# Space-Time Code Design and Its Applications in Wireless Networks

Thesis by

Yindi Jing

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy


California Institute of Technology

Pasadena, California

2004

(Submitted September 7, 2004)

# Acknowledgments

I owe a debt of gratitude to many people who have helped me with my graduate study and research in diverse ways. Without their generosity and assistance, the completion of this thesis would not have been possible.

First of all, I would like to express my deepest gratitude and appreciation to my advisor, Professor John C. Doyle, for his excellent guidance and generous support. He allowed me to initiate my graduate studies at California Institute of Technology, one of the most elite graduate universities in the world. John has incredible vision and boundless energy. He is also an endless source of creative ideas. Often times, I have realized how truly fortunate I am to have such an open-minded advisor who allowed me to choose my research subject freely.

My greatest and heartfelt thanks must also go to Professor Babak Hassibi, my associate advisor and mentor, for his constant encouragement, inspiration, and guidance both in completing this thesis and in my professional development. He led me to the exciting world of wireless communications. He not only always has great insights but also shows his students how to start from an ultimate vision of the world and reduce it to a tractable problem. It is hard to imagine having done my Ph.D. without his help.

I would like to thank the other members of my dissertation committee, Professor Robert J. McEliece, Professor P. P. Vaidyanathan, Professor Steven Low, and my other candidacy committee member Professor Michael Aschbacher, for their valuable time, comments, feedback and interest in this work.

I would like to express my gratitude to my parents Zhaojin Jing and Yufang Lei for their endless emotional support. I also credit my husband, Xinwei Yu, for

me surviving my last two years of graduate study. It is his unwavering love and unconditional support that inspire my life and work. I also learned a lot from him.

I am grateful to my officemates of Moore 155C, Radhika Gowaikar and Chaitanya Rao, for making my graduate school experience both memorable and fun. Chaitanya also kindly helped me proofread this thesis. I would also like to thank other members of the wireless communications group, Amir F. Dana, Masoud Sharif, Mihailo Stojnic, Vijay Gupta, and Haris Vikalo. Great thanks to Maralle Fakhereddin, a summer intern, who spent a lot of time and energy proofreading this thesis.

Special thanks to my friends Lun Li and Min Tao for their help, support, and friendship during my darkest time. My lifetime friend Bing Liu deserves special mention for his support and concern. He is like a family member to me.

I would also like to acknowledge my officemates of Steele 7, Xin Liu and Domitilla del Vecchio, my classmate and former officemate of Steele 135 and Steele 4, Jiantao Wang, for discussing homework and research problems during my first two years at Caltech, Jim Endrizzi, the international student advisor of Caltech International Students Program office, for helping me with international student issues, Caltech Chinese Association, and Caltech Badminton Club for entertaining my stay at Caltech.

# Abstract

This thesis has two main contributions: the designs of differential/non-differential unitary space-time codes for multiple-antenna systems and the analysis of the diversity gain when using space-time coding among nodes in wireless networks.

Capacity has long been a bottleneck in wireless communications. Recently, multiple-antenna techniques have been used in wireless communications to combat the fading effect, which improves both the channel capacity and performance greatly. A recently proposed method for communicating with multiple antennas over block-fading channels is unitary space-time modulation, which can achieve the channel capacity at high SNR. However, it is not clear how to generate well performing unitary space-time codes that lend themselves to efficient encoding and decoding. In this thesis, the design of unitary space-time codes using Cayley transform is proposed. The codes are designed based on an information-theoretic criterion and have a polynomial-time near-maximum-likelihood decoding algorithm. Simulations suggest that the resulting codes allow for effective high-rate data transmissions in multiple-antenna communication systems without knowing the channel. Another well-known transmission scheme for multiple-antenna systems with unknown channel information at both the transmitter and the receiver is differential unitary space-time modulation. It can be regarded as a generalization of DPSK and is suitable for continuous fading. In differential unitary space-time modulation, fully diverse constellations, i.e., sets of unitary matrices whose pairwise differences are non-singular, are wanted for their good pairwise error properties. In this thesis, Lie groups and their representations are used in solving the design problem. Fully diverse differential unitary space-time codes for systems with four and three transmit antennas are constructed based on the

Lie groups $Sp(2)$ and $SU(3)$. The designed codes have high diversity products, lend themselves to a fast maximum-likelihood decoding algorithm, and simulation results show that they outperform other existing codes, especially at high SNR.

Then the idea of space-time coding devised for multiple-antenna systems is applied to communications over wireless networks. In wireless relay networks, the relay nodes encode the signals they receive from the transmit node into a distributed space-time code and transmit the encoded signals to the receive node. It is shown in this thesis that at very high SNR, the diversity gain achieved by this scheme is almost the same as that of a multiple-antenna system whose number of transmit antennas is the same as the number of relay nodes in the network, which means that the relay nodes work as if they can cooperate fully and have full knowledge of the message. However, at moderate SNR, the diversity gain of the wireless network is inferior to that of the multiple-antenna system. It is further shown that for a fixed total power consumed in the network, the optimal power allocation is that the transmitter uses half the power and the relays share the other half fairly. This result addresses the question of what performance a relay network can achieve. Both it and its extensions have many applications to wireless ad hoc and sensory network communications.

# Contents

# List of Figures

# List of Tables

# List of Symbols

| | |
|---|---|
| $A^t$ | transpose of $A$ |
| $\bar{A}$ | conjugate of $A$ |
| $A^*$ | conjugate transpose of $A$ |
| $A^\perp$ | unitary complement of $A$ |
| | If $A$ is $m \times n$, $A^\perp$ is the $m \times (m-n)$ matrix such that $[A \;\; A^\perp]$ is unitary. |
| $\operatorname{tr} A$ | trace of $A$ |
| $\det A$ | determinant of $A$ |
| $\operatorname{rank} A$ | rank of $A$ |
| $\|A\|_F$ | Frobenius norm of $A$ |
| $A_{Re}$ | real part of $A$ |
| $A_{Im}$ | imaginary part of $A$ |
| $a_{ij}$ | $(i,j)$-th entry of $A$ |
| $\log$ | natural logarithm |
| $\log_{10}$ | base-10 logarithm |
| $I_n$ | $n \times n$ identity matrix |
| $\mathbf{0}_{mn}$ | $m \times n$ matrix with all zero entries |
| $\operatorname{diag}\{\lambda_1, \cdots, \lambda_n\}$ | diagonal matrix with diagonal entries $\lambda_1, \cdots, \lambda_n$. |
| P | probability |

| | |
|---|---|
| $\mathrm{E}$ | expected value |
| $\mathrm{Var}$ | variance |
| $\mathrm{sgn}$ | sign function |
| $\lceil x \rceil$ | smallest integer that is larger than $x$ |
| $\lfloor x \rfloor$ | largest integer that is smaller than $x$ |
| $\gcd(m, n)$ | greatest common divisor of $m$ and $n$ |
| $\lvert x \rvert$ | absolute value of $x$ |
| $\min\{x_1, x_2\}$ | minimum of $x_1$ and $x_2$ |
| $\max\{x_1, x_2\}$ | maximum of $x_1$ and $x_2$ |
| $a^+$ | maximum of $a$ and 0 |
| $\mathrm{Arg}\, x$ | angument of the complex scalar $x$ |
| $\Re x$, $[x]_{Re}$ | real part of $x$ |
| $\Im x$, $[x]_{Im}$ | imaginary part of $x$ |

| | |
|---|---|
| $\mathbb{Z}$ | set of integers |
| $\mathbb{R}$ | set of real numbers |
| $\mathbb{C}$ | set of complex numbers |
| $\mathbb{R}^n$ | set of $n$-dimensional real vectors |
| $\mathbb{C}^n$ | set of $n$-dimensional complex vectors |
| $\mathbb{C}^{n \times n}$ | set of $n \times n$ complex matrices |
| $\lvert \mathcal{C} \rvert$ | cardinality of the set $\mathcal{C}$ |
| $O(x)$ | order of $x$ |
| $o(x)$ | lower order of $x$ |

| | |
|---|---|
| AWGN | additive white Gaussian noise |
| BER | bit error rate |
| BLER | block error rate |
| DPSK | differential phase-shift keying |
| iid | identical independent distribution |
| ISI | inter-symbol interference |
| LD | linear dispersion |
| MIMO | multiple-input-multiple-output |
| ML | maximum-likelihood |
| OFDM | orthogonal frequency division multiplexing |
| PEP | pairwise error probability |
| PSK | phase shift keying |
| QAM | quadrature amplitude modulation |
| SNR | signal-to-noise ratio |
| TDMA | time division multiple access |
| USTM | unitary space-time modulation |

# Chapter 1 Introduction to Multiple-Antenna Communication Systems

## 1.1 Introduction

Wireless communications first appeared in 1897, when Guglielmo Marconi demonstrated radio's ability to provide contact with ships sailing the English channel. During the following one hundred years, wireless communications has experienced remarkable evolution, for example, the appearance of AM and FM communication systems for radios [Hay01] and the development of the cellular phone system from its first generation in the 1970s to the third generation, which we are about to use soon [Cal03, Stu00, Rap02]. The use of wireless communications met its greatest increase in the last ten years, during which new methods were introduced and new devices invented. Nowadays, we are surrounded by wireless devices and networks in our everyday lives: cellular phone, handheld PDA, wireless INTERNET, walkie-talkie, etc. The ultimate goal of wireless communications is to communicate with anybody from anywhere at anytime for anything.

