with variance N. Mathematically we have

$$Y_i = X_i + Z_i, \qquad Z_i \sim \mathcal{N}(0, N).$$

The noise Z_i is assumed to be independent of the signal X_i. Without further conditions, the capacity of this channel may be infinite. For instance if the noise variance is zero, then the receiver will receive the transmitted symbol perfectly. Since X can take on any real value, an arbitrary real number can be transmitted through the channel without any error.

We now define the information capacity of the channel to be the maximum of the mutual information between the input and output over all distributions on the input that satisfy the average power constraint. Therefore the information capacity of the Gaussian channel with power constraint P is

$$C = \max_{p(x):EX^2 \leq P} I(X;Y).$$

The information capacity for a Gaussian channel can be calculated as follows. Expand $I(X;Y)$, then according to Theorem 1.3 we have

$$I(X;Y) = h(Y) - h(Y|X) = h(Y) - h(X+Z|X) = h(Y) - h(Z|X) = h(Y) - h(Z),$$

since Z is independent of X. The fact that Z is Gaussian random variable will lead to

$$h(Z) = \frac{1}{2}\log 2\pi eN.$$

Since X and Z are independent and $EZ = 0$

$$EY^2 = E(X+Z)^2 = EX^2 + 2EXEZ + EZ^2 = P + N.$$

By the fact that the normal distribution maximizes the entropy for a given variance, we conclude that the entropy of Y is bounded by $\frac{1}{2}\log 2\pi e(P+N)$ (note that $EY^2 = P+N$). Applying this result to bound the mutual information $I(X,Y)$, we obtain

$$I(X;Y) = h(Y) - h(Z) \leq \frac{1}{2}\log 2\pi e(P+N) - \frac{1}{2}\log 2\pi eN = \frac{1}{2}\log\left(1+\frac{P}{N}\right).$$

One can check the equality can be reached if X assumes the normal distribution. Hence the information capacity of the Gaussian channel is

$$C = \max_{EX^2 \leq P} I(X;Y) = \frac{1}{2}\log(1+\frac{P}{N}),$$

and the capacity is attained when $X \sim \mathcal{N}(0,P)$.

1.3.2 Parallel Gaussian Channel

In the sequel, k independent Gaussian channels in parallel with a common power constraint will be analyzed. The objective is to distribute the total power among the channels to maximize the capacity. This channel models a non-white additive Gaussian noise channel where each parallel component represents a different frequency.

Assume that we have a set of Gaussian channels in parallel as illustrated in Figure 1.3. The output of each channel is the sum of the input and Gaussian noise.

Mathematically for each channel j,

$$Y_j = X_j + Z_j, \qquad j = 1, 2, \cdots, k,$$

with

$$Z_j \sim \mathcal{N}(0, N_j),$$

and the noise is assumed to be independent from channel to channel. There is a common power constraint on the total power to be used, *i.e.*,

$$E \sum_{j=1}^{k} X_j^2 \leq P.$$

We wish to distribute the power among the various channels to to maximize the total capacity C:

$$C = \max_{f(x_1, x_2, \cdots, x_k): \sum EX_i^2 \leq P} I(X_1, X_2, \cdots, X_k; Y_1, Y_2, \cdots, Y_k).$$

We calculate the distribution that achieves the information capacity for this channel. Since $Z_1, Z_2, \cdots, Z_k$ are independent,

$$I(X_1, X_2, \cdots, X_k; Y_1, Y_2, \cdots, Y_k) = h(Y_1, Y_2, \cdots, Y_k) - h(Y_1, Y_2, \cdots, Y_k | X_1, X_2, \cdots, X_k)$$

$$= h(Y_1, Y_2, \cdots, Y_k) - h(Z_1, Z_2, \cdots, Z_k | X_1, X_2, \cdots, X_k)$$

$$= h(Y_1, Y_2, \cdots, Y_k) - h(Z_1, Z_2, \cdots, Z_k)$$

$$= h(Y_1, Y_2, \cdots, Y_k) - \sum_i h(Z_i) \leq \sum_i h(Y_i) - h(Z_i) \leq \sum_i \frac{1}{2} \log(1 + \frac{P_i}{N_i}),$$

where $P_i = EX_i^2$ and $\sum_i P_i = P$. The above equality is achieved by

$$(X_1, X_2, \cdots, X_k) \sim \mathcal{N}\left(0, \begin{bmatrix} P_1 & 0 & \cdots & 0 \\ 0 & P_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_k \end{bmatrix}\right).$$

So the problem is reduced to finding the power allocation that maximizes the capacity subject to the constraint that $\sum_i P_i = P$. This is a standard optimization problem and can be solved using Lagrange multipliers. Routine analysis will show that the solution

$$P_i = (v - N_i)^+$$

is the assignment that maximize capacity, where v is chosen so that

$$\sum (v - N_i)^+ = P.$$

Here $(x)^+$ denotes the positive part of x, *i.e.*,

$$(x)^+ = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}.$$

1.3.3 Multiple Input Multiple Output (MIMO) Channel

In this section, we analyze a MIMO channel. Figure 1.4 depicts a wireless communication channel with multiple transmit antennas and multiple receive antennas. We use a block-fading model with coherence interval T. The transmitter is equipped with M transmit antennas, and the signal are simultaneously transmitted to a single receiver equipped with N antennas in a flat-fading environment, where channel state information (CSI) may or may not be known by the receiver. The fading is described by an $M \times N$ complex-valued propagation matrix H, which remains constant for T symbol periods, after which it jumps to a new independent value for another T symbols, and so on. During a coherence interval, the M antennas collectively transmit a $T \times M$ complex matrix S, whose columns represent different antennas, and the receiver records a $T \times N$ complex matrix X. The basic equation describing this channel is as follows:

$$X = \sqrt{\frac{\rho}{M}} SH + W,$$

where W is a $T \times N$ matrix of additive receiver noise, whose components are independent, zero-mean complex Gaussian with unit variance distributed with density

$$p(\mathbf{w}_{ij}) = \frac{1}{\pi} exp[-|\mathbf{w}_{ij}|^2].$$

A complex random variable with the above density function is called $\mathcal{CN}(0,1)$ distributed. The components of H are assumed independent with distribution $\mathcal{CN}(0,1)$ as well. We enforce an expected power constraint

$$\mathrm{tr}\, E\{SS^*\} = TM.$$

Combined with the normalization, this constraint $1/\sqrt{M}$ implies that the total transmitted power remains constant as the number of transmit antennas changes, and here the ρ denotes the expected SNR at each receive antenna.

We wish to choose the joint probability density of the components of S, subject to the independence of its M columns, and subject to the power constraint to maximize the mutual information

$$I(X;S) := E\{\log\left(\frac{p(X|S)}{p(X)}\right)\}.$$

This maximization yields the total throughout or capacity.

It is convenient to assume Rayleigh fading, where the independent components of H are distributed as $\mathcal{CN}(0,1)$, although this assumption can be relaxed for some of the asymptotic results. For Rayleigh fading, the conditional density takes the form

$$p(X|S) = \frac{\exp(-X^*\Lambda^{-1}X)}{\pi^{TN}\det{}^N\Lambda}, \tag{1.1}$$

where $\Lambda = I_T + (\rho/M)SS^*$ and I_T denotes the $T \times T$ identity matrix.

If the receiver somehow knew the random propagation coefficients, the capacity would be greater than for the case of interest where the receiver does not know

propagation coefficients. Telatar [45] computes the perfect-knowledge capacity for the case $T = 1$. It is straightforward to extend his analysis to $T > 1$. The perfect-knowledge fading link is completely described by the conditional probability density

$$p(X, H|S) = p(X|H, S)p(H).$$

The perfect-knowledge capacity is obtained by maximizing the mutual information between (X, H) and S with respect to $p(S)$. The mutual information is

$$I(X, H; S) = E \log \frac{p(X, H|S)}{p(X, H)} = E \log \frac{p(X|H, S)}{p(X|H)} = E\{E\{\log \left(\frac{p(X|H, S)}{p(X|H)} \right) |H\}\}.$$

The inner expectation, conditioned on H, is simply the mutual information for the classical additive Gaussian noise case and is maximized by making the components of S independent $\mathcal{CN}(0, 1)$. The resulting perfect-knowledge capacity is

$$C = E \log \det[I_N + \frac{\rho}{M} H^* H].$$

The asymptotic capacity can be described as follows:

$$\lim_{M \to \infty} C = N \log(1 + \rho).$$

Using the identity $\det[I_N + \frac{\rho}{M} H^* H] = \det[I_M + \frac{\rho}{M} H H^*]$, we can derive the following limiting result:

$$\lim_{M \to \infty} C = M \log(1 + \frac{\rho N}{M}).$$

For a slow fading channel, for instance a fixed wireless communication system, the coherent assumption is reasonable because the transmitter can send pilot signals which allow the receiver to estimate the fading coefficients accurately. However in certain situations, due to the difficulty of measuring the fading coefficients under the practical communication environment (for instance the limited resources or a fast fading link) the above coherent model would be questionable. Because of this, researchers put their efforts into the case when the channel is non-coherent, *i.e.*,

the CSI is not known by either the transmitter or the receiver. The basic equation will still be still the same, however in non-coherent scenarios it is assumed that the receiver does not know the exact values of the entries of H (other than their statistical distribution).

Theorem 1.8. [34] *For any coherence interval T and any number of receive antennas, the capacity obtained with $M > T$ transmitter antennas is the same as the capacity obtained with $M = T$ transmitter antennas.*

Sketch of the proof. Suppose that a particular joint probability density of the elements of SS^* achieves capacity with $M > T$ antennas. We can perform the Cholesky factorization $SS^* = LL^*$, where L is a $T \times T$ lower triangular matrix. Using T transmitter antennas, with a signal matrix that has the same joint probability density as the joint probability density of L, we may therefore also achieve the same probability density on SS^*, thus we get the same channel conditional probability (1.1). Moreover, if S satisfies the power constraint, then so does L. □

A $T \times T$ unitary matrix Φ is called isotropically distributed if its probability density is unchanged when pre-multiplied by a deterministic unitary matrix:

$$p(\Phi) = p(\Theta\Phi), \qquad \forall\Theta : \Theta^*\Theta = I.$$

One will see the capacity approaching signal sets are closely related to isotropically distributed unitary random matrices. The following theorem tell us it is important to study the unitary constellations for a MIMO fading channel.

Theorem 1.9. [34] *The signal matrix that achieves capacity can be written as $S = \Phi V$, where Φ is an $T \times T$ isotropically distributed unitary matrix, and V is an independent $T \times M$ real, nonnegative, diagonal matrix. Furthermore we can choose the joint density of the diagonal elements of V to be unchanged by rearrangements of its arguments.*

Sketch of the proof. Combine the singular value decomposition $S = U\Phi V$ and channel conditional density function (1.1), one can prove the same capacity can be

achieved by $S' = \Phi V$. Suppose S' has probability density function $p(S')$ and generate mutual information I_0. Let Θ be an isotropically distributed unitary matrix that is statistically independent of Φ and V, and define a new signal matrix $S'' = \Theta S$, generating mutual information I_1. One can check conditioned on Θ, the mutual information generated by S'' equals I_0. The concavity of mutual information as a function as a function of $p(S)$ and Jensen's inequality imply that $I_1 \geq I_0$. The same argument can be applied to the distribution of diagonal matrix. □

The exact closed for the capacity of a non-coherent channel is difficult to calculate. However one can analyze the limiting behavior of the capacity with the following theorem.

Theorem 1.10. [48]

$$\lim_{\rho \to \infty} \frac{C(M, N, T; \rho)}{\log \rho} = TM^* \left(1 - \frac{M^*}{T}\right),$$

where $M^* = \min\{M, N, \lfloor \frac{T}{2} \rfloor\}$.

Intuitively if we fix the number of antennas M and let the coherence time T increase to infinity, this would corresponds to the case with perfect knowledge of fading coefficients. The following theorem characterize the connection between a non-coherent and a coherent channel.

Theorem 1.11. [48]

$$C_{M,M}(\rho) = (1 - \frac{M}{T}) C_{coherent}(\rho) + \frac{1}{T} \log_2 |G(T, M)| + M(1 - \frac{M}{T}) \log_2 \frac{T}{\pi e} + o(1),$$

where $G(T, M)$ *represent the* $T \times M$ *complex Grassmannian manifold, which is the set of* M *complex dimensional subspaces of* $\mathbb{C}^T$.

1.4 Overview of this Dissertation

This dissertation is devoted to constructing good-performing signal constellations for non-coherent MIMO fading channels.

Chapter 2 will briefly review the mathematical tools employed in this dissertation. Since we consider the construction of unitary constellations, we attempt to look at $U(M)$ from different perspectives. Basic concepts from differential geometry will be introduced first. As a homogeneous space, $U(M)$ also has group structure, namely $U(M)$ is also a Lie group. Haar measure on this compact Lie group can be explicitly described. That will lead to the solution to the sphere packing problem on $U(M)$. The Riemannian geometry will follow the discussion since we will consider the geodesics, tangent space, normal space of $U(M)$, which has the naturally induced Riemannian metric from the embedding Euclidean space. Full diversity of a unitary constellations can be easily explained by the Zariski topology of an algebraic variety, therefore elements of algebraic geometry will be presented as well.

Design criteria will be discussed in Chapter 3. Basically the diversity product and the diversity sum describe the limiting behavior of the diversity function, which is somehow related to the signal-to-noise ratio. Specifically the diversity product can be viewed as the design criterion of unitary constellations at high SNR. Although one may argue unitary constellations are not capacity approaching signal sets anymore at low SNR case, our simulations show the constellations aiming for a large diversity sum perform better compared to the compatible constellations designed for large diversity product.

Chapter 4 will use the homogeneity of $U(M)$ to derive an upper bound for the diversity. Three upper bounds will be given, however one can see that the basic idea is to pack the small 'balls' into the whole manifold. For 2 dimension, the new upper bound is tighter than the previously derived one in the literature when the constellation size is greater than 100. For higher dimensions, we believe the upper bound should be very tight when the constellation size gets very large.

In Chapter 5, the reader will see a generalization of the PSK signal sets, which

is a popular choice for a traditional communication channel. When it comes to a multiple input and multiple output channel, one can consider the counterpart in higher dimensions. The proposed GPSK constellations use the space diversity more efficiently by reallocating the power among different transmit antennas. Since the demodulation of GPSK signals are able to be decomposed to that of one complex dimensional PSK (or 2 real dimensional PSK), they feature very fast decoding algorithm.

Inspired by the group structure, which has been considered to construct fully diverse constellations in the literature, we present structured constellations in Chapter 6. Our idea is to relax the restrictive constraint imposed on the group structure and to consider a subset of a freely generated group. The suitability of the structure allows one to search for good-performing constellations with large diversity. Meanwhile the fast decoding algorithm can be applied to these structures as well. To the best of our knowledge, the structured constellations are the first series of constellations which feature large diversity and fast decoding for any parameters M, N, T, L.

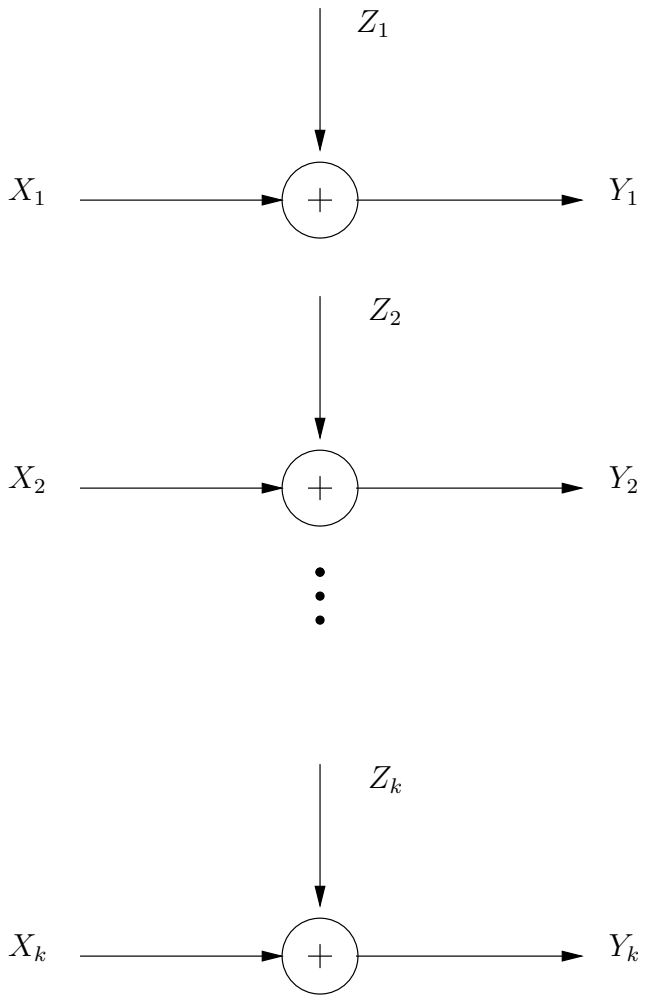


Figure 1.3. Discrete time parallel AWGN channel

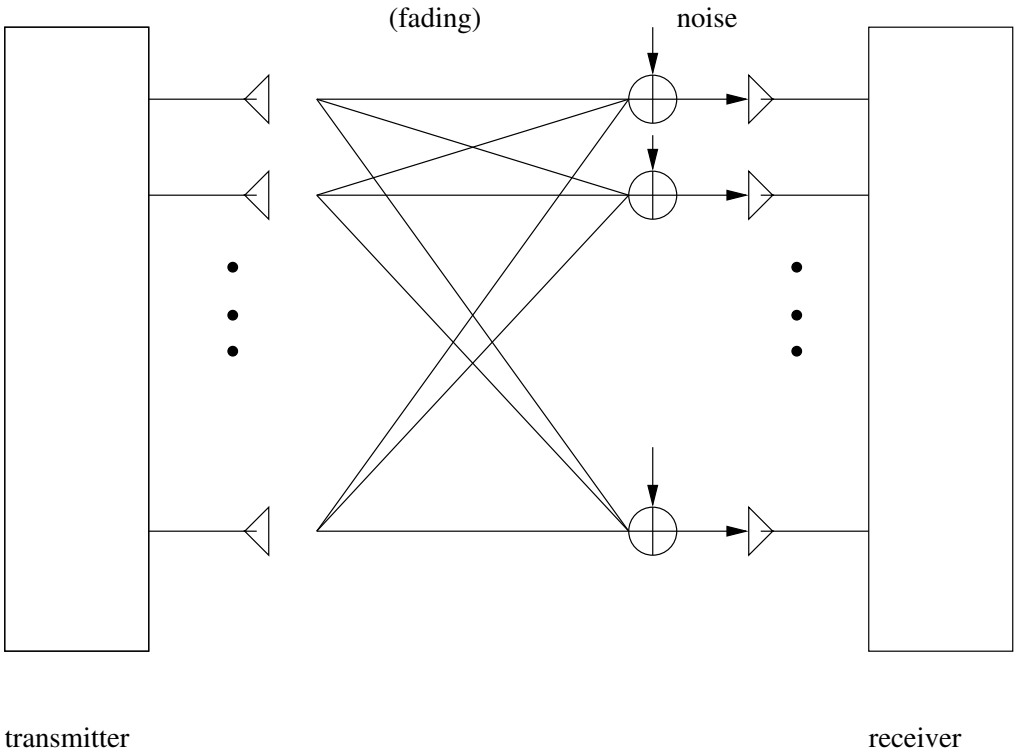


Figure 1.4. A wireless communication channel with multiple antennas

CHAPTER 2

PRELIMINARY MATHEMATICS

This chapter is intended for engineers who are not very familiar with the preliminary mathematical tools which will be used to develop this thesis. For the purpose of this dissertation, we are mainly interested in the geometrical aspects of the unitary group $U(n)$. First we present some general basic elements about differential geometry. Based on these knowledge, Lie groups will be discussed. Some basic elementary concepts in algebraic geometry will be introduced as well. Finally optimization techniques on a Riemannian manifold will be presented. We will only introduce some elementary concepts and intuitive description in this chapter, no complete characterization of the above theory will be attempted. For a more detailed and complete reference of the relevant theory, one can go to [2, 7, 26].

2.1 Differential Geometry

In this section, we introduce some basic concepts from differential geometry. A differentiable structure on a topological manifold M is a family of coordinate neighborhoods $\mathcal{U} = \{(U_\alpha, \phi_\alpha) : \alpha \in I\}$ such that:

1. $\{U_\alpha\}$ cover M;
2. Any two neighborhoods in $\mathcal{U}$ are C^∞-compatible;
3. $\mathcal{U}$ is maximal: any C^∞-coordinate neighborhood which is compatible with all coordinate neighborhoods of $\mathcal{U}$ is in $\mathcal{U}$.

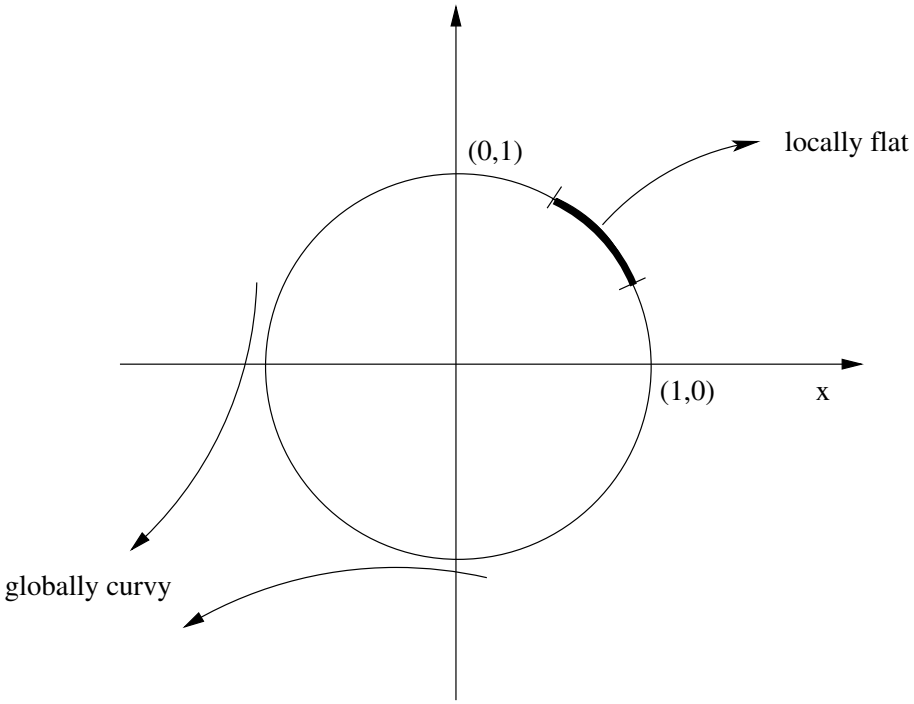


Figure 2.1. A simple example of manifold

A topological manifold together with a C^∞-differentiable structure is called a differentiable manifold .

From the above definition, one can see that locally a differentiable manifold behaves like a Euclidean space, although globally they may have totally different properties compared to a Euclidean space. As shown in Figure 2.1, a very simple example of a differentiable manifold is the unit circle in $\mathbb{R}^2$.

Let M be a C^∞ manifold of dimension n. For a point $p \in M$, consider the set of all the C^∞ functions whose domain of definition includes some open neighborhood of p. Two functions are equivalent if they agree on some neighborhood of p. The equivalence class is denoted by $C^\infty(p)$. Let $U \subset M$ be an open set, then $C^\infty(U) = \cup_{p\in U} C^\infty(p)$ is a commutative algebra over $\mathbb{R}$.

The tangent space $T_p(M)$ to M at p is defined to be the set of all mappings $X_p : C^\infty(p) \to \mathbb{R}$ satisfying for all $\alpha, \beta \in \mathbb{R}$ and $f, g \in C^\infty(p)$ the two conditions:

1. $X_p(\alpha f + \beta g) = \alpha(X_p f) + \beta(X_p g)$ (linearity),

2. $X_p(fg) = (X_p f)g(p) + f(p)(X_p g)$ (Leibniz rule),

with the vector space operations in $T_p(M)$ defined by

$$(X_p + Y_p)f = X_p f + Y_p f,$$

$$(\alpha X_p)f = \alpha(X_p f).$$

A tangent vector to M at p is any $X_p \in T_p(M)$. A tagent vector can be viewed as a generalizatin of a directional derivative in Euclidean space. Consider a vector $\mathbf{d} = (d_1, d_2, d_3)$ in $\mathbb{R}^3$ and a smooth function $f : \mathbb{R}^3 \to \mathbb{R}$, then the derivative of f at $p \in \mathbb{R}^3$ along the direction $\mathbf{d}$ is defined as:

$$\mathbf{d}_p f = \frac{d}{dt}(f(p + t\mathbf{d}))|_{t=0} = (d_1, d_2, d_3) \cdot (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}),$$

or in another way, one sees

$$\mathbf{d}_p f = (d_1 \frac{\partial}{\partial x_1} + d_1 \frac{\partial}{\partial x_1} + d_1 \frac{\partial}{\partial x_1})f.$$

The derivative $\mathbf{d}_p$ defined above satisfies linearity requirement and Leibniz rule, therefore it is a tangent vector at p in $\mathbb{R}^3$.

The union $T(M) = \cup_{p \in M} T_p(M)$, with a suitable C^∞ structure of a manifold, is called the tangent bundle of M. A C^∞ vector field X on M is a mapping from M to $T(M)$, which assigns to each point p of M a vector $X_p \in T_p(M)$ whose components in the frames of any local coordinates (U, ϕ) are C^∞ functions on the domain U of the coordinates.

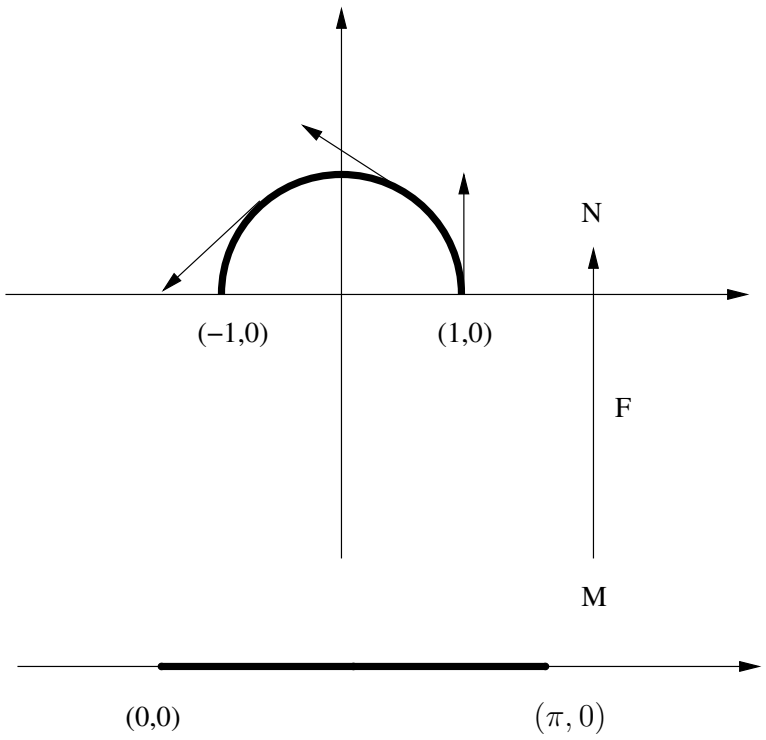


Figure 2.2. An example of the differential of a C^∞ mapping

Let $F : M \to N$ be a C^∞ map of manifolds. For $p \in M$, $f \in C^\infty(F(p))$, the differential of F at p is the homeomorphism $F_* : T_p(M) \to T_{F(p)}(N)$ defined by

$$F_*(X_p)f = X_p(f \circ F).$$

F_* is also denoted by dF. When $F : M \to N$ is the identity, F_* is the identity isomorphism.

Example 2.1. In this example, let $M = [0, \pi]$, $N = \{(\cos t, \sin t) \mid 0 \leq t \leq \pi\}$, and F is defined as a mapping from M to N with $F(t) = (\cos t, \sin t)$, where $0 \leq t \leq \pi$ (See Figure 2.2). Let $X(t)$ be a vector field on M such that $X(t) = 1$ for $0 \leq t \leq \pi$, then the $Y(\cos t, \sin t) = F_*(X(t)) = (-\sin t, \cos t)$.

For the tangent bundle, there is a natural projection $\pi : T(M) \to M$ taking the vector $X_p \in T(M)$ to p. $\pi^{-1}(p) = T_p(M)$ is called the fiber at p. For a vector field X as a mapping $X : M \to T(M)$, we have $\pi \circ X = i_M$, the identity on M. Then a vector field on M is a section of $T(M)$. Let $T_p^*(M)$ denote the dual space of $T_p(M)$. The union $T^*(M) = \cup_{p \in M} T_p^*(M)$ with the structure of a manifold is called the cotangent bundle or dual tangent bundle of M. An element of $T_p^*(M)$ is called a covector at p. A section of $T_p^*(M)$

$$\omega : M \to T^*(M)$$

$$p \to \omega_p \in T_p^*(M)$$

is called a covector field, or 1-form on M. If M is a vector field on an open set $U \subset M$, and ω is a covector field on M, we can define a function $\omega(X) : U \to \mathbb{R}$ by $\omega(X)(p) = \omega_p(X_p)$. Let (U, ϕ) be a coordinate neighborhood of M with coordinate frames $E_1, E_2, \cdots, E_n$. If for any (U, ϕ) the functions $\omega(E_i)$, $i = 1, 2, \cdots, n$, are C^∞, we call ω a C^∞ covector field.

Let $F : M \to N$ be a C^∞ mapping and suppose $p \in M$. Then there is an induced linear map $F_* : T_p(M) \to T_{F(p)}(N)$. Suppose $\omega_{F(p)} \in T_{F(p)}^*(N)$, $X_p \in T_p(M)$. F_* determines a linear map

$$F^* : T_{F(p)}^*(N) \to T_p^*(M),$$

given by the formula

$$F^*(\omega_{F(p)})(X_p) = \omega_{F(p)}(F_*(X_p)).$$

In the following we will introduce the concept of tensor field on a vector space. Let V be a vector space of dimension n and V^* its dual space. A tensor Φ on V is a multilinear map

$$\Phi : \underbrace{V \times \cdots \times V}_{r} \times \underbrace{V^* \times V^*}_{s} \to \mathbb{R},$$

r is the tensor covariant order, and s is the contravariant order. With the natural definitions of linear operation, the set of all tensors of order (r, s) on V form a vector space of dimension n^{r+s}. When $s = 0$, the corresponding is called covariant tensors of order r. It is easy to extend these ideas to manifolds if we let V be the tangent space of a manifold at each point. A C^∞-covariant tensor field of order r on a C^∞ manifold M is a mapping Φ which assigns to each $p \in M$ an element Φ_p of $T_p^*(M)$ and which has the additional property that given any C^∞ vector fields $X_1, X_2, \cdots, X_r$ on an open subset U of M, then $\Phi(X_1, X_2, \cdots, X_r)$ is a C^∞ function on U, defined by $\Phi(X_1, \cdots, X_r)(p) = \Phi_p(X_{1p}, \cdots, X_{rp})$. We denote by $\mathcal{T}^r(M)$ the set of all C^∞ covariant tensor fields of order r on M.

If $r = 1$, the covariant tensor fields are the covector fields. If $r = 2$, they are the so-called field of bilinear forms, which are important in Riemannian geometry.

Definition 2.2. A differentiable manifold M is called Riemannian manifold if there is a field of symmetric, positive definite, bilinear form g defined on M, here g is called the Riemannian metric.

A very simple example of a Riemannian manifold is Euclidean space $\mathbb{R}^n$. Note that in this case, the tangent space of $\mathbb{R}^n$ is identical to $\mathbb{R}^n$ itself. For any two vectors $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ and $\mathbf{y} = (y_1, y_2, \cdots, y_n)$, we define

$$g(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}^T = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

It can be checked that g, as defined above, is a symmetric, postive definite bilinar form.

In integral geometry, the alternating covariant tensor is for us the most interesting tensor. Let δ denote a permutation of $(1, 2, \cdots, r)$ with

$$(1, 2, \cdots, r) \to (\delta(1), \delta(2), \cdots, \delta(r)).$$

Let $\text{sgn}\,(\delta) = +1$ if δ is an even permutation of $(1, 2, \cdots, r)$ and let $\text{sgn}\,(\delta) = -1$ otherwise. $\Phi \in \mathcal{T}^r(V)$ is symmetric if

$$\Phi(V_1, V_2, \cdots, V_r) = \Phi(V_{\delta(1)}, V_{\delta(2)}, \cdots, V_{\delta(r)})$$

for every $V_1, V_2, \cdots, V_r$ and every permutation δ, and Φ is alternating if

$$\Phi(V_1, V_2, \cdots, V_r) = \text{sgn}\,(\delta)\Phi(V_{\delta(1)}, V_{\delta(2)}, \cdots, V_{\delta(r)})$$

for every $V_1, V_2, \cdots, V_r$ and every permutation δ. Let $\wedge^r(V)$ denote the set of alternating covariant tensors of order r.

Consider a transformation on the vector space $\mathcal{T}^r(V)$

$$\mathcal{A} : \mathcal{T}^r(V) \to \mathcal{T}^r(V),$$

which is given by the formula

$$(\mathcal{A}\Phi)(V_1, V_2, \cdots, V_r) = \frac{1}{r!} \sum_{\delta} \text{sgn}\,\delta\ \Phi(V_{\delta(1)}, V_{\delta(2)}, \cdots, V_{\delta(r)}).$$

One checks that for any tensor $\Phi \in \mathcal{T}^r(V)$, $\mathcal{A}\Phi$ is alternating. Then $\mathcal{A}$ is called the alternating transformation. It can be checked that $\mathcal{A}(\mathcal{T}^r(V)) = \wedge^r(V)$.

Assume that $\wedge^0(V) = \mathcal{R}$, then $\wedge^1(V) = \mathcal{T}^1(V)$ and $\wedge^r(V)$ is properly contained in $\mathcal{T}^r(V)$ for $r > 1$. Consider the direct sum

$$\wedge(V) = \wedge^0(V) \oplus \wedge^1(V) \oplus \cdots \wedge^r(V) \oplus \cdots .$$

Let $\phi \in \wedge^r(V)$ and $\psi \in \wedge^s(V)$, the exterior product of ϕ and ψ, denoted by $\phi \wedge \psi$, is defined by

$$\phi \wedge \psi = \frac{(r+s)!}{r!s!}\mathcal{A}(\phi \otimes \psi).$$

One checks that the exterior product is bilinear, skew-commutative and associative. Let $\phi, \psi \in \wedge(V)$,

$$\phi = \phi_1 + \phi_2 + \cdots + \phi_k, \qquad \phi_i \in \wedge^{r_i}(V), \quad i = 1, 2, \cdots, k,$$

$$\psi = \psi_1 + \psi_2 + \cdots + \psi_k, \qquad \psi_i \in \wedge^{r_i}(V), i = 1, 2, \cdots, l.$$

The exterior product of ϕ and ψ is defined by

$$\phi \wedge \psi = \sum_{i=1}^{k} \sum_{j=1}^{l} \phi_i \wedge \psi_j.$$

$\wedge(V)$ with the exterior product defined above is an associative algebra over $\mathcal{R}$. It is called the exterior algebra or Grassman algebra over V.

In the smiple case that $V = \mathbb{R}^2$. All the elements in $\wedge(\mathbb{R}^2)$ are in fact smooth functions on $\mathbb{R}^2$. Every element (1-form) in $\wedge(\mathbb{R}^2)$ can be written as

$$f(x, y)dx + g(x, y)dy,$$

where f and g are two smooth functions on $\mathbb{R}^2$. The $\wedge$ product of two 1-forms can be calculated as follows:

$$(f(x, y)dx + g(x, y)dy) \wedge (f_1(x, y)dx + g_1(x, y)dy)$$

$$= f(x, y)f_1(x, y)dx \wedge dx + f(x, y)g_1(x, y)dx \wedge dy + g(x, y)dy \wedge dx + g(x, y)dy \wedge dy.$$

Since $\wedge$ product is anti-symmetric, we have $dx \wedge dx = 0$, $dy \wedge dy = 0$, and $dx \wedge dy = -dy \wedge dx$. Consequently

$$(fdx + gdy) \wedge (f_1dx + g_1)dy) = (fg_1 - gf_1)dx \wedge dy.$$

Therefore the result is a 2-form in $\wedge^2(\mathbb{R}^2)$.

The concept of tensor can be extended to the case on a C^∞ manifold M. An alternating covariant tensor field of order r on M is called an exterior differential form of degree r, or simply r-form. A C^∞ function on M is called 0-form. $\wedge^r(M)$ denotes the set of r-forms. $\wedge^n(M)$ denotes the set of volume elements of M. We can similarly define $\wedge(M)$ which is called the exterior algebra on a manifold M.

For a differential function on the Euclidean space $\mathbb{R}^n$, derivative and differential can be defined. Such concepts can be generalized to a general Riemannian manifold.

Let M be an n-dimensional differentiable manifold and $T(M)$ be the Lie algebra of C^∞ vector fields on M. A function

$$\bigtriangledown : T(M) \times T(M) \to T(M), (X, Y) \to \bigtriangledown_X Y$$

with the properties

$$\bigtriangledown_{fX+gY} Z = f \bigtriangledown_X Z + g \bigtriangledown_Y Z,$$

$$\bigtriangledown_X (aY + bZ) = a \bigtriangledown_X Y + b \bigtriangledown_X Z,$$

$$\bigtriangledown_X (fY) = X(f)Y + f \bigtriangledown_X Y,$$

where $f, g : M \to \mathbb{R}$, $a, b \in \mathbb{R}$, $X, Y, Z \in T(M)$, is called linear connection or covariant derivative on M.

Theorem 2.3. *On a Riemannian manifold (M, g) there exists a unique symmetric linear connection $\bigtriangledown$ with the property*

$$\bigtriangledown_X (g(Y, Z)) = g(\bigtriangledown_X Y, Z) + g(Y, \bigtriangledown_X Z), \qquad \forall X, Y, Z \in T(M),$$

called the Riemannian connection.

Consider two points p and q in the Euclidean space $\mathbb{R}^n$. The straight line segment connecting p and q is the curve which has the shortest distance between these two points over all the possible curves connecting p and q. Such a line segment is called the geodesic from p to q in $\mathbb{R}^n$. One can extend the definition of geodesic to a Riemannian manifold. Let (M, g) be connected n-dimensional Riemannian manifold and I an interval in $\mathbb{R}$. A C^∞ curve $\gamma : I \to M$ whose acceleration vector field

$$\frac{\bigtriangledown}{dt} \frac{d\gamma}{dt}$$

vanishes identically is called a geodesic. It can be shown that the geodesic defined above generally is the shortest curve which connects two points, as in the Euclidean space case.

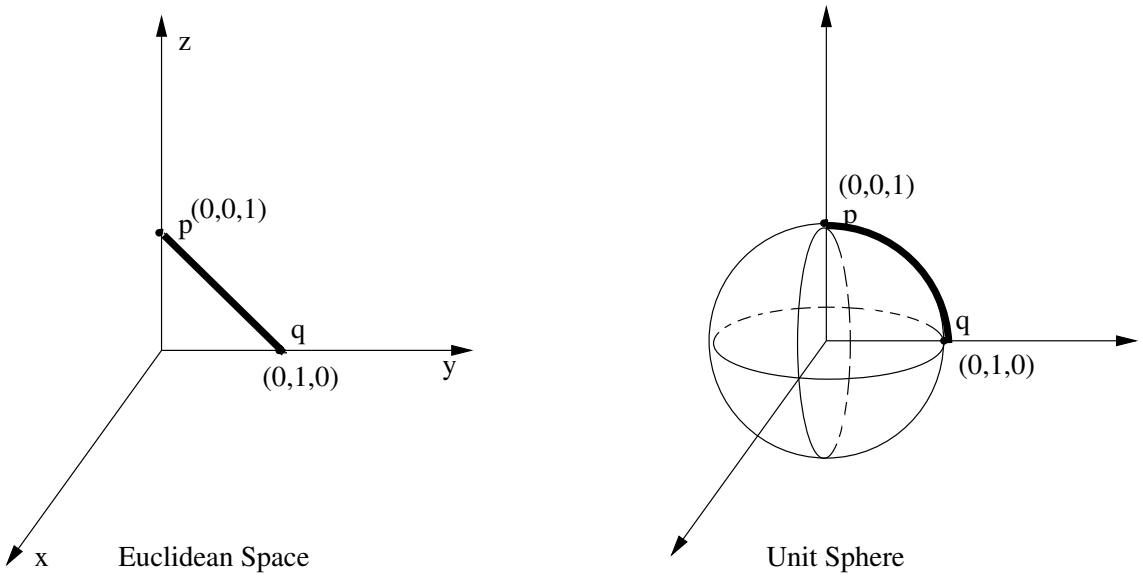


Figure 2.3. Geodesics connecting the same two point in different metric

Figure 2.3 illustrates the concept of geodesic. For two pionts $p = (0, 0, 1)$ and $q = (0, 1, 0)$ in Euclidean space $\mathbb{R}^3$, the geodesic connecting p and q is the line segment $(0, t, 1-t)$ with $0 \leq t \leq 1$. Meanwhile p and q can be viewed as two points on the unit sphere $x^2 + y^2 + z^2 = 1$, which is a Riemannian manifold with induced metric from $\mathbb{R}^3$. In terms of the metric on the unit sphere, the geodesic from p to q is described by the arc $(0, \cos t, \sin t)$ with $0 \leq t \leq \pi/2$.

2.2 Introduction to Lie Theory

According to the previous section, the space $\mathbb{R}^n$ is a C^∞ manifold. At the same time it is an abelian group with group operation. The group algebraic and differentiable structures are related by

$$(x, y) \mapsto x + y(\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n);$$

$$x \mapsto -x(\mathbb{R}^n \to \mathbb{R}^n)$$

and the above mapping are C^∞. Now let G be a group which is at the same time a differentiable manifold. For $x, y \in G$ let xy denote their product and x^{-1} the inverse of x.

Definition 2.4. G is a Lie group provided that the mapping of $G \times G \to G$ defined by $(x, y) \mapsto xy$ and the mapping of $G \to G$ defined by $x \mapsto x^{-1}$ are both C^∞ mappings.

Examples of Lie groups:

1. The Euclidean space $\mathbb{R}^n$ is a Lie group under the vector addition.
2. The non-zero complex numbers $\mathbb{C}^*$ forms a Lie group under multiplication.
3. The unit circle $S^1 \subset \mathbb{C}^*$ is a Lie group with the multiplication induced from $\mathbb{C}^*$.
4. The matrix group is a Lie group. Particularly $U(n)$ is a Lie group.

Let G be a Lie group. For each fixed $g_0 \in G$ we have two differentiable mappings $G \to G$:

1. the left translation $L_{g_0} : g \mapsto g_0 g$;
2. the right translation $R_{g_0} : g \mapsto g g_0$.

with differentials $(L_{g_0})_*$ and $(R_{g_0})_*$, or denoted by dL_{g_0} and dR_{g_0}. A vector field X on G is said to be left invariant if for every g, $g_0 \in G$, we have

$$d(L_{g_0})X_g = X_{g_0 g}.$$

The right invariant vector field can be similarly defined. A 1-form ω is left invariant if for all g, $g_0 \in G$

$$L_{g_0}^* \omega(g_0 g) = \omega(g).$$

The right invariant 1-form is similarly defined.

Definition 2.5. Any set of n linearly independent left invariant 1-forms are called Maurer-Cartan forms on G.

If G is of dimension n, then all the left invariant n-forms are equal up to a constant factor. Thus they are called volume elements.

Now we construct the Maurer-Cartan forms for matrix groups. Let G be a group of matrices $g = (g_{ij})_{k\times k}$. Consider the matrix

$$\Omega = g^{-1}dg = (\omega_{ij}),$$

that is

$$\omega_{ij} = \sum_{h=1}^{k} \gamma_{ih} dg_{hj},$$

where $g^{-1} = (\gamma_{ij})$ is the inverse matrix of g. Since

$$\Omega(g_0 g) = (g_0 g)^{-1} d(g_0 g) = g^{-1} g_0^{-1} g_0 dg = \Omega(g),$$

ω_{ij} are the left invariant 1-forms. If G is of dimension n, then there exist exactly n linear independent invariant forms among the k^2 1-forms. We can select a set of Maurer-Cartan forms for G.

Example 2.6. Consider the 2-dimensional matrix G of elements

$$g = \begin{pmatrix} x & 0 \\ e^y & x \end{pmatrix}, x \neq 0.$$

The group manifold in $\mathbb{R}^2$ deleted the line $x = 0$. It is easy to see that

$$\Omega = g^{-1}dg = \begin{pmatrix} \frac{dx}{x} & 0 \\ \frac{e^y}{x^2}dx + \frac{e^y}{x}dy & \frac{dx}{x} \end{pmatrix} = \begin{pmatrix} \omega_1 & 0 \\ \omega_2 & \omega_1 \end{pmatrix}.$$

Therefore a set of Maurer-Cartan forms is

$$\omega_1 = \frac{dx}{x}, \omega_2 = \frac{e^y}{x^2}dx + \frac{e^y}{x}dy.$$

Let $\{\omega_i\}$ be a set of Maurer-Cartan 1-forms of the Lie group G,

$$d_L G = \omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_n$$

is a left invariant n-form. Up to a constant factor, it is the unique left invariant n-form on G. The n-form $d_L G$ is called the left invariant volume elements of G. In integral geometry it is known as the kinematic density of G. The right invariant volume element of G is similarly defined as

$$d_R G = \bar{\omega}_1 \wedge \bar{\omega}_2 \wedge \cdots \wedge \bar{\omega}_n,$$

where $\{\bar{\omega}_i\}$ is a basis for the space of the right invariant 1-forms. A Lie group G is called unimodular if its left invariant volume element is also right invariant.

Theorem 2.7. *The Lie groups of finite volume are unimodular. Consequently compact Lie groups are unimodular.*

2.3 Basic Concepts in Algebraic Geometry

In this section, k will denote a fixed algebraically closed field. Affine n-space over k, denoted by $\mathbb{A}_k^n$ or simply by $\mathbb{A}^n$, is defined to be the set of all n-tuples of elements in k. An element $P \in \mathbb{A}^n$ will be called a point, and if $P = (a_1, a_2, \cdots, a_n)$ with $a_i \in k$, then the a_i's will be called the coordinates of P.

Let $\mathbb{A} = k[x_1, x_2, \cdots, x_n]$ be the polynomial ring in n variables over k. If we define $f(P) = f(a_1, a_2, \cdots, a_n)$, where $f \in \mathbb{A}$ and $P \in \mathbb{A}^n$, the elements of $\mathbb{A}$ can be viewed as functions from the affine n-space to k. Thus if $f \in \mathbb{A}$ is a polynomial, we can talk about the set of zeros of f, namely $Z(f) = \{P \in \mathbb{A}^n | f(P) = 0\}$. More generally, if T is any subset of $\mathbb{A}$, we define the zero set of T to be the common zeros of all the elements of T, namely

$$Z(T) = \{P \in \mathbb{A}^n | f(P) = 0 \text{ for all } f \in T\}.$$

Clearly if a is the ideal of $\mathbb{A}$ generated by T, then $Z(T) = Z(a)$. Furthermore, since $\mathbb{A}$ is a Noetherian ring, any ideal a has a finite set of generators $f_1, f_2, \cdots, f_r$.

Thus $Z(T)$ can be expressed as the common zeros of the finite set of polynomials $f_1, f_2, \cdots, f_r$.

We call a subset Y of $\mathbb{A}^n$ an algebraic set if there exists a subset $T \subset \mathbb{A}$ such that $Y = Z(T)$. One can easily verify that the union of two algebraic sets is an algebraic set. The intersection of any family of algebraic sets is an algebraic set. The empty set and the whole space are algebraic sets. Thus a topology on $\mathbb{A}^n$ can be defined using algebraic sets. We define the Zariski topology on $\mathbb{A}^n$ by taking the open subsets to be the complements of the algebraic sets. This is a topology, because the intersection of two open sets is open, and the union of any family of open sets is open. Furthermore, the empty set and the whole space are both open.

Example 2.8. Let us consider the Zariski topology on the affine line $\mathbb{A}^1$. Every ideal in $R = k[x]$ is principal, so every algebraic set is the set of zeros of a single polynomial. Since k is algebraically closed, every nonzero polynomial $f(x)$ can be written $f(x) = c(x - a_1)(x - a_2)\cdots(x - a_n)$ with $c, a_1, a_2, \cdots, a_n \in k$. Then $Z(f) = a_1, a_2, \cdots, a_n$. Thus the algebraic sets in $\mathbb{A}^1$ are just the finite subsets and the whole space. Thus the open sets are the empty set and the complements of finite subsets.

A nonempty subset Y of a topological space X is irreducible if it cannot be expressed as the union $Y = Y_1 \cup Y_2$ of two proper subsets, each one of which is closed in Y. The empty set is not considered to be irreducible. For example, $\mathbb{A}^1$ is irreducible, because its only proper closed subsets are finite, yet it is finite (because k is algebraically closed, hence infinite). Any nonempty open subset of an irreducible space is irreducible and dense. If Y is an irreducible subset of X, then its closure $\bar{Y}$ in X is also irreducible.

Definition 2.9. An affine algebraic variety is an irreducible closed subset of $\mathbb{A}^n$ (with the induced topology).

Example 2.10. Consider

$$U(1) = \{z \in \mathbb{C} \mid |z| = 1\}.$$

In terms of real coordinates, $U(1)$ can be rewritten as

$$U(1) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

Further argument can show that $U(1)$ is also irreducible. Thus $U(1)$ is a real algebraic variety. Since $U(1)$ is a bounded and closed subset of $\mathbb{R}^2$, $U(1)$ is also compact. A complex Stiefel manifold $S(T, M)$ (see Chapter 6 for more details) is a real algebraic variety. In particular $U(n)$ is a connected real algebraic variety with real dimension n^2. We will investigate the relationship between a complex Stiefel manifold and $U(n)$ later on.

2.4 Optimization Techniques on a Riemannian Manifold

The steepest descent method and Newton's method on Euclidean space are used extensively in finding a local minimum of a target function. The idea can be generalized to a Riemannian manifold. In fact the method of steepest descent on a Riemannian manifold is conceptually identical to the method of steepest descent on Euclidean space. The algorithm consists of iterative search for a "better" point along the geodesic (See Figure 2.4). Each iteration involves a gradient computation and minimization along the geodesic determined by the gradient. In this section we restate the method of steepest descent described in the literature and provide an alternative formalism that will be useful in the development of Newton's method on a Riemannian mainfold.

Algorithm (the method of steepest descent) Let M be a complete Riemannian manifold with Riemannian metric g and Levi-Civita connection $\triangle$, and let $f \in C^\infty(M)$.

1. Select $p_0 \in M$, compute $G_0 = -(\text{grad}\, f)_{p_0}$, and set $i = 0$.

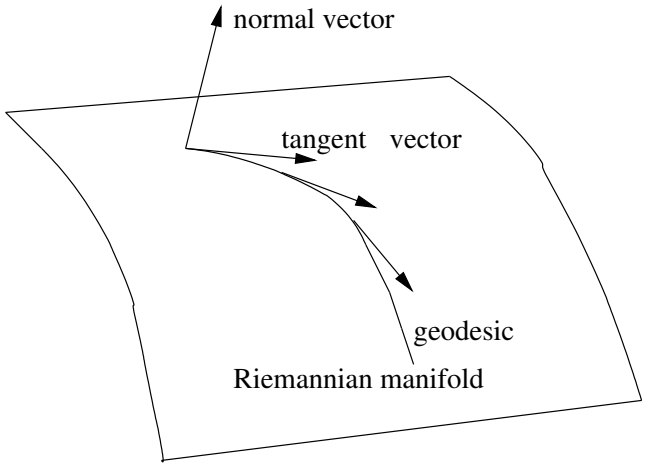


Figure 2.4. Local optimization methods on a Riemannian manifold

2. Compute λ_i such that

$$f(\exp_{p_i} \lambda_i G_i) \leq f(\exp_{p_i} \lambda G_i) \qquad \text{for} \quad \text{all} \ \ \lambda \geq 0.$$

3. Set

$$p_{i+1} = \exp_{p_i} \lambda_i G_i, \qquad G_{i+1} = -(\operatorname{grad} f)_{p_{i+1}},$$

increase i by 1, and go to step 1.

It is easy to verify that $< G_{i+1}, \tau G_i >= 0$, for $i \geq 0$, where τ is the parallelism with respect to the geodesic from p_i to p_{i+1}. By assumption,the function $\lambda \rightarrow f(\exp \lambda G_i)$ is minimized at λ_i. Therefore, we have $(d/dt)|_{t=0} f(\exp(\lambda_i + t)G_i) = df_{p_{i+1}}(\tau G_i) =< (\operatorname{grad} f)_{p_{i+1}}, \tau G_i >$. Thus the method of steepest descent on a Riemaniann manifold has the same deficiency as its counterpart on a Euclidean space, *i.e.*, it makes a ninety degree turn at every step.

Theorem 2.11. *If the above algorithm converges to a local minimum, it converges linearly.*

Algorithm (Newton's method) Let M be a complete Riemannian manifold with Riemannian structure G and Levi-Civita connection $\triangle$, and let μ be a C^∞ one-form on M.

1. Select $p_0 \in M$ such tat $(\triangle)_{p_0}$ is nondegenerate, and set $i = 0$.

2. Compute

$$H_i = -(\triangle\mu)_{p_i}^{-1}\mu_{p_i}p_{i+1} = \exp_{p_i} H_i,$$

   increase i by 1 and repeat.

Theorem 2.12. *If $(\triangle^2 f)_{\tilde{p}}$ is positive definite and the above algorithm converges to $\hat{p}$, then the algorithm converges quadratically to a local minimum of f.*

CHAPTER 3

UNITARY SPACE TIME CONSTELLATION DESIGN CRITERIA

Multiple antenna communication systems have been actively studied recently [3, 28, 29, 44]. By exploiting the temporal and spatial diversity at both transmitter side and receiver side, such systems can increase the channel capacity compared to single antenna communication systems, consequently promising more reliable data transmission for high rate applications.

As already discussed in Subsection 1.3.3, we investigate a multiple antenna communication system under Rayleigh flat fading assumptions and we assume fading is quasi-static over a time period of length T. Denote by M and N the number of transmit and receive antennas, respectively. Let ρ represent the expected signal-to-noise ratio (SNR) at each receiving antenna. For the system above, the basic equation between the received signal R, which is a $T \times N$ matrix, and the transmitted signal Φ, which is chosen from a $T \times M$ matrix constellation $\mathcal{V} = \{\Phi_1, \Phi_2, \cdots, \Phi_L\}$ (L is the constellation size), is given through:

$$R = \sqrt{\frac{\rho}{M}}\Phi H + W,$$

where the $M \times N$ matrix H accounts for the multiplicative complex Gaussian fading coefficients and the $T \times N$ matrix W accounts for the additive white Gaussian noise. The entries $h_{m,n}$ of the matrix H as well as the entries $w_{t,n}$ of the matrix W are assumed to have a statistically independent complex normal distribution $\mathcal{CN}(0, 1)$.

One can verify that the transmission rate is determined by L and T:

$$\mathtt{R} = \frac{\log_2(L)}{T}.$$

When the fading coefficients are unknown to the transmitter however known to the receiver, it is proved [14, 45] that the channel capacity will increase linearly with $\min\{M, N\}$ (See Subsection 1.3.3 for the brief description). Since H is known to the receiver, for a received signal R the maximum likelihood (ML) decoder will take the following evaluation to resolve the most likely sent signal (codeword):

$$\hat{\Phi} = \arg\min_{\Phi \in \mathcal{V}} \left\| R - \sqrt{\frac{\rho}{M}} \Phi H \right\|,$$

where $\|\cdot\|$ represents Frobenius norm, *i.e.*

$$\|A\| = \sqrt{\mathrm{tr}\,(AA^*)}.$$

Let $P_{\Phi_l, \Phi_{l'}}$ denote the probability that ML decoder mistakes Φ_l for $\Phi_{l'}$. An upper bound of this error probability has been derived in [44]:

$$P_{\Phi_l, \Phi_{l'}} \leq \frac{1}{2} \prod_{m=1}^{M} \left[1 + \frac{\rho T}{4M} \delta_m^2 (\Phi_l - \Phi_{l'}) \right]^{-N}.$$

Through the analysis of the above upper bound, design criteria for designing codes for the coherent channel with ideal channel state information (CSI) are proposed in [44]:

The Rank Criterion: In order to achieve the maximum diversity MN, the matrix $B(\Phi_l, \Phi_l') = \Phi_l - \Phi_{l'}$ must have full rank for any codewords Φ_l and $\Phi_{l'}$. If $B(\Phi_l, \Phi_{l'})$ has minimum rank r over the set of two tuples of distinct codewords, then a diversity rN is achieved.

The Determinant Criterion: Suppose that a diversity benefit of rM is our target. The minimum of rth roots of the sum of determinants of all $r \times r$ principal cofactors

of $A(\Phi_l, \Phi_{l'}) = B(\Phi_l, \Phi_{l'})B^*(\Phi_l, \Phi_{l'})$ taken over all pairs of distinct codewords Φ_l and $\Phi_{l'}$ corresponds to the coding advantage, where r is the rank of $A(\Phi_l, \Phi_{l'})$.

Note that for the case when $T = M$ and full diversity is achieved, the above design criteria can be simplified as follows:

Construct a constellation of matrices $\mathcal{V} = \{\Phi_1, \Phi_2, \cdots, \Phi_L\}$ such that the diversity product [28]

$$\prod \mathcal{V} = \min_{l \neq l'} \frac{1}{2} |\det(\Phi_l - \Phi_{l'})|^{1/M}$$

is as large as possible.

In [29], Hochwald and Marzetta study unitary space-time modulation for a non-coherent channel. We will use the same notations as in the coherent case, so the basic equation will be still the same:

$$R = \sqrt{\frac{\rho}{M}}\Phi H + W,$$

however in non-coherent scenarios it is assumed that the receiver does not know the exact values of the entries of H (other than their statistical distribution). Another difference is that the signal constellation $\mathcal{V} := \{\Phi_1, \Phi_2, \cdots, \Phi_L\}$ has unitary constraints: $\Phi_k^*\Phi_k = I_M$ for $k = 1, 2, \cdots, L$. The last equation simply states that the columns of Φ_k form a "unitary frame", *i.e.*, the column vectors all have unit length in the complex vector space $\mathbb{C}^T$ and the vectors are pairwise orthogonal. The scaled matrices $\sqrt{T}\Phi_k$, represent the codewords used during the transmission.

The decoding task asks for the computation of the most likely sent code word Φ given the received signal R. If $A = (a_{i,j})$ then the Frobenius norm is defined through $\|A\| = \sqrt{\sum_{i,j} |a_{i,j}|^2}$. Under the assumption of the above model the maximum likelihood (ML) decoder will have to compute:

$$\Phi_{ML} = \arg \max_{\Phi_l \in \{\Phi_1, \Phi_2, \cdots, \Phi_L\}} \|R^*\Phi_l\|$$

for each received signal R. (See [29]).

Let $\delta_m(\Phi_l^*\Phi_{l'})$ be the m-th singular value of $\Phi_l^*\Phi_{l'}$. It has been shown in [29] that the pairwise probability of mistaking Φ_l for $\Phi_{l'}$ using maximum likelihood decoding satisfies:

$$\begin{aligned}
P_{\Phi_l,\Phi_{l'}} &= \text{Prob}\,(\text{ choose } \Phi_{l'} \mid \Phi_l \text{ transmitted }) \,(\rho) \\
&= \text{Prob}\,(\text{ choose } \Phi_l \mid \Phi_{l'} \text{ transmitted }) \,(\rho) \\
&= \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{4}{4w^2+1} \prod_{m=1}^{M} \left[1 + \frac{(\rho T/M)^2(1-\delta_m^2(\Phi_l^*\Phi_{l'}))}{4(1+\rho T/M)}(4w^2+1)\right]^{-N} dw \quad (3.1) \\
&\leq \frac{1}{2}\prod_{m=1}^{M}\left[1 + \frac{(\rho T/M)^2(1-\delta_m^2(\Phi_l^*\Phi_{l'}))}{4(1+\rho T/M)}\right]^{-N}. \quad (3.2)
\end{aligned}$$

It is a basic design objective to construct constellations $\mathcal{V} = \{\Phi_1, \ldots, \Phi_L\}$ such that the pairwise probabilities $P_{\Phi_l,\Phi_{l'}}$ are as small as possible. Mathematically we are dealing with an optimization problem with unitary constraints:

Minimize $\max_{l \neq l'} P_{\Phi_l,\Phi_{l'}}$ with the constraints $\Phi_i^*\Phi_i = I$ where $i = 1, 2, \cdots, L$.

Formula (3.2) is sometimes referred to as "Chernoff's bound". This formula is easy to work with, the exact formula (3.1) is in general not easy to work with, although it could be useful in the numerical search of good constellations as well. Researchers have been searching for constructions where the maximal pairwise probability of $P_{\Phi_l,\Phi_{l'}}$ is as small as possible. Of course the pairwise probabilities depend on the chosen signal to noise ratio ρ, and the construction of constellations has therefore to be optimized for particular values of the SNR.

The design objective is slightly simplified if one assumes that transmission operates at high SNR situations. In [28], a design criterion for high SNR is presented and the problem has been converted to the design of a finite set of unitary matrices whose diversity product is as large as possible. In this special situation several researchers [3, 42, 41, 40] came up with algebraic constructions and we will say more about this later.

3.1 The Diversity Function, the Diversity Product and the Diversity Sum

In this dissertation we will be concerned with the construction of constellations where the right hand sides in (3.1) and (3.2), maximized over all pairs l, l' are as small as possible for fixed numbers of T, M, N, L. As already mentioned this task depends on the signal to noise ratio at which the system is operating. For this purpose we define the *exact diversity function* dependent on the constellation $\mathcal{V} = \{\Phi_1, \ldots, \Phi_L\}$ and a particular SNR ρ through:

$$\mathcal{D}_e(\mathcal{V}, \rho) := \max_{l \neq l'} \text{Prob}\,(\text{ choose } \Phi_{l'} \mid \Phi_l \text{ transmitted }) (\rho) \tag{3.3}$$

For a particular constellation with a large number L of elements, with many transmit and receive antennas, the function $\mathcal{D}_e(\mathcal{V}, \rho)$ is very difficult to compute. Indeed for each pair $\Phi_{l'}, \Phi_l$ it is required to compute the singular values of the $M \times M$ matrix $\Phi_l^* \Phi_{l'}$, one has to evaluate up to $L(L-1)/2$ integrals of the form (3.1), and this has to be done for each value of ρ. Although this task is formidable it can be done in cases where T, M, L are all in the single digits using *e.g.* Maple.

Using Chernoff's bound (3.2) we define a simplified function called the *diversity function* through:

$$\mathcal{D}(\mathcal{V}, \rho) := \max_{l \neq l'} \frac{1}{2} \prod_{m=1}^{M} \left[1 + \frac{(\rho T/M)^2}{4(1 + \rho T/M)} (1 - \delta_m^2(\Phi_l^* \Phi_{l'})) \right]^{-N}. \tag{3.4}$$

The computation of $\mathcal{D}(\mathcal{V}, \rho)$ does not require the evaluation of an integral and the computation requires essentially the computation of $ML(L-1)/2$ singular values. The singular values $\delta_m(\Phi_l^* \Phi_{l'})$ are by definition all real numbers in the interval $[0, 1]$ as we assume that the columns of $\Phi_l, \Phi_{l'}$ form both orthonormal frames. The functions $\mathcal{D}_e(\mathcal{V}, \rho)$ and $\mathcal{D}(\mathcal{V}, \rho)$ are the smallest if the singular values $\delta_m(\Phi_l^* \Phi_{l'})$ are as small as possible. These numbers are all equal to zero if and only if the column spaces of $\Phi_l, \Phi_{l'}$ are pairwise perpendicular. We call such a constellation

fully orthonormal. Since the columns of Φ_l generate an M-dimensional subspace this can only happen if $L \leq T/M$. On the other hand if $L \leq T/M$ it is easy to construct a constellation where the singular values of $(\Phi_l^* \Phi_{l'})$ are all zero. Just pick LM different columns from a $T \times T$ unitary matrix. Figure 3.1 depicts the functions $\mathcal{D}_e(\mathcal{V}, \rho)$ and $\mathcal{D}(\mathcal{V}, \rho)$ for a fully orthonormal constellation with $T = 10$ and $M = N = 2$. Fully orthonormal constellations can achieve good diversity, however, they impose a restriction to the transmission rate.

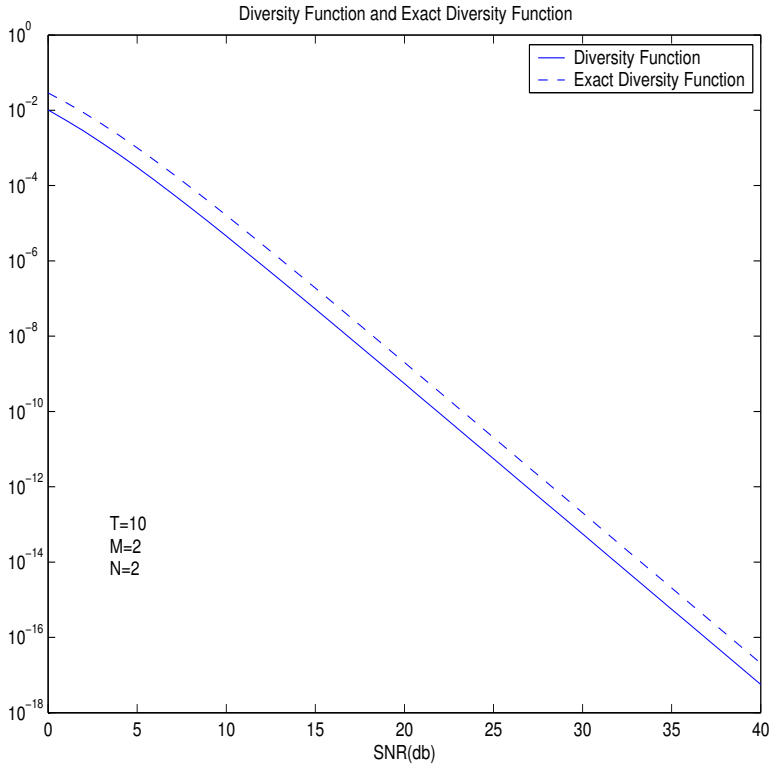


Figure 3.1. Diversity function $\mathcal{D}(\mathcal{V}, \rho)$ and exact diversity function $\mathcal{D}_e(\mathcal{V}, \rho)$ of a fully orthonormal constellation

In order to study the function $\mathcal{D}(\mathcal{V}, \rho)$ more carefully let

$$\tilde{\rho} := \frac{(\rho T/M)^2}{4(1 + \rho T/M)}. \tag{3.5}$$

In some small interval $[\rho_1, \rho_2]$ the maximum in (3.4) is achieved for some fixed indexes l, l' and in terms of $\tilde{\rho}$ the function $\mathcal{D}(\mathcal{V}, \rho)$ is of the form:

$$\mathcal{D}(\mathcal{V}, \tilde{\rho}) = \frac{1}{2\left(1 + c_1\tilde{\rho} + \cdots + c_M\tilde{\rho}^M\right)^N},$$

where the coefficients $c_1, \ldots, c_M$ depend on the particular constellation and on the chosen interval $[\rho_1, \rho_2]$. For an interval close to zero the dominating term will be the coefficient c_1. Up to some factor this term will define the *diversity sum* (DS) of the constellation. When $\tilde{\rho} >> 0$ (ρ is large enough number) then the dominating term will be the coefficient c_M and up to some scaling this term will define the *diversity product* (DP) of the constellation. A constellation will have a small diversity function for small values of ρ (and presumably performs well in this range) when the constellation is chosen having a large diversity sum. A constellation will have a small diversity function for large values of ρ (and presumably performs well in this range) when the constellation is chosen having a large diversity product. In the next two subsections we will study the limiting behavior of $\mathcal{D}(\mathcal{V}, \rho)$ as ρ goes to zero and to infinity.