In reaching this ultimate goal, the bottleneck lies in the capacity of wireless communication systems, that is, how much information can go through the system. With the increasing use of diverse wireless facilities, the demand for bandwidth or capacity becomes more and more urgent, especially for power and complexity limited systems. This means that we can not increase capacity by simply increasing the transmit power. Communication systems in use today are predominantly single-antenna sys-

tems. Because of the multiple-path propagation in wireless channels, the capacity of a single wireless channel can be very low. Research efforts have focused on ways to make more efficient use of this limited capacity and have accomplished remarkable progresses. On the one hand, efficient techniques, such as frequency reuse [Rap02] and OFDM [BS99], have been invented to increase the bandwidth efficiency; on the other hand, advances in coding such as turbo codes [BGT93] and low density parity check codes [Gal62, MN96, McE02] make it feasible to almost reach Shannon capacity [CT91, McE02], the theoretical upper bound for the capacity of the system. However, a conclusion that the capacity bottleneck has been broken is still far-fetched.

Other than low Shannon capacity, single-antenna systems suffer another great disadvantage: its high error rate. In an additive white Gaussian noise (AWGN) channel, which models a typical wired channel, the pairwise error probability (PEP), the probability of mistaking the transmitted signal with another one, decreases exponentially with the signal-to-noise ratio (SNR), while due to the fading effect, the average PEP for wireless single-antenna systems only decreases linearly with SNR. Therefore, to achieve the same performance, a much longer code or much higher transmit power is needed for single-antenna wireless communication systems.

Given the above disadvantages, single-antenna systems are unpromising candidates to meet the needs of future wireless communications. Therefore, new communication systems superior in capacity and error rate must be introduced and consequently, new communication theories for these systems are of great importance at the present time.

Recently, one such new systems, digital communication systems using multiple-input-multiple-output (MIMO) wireless links, that is, using multiple antennas at the transmitter and the receiver, has emerged. This is one of the most significant technical breakthroughs in modern communications. The key feature of a multiple-antenna system is its ability to turn multiple-path propagation, which is traditionally

regarded as a disadvantage to wireless communications, into a benefit to the users.

In 1996 and 1999, Foschini and Telatar proved in [Fos96] and [Tel99] that communication systems with multiple antennas have a much higher capacity than single-antenna systems. They showed that the capacity improvement is almost linear in the number of transmit antennas or the number of receive antennas, whichever is smaller. This result indicated the superiority of multiple-antenna systems and ignited great interest in this area. In few years, much work has been done generalizing and improving their results. On the one hand, for example, instead of assuming that the channels have rich scattering so that the propagation coefficients between transmit and receive antennas are independent, it was assumed that correlation can exist between the channels; on the other hand, unrealistic assumptions, such as perfect channel knowledge at both the transmitter and the receiver are replaced by more realistic assumptions of partial or no channel information at the receiver. Information theoretic capacity results have been obtained under these and other new assumptions, for example, [ZT02, SFGK00, CTK02, CFG02].

These results indicate that multiple-antenna systems have much higher Shannon capacity than single-antenna ones. However, since Shannon capacity can only be achieved by codes with unbounded complexity and delay, the above results do not reflect the performance of real transmission systems. For example, in a system with two transmit antennas, if identical signals are transmitted from both antennas at a time, a PEP that is inversely proportional to SNR is obtained, which is the same as that of single-antenna systems although the coding gain is improved. However, if Alamouti's scheme [Ala98] is used, a PEP that behaves as $\mathrm{SNR}^{-2}$ is obtained. Therefore, it is important to develop algorithms that take advantage of the spatial diversity provided by multiple antennas. Many algorithms with reasonable complexity and performance have been proposed, for example, the diversity techniques and diversity combining methods (see [Win83, Wit93, Stu00, Rap02, Pro00]). Among

them, the most successful one is space-time coding.

In space-time coding, the signal processing at the transmitter is done not only in the time dimension, as what is normally done in many single-antenna communication systems, but also in the spatial dimension. Redundancy is added coherently to both dimensions. By doing this, both the data rate and the performance are improved by many orders of magnitude with no extra cost of spectrum. This is also the main reason that space-time coding attracts much attention from academic researchers and industrial engineers alike.

The idea of space-time coding was first proposed by Tarokh, Seshadri and Calderbank in [TSC98]. They proved that space-time coding achieves a PEP that is inversely proportional to $\mathrm{SNR}^{MN}$, where $M$ is the number of transmit antennas and $N$ is the number of receive antennas. The number $MN$ is called the *diversity* of the space-time code. Comparing with the PEP of single-antenna systems, which is inversely proportional to the SNR, the error rate is reduced dramatically. It is also shown in [TSC98] that by using space-time coding, some coding gain can be obtained. The first practical space-time code is proposed by Alamouti in [Ala98], which works for systems with two transmit antennas. It is also one of the most successful space-time codes because of its great performance and simple decoding.

The result in [TSC98] is based on the assumption that the receiver has full knowledge of the channel, which is not a realistic assumption for systems with fast-changing channels. Hochwald and Marzetta studied the much more practical case where no channel knowledge is available at either the transmitter or the receiver. They first found a capacity-achieving space-time coding structure in [MH99] and based on this result, they proposed unitary space-time modulation [HM00]. In [HM00], they also proved that unitary space-time coding achieves the same diversity, $MN$, as general space-time coding.

Based on unitary space-time modulation, a transmission scheme that is better

tailored for systems with no channel information at both the transmitter and the receiver is proposed by Hochwald and Sweldens in [HS00] and Hughes in [Hug00a], which is called differential unitary space-time modulation. Differential unitary space-time modulation can be regarded as an extension of differential phase-shift keying (DPSK), a very successful transmission scheme for single-antenna systems.

During the last few years, the technology of multiple antennas and space-time coding has been improved greatly. There are many papers on the design of differential and non-differential unitary space-time codes, for example, [TJC99, HH02b, MHH02, HH02a, JH03e, JH03b, JH04e, GD03, DTB02]. There is also much effort in trying to improve the coding gain by combining space-time codes with other error-correcting codes or modulations, for example, [SD01, SG01, LFT01, Ari00, BBH00, FVY01, GL02, BD01, LB02, JS03]. Today, this area is still under intensive theoretical study.

In this thesis, the design of space-time codes is investigated in order to exploit the transmit diversity provided by the transmit antennas at the transmitter along with the applications of space-time coding in wireless networks in order to exploit the distributed spatial diversity provided by antennas of the distributed nodes in a network. The thesis has five parts. The first part includes Chapters 1 and 2, in which a brief but broad introduction of multiple-antenna systems and space-time coding is provided. Chapter 3 is Part II, which describes the design of unitary space-time codes using Cayley transform. Part III includes Chapters 4, 5 and 6, where the design of differential unitary space-time codes based on groups is discussed. In Chapter 4, concepts and background materials of groups and Lie groups are listed, along with motivations to the use of groups in differential unitary space-time code design. Chapters 5 and 6 are on the design of differential unitary space-time codes for systems with four transmit antennas and three transmit antennas based on the Lie groups $Sp(2)$ and $SU(3)$, respectively. Part IV is Chapter 7, in which the idea of space-time coding is used in wireless networks to exploit the distributed diversity

among the relay nodes. The last part, Chapter 8, is the summary and discussion.

## 1.2 Multiple-Antenna Communication System Model

Consider a wireless communication system with two users. One is the transmitter and the other is the receiver. The transmitter has $M$ transmit antennas and the receiver has $N$ receive antennas as illustrated in Figure 1.1. There exists a wireless channel between each pair of transmit and receive antennas. The channel between the $m$-th transmit antenna and the $n$-th receive antenna can be represented by the random propagation coefficient $h_{mn}$, whose statistics will be discussed later.