3.2 The Diversity Product: Design Criterion for High SNR

When the SNR ρ is very large then $\mathcal{D}(\mathcal{V}, \rho)$ can be approximated via:

$$\mathcal{D}(\mathcal{V}, \rho) \simeq \max_{l \neq l'} \frac{1}{2}\left(\frac{(\rho T/M)^2}{4(1+\rho T/M)}\right)^{-NM} \prod_{m=1}^{M} \frac{1}{(1 - \delta_m^2(\Phi_l^*\Phi_{l'}))^N}. \tag{3.6}$$

It is the design objective to construct a constellation $\Phi_1, \Phi_2, \cdots, \Phi_n$ such that

$$\min_{l \neq l'} \prod_{m=1}^{M} (1 - \delta_m^2(\Phi_l^*\Phi_{l'}))$$

is as large as possible. This last expression defines in essence the diversity product. In order to compare different dimensional constellations it is customary to use the definition:

Definition 3.1. (See [28]) The *diversity product* of a unitary constellation $\mathcal{V}$ is defined as

$$\prod \mathcal{V} = \min_{l \neq l'} \left(\prod_{m=1}^{M} (1 - \delta_m(\Phi_l^* \Phi_{l'})^2) \right)^{\frac{1}{2M}}.$$

An important special case occurs when $T = 2M$. In this situation it is customary to represent all unitary matrices Φ_k in the form:

$$\Phi_k = \frac{\sqrt{2}}{2} \begin{pmatrix} I \\ \Psi_k \end{pmatrix}. \tag{3.7}$$

Note that by definition of Φ_k the matrix Ψ_k is a $M \times M$ unitary matrix. The diversity product as defined in Definition 3.1 has then a nice form in terms of the unitary matrices. For this let λ_m be the mth eigenvalue of a matrix, then

$$1 - \delta_m^2(\Phi_{l'}^* \Phi_l) = \frac{1}{4} \lambda_m(2I_M - \Phi_l^* \Phi_{l'} - \Phi_{l'}^* \Phi_l) = \frac{1}{4} \delta_m^2(I_M - \Psi_{l'}^* \Psi_l) = \frac{1}{4} \delta_m^2(\Psi_{l'} - \Psi_l).$$

So we have

$$\prod_{m=1}^{M} (1 - \delta_m^2(\Phi_{l'}^* \Phi_l))^{\frac{1}{2M}} = \frac{1}{2} \prod_{m=1}^{M} \delta_m(\Psi_{l'} - \Psi_l)^{\frac{1}{M}} = \frac{1}{2} |\det(\Psi_{l'} - \Psi_l)|^{\frac{1}{M}}.$$

When $T = 2M$ and the constellation $\mathcal{V}$ is defined as above, then the formula of the diversity product assumes the simple form:

$$\prod \mathcal{V} = \frac{1}{2} \min_{0 \leq l < l' \leq L} |\det(\Psi_l - \Psi_{l'})|^{\frac{1}{M}}. \tag{3.8}$$

We call a constellation $\mathcal{V}$ a fully diverse constellation if $\prod \mathcal{V} > 0$. A lot of efforts have been taken to construct constellations with large diversity product. (See *e.g.* [28, 32, 20, 19, 41, 40, 42]). For the particular situation $T = 2M$ with special form (3.7) the design asks for the construction of a discrete subset $\mathcal{V} = \{\Psi_1, \ldots, \Psi_L\}$ of the set of $M \times M$ unitary matrices $U(M)$. When this discrete subset has the structure of a discrete subgroup of $U(M)$ then the condition that $\mathcal{V}$ is fully diverse is equivalent to the condition that the identity matrix is the only element of $\mathcal{V}$

having an eigenvalue of 1. In other words the constellation $\mathcal{V}$ is required to operate fixed point free on the vector space $\mathbb{C}^M$. Using a classical classification result of fixed point free unitary representations by Zassenhaus [47], Shokrollahi *et. al.* [41, 40] were able to study the complete list of fully diverse finite group constellations inside the unitary group $U(M)$. Some of these constellations have the best known diversity product for given fixed parameters M, N, L. Unfortunately the possible configurations derived in this way is somewhat limited. The constellations are also optimized for the diversity product and as we demonstrate in this dissertation for unitary space time modulation maybe attention should be given to the diversity sum.

In most of the literature mentioned above researchers focus their attention to constellations having the special form (3.7). Unitary differential modulation [28] is used to avoid sending the identity (upper part of every element in the constellation) redundantly. A differential modulation scheme is discussed in [28] and this special form is used. In this scheme, one does not send the identity matrix (the upper part of the signal) every time. Instead of sending

$$\begin{pmatrix} I_M \\ \Psi_1 \end{pmatrix}, \begin{pmatrix} I_M \\ \Psi_2 \end{pmatrix}, \begin{pmatrix} I_M \\ \Psi_3 \end{pmatrix}, \cdots,$$

one sends

$$I_M, \Psi_1, \Psi_2\Psi_1, \Psi_3\Psi_2\Psi_1, \cdots.$$

This increases the transmission rate by a factor of 2 to:

$$\mathtt{R} = \frac{\log_2(L)}{M} = 2\frac{\log_2(L)}{T}.$$

Let Ψ_τ and R_τ denote the sent and received signals at time τ, respectively, the ML decoder of the above differential space time modulation scheme will have to

compute:

$$\hat{\Psi}_\tau = \arg\min_{\Psi\in\mathcal{V}} \|R_\tau - \Psi R_{\tau-1}\|.$$

For the constellations with the special form (3.7), the pairwise error probability satisfies:

$$P_{\Phi_l,\Phi_{l'}} \leq \frac{1}{2}\prod_{m=1}^{M}\left[1+\frac{\rho^2\delta_m^2(\Psi_l-\Psi_{l'})}{4(1+2\rho)}\right]^{-N}. \tag{3.9}$$

Because of this reason we will also focus ourselves in the latter part of the chapter to the special form (3.7) as well. Nonetheless it will become obvious that the numerical techniques also work in the general situation.

3.3 The Diversity Sum: Design Criterion for Low SNR

As we mentioned before a constellation with a large diversity sum will have a small diversity function at small values of the signal to noise ratio. This is particularly suitable when the system operates in a very noisy environment. When ρ is small, using Formula (3.5), one has the following expansion:

$$\begin{aligned}\prod_{m=1}^{M}[1+\frac{(\rho T/M)^2}{4(1+\rho T/M)}(1-\delta_m^2(\Phi_l^*\Phi_{l'}))] &= \prod_{m=1}^{M}[1+\tilde{\rho}(1-\delta_m^2(\Phi_l^*\Phi_{l'}))] \\ &= 1+\tilde{\rho}\sum_{m=1}^{M}(1-\delta_m^2(\Phi_l^*\Phi_{l'}))+O(\tilde{\rho}^2).\end{aligned}$$

When $\rho\to 0$, *i.e.* $\tilde{\rho}\to 0$, we can omit the higher order terms $O(\tilde{\rho}^2)$ and the upper bound of $P_{\Phi_l,\Phi_{l'}}$ requires that

$$\sum_m(1-\delta_m^2(\Phi_l^*\Phi_{l'})) = M - \|\Phi_l^*\Phi_{l'}\|^2$$

is large. In order to lower the pairwise error probability, it is the objective to make $\|\Phi_l^*\Phi_{l'}\|^2$ as small as possible for every pair of l, l'. It follows that at high SNR, the probability primarily depends on $\prod_{m=1}^{M}(1-\delta_m^2(\Phi_l^*\Phi_{l'}))$, but at low SNR, the probability primarily depends on $\sum_{m=1}^{M}(1-\delta_m^2(\Phi_l^*\Phi_{l'}))$. In order to be able to compare the constellation of different dimensions, we define:

Definition 3.2. The *diversity sum* of a unitary constellation $\mathcal{V}$ is defined as

$$\sum \mathcal{V} = \min_{l \neq l'} \sqrt{1 - \frac{\|\Phi_l^* \Phi_{l'}\|^2}{M}}.$$

Again one has the important special case where $T = 2M$ and the matrices Φ_k take the special form (3.7). In this case one verifies that

$$\begin{aligned}\|\Phi_l^* \Phi_{l'}\|^2 &= \frac{1}{4}\|I + \Psi_l^* \Psi_{l'}\|^2 = \frac{1}{4}\mathrm{tr}\left((I + \Psi_{l'}^* \Psi_l)(I + \Psi_l^* \Psi_{l'})\right) \\ &= \frac{1}{4}\mathrm{tr}\left(2I + \Psi_{l'}^* \Psi_l + \Psi_l^* \Psi_{l'}\right) = \frac{1}{4}(4M - (2M - \mathrm{tr}\left(\Psi_{l'}^* \Psi_l + \Psi_l^* \Psi_{l'}\right))) \\ &= \frac{1}{4}(4M - \mathrm{tr}\left((\Psi_l - \Psi_{l'})^*(\Psi_l - \Psi_{l'})\right)) = \frac{1}{4}(4M - \|\Psi_l - \Psi_{l'}\|^2).\end{aligned}$$

For the form (3.7) the diversity sum assumes the following simple form:

$$\sum \mathcal{V} = \min_{l,l'} \frac{1}{2\sqrt{M}}\|\Psi_l - \Psi_{l'}\|. \tag{3.10}$$

Without mentioning the term the concept of diversity sum was used in [27]. Liang and Xia [32, p. 2295] explicitly defined the diversity sum in the situation when $T = 2M$ using equation (3.10). Definition 3.2 naturally generalizes the definition to arbitrary constellations.

We want to point out that the diversity sum is the design criterion only for unitary space time modulation. Hochwald and Marzetta [29] calculate the non-coherent space time channel capacity and indicate that unitary signal constellation are capacity achieving signal sets only for high SNR scenarios. For low SNR case, the transmitting power should be allocated asymmetrically, *i.e.*, unitary constellations are not capacity achieving in the first place. However unitary signal sets are easily manageable and one can take advantage of differential modulation technique [28] to speed up the transmission. Moreover, our simulation results indicate that codes with near optimal diversity sum tend to perform significantly better compared to the currently existing ones optimized for the diversity product for low and even

moderate SNR scenarios. So it is quite reasonable and more toward the practical use to construct unitary constellations with good diversity sum.

As the formulas make it clear the diversity sum and the diversity product are in general very different. There is however an exception: when $T = 4$, $M = 2$ and the constellation $\mathcal{V}$ is in the special (3.7). If in addition all the 2×2 matrices $\{\Psi_1, \ldots, \Psi_L\}$ are a subset of the special unitary group

$$SU(2) = \{A \in \mathbb{C}^{2\times 2} \mid A^*A = I \text{ and } \det A = 1\}$$

then it turns out that the diversity product $\prod \mathcal{V}$ and the diversity sum $\sum \mathcal{V}$ of such a constellation are the same. For this note that elements $\Psi_l, \Psi_{l'}$ of $SU(2)$ have the special form:

$$\Psi_l = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \ \Psi_{l'} = \begin{pmatrix} c & d \\ -\bar{d} & \bar{c} \end{pmatrix}.$$

Through a direct calculation one verifies that $\det(\Psi_l - \Psi_{l'}) = |a-c|^2 + |b-d|^2$ and $\|\Psi_l - \Psi_{l'}\|^2 = 2(|a-c|^2 + |b-d|^2)$. But this means that $\prod \mathcal{V} = \sum \mathcal{V}$ for constellations inside $SU(2)$.

CHAPTER 4

AN UPPER BOUND FOR THE DIVERSITY

4.1 Introduction

Let A be a matrix with complex entries. A^* denotes the conjugate transpose of A. Let $\| \ \|$ denote the Frobenius norm of a matrix. A square matrix A is called unitary if $A^*A = AA^* = I$, where I denotes the identity matrix. We denote by $U(n)$ the set of all $n \times n$ unitary matrices. $U(n)$ is a real algebraic variety and a smooth manifold of real dimension n^2. As defined in Chapter 3, a unitary space time constellation (or code) $\mathcal{V}$ is simply a finite subset of $U(n)$,

$$\mathcal{V} = \{A_1, A_2, \cdots, A_m\} \subset U(n).$$

We say $\mathcal{V}$ has dimension n and size m. Unitary space time codes have been intensely studied in recent years and we refer the interested readers to [3, 28, 29, 41] and the references of these papers. As explained in Chapter 3 the quality of a unitary space time code is governed by two important parameters, the diversity product and the diversity sum. Recall that the *diversity product* [28] of a unitary space time code $\mathcal{V}$ is defined through

$$\prod \mathcal{V} := \frac{1}{2} \min\{|\det(A-B)|^{\frac{1}{n}} \mid A, B \in \mathcal{V}, A \neq B\}.$$

The *diversity sum* [32] is defined as

$$\sum \mathcal{V} := \frac{1}{2\sqrt{n}} \min\{\|A-B\| \mid A, B \in \mathcal{V}, A \neq B\}.$$

$\mathcal{V}$ is called fully diverse if $\prod \mathcal{V} > 0$. As explained in Chapter 3, a space time code with large diversity sum tends to perform well at low signal to noise ratios whereas a code with a large diversity product tends to perform well at high signal to noise ratios. A major coding design problem is the construction of unitary space time codes where the diversity sum (or product) is optimal or near optimal inside the set of all the space time codes with the same parameters n, m.

The purpose of this chapter is to derive for n and m tight upper bounds for the diversity product $\prod \mathcal{V}$ and the diversity sum $\sum \mathcal{V}$. When $n = 1$ then trivially $|\det(A - B)| = \|A - B\|$ and it follows that $\sum \mathcal{V} = \prod \mathcal{V}$ in this situation. The following lemma states that for every space time code $\mathcal{V}$, $\sum \mathcal{V}$ is an upper bound for $\prod \mathcal{V}$ and by having an upper bound for $\sum \mathcal{V}$ we immediately also have an upper bound for $\prod \mathcal{V}$. The readers can find the statements about the relationship between $\prod \mathcal{V}$ and $\sum \mathcal{V}$ in [32], for completeness we include a detailed proof.

Lemma 4.1. *For any unitary space time code $\mathcal{V}$,*

$$\prod \mathcal{V} \leq \sum \mathcal{V}.$$

Proof. Let C be an $n \times n$ complex matrix with singular value decomposition

$$C = U \operatorname{diag}(c_1, c_2, \cdots, c_n) V,$$

where U, V are unitary matrices and $c_j \geq 0$ for $j = 1, 2, \cdots, n$ are the singular values of C. First we prove

$$\frac{1}{2} |\det(C)|^{\frac{1}{n}} \leq \frac{1}{2\sqrt{n}} \|C\|.$$

If $c_j = 0$ for some j, then the inequality is trivial. Hence we assume $c_j > 0$ for all j's. Because U, V are unitary matrices, it follows that

$$\frac{1}{2} |\det(C)|^{\frac{1}{n}} = \frac{1}{2} \left(\prod_{j=1}^{n} c_j \right)^{\frac{1}{n}}.$$

Similarly one verifies that

$$\frac{1}{2\sqrt{n}}\|C\| = \frac{1}{2\sqrt{n}}\sqrt{\sum_{j=1}^{n} c_j^2}.$$

Applying Cauchy-Schwarz inequality, we have

$$\left(\prod_{j=1}^{n} c_j\right)^{\frac{1}{n}} \leq \frac{\sum_{j=1}^{n} c_j}{n} \leq \frac{1}{\sqrt{n}}\sqrt{\sum_{j=1}^{n} c_j^2}.$$

Hence one concludes that for an $n \times n$ square matrix C,

$$\frac{1}{2}|\det(C)|^{\frac{1}{n}} \leq \frac{1}{2\sqrt{n}}\|C\|.$$

By the definition of $\prod \mathcal{V}$, $\sum \mathcal{V}$ and the above inequality one gets

$$\prod \mathcal{V} \leq \sum \mathcal{V}.$$

□

Of course it would be desirable to know for every n and m what the largest possible value of $\sum \mathcal{V}$ is. This is the motivation of the following definition.

Definition 4.2. Let $\Delta(n, m)$ be the infimum of all numbers such that for every unitary space time code $\mathcal{V}$ of dimension n and size m, one has

$$\sum \mathcal{V} \leq \Delta(n, m).$$

Remark 4.3. As pointed out by Liang and Xia [32] there exists a constellation $\mathcal{V}$ of dimension n and size m with $\sum \mathcal{V} = \Delta(n, m)$. This is due to the fact that $U(n)^m$ is a compact manifold.

The exact values of $\Delta(n, m)$ are only known in a few special cases. In the case $n = 1$, one checks that $\Delta(1, m) = \sin\frac{\pi}{m}$ for $m \geq 2$. When $n \geq 2$ and $m = 3$, one has $\Delta(n, 3) = \frac{\sqrt{3}}{2}$. When $m = 2$, we have $\Delta(n, 2) = 1$ for $n \geq 2$. For $n = 2$, the following values were computed in [32].

Table 4.1. Known results for the $\Delta(n, m)$

m	2	3	4	5	6	7	8	9	10 through 16
$\Delta(2, m)$	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{3}\sqrt{6}$	$\frac{1}{4}\sqrt{10}$	$\frac{1}{5}\sqrt{15}$	$\frac{1}{6}\sqrt{21}$	$\frac{1}{7}\sqrt{28}$	$\frac{1}{8}\sqrt{36}$	$\frac{1}{2}\sqrt{2}$

Liang and Xia [32] observed the connection between a unitary constellation and an Euclidean sphere code and beautifully derived an upper bound for 2 dimensional unitary constellations which is very tight when $m \leq 100$. In this chapter we present a new general upper bound for $\Delta(n, m)$ for every dimension n and every size m while improving certain results in [32]. To the best of our knowledge the new upper bounds we derived are tighter than any previously published bounds as soon as m is sufficiently large.

4.2 Upper Bound Analysis

In this section we shall study the packing problem on $U(n)$ and derive three upper bounds for the numbers $\Delta(n, m)$. All the resulting bounds are derived by differential geometric means and all bounds can be viewed as certain sphere packing bounds.

From a differential geometry point of view we can view $U(n)$ as a n^2-dimensional compact Lie group. $U(n)$ is also naturally a submanifold of the Euclidean space $\mathbb{R}^{2n^2}$. In this way $U(n)$ will have the induced geometry of the standard Euclidean geometry of $\mathbb{R}^{2n^2}$. Finally there is a third way to see $U(n)$ as a submanifold of another Riemannian manifold $S(n)$ and we will say more about this later.

The basic strategy for computing the upper bounds for $\Delta(n, m)$ is as follows (See also Figure 4.1). Given a unitary space time code $\mathcal{V} = \{A_1, A_2, \cdots, A_m\}$, around each matrix A_j we can choose a neighborhood $N_r(A_j)$ with radius r (the radius will be specified later). Let $V_j = V(N_r(A_j))$ be the volume of the neighborhood $N_r(A_j)$.

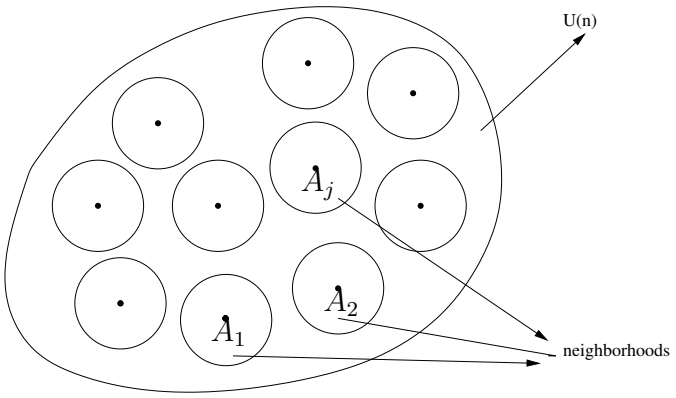


Figure 4.1. Sphere packing on the compact Lie group $U(n)$

If all the neighborhoods are non-overlapping, then necessarily we will have

$$\sum_{j=1}^{m} V_j \leq V(U(n)),$$

where $V(U(n))$ denotes the total volume of unitary group $U(n)$. This inequality in turn will result in an upper bound for the numbers $\Delta(n, m)$. By employing different metrics (Euclidean or Riemannian) and by considering different embeddings of $U(n)$, we derive three different upper bounds for $\Delta(n, m)$.

Let $\mathcal{M}_1$ be the manifold consisting of all the $n \times n$ Hermitian matrices, *i.e.*

$$\mathcal{M}_1 = \{H | H = H^*\}.$$

$\mathcal{M}_1$ has dimension n^2 and can be viewed isometrically as Euclidean space $\mathbb{R}^{n^2}$. Assume that $H = (H_{jk})$ and assume that $H_{jk} = x_{jk} + iy_{jk}$. We use (dH) to denote the volume element of $\mathcal{M}_1$, where

$$(dH) = \left(\frac{i}{2}\right)^{n(n-1)/2} \bigwedge_{l=1}^{n} dH_{ll} \bigwedge_{j<k} dH_{jk} \bigwedge_{j<k} d\bar{H}_{jk} = \bigwedge_{l=1}^{n} dx_{ll} \bigwedge_{j<k} dx_{jk} \bigwedge_{j<k} dy_{jk}. \qquad (4.1)$$

With a small abuse of notation, one can check that the volume element of $\mathcal{M}_2$, the

manifold consisting of all the $n \times n$ skew-Hermitian matrices, can be written as

$$(dH) = \left(\frac{i}{2}\right)^{n(n-1)/2} \left(\frac{1}{i}\right)^{n} \bigwedge_{l=1}^{n} dH_{ll} \bigwedge_{j<k} dH_{jk} \bigwedge_{j<k} d\bar{H}_{jk} = \bigwedge_{l=1}^{n} dy_{ll} \bigwedge_{j<k} dx_{jk} \bigwedge_{j<k} dy_{jk}. \tag{4.2}$$

For a unitary matrix U, if we differentiate $U^*U = I$, we will have

$$U^*dU + dU^*U = 0.$$

Therefore U^*dU is skew-Hermitian. The following lemma will characterize the volume element of $U(n)$. For the terminologies in this lemma, we refer to Chapter 2 or a standard differential geometry or integral geometry book, *e.g.* [26, 38].

Lemma 4.4. *The volume element of $U(n)$ induced by the Euclidean space $\mathbb{R}^{2n^2}$ is bi-invariant and the volume element can be written as (U^*dU) up to a scalar constant.*

Proof. The bi-invariance comes from the orthonormality of $U(n)$. (U^*dU) is left-invariant according to the definition. Indeed for a fixed yet arbitrary unitary matrix V,

$$(VU)^*d(VU) = U^*V^*VdU = U^*dU.$$

Since $U(n)$ is a compact Lie group and any compact Lie group is unimodular, (U^*dU) is also right-invariant. Because the bi-invariant n^2 differential forms are unique up to a scalar, one concludes that the volume element can be written as (U^*dU). □

The following theorem will represent the volume element of $U(n)$ in another way. One will see that it is closely related to the eigenvalues of unitary matrices.

Theorem 4.5. *Consider the eigenvalue decomposition of a unitary matrix Θ:*

$$\Theta = U\mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})U^*, \tag{4.3}$$

In order to make this decomposition unique, we assume that $U \in \tilde{U}(n)$, where $\tilde{U}(n)$ denotes the set of all the unitary matrices with nonnegative real diagonal elements and that $\theta_1 > \theta_2 > \cdots > \theta_n$. We have

$$(\Theta^*d\Theta) = \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 \wedge d\theta_2 \wedge \cdots \wedge d\theta_n \wedge (U^*dU - \mathrm{diag}\,(U^*dU)). \tag{4.4}$$

Proof. Let $D = \text{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})$ and take the differential of Equation (4.3),

$$d\Theta = dUDU^* + UdDU^* + UDdU^*.$$

It follows that,

$$\Theta^* d\Theta = UD^*U^*dUDU^* + UD^*dDU^* + UdU^* = U(D^*U^*dUD + D^*dD)U^* + UdU^*.$$

Due to the right-invariance of the volume element in $U(n)$, it follows that

$$(\Theta^* d\Theta) = (U^*\Theta^* d\Theta U) = (D^*U^*dUD - U^*dU + i\text{diag}\,(d\theta_1, d\theta_2, \cdots, d\theta_n)).$$

Note that $(D^*U^*dUD - U^*dU)_{jk} = (e^{i\theta_j} - e^{i\theta_k})U_{jk}$, therefore the diagonal elements of $D^*U^*dUD - U^*dU$ are all zeros and the off diagonal elements are scaled version of the ones of U^*dU. According to formula (4.2), the claim in the theorem follows. □

The following theorem calculates the volume of a small neighborhood with Euclidean distance r. Because of the homogeneity of $U(n)$, the center of this small "ball" is chosen to be I without loss of generality. For a unitary matrix U, we assume $e^{i\theta_j}$'s are its eigenvalues, *i.e.*, $U \sim \text{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})$. For a fixed unitary matrix A, let

$$U_r^E(n, A) = \{U \in U(n) | \|U - A\| \leq r\}.$$

Again because of the homogeneity of $U(n)$, $V(U_r^E(n, A))$ does not depend on the choice of A. In the sequel $V(U_r^E(n))$ will be used to denote $V(U_r^E(n, A))$ for any unitary matrix A. Let $S(n)$ denote a $2n^2 - 1$ dimensional sphere centered at the origin with radius $\sqrt{n}$, *i.e.*,

$$S(n) = \{(x_1, x_2, \cdots, x_{2n^2}) | x_1^2 + x_2^2 + \cdots + x_{2n^2}^2 = n\}.$$

Observe that $U(n)$ is a submanifold of $S(n)$. For a given point $S_0 \in S(n)$, let

$$S_r(n, S_0) = \{S \in S(n) | \|S - S_0\| \leq r\}.$$

Theorem 4.6. *Let*

$$D_1 = \{(\theta_1, \theta_2, \cdots, \theta_n) | -\pi \leq \theta_j < \pi \text{ for } j = 1, 2, \cdots, n\} \tag{4.5}$$

and

$$D_2 = \left\{ (\theta_1, \theta_2, \cdots, \theta_n) | \sum_{j=1}^{n} \sin^2 \frac{\theta_j}{2} \leq \frac{r^2}{4} \right\}, \tag{4.6}$$

then

$$V(U_r^E(n)) = \frac{\iint_{D_1 \cap D_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n}{\iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n} V(U(n)). \tag{4.7}$$

Proof. Note that $\|I - U\|_2 \leq r$ is equivalent to $\sum_{j=1}^{n} \sin^2 \frac{\theta_j}{2} \leq \frac{r^2}{4}$. For a given unitary matrix Θ, the eigenvalue decomposition $\Theta = U^* \text{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n}) U$ is unique if θ_j's are strictly ordered. So if we take the integral of formula (4.4) over the integration region disregarding the order of θ_j's, we will obtain $n!$ times the volume of $V(U_r^E(n))$. Thus the volume of $U_r^E(n)$ is

$$V(U_r^E(n)) = \frac{1}{n!} \iint_{D_1 \cap D_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n \iint_{U(n)} (U^* dU - \text{diag}\,(U^* dU)).$$

Using the same argument, we derive the volume of $U(n)$:

$$V(U(n)) = \frac{1}{n!} \iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n \iint_{U(n)} (U^* dU - \text{diag}\,(U^* dU)).$$

Comparing the two derived volume formulas, the claim in the theorem follows. □

Remark 4.7. By the Weyl denominator formula [16] one can replace

$$\iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n$$

with $(2\pi)^n n!$. We keep it as it is to make the formula literally understandable.

There are several approaches to derive upper bounds for the diversity sum. The first approach considers $U(n)$ as a submanifold of $S(n)$, then chooses the non-overlapping neighborhoods to be small balls with radius r (with regard to the Euclidean distance). This will result in the first upper bound (B1) which we derive in this chapter.

Theorem 4.8. *Let D_1 and D_2 be defined as in* (4.5) *and* (4.6). *Assume $r_0^E = r_0^E(n,m)$ is the solution to the following equation (with variable r):*

$$m \iint_{D_1 \cap D_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n = \iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n, \tag{4.8}$$

then

$$\Delta(n,m) \leq \sqrt{\frac{(r_0^E)^2}{n} - \frac{(r_0^E)^4}{4n^2}}. \tag{B1}$$

Proof. For a fixed yet arbitrary unitary constellation $\mathcal{V} = \{A_1, A_2, \cdots, A_m\}$, consider m small non-overlapping neighborhoods $S_r(n, A_j)$ in $S(n)$. We can increase r such that there exist l, k such that $S_r(n, A_l)$ and $S_r(n, A_k)$ are tangent to each other. Apparently

$$U_r^E(n, A_j) = S_r(n, A_j) \cap U(n),$$

for any j. Since $S_r(n, A_j)$'s are non-overlapping, we conclude that $U_r^E(n, A_j)$'s are non-overlapping. Therefore we have

$$\sum_{j=1}^{m} V(U_r^E(n, A_j)) \leq V(U(n)),$$

that is

$$mV(U_r^E(n)) \leq V(U(n)).$$

One can check that $V(U_r^E(n))$ is an increasing function of r, so any r satisfying the above inequality will be less than the solution to the equality:

$$mV(U_r^E(n)) = V(U(n)),$$

which is essentially Equality (4.8). So we conclude that $r \leq r_0^E$.

Note that any two points $S_0, S_1 \in S(n)$ with two non-overlapping neighborhoods $S_r(n, S_0)$ and $S_r(n, S_1)$ will have distance $\|S_0 - S_1\| \geq 2\sqrt{r^2 - r^4/(4n)}$, where the equality holds only if $S_r(n, S_0)$ and $S_r(n, S_1)$ are tangent to each other. Apply the argument to A_j's and note that A_l and A_k are the closest pair of points with $\|A_l - A_k\| = 2\sqrt{r^2 - r^4/(4n)}$, we reach the conclusion of the theorem. □

For a fixed $S_0 \in S(n)$, consider $S_r(n, S_0) \subset S(n)$. Let $\tau = \tau(n, r)$ denote the maximal number τ such that

$$S_r(n, S_1), S_r(n, S_2), \cdots, S_r(n, S_\tau)$$

are non-overlapping and $S_r(n, S_j)$ is tangent to $S_r(n, S_0)$ for $j = 1, 2, \cdots, n$. One checks that $\tau(n, r)$ does not depend on the choice of S_0. In this sense $\tau(n, r)$ can be viewed as generalized kissing number [9] on an Euclidean sphere. For a fixed n dimensional unitary constellation $\mathcal{V} = \{A_1, A_2, \cdots, A_m\}$, let $r(\mathcal{V})$ denote the maximal radius r such that

$$S_r(n, A_1), S_r(n, A_2), \cdots, S_r(n, A_m)$$

are non-overlapping. Let $r_{opt} = r_{opt}(n, m)$ denote the maximal $r(\mathcal{V})$ over all possible n dimensional unitary constellation $\mathcal{V}$ with cardinality m. One checks $\Delta(n, m) = r_{opt}(n, m)/\sqrt{2n}$. The following theorem and corollary give a lower bound for the optimal diversity sum $\Delta(n, m)$.

Theorem 4.9. *Let D_1 be defined as in* (4.5) *and assume that $r_0^E = r_0^E(n, m)$ is the solution to the equation* (4.8)*. Let*

$$\tilde{D}_2 = \left\{ (\theta_1, \theta_2, \cdots, \theta_n) | \sum_{j=1}^{n} \sin^2 \frac{\theta_j}{2} \leq \frac{(r_0^E)^2}{4} \right\}$$

and let

$$D_3 = \left\{ (\theta_1, \theta_2, \cdots, \theta_n) | \sum_{j=1}^{n} \sin^2 \frac{\theta_j}{2} \leq r_{opt}(n, m)^2 - r_{opt}(n, m)^4/(4n) \right\}.$$

Then

$$\iint_{D_1 \cap \tilde{D}_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n$$
$$\leq (\tau(n, r_{opt}(n, m)) + 1) \iint_{D_1 \cap D_3} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n.$$

Proof. According to the derivation of r_0^E, we have

$$m \iint_{D_1 \cap \tilde{D}_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n = \iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n. \tag{4.9}$$

Assume that $\mathcal{V} = \{A_1, A_2, \cdots, A_m\}$ is an n dimensional unitary constellation reaching $r_{opt}(n,m)$, *i.e.*, $r(\mathcal{V}) = r_{opt}(n,m)$. For simplicity let $r = r(\mathcal{V})$. Let m' denote the maximal number such that $S_r(n,A_1), S_r(n,A_2), \cdots, S_r(n,A_m), \cdots, S_r(n,A_{m'})$ are non-overlapping. Let $r_1 = 2\sqrt{r^2 - r^2/(4n)}$, we claim that

$$U(n) \subset \bigcup_{j=1}^{m'} U_{r_1}^E(n, A_j).$$

Otherwise suppose there is a unitary matrix $A_0 \notin \bigcup_{j=1}^{m'} U_{r_1}^E(n,A_j)$, then $\|A_0 - A_j\| > r_1$ (see Theorem 4.8). Thus $S_r(n,A_0)$ does not intersect with $S_r(n,A_j)$ for $j = 1, 2, \cdots, m'$. Therefore one can find $m'+1$ small balls with radius r which are non-overlapping. This contradicts the maximality of m'. Thus we have $\sum_{j=1}^{m'} V(U_{r_1}^E(n,A_j)) \geq V(U(n))$, that is

$$m' \iint_{D_1 \cap D_3} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n \geq \iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n. \tag{4.10}$$

We further claim that

$$m' \leq (m-1)(\tau(n,r)+1). \tag{4.11}$$

By contradiction assume that $m' \geq (m-1)(\tau(n,r)+1)+1$. Let

$$\text{tang}\,(j) = \{l | 1 \leq l \leq m', S_r(n,A_l) \quad \text{tangent to} \quad S_r(n,A_j)\}.$$

According to the definition of $\tau(n,r)$, we know the cardinality of $\text{tang}\,(j)$ is less than $\tau(n,r)$. We first pick j_1 from $\{0,1,\cdots,m'\}$, then pick j_2 from $\{0,1,\cdots,m'\} - \text{tang}\,(j_1)$. And we continue this process by always picking j_{k+1} from

$$\{0,1,\cdots,m'\} - \bigcup_{l=1}^{k} \text{tang}\,(j_l).$$

Since the cardinality of the above set is strictly greater than 0 when $k \leq m-1$, we can pick $j_1, j_2, \cdots, j_m$ from the index set $\{1,2,\cdots,m'\}$ such that $S_r(n,A_{j_1}), S_r(n,A_{j_2}), \cdots, S_r(n,A_{j_m})$ are non-overlapping and every two of them are not tangent to each other. Then we can find a small enough real number $\varepsilon > 0$ and increase the radius r to $r+\varepsilon$ such that

$$S_{r+\varepsilon}(n,A_{j_1}), S_{r+\varepsilon}(n,A_{j_2}), \cdots, S_{r+\varepsilon}(n,A_{j_m})$$

are still non-overlapping. However this contradicts the maximality of $r = r_{opt}(n, m)$.