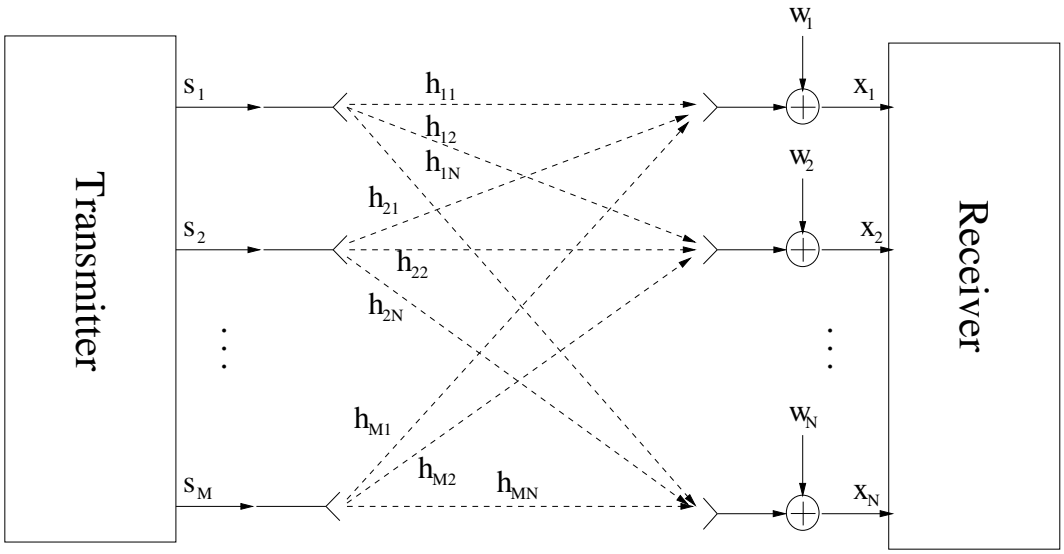


Figure 1.1: Multiple-antenna communication system

To send information to the receiver, at every transmission time, the transmitter feeds signals $s_1, \cdots, s_M$ to its $M$ antennas respectively. The antennas then send the signals simultaneously to the receiver. Every receive antenna at the receiver obtains a signal that is a superposition of the signals from every transmit antenna through the fading coefficient. The received signal is also corrupted by noise. If we denote the noise at the $n$-th receive antenna by $w_n$, the received signal at the $n$-th receive

antenna is

$$x_n = \sum_{m=1}^{M} h_{mn} s_m + w_n.$$

This is true for $n = 1, 2, \cdots, N$. If we define the vector of the trasmitted signal as $\mathbf{s} = [s_1, s_2, \cdots, s_M]$, the vector of the received signal as $\mathbf{x} = [x_1, x_2, \cdots, x_M]$, the vector of noise as $\mathbf{w} = [w_1, w_2, \cdots, w_M]$ and the channel matrix as

$$H = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1N} \\ h_{21} & h_{22} & \cdots & h_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1} & h_{M2} & \cdots & h_{MN} \end{bmatrix},$$

the system equation can be written as

$$\mathbf{x} = \mathbf{s}H + \mathbf{w}. \tag{1.1}$$

The total transmit power is $P = \mathbf{s}\mathbf{s}^* = \operatorname{tr} \mathbf{s}^*\mathbf{s}$.

## 1.3 Rayleigh Flat-Fading Channel

The wireless characteristic of the channel places fundamental limitations on the performance of wireless communication systems. Unlike wired channels that are stationary and predictable, wireless channels are extremely random and are not easily analyzed due to the diverse environment, the motion of the transmitter, the receiver, and the surrounding objects. In this section, characteristics of wireless channels are discussed and the Rayleigh flat-fading channel model is explained in detail.

In a mobile wireless environment, the surrounding objects, such as buildings, trees, and houses act as reflectors of electromagnetic waves. Due to these reflections, electromagnetic waves travel along different paths of varying lengths and therefore have

various amplitudes and phases. The interaction between these waves causes multiple fading at the receiver location, and the strength of the waves decreases as the distance between the transmitter and the receiver increases. Traditionally, propagation modeling focuses on two aspects. Propagation models that predict the mean signal strength for an arbitrary transmitter-receiver separation distance are called *large-scale* propagation models since they characterize signal strength over large transmitter-receiver distances. Propagation models that characterize the rapid fluctuations of the received signal strength over very short travel distances or short time durations are called *small scale* or *fading* models. In this thesis, the focus is on fading models, which are more suitable for indoor and urban areas.

Small-scale fading is affected by many factors, such as multiple-path propagation, speed of the transmitter and receiver, speed of surrounding objects, and the transmission bandwidth of the signal. In this work, narrowband systems are considered, in which the bandwidth of the transmitted signal is smaller than the channel's *coherence bandwidth*, which is defined as the frequency range over which the channel fading process is correlated. This type of fading is referred to as *flat fading* or *frequency nonselective fading*.

The Rayleigh distribution is commonly used to describe the statistical time-varying nature of the received envelope of a flat-fading signal. It is also used to model fading channels in this thesis. For a typical mobile wireless channel in indoor or urban areas, we may assume that the direct line-of-sight wave is obstructed and the receiver obtains only reflected waves from the surrounding objects. When the number of reflected waves is large, according to central limit theory, two quadrature components of the received signal are uncorrelated Gaussian random processes with mean zero and variance $\sigma^2$. As a result, the envelope of the received signal at any time instant has a Rayleigh distribution and its phase is uniform between $-\pi$ and $\pi$.

The probability density function of the Rayleigh distribution is given by

$$p(r) = \begin{cases} \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} & r \geq 0 \\ 0 & r < 0 \end{cases}.$$

If the fading coefficients in the multiple-antenna system model given in (1.1) are normalized by

$$\sum_{m=1}^{M} |h_{mn}^2| = M, \quad \text{for} \quad i = 1, 2, \cdots, N, \tag{1.2}$$

we have $\sigma^2 = \frac{1}{2}$. Therefore, the fading coefficient $h_{mn}$ has a complex Gaussian distribution with zero-mean and unit-variance, or equivalently, the real and imaginary parts of $h_{mn}$ are independent Gaussians with mean zero and variance $\frac{1}{2}$. Note that with (1.2),

$$\mathrm{E} \left| \sum_{m=1}^{M} h_{mn} s_n \right|^2 = \sum_{m=1}^{M} \mathrm{E} \, |h_{mn}|^2 |s_n|^2 = \sum_{m=1}^{M} \mathrm{E} \, |s_n|^2 = P,$$

which indicates that the normalization in (1.2) makes the received signal power at every receive antenna equals the total transmit power.

Another widely used channel model is the Ricean model which is suitable for the case when there is a dominant stationary signal component, such as a line-of-sight propagation path. The small-scale fading envelope is Ricean, with probability density function,

$$p(r) = \begin{cases} \frac{r}{\sigma^2} e^{-\frac{r^2 + A^2}{2\sigma^2}} \mathbf{I}_0 \left( \frac{Ar}{\sigma^2} \right) & \text{if} \quad r \geq 0 \\ 0 & \text{if} \quad r < 0 \end{cases}.$$

The parameter $A$ is always positive and denotes the peak amplitude of the dominant signal, and $\mathbf{I}_0(\cdot)$ is the zeroth-order modified Bessel function of the first kind [GR00]. For more information on propagation models, see [Rap02].

## 1.4 Capacity Results

As discussed in Section 1.1, communication systems with multiple antennas can greatly increase capacity, which is one of the main reasons that multiple-antenna systems are of great interest. This section is about the capacity of multiple-antenna communication systems with Rayleigh fading channels. Three cases are discussed: both the transmitter and the receiver know the channel, only the receiver knows the channel, and neither the transmitter nor the receiver knows the channel. The results are based on Telatar's results in [Tel99].

It is obvious that the capacity depends on the transmit power. Therefore, assume that the power constraint on the transmitted signal is

$$\mathrm{E}\,\mathrm{tr}\,\mathbf{s}^*\mathbf{s} \le P, \quad \text{or equivalently,} \quad \mathrm{E}\,\mathrm{tr}\,\mathbf{s}\mathbf{s}^* \le P.$$

In the first case, assume that both the transmitter and receiver know the channel matrix $H$. Note that $H$ is deterministic in this case. Consider the singular value decomposition of $H$ : $H = UDV^*$, where $U$ is an $M \times M$ unitary matrix, $V$ is an $N \times N$ unitary matrix, and $D$ is an $M \times N$ diagonal matrix with non-negative diagonal entries.[1] By defining $\tilde{\mathbf{x}} = V\mathbf{x}, \tilde{\mathbf{s}} = \mathbf{s}U$, and $\tilde{\mathbf{v}} = V\mathbf{v}$, the system equation (1.1) is equivalent to

$$\tilde{\mathbf{x}} = D\tilde{\mathbf{s}} + \tilde{\mathbf{v}}.$$

Since $\mathbf{v}$ is circularly symmetric complex Gaussian[2] with mean zero and variance $I_N$,

[1] An $M \times N$ matrix, $A$, is diagonal if its off-diagonal entries, $a_{ij}, i \neq j, i = 1, 2, \cdots, M, j = 1, 2, \cdots, N$, are zero.

[2] A complex vector $\mathbf{x} \in \mathbb{C}^n$ is said to be Gaussian if the real random vector $\hat{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_{Re} \\ \mathbf{x}_{Im} \end{bmatrix} \in \mathbb{R}^{2n}$ is Gaussian. $\mathbf{x}$ is circularly symmetric if the variance of $\hat{\mathbf{x}}$ has the structure

$$\begin{bmatrix} Q_{Re} & -Q_{Im} \\ Q_{Im} & Q_{Re} \end{bmatrix}$$

for some Hermitian non-negative definite matrix $Q \in \mathbb{C}^{n \times n}$. For more on this subject, see [Tel99, Ede89].