The combination of the three formulas (4.9), (4.10), (4.11) will lead to

$$\frac{\iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n}{\iint_{D_1 \cap D_3} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n} \leq \left(\frac{\iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n}{\iint_{D_1 \cap \tilde{D}_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n} - 1 \right) (\tau(n, r) + 1). \quad (4.12)$$

Note that the inequality above is in fact stronger than the claim in the theorem. We can reach the conclusion of the theorem by relaxing the right hand side of the inequality (by ignoring -1). □

Corollary 4.10. *When $m \to \infty$, asymptotically we have*

$$\Delta(n, m) \geq 2\sqrt{n} r_0^E(n, m) \frac{1}{2} (\tau(2n^2 - 1) + 1)^{-1/n^2}.$$

Proof. We only sketch the idea of the proof. Intuitively $U_r^E(n, A_0)$ looks more "flat" when $m \to \infty$ (consequently $r \to 0$), so $V(U_r^E(n, A_0))$ can be approximated by the volume of $U_r^E(n, A_0)$'s projection to the tangent space of $U(n)$ at A_0:

$$\iint_{D_1 \cap \tilde{D}_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n \sim C(r_0^E)^{n^2}$$

for some constant C. The same argument will lead to

$$\iint_{D_1 \cap D_3} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n \sim C(2r_{opt})^{n^2}$$

for the same constant C. For any fixed n, $\tau(n, r)$ will approach to the standard kissing number in Euclidean space $\tau(2n^2 - 1)$ when r goes to zero. Combining the three approximations, we reach the claim according to the previous theorem.

□

$U(n)$ is a compact Lie group equipped with a Riemannian metric. Given two points $A_0, A_1 \in U(n)$, one can always find a geodesic $\gamma(t)$ (mapping from $[0, 1]$ to $U(n)$) which will connect these two points, *i.e.* $\gamma(0) = A_0$ and $\gamma(1) = A_1$. Recall

that the Euclidean distance between A_0 and A_1 is defined to be $\|A_0 - A_1\|$. We further define the Riemannian distance between A_0 and A_1 to be:

$$\text{dist}\,(A_0, A_1) = \int_0^1 \|\gamma'(t)\| dt.$$

As a Lie group $U(n)$ is homogeneous. In particular one has that

$$\text{dist}\,(A_0, A_1) = \text{dist}\,(UA_0, UA_1) = \text{dist}\,(A_0U, A_1U)$$

for any $U \in U(n)$. The following theorem use the homogeneity and the relationship between the Riemannian distance and Euclidean distance to derive another upper bound for the diversity sum in general and it is the base of the second approach.

Theorem 4.11. *Let $f(\cdot)$ and $g(\cdot)$ be two fixed monotone increasing real functions. If*

$$g(\|A_0 - A_1\|) \leq \text{dist}\,(A_0, A_1) \leq f(\|A_0 - A_1\|)$$

for any two unitary matrices A_0 and A_1, then

$$\Delta(n, m) \leq g^{-1}(2f(r_0^E(n, m)))/(2\sqrt{n}).$$

Proof. For a fixed unitary constellation $\mathcal{V} = \{A_1, A_2, \cdots, A_m\}$, consider

$$U_r^E(n, A_1), U_r^E(n, A_2), \cdots, U_r^E(n, A_m)$$

for $r > 0$. We can increase r until there exist j and k such that $U_r^E(n, A_j)$ and $U_r^E(n, A_k)$ are tangent to each other at a point A_0. As examined in Theorem 4.8, one can make a conclusion that $r \leq r_0^E(n, m)$. Accordingly we have

$$\begin{aligned}\text{dist}\,(A_j, A_k) &\leq \text{dist}\,(A_j, A_0) + \text{dist}\,(A_k, A_0) \\ &\leq f(\|A_j - A_0\|) + f(\|A_k - A_0\|) = 2f(r) \leq 2f(r_0^E(n, m)).\end{aligned}$$

On the other hand since g is monotonically increasing one has:

$$\|A_j - A_k\| \leq g^{-1}(\text{dist}\,(A_j, A_k)).$$

The combination of the above two inequalities will lead to

$$\|A_j - A_k\| \leq g^{-1}(2f(r_0^E(n,m))).$$

Immediately we will have

$$\sum \mathcal{V} \leq g^{-1}(2f(r_0^E(n,m)))/(2\sqrt{n}).$$

Since $\mathcal{V}$ is an arbitrary unitary constellation, the claim in the theorem follows. □

Based on the above theorem, the following corollary gives the second upper bound (B2).

Corollary 4.12. *For a real number r, let $\lfloor r \rfloor$ denote the greatest integer less than or equal to r. Then*

$$\Delta(n,m) \leq \sin \sqrt{\frac{\pi^2}{n}\left\lfloor \frac{(r_0^E)^2(n,m)}{4} \right\rfloor + \frac{4}{n}\arcsin^2 \sqrt{\frac{(r_0^E)^2(n,m)}{4} - \left\lfloor \frac{(r_0^E)^2(n,m)}{4} \right\rfloor}}. \tag{B2}$$

Proof. Consider I and another point $U = V\mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})V^*$, where $-\pi \leq \theta_j < \pi$. It is known that [12] the geodesic from I to U can be parameterized by

$$\gamma(t) = V\mathrm{diag}\,(e^{i\theta_1 t}, e^{i\theta_2 t}, \cdots, e^{i\theta_n t})V^*,$$

where $0 \leq t \leq 1$. The Riemannian distance from I to U is

$$\mathrm{dist}\,(I,U) = \sqrt{\theta_1^2 + \theta_2^2 + \cdots + \theta_n^2}.$$

We want to derive $g(\cdot), f(\cdot)$ as in Theorem 4.11. Suppose the Euclidean distance between I and U is r, *i.e.*,

$$\sin^2\frac{\theta_1}{2} + \sin^2\frac{\theta_2}{2} + \cdots + \sin^2\frac{\theta_n}{2} = r^2/4.$$

After substituting with $x_j = \sin^2\theta_j/2$ and denoting $G(x) = \arcsin^2\sqrt{x}$, we convert the above problem to the following optimization problem:

Find the minimum and maximum of the function

$$F(x_1, x_2, \cdots, x_n) = \theta_1^2 + \theta_2^2 + \cdots + \theta_n^2 = 4(G(x_1) + G(x_2) + \cdots + G(x_n))$$

with the constraints $x_1+x_2+\cdots+x_n = r^2/4$ and $0 \le x_j \le 1$ for $j = 1, 2, \cdots, n$. Since $G(x)$ is a convex function on $[0, 1]$, we derive the lower bound of $F(x_1, x_2, \cdots, x_n)$,

$$4n \arcsin^2(r/(2\sqrt{n})) \le F(x_1, x_2, \cdots, x_n). \tag{4.13}$$

In the sequel we will calculate an upper bound of $F(x_1, x_2, \cdots, x_n)$. Without loss of generality, we assume $0 \le x_1 \le x_2 \le \cdots \le x_n \le 1$. Let $k = \lfloor r^2/4 \rfloor$ and $\alpha = r^2/4-k$, we claim that $F(x_1, x_2, \cdots, x_n)$ will reach its maximum when

$$x_j = \begin{cases} 0, & 1 \le j \le n-k-1 \\ \alpha, & j = n-k \\ 1, & n-k+1 \le j \le n \end{cases}$$

Suppose by contradiction that F reaches its maximum at $(x_1, x_2, \cdots, x_n)$ with $x_1 > 0$. Now from

$$x_1 + x_{n-k} + x_{n-k+1} + \cdots + x_n \le r^2/4 = k + \alpha,$$

surely one can find $x'_{n-k}, x'_{n-k+1}, \cdots, x'_n$ such that

$$x_1 + x_{n-k} + x_{n-k+1} + \cdots + x_n = x'_{n-k} + x'_{n-k+1} + \cdots + x'_n,$$

with $x'_j \ge x_j$ for $j = n-k, n-k+1, \cdots, n$. Now set $x_1^* = 0$, $x_j^* = x_j$ for $j = 2, 3, \cdots, n-k-1$ and $x_j^* = x'_j$ for $j = n-k, n-k+1, \cdots, n$. By the mean value theorem, there exist ζ_j's with $x_1^* = 0 \le \zeta_1 \le x_1$ and $x_j \le \zeta_j \le x_j^*$ for $j = 2, 3, \cdots, n$ such that

$$F(x_1^*, x_2^*, \cdots, x_n^*) - F(x_1, x_2, \cdots, x_n) = \sum_{j=1}^{n} G'(\zeta_j)(x_j^* - x_j).$$

Since $G(x)$ is a strictly convex function, we have

$$0 < G'(\zeta_1) < G'(\zeta_2) < \cdots < G'(\zeta_n).$$

Now

$$\begin{aligned} F(x_1^*, x_2^*, \cdots, x_n^*) - F(x_1, x_2, \cdots, x_n) &\ge G'(\zeta_2)(\sum_{j=2}^{n}(x_j^* - x_j)) - G'(\zeta_1)(x_1 - x_1^*) \\ &= (G'(\zeta_2) - G'(\zeta_1))(x_1 - x_1^*) = (G'(\zeta_2) - G'(\zeta_1))x_1 > 0. \end{aligned}$$

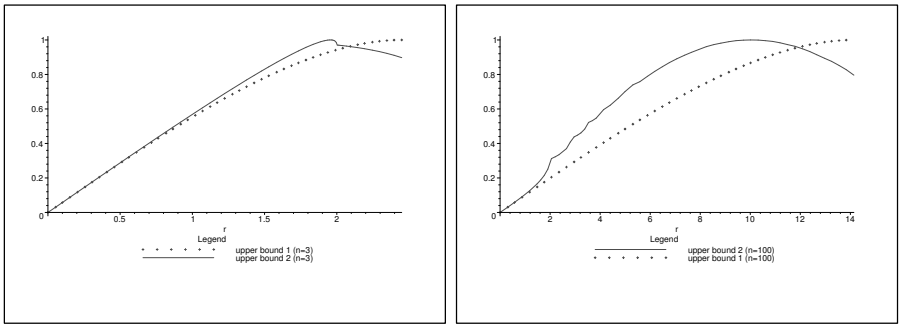


Figure 4.2. The comparisons of two upper bounds as functions for $n = 3$ and $n = 100$

This contradicts the maximality of F at $(x_1, x_2, \cdots, x_n)$. Applying exactly the same analysis to $x_2, x_3, \cdots, x_{n-k-1}, x_{n-k}$ we deduce that $x_j = 0$ for $j = 2, 3, \cdots, n-k-1$ and $x_{n-k} = \alpha$. So the upper bound of F can be given as

$$F(x_1, x_2, \cdots, x_n) \leq 4\left(k\frac{\pi^2}{4} + \arcsin^2(\sqrt{\alpha})\right).$$

Take $g(r) = 2\sqrt{n}\arcsin(r/(2\sqrt{n}))$ and $f(r) = 2\sqrt{k\pi^2/4 + \arcsin^2\sqrt{\alpha}}$, the corollary follows according to the previous theorem. □

Note that both upper bound (B1) and upper bound (B2) depend on $r_0^E(n, m)$. In Figure 4.2 we plot both upper bounds as functions of $r_0^E(n, m)$ for 3 and 100 dimensions. One can see that if and only if $r_0^E(3, m) > 2.0881$, the upper bound (B2) is tighter than the upper bound (B1). While for the 100 dimension case, the upper bound (B1) is tighter than the upper bound (B2) if and only if $r_0^E(100, m) > 11.9155$. In fact it can be checked that asymptotically when n is large enough, upper bound (B2) is tighter than upper bound (B1) if and only if $r_0^E(n, m) > 1.1892\sqrt{n}$.

For a packing problem on a manifold, alternatively one can choose the neighborhood to be a small "ball" with Riemannian radius r. This will be our third approach

to derive an upper bound for the diversity sum. For a particular $A \in U(n)$, let

$$U_r^R(n, A) = \{U \in U(n) | \text{dist}\,(U, A) \leq r\}.$$

Note that the constraint $\text{dist}\,(U, I) \leq r$ is equivalent to

$$\theta_1^2 + \theta_2^2 + \cdots + \theta_n^2 \leq r^2.$$

Therefore we apply the same argument as in the proof of Theorem 4.6 and conclude that:

$$V(U_r^R(n)) = \frac{\iint_{D_1 \cap D_4} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n}{\iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n} V(U(n)),$$

where D_1 was defined in (4.5) and

$$D_4 := \{(\theta_1, \theta_2, \cdots, \theta_n) | \sum_{j=1}^{n} \theta_j^2 \leq r^2\}. \tag{4.14}$$

Instead of considering the Euclidean neighborhoods

$$U_r^E(n, A_1), U_r^E(n, A_2), \cdots, U_r^E(n, A_m),$$

we can consider the Riemannian neighborhood $U_r^R(n, A_1), U_r^R(n, A_2), \cdots, U_r^R(n, A_m)$. Utilizing the fact that the Euclidean distance $\|A_j - A_k\|$ and the Riemannian distance $\text{dist}\,(A_j, A_k)$ are related (compare with Formula (4.13)):

$$4n \arcsin^2(\|A_j - A_k\|/(2\sqrt{n})) \leq \text{dist}\,(A_j, A_k)$$

for any two unitary matrices A_j and A_k, we can derive the third upper bound (B3). The proof of the following theorem is very similar to the one of Theorem 4.11 and for the sake of brevity we omit it.

Theorem 4.13. *Let D_1 and D_4 be defined as in (4.5) and (4.14) and assume $r_0^R(n, m)$ is the solution to the following equation (with variable r):*

$$m \iint_{D_1 \cap D_4} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n = \iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n, \tag{4.15}$$

then

$$\Delta(n, m) \leq \sin\left(\frac{r_0^R(n, m)}{\sqrt{n}}\right). \tag{B3}$$

4.3 Numerical Results and Conclusions

We gave three approaches to derive upper bounds for the diversity sum and hence also for the diversity product. All of them involve the calculation of $r_0^E(n,m)$ or $r_0^R(n,m)$, which are the solutions of equation (4.8) and equation (4.15), respectively. Fortunately we are dealing with finding a root of a monotone increasing function (recall that both $V(U_r^E(n,m))$ and $V(U_r^R(n,m))$ are monotone increasing functions with respect to r), the bisection method [5] will be highly effective to solve this kind of problem. Our numerical experiments for small size constellations with small dimensions show that upper bound (B3) is looser than the first two upper bounds. However, when m goes to infinity, these three upper bounds give almost the same estimation. This makes sense because asymptotically the small balls look like a n^2 dimensional ball in Euclidean space. One can see the derived upper bounds for 2 and 3 dimensional constellations in Figure 4.3.

We compare the derived upper bounds with the currently existing one presented in [32]. For $n = 2$ the upper bounds derived by Liang and Xia [32] tend to be better when $m \leq 100$ and our bounds become tighter when $m \geq 100$ (see the following Table 4.2). For $n \geq 3$ Liang and Xia [32] outlined a method by considering a sphere packing computation in $S(n)$. It is our belief that this method will result in a weaker bound than the upper bounds we derived in this chapter. For the sample programs to do the upper bound calculation, we refer to [18].

One interesting fact about the limiting behavior of $\Delta(n,m)$ (when $m \to \infty$) is its connection to the Kepler problem [9]. Certainly one can use Kepler density [9] to obtain a tighter bound of the diversity sum asymptotically.

We presented three approaches to derive upper bounds for the diversity sum of unitary constellations of any dimension n and any size m. The derived bounds seem to improve the existing bounds when $n = 2$ and $m \geq 100$. When n is large the exact

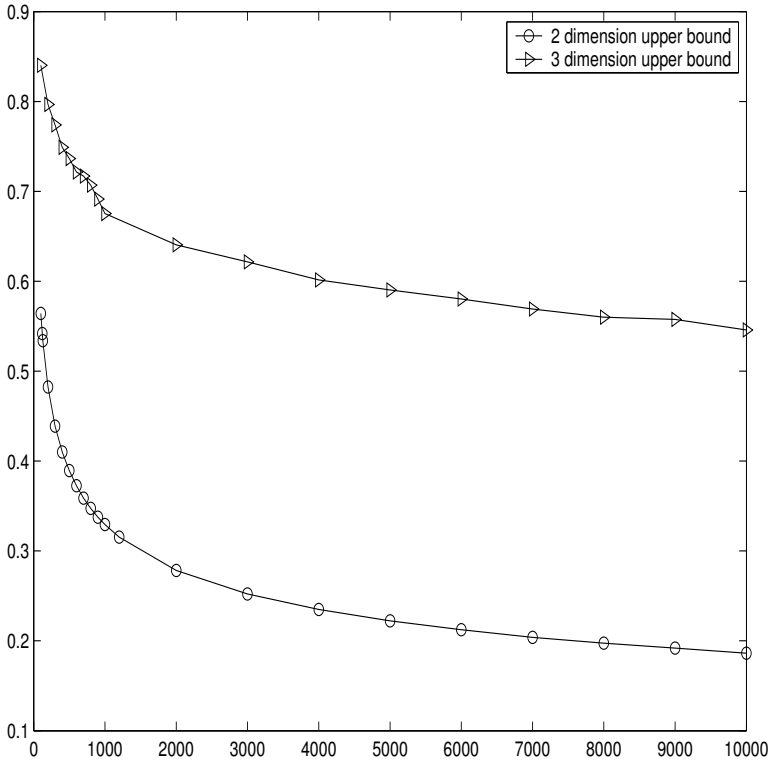


Figure 4.3. Upper bounds for 2 and 3 dimensional constellations

For $n = 2$ the following table compares the upper bounds in [32] with our new bounds (B1) and (B2).

Table 4.2. Two dimensional constellation diversity upper bound

m	24	48	64	80	100	120	128
upper bounds in [32]	0.6746	0.6193	0.5969	0.5799	0.5632	0.5499	0.5452
upper bound (B1)	0.7598	0.6603	0.6131	0.5932	0.5578	0.5425	0.5347
upper bound (B2)	0.7794	0.6734	0.6235	0.6026	0.5654	0.5496	0.5415

computation of r_0^E is rather involved and hence it is also computationally difficult to compute the bounds (B1) and (B2). Nonetheless it is our belief that the resulting

upper bounds (B1) and (B2) become fairly tight as soon as m is sufficiently large. It is interesting to see that if one can apporoach the diveristy upper bound as tight as possible. In the following chapters, constellations with large diversity (product and sum) featuing fast decoding algorithms will be presented.

CHAPTER 5

GENERALIZED PSK IN SPACE TIME CODING

5.1 Introduction

Orthogonal designs have been investigated for the constellation construction for coherent channels. Using a complex orthogonal design, Alamouti [3] proposed a very simple transmitter diversity scheme with 2 transmitting antennas. A very interesting coding scheme from the real orthogonal designs is presented in [43]. Applying the similar idea as in [3], the authors also explain how the symmetric structure of the orthogonal design codes leads to a much simpler decoding algorithm. However no explicit constellations have been studied in details in these work.

Interestingly enough the normalized complex orthogonal design Alamouti codes and real orthogonal design codes can be used in non-coherent scenarios as well. Our work is mainly about non-coherent channels, however without any doubts the resulting codes can be used for coherent channels as well. In this chapter we will show how to construct space time codes from these schemes using generalized phase shift keying (GPSK) signals. Our results can be applied to generalized orthogonal design in [30].

This chapter is organized as follows: in Section 5.2, three series of 2 dimensional GPSK constellations from complex orthogonal designs will be presented. An algebraic calculation will show that they have larger diversity products than the original orthogonal design constellations. As a consequence of larger diversity, the

ML decoding of a GPSK constellation gives better performance. In Section 5.3, we explicitly construct 4 and 8 dimensional GPSK constellations from the real orthogonal designs. Fast decoding algorithms for the proposed constellations will be presented in Section 5.2 and Section 5.3.

5.2 GPSK Constellations from the Complex Orthogonal Designs

A very simple yet interesting complex orthogonal scheme is described in [3]. We will consider the normalized version of this proposed code, *i.e.*, every element of this code is a matrix given by

$$\mathcal{O}(a,b) = \frac{1}{\sqrt{2}} \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix},$$

where $|a|^2 = |b|^2 = 1$. Observe that $\mathcal{O}(a,b)$ is a unitary matrix and a constellation $\mathcal{O}(n)$ with size $L = n^2$ is obtained by letting a and b range over the set of n-th roots of unity

$$\vartheta = \{1, e^{2\pi i/n}, \cdots, e^{2\pi i(n-1)/n}\},$$

namely the entries of the matrix are chosen from scaled one dimensional PSK signal set ϑ (Note that a and b have the same energy, see an example when $n = 8$ in Figure 5.1). This would be the most commonly implemented Alamouti's scheme. As shown at the end of Chapter 3, the diversity product and the diversity sum of such constellations are equal, thus in the following we only consider the diversity product. The diversity product of the constellation $\mathcal{O}(n)$ is

$$\prod \mathcal{O}(n) = \frac{\sqrt{2}}{2} \sin \frac{\pi}{n}.$$

Note that $\mathcal{O}(n)$ is similar to ϑ in the sense that all the elements in $\mathcal{O}(n)$ have unit energy and every pair of elements differ only by the phases. $\mathcal{O}(n)$ is a subset of the

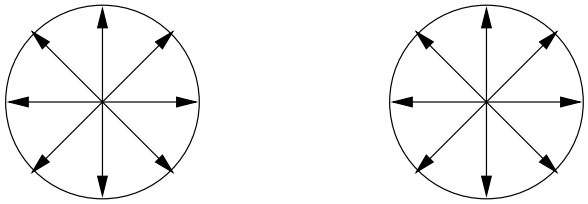


Figure 5.1. The values of a and b

special unitary group

$$SU(2) = \left\{ \left(\begin{array}{cc} a & b \\ -b^* & a^* \end{array} \right) \middle| \, |a|^2 + |b|^2 = 1 \right\}.$$

In this section we will present three series of unitary constellations as finite subsets of $SU(2)$.

The basic principle to decode the constellation $\mathcal{O}(n)$ has been discussed in [3]. We shall describe this decoding process in another way with more details and generalize it to our constellations. Consider a non-coherent wireless communication system modulated by $\mathcal{O}(n)$ with 2 transmitting antennas and N receiving antennas and assume the differential unitary space time modulation [28] is used. Let X and Y denote the received matrices at time block τ and $\tau + 1$, respectively, then the ML decoder will perform the following decoding task

$$(\hat{a}, \hat{b}) = \arg \min_{a,b \in \vartheta} \left\| Y - \frac{1}{\sqrt{2}} \left(\begin{array}{cc} a & b \\ -b^* & a^* \end{array} \right) X \right\|^2.$$

With simple matrix manipulations, one can check that

$$(\hat{a},\hat{b}) = \arg\min_{a,b\in\vartheta} \sum_{i=1}^{N} \left\| \begin{pmatrix} Y_{1i} \\ Y_{2i} \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \begin{pmatrix} X_{1i} \\ X_{2i} \end{pmatrix} \right\|^2$$

$$= \arg\min_{a,b\in\vartheta} \sum_{i=1}^{N} \left\| \begin{pmatrix} Y_{1i} \\ Y_{2i}^* \end{pmatrix} - \frac{1}{\sqrt{2}} \begin{pmatrix} X_{1i} & X_{2i} \\ X_{2i}^* & -X_{1i}^* \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \right\|^2$$

$$= \arg\min_{a,b\in\vartheta} \sum_{i=1}^{N} \frac{2}{|X_{1i}|^2 + |X_{2i}|^2} \left\| \frac{1}{\sqrt{2}} \begin{pmatrix} X_{1i} & X_{2i} \\ X_{2i}^* & -X_{1i}^* \end{pmatrix} \begin{pmatrix} Y_{1i} \\ Y_{2i}^* \end{pmatrix} - \frac{1}{2}(|X_{1i}|^2 + |X_{2i}|^2) \begin{pmatrix} a \\ b \end{pmatrix} \right\|^2$$

$$= \arg\min_{a,b\in\vartheta} \sum_{i=1}^{N} \frac{2}{|X_{1i}|^2 + |X_{2i}|^2} \left(\left| \frac{1}{\sqrt{2}} X_{1i}^* Y_{1i} + \frac{1}{\sqrt{2}} X_{2i} Y_{2i}^* - \frac{1}{2}(|X_{1i}|^2 + |X_{2i}|^2) a \right|^2 \right.$$

$$\left. + \left| \frac{1}{\sqrt{2}} X_{2i}^* Y_{1i} + \frac{1}{\sqrt{2}} X_{1i} Y_{2i}^* - \frac{1}{2}(|X_{1i}|^2 + |X_{2i}|^2) b \right|^2 \right).$$

Further algebraic simplifications show that ML decoding is very simple:

$$(\hat{a},\hat{b}) = \arg\max_{a,b\in\vartheta} \left(\mathrm{Re}\left(a \sum_{i=1}^{N} \bar{Z}_i \right) + \mathrm{Re}\left(b \sum_{i=1}^{N} \bar{W}_i \right) \right), \tag{5.1}$$

where $Z_i = X_{1i}^* Y_{1i} + X_{2i} Y_{2i}^*$ and $W_i = X_{2i}^* Y_{1i} + X_{1i} Y_{2i}^*$. Since a and b are independent of each other (a and b can be chosen freely in the set ϑ), the evaluations above amount to

$$\hat{a} = \arg\max_{a\in\vartheta} \mathrm{Re}\left(a \sum_{i=1}^{n} \bar{Z}_i \right),$$

and

$$\hat{b} = \arg\max_{b\in\vartheta} \mathrm{Re}\left(b \sum_{i=1}^{n} \bar{W}_i \right).$$

Rewrite a as $a = e^{2\pi ji/n}$ and b as $b = e^{2\pi ki/n}$ and let

$$\lfloor r \rceil = \lfloor r + 1/2 \rfloor,$$

i.e., $\lfloor r \rceil$ denotes the smaller of the closest integers to r. The ML decoder will take the following simple form:

$$\hat{j} = \left\lfloor \frac{n \arg Z}{2\pi} \right\rceil, \qquad \hat{k} = \left\lfloor \frac{n \arg W}{2\pi} \right\rceil, \tag{5.2}$$

where $Z = \sum_{i=1}^{n} Z_i$ and $W = \sum_{i=1}^{n} W_i$.

Assume that a communication channel is modulated by one dimensional PSK signals ϑ. Let $a = e^{2\pi l/n} \in \vartheta$ denote the sent signal at one time and Z denote the corresponding corrupted signal. The ML decoding will look for the closest signal $\hat{a} = e^{2\pi \hat{j}/n}$ to a in the signal set ϑ, *i.e.*,

$$\hat{j} = \left\lfloor \frac{n \arg Z}{2\pi} \right\rceil. \tag{5.3}$$

Compare Formula (5.2) and Formula (5.3), one can conclude that the decoding of space time code $\mathcal{O}(n)$ is decomposable and can be implemented by taking one dimensional PSK demodulation twice. The normalized Alamouti codes $\mathcal{O}(n)$'s admit very simple decoding algorithms, since roughly only $4N$ complex multiplications and $4N$ complex additions are needed. We introduce other complex orthogonal codes using GPSK signals in the sequel. Instead of allocating the same energy to every transmitting antenna, we attempt to maximize the diversity product, consequently optimize the power allocation. Compared to $\mathcal{O}(n)$, these codes have larger diversity, which promise good performance with ML decoding. Due to the symmetry from GPSK signals, these codes are also decomposable, therefore fast decoding algorithms can be also applied to these codes. In the following three subsections we present the constructions of these three series of constellations.

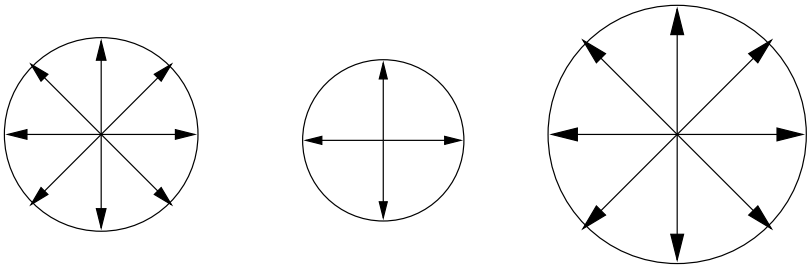
5.2.1 Construction 1

Let n be an even number and let $0 < r < \frac{\sqrt{2}}{2}$ be the root of the following equation

$$\left(\frac{\sqrt{2}}{2} - r\right)^2 + \left(\frac{\sqrt{2}}{2} - \sqrt{1-r^2}\right)^2 = 4r^2 \sin^2 \frac{2\pi}{n}. \tag{5.4}$$

Consider the following sets of the scaled one dimensional PSK signals:

$$A_1(n) = \left\{ \sqrt{2}/2 e^{i\frac{2\pi}{n}k} \middle| \, k = 0, 1, \cdots, n-1 \right\},$$



$A_1(8)$ $A_2(8)$ $A_3(8)$

Figure 5.2. The set of A_j's when n=8

$$A_2(n) = \left\{ re^{i\frac{4\pi}{n}k} \middle| \, k = 0, 1, \cdots, \frac{n}{2} - 1 \right\},$$
$$A_3(n) = \left\{ \sqrt{1-r^2} e^{i\frac{2\pi}{n}k} \middle| \, k = 0, 1, \cdots, n-1 \right\}.$$

When $n = 8$, the above three sets can be illustrated as in Figure 5.2. Consider the following subsets of $SU(2)$:

$$C_1(n) = \left\{ \left(\begin{array}{cc} a & b \\ -\bar{b} & \bar{a} \end{array} \right) \middle| \, a \in A_1(n), b \in A_1(n) \right\},$$
$$C_2(n) = \left\{ \left(\begin{array}{cc} a & b \\ -\bar{b} & \bar{a} \end{array} \right) \middle| \, a \in A_2(n), b \in A_3(n) \right\},$$
$$C_3(n) = \left\{ \left(\begin{array}{cc} a & b \\ -\bar{b} & \bar{a} \end{array} \right) \middle| \, a \in A_3(n), b \in A_2(n) \right\}.$$

Theorem 5.1. *Let*

$$\mathcal{V}_1(n) := C_1(n) \bigcup C_2(n) \bigcup C_3(n),$$

$\mathcal{V}_1(n)$ *is a fully diverse constellation of* $2n^2$ *elements with diversity product:*

$$\prod \mathcal{V}_1(n) = \min \left\{ \frac{\sqrt{2}}{2} \sin \frac{\pi}{n}, \frac{1}{2} \sqrt{ \left(\frac{\sqrt{2}}{2} - r \right)^2 + \left(\frac{\sqrt{2}}{2} - \sqrt{1-r^2} \right)^2 } \right\}.$$

Proof. Take two elements $A \in C_i$ and $B \in C_j$. Without loss of generality, we further assume $i \leq j$. If $A, B \in C_1(n)$, then

$$|\det(A-B)| \geq \left(\sqrt{2}\sin\frac{\pi}{n}\right)^2 = 2\sin^2\frac{\pi}{n}.$$

If $A \in C_1(n), B \in C_2(n)$ or $A \in C_1(n), B \in C_3(n)$ then

$$|\det(A-B)| \geq \left(\frac{\sqrt{2}}{2}-r\right)^2 + \left(\frac{\sqrt{2}}{2}-\sqrt{1-r^2}\right)^2 = 2-\sqrt{2}r-\sqrt{2}\sqrt{1-r^2}.$$

If $A \in C_2(n)$ and $B \in C_3(n)$ then

$$|\det(A-B)| \geq 2(\sqrt{1-r^2}-r)^2.$$

If $A, B \in C_3(n)$ or $A, B \in C_2(n)$ then

$$|\det(A-B)| \geq \left(2r\sin\frac{2\pi}{n}\right)^2.$$

Use the fact that r is the root of Equation (5.4) and compare the lower bounds of the diversity product in all the above cases, the claim in the theorem can be established. □

The following table shows how the diversity product of $\mathcal{V}_1(n)$ compares with the diversity product of $\mathcal{O}(n)$ when $n \leq 12$.

n	r	$\frac{\sqrt{2}}{2}\sin\frac{\pi}{n} = \prod\mathcal{O}(n)$	$\frac{1}{2}\sqrt{(\frac{\sqrt{2}}{2}-r)^2+(\frac{\sqrt{2}}{2}-\sqrt{1-r^2})^2}$	$\prod\mathcal{V}_1(n)$
4	0.259	0.5	0.259	0.259
6	0.284	0.353	0.246	0.246
8	0.321	0.271	0.227	0.227
10	0.360	0.219	0.209	0.209
12	0.386	0.183	0.193	0.183

Corollary 5.2. *For $n \geq 12$,*

$$\prod\mathcal{V}_1(n) = \frac{\sqrt{2}}{2}\sin\frac{\pi}{n}.$$

Proof. For $n \geq 12$,

$$\frac{1}{2}\sqrt{\left(\frac{\sqrt{2}}{2}-r\right)^2+\left(\frac{\sqrt{2}}{2}-\sqrt{1-r^2}\right)^2} = r\sin\frac{2\pi}{n} = 2r\sin\frac{\pi}{n}\cos\frac{\pi}{n}$$

$$\geq 2\times 0.386\cos\frac{\pi}{12}\sin\frac{\pi}{n} \geq \frac{\sqrt{2}}{2}\sin\frac{\pi}{n}.$$

Consequently we have

$$\prod \mathcal{V}_1(n) = \min\left\{\frac{\sqrt{2}}{2}\sin\frac{\pi}{n}, \frac{1}{2}\sqrt{\left(\frac{\sqrt{2}}{2}-r\right)^2+\left(\frac{\sqrt{2}}{2}-\sqrt{1-r^2}\right)^2}\right\} = \frac{\sqrt{2}}{2}\sin\frac{\pi}{n}.$$

□

The above corollary indicates that for $n \geq 12$, the GPSK constellation $\mathcal{V}_1(n)$ has the same diversity product as the orthogonal constellation $\mathcal{O}(n)$, while it has twice as many elements: $\mathcal{V}_1(n)$ has $2n^2$ elements whereas $\mathcal{O}(n)$ has n^2 elements.

Similar to the case for $\mathcal{O}(n)$, the ML decoding of $\mathcal{V}_1(n)$ boils down to Formula (5.1). However we can not separate the estimation of a, b using the simple Formula (5.2), because generally a and b are not independent anymore. If we restrict the decoding evaluation in a particular $C_i(n)$, then $\arg(a)$ and $\arg(b)$ can be chosen freely. Namely within the restricted searching area $C_i(n)$, a and b are independent of each other. Assume the originally sent codeword falls in $C_1(n)$, one can use Formula (5.2) to resolve a candidate $(\hat{a}_1, \hat{b}_1)$:

$$\hat{j} = \left\lfloor \frac{n\arg Z}{2\pi} \right\rceil, \qquad \hat{k} = \left\lfloor \frac{n\arg W}{2\pi} \right\rceil.$$

Similarly we can have candidates $(\hat{a}_i, \hat{b}_i)$ for $C_i(n)$, where $i = 2, 3$. The final ML decoder will resolve the most likely sent codeword:

$$(\hat{a}, \hat{b}) = \arg\max_{\hat{a}_i, \hat{b}_i} \left(\mathrm{Re}(a\bar{Z}) + \mathrm{Re}(b\bar{W})\right).$$

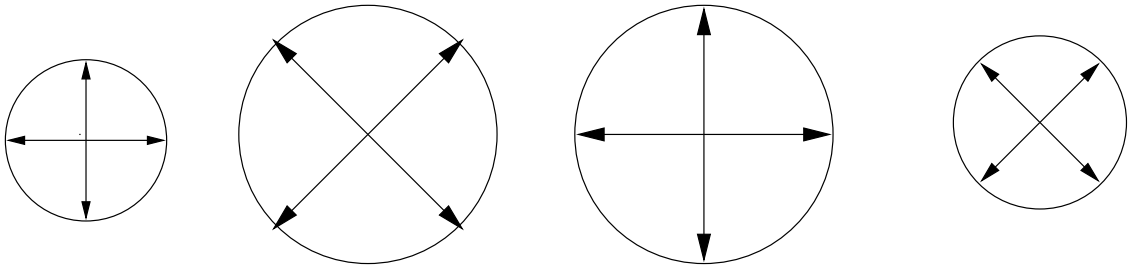


Figure 5.3. The values of A_j's when $n = 8$

Again we conclude that the decoding of $\mathcal{V}_1(n)$ is decomposable and therefore $\mathcal{V}_1(n)$ admits a simple decoding. The above evaluations require roughly $4N$ complex multiplications and $4N$ complex additions, therefore the decoding of the GPSK constellation $\mathcal{V}_1(n)$ has the same complexity as that of $\mathcal{O}(n)$.

5.2.2 Construction 2

Let $n = 2m$ and consider the following sets consisting of scaled one dimensional PSK signals:

$$\begin{aligned}
A_1(n) &= \left\{ re^{i\frac{2\pi}{m}j} | j = 0, 1, \cdots, m-1 \right\}, \\
A_2(n) &= \left\{ \sqrt{1-r^2} e^{i(\frac{2\pi}{m}j + \frac{\pi}{m})} | j = 0, 1, \cdots, m-1 \right\}, \\
A_3(n) &= \left\{ \sqrt{1-r^2} e^{i\frac{2\pi}{m}j} | j = 0, 1, \cdots, m-1 \right\}, \\
A_4(n) &= \left\{ re^{i(\frac{2\pi}{m}j + \frac{\pi}{m})} | j = 0, 1, \cdots, m-1 \right\},
\end{aligned}$$

where

$$r = \frac{1}{\sqrt{2\sin^2\frac{\pi}{m} + 2\sqrt{2}\sin\frac{\pi}{m} + 2}}.$$

Based on the above signal sets, construct the following subsets of $SU(2)$:

$$C_1(n) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} | a \in A_1(n), b \in A_2(n) \right\},$$

$$C_2(n) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} | a \in A_2(n), b \in A_1(n) \right\},$$

$$C_3(n) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} | a \in A_3(n), b \in A_4(n) \right\},$$

$$C_4(n) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} | a \in A_4(n), b \in A_3(n) \right\}.$$

Theorem 5.3. *Let*

$$\mathcal{V}_2(n) := C_1(n) \bigcup C_2(n) \bigcup C_3(n) \bigcup C_4(n),$$

$\mathcal{V}_2(n)$ *is a fully diverse constellation of* n^2 *elements with diversity product*

$$\prod \mathcal{V}_2(n) = \min \left\{ r \sin \frac{2\pi}{n}, \sin \frac{\pi}{n} \right\}.$$

Proof. Take two elements $A \in C_i(n)$ and $B \in C_j(n)$. Without loss of generality, we can further assume $i \leq j$. If $i = j$, we have

$$|\det(A-B)| \geq |re^{i\frac{2\pi}{m}k} - re^{i\frac{2\pi}{m}(k+1)}|^2 = 4r^2 \sin^2 \frac{\pi}{m}.$$

If $A \in C_1(n), B \in C_2(n)$ or $A \in C_3(n), B \in C_4(n)$, one can check that

$$|\det(A-B)| \geq 2|re^{i\frac{2\pi}{m}k} - \sqrt{1-r^2}e^{i(\frac{2\pi}{m}k+\frac{\pi}{m})}| = 2(1 - 2\sqrt{1-r^2}r\cos\frac{\pi}{m}).$$

Similarly if $A \in C_1(n)$, $B \in C_3(n)$ or $A \in C_2(n)$, $B \in C_4(n)$,

$$|\det(A-B)| \geq 2(\sqrt{1-r^2} - r)^2 = 2(1 - 2\sqrt{1-r^2}r).$$

If $A \in C_1(n)$, $B \in C_4(n)$ or $A \in C_2(n)$, $B \in C_3(n)$,

$$|\det(A-B)| \geq |re^{i\frac{2\pi}{m}k} - re^{i\frac{2\pi k}{m}+\frac{\pi}{m}}|^2 + |\sqrt{1-r^2}e^{i\frac{2\pi}{m}k} - \sqrt{1-r^2}e^{i\frac{2\pi k}{m}+\frac{\pi}{m}}|^2 = 4\sin^2\frac{\pi}{2m}.$$

It follows from the definition of r that

$$2(1-2\sqrt{1-r^2}r) = 4r^2\sin^2\frac{\pi}{m},$$

and naturally we will have

$$2\left(1-2\sqrt{1-r^2}r\cos\frac{\pi}{m}\right) \geq 2(1-2\sqrt{1-r^2}r).$$

Compare the lower bounds in all the cases and take the minimum of them, we establish the claim in the theorem. □

The following table shows how the diversity product of $\mathcal{V}_2(n)$ compares with that of $\mathcal{O}(n)$ when $n \leq 14$.

n	r	$r\sin\frac{2\pi}{n}$	$\sin\frac{\pi}{n}$	$\prod\mathcal{V}_2(n)$	$\prod\mathcal{O}(n)$
4	0.383	0.383	0.707	0.383	0.5
6	0.410	0.355	0.500	0.355	0.354
8	0.447	0.316	0.355	0.316	0.271
10	0.479	0.282	0.309	0.282	0.219
12	0.505	0.253	0.259	0.253	0.183
14	0.527	0.229	0.222	0.222	0.157

Corollary 5.4. *For $n \geq 14$,*

$$\prod\mathcal{V}_2(n) = \sin\frac{\pi}{n}.$$

Proof. According to the definition of r,

$$2(1-2\sqrt{1-r^2}r) = 4r^2\sin^2\frac{\pi}{m}.$$

One can also check that

$$4r^2\sin^2\frac{\pi}{m} = 16r^2\sin^2\frac{\pi}{2m}\cos^2\frac{\pi}{2m} = 4r^2\cos^2\frac{\pi}{2m}\left(4\sin^2\frac{\pi}{2m}\right)$$

$$= \frac{4\cos^2\frac{\pi}{2m}}{2\sin^2\frac{\pi}{m}+2\sqrt{2}\sin\frac{\pi}{m}+2}\left(4\sin^2\frac{\pi}{2m}\right) \geq \frac{4\cos^2\frac{\pi}{14}}{2\sin^2\frac{\pi}{7}+2\sqrt{2}\sin\frac{\pi}{7}+2}\left(4\sin^2\frac{\pi}{2m}\right)$$
$$\geq 4\sin^2\frac{\pi}{2m}.$$

Consequently one concludes that

$$\prod \mathcal{V}_2(n) = \min\left\{r\sin\frac{\pi}{m}, \sin\frac{\pi}{2m}\right\} = \sin\frac{\pi}{n}.$$

□

For $n \geq 14$, the constellation $\mathcal{V}_2(n)$ has as many elements as the constellation $\mathcal{O}(n)$, however its diversity product is larger than that of $\mathcal{O}(n)$ by a factor of $\sqrt{2}$, *i.e.*, $\prod \mathcal{V}_2(n) = \sqrt{2}\prod \mathcal{O}(n)$. The corollary indicates that we could allocate different power to the transmitting antennas to achieve more reliable transmission, while still keeping the total energy. Similar to the case for $\mathcal{V}_1(n)$, one can apply exactly the same algorithm to achieve the ML decoding for $\mathcal{V}_2(n)$. It can be easily seen that these two algorithms have the same complexity.

5.2.3 Construction 3

We will take further efforts to explore the subsets of $SU(2)$. In the following we will describe a series of unitary constellation $\mathcal{V}_3(n)$. For the sample program implemented to construct $\mathcal{V}_3(n)$, we refer to [18]. Now for given integers $n > 0$ and $0 \leq k \leq n$, we define

$$N_0 = 1, N_k = \frac{\pi}{\left\lfloor \arcsin\frac{\sin\frac{\pi}{4n}}{\cos\frac{(n-k)\pi}{2n}}\right\rfloor}, k = 1, 2, \cdots, n,$$

$$M_k = \frac{\pi}{\left\lfloor \arcsin\frac{\sin\frac{\pi}{4n}}{\sin\frac{(n-k)\pi}{2n}}\right\rfloor}, k = 0, 1, \cdots, n-1, M_n = 1.$$

An example (see Figure 5.4) shows how to choose "multi-layered" vectors $a_{k,j}$'s and $b_{k,l}$'s.

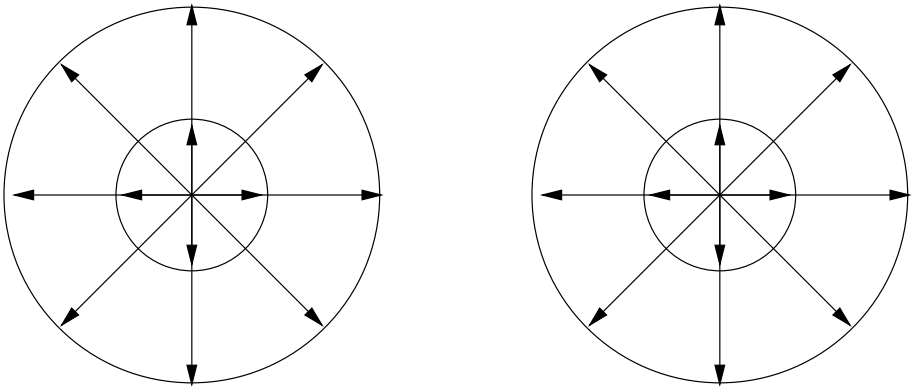


Figure 5.4. The values of $a_{k,j}$ and $b_{k,l}$

$$C_k(n) = \left\{ \left(\begin{array}{cc} a_{k,j} & b_{k,l} \\ -\bar{b}_{k,l} & \bar{a}_{k,j} \end{array} \right) \middle| \; a_{k,j} = \cos\frac{(n-k)\pi}{2n} e^{i\frac{2j\pi}{N_k}}, b_{k,l} = \sin\frac{(n-k)\pi}{2n} e^{i\frac{2l\pi}{M_k}} \right\}.$$

Theorem 5.5. *Let*

$$\mathcal{V}_3(n) = \bigcup_{k=0}^{n} C_k(n),$$

$\mathcal{V}_3(n)$ *is a fully diverse constellation of* $\sum_{k=0}^{n} M_k N_k$ *elements with diversity product*

$$\prod \mathcal{V}_3(n) = \sin\frac{\pi}{4n}.$$

Proof. Pick two distinct elements $A, B \in \mathcal{V}$,

$$A = \left(\begin{array}{cc} a & b \\ -\bar{b} & \bar{a} \end{array} \right), \qquad B = \left(\begin{array}{cc} c & d \\ -\bar{d} & \bar{c} \end{array} \right).$$

One can verify that

$$|\det(A-B)| = \det(A-B) = |a-c|^2 + |b-d|^2.$$

So if $|a| \neq |c|$, then we have

$$|\det(A-B)| = |a-c|^2 + |b-d|^2 \geq (|a|-|c|)^2 + (|b|-|d|)^2 \geq 2 - 2\cos\frac{\pi}{2n},$$

and it can be verified that the equality holds if there is a $k \in \{0, 1, \cdots, n\}$ such that $A \in C_k(n)$ and $B \in C_{k+1}(n)$ or alternatively $B \in C_k(n)$ and $A \in C_{k+1}(n)$. In the case that $|a| = |c|$, we will have

$$|\det(A-B)| = |a-c|^2 + |b-d|^2 \geq \max\{|a-c|^2, |b-d|^2\} \geq 2 - 2\cos\frac{\pi}{2n}.$$

Therefore for all the cases, we have

$$|\det(A-B)| \geq 2 - 2\cos\frac{\pi}{2n}.$$

One checks that the lower bound for each case can be reached. So for constellation $\mathcal{V}_3(n)$, it follows that

$$\prod \mathcal{V}_3(n) = \frac{1}{2}(2 - 2\cos\frac{\pi}{2n})^{\frac{1}{2}} = \sin\frac{\pi}{4n}.$$

□

Corollary 5.6. *When $n \to \infty$, $\mathcal{V}_3(n)$ will have $O(n^3)$ elements and have the diversity product $O(\frac{1}{n})$.*

Proof. For $k = 1, 2, \cdots, n$,

$$N_k = \frac{\pi}{\left\lfloor \arcsin \frac{\sin\frac{\pi}{4n}}{\cos\frac{(n-k)\pi}{2n}} \right\rfloor} \leq \frac{\pi}{\left\lfloor \arcsin\sin\frac{\pi}{4n} \right\rfloor} \leq 4n.$$

Similarly for $k = 0, 1, \cdots, n-1$,

$$M_k = \frac{\pi}{\left\lfloor \arcsin \frac{\sin\frac{\pi}{4n}}{\sin\frac{(n-k)\pi}{2n}} \right\rfloor} \leq \frac{\pi}{\left\lfloor \arcsin\sin\frac{\pi}{4n} \right\rfloor} \leq 4n.$$

Hence we derive an asymptotic upper bound for the cardinality of $\mathcal{V}_3(n)$,

$$|\mathcal{V}_3(n)| = \sum_{k=0}^{n} |C_k(n)| = \sum_{k=0}^{n} M_k N_k \leq 2n + 16(n-1)n^2 \leq O(n^3).$$

In the following we shall derive an asymptotic lower bound for $|\mathcal{V}_3(n)|$. First we pick two real numbers α, β such that

$$0 < \alpha < \beta < 1.$$

For k such that $\alpha n < k < \beta n$ (such k always exists when n is large enough), we have

$$N_k = \frac{\pi}{\left\lfloor \arcsin \frac{\sin \frac{\pi}{4n}}{\cos \frac{(n-k)\pi}{2n}} \right\rfloor} \geq \frac{\pi}{\left\lfloor \arcsin \frac{\sin \frac{\pi}{4n}}{\cos \frac{(1-\alpha)\pi}{2}} \right\rfloor}.$$

Utilizing the fact that

$$\sin x \sim x \sim \arcsin x,$$

when x is close to 0. When n is large enough, one derives

$$N_k \geq 4 \cos \frac{(1-\alpha)\pi}{2} n.$$

Similarly

$$M_k \geq 4 \sin \frac{(1-\beta)\pi}{2} n.$$

Therefore we have an asymptotic lower bound for the cardinality of $\mathcal{V}_3(n)$,

$$|\mathcal{V}_3(n)| \geq \sum_{\alpha n < k < \beta n} M_k * N_k \geq 16(\beta - \alpha) \sin \frac{(1-\beta)\pi}{2} \cos \frac{(1-\alpha)\pi}{2} n^3 \geq O(n^3).$$

Combining the upper bound and lower bound, we conclude that $\mathcal{V}_3(n)$ has $O(n^3)$ elements with diversity product $\sin(\pi/(4n)) = O(1/n)$. □

This corollary indicates that asymptotically $\mathcal{V}_3(n)$ will have much better diversity product compared to $\mathcal{O}(n), \mathcal{V}_1(n), \mathcal{V}_2(n)$, because the other three constellations asymptotically will have $O(n^2)$ elements and have diversity product $O(\frac{1}{n})$. Observe that $SU(2)$ in fact can be viewed as a 3 dimensional unit sphere. Finding a constellation with the optimal diversity product can be converted to a sphere packing problem on this unit sphere. For n points on a 3 dimensional unit sphere, the largest minimum distance one can hope for is asymptotically $O(1/n^{1/3})$. Therefore when n becomes large, asymptotically $\mathcal{V}_3(n)$ is the best constellation over all the subsets of $SU(2)$. Table 5.1 shows the diversity product of constellation $\mathcal{V}_3(n)$.

The ML decoding of $\mathcal{V}_3(n)$ is also decomposable and it roughly requires $4N + 2n$ complex multiplications and $4N + 2n$ complex additions, therefore asymptotically it requires $O(N) + O(L^{1/3})$ complex multiplications and additions. One can see

Table 5.1. The diversity product of $\mathcal{V}_3(n)$

n	size	diversity product
3	124	0.259
4	293	0.195
5	582	0.156
6	974	0.131
7	1640	0.112
8	2438	0.098
9	3510	0.087
10	4898	0.078
11	6516	0.071
12	8433	0.065
13	10770	0.060

from Table 5.1 that for reasonably small values of n the constellation size grows rapidly. For example when the constellation size L is already huge (10770), n is still reasonably small (13). So even if we are dealing with a high rate constellation, the decoding process of $\mathcal{V}_3(n)$ is still very simple.

We compare different codes from GPSK signals with transmission rate around 4.5 in Figure 5.5. One can see that $\mathcal{V}_3(5)$ outperforms other constellations even with the highest transmission rate. Of course the decoding of $\mathcal{V}_3(5)$ is a little more complex than other constellations, however the sacrifice in decoding efficiency is worthwhile for the remarkably gained performance.

We compare the GPSK constellations with the Cayley codes [24] in Figure 5.6. One can see that $\mathcal{V}_1(44)$ with transmission rate $R = 5.9594$ has a gain of about 2dB compared to the Cayley code, and the performance of $\mathcal{V}_3(9)$ with $R = 5.8886$ and $\mathcal{V}_3(10)$ with $R = 6.1290$ are even more remarkable. Note that all the GPSK constellations admit the presented simple decoding algorithms. $\mathcal{V}_1(44)$ or $\mathcal{V}_2(64)$ only needs about 8 complex multiplications and 8 complex additions to decode one

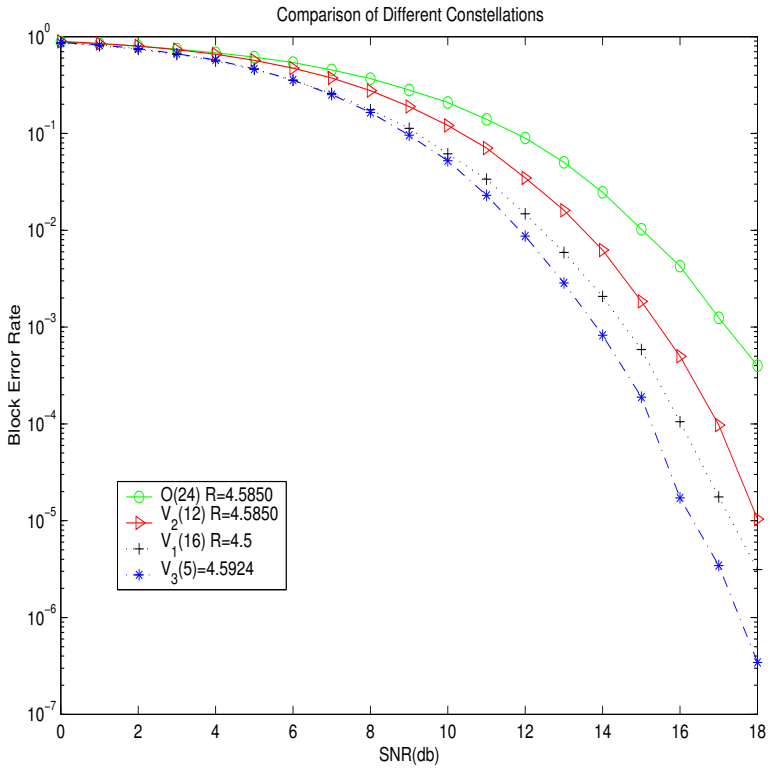


Figure 5.5. Performance of different codes from GPSK signals, M=2, N=12

codeword, $\mathcal{V}_3(9)$ or $\mathcal{V}_3(10)$ needs about 28 complex multiplications and 28 complex additions to decode.

5.3 GPSK Constellations from the Real Orthogonal Designs

In [43] space-time block coding has been discussed for coherent channels. The authors classify all the real orthogonal designs and show that orthogonal design codes only exist for $2, 4$ or 8 dimension. We show that the normalized version of the real orthogonal codes can be used in unitary space time modulation for non-coherent scenarios as well. We will present explicit implementations of the real orthogonal designs, which feature simple decoding algorithms. For simplicity we describe our

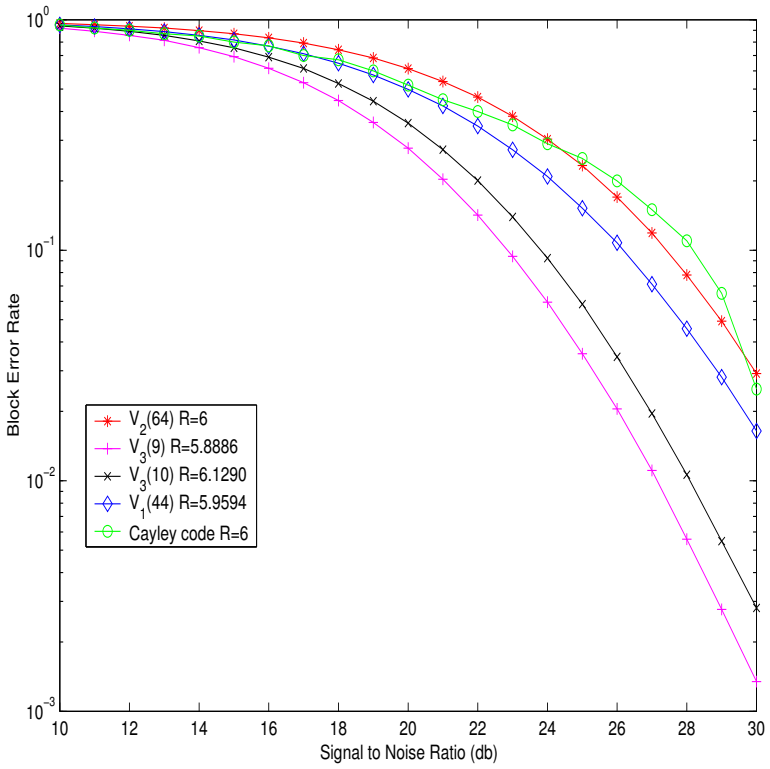


Figure 5.6. Performance of GPSK constellations and a Cayley code, M=2, N=2

idea using a 4 dimensional real orthogonal design:

$$S = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 \\ -s_2 & s_1 & -s_4 & s_3 \\ -s_3 & s_4 & s_1 & -s_2 \\ -s_4 & -s_3 & s_2 & s_1 \end{pmatrix}. \tag{5.5}$$

One natural idea is to transform a 2 dimensional complex orthogonal code $\mathcal{O}(n)$ to a 4 dimensional real orthogonal code. Recall that $\mathcal{O}(n)$ is a set of matrices

$$\mathcal{O}(a,b) = \frac{1}{\sqrt{2}} \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}$$

with a, b ranging over $\vartheta = \{1, e^{2\pi i/n}, \cdots, e^{2\pi i(n-1)/n}\}$. Let

$$s_1 = \mathrm{Re}(a), s_2 = Im(a), s_3 = \mathrm{Re}(b), s_4 = Im(b),$$

we will have a 4 dimensional real orthogonal code with the same diversity product as $\mathcal{O}(n)$.

This code also features a simple decoding algorithm. The simple decoding of real orthogonal design codes has been explained in [43], in the following we will discuss the decoding process in more details such that it can be generalized to our constellations straightforwardly. First we introduce some notations to be used. The superscript T will be the notation for a matrix transpose. For two vectors F and G, let $F \cdot G$ denote the dot product of F and G. For a vector F, F_j denotes the j-th element of F. For a matrix F, F_j denotes the j-th column vector. Let B be the mapping

$$B(x_1, x_2, x_3, x_4) = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & -x_1 & x_4 & -x_3 \\ x_3 & -x_4 & -x_1 & x_2 \\ x_4 & x_3 & -x_2 & -x_1 \end{pmatrix},$$

immediately one checks that

$$S \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = B(x_1, x_2, x_3, x_4) \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix}. \tag{5.6}$$

Assume that the differential unitary space time modulation is used for a wireless communication system with M transmitting antennas and N receiving antennas. Let X, Y denote the received $M \times N$ matrices at time block $\tau - 1$ and τ, respectively. The ML decoder will make the following estimation:

$$\hat{S} = \arg\min_S \|Y - SX\|^2.$$

We separate X, Y into the real parts and the imaginary parts with

$$X = F + iG, \qquad Y = P + iQ,$$

then we have

$$\hat{S} = \arg\min_{S} \left(\|P - SF\|^2 + \|Q - SG\|^2 \right)$$
$$= \arg\min_{S} \left(\sum_{j=1}^{N} \|P_j - SF_j\|^2 + \sum_{j=1}^{N} \|Q_j - SG_j\|^2 \right).$$

Use Equation (5.6) and let $\mathbf{s} = (s_1, s_2, s_3, s_4)$, we have

$$\hat{S} = \arg\min_{\mathbf{s}} \left(\sum_{j=1}^{N} \left\|P_j - B(F_j)\mathbf{s}^T\right\|^2 + \sum_{j=1}^{N} \left\|Q_j - B(G_j)\mathbf{s}^T\right\|^2 \right).$$

Since $B(F_j), B(G_j)$ are orthogonal matrices, simple algebraic manipulations will lead the above evaluation to

$$\arg\min_{\mathbf{s}} \sum_{j=1}^{N} \left(\frac{1}{|F_j|^2} \left\|B^T(F_j)P_j - |F_j|^2\mathbf{s}^T\right\|^2 + \frac{1}{|G_j|^2} \left\|B^T(G_j)Q_j - |G_j|^2\mathbf{s}^T\right\|^2 \right)$$

$$= \arg \max_{s_1,s_2,s_3,s_4} (s_1, s_2, s_3, s_4) \cdot (\sum_{j=1}^{N} B^T(F_j)P_j + \sum_{j=1}^{N} B^T(G_j)Q_j) = \arg\max_{s_j} \sum_{j=1}^{N} s_j U_j,$$

where $U_j = \sum_{j=1}^{N} B^T(F_j)P_j + \sum_{j=1}^{N} B^T(G_j)Q_j$. Since the construction of our code is based on $\mathcal{O}(n)$ as described above, one can check that (s_1, s_2) in fact is independent of (s_3, s_4). Rewrite $s_1 + is_2 = \frac{1}{\sqrt{2}} e^{2k\pi i/n}$ and $s_3 + is_4 = \frac{1}{\sqrt{2}} e^{2l\pi i/n}$, then from

$$\hat{S} = \arg \max_{s_1,s_2,s_3,s_4} \left((s_1, s_2) \cdot (U_1, U_2)^T + (s_3, s_4) \cdot (U_3, U_4)^T \right),$$

we conclude that the ML decoding of this GPSK constellation is decomposable and can be boiled down to the following simple form:

$$\hat{k} = \left\lfloor \frac{2\pi \arg(U_1 + iU_2)}{n} \right\rceil, \qquad \hat{l} = \left\lfloor \frac{2\pi \arg(U_3 + iU_4)}{n} \right\rceil.$$

Generally speaking the use of the proposed codes above for a wireless communication system with M transmitting antennas and N receiving antennas will take $8M^2N$

real multiplications and $8M^2N$ real additions to decode one codeword, which is very simple.

Apparently we can apply the same idea for $\mathcal{V}_i(n)$ $(i = 1, 2, 3)$ to construct GPSK real orthogonal constellations. The ML decoder of the corresponding codes will take similar approaches as in the $\mathcal{O}(n)$ case, except that one should notice that (s_1, s_2) is not independent of (s_3, s_4) anymore. In this case one can restrict the searching area to be the subsets $C_i(n)$ "locally" and apply the similar techniques as we did in the complex orthogonal design case to achieve the ML decoding, then the corresponding codes will admit simple decoding algorithms too. Also we conclude that codes from $\mathcal{V}_1(n)$ or $\mathcal{V}_2(n)$ are of the same complexity as codes from $\mathcal{O}(n)$. For the codes from $\mathcal{V}_3(n)$, the complexity is increased for high transmission rates, however they will have more pronounced performance.

Based on the idea discussed above, we shall present a series of 8 dimensional GPSK real orthogonal constellation $\mathcal{V}_4(n)$. Consider the 8 dimensional orthogonal design:

$$\begin{pmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ -s_2 & s_1 & s_4 & -s_3 & s_6 & -s_5 & -s_8 & s_7 \\ -s_3 & -s_4 & s_1 & s_2 & s_7 & s_8 & -s_5 & -s_6 \\ -s_4 & s_3 & -s_2 & s_1 & s_8 & -s_7 & s_6 & -s_5 \\ -s_5 & -s_6 & -s_7 & -s_8 & s_1 & s_2 & s_3 & s_4 \\ -s_6 & s_5 & -s_8 & s_7 & -s_2 & s_1 & -s_4 & s_3 \\ -s_7 & s_8 & s_5 & -s_6 & -s_3 & s_4 & s_1 & -s_2 \\ -s_8 & -s_7 & s_6 & s_5 & -s_4 & -s_3 & s_2 & s_1 \end{pmatrix}. \tag{5.7}$$

Similar to the 4 dimensional real orthogonal constellation case, one can obtain 8 dimensional unitary codes by transforming $V(n) \times W(n)$, where V or W denotes any one of $\mathcal{O}, \mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3$ and $\times$ denotes the Cartesian product. For this implementation,

s_i can be assigned to be the scaled version of the real or imaginary part of "a" or "b" as in 4 dimensional case. However as in the complex orthogonal case $\mathcal{V}_3(n)$ motivates us to explore more densely packed constellations.

The problem of constructing an 8 dimensional real orthogonal code with the maximal diversity product is equivalent to the packing problem on a 7 dimensional unit sphere. Therefore, any currently existing results in packing problem on a 7 dimensional unit sphere can be "borrowed" for the real orthogonal code construction. However decoding such codes using exhaustive search will be unpractical for high transmission rates. In the sequel we will present a series of 8 dimensional GPSK orthogonal codes $\mathcal{V}_4(n)$ featuring simple decoding algorithms.

The basic idea is that instead of considering the problem of packing the real vector $\mathbf{s} = (s_1, s_2, \cdots, s_8)$ on a unit sphere, we consider the problem of packing the complex vector $\mathbf{z} = (z_1, z_2, z_3, z_4)$ with constraints $\sum_{j=1}^{4} |z_j|^2 = 1$, where $z_j = s_{2j-1} + is_{2j}$. Since the distance of two complex vectors satisfies

$$\sum_{j=1}^{4} |z_j - w_j|^2 \geq \sum_{j=1}^{4} (|z_j| - |w_j|)^2,$$

we can pack the amplitude vector $(|z_1|, |z_2|, |z_3|, |z_4|)$ first, then pack the argument vector $(\arg(z_1), \arg(z_2), \arg(z_3), \arg(z_4))$. To explain this implementation in more details, we represent $\mathbf{z}$ using the polar coordinates:

$$\begin{aligned}
z_1 &= \cos\theta_1 e^{i\gamma_1}, \\
z_2 &= \sin\theta_1 \cos\theta_2 e^{i\gamma_2}, \\
z_3 &= \sin\theta_1 \sin\theta_2 \cos\theta_3 e^{i\gamma_3}, \\
z_4 &= \sin\theta_1 \sin\theta_2 \sin\theta_3 e^{i\gamma_4}.
\end{aligned}$$

Let θ_j run over $m_j + 1$ evenly distributed discrete values from 0 to $\pi/2$:

$$\theta_j \in \left\{ k\frac{\pi}{2m_j} \mid k = 0, 1, \cdots, m_j \right\},$$

and let γ_j run over n_j evenly distributed discrete values from 0 to 2π:

$$\gamma_j \in \left\{ k\frac{2\pi}{n_j} \mid k = 0, 1, \cdots, n_j - 1 \right\}.$$

Now we have a finite set of complex vectors.