$\tilde{\mathbf{v}}$ is also circularly symmetric complex Gaussian with mean zero and variance $I_N$. Since the rank of $H$ is $\min\{M,N\}$, at most $\min\{M,N\}$ of its singular values are non-zero. Denote the non-zero singular values of $H$ as $\sqrt{\lambda_i}$. The system equation can be written component-wisely to get

$$\tilde{x}_i = \sqrt{\lambda_i}\tilde{s}_i + \tilde{n}_i, \quad \text{for} \quad 1 \le i \le \min\{M,N\}.$$

Therefore, the channel is decoupled into $\min\{M,N\}$ uncorrelated channels, which is equivalent to $\min\{M,N\}$ single-antenna systems. It is proved in [Tel99] that the capacity achieving distribution of $\tilde{s}_i$ is circularly symmetric Gaussian and the capacity for the $i$-th independent channel is $\log(1+\lambda_i P_i)$, where $P_i = \mathrm{E}\,\tilde{s}_i\tilde{s}_i^*$ is the power consumed in the $i$-th independent channel. Therefore, to maximize the mutual information, $\tilde{s}_i$ should be independent circularly symmetric Gaussian distributed and the transmit power should be allocated to the equivalent independent channels optimally. It is also proved in [Tel99] that the power allocation should follow "water-filling" mechanism. The power for the $i$-th sub-channel should be $\mathrm{E}\,\tilde{s}_i^*\tilde{s}_i = (\mu - \lambda_i^{-1})^+$, where $\mu$ is chosen such that $\sum_{i=1}^{\min\{M,N\}}(\mu-\lambda_i^{-1})^+ = P$.[3] The capacity of the system is thus

$$C = \sum_{i=1}^{\min\{M,N\}} \log(\mu\lambda_i),$$

which increases linearly in $\min\{M,N\}$.

When only the receiver knows the channel, the transmitter cannot perfrom the "water-filling" adaptive transmission. It is proved in [Tel99] that the channel capacity is given by

$$C = \log\det(I_N + (P/M)H^*H),$$

which is achieved when $\mathbf{s}$ is circularly symmetric complex Gaussian with mean zero

[3] $a^+$ denotes $\max\{o,a\}$.

and variance $(P/M)I_M$. When the channel matrix is random according to Rayleigh distribution, the expected capacity is just

$$C = \mathrm{E} \log \det(I_N + (P/M)H^*H),$$

where the expectation is over all possible channels.

For a fixed $N$, by the law of large numbers, $\lim_{M\to\infty} \frac{1}{M} H^*H = I_N$ with probability 1. Thus the capacity behaves, with probability 1, as

$$N \log(1 + P),$$

which grows linearly in $N$, the number of receive antennas. Similarly, for a fixed $M$, $\lim_{N\to\infty} \frac{1}{N} HH^* = I_M$ with probability 1. Since $\det(I_N + (P/M)H^*H) = \det(I_M + (P/M)HH^*)$, the capacity behaves, with probability 1, as

$$M \log\left(1 + \frac{PN}{M}\right),$$

which increases almost linearly in $M$, the number of transmit antennas. Therefore, comparing with the single antenna capacity

$$\log(1 + P),$$

the capacity of multiple-antenna systems increases almost linearly in $\min\{M, N\}$. Multiple-antenna systems then give significant capacity improvement than single-antenna systems.

The capacity for the case when neither the transmitter nor the receiver knows the channel is still an open problem. Zheng and Tse [ZT02] have some results based on the block-fading channel model, which will be discussed in the next chapter.

## 1.5 Diversity

Another prominent advantage of multiple-antenna systems is that they provide better reliability in transmissions by using diversity techniques without increasing transmit power or sacrificing bandwidth. The basic idea of diversity is that, if two or more independent samples of a signal are sent and then fade in an uncorrelated manner, the probability that all the samples are simultaneously below a given level is much lower than the probability of any one sample being below that level. Thus, properly combining various samples greatly reduces the severity of fading and improves reliability of transmission. We give a very simple analysis below. For more details, please refer to [Rap02, Stu00, VY03].

The system equation for a single-antenna communication system is

$$x = \sqrt{\rho} s h + v,$$

where $h$ is the Rayleigh flat-fading channel coefficient. $\rho$ is the transmit power. $v$ is the noise at the receiver, which is Gaussian with zero-mean and unit-variance. $s$ satisfies the power constraint $\mathrm{E}\,|s|^2 = 1$. Therefore, the SNR at the receiver is $\rho|h|^2$. Since $h$ is Rayleigh distributed, $|h|^2$ is exponentially distributed with probability density function

$$p(x) = e^{-x}, \quad x > 0.$$

Thus, the probability that the receive SNR is less than a level $\epsilon$ is,

$$\mathrm{P}\left(\rho|h|^2 < \epsilon\right) = \mathrm{P}\left(|h|^2 < \frac{\epsilon}{\rho}\right) = \int_0^{\frac{\epsilon}{\rho}} e^{-x} dx = 1 - e^{-\frac{\epsilon}{\rho}}.$$

When the transmit power is high ($\rho \gg 1$),

$$\mathrm{P}\left(\rho|h|^2 < \epsilon\right) \approx \frac{\epsilon}{\rho},$$

which is inversely proportional to the transmit power. For a multiple-antenna system, with the same transmit power, the system equation is

$$\mathbf{x} = \sqrt{\rho}\mathbf{s}H + \mathbf{v},$$

where $\mathrm{E}\,\mathbf{s}\mathbf{s}^* = 1$. Further assume that the elements of $\mathbf{s}$ are iid, in which case $\mathrm{E}\,|s_i|^2 = 1/M$. Since $h_{ij}$ are independent, the expected SNR at the receiver is

$$\rho \mathrm{E}\,\mathbf{s}HH^*\mathbf{s}^* = \rho \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{N} \mathrm{E}\, s_i\overline{s_j}h_{ik}\overline{h_{jk}} = \rho \sum_{i=1}^{M} \mathrm{E}\,|s_i|^2 \sum_{k=1}^{N} |h_{ik}|^2 = \frac{\rho}{M}\sum_{i=1}^{M}\sum_{k=1}^{N}|h_{ik}|^2.$$

The probability that the SNR at the receiver is less than the level $\epsilon$ is then

$$\begin{aligned}
\mathrm{P}\left(\frac{\rho}{M}\sum_{i=1}^{M}\sum_{k=1}^{N}|h_{ik}|^2 < \epsilon\right) &= \mathrm{P}\left(\sum_{i=1}^{M}\sum_{k=1}^{N}|h_{ik}|^2 < \frac{\epsilon M}{\rho}\right) \\
&< \mathrm{P}\left(|h_{11}|^2 < \frac{cM}{\rho}, \cdots, |h_{MN}|^2 < \frac{\epsilon M}{\rho}\right) \\
&= \prod_{i=1,k=1}^{M,N} \mathrm{P}\left(|h|^2 < \frac{\epsilon M}{\rho}\right) \\
&= \left(1 - e^{-\frac{\epsilon M}{\rho}}\right)^{MN}.
\end{aligned}$$

When the transmit power is high ($\rho \gg 1$),

$$\mathrm{P}\left(\frac{\rho}{M}\sum_{i=1}^{M}\sum_{k=1}^{N}|h_{ik}|^2 < \epsilon\right) \lesssim \left(\frac{\epsilon M}{\rho}\right)^{MN},$$

which is inversely proportional to $\rho^{MN}$. Therefore, multiple-antenna systems have much lower error probability than single-antenna systems at high transmit power.

There are a lot of diversity techniques. According to the domain where diversity is introduced, they can be classified into *time diversity*, *frequency diversity* and *antenna diversity* (*space diversity*). Time diversity can be achieved by transmitting

identical messages in different time slots, which results in uncorrelated fading signals at the receiver. Frequency diversity can be achieved by using different frequencies to transmit the same message. The issue we are interested in is space diversity, which is typically implemented using multiple antennas at the transmitter or the receiver or both. The multiple antennas should be separated physically by a proper distance to obtain independent fading. Typically a separation of a few wavelengths is enough.

Depending on whether multiple antennas are used for transmission or reception, space diversity can be classified into two categories: *receive diversity* and *transmit diversity*. To achieve receive diversity, multiple antennas are used at the receiver to obtain independent copies of the transmitted signals. The replicas are properly combined to increase the overall receive SNR and mitigate fading. There are many combining methods, for example, selection combining, switching combining, maximum ratio combining, and equal gain combining. Transmit diversity is more difficult to implement than receive diversity due to the need for more signal processing at both the transmitter and the receiver. In addition, it is generally not easy for the transmitter to obtain information about the channel, which results in more difficulties in the system design.

Transmit diversity in multiple-antenna systems can be exploited by a coding scheme called *space-time coding*, which is a joint design of error-control coding, modulation, and transmit diversity. The idea of space-time coding is discussed in the next chapter.

# Chapter 2 Space-Time Block Codes

## 2.1 Block-Fading Model

Consider the wireless communication system given in Figure 1.1 in Section 1.2. We use *block-fading* model by assuming that the fading coefficients stay unchanged for $T$ consecutive transmissions, then jump to independent values for another $T$ transmissions and so on. This piecewise constant fading process mimics the approximate coherence interval of a continuously fading process. It is an accurate representation of many TDMA, frequency-hopping, and block-interleaved systems.

The system equation for a block of $T$ transmissions can be written as

$$X = \sqrt{\frac{\rho T}{M}} SH + W, \tag{2.1}$$

where

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1M} \\ \vdots & \ddots & \vdots \\ s_{T1} & \cdots & s_{TM} \end{bmatrix}, \qquad X = \begin{bmatrix} x_{11} & \cdots & x_{1M} \\ \vdots & \ddots & \vdots \\ x_{T1} & \cdots & x_{TM} \end{bmatrix},$$

$$H = \begin{bmatrix} h_{11} & \cdots & h_{1N} \\ \vdots & \ddots & \vdots \\ h_{M1} & \cdots & h_{MN} \end{bmatrix}, \qquad W = \begin{bmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{T1} & \cdots & x_{TN} \end{bmatrix}.$$

$S$ is the $T \times M$ transmitted signal matrix with $s_{tm}$ the signal sent by the $m$-th transmit

antenna at time $t$. The $t$-th row of $S$ indicates the row vector of the transmitted values from all the transmitters at time $t$ and the $m$-th column indicates the transmitted values of the $m$-th transmit antenna across the coherence interval. Therefore, the horizontal axis of $S$ indicates the spatial domain and the vertical axis of $S$ indicates the temporal domain. This is why $S$ is called a space-time code. In the design of $S$, redundancy is added in both the spatial and the temporal domains. $H$ is the $M \times N$ complex-valued matrix of propagation coefficients which remains constant during the coherent period $T$ and $h_{mn}$ is the propagation coefficient between the $m$-th transmit antenna and the $n$-th receive antenna. $h_{mn}$ have a zero-mean unit-variance circularly-symmetric complex Gaussian distribution $\mathcal{CN}(0,1)$ and are independent of each other. $V$ is the $T \times N$ noise matrix with $v_{tn}$ the noise at the $n$-th receive antenna at time $t$. The $v_{tn}$s are iid with $\mathcal{CN}(0,1)$ distribution. $X$ is the $T \times N$ matrix of the received signal with $x_{tn}$ the received value by the $n$-th receive antenna at time $t$. The $t$-th row of $X$ indicates the row vector of the received values at all the receivers at time $t$ and the $n$-th column indicates the received values of the $n$-th transmit antenna across the coherence interval.

If the transmitted signal is further normalized as

$$\frac{1}{M}\sum_{m=1}^{M} \mathrm{E}\,|s_{tm}|^2 = \frac{1}{T}, \quad \text{for} \quad t = 1, 2, ..., T, \tag{2.2}$$

which means that the average expected power over the $M$ transmit antennas is kept constant for each channel use, the expected received signal power at the $n$-th receive antenna and the $t$-th transmission is as follows.

$$\mathrm{E}\,\frac{\rho T}{M}\left|\sum_{m=1}^{M} s_{tm}h_{mn}\right|^2 = \frac{\rho T}{M}\sum_{m=1}^{M}\mathrm{E}\,|s_{tm}|^2\mathrm{E}\,|h_{mn}|^2 = \frac{\rho T}{M}\sum_{m=1}^{M}\mathrm{E}\,|s_{tm}|^2 = \rho.$$

The expected noise power at the $n$-th receive antenna and the $t$-th transmission is

$$\mathrm{E}\,|w_{tn}|^2 = 1.$$

Therefore $\rho$ represents the expected SNR at every receive antenna.

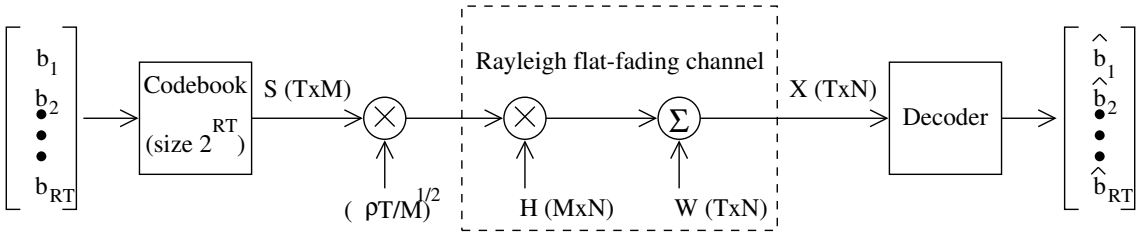


Figure 2.1: Space-time block coding scheme

The space-time coding scheme for multiple-antenna systems can be described by the diagram in Figure 2.1. For each block of transmissions, the transmitter selects a $T \times M$ matrix in the codebook according to the bit string $[b_1, b_2, \cdots, b_{2^{RT}}]$ and feeds columns of the matrix to its transmit antennas. The receiver decodes the $R$ bits based on its received signals which are attenuated by fading and corrupted by noise. The space-time block code design problem is to design the set, $\mathcal{C} = \{S_1, S_2, \cdots, S_{2^{RT}}\}$, of $2^{RT}$ transmission matrices in order to obtain low error rate.

## 2.2 Capacity for Block-Fading Model

In this section, the capacity of multiple-antenna systems using block-fading channel model is discussed. Note that the results in Section 1.4 are actually included in the results here since the system model used in Section 1.4 is a special case of the block-fading model used here with $T = 1$. As before, three cases are discussed: both the transmitter and the received know the channel, only the receiver knows the channel, and neither the transmitter nor the receiver knows the channel. The results are based

on [Tel99, MH99] and [ZT02].

When both the transmitter and the receiver know the channel, the capacity is the same using block-fading model or not since in this case, $H$ is deterministic. When only the receiver has perfect knowledge of the channel, it is proved in [MH99] that the average capacity per block of $T$ transmissions is

$$C = T \cdot \mathrm{E} \log\det\left(I_N + \frac{\rho}{M} H^* H\right),$$

where the expectation is over all possible channel realizations. Therefore, the average capacity per channel use is just

$$C = \mathrm{E} \log\det\left(I_N + \frac{\rho}{M} H^* H\right),$$

which is the same as the result in Section 1.4. Thus, the capacity increases almost linearly in $\min\{M, N\}$.

Now, we discuss the case that neither the transmitter nor the receiver knows the channel. It is proved in [MH99] that for any coherence interval $T$ and any number of receive antennas, the capacity with $M > T$ transmit antennas is the same as the capacity obtained with $M = T$ transmit antennas. That is, according to capacity, there is no point in having more transmit antennas than the length of the coherence interval. Therefore, in the following text, we always assume that $T \geq M$.

The structure of the signal that achieves capacity is also given in [MH99], which will be stated in Section 2.5. Although the structure of capacity-achieving signal is given in [MH99], the formula for the capacity is still an open problem. In [ZT02], the asymptotic capacity of Rayleigh block-fading channels at high SNR is computed. The capacity formula is given up to the constant term according to SNR. Here is the main result.

Define $G(T, M)$ as the set of all $M$ dimensional subspaces of $\mathbb{C}^T$. Let $K =$

$\min\{M, N\}$. If $T \geq K + N$, then at high SNR, the asymptotic optimal scheme is to use $K$ of the transmit antennas to send signal vectors with constant equal norm. The resulting capacity in bits per channel use is

$$C = K\left(1 - \frac{K}{T}\right)\log\rho + c_{k,n} + o(1), \tag{2.3}$$

where

$$c_{k,n} = \frac{1}{T}\log|G(T,M)| + M\left(1 - \frac{M}{T}\right)\log\frac{T}{\pi e} + \left(1 - \frac{M}{T}\right)\sum_{i=N-M+1}^{N} \mathrm{E}\,\log\chi^2_{2i},$$

and

$$|G(T,M)| = \frac{|S(T,M)|}{|S(M,M)|} = \frac{\prod_{i=T-M+1}^{T}\frac{2\pi^i}{(i-1)!}}{\prod_{i=1}^{M}\frac{2\pi^i}{(i-1)!}}$$

is the volume of the Grassmann manifold $G(T,M)$. $\chi^2_{2i}$ is a chi-square random variable (see [EHP93]) of dimension $2i$. Formula (2.3) indicates that the capacity is linear in $\min\{M, N, \frac{T}{2}\}$ at high SNR. This capacity expression also has a geometric interpretation as sphere packing in Grassmann manifold.

In [HM02], the probability density of the received signal when transmitting isotropically distributed unitary matrices is obtained in closed form, from which capacity of multiple-antenna systems can be computed. Also, simulated results in [HM02] show that at high SNR, the mutual information is maximized when $M = \min\{N, \frac{T}{2}\}$, whereas at low SNR, the mutual information is maximized by allocating all transmit power to a single antenna.