Let $\mathbf{v} = (v_1, v_2, v_3, v_4)$ and $\mathbf{w} = (w_1, w_2, w_3, w_4)$ be two distinct resulting complex vectors. If $|v_1| \neq |w_1|$, one can check that

$$|\mathbf{v} - \mathbf{w}|^2 \geq 4\sin^2(\pi/(4m_1)).$$

If $|v_1| = |w_1|$ but $|v_2| \neq |w_2|$, we will have

$$|\mathbf{v} - \mathbf{w}|^2 \geq 4(1 - |v_1|^2)\sin^2(\pi/(4m_2)).$$

In the case that $|v_1| = |w_1|$ and $|v_2| = |w_2|$ but $|v_3| \neq |w_3|$, similar algebraic calculations will give a lower bound

$$|\mathbf{v} - \mathbf{w}|^2 \geq 4(1 - |v_1|^2 - |v_2|^2)\sin^2(\pi/(4m_3)).$$

We further consider the case $|v_j| = |w_j|$ for $j = 1, 2, 3, 4$. In this case, if $v_j \neq w_j$ for some j, we can have

$$|\mathbf{v} - \mathbf{w}|^2 \geq 4|v_j|^2\sin^2(\pi/n_j).$$

It is easy to check that the lower bound can be reached for some special pair of $\mathbf{z}, \mathbf{w}$ in all the cases. Based on the 8 dimensional real orthogonal design (5.7), assign $s_{2j-1} = \mathrm{Re}(z_j)$ and $s_{2j} = Im(z_j)$ for $j = 1, 2, 3, 4$. Now we have a finite unitary constellation whose diversity product is

$$\min_{\mathbf{v},\mathbf{w}} \sqrt{|\mathbf{v} - \mathbf{w}|}/2.$$

However if we choose m_j, n_j arbitrarily, often we will end up having a constellation with a small diversity. In the sequel we add suitable constraints to the choices

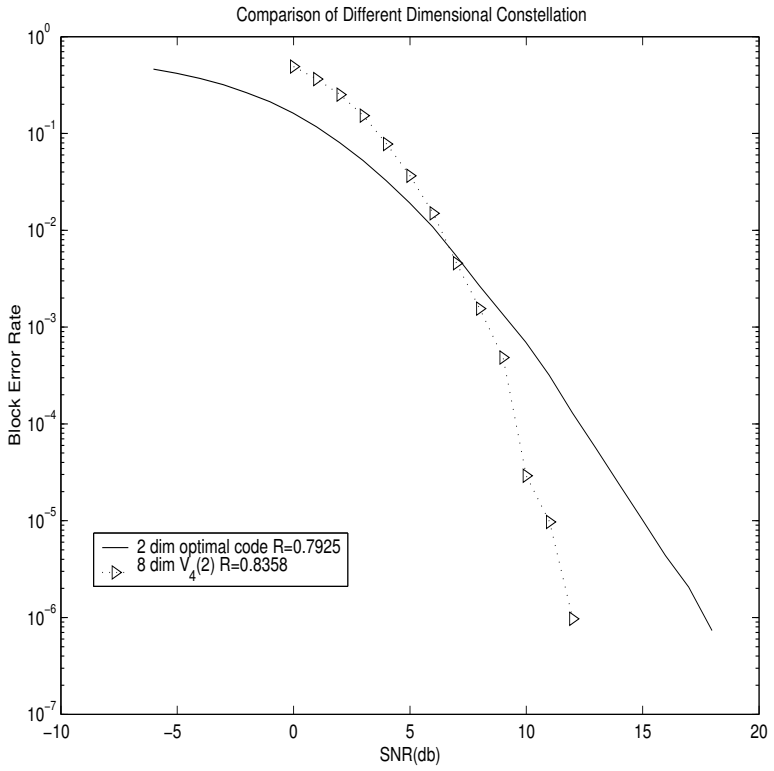


Figure 5.7. Performance of a GPSK constellation and a code with optimal diversity product, N=2

of m_j and n_j to guarantee that we have a large diversity product. We present the algorithm step by step. For the sample program of this algorithm, we refer to [18].

Algorithm to Construct $\mathcal{V}_4(n)$ (given the input n):

1. Fix $m_1 = n$ and let θ_1 run over $n+1$ even distributed discrete values from 0 to $\pi/2$ (consequently $|z_1|$ runs over $n+1$ discrete values from 0 to 1).

2. For any fixed $|z_1|$, take m_2 to be the largest integer such that

$$4(1-|z_1|^2)\sin^2(\pi/(4m_2)) \geq 4\sin^2(\pi/(4n)).$$

Let θ_2 run over m_2+1 evenly distributed discrete values from 0 to $\pi/2$.

3. For any fixed $|z_1|, |z_2|$, take m_3 to be the largest integer such that

$$4(1 - |z_1|^2 - |z_2|^2) \sin^2(\pi/(4m_3)) \geq 4 \sin^2(\pi/(4n)).$$

Let θ_3 run over $m_3 + 1$ evenly distributed discrete values from 0 to $\pi/2$.

4. For any fixed $|z_1|, |z_2|, |z_3|$, take n_j to be the largest integer such that

$$4|z_j|^2 \sin^2(\pi/n_j) \geq 4 \sin^2(\pi/(4n)).$$

Let γ_j run over n_j evenly distributed discrete values from 0 to 2π.

5. The steps above result in a finite set of complex vectors $\mathbf{z} = (z_1, z_2, z_3, z_4)$. Based on the 8 dimensional real orthogonal design (5.7), assign $s_{2j-1} = \mathrm{Re}(z_j)$ and $s_{2j} = Im(z_j)$ with $j = 1, 2, 3, 4$. Now we have a fully diverse unitary constellation with diversity product $\sin(\pi/(4n))$.

We call this constellation $\mathcal{V}_4(n)$. Recall that for the packing problem with n points on a 7 dimensional unit sphere, asymptotically the largest minimum distance one can hope for is $O(1/n^7)$. The following theorem indicates that asymptotically $\mathcal{V}_4(n)$ are the best constellations from 8 dimensional real orthogonal designs. Although the proof of this theorem is quite tedious, the basic idea is very similar to Corollary 5.6, thus we skip the proof.

Theorem 5.7. *$\mathcal{V}_4(n)$ is a fully diverse constellation with diversity product $\sin(\pi/(4n))$. When $n \to \infty$, $\mathcal{V}_4(n)$ will have have $O(n^7)$ elements and have the diversity product $O(1/n)$.*

Applying the same analysis for the constellation $\mathcal{V}_3(n)$, we conclude $\mathcal{V}_4(n)$ is also decomposable and the decoding complexity for one codeword will be $O(N) + O(L^{3/7})$. In Figure 5.7 we compare the performance of $\mathcal{V}_4(2)$ of 103 elements with a 2 dimensional constellation of 3 elements:

$$\{I_2, A, B\},$$

where $A = \mathrm{diag}\,(e^{i2\pi/3}, e^{i2\pi/3})$ and $B = A^2$. It is well known that this 2 dimensional constellation has the optimal diversity product over all the 2 dimensional unitary constellations with 3 elements. Since a large number of transmitting antennas guarantee that the full diversity at the transmitter side can be utilized more efficiently (see Inequality (3.9)), it is not too surprising to see that the 8 dimensional GPSK constellation performs better at high SNR region. The figure shows that from 6dB, 8 dimensional constellation begins to outperform the 2 dimensional one. Around Block Error Rate (BLER) of 10^{-6}, the performance gain is about 5dB, which is very pronounced considering the decoding of $\mathcal{V}_4(2)$ is relatively easy.

5.4 Conclusions

The complex and real orthogonal coding schemes admit simple decoding algorithms. Based on these schemes, we generalize one dimensional PSK signals and explicitly construct GPSK unitary space time constellations. These constellations can be viewed as higher dimensional generalizations of one dimensional PSK signals and theoretical analysis shows that their decoding procedures are decomposable, *i.e.*, the demodulation of these codes can be boiled down to one dimensional PSK demodulation. Therefore our constellations have very simple decoding procedures. For some of the resulting codes (for example, $\mathcal{V}_1(n), \mathcal{V}_2(n)$), the complexity of ML decoding does not even depend on the transmission rate. We reallocate the power among the antennas (meanwhile keeping the total energy) to optimize the diversity product. Numerical experiments show that our codes perform better than some of the currently existing comparable ones. For the sample programs regarding how to construct the proposed constellations, we refer to [18].

CHAPTER 6

GEOMETRICAL AND NUMERICAL DESIGN OF STRUCTURED UNITARY SPACE TIME CONSTELLATION

6.1 Introduction

The main purpose of this chapter is to present structured constellation and to develop geometrical and numerical procedures that allow one to construct unitary constellations with excellent diversity for any set of parameters M, N, T, L and for any signal to noise ratio ρ. The chapter is structured as follows. In Section 6.2, four illustrative examples of constellations designed from different approaches will be presented. We analyze their performance using the concepts of diversity product and diversity sum, which were introduced in Chapter 3.

In Section 6.3 we first show that randomly constructed codes are fully diverse with probability one. Then we start the main task of this chapter, namely to parameterize constellations which will be efficient for numerical search algorithms. For this purpose we introduce the concept of a *weak group structure* and we classify all weak group structures whose elements are normal and positive.

In Section 6.4 we investigate an algebraic structure which led to some of the best constellations which we were able to derive. Not surprisingly, we also show that in the good-performing codes the distance spectrum profile for both the diversity sum and the diversity product are important.

Section 6.5 is one of the main sections of this chapter. We first explain a general method for how one can efficiently design excellent constellations for any set of

parameters M, N, T, L and ρ. For this we review the properties of the complex Stiefel manifold and the Cayley transform. We conclude this section with an extensive table where we publish a large set of codes having some of the best diversity sums and diversity products in their parameter range. More extensive lists of codes with large diversity can be found on the website [18].

Finally in Section 6.6 we explain how the algebraic structure which underlies most of the derived codes can be used to have a fast decoding algorithm. Our simulations indicate that in the design of codes more attention should be given to the diversity sum (more generally diversity function) which previously has not been fully studied.

6.2 Four Illustrative Examples

The diversity sum and the diversity product govern the diversity function at low SNR and at high SNR, respectively. Codes optimized at these extreme values of the SNR-axis do not necessarily perform well on the "other side of the spectrum". In this subsection we illustrate the introduced concepts on four examples. All examples have about equal parameters, namely $T = 4$, $M = 2$ and the size L is 121 respectively 120. The first two examples are well studied examples from the literature. We derived the third and the fourth examples by geometrical design and numerical methods, respectively.

Orthogonal Design: This constellation has been considered by several authors [3, 41]. For our purpose we simply define this code as a subset of $SU(2)$:

$$\left\{ \frac{\sqrt{2}}{2} \begin{pmatrix} e^{\frac{2m\pi i}{11}} & e^{\frac{2n\pi i}{11}} \\ -e^{-\frac{2n\pi i}{11}} & e^{-\frac{2m\pi i}{11}} \end{pmatrix} | m, n = 0, 1, \cdots, 10 \right\}.$$

The constellation has 121 elements ($R = 3.45$) and the diversity sum and the diversity product are both equal to 0.1992.

Unitary Representation of $SL_2(\mathbb{F}_5)$: Shokrollahi et al. [41] derived a constellation using the theory of fixed point free representations whose diversity product is near optimal. This constellation appears as a unitary representation of the finite group $SL_2(\mathbb{F}_5)$ and we will refer to this constellation as the $SL_2(\mathbb{F}_5)$-constellation. The finite group $SL_2(\mathbb{F}_5)$ has 120 elements and this is also the size of the constellation. In order to describe the constellation let $\eta = e^{\frac{2\pi i}{5}}$ and define

$$P = \frac{1}{\sqrt{5}} \begin{pmatrix} \eta^2 - \eta^3 & \eta^1 - \eta^4 \\ \eta^1 - \eta^4 & \eta^3 - \eta^2 \end{pmatrix}, \quad Q = \frac{1}{\sqrt{5}} \begin{pmatrix} \eta^1 - \eta^2 & \eta^2 - \eta^1 \\ \eta^1 - \eta^3 & \eta^4 - \eta^3 \end{pmatrix}.$$

Then the constellation is given by the set of matrices $(PQ)^j X$,where $j = 0, 1, \cdots, 9$, and X runs over the set

$$\{I_2, P, Q, QP, QPQ, QPQP, QPQ^2, QPQPQ, QPQPQ^2,$$
$$QPQPQ^2P, QPQPQ^2PQ, QPQPQ^2PQP\}.$$

The constellation has rate $R = 3.45$ and $\prod SL_2(\mathbb{F}_5) = \sum SL_2(\mathbb{F}_5) = \frac{1}{2}\sqrt{\frac{(3-\sqrt{5})}{2}} \sim 0.3090$. The diversity product of this constellation is truly outstanding. For illustrative purposes we plotted in Figure 6.1 the exact diversity functions and the diversity function of this constellation.

Numerically Derived Constellation: Using simulated annealing algorithm we found after short computation a constellation with very good diversity sum. The constel-

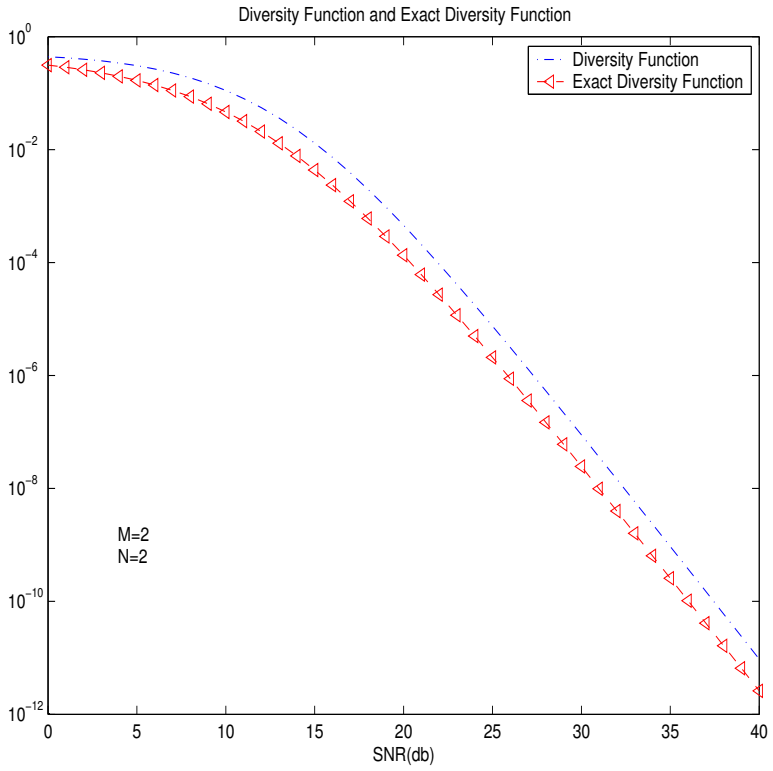


Figure 6.1. Diversity function $\mathcal{D}(\mathcal{V}, \rho)$ and exact diversity function for the group constellation $SL_2(\mathbb{F}_5)$.

lation is given through a set of 121 matrices

$$\left\{ \Psi_{k,l} := A^k B^l \mid A = \begin{pmatrix} -0.9049 + 0.3265 * i & 0.1635 + 0.2188 * i \\ 0.0364 + 0.2707 * i & -0.8748 + 0.4002 * i \end{pmatrix}, \right.$$
$$\left. B = \begin{pmatrix} -0.1596 + 0.9767 * i & -0.1038 + 0.0994 * i \\ 0.0833 - 0.1171 * i & -0.9432 + 0.2995 * i \end{pmatrix}, k, l = 0, 1, \cdots, 10 \right\}.$$

As we explain in Section 6.6, the maximum likelihood decoding of this constellation admits a simple decoding algorithm: sphere decoding.

Table 6.1. Four unitary constellations

	Orthogonal design	$SL_2(\mathbb{F}_5)$	Numerically derived	Geometrically designed
Number of elements	121	120	121	120
diversity sum	0.1992	0.309	0.3886	0.4156
diversity product	0.1992	0.309	0.0278	0.1464

Geometrically Designed Constellation: Based on the algebraic structure we propose in this chapter, we further implement the geometrical symmetry into this structure. A geometrically designed constellation can be described as follows:

$$\left\{ \Psi_k := A^k B^k | A = \begin{pmatrix} e^{17\pi/60} & 0 \\ 0 & e^{13\pi/60} \end{pmatrix}, \right.$$
$$\left. B = \begin{pmatrix} \cos(22\pi/60) & \sin(22\pi/60) \\ -\sin(22\pi/60) & \cos(22\pi/60) \end{pmatrix}, k = 0, 1, \cdots, 119 \right\}.$$

This constellation has superb diversity sum and reasonably good diversity product. One can also use sphere decoding to implement ML decoding of this constellation. Table 6.1 summarizes the parameters of the four constellations.

Of course we were curious about the performances of these four different codes. Figure 6.2 provides simulation results for each of the four constellations. Note that the numerically designed code which has a very bad diversity product is performing very well nevertheless due to the exceptional diversity sum. One can see that up to 12dB numerically derived codes outperform the group code by about 1dB. In fact, our simulation results show that until 35dB the numerical one is still performing much better than the orthogonal one. However at around 18dB, the group constellation surpasses the numerical one due to exceptional diversity product. The

geometrically designed constellation has better diversity sum and diversity product than the numerical one, therefore its performance is better than the numerical one (our results show that their performance curves are quite close, although the geometrical one is slightly better). These simulation results give an indication that the diversity sum is a very important parameter for a unitary constellation in low SNR regime.

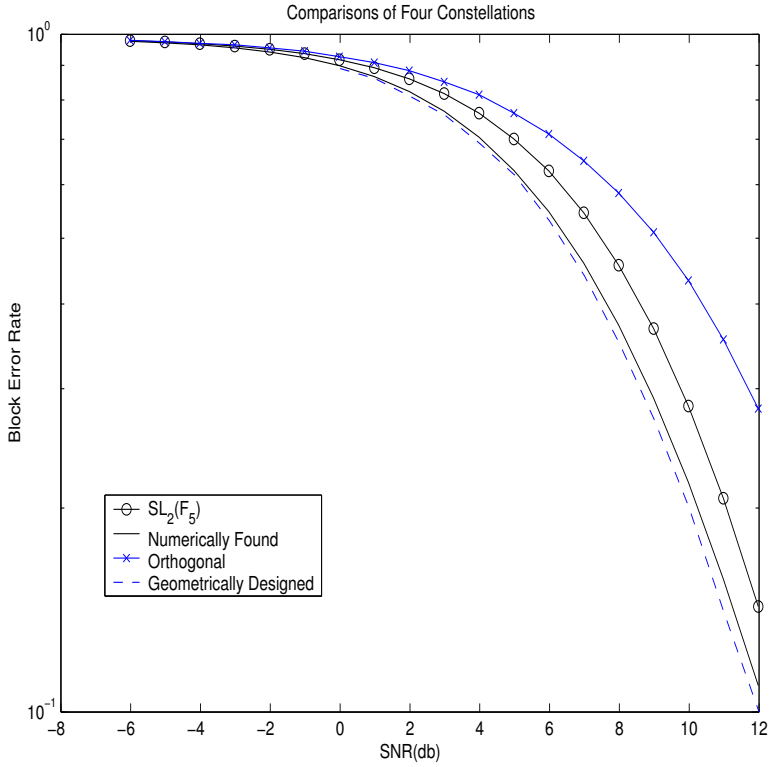


Figure 6.2. Simulations of four constellations having sizes $T = 4$, $M = 2$ and $L = 120$ respectively $L = 121$.

6.3 Constellations with Algebraic Structure

Before we venture into the realm of structured constellation, we would like to explore random unitary space time constellations first. We introduce the Haar distributed random matrix, which in some sense can be viewed as a high dimensional generalization of a complex random variable with circular symmetric distribution $\mathcal{CN}(0,1)$.

Definition 6.1. The Haar measure on $U(M)$ is defined to be a probability measure $\mathcal{H}$ on $U(M)$ which is translate invariant: for any mensurable set S in $U(M)$ and any fixed element U_0 in $U(M)$

$$\mathcal{H}(S) = \mathcal{H}(U_0 S).$$

A unitary random matrix $\mathbf{U}$ is Haar distributed (h.d.) if for any measurable set S we have

$$\Pr(\mathbf{U} \in S) = \mathcal{H}(S).$$

Remark 6.2. Note that h.d. matrix can be also proved to be a special case of isotropically distributed matrix in Chapter 1 or in [34]. We want to point out that Haar measure can be defined more generally. In fact every compact Lie group admits a unique (up to scalar) translate invariant measure: Haar measure [7].

A well known yet non-trivial fact is that for any measurable set $S \subset U(M)$, we have

$$\mathcal{H}(S) = \mathcal{H}(S^*),$$

where S^* consists of the conjugate transpose of all the elements in S. Thus for a h.d. matrix $\mathbf{U}$, one can verify

$$\Pr(\mathbf{U}^* \in S) = \Pr(U \in S^*) = \mathcal{H}(S^*) = \mathcal{H}(S).$$

Immediately we conclude $\mathbf{U}^*$ is also h.d. matrix. Also one can verify that the product of two h.d. matrices is still h.d. Another very interesting property about a h.d. matrix is its spectrum. As derived in [16], the joint probability density for the

eigenvalues of a h.d. random matrix $\mathbf{U} \sim \text{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_M})$ in $U(M)$ is given by the Weyl denominator formula:

$$f(\theta_1, \theta_2, \cdots, \theta_M) = \frac{1}{(2\pi)^M M!} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2.$$

The properties of h.d. matrices lead to the following theorem about random unitary space time constellations:

Theorem 6.3. *For a random unitary space time constellation $\mathcal{V}$ consisting of L h.d. independent random matrices $\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_L$, we have*

$$\Pr(\prod \mathcal{V} = 0) = 0.$$

In other words, the probability of $\mathcal{V}$ being fully diverse is 1.

Proof. First we can rewrite

$$\Pr(\prod \mathcal{V} = 0) = \Pr(\bigcup_{j<k} |\det(\mathbf{U}_j - \mathbf{U}_k)| = 0) \leq \sum_{j<k} \Pr(|\det(\mathbf{U}_j - \mathbf{U}_k)| = 0).$$

Next we shall show that the probability of the event $|\det(\mathbf{U}_j - \mathbf{U}_k)| = 0$ happening is 0. Now,

$$\Pr(|\det(\mathbf{U}_j - \mathbf{U}_k)| = 0) = \Pr(|\det(I - \mathbf{U}_j{}^*\mathbf{U}_k)| = 0).$$

Let $\mathbf{U}$ denote $\mathbf{U}_j{}^*\mathbf{U}_k$, we know $\mathbf{U}$ is h.d. matrix. Using the Weyl denominator formula, one computes

$$\Pr(|\det(I - \mathbf{U})| = 0) = \Pr\left(\bigcup_{l=1}^{M} \theta_l = 0\right)$$

$$\leq \frac{1}{(2\pi)^M M!} \sum_{l=1}^{M} \iint_{\theta_l=0} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_M.$$

Since

$$\iint_{\theta_j=0} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_M \leq 2^{M(M-1)} \iint_{\theta_j=0} d\theta_1 d\theta_2 \cdots d\theta_M = 0,$$

we conclude that

$$\Pr(|\det(\mathbf{U}_j - \mathbf{U}_k)| = 0) = 0.$$

Consequently

$$\Pr(\prod \mathcal{V} = 0) = 0,$$

that is the probability of $\mathcal{V}$ being fully diverse is 1. □

Note that if an $M \times M$ matrix G of independent complex Gaussian entries is input to QR algorithm, the resulting unitary matrix Q is Haar distributed [11]. For simplicity we sketch the proof as follows: First one can write $Q = GR^{-1}$, then for a fixed unitary matrix U_0, it can be checked that $U_0 G$ has the same distribution as G. Consequently $U_0 Q$ has the same distribution as Q, *i.e.*, the distribution of Q is translate invariant. Therefore the uniqueness of translate invariant measure on a compact Lie group guarantees that Q is Haar distributed. As a consequence of the above theorem, an algorithm which produces a fully diverse unitary constellation with probability 1 can be given as follows: take L instance of complex Gaussian matrices and feed them through the QR algorithm, the resulting L unitary matrices constitute a fully diverse constellation with probability 1.

From an algebraic geometry point of view one easily shows that the set of constellations with $\prod \mathcal{V} = 0$ forms a lower dimensional proper algebraic sub-variety of $U(M)^L$. In particular the set of all the fully diverse constellations is Zariski open [23] in $U(M)^L$, *i.e.*, fully diverse constellations are dense in $U(M)^L$. Haar distributed random constellations will not be practical for maximum likelihood decoding in high transmission rate scenario because no algebraic structure is assumed for random constellation and therefore the decoding process will be too complex. In the sequel we investigate structured constellations and explain how one can restrict the parameter space to judiciously chosen subsets and how one can convert maximum likelihood decoding to lattice decoding by using structured constellations.

Note that Haar measure and the corresponding distribution can be defined on Stiefel manifold $\mathcal{S}(T, M)$ (see Section 6.5) and Grassmannian $G(T, M)$, which are

quotient spaces of $U(T)$. Therefore the argument can be applied to non-square constellation case with $T \geq 2M$, although it is hard to get a closed form expression for the joint probability density function of the θ_i's. The intuition behind the fully diversity is: if $T \geq 2M$, the probability of two M-dimensional subspace intersecting with each other non-trivially is 0. Namely the column space of U_i and U_j (here they are $T \times M$ non-square constellations) will not intersect with probability 1, that means the full diversity of the whole constellation.

Theorem 6.4. *Given $T \geq 2M$ and a random unitary space time constellation $\mathcal{V}$ consisting of L h.d. independent random $T \times M$ matrices $\mathbf{\Phi}_1, \mathbf{\Phi}_2, \cdots, \mathbf{\Phi}_L$, we have*

$$\Pr(\prod \mathcal{V} = 0) = 0,$$

that is the probability of $\mathcal{V}$ being fully diverse is 1.

Consider a general constellation of square unitary matrices,

$$\mathcal{V} = \{\Psi_1, \Psi_2, \cdots, \Psi_L\}.$$

In order to calculate the diversity product, one needs to do $\frac{L(L-1)}{2}$ calculations: $|\det(\Psi_i - \Psi_j)|$ for every different pair i, j. The same statement can be said about the diversity sum, however for simplicity we only show the diversity product case in the sequel unless specified otherwise.

If one deals with a group constellation then one only needs to calculate $L - 1$ such determinant calculations and this is one of the remarkable advantages of group constellations. This is a direct consequence of

$$|\det(\Psi_i - \Psi_j)| = |\det(\Psi_i)\det(I - \Psi_i^*\Psi_j)| = |\det(I - \Psi_i^*\Psi_j)|,$$

where $\Psi_i^*\Psi_j$ is still in the group.

As we mentioned before group constellations are however very restrictive about what the algebraic structure is concerned. In the following we will present some

constellations which have some small number of generators and whose diversity can be efficiently computed. This will ensure that the total parameter space to be searched is limited as well. We start with an example:

Example 6.5. Consider the constellation

$$\mathcal{V} = \{A^k B^l | A, B \in U(M), k = 0, \cdots, p, l = 0, \cdots, q\}.$$

The parameter space for this constellation is $U(M) \times U(M)$, this is a manifold of dimension $2M^2$ and the number of elements in $\mathcal{V}$ is $(p+1)(q+1)$. If one has to compute $|\det(\Psi_i - \Psi_j)|$ for every distinct pair this would require $\begin{pmatrix} (p+1)(q+1) \\ 2 \end{pmatrix} \sim (pq)^2$ determinant calculations. We will show in the following that the same result can be obtained by just doing $2pq + p + q \sim 2pq$ determinant computations.

Let Ψ_i and Ψ_j be two distinct elements having the form $A^{k_1} B^{l_1}$ and $A^{k_2} B^{l_2}$, respectively. We have now several cases. When $k_1 = k_2$, then necessarily $l_1 \neq l_2$ and the distance is computed as

$$|\det(A^{k_1} B^{l_1} - A^{k_2} B^{l_2})| = |\det(I - B^{|l_2 - l_1|})|,$$

where $|l_2 - l_1|$ is an integer between 1 and q. If $l_1 = l_2$, then we have $k_1 \neq k_2$ and the distance is computed as

$$|\det(A^{k_1} B^{l_1} - A^{k_2} B^{l_2})| = |\det(I - A^{|k_2 - k_1|})|,$$

where $|k_2 - k_1|$ is an integer between 1 and p. If $(k_1 < k_2 \text{ and } l_1 < l_2)$ or $(k_1 > k_2 \text{ and } l_1 > l_2)$, we have

$$|\det(A^{k_1} B^{l_1} - A^{k_2} B^{l_2})| = |\det(I - A^{|k_2 - k_1|} B^{|l_2 - l_1|}),$$

where $1 \leq |k_2 - k_1| \leq p$ and $1 \leq |l_2 - l_1| \leq q$. Similarly if $(k_1 < k_2 \text{ and } l_1 > l_2)$ or $(k_1 > k_2 \text{ and } l_1 < l_2)$ then

$$|\det(A^{k_1} B^{l_1} - A^{k_2} B^{l_2})| = |\det(A^{|k_2 - k_1|} - B^{|l_2 - l_1|})|,$$

with $1 \leq |k_2 - k_1| \leq p$ and $1 \leq |l_2 - l_1| \leq p$. The total number of distances to be computed is in total equal to $2pq + p + q$.

The number of distances to be computed indicates how complex the calculation for the diversity is. In fact the smaller this number is, intuitively the larger possibility of finding a unitary constellation with good diversity we will have. An immediate observation is that for two pairs of unitary matrices (A, B) and (C, D), if $(C, D) = (UAV, UBV)$ or $(C, D) = (UA^{-1}V, UB^{-1}V)$, then $|\det(C - D)| = |\det(A - B)|$. We consider several constellations starting with this observation.

Example 6.6. Consider the case that $G \subset U(n)$ is a subgroup with L elements, then for any two distinct elements $A, B \in G$ we have $|\det(A-B)| = |\det(I-A^{-1}B)|$ with $A^{-1}B \in G$. Therefore at most $L - 1$ distance calculations are needed to derive the diversity product. The product of two group constellations has a similar property. Consider $G_i \subset U(M)$ with order l_i, where $i = 1, 2$. Let

$$G = \{AB | A \in G_1, B \in G_2\}.$$

Since $|\det(A_1B_1 - A_2B_2)| = |\det(I - A_1^{-1}A_2B_2B_1^{-1})|$ with $A_1^{-1}A_2B_2B_1^{-1} \in G$, at most $L - 1$ calculations are needed in this case, where $L = l_1l_2$.

Example 6.7. Consider a constellation with the following form:

$$\{A^iB^j | i = 0, \cdots, l_1 - 1, j = 0, \cdots, l_2 - 1 \text{ and } A, B \in U(M), A^{l_1} = I, B^{l_2} = I\}.$$

It can be checked that for the above constellation at most $L - 1$ calculations are needed, where $L = l_1l_2$.

Group structures do have certain advantages for constructing unitary constellations: it is less complex to calculate the diversity product (or sum); the possibility of finding a large diversity constellation intuitively may be increased. However the constellations found by this approach [41] are really few and far between. Somehow one wonders if the group structure is too restrictive to find a good-performing constellation.

In the sequel we will loosen the constraints imposed by the group structures. As demonstrated in Example 6.5 it is desirable to have a small dimensional manifold (in

Example 6.5 it was $U(M)\times U(M)$) that parameterizes a set of potentially interesting constellations. Having such a parameterization will help to avoid the problem of "dimension explosion". The set of constellations parameterized by $U(M)\times U(M)$ in Example 6.5 are interesting as we are not required to compute all pairwise distances in order to compute the diversity product (sum).

Definition 6.8. Let X be the set $\{x_1, x_2, \cdots, x_n\}$ and F be the free group on the set X. A subset $G \subset U(M)$ is called *freely generated* if there are elements $\{g_1, g_2, \cdots, g_n\} \subset G$ such that the homomorphism $\phi : F \longrightarrow G$ with $\phi(x_i) = g_i$ is an isomorphism.

An immediate consequence of this definition is that every element in G can be uniquely written as a product of g_i's and g_i^{-1}'s. The elements g_i are called the generators of G. A freely generated subset G is simply parameterized by the set:

$$\{a_1^{p_1} a_2^{p_2} \cdots a_k^{p_k} \mid a_i \text{ is one of } g_i's,\ p_i \in \mathbb{Z}\}.$$

Take an element $g \in G$ with its presentation $g = \prod_{i=1}^{k} a_i^{p_i}$, we say that the presentation is *reduced* whenever $a_i \neq a_{i+1}$ for $i = 1, \ldots, n-1$. Observe that taking the product of distinct matrices $\prod_{i=1}^{n} A_i$ is numerically expensive, however taking the power of one matrix A^k is much easier (note that for $A = U \sum U^{-1}$ with $\sum$ diagonal, we have $A^k = U \sum^k U^{-1}$). Moreover, by considering the powers of matrices, we are able to impose the lattice structure to the constellation, which makes sphere decoding of structured constellations possible. (see Section 6.6) Therefore we are interested in "normal" elements of G.

Definition 6.9. We say that an element $g = \prod_{i=1}^{k} a_i^{p_i}$ in reduced form is a *normal element* whenever $a_i \neq a_j$ for $i \neq j$. A subset $\mathcal{V}$ of the freely generated set G is said to be a *normal constellation* if every non-identity element in $\mathcal{V}$ is normal.

In the following we limit our searches to positive constellations:

Definition 6.10. An element g in G with the reduced form $g = \prod_{i=1}^{k} a_i^{p_i}$ is said to be a *positive element* if $p_i > 0$ for $i = 1, 2, \cdots, k$. A subset $\mathcal{V}$ of the freely generated set G is said to be a *positive constellation* if every non-identity element in $\mathcal{V}$ is positive.

Positive normal constellations are desirable for numerical searches as they can be efficiently parameterized and searched. If one wants to compute the diversity product (or sum) of an arbitrary positive constellation with L elements one still has to compare a total of $\binom{L}{2}$ pairs of matrices. In the sequel we will impose more structure on a constellation $\mathcal{V} \subset G$ which will guarantee that only $L-1$ pair of elements have to be compared during the diversity product (sum) computation.

Definition 6.11. Two unitary matrices $A, B \in G$ are said to be *equivalent* (denoted by $A \sim B$) if there is unitary matrix $U \in G$ such that $A = UBU^{-1}$ or $A = UB^{-1}U^{-1}$. $[A]$ will denote all the matrices that are equivalent to A. For a constellation $\mathcal{V} \subset G$, we say $\mathcal{V} = \{\Psi_1, \Psi_2, \cdots, \Psi_L\}$ has a *weak group structure* if for any two distinct elements Ψ_i, Ψ_j the product $\Psi_i^{-1}\Psi_j$ is equivalent to some Ψ_k.

The reader verifies that we indeed defined an equivalence relation. Note also that $\mathcal{V}$ has a group structure as long as $\Psi_i^{-1}\Psi_j$ is always another element of $\mathcal{V}$ and this explains our wording.

Lemma 6.12. *Let $\mathcal{V} = \{\Psi_0 = I, \Psi_1, \Psi_2, \cdots, \Psi_{L-1}\}$ be a constellation with a weak group structure. In order to compute the diversity product (sum) it is enough to do $L-1$ distance computations.*

Proof.

$$|\det(\Psi_i - \Psi_j)| = |\det(I - \Psi_i^{-1}\Psi_j)| = |\det(I - B)|,$$

where $B \in \mathcal{V}$ is an element in $\mathcal{V}$ equivalent to $\Psi_i^{-1}\Psi_j$. This shows the result for the diversity product. If one is concerned with the diversity sum, then the same argument still holds if the absolute value of the determinant $|\ \det(\cdot)\ |$ is replaced by the Frobenius norm $\| \ . \ \|_F$. □

Based on this lemma we are interested in finite constellations inside G whose elements have a weak group structure and are all normal. The following theorem provides a complete characterization of all these constellations:

Theorem 6.13. *Let $\mathcal{V} \subset G$ be a finite positive normal constellation (including identity element) with $L \geq 3$ elements. If $\mathcal{V}$ has a weak group structure then $\mathcal{V}$ takes one of the following forms:*

- $\{I, A, A^2, \cdots, A^{L-1}\}$
- $\{I, AB, A^2B^2, \cdots, A^{L-1}B^{L-1}\}$

where $A = g_i^{p_i}$, $B = g_j^{p_j}$ for some $i \neq j$.