## 2.3 Performance Analysis of Systems with Known Channels

When the receiver knows the channel $H$, it is proved in [TSC98] and [HM00] that the maximum-likelihood (ML) decoding is

$$\arg \max_{i=1,2,\cdots,L} \mathrm{P}\left(X|S_i\right) = \min_{i=1,2,\cdots,L} \left\| X - \sqrt{\rho T/M} S_i H \right\|_F^2 .$$

Since the exact symbol error probability and bit error probability are very difficult to calculate, research efforts focus on the *pairwise error probability (PEP)* instead in order to get an idea of the error performance. The PEP of mistaking $S_i$ by $S_j$ is the probability that $S_j$ is decoded at the receiver while $S_i$ is transmitted. In [TSC98] and [HM00], it is proved that the PEP of mistaking $S_i$ and $S_j$, averaged over the channel distribution, has the following upper bound:

$$P_e \leq \det{}^{-N} \left[ I_M + \frac{\rho T}{4M} (S_i - S_j)^* (S_i - S_j) \right].$$

If $S_i - S_j$ is full rank, at high SNR ($\rho \gg 1$),

$$P_e \lesssim \det{}^{-N} (S_i - S_j)^* (S_i - S_j) \left( \frac{4M}{\rho T} \right)^{MN} . \tag{2.4}$$

We can see that the average PEP is inversely proportional to $SNR^{MN}$. Therefore, diversity $MN$ is obtained. The coding gain is $\det{}^{N}(S_i - S_j)^*(S_i - S_j)$. If $S_i - S_j$ is not full rank, the diversity is $\mathrm{rank}\,(S_i - S_j)N$.

Researchers also have worked on the exact PEP and other upper bounds of PEP, from which estimations of the bit error probability are made. For more, see [HM00, ZAS00, UG00, TB02].

## 2.4 Training-Based Schemes

In wireless communication systems, for the receiver to learn the channel, training are needed. Then, data information can be sent, and the ML decoding and performance analysis follow the discussions in the previous section. This scheme is called training-based scheme.

Training-based schemes are widely used in multiple-antenna wireless communications. The idea of training-based schemes is that when the channel changes slowly, the receiver can learn the channel information by having the transmitter send pilot signals known to the receiver. Training-based schemes dedicate part of the transmission matrix $S$ to be a known training signal from which $H$ can be learned. In particular, training-based schemes are composed of two phases: the training phase and the data-transmission phase. The following discussion is based on [HH03].

The system equation for the training phase is

$$X_t = \sqrt{\frac{\rho_t}{M}} S_t H + V_t,$$

where $S_t$ is the $T_t \times M$ complex matrix of training symbols sent over $T_t$ time samples and known to the receiver, $\rho_t$ is the SNR during the training phase, $X_t$ is the $T_t \times N$ complex received matrix, and $V_t$ is the noise matrix. $S_t$ is normalized as $\operatorname{tr} S_t S_t^* = MT_t$.

Similarly, the system equation for the data-transmission phase is

$$X_d = \sqrt{\frac{\rho_d}{M}} S_d H + V_d,$$

where $S_d$ is the $T_d \times M$ complex matrix of data symbols sent over $T_d = T - T_\tau$ time samples, $\rho_d$ is the SNR during the data-transmission phase, $X_d$ is the $T_d \times N$ complex received matrix, and $V_d$ is the noise matrix. $S_d$ is normalized as $\mathrm{E} \operatorname{tr} S_d S_d^* = MT_d$.

The normalization formula has an expectation because $S_d$ is random and unknown. Note that $\rho T = \rho_d T_d + \rho_t T_t$.

There are two general methods to estimate the channel: the ML (maximum-likelihood) and the LMMSE (linear minimum-mean-square-error) estimation whose channel estimations are given by

$$\hat{H} = \sqrt{\frac{M}{\rho_t}}(S_t^* S_t)^{-1} S_t^* X_t \quad \text{and} \quad \hat{H} = \sqrt{\frac{M}{\rho_t}} \left( \frac{M}{\rho_t} T_M + S_t^* S_t \right)^{-1} S_t^* X_t,$$

respectively.

In [HH03], an optimal training scheme that maximizes the lower bound of the capacity for MMSE estimation is given. There are three parameters to be optimized. The first one is the training data $S_t$. It is proved that the optimal solution is to choose the training signal as a multiple of a matrix with orthonormal columns. The second one is the length of the training interval. Setting $T_t = M$ is optimal for any $\rho$ and $T$. Finally, the third parameter is the optimal power allocation, which should satisfy the following,

$$\begin{cases} \rho_d < \rho < \rho_t & \text{if } T > 2M \\ \rho_d = \rho = \rho_t & \text{if } T = 2M \\ \rho_d > \rho > \rho_t & \text{if } T < 2M \end{cases} .$$

Combining the training-phase equation and the data-transmission-phase equation, the system equation can be written as

$$\begin{bmatrix} X_\tau \\ X_d \end{bmatrix} = \sqrt{\rho} \begin{bmatrix} I_M \\ S_d \end{bmatrix} H + \begin{bmatrix} V_\tau \\ V_d \end{bmatrix}. \tag{2.5}$$

Therefore, the transmitted signal is

$$S = \frac{1}{\sqrt{2}} \begin{bmatrix} I_M \\ S_d \end{bmatrix}.$$

If we further assume that the $(T - M) \times M$ data matrix $S_d$ is unitary, the unitary complement of $S$ can be easily seen to be

$$S^{\perp} = \frac{1}{\sqrt{2}} \begin{bmatrix} -S_d^* \\ I_{T-M} \end{bmatrix}. \tag{2.6}$$

If $S_d$ is not unitary, the matrix given in (2.6) is only the orthogonal complement of $S$.[1] Note that the unitary complement of $S$ may not exist in this case.

There are other training schemes according to other design criterions. For example, in [Mar99], it is shown that, under certain conditions, by choosing the number of transmit antennas to maximize the throughput in a wireless channel, one generally spends half the coherence interval training.

## 2.5 Unitary Space-Time Modulation

As discussed in the previous section, training-based scheme allocates part of the transmission interval and power to training, which causes both extra time delay and power consumption. For systems with multiple transmit and receive antennas, since there are $MN$ channels in total, to have a reliable estimation of the channels, considerably long training interval is needed. Also, when the channels change fast because of the movings of the transmitter, the receiver, or surrounding objects, training is not possible. In this section, a transmission scheme called *unitary space-time modulation*

[1] A $T \times (T - M)$ matrix $\tilde{A}$ is the orthogonal complement of a $T \times M$ matrix $A$ if and only if $\tilde{A}^* A = 0$.

*(USTM)* is discussed, which is suitable for trainsmissions in multiple-antenna systems when the channel is unknown to both the transmitter and the receiver without training. This scheme was proposed in [HM00].

### 2.5.1 Transmission Scheme

When the receiver does not know the channel, it is not clear how to design the signal set and decode. In [MH99], the capacity-achieving signal is given, and the main result is as follows in Theorem 2.1.

**Theorem 2.1 (Structure of capacity-achieving signal).** *[MH99] A capacity-achieving random signal matrix for (2.1) may be constructed as a product $S = VD$, where $V$ is a $T \times T$ isotropically distributed unitary matrix, and $D$ is an independent $T \times M$ real, nonnegative, diagonal matrix. Furthermore, for either $T \gg M$, or high SNR with $T > M$, $d_{11} = d_{22} = \cdots = d_{MM} = 1$ achieves capacity where $d_{ii}$ is the $i$-th diagonal entry of $D$.*

An isotropically distributed $T \times T$ unitary matrix has a probability density that is unchanged when the matrix is left or right multiplied by any deterministic unitary matrix. It is the uniform distribution on the space of unitary matrices. In a natural way, an isotropically distributed unitary matrix is the $T \times T$ counterpart of a complex scalar having unit magnitude and uniformly distributed phase. For more on the isotropic distribution, refer to [MH99] and [Ede89].

Motivated by this theorem, in [HM00], it is proposed to design the transmitted signal matrix $S$ as $S = \Phi \begin{bmatrix} I_M & \mathbf{0}_{T-M,M} \end{bmatrix}^t$ with $\Phi$ a $T \times T$ unitary matrix. That is, $S$ is designed as the first $M$ columns of a $T \times T$ unitary matrix. This is called unitary space-time modulation, and such an $S$ is called a $T \times M$ unitary matrix since its $M$ columns are orthonormal. In USTM, the transmitted signals are chosen from a constellation $\mathcal{V} = \{S_1, \ldots, S_L\}$ of $L = 2^{RT}$ (where $R$ is the transmission rate in bits

per channel use) $T \times M$ unitary matrices.

### 2.5.2 ML Decoding and Performance Analysis

It is proved in [HM00] that the ML decoding of USTM is

$$\hat{\ell} = \arg \max_{\ell=1,\ldots,L} \|X^* S_\ell\|_F^2 = \arg \min_{\ell=1,\ldots,L} \|X^* S_\ell^\perp\|_F^2. \tag{2.7}$$

With this ML decoding, the PEP of mistaking $S_i$ by $S_j$, averaged over the channel distribution, has the following Chernoff upper bound

$$Pe \leq \frac{1}{2} \prod_{m=1}^{M} \left[ \frac{1}{1 + \frac{(\rho T/M)^2 (1-d_m^2)}{4(1+\rho T/M)}} \right]^N,$$

where $1 \geq d_1 \geq ... \geq d_M \geq 0$ are the singular values of the $M \times M$ matrix $S_j^* S_i$. The formula shows that the PEP behaves as $|\det(S_j^* S_i)|^{-2N}$. Therefore, many design schemes have focused on finding a constellation that maximizes $\min_{j \neq i} |\det(S_j^* S_i)|$, for example, [HMR$^+$00, ARU01, TK02]. Since $L$ can be quite large, this calls into question the feasibility of computing and using this performance criterion. The large number of possible signals also rules out the possibility of decoding via an exhaustive search. To design constellations that are huge, effective, and yet still simple, so that they can be decoded in real time, some structure needs to be introduced to the signal set. In Chapter 3, it is shown how Cayley transform can be used for this purpose.

## 2.6 Differential Unitary Space-Time Modulation

### 2.6.1 Transmission Scheme

Another way to communicate with unknown channel information is differential unitary space-time modulation, which can be seen as a natural higher-dimensional exten-

sion of the standard differential phase-shift keying (DPSK) commonly used in signal-antenna unknown-channel systems (see [HS00, Hug00a]). In differential USTM, the channel is used in blocks of $M$ transmissions, which implies that the transmitted signal, $S$, is an $M \times M$ unitary matrix. The system equation at the $\tau$-th block is

$$X_\tau = \sqrt{\rho} S_\tau H_\tau + V_\tau, \tag{2.8}$$

where $S_\tau$ is $M \times M$, $H_\tau$, $X_\tau$, and $V_\tau$ are $M \times N$. Similar to DPSK, the transmitted signal $S_\tau$ at the $\tau$-th block equals the product of a unitary data matrix, $U_{z_\tau}, z_\tau \in 0, ..., L-1$, taken from our signal set $\mathcal{C}$ and the previously transmitted matrix, $S_{\tau-1}$. In other words,

$$S_\tau = U_{z_\tau} S_{\tau-1} \tag{2.9}$$

with $S_0 = I_M$. To assure that the transmitted signal will not vanish or blow up to infinity, $U_{z_\tau}$ must be unitary. Since the channel is used $M$ times, the corresponding transmission rate is $R = \frac{1}{M} \log_2 L$, where $L$ is the cardinality of $\mathcal{C}$. If the propagation environment keeps approximately constant for $2M$ consecutive channel uses, that is, $H_\tau \approx H_{\tau-1}$, then from the system equation in (2.8),

$$X_\tau = \sqrt{\rho} U_{z_\tau} S_{\tau-1} H_{\tau-1} + V_\tau = U_{z_\tau}(X_\tau - V_{\tau-1}) + V_\tau = U_{z_\tau} X_\tau + V_\tau - U_{z_\tau} V_{\tau-1}.$$

Therefore, the following fundamental differential receiver equation is obtained [HH02a],

$$X_\tau = U_{z_\tau} X_{\tau-1} + W'_\tau, \tag{2.10}$$

where

$$W'_\tau = V_\tau - U_{z_\tau} V_{\tau-1}. \tag{2.11}$$

The channel matrix $H$ does not appear in (2.10). This implies that, as long as the channel is approximately constant for $2M$ channel uses, differential transmission permits decoding at the receiver without knowing the channel information.

### 2.6.2 ML Decoding and Performance Analysis

Since $U_{z_\tau}$ is unitary, the additive noise term in (2.11) is statistically independent of $U_{z_\tau}$ and has independent complex Gaussian entries. Therefore, the maximum-likelihood decoding of $z_\tau$ can be written as

$$\hat{z}_\tau = \arg \max_{l=0,\ldots,L-1} \|X_\tau - U_l X_{\tau-1}\|_F. \tag{2.12}$$

It is shown in [HS00, Hug00a] that, at high SNR, the average PEP of transmitting $U_i$ and erroneously decoding $U_j$ has the upper bound

$$Pe \lesssim \frac{1}{2}\left(\frac{8}{\rho}\right)^{MN} \frac{1}{|\det(U_i - U_j)|^{2N}},$$

which is inversely proportional to $|\det(U_i - U_j)|^{2N}$. Therefore the quality of the code is measured by its *diversity product* defined as

$$\xi_{\mathcal{C}} = \frac{1}{2} \min_{0 \le i < j \le L} |\det(U_i - U_j)|^{\frac{1}{M}}. \tag{2.13}$$

From the definition, the diversity product is always non-negative. A code is said to be *fully diverse* or have *full diversity* if its diversity product is not zero. Fully diverse physically means that the receiver will always decode correctly if there is no noise.

The power $\frac{1}{M}$ and the coefficient $\frac{1}{2}$ in formula (2.13) are used for normalization. With this normalization, the diversity product of any set of unitary matrices is between 0 and 1. From the definition of diversity product, it is easy to see that the set with the largest diversity product is $\{I_M, -I_M\}$ since it has the minimum number of

elements with the maximum determinant difference. Since

$$|\det(I_M - (-I_M))| = \det 2I_M = 2^M,$$

to normalize the diversity product of the set to 1, (2.13) is obtained. The differential unitary space-time code design problem is thus the following: Let $M$ be the number of transmitter antennas, and $R$ be the transmission rate. Construct a set $\mathcal{C}$ of $L = 2^{MR}$ $M \times M$ unitary matrices such that its diversity product, as defined in (2.13), is as large as possible.

Many design schemes [HS00, Hug00a, SHHS01, Hug00b, GD03, DTB02] have focused on finding a constellation $\mathcal{C} = \{U_0, ..., U_L\}$ of $L = 2^{MR}$ unitary $M \times M$ matrices that maximizes $\xi_{\mathcal{C}}$ defined in (2.13). Similar to USTM, in general, the number of unitary $M \times M$ matrices in $\mathcal{C}$ can be quite large. This huge number of signals calls into question the feasibility of computing $\xi_{\mathcal{C}}$ and also rules out the possibility of decoding via an exhaustive search. To design constellations that are huge, effective, and yet still simple so that they can be decoded in real time, some structure should be imposed upon the signal set. In Chapter 4 of this thesis, the idea of design differential unitary space-time code with group structure is introduced. In Chapters 5 and 6, our work on the designs of $4 \times 4$ and $3 \times 3$ differential unitary space-time codes based on Lie groups $Sp(2)$ and $SU(3)$ are explained in detail. The codes proposed not only have great performance at high data rates but also lend themselves to a fast decoding algorithm using sphere decoding.

## 2.7 Alamouti's $2 \times 2$ Orthogonal Design and Its Generalizations

The Alamouti's scheme [Ala98] is historically the first and the most well-known space-time code which provides full transmit diversity for systems with two transmit antenna. It is also well-known for its simple structure and fast ML decoding.

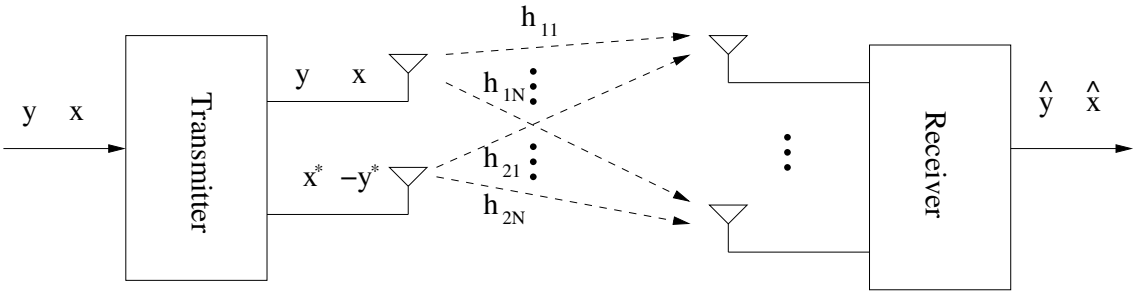


Figure 2.2: Transmission of Alamouti's scheme

The transmission scheme is shown in Figure 2.2. The channel is used in blocks of two transmissions. During the first transmission period, two signals are transmitted simultaneously from the two antennas. The first antenna transmits signal $x$ and the second antenna transmits signal $-y^*$. During the second transmission period, the first antenna transmits signal $y$ and the second antenna transmits signal $x^*$. Therefore, the transmitted signal matrix $S$ is

$$S = \begin{bmatrix} x & y \\ -y^* & x^* \end{bmatrix}.$$

It is easy to see that the two columns/rows of $S$ are orthogonal. This design scheme is also called the $2 \times 2$ orthogonal design. Further more, with the power constraint $|x|^2 + |y|^2 = 1$, $S$ is actually a unitary matrix with determinant 1. In general, the

Alamouti's code can be written as

$$\mathcal{C} = \left\{ \frac{1}{\sqrt{|x|^2+|y|^2}} \begin{bmatrix} x & y \\ -y^* & x^* \end{bmatrix} \middle| x \in \mathcal{S}_1, y \in \mathcal{S}_2 \right\},$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ are two sets in $\mathbb{C}$. If $\mathcal{S}_1 = \mathcal{S}_2 = \mathbb{C}$, the code is exactly the Lie group $SU(2)$. To obtain finite codes, $\mathcal{S}_1$ and $\mathcal{S}_2$ should be chosen as finite sets, and therefore the codes obtained are finite samplings of the infinite Lie group.