The proof of Theorem 6.13 is rather involved. In order to make it more understandable we will divide it in several definitions and lemmas.

Definition 6.14. For any element $\Psi \in G$, we define the length of $\Psi = \prod_{i=1}^{k} {a_i}^{p_i}$ to be

$$\text{length}\,(\Psi) = \sum_{i=1}^{k} p_i.$$

It is routine to check that the definition is well-defined and does not depend on the representation of the element. For the identity element one will have length $(I) = 0$. One immediate consequence from this definition is that if $A \sim B$, one will have $|\text{length}\,(A)| = |\text{length}\,(B)|$. The following lemma claims that any freely generated positive weak group constellation "approximately" takes cyclic form.

Lemma 6.15. *Let $\mathcal{V} = \{\Psi_0 = I, \Psi_1, \Psi_2, \cdots, \Psi_{L-1}\} \subset G$ be a positive constellation of the freely generated set $G \subset U(M)$. Suppose* length $(\Psi_i) \leq$ length (Ψ_j) *for $i < j$. If $\mathcal{V}$ is a weak group constellation, then*

$$\Psi_i \in [\Psi_1]^i,$$

where $[\Psi_1]^i = \{a_1 a_2 \cdots a_i | a_1, a_2, \cdots, a_i \in [\Psi_1]\}$.

Proof. We first show that length $(\Psi_i) <$ length (Ψ_j) for $i < j$: Indeed, if length $(\Psi_i) =$ length (Ψ_j), then length $(\Psi_i^{-1}\Psi_j) =$ length $(\Psi_j) -$ length $(\Psi_i) = 0$. That means $\Psi_i^{-1}\Psi_j \sim I$, equivalently one will have $\Psi_i^{-1}\Psi_j = I$, *i.e.* $\Psi_i = \Psi_j$. That contradict the fact that Ψ_i and Ψ_j are distinct.

Consider $\Psi_1^{-1}\Psi_2$. Since $0 <$ length $(\Psi_1^{-1}\Psi_2) =$ length $(\Psi_2) -$ length $(\Psi_1) <$ length (Ψ_2), therefore $\Psi_1^{-1}\Psi_2 = \bar{\Psi}_1$ where $\bar{\Psi}_1 \sim \Psi_1$. So $\Psi_2 = \Psi_1\bar{\Psi}_1 \in [\Psi_1]^2$. Proceed by induction, one can show $\Psi_k^{-1}\Psi_{k+1} = \bar{\Psi}_2$ where $\bar{\Psi}_2 \sim \Psi_1$. So $\Psi_{k+1} = \Psi_k\bar{\Psi}_2 \in [\Psi_1]^{k+1}$ by induction. □

Remark 6.16. An immediate observation is that

$$\text{length}\,(\Psi_i) = i * \text{length}\,(\Psi_1).$$

Take two positive normal elements in G with their reduced forms:

$$\Psi_1 = a_1^{p_1}a_2^{p_2}\cdots a_m^{p_m}, \qquad \Psi_2 = b_1^{q_1}b_2^{q_2}\cdots b_n^{q_n}.$$

We define the shift operator S_k on the reduced form of a positive normal element Ψ by induction: $S_1(\Psi) = S_1(a_1^{p_1}a_2^{p_2}\cdots a_m^{p_m}) = a_2^{p_2}\cdots a_m^{p_m}a_1^{p_1}$ and $S_{k+1} = S_k \circ S_1$. We assume that $S_0(\Psi) = \Psi$, then apparently for a fixed element Ψ shift operator is periodic. We have the following lemma.

Lemma 6.17. *$\Psi_1 \sim \Psi_2$ if and only if $\Psi_1 = S_k(\Psi_2)$ for some k.*

Proof. The sufficiency part of this lemma is straightforward. So we have to prove the necessity part. Since $\Psi_1 \sim \Psi_2$, according to the definition of equivalence there exists c such that $c\Psi_1c^{-1} = \Psi_2$ or $c\Psi_1c^{-1} = \Psi_2^{-1}$. However since length $(c\Psi_1c^{-1}) =$ length $(\Psi_2) > 0$ and length $(\Psi_2^{-1}) < 0$, the second case will not happen. The only possibility is $c\Psi_1c^{-1} = \Psi_2$. We assume that c is generated by only one generator and further assume $c = c_1^{l_1}$ with $l_1 > 0$, then we will have

$$c_1^{l_1}a_1^{p_1}a_2^{p_2}\cdots a_m^{p_m}c_1^{-l_1} = b_1^{q_1}b_2^{q_2}\cdots b_n^{q_n}.$$

So $c_1 = a_m$ and $l_1 \leq p_m$ follows, otherwise the left hand side of the equation above will have negative power, while the right hand side only has positive power. This

will contradict the uniqueness of the representation of the same element. In fact $l_1 = p_m$, since otherwise $\Psi_2 = c_1^{l_1} a_1^{p_1} a_2^{p_2} \cdots c_1^{p_m - l_1}$. This will contradict the fact that Ψ_2 is a normal element. So with

$$a_m^{p_m} a_1^{p_1} \cdots a_{m-1}^{p_{m-1}} = b_1^{q_1} b_2^{q_2} \cdots b_n^{q_n},$$

one can check $m = n$ and $\Psi_2 = S_{m-1}(\Psi_1)$.

Proceed by induction, suppose c has the reduced form $c = c_1^{l_1} c_2^{l_2} \cdots c_{k+1}^{l_{k+1}}$, then the following equation follows:

$$c_1^{l_1} c_2^{l_2} \cdots c_{k+1}^{l_{k+1}} a_1^{p_1} a_2^{p_2} \cdots a_m^{p_m} c_{k+1}^{-l_{k+1}} \cdots c_2^{-l_2} c_1^{-l_1} = b_1^{q_1} b_2^{q_2} \cdots b_n^{q_n}.$$

Without loss of generality, we assume $l_{k+1} > 0$ and apply the same argument as in the one generator case. One proves $a_m = c_{k+1}$ and $l_{k+1} = p_m$. Therefore we reach the following equation:

$$c_1^{l_1} c_2^{l_2} \cdots c_k^{l_k} S_{m-1}(\Psi_1) c_k^{-l_k} \cdots c_2^{-l_2} c_1^{-l_1} = b_1^{q_1} b_2^{q_2} \cdots b_n^{q_n}.$$

By induction, $\Psi_2 = S_{k_1} \circ S_{m-1}(\Psi_1) = S_{k_1+m-1}(\Psi_1)$ for some k_1. □

Proof of Theorem 6.13. Pick any two distinct elements $\Psi_i, \Psi_j \in \mathcal{V}$ having

$$\text{length}\,(\Psi_i) < \text{length}\,(\Psi_j).$$

We claim that if $\Psi_i = a_1 a_2 \cdots a_m$, then either there exists $1 \leq k \leq m-1$ such that $\Psi_j = a_1 a_2 \cdots a_k b_1 b_2 \cdots b_l a_{k+1} \cdots a_m$, or $\Psi_j = b_1 b_2 \cdots b_l a_1 a_2 \cdots a_m$ or $\Psi_j = a_1 a_2 \cdots a_m b_1 b_2 \cdots b_l$ for some $l > 0$.

Suppose that the claim is not true, then for $\Psi_j = c_1 c_2, \cdots c_p$, there exist k_1, k_2 such that $0 \leq k_1 \leq m$, $1 \leq k_2 \leq m+1$ and $k_1 < k_2 - 1$ and Ψ_j will take the following form:

$$\Psi_j = a_1 a_2 \cdots a_{k_1} b_1 b_2 \cdots b_l a_{k_2} \cdots a_m,$$

where $b_1 \neq a_{k_1+1}$ and $b_l \neq a_{k_2-1}$. (For the special case $k_1 = 0$, we assume $c_1 \neq a_1$. For the special case $k_2 = m+1$, we assume $c_p \neq a_m$.) Then $\Psi_i^{-1} \Psi_j$ would be equivalent to $a_{k_2-1}^{-1} \cdots a_{k_1+1}^{-1} b_1 b_2 \cdots b_l$, which in any case will not be equivalent to any positive element $\Psi_k = d_1 d_2 \cdots d_q$ or I. That contradicts the fact that $\mathcal{V}$ is equipped with a weak group structure.

As explained above we can further assume that

$$\text{length}\,(I) < \text{length}\,(\Psi_1) < \cdots < \text{length}\,(\Psi_{L-1}).$$

If Ψ_1 is generated by only one generator, *i.e.* $\Psi_1 = g_i^{p_i}$ for some i. Since Ψ_2 is a normal element, according to the claim, either $\Psi_2 = \Psi_1\tilde{\Psi}_2$ or $\Psi_2 = \tilde{\Psi}_2\Psi_1$ for some $\tilde{\Psi}_2$. In either case $\tilde{\Psi}_2$ will be equivalent to Ψ_1, while Lemma 6.17 will guarantee $\tilde{\Psi}_2 = \Phi_1$. Therefore we will have $\Psi_2 = g_i^{2p_i}$. Proceed by induction, it can be checked that $\Psi_l = g_i^{lp_i}$ for every l. So the constellation will take the first form in the theorem.

If Φ_1 is generated by two generators, *i.e.* $\Psi_1 = g_i^{p_i} g_j^{p_j}$ for some i, j. According to the claim, we will have $\Psi_2 = \Psi_1\tilde{\Psi}_2$ or $\Psi_2 = \tilde{\Psi}_2\Psi_1$ or $\Psi_2 = g_i^{p_i}\tilde{\Psi}_2 g_j^{p_j}$. Because $\tilde{\Psi}_2$ is equivalent to Ψ_1, $\tilde{\Psi}_2$ is a shifted version of Ψ_1. Exhausting all the possibilities, the first two cases would make Ψ_2 a non-normal element, so the only possibility is the third case. Consider two shifted version of Ψ_1: $S_0(\Psi_1) = g_i^{p_i} g_j^{p_j}$ and $S_1(\Psi_1) = g_j^{p_j} g_i^{p_i}$. Only $S_0(\Psi_1)$ will satisfy the condition that Ψ_2 is a normal element. So the analysis above shows that

$$\Psi_2 = g_i^{p_i}\Psi_1 g_j^{p_j} = g_i^{2p_i} g_j^{2p_j}.$$

By induction it can be shown that

$$\Psi_{k+1} = g_i^{p_i}\Psi_k g_j^{p_j} = g_i^{(k+1)p_i} g_j^{(k+1)p_j}.$$

So in this case, the constellation will take the second form in the theorem.

However the constellation does not exist if Ψ_1 is generated by more than 3 elements. Indeed suppose with the reduced form $\Psi_1 = a_1^{p_1} a_2^{p_2} \cdots a_m^{p_m}$ with $m \geq 3$, then Ψ_2 will take one of the following form:

$$\tilde{\Psi}_2 a_1^{p_1} a_2^{p_2} \cdots a_m^{p_m}, a_1^{p_1}\tilde{\Psi}_2 a_2^{p_2} \cdots a_m^{p_m}, \cdots, a_1^{p_1} a_2^{p_2} \cdots a_m^{p_m}\tilde{\Psi}_2$$

with $\tilde{\Psi}_2$ being a shifted version of Ψ_1. But Ψ_2 would not be a normal element for any of the above form, so there does not exist weak group constellation for this case.

□

A weak group constellation is very group like, while it is not exactly a group. It does keep the advantage of a group constellation: for example, for any weak group constellation $\mathcal{V}$ taking the second form in the theorem, only $L-1$ computations

$|\det(I - A^k B^k)|$ for $k = 1, 2, \cdots, L-1$ are needed to calculate the diversity product. It also overcome the disadvantage of group codes: one can freely choose the generators, while in group structures, the generators have to satisfy certain relations to be a group. Last but not least it turns out that the restriction to code elements in normal form is very advantageous during sphere decoding. In the next section we will mainly use the second weak group structure as described in Theorem 6.13. Before we describe these search procedures we would like to illustrate some alternative methods.

It is possible to increase the number of generators to obtain new structures. For instance, $\mathcal{V} = \{A^k B^l C^m | A, B, C \in U(M), k = 0, \cdots, p, l = 0, \cdots, q, m = 0, \cdots, r\}$.

For a unitary constellation $\mathcal{V} = \{\Phi_i | i = 1, \cdots, L\}$, we call $\mathcal{V}_s = \{U\Phi_i V | i = 1, \cdots, L\}$ shifted version of $\mathcal{V}$. It will be straightforward to prove that $\mathcal{V}_s$ has the same complexity as $\mathcal{V}$ when one calculates the diversity. $\{A^k C B^k | A, B, C \in U(M), k = 0, \cdots, L-1\}$ is a shifted copy of the second weak group structure in Theorem 6.13. To see this, note that $A^k C B^k = A^k C B^k C^{-1} C = A^k (C B C^{-1})^k C$. It can checked that $A^k B^{L+1-k} = A^k (B^{-1})^k B^{L+1}$, therefore $\{A^k B^{L+1-k} | A, B \in U(M), k = 1, \cdots, L\}$ is also a shifted version of the second form weak group structure.

Also we can consider the "combination" or the "product" of two structures. For example, $\{I, A, AB, ABA, ABAB, ABABA, \cdots\}$ is the union of $\{(AB)^k | k = 0, \cdots\}$ and its shifted version $\{(AB)^k A | k = 0, \cdots\}$. Another example is the product case: let $\mathcal{V}_1 = \{I, C, C^2, C^3, \cdots\}$ and $\mathcal{V}_2 = \{I, A, AB, ABA, \cdots\}$ and consider the Cartesian product constellation

$$\mathcal{V} = \mathcal{V}_1 \times \mathcal{V}_2 = \{AB | A \in \mathcal{V}_1, B \in \mathcal{V}_2\}.$$

One may wonder how restrictive the proposed structures are. We all know a compact Lie group can be generated by any open neighborhood of any element in the Lie group. So with the above structure, even if one chooses the generators locally,

the elements in the constellation could be spreading out on the whole manifold. Somehow this indicates that the proposed structure will not be too restrictive.

6.4 Geometrical Design of Unitary Constellations with Good Diversity

For low dimensional constellations, one may further specify the generators in the proposed structure. Observe that for the second form weak group constellation, one can always assume A is diagonal. In the sequel, we further assume that B is real orthogonal, *i.e.* based on the weak group structure we consider the following 2 dimensional constellation:

$$\mathcal{V} = \{A^k B^k | A = \begin{pmatrix} e^{ix} & 0 \\ 0 & e^{iy} \end{pmatrix}, B = \begin{pmatrix} \cos z & \sin z \\ -\sin z & \cos z \end{pmatrix}, k = 0, 1, \cdots, L-1\}. \tag{6.1}$$

There are several ways to design constellations with good diversity from this specific structure. A natural idea is to do Brute Force search using fine step size. Another approach is to design the constellation with the help of geometrical intuition. Note that a 2×2 complex matrix can be viewed as a vector in $\mathbb{C}^4$. In this context A and B can be viewed as "rotation" transforms (induced by regular matrix multiplication) acting on $\mathbb{C}^4$. A constellation of form (6.1) can be viewed as a set of rotated vectors under the transforms $A^k B^k$, $k = 0, 1, \cdots, L-1$. Intuition says that good constellations can be found if the rotation angle is symmetrical. Based on the idea above we assume that x, y, z to be the multiples of $2\pi/L$, we found a lot of good codes resulted from this geometrical symmetry (see tables in Section 6.5).

2 dimensional constellation design has been studied in [32]. In this paper Liang proposed very interesting parametric codes and many codes with excellent diversity are found. The codes shown in [32] can be achieved by our design as well. In fact, most of Liang's codes belong to a special form of our parameterization (6.1). To

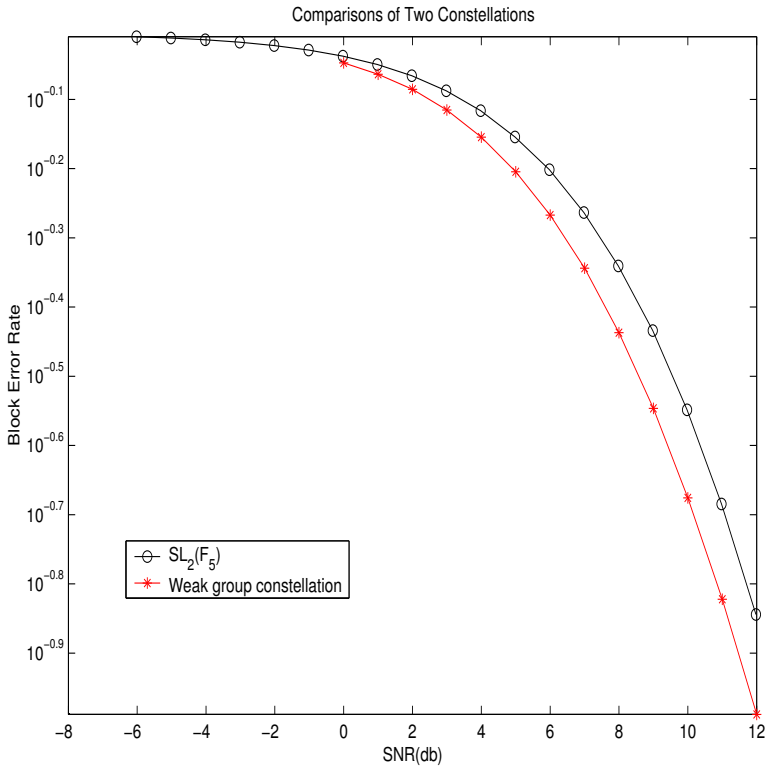


Figure 6.3. 2 dimensional weak group constellations and group constellation

our best knowledge, most of our codes shown on the web site [18] are the best codes ever found or never found before.

Example 6.18. A very interesting code with 120 elements is found using this approach:

$$\mathcal{V} = \{A^k B^k | A = \begin{pmatrix} e^{\pi/30i} & 0 \\ 0 & e^{11\pi/30i} \end{pmatrix}, B = \begin{pmatrix} \cos\pi/4 & \sin\pi/4 \\ -\sin\pi/4 & \cos\pi/4 \end{pmatrix}, k = 0, 1, \cdots, 119\}.$$

It can be checked that $\prod \mathcal{V} = \sum \mathcal{V} = \frac{1}{2}\sqrt{\frac{(3-\sqrt{5})}{2}}$, *i.e.* the diversity product and the diversity sum are identical to the ones of the $SL_2(\mathbb{F}_5)$-constellation. We simulated the performance of this code and compared it with the performance of the $SL_2(\mathbb{F}_5)$-constellation. To our big surprise our new code performed considerably better than the $SL_2(\mathbb{F}_5)$-constellation. The constellation $\mathcal{V}$ with sphere decoding

Table 6.2. Distance spectrum of the weak group constellation

Weak group constellation
DP distance spectrum

distance	distribution
0.3090	360
0.3136	480
0.3895	480
0.3931	1440
0.4402	240
0.5000	120
0.5878	120
0.6360	1440
0.6787	480
0.7071	600
0.8090	360
0.8430	480
0.8660	120
0.8979	240
0.9511	120
1	60

Weak group constellation
DS distance spectrum

distance	distribution
0.3090	120
0.4402	240
0.5000	120
0.5023	480
0.5457	240
0.5878	120
0.6367	480
0.6502	240
0.7071	3000
0.7598	240
0.7711	240
0.8090	120
0.8380	240
0.8647	480
0.8660	120
0.8979	240
0.9511	120
1	60

outperformed the $SL_2(\mathbb{F}_5)$-constellation by about 1dB up to about 20dB (see Figure 6.4). As the SNR increases, the two curves are getting closer though.

In order to understand the difference in the performance of the two seemingly similar constellations we investigated the diversity product (DP) and diversity sum (DS) *distance spectrum* for each of them. As we explained before, for a unitary constellation with L elements, $L(L-1)/2$ distance calculations may produce distances with multiplicities. For example consider $\mathcal{V}$ as above, 360 out of 7140 pairs of elements have distance 0.3090 (see DP distance spectrum in Table 6.2). So one can explain the behavior difference of the two codes using their distance spectrum. Table 6.2 shows that the DP and DS distance spectrum of our weak group constellation.

One can check that the DP distance spectrum of the $SL_2(\mathbb{F}_5)$-constellation is

Table 6.3. Distance spectrum of the group constellation

$SL_2(\mathbb{F}_5)$-constellation
DP (DS) distance spectrum

distance	distribution
0.3090	720
0.5000	1200
0.5878	720
0.7071	1800
0.8090	720
0.8660	1200
0.9511	720
1	60

identical to the DS distance spectrum. The following Table 6.3 shows that the DS distance spectrum for the $SL_2(\mathbb{F}_5)$-constellation has denser small distance distribution compared to DS spectrum of our constellation and this explains the considerable worse performance of this constellation in our simulations.

Although we have concentrated so far on the design of 2-dimensional constellations, there is actually no restriction with our approach. The similar "rotation" idea can be applied to other low dimensional constellation design. For instance, we can make further specifications to 3 dimensional weak group constellations:

$$\mathcal{V} = \{A^k B^k | A = \begin{pmatrix} \cos x & \sin x & 0 \\ -\sin x & \cos x & 0 \\ 0 & 0 & e^{iy} \end{pmatrix}, B = \begin{pmatrix} e^{iz} & 0 & 0 \\ 0 & \cos w & \sin w \\ 0 & -\sin w & \cos w \end{pmatrix}, k = 0, \cdots, L-1\}.$$

where x, y, z, w are assumed to take the multiple of $2\pi/L$. Apparently algebraic design based on geometrical symmetry can be applied to any other structure as well. For instance consider the following specified structures:

$$\mathcal{V} = \{A^k B^l | A = \begin{pmatrix} e^{ix} & 0 \\ 0 & e^{iy} \end{pmatrix}, B = \begin{pmatrix} \cos z & \sin z \\ -\sin z & \cos z \end{pmatrix}, k = 0, \cdots, p-1, l = 0, \cdots, q-1\}.$$

where we can take x, y to be multiple of $2\pi/p$ and z to be multiple of $2\pi/q$. We refer to [18] for the designed low dimensional constellations from these approaches.

6.5 Numerical Design of Unitary Constellations with Good Diversity

In order to numerically design constellations, it will be necessary to have a good parameterization for the set of unitary constellations having size L, operating with M transmit antennas. In this section we show how one can use the theory of complex Stiefel manifolds and the classical Cayley transform to obtain such a parameterization.

6.5.1 The Complex Stiefel Manifold

Definition 6.19. The subset of $T \times M$ complex matrices

$$\mathcal{S}_{T,M} := \left\{ \Phi \in \mathbb{C}^{T \times M} \mid \Phi^* \Phi = I_M \right\}$$

is called the *complex Stiefel manifold*.

From an abstract point of view a constellation $\mathcal{V} := \{\Phi_1, \ldots, \Phi_L\}$ having size L, block length T and operating with M antennas can be viewed as a point in the complex manifold

$$\mathcal{M} := (\mathcal{S}_{T,M})^L = \underbrace{\mathcal{S}_{T,M} \times \cdots \times \mathcal{S}_{T,M}}_{L \text{ copies}}.$$

The search for good constellations $\mathcal{V}$ requires hence the search for points in $\mathcal{M}$ whose diversity is excellent in some interval $[\rho_1, \rho_2]$.

Stiefel manifolds have been intensely studied in the mathematics literature since their introduction by Eduard Stiefel some 50 years ago. A classical paper on complex Stiefel manifolds is [4], a paper with a point of view toward numerical algorithms is [12]. The major properties are summarized by the following theorem:

Theorem 6.20. *$\mathcal{S}_{T,M}$ is a smooth, real and compact sub-manifold of $\mathbb{C}^{MT} = \mathbb{R}^{2MT}$ of real dimension $2TM - M^2$.*

Some of the stated properties will follow from our further development. The following two examples give some special cases.

Example 6.21.

$$\mathcal{S}_{T,1} = \left\{ x \in \mathbb{C}^T \mid ||x|| = \sqrt{\sum_{i=1}^{M} x_i \bar{x}_i} = 1 \right\} \subset \mathbb{R}^{2T}$$

is isomorphic to the $2T-1$ dimensional unit sphere S^{2T-1}.

Example 6.22. When $T = M$ then $\mathcal{S}_{T,M} = U(M)$, the group of $M \times M$ unitary matrices. It is well known that the Lie algebra of $U(M)$, *i.e.* the tangent space at the identity element, consists of all $M \times M$ skew-Hermitian matrices. This linear vector space has real dimension M^2, in particular the dimension of $U(M)$ is M^2 as well.

A direct consequence of Theorem 6.20 is:

Corollary 6.23. *The manifold $\mathcal{M}$ which parameterizes the set of all constellations $\mathcal{V}$ having size L, block length T and operating with M antennas forms a real compact manifold of dimension $2LTM - LM^2$.*

As this corollary makes it clear a full search over the total parameter space is only possible for very moderate sizes of M, L, T. It is also required to have a good parameterization of the complex Stiefel manifold $\mathcal{S}_{T,M}$ and we will go after this task next.

The unitary group is closely related to the complex Stiefel manifold and the problem of parameterization ultimately boils down to the parameterization of unitary matrices. For this assume that Φ is a $T \times M$ matrix representing an element of the complex Stiefel manifold $\mathcal{S}_{T,M}$. Using Gramm-Schmidt one constructs a $T \times (T-M)$ matrix V such that the $T \times T$ matrix $[\Phi \mid V]$ is unitary. Define two $T \times T$ unitary matrices $[\Phi_1 \mid V_1]$ and $[\Phi_2 \mid V_2]$ to be equivalent whenever $\Phi_1 = \Phi_2$. A direct calculation shows that two matrices are equivalent if and only if there is $(T-M) \times (T-M)$

matrix Q such that:

$$[\Phi_2 \mid V_2] = [\Phi_1 \mid V_1] \begin{pmatrix} I & 0 \\ 0 & Q \end{pmatrix}. \tag{6.2}$$

Identifying the set of matrices Q appearing in (6.2) with the unitary group $U(T-M)$ we get the result:

Lemma 6.24. *The complex Stiefel manifold $\mathcal{S}_{T,M}$ is isomorphic to the quotient group*

$$U(T)/U(T-M).$$

This lemma verifies the dimension formula for $\mathcal{S}_{T,M}$ stated in Theorem 6.20:

$$\dim \mathcal{S}_{T,M} = \dim U(T) - \dim U(T-M) = T^2 - (T-M)^2 = 2TM - M^2.$$

This section makes it clear that a good parameterization of the set of constellations $\mathcal{V}$ requires a good parameterization of the manifold $\mathcal{M}$ and this in turn requires a good parameterization of the unitary group $U(M)$.

Once one has a nice parameterization of the unitary group $U(M)$ then Lemma 6.24 provides a way to parameterize the Stiefel manifold $\mathcal{S}_{T,M}$ as well. Parameterizing $U(T)$ modulo $U(T-M)$ is however an "over parameterization". Edelman, Arias and Smith [12] explained a way on how to describe a local neighborhood of a (real) Stiefel manifold $\mathcal{S}_{T,M}$. The method can equally well be applied in the complex case. We do not pursue this parameterization in this chapter and leave this for future work.

In the remainder of this chapter we will concentrate on constellations having the special form (3.7). From a numerical point of view we require for this a good parameterization of the unitary group and the next subsection provides an elegant way to do this.

6.5.2 Cayley Transformation

There are several ways to represent a unitary matrix in a very explicit way. One elegant way makes use of the classical Cayley transformation. In order that this chapter is self contained we provide a short summary. More details are given in [37, Section 22] and [24].

Definition 6.25. For a complex $M \times M$ matrix Y which has no eigenvalues at -1, the Cayley transform of Y is defined to be

$$Y^c = (I + Y)^{-1}(I - Y),$$

where I is the $M \times M$ identity matrix.

Note that $(I + Y)$ is nonsingular whenever Y has no eigenvalue at -1. One immediately verifies that $(Y^c)^c = Y$. This is in analogy to the fact that the linear fractional transformation $f(z) = \frac{1-z}{1+z}$ has the property that $f(f(z)) = z$. Recall that a matrix M is skew-Hermitian whenever $A^* = -A$. The set of $M \times M$ skew-Hermitian matrices forms a linear subspace of $\mathbb{C}^{M\times M} \cong \mathbb{R}^{2M^2}$ having real dimension M^2. This is the Lie algebra of the unitary group $U(M)$. The main property of the Cayley transformation is summarized in the following theorem. (See *e.g.* [24, 37]).

Theorem 6.26. *When A is a skew-Hermitian matrix then $(I + A)$ is nonsingular and the Cayley transform $V := A^c$ is a unitary matrix. When V is a unitary matrix which has no eigenvalues at -1 then the Cayley transform V^c is skew-Hermitian.*

This theorem allows one to parameterize the open set of $U(M)$ consisting of all unitary matrices whose eigenvalues do not include -1 through the linear vector space of skew-Hermitian matrices. The Cayley transformation is very important for the numerical design of constellations because it makes the local topology of $U(M)$ clear. One can see that most optimization methods require us to consider the neighborhood of one element in $U(M)$.

6.5.3 Simulated Annealing (SA) Algorithm

In our numerical experiments we have considered several methods. Because there are a large number of target functions the best known optimization algorithms, such as the steepest descent method, Newton's method and the conjugate gradient method on a general Riemannian manifold, are difficult to implement. Surprisingly the *Simulated Annealing Algorithm* turned out to be very practical for this problem.

Simulated Annealing (SA) is a method which mimics the process of melted metal cooling off. In the annealing process of the melted metal, first the metal is heated to melt, then the temperature decreases gradually. The metal will get to a minimized energy state if the temperature is lowering slow enough. For more details about this algorithm, we refer to [1, 46, 36].

In fact, we would rather call it a general method instead of a concrete algorithm. Generally speaking, for a given optimization problem we always take an initial solution in some certain way, then consider a second solution in the "neighborhood" of this solution. We will accept the solution according to some predefined criterion which might involve a probability threshold.

Combining with good algebraic structure and Cayley transform, which is a good representation of any dimensional unitary matrix, one can see that numerical method can be applied to any dimensional and any size constellation design. Our implementation of the algorithm can be summarized in the following way, one can find simple sample program on our web site [18].

1. Choose a proposed algebraic structure for the constellation.

2. Generate initial generators of the whole constellation. One can either take an existing constellation as the start point or just take the initial point randomly.

3. Generate randomly a new constellation using Cayley transform in the neigh-

borhood of the old constellation where the selection is done using a Gaussian distribution with decreasing variances as the algorithm progresses.

4. Calculate the diversity function (product, sum) of the newly constructed constellation.

5. If the new constellation has better diversity function (product, sum), then accept the new constellation. If not, reject the new constellation and keep the old constellation (or accept it according to Metropolis's criterion [35]).

6. Check the stopping criterion, if satisfied, then stop, otherwise go to 2 and continue the iteration.

Example 6.27. As we mentioned before, one can either choose an existing constellation as the starting point for our numerical method or just take the initial point randomly. In the sequel, we use the group constellation $G_{21,4}$ in [41]:

$$\mathcal{V}_1 = \{A^k B^l | A = \begin{pmatrix} \eta & 0 & 0 \\ 0 & \eta^4 & 0 \\ 0 & 0 & \eta^{16} \end{pmatrix}, B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \eta^7 & 0 & 0 \end{pmatrix}, k = 0, 1, \cdots, 20, l = 0, 1, 2\}$$

One can verify that

$$\prod \mathcal{V}_1 = 0.3851.$$

It seems that $G_{21,4}$ is already a very good constellation, our algorithm only improves a little (see $\mathcal{V}_2$ below). However one can check for most of the cases, the algorithm will improve much compared to the original group constellation.

$$\mathcal{V}_2 = \{A^k B^l | k = 0, 1, \cdots, 20, l = 0, 1, 2\},$$

where

$$A = \begin{pmatrix} 0.9415 + 0.3155 * i & 0.0573 - 0.0222 * i & 0.0496 + 0.0882 * i \\ 0.0160 - 0.0555 * i & 0.4005 + 0.9136 * i & 0.0326 - 0.0212 * i \\ 0.0579 + 0.0855 * i & -0.0312 - 0.0099 * i & 0.1384 - 0.9844 * i \end{pmatrix},$$

$$B = \begin{pmatrix} 0.0175 + 0.0095 * i & 0.9997 + 0.0111 * i & 0.0079 + 0.0042 * i \\ 0.0086 + 0.0100 * i & -0.0082 + 0.0040 * i & 0.9999 + 0.0036 * i \\ -0.4836 + 0.8750 * i & 0.0004 - 0.0198 * i & -0.0045 - 0.0126 * i \end{pmatrix}.$$

One verifies that

$$\prod \mathcal{V}_2 = 0.3874.$$

Example 6.28. Different industrial applications require different level of reliability of the communication channels. One may want to optimize the constellation at certain block error rate (BER) or signal to noise ratio (SNR). It can also be shown theoretically that numerical methods together with the proposed structure works in the same way if one wants to optimize the diversity function at a certain SNR. This is essentially the case because for a complex matrix A and unitary matrices U, V one has that

$$\delta_m(UAV) = \delta_m(A), \tag{6.3}$$

for $m = 1, 2, \cdots, M$. With the constellation structures as above we are able to reduce the dimension of the parameter space and at the same time we have a considerable reduction in the number of targets to be checked. Intuitively, algebraically designing codes for this purpose seems to be impossible.

The following graph shows the comparison of three constellations with different dimensions with 2 receiver antennas. The first one is a 2 dimensional constellation with 3 elements ($R = 0.7925$) and optimal diversity product 0.8660 and optimal diversity sum 0.8660. The second constellation is a 3 dimensional constellation which has 5 elements ($R = 0.7740$) with diversity product 0.7183 and diversity sum 0.7454. The third constellation is a 4 dimensional one consisting of 9 elements ($R = 0.7925$) with diversity product 0.5904 and diversity sum 0.6403. Here based on the structure $A^k B^k$ we used Simulated Annealing to optimize the diversity function at 6dB to acquire the last two constellations.

One can see that around 5dB, the second constellation surpasses the first one and is getting better and better as the SNR becomes larger. This can be easily understood since the diversity function of the first constellation is approximately dominated by $1/\rho^4$ at high SNR, while the diversity function of the second constellation is dominated by $1/\rho^6$. The same explanation can be applied to the third constellation's performance. One can even foresee that higher dimensional constellations will perform even better and the BER curve will be sharper than the lower dimensional ones. It is believable that higher dimensional constellations will achieve much more diversity gain compared to lower dimensional ones.

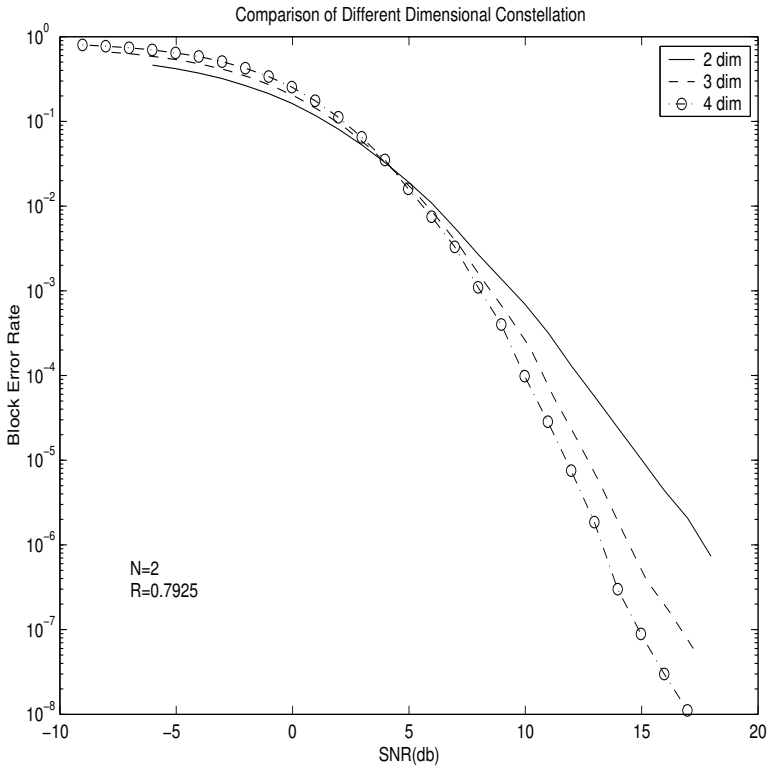


Figure 6.4. 2 dimensional weak group constellations and group constellation

Surprisingly SA works very well when it is applied to an algebraic structure with symmetry. Like all the other numerical methods, one has to suffer the loss of performance due to the increasing complexity as the size and dimension go up and due to the limited computational resources. However without any doubts, the numerical approach is very flexible and can be used for any dimensional and any size constellation, producing very good diversity. So a lot of good-performing unitary constellations are found this way, which were never found by any algebraic method. At the end of this section we will show some 2 dimensional constellations we found using various methods based on the proposed structure. We skip, however, our numerical results on the higher dimensional unitary constellation design, since one can check them on the web site [18].

One very interesting fact is the numerical results for diversity sum from 3 dimensional structured constellation are even better than the corresponding upper bound for 2 dimensional constellations. Somehow it will not be too surprising if one notices that from $U(2)$ to $U(3)$, we have 5 more dimensions to maneuver.

In Chapter 4, packing problems on compact Lie groups are analyzed and the upper bound for the diversity sum and the diversity product are derived. In Figure 6.5 one can see the limiting behavior of 2 dimensional structured constellations compared to the upper bound. One can check [18] for the comparisons for other dimensions.

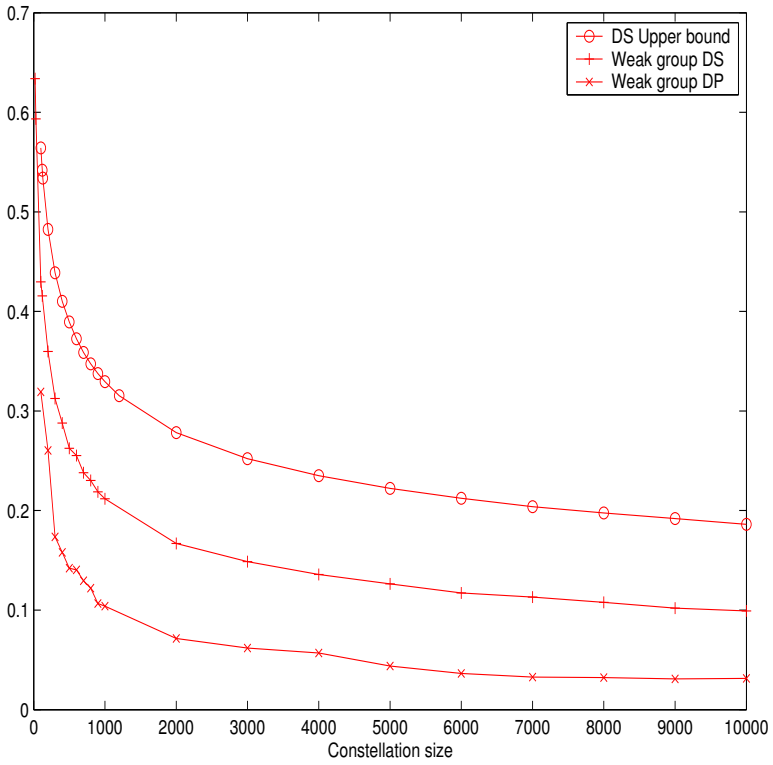


Figure 6.5. 2 dimensional weak group constellations and upper bound

6.5.4 Constellations with Extremely Large Diversity

In this subsection we list the best 2-dimensional constellations we found with the techniques described in Sections 6.4 and 6.5. The tabulated constellations have some of the best diversity sums and diversity products published so far. All the constellations searched by simulated annealing (SA) were based on the $A^k B^k$ structure. For the constellations with L elements and parameters x, y, z being multiples of $2\pi/L$, they are found by geometrical methods using the parameterization (6.1). For the constellations with L elements and parameters x, y, z being decimals, they are found by Brute Force with step size 0.1000 based on the same parameterization (6.1). In Table 6.4, 6.5 and 6.6, we show the diversity product and sum of 2 dimensional constellation based on weak group structure:

Table 6.4. The diversity product of 2 dimensional constellations

Number of elements	Diversity Product	Codes and Comments
2	1	$x = \pi, y = \pi, z = 0$ (optimal)
3	$\sqrt{3}/2$	$x = 2\pi/3, y = 2\pi/3, z = 0$ (optimal)
4	0.7831	$x = 0.6000, y = 6.0000, z = 4.4000$
5	$\sqrt{5/8}$	$x = 2\pi/5, y = 8\pi/5, z = 4\pi/5$ (optimal)
8	0.7071	$x = 2.3562, y = 3.9270, z = 4.7124$
9	0.6524	SA searched code
10	0.6124	$x = 2\pi/5, y = 8\pi/5, z = \pi/5$
16	$\sqrt[4]{2}/2$	$x = \pi/4, y = 5\pi/4, z = 13\pi/8$
17	0.5255	SA searched code
18	0.5207	SA searched code
19	0.5128	SA searched code
20	0.5011	$x = 1.6500, y = 3.7500, z = 4.0500$
24	0.5000	$x = \pi/12, y = 5\pi/12, z = \pi/2$
37	0.4461	$x = 2\pi/37, y = 6\pi/37, z = 12\pi/37$
39	0.3984	$x = 8\pi/39, y = 34\pi/39, z = 36\pi/39$
40	0.3931	$x = 3\pi/10, y = 11\pi/10, z = 3\pi/4$
55	0.3874	$x = 2\pi/55, y = 68\pi/55, z = 6\pi/11$
57	0.3764	$x = 2\pi/57, y = 40\pi/57, z = 48\pi/57$
75	0.3535	$x = 2\pi/75, y = 98\pi/75, z = 96\pi/75$
85	0.3497	$x = 26\pi/85, y = 94\pi/85, z = 18\pi/17$
91	0.3451	$x = 2\pi/91, y = 128\pi/91, z = 42\pi/91$
96	0.3192	$x = 7\pi/16, y = 29\pi/16, z = \pi/6$
105	0.3116	$x = 2\pi/105, y = 68\pi/105, z = 84\pi/105$
120	0.3090	$x = \pi/30, y = 11\pi/30, z = \pi/4$
135	0.2869	$x = 2\pi/135, y = 28\pi/135, z = 68\pi/135$
145	0.2841	$x = 2\pi/145, y = 64\pi/145, z = 76\pi/145$
165	0.2783	$x = 2\pi/33, y = 20\pi/33, z = 2\pi/5$
203	0.2603	$x = 2\pi/203, y = 290\pi/203, z = 70\pi/203$
225	0.2499	$x = 82\pi/225, y = 118\pi/225, z = 126\pi/225$
217	0.2511	$x = 2\pi/217, y = 250\pi/217, z = 168\pi/217$
225	0.2499	$x = 82\pi/225, y = 118\pi/225, z = 126\pi/225$
240	0.2239	$x = \pi/40, y = 9\pi/40, z = \pi/6$
273	0.2152	$x = 2\pi/273, y = 208\pi/273, z = 142\pi/273$
295	0.2237	$x = 14\pi/295, y = 104\pi/295, z = 22\pi/59$
297	0.1910	$x = 242\pi/297, y = 548\pi/297, z = 54\pi/297$
299	0.1858	$x = 8\pi/299, y = 220\pi/299, z = 18\pi/299$
300	0.1736	$x = \pi/150, y = 51\pi/150, z = 5\pi/6$

6.5.5 General Form Constellation Numerical Design

The connection between the complex Stiefel manifold and $U(M)$ (see the beginning of this section) makes clear that the techniques used above for square unitary constellations can be applied to design general form unitary constellations too. For simplicity we describe the idea with assumption $T = 2M$ and consider the following structure:

$$\{A^k B | A \in U(T), B = \begin{pmatrix} I_M \\ 0 \end{pmatrix}, k = 0, 1, \cdots, L-1\}.$$

One can check at most $2L-1$ distance calculations are needed to derive the diversity product (sum or function) with this algebraic structure. The numerical results for non-square constellations ($T = 5, M = 2$) can be found in Table 6.7. More results can be found in [18].

6.6 Fast Decoding of the Structured Constellations

The complexity of ML decoding for unitary space time constellations increases exponentially with the number of antennas or the transmission rate. This will preclude its practical use for high transmission rates or for large number of antennas. Basically our structured constellations can convert the ML decoding to lattice decoding naturally, consequently they admit fast decoding algorithms.

The principle of sphere decoding [13] is as follows: instead of doing an exhaustive search over all the lattice points, one can limit its search area to a sphere with given radius $\sqrt{C}$ centered at received point. One can check the complexity of this approach in [13] and in [25].

We will use the $A^k B^l$ structure to describe how one can apply sphere decoding algorithm for the demodulation based on our constellations. Suppose A has Schur decomposition $A = U\text{diag}\,(e^{i\alpha_1}, e^{i\alpha_2}, \cdots, e^{i\alpha_M})U^*$, similarly assume $B =$

Table 6.5. The diversity sum of 2 dimensional constellations

number of elements	Diversity Sum	Codes and Comments
2	1	$x = \pi, y = \pi, z = 0$ (optimal)
3	$\sqrt{3}/2$	$x = 2\pi/3, y = 2\pi/3, z = 0$ (optimal)
5	$\sqrt{5/8}$	$x = 2\pi/5, y = 8\pi/5, z = 4\pi/5$ (optimal)
9	$3/4$	$x = 10\pi/9, y = 4\pi/3, z = 4\pi/9$ (optimal)
16	$\sqrt{2}/2$	$x = \pi/4, y = 5\pi/4, z = 13\pi/8$ (optimal)
18	0.6614	$x = 4\pi/9, y = 2\pi/3, z = 7\pi/9$
19	0.6391	SA searched code
20	0.6338	SA searched code
21	0.6307	SA searched code
22	0.6154	SA searched code
24	0.6124	$x = \pi/6, y = \pi/4, z = 5\pi/12$
28	0.5996	$x = 3\pi/8, y = \pi/2, z = 2\pi/7$
30	0.5934	$x = 4\pi/15, y = \pi/3, z = 7\pi/15$
31	0.5739	SA searched code
32	0.5734	SA searched code
39	0.5726	$x = 14\pi/39, y = 40\pi/39, z = 18\pi/39$
40	0.5499	$x = 3\pi/20, y = 7\pi/20, z = 3\pi/10$
42	0.5371	$x = 4\pi/7, y = 13\pi/21, z = \pi/3$
45	0.5342	$x = 2\pi/9, y = 4\pi/9, z = 14\pi/15$
52	0.5332	$x = \pi/13, y = 2\pi/13, z = 9\pi/26$
57	0.5053	$x = 4\pi/57, y = 8\pi/57, z = 40\pi/57$
60	0.5000	$x = \pi/15, y = 4\pi/15, z = 3\pi/10$
64	0.4852	$x = 3\pi/16, y = 53\pi/32, z = 55\pi/32$
75	0.4850	$x = 32\pi/75, y = 14\pi/75, z = 2\pi/75$
76	0.4672	$x = 3\pi/19, y = 4\pi/19, z = 11\pi/38$
77	0.4595	$x = 52\pi/77, y = 82\pi/77, z = 60\pi/77$
85	0.4540	$x = 2\pi/17, y = 8\pi/17, z = 14\pi/85$
87	0.4460	$x = 52\pi/87, y = 98\pi/87, z = 82\pi/87$
95	0.4418	$x = 6\pi/19, y = 2\pi/95, z = 36\pi/95$
96	0.4390	$x = 39\pi/48, y = 5\pi/12, z = 11\pi/24$
99	0.4297	$x = 62\pi/99, y = 192\pi/99, z = 142\pi/99$
105	0.4295	$x = 2\pi/105, y = 16\pi/105, z = 28\pi/105$
106	0.4161	$x = 2\pi/53, y = 13\pi/53, z = 12\pi/53$
120	0.4156	$x = \pi/10, y = \pi/6, z = 5\pi/4$
123	0.4077	$x = 188\pi/123, y = 38\pi/123, z = 182\pi/123$

Table 6.6. The diversity sum of 2 dimensional constellations (cont.)

number of elements	Diversity Sum	Codes and Comments
130	0.4071	$x = 26\pi/65, y = 5\pi/13, z = 2\pi/13$
133	0.3971	$x = 2\pi/133, y = 212\pi/133, z = 206\pi/133$
138	0.3963	$x = 16\pi/69, y = 19\pi/69, z = 4\pi/69$
145	0.3949	$x = 138\pi/145, y = 22\pi/145, z = 40\pi/29$
148	0.3840	$x = 5\pi/74, y = 13\pi/37, z = 2\pi/37$
150	0.3758	$x = \pi/15, y = 8\pi/75, z = 19\pi/75$
155	0.3828	$x = 2\pi/5, y = 26\pi/31, z = 58\pi/31$
156	0.3824	$x = 5\pi/39, y = 8\pi/39, z = 15\pi/78$
158	0.3823	$x = 58\pi/79, y = 81\pi/79, z = 64\pi/79$
159	0.3814	$x = 8\pi/159, y = 64\pi/159, z = 30\pi/159$
160	0.3802	$x = 69\pi/80, y = 59\pi/80, z = 37\pi/20$
162	0.3770	$x = 53\pi/21, y = 10\pi/9, z = 19\pi/81$
165	0.3760	$x = 24\pi/165, y = 26\pi/165, z = 34\pi/165$
166	0.3699	$x = 14\pi/83, y = 21\pi/83, z = 10\pi/83$
169	0.3696	$x = 56\pi/169, y = 76\pi/169, z = 284\pi/169$
171	0.3678	$x = 32\pi/171, y = 294\pi/171, z = 6\pi/171$
178	0.3664	$x = 145\pi/89, y = 26\pi/89, z = 10\pi/89$
180	0.3636	$x = \pi/9, y = 97\pi/90, z = 127\pi/90$
193	0.3598	$x = 90\pi/193, y = 98\pi/193, z = 26\pi/193$
204	0.3566	$x = 13\pi/51, y = 4\pi/51, z = 5\pi/34$
208	0.3501	$x = \pi/13, y = 8\pi/13, z = 65\pi/104$
214	0.3476	$x = 98\pi/107, y = 67\pi/107, z = 59\pi/107$
220	0.3459	$x = 19\pi/11, y = 163\pi/110, z = 121\pi/110$
222	0.3438	$x = 19\pi/111, y = 22\pi/111, z = 15\pi/111$
225	0.3420	$x = 2\pi/225, y = 52\pi/225, z = 414\pi/225$
234	0.3410	$x = 4\pi/117, y = 24\pi/117, z = 43\pi/117$
240	0.3371	$x = 71\pi/120, y = 11\pi/10, z = 187\pi/120$
244	0.3335	$x = 39\pi/122, y = 14\pi/61, z = 20\pi/61$
245	0.3305	$x = 16\pi/245, y = 186\pi/245, z = 46\pi/245$
248	0.3291	$x = 103\pi/124, y = 39\pi/31, z = 179\pi/124$
259	0.3288	$x = 30\pi/259, y = 44\pi/259, z = 42\pi/259$
262	0.3274	$x = 142\pi/131, y = 215\pi/131, z = 87\pi/131$
264	0.3247	$x = 79\pi/66, y = 129\pi/66, z = 215\pi/132$
276	0.3237	$x = 23\pi/138, y = 15\pi/69, z = 6\pi/69$
287	0.3188	$x = 6\pi/287, y = 76\pi/287, z = 28\pi/287$
292	0.3164	$x = 65\pi/146, y = 14\pi/73, z = 82\pi/73$
295	0.3147	$x = \pi/5, y = 50\pi/59, z = 22\pi/59$
300	0.3126	$x = \pi/75, y = 17\pi/150, z = 9\pi/25$

Table 6.7. The diversity of non-square constellation ($T = 5$, $M = 2$)

Size	DP
3	0.8527
4	0.8152
5	0.7171
6	0.7668
7	0.7493
8	0.7418
9	0.7183
10	0.6608
20	0.6240
30	0.5985
40	0.5552
50	0.5556
60	0.5088
70	0.4976
80	0.4978
90	0.4487
100	0.4450
200	0.3564
300	0.3563
400	0.3205
500	0.2955
600	0.2821
700	0.2791
800	0.2723
900	0.2472
1000	0.2643
2000	0.2086
3000	0.1867
4000	0.1740
5000	0.1770
6000	0.1545
7000	0.1466
8000	0.1245
9000	0.1296
10000	0.1426

Size	DS
3	0.8693
4	0.8589
5	0.8243
6	0.7976
7	0.7960
8	0.7844
9	0.7659
10	0.7737
20	0.7243
30	0.6837
40	0.6576
50	0.6392
60	0.6237
70	0.5987
80	0.5801
90	0.5775
100	0.5555
200	0.4909
300	0.4369
400	0.4088
500	0.3877
600	0.3687
700	0.3566
800	0.3504
900	0.3358
1000	0.3352
2000	0.2796
3000	0.2461
4000	0.2351
5000	0.1863
6000	0.2163
7000	0.2005
8000	0.2000
9000	0.1874
10000	0.1735

$V\text{diag}\,(e^{i\beta_1}, e^{i\beta_2}, \cdots, e^{i\beta_M})V^*$. Consider unitary differential modulation [28] and denote with X_τ the received signal at time block τ. The ML demodulation algorithm involves the following minimization problem:

$$(\hat{k}, \hat{l}) = \arg\min_{k,l} \|X_\tau - A^k B^l X_{\tau-1}\|.$$

Algebraically one can check that

$$\|X_\tau - A^k B^l X_{\tau-1}\| = \|A^{-k} X_\tau - B^l X_{\tau-1}\|$$

$$= \|U\text{diag}\,(e^{-ik\alpha_1}, e^{-ik\alpha_2}, \cdots, e^{-ik\alpha_M})U^* X_\tau - V\text{diag}\,(e^{-il\beta_1}, e^{-il\beta_2}, \cdots, e^{-il\beta_M})V^* X_{\tau-1}\|$$

So every entry of $X_\tau - A^k B^l X_{\tau-1}$ is a linear combination of trigonometric functions cos or sin in the variables k, l, which can be viewed as lattice points. As demonstrated in [31] and [25], the whole demodulation task has been converted to a least-square problem. The same fast conversion from ML decoding to lattice decoding can applied to general form constellations in 6.5.5 too. Consequently our structured constellation will admit sphere decoding algorithm. In [31] a detailed study of the sphere decoding algorithm applied to constellations from $Sp(2)$ was undertaken.

The complexity (either upper bound or average complexity) of sphere decoding will depend on the dimension of the lattice. This will make the weak group structure $A^k B^k$ more remarkable, because in this case the algorithm requires considering finding the closest point in a one dimensional lattice, which is very simple.

In [8] a very interesting fast demodulation approach is proposed for diagonal space time constellations. The authors use numerical approximation and LLL basis reduction technique to reduce the decoding complexity. Note that a constellation with the weak group structure A^k essentially is a diagonal constellation (straightforward Schur decomposition will show this), therefore the same technique can be

applied to this structure. Most importantly some other algebraic structures can employ the techniques too. For instance, consider the $A^k B^l C^m$ structure. If we let l go over a large interval and let k, m stay within a small interval, the structure will become "almost" diagonal. For efficient decoding, one only has to do exhaustive search for k, m and apply the techniques for diagonal constellations to decode l. Although the decoding complexity will increase a little, our experiments show the performance will outperform the diagonal one remarkably. Exactly the same "almost" diagonal idea can be applied to other proposed structures.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

In this dissertation, we study the limiting behavior of the *diversity function* by letting the SNR go to either infinity or zero. Respectively the *diversity product* and the *diversity sum* as unitary constellation design criteria are derived from the analysis of the limiting behavior.

Using the packing techniques on the compact Lie group $U(n)$ we presented three approaches to derive upper bounds for the diversity sum of unitary constellations with any dimension n and any size m. At certain situations the derived upper bounds give tighter results compared to the currently existing ones.

The complex and real orthogonal coding schemes admit simple decoding algorithms. Based on these schemes, we generalize one dimensional PSK signals and explicitly construct GPSK unitary space time constellations. These constellations can be viewed as higher dimensional generalizations of one dimensional PSK signals and theoretical analysis shows that their decoding procedures are decomposable, *i.e.*, the demodulation of these codes can be boiled down to one dimensional PSK demodulation. Therefore our constellations have very simple decoding procedures. For some of the resulted codes (for example, $\mathcal{V}_1(n), \mathcal{V}_2(n)$), the complexity of ML decoding does not even depend on the transmission rate. We use scaled one dimensional PSK signals to optimize the diversity product and reallocate the power for each transmit antenna (meanwhile keeping the same total energy). Numerical

experiments show that our codes perform better than some of the currently existing compatible ones. For the sample programs regarding to how to construct the proposed constellations, we refer to [18].

We propose algebraic structures, which are suitable for constructing unitary space time constellations and feature fast decoding algorithms. Based on the presented structures we construct unitary constellations using geometrical symmetry and numerical methods. For 2 dimension most of our codes are better or equal to the currently existing ones. For higher dimensions many codes with excellent diversity are found, which were never found before. Combined with the proposed algebraic structure the numerical methods can also be employed to optimize the diversity function at a certain SNR.

There are many problems in space time coding that need to be answered. In the following several research directions will be discussed.

- **Upper bound analysis of diversity product:** Although the resulted upper bound [17] is also the one of diversity product of the same constellation, intuitively we can improve the margin between the bound and the maximal possible value in terms of diversity product. The future work may involve the derivation of a tighter upper bound analysis of diversity product of a unitary constellation. In contrast to diversity sum, diversity product is highly non-convex. We are considering packing some non-typical small "sphere"s into $U(n)$ instead of packing standard Euclidean spheres into $U(n)$.

- **GPSK constellations for any dimension** The theoretical analysis [43] shows that the complex and real orthogonal designed constellations only exist for $2, 4$ or 8 dimension. Based on these schemes we only propose $2, 4$ and 8 dimensional GPSK constellations. We are trying to generalize the idea of decomposable GPSK constellations for any dimension. Constellations with

simple decoding algorithms based on the generalized orthogonal design [30] is under close investigations.

- **Optimization of structured constellation:** By exploiting the algebraic structure proposed in [21], one can greatly reduce the dimension of the parameter space and increase the possibility of obtaining constellations with large diversity. We used the Simulated Annealing algorithm in [21] to design such structured constellations. However, it would be interesting to investigate some gradient-based optimization methods on the underlying Riemannian manifold. As a special case of Riemannian manifold, $U(n)$ is also a compact Lie group with homogeneity and nice symmetry. Consequently a lot of geometrical aspects of $U(n)$, such as geodesics, tangents, gradients and Hessians can be described in closed forms. Therefore the steepest descent gradient method, Newton method or conjugate gradient method can be applied to search for good diversity constellations and they should be very promising. This direction is under close investigation now.

- **Differential space time modulation:** Presented in [28], Unitary Space Time Modulation can be viewed as higher dimensional generalization of DPSK. It is well known that PSK signals have unit energy and only differs by their phases. On the contrary, QAM signals differ by both their energies and phases. Since for certain situations QAM signals fit better than PSK signals, the question is that how the higher dimensional generalization of QAM signals perform. Since we still want to differentially modulate the matrix signals, we must choose finite matrix group to avoid the energy of the modulated signals going to infinity. We are trying to analyze non-coherent channels with generalized QAM signals and the representations of certain types of finite groups

are under investigation to design good-performing constellations.

- **Lattice decoding and Cayley transform:**

  Constellations with lattice structure usually feature fast decoding scheme and they are heavily explored in the literature. On the other hand, unitary constellation are relatively new and few results have been known so far. Can we convert the unitary constellation design problem to lattice design problem? Cayley transform [24] might be a very powerful tool to explore this problem. Cayley transform essentially maps $U(n)$ to standard Euclidean space, which is much easier to impose lattice structure. What type of lattice should we choose? How will these unitary constellations with lattice structure perform? A lot of questions will be addressed in the future research work.

In fact there are many other questions to explore in space time coding. For instance, how to combine the coding gain and diversity gain together to design an efficient coded modulation schemes? We are considering to combine LDPC codes over Galois field with unitary constellations we proposed to achieve this goal. Is there other transmission scheme, which feature low complexity decoding scheme for non-coherent channel? Many problems seem to be open in the study of space time coding.

BIBLIOGRAPHY

[1] E. H. L. Aarts and J. Korst. *Simulated Annealing and Boltzmann Machines*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Ltd., Chichester, 1989. A stochastic approach to combinatorial optimization and neural computing.

[2] S. Abhyankar. *Algebraic Geometry for Scientists and Engineers*. American Mathematical Society, Providence, R.I., 1990.

[3] S. M. Alamouti. A simple transmitter diversity scheme for wireless communications. *IEEE J. Selected Areas of Commun.*, pages 1451–1458, October 1998.

[4] M. F. Atiyah and J. A. Todd. On complex Stiefel manifolds. *Proc. Cambridge Philos. Soc.*, 56:342–353, 1960.

[5] K. E. Atkinson. *An Introduction to Numerical Analysis*. J. Wiley, New York, 1978.

[6] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon limit error-correcting coding and decoding: Turbo-codes. In *Proc. of IEEE Int. Conference on Communication*, pages 1064–1070, Geneva, Switzerland, May 1993.

[7] W. M. Boothby. *An introduction to differentiable manifolds and Riemannian geometry*, volume 120 of *Pure and applied mathematics, a series of monographs and textbooks*. Academic Press, Orlando, Fla, 1986.

[8] K. L. Clarkson, W. Sweldens, and A. Zheng. Fast multiple-antenna differential decoding. *IEEE Trans. on communications*, 49(2):253–261, February 2001.

[9] J. H. Conway and N. J. A. Sloane. *Sphere packings, lattices and groups*. Springer-Verlag, New York, second edition, 1993. With additional contributions by E. Bannai, R. E. Borcherds, J. Leech, S. P. Norton, A. M. Odlyzko, R. A. Parker, L. Queen and B. B. Venkov.

[10] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley & Sons, New York, 1991.

[11] M. Eaton. *Multivariate statistics*. Wiley, 1983.

[12] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20(2):303–353 (electronic), 1999.

[13] U. Fincke and M. Pohst. Improved methods for calculating vectors of short length in a lattice, including a complexity analysis. *Mathematics of Computation*, 44:463–471, April 1985.

[14] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Tech. J.*, 1(2):41–59, 1996.

[15] R.G. Gallager. Low-density parity-check codes. *IRE Trans. on Info. Theory*, IT-8:21–28, 1962.

[16] R. Goodman and N. R. Wallach. *Representations and invariants of the classical groups*. Cambridge University Press, Cambridge, U.K. ; New York, NY, USA, 1998.

[17] G. Han and J. Rosenthal. Unitary Space Time Constellation Analysis: An Upper Bound for the Diversity, January 2004. http://front.math.ucdavis.edu/math.CO/0401045.

[18] G. Han and J. Rosenthal. A website of unitary space time constellations with large diversity. http://www.nd.edu/~eecoding/space-time/.

[19] G. Han and J. Rosenthal. Unitary constellation design and its application to space-time coding. In D. Gilliam and J. Rosenthal, editors, *Proceedings of the 15-th International Symposium on the Mathematical Theory of Networks and Systems*, University of Notre Dame, August 2002.

[20] G. Han and J. Rosenthal. Unitary constellations with large diversity sum and good diversity product. In *Proc. of the 40-th Allerton Conference on Communication, Control, and Computing*, pages 48–57, 2002.

[21] G. Han and J. Rosenthal. Geometrical and numerical design of structured unitary space time constellations, December 2003. http://front.math.ucdavis.edu/math.OC/0312170.

[22] G. Han. Generalized PSK in space time coding. January 2004. http://front.math.ucdavis.edu/math.OC/0401157.

[23] R. Hartshorne. *Algebraic Geometry*. Springer Verlag, Berlin, 1977.

[24] B. Hassibi and B. M. Hochwald. Cayley differential unitary space-time codes. *IEEE Trans. Inform. Theory*, 48(6):1485–1503, 2002. Special issue on Shannon theory: perspective, trends, and applications.

[25] B. Hassibi and H. Vikalo. On the expected complexity of integer least-squares problems. In *Acoustics, Speech and Signal Processing, 2002 IEEE International conference*, pages 1497–1500, April 2002.

[26] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces*. Graduate studies in mathematics. Providence, R.I. : American Mathematical Society, 2001.

[27] B. Hochwald, T. Marzetta, T. Richardson, W. Sweldens, and R. Urbanke. Systematic design of unitary space-time constellations. *IEEE Trans. Inform. Theory*, 46(6):1962–1973, 2000.

[28] B. Hochwald and W. Sweldens. Differential unitary space-time modulation. *IEEE Trans. Comm.*, pages 2041–2052, December 2000.

[29] B. M. Hochwald and T. L. Marzetta. Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading. *IEEE Trans. Inform. Theory*, 46(2):543–564, 2000.

[30] H. Jafarkhani and V. Tarokh. Multiple transmit antenna differential detection from generalized orthogonal designs. *IEEE Trans. Inform. Theory*, 47(6):2626–2631, 2001.

[31] Y. Jing and B. Hassibi. Fully-diverse sp(2) code design. In *Proceedings of the 2003 IEEE International Symposium on Information Theory*, page 299, Yokohoma, Japan, 2003.

[32] X.-B. Liang and X.-G. Xia. Unitary signal constellations for differential space-time modulation with two transmit antennas: Parametric codes, optimal designs and bounds. *IEEE Trans. Inform. Theory*, 48(8):2291–2322, August 2002.

[33] S. Lin and D. J. Costello Jr. *Error Control Coding: Fundamentals and Applications*. Prentice-Hall, Englewood Cliffs, NJ, 1983.

[34] T. L. Marzetta and B. M. Howchwald. Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading. *IEEE Trans. Inform. Theory*, 45(1):139–157, 1999.

[35] N. Metropolis, A.W. Rosenbluth, M.N.Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. of Chem. Phys.*, 21(6):1087–1092, 1953.

[36] R. H. J. M. Otten and L. P. P. P. van Ginneken. *The Annealing Algorithm*. The Kluwer International Series in Engineering and Computer Science. VLSI, Computer Architecture and Digital Signal Processing. Kluwer Academic Publishers, Boston, MA, 1989.

[37] V. V. Prasolov. *Problems and theorems in linear algebra*, volume 134 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1994. Translated from the Russian manuscript by D. A. Leĭtes.

[38] L. A. Santaló. *Integral geometry and geometric probability*, volume 1 of *Encyclopedia of mathematics and its applications ; v. 1 Section, Probability*. Addison-Wesley Pub. Co., 1976.

[39] C.E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423 and 623–656, 1948.

[40] A. Shokrollahi. Computing the performance of unitary space-time group codes from their character table. *IEEE Trans. Inform. Theory*, 48(6):1355–1371, 2002. Special issue on Shannon theory: perspective, trends, and applications.

[41] A. Shokrollahi, B. Hassibi, B. M. Hochwald, and W. Sweldens. Representation theory for high-rate multiple-antenna code design. *IEEE Trans. Inform. Theory*, 47(6):2335–2367, 2001.

[42] V. Tarokh and H. Jafarkhani. A differential detection scheme for transmit diversity. *IEEE J. Sel. Area Comm.*, 18(7):1169–1174, 2000.

[43] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block codes from orthogonal designs. *IEEE Trans. Inform. Theory*, 45(5):1456–1467, 1999.

[44] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: performance criterion and code construction. *IEEE Trans. Inform. Theory*, 44(2):744–765, 1998.

[45] İ. E. Telatar. Capacity of multi-antenna Gaussian channels. *European Trans. Telecommun.*, pages 585–595, 1999.

[46] P. J. M. van Laarhoven and E. H. L. Aarts. *Simulated Annealing: Theory and Applications*, volume 37 of *Mathematics and its Applications*. D. Reidel Publishing Co., Dordrecht, 1987.

[47] H. Zassenhaus. über endliche Fastkörper. *Abh. Math. Sem. Hamburg*, 11:187–220, 1936.

[48] L. Zheng and D. N. C. Tse. Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel. *IEEE Trans. Inform. Theory*, 48(2):359–383, 2002.

[49] R.E. Ziemer and R.L. Peterson. *Introduction to Digital Communication*. MacMillan, 2nd ed. edition, 2000.