The Alamouti's scheme not only has the properties of simple structure and full rate (it's rate is 1 symbol per channel use), it also has an ML decoding method with very low complexity. With simple algebra, the ML decoding of Alamouti's scheme is equivalent to

$$\arg\min_x \left| x - \sqrt{\frac{1}{\rho}} \sum_{i=1}^{N} (x_{1i} h_{1i}^* + x_{2i}^* h_{2i}) \right| \text{ and } \arg\min_y \left| y - \sqrt{\frac{1}{\rho}} \sum_{i=1}^{N} (x_{1i} h_{2i} - x_{2i}^* h_{1i}^*) \right|,$$

which actually shows that the decodings of the two signals $x$ and $y$ can be decoupled. Therefore, the complexity of this decoding is very small. This is one of the most important features of Alamouti's scheme. For the unknown channel case, this transmission scheme can also be used in differential USTM, whose decoding is very similar to the one shown above and thus can be done very fast.

We now turn our attention to the performance of this space-time code. For any two non-identical signal matrices in the codes,

$$S_1 = \begin{bmatrix} x_1 & y_1 \\ -y_1^* & x_1^* \end{bmatrix} \text{ and } S_2 = \begin{bmatrix} x_2 & y_2 \\ -y_2^* & x_2^* \end{bmatrix},$$

we have

$$\det(S_1 - S_2) = \det \begin{bmatrix} x_1 - y_1 & x_2 - y_1 \\ -(x_2 - y_2)^* & (x_1 - y_1)^* \end{bmatrix} = |x_1 - x_2|^2 + |y_1 - y_2|^2,$$

which is always positive since either $x_1 \neq x_2$ or $y_1 \neq y_2$. Therefore, the rank of $S_1 - S_2$ is 2, which is the number of transmit antennas. Full transmit diversity is obtained. The diversity product of the code is $\min_{(x_1,y_1)\neq(y_1,y_2)} |x_1 - x_2|^2 + |y_1 - y_2|^2$. If $x_i$ and $y_i$ are chosen from the $P$-PSK signal set $\{1, e^{2\pi j \frac{1}{P}}, \cdots, e^{2\pi j \frac{P-1}{P}}\}$, it is shown in [SHHS01] that the diversity product of the code is $\frac{\sin(\pi/P)}{\sqrt{2}}$.

Because of its great features, much attention has been dedicated to finding methods to generalize Alamouti's scheme for higher dimensions. A real orthogonal design of size $n$ is an $n \times n$ orthogonal matrix whose entries are the indeterminants $\pm x_1, \cdots, \pm x_n$. The existence problem for real orthogonal designs is known as the Hurwitz-Radon problem [GS79] and has been solved by Radon. In fact, real orthogonal designs exist only for $n = 2, 4, 8$.

A complex orthogonal design of size $n$ is an $n \times n$ unitary matrix whose entries are the indeterminants $\pm x_1, \cdots, \pm x_n$ and $\pm x_1^*, \cdots, \pm x_n^*$. In [TJC99], Tarokh, Jafarkhani and Calderbank proved that complex orthogonal design only exists for the two-dimensional case, that is, the Alamouti's scheme is unique. They then generalized the complex orthogonal design problem to non-square case also. They proved the existence of complex orthogonal designs with rate no more than $\frac{1}{2}$ and gave a $4 \times 3$ complex orthogonal design with rate $\frac{3}{4}$. In [WX03], Wang and Xia proved that the rate of complex orthogonal designs is upper-bounded by $\frac{3}{4}$ for systems with more than two transmit antennas and the rate of generalized orthogonal designs (non-square case) is upper-bounded by $\frac{4}{5}$. The restricted (generalized) complex orthogonal design is also discussed in [WX03].

## 2.8 Sphere Decoding and Complex Sphere Decoding

To accomplish transmissions in real time, fast decoding algorithm at the receiver is required. A natural way to decode is exhaustive search, which finds the optimal decoding signal by searching over all possible signals. However, this algorithm has a complexity that is exponential in both the transmission rate and the dimension. Therefore, it may take long time and cannot fulfill the real time requirement especially when the rate and dimension is high. There are other decoding algorithms, such as nulling-and-canceling [Fos96], whose complexity is polynomial in rate and dimension, however, they only provide approximate solutions. In this section, an algorithm called sphere decoding is introduced which not only provides the exact ML solutions for many communication systems but also has a polynomial complexity for almost all rates.

*Sphere decoding* algorithm was first proposed to find vectors of shortest length in a given lattice [Poh81], and has been tailored to solve the so-called *integer least-square problem*:

$$\min_{\mathbf{s}\in\mathbb{Z}^n} \|\mathbf{x} - H\mathbf{s}\|_F^2,$$

where $\mathbf{x} \in R^{m\times 1}, H \in R^{m\times n}$ and $\mathbb{Z}^n$ denotes the $m$-dimensional integer lattice, i.e., $\mathbf{s}$ is an $n$-dimensional vector with integer entries. The geometric interpretation of the integer least-square problem is this: as the entries of $\mathbf{s}$ run over $\mathbb{Z}$, $\mathbf{s}$ spans the "rectangular" $n$-dimensional lattice. For any $H$, which we call the *lattice-generating matrix*, $H\mathbf{s}$ spans a "skewed" lattice. Therefore, given the skewed lattice and a vector $\mathbf{x}$, the integer least-square problem is to find the "closest" lattice point (in Euclidean sense) to $\mathbf{x}$. We can generalize this problem by making $\mathbf{s} \in \mathcal{S}^n$ where $\mathcal{S}$ is any discrete

set.

Many communication decoding problems can be formulated into this problem with little modification since many digital communication problems have a lattice formulation [VB93, HH02b, DCB00]. The system equation is often

$$\mathbf{x} = H\mathbf{s} + \mathbf{v},$$

where $\mathbf{s} \in \mathbb{R}^{n\times 1}$ is the transmit signal, $\mathbf{x} \in \mathbb{R}^{m\times 1}$ is the received signal, $H \in \mathbb{R}^{m\times n}$ is the channel matrix and $\mathbf{v} \in \mathbb{R}^{m\times 1}$ is the channel noise. Note that here all the matrix and vectors are real. The decoding problem is often

$$\min_{\mathbf{s}\in\mathcal{S}^n} \|\mathbf{x} - H\mathbf{s}\|_F^2. \tag{2.14}$$

To obtain the exact solution to this problem, as mentioned before, an obvious method is exhaustive search, which searches over all $\mathbf{s} \in \mathcal{S}^n$ and finds the one with the minimum $\|\mathbf{x} - H\mathbf{s}\|_F^2$. However, this method is not feasible when the number of possible signals is infinite. Even when the cardinality of the lattice is finite, the complexity of exhaustive search is usually very high especially when the cardinality of the lattice is huge. It often increases exponentially with the number of antennas and transmission rate. Sphere decoding gives the exact solution to the problem with a much lower complexity. In [HVa], it is shown that sphere decoding has an average complexity that is cubic in the transmission rate and number of antennas for almost all practical SNRs and rates. It is a convenient fast ML decoding algorithm.

The idea of sphere decoding is to search over only lattice points that lie in a certain sphere of radius $d$ around the given vector $\mathbf{x}$. Clearly, the closest lattice point inside the sphere is the closest point in the whole lattice. The main problem is how to find the vectors in the sphere.

A lattice point $H\mathbf{s}$ is in a sphere of radius $d$ around $\mathbf{x}$ if and only if

$$\|\mathbf{x} - H\mathbf{s}\|_F^2 \leq d^2. \tag{2.15}$$

Consider the Cholesky or QR factorization of $H$: $H = Q \begin{bmatrix} R \\ \mathbf{0}_{m-n,n} \end{bmatrix}$, where $R$ is an $n \times n$ upper triangular matrix with positive diagonal entries and $Q$ is an $m \times m$ orthogonal matrix.[2] If we decompose $Q$ as $[Q_1 \ Q_2]$, where $Q_1$ is the first $n$ columns of $Q$, (2.15) is equivalent to

$$\|Q_1^*\mathbf{x} - R\mathbf{s}\|_F^2 + \|Q_2^*\mathbf{x}\|_F^2 \leq d^2.$$

Define $d_n^2 = d^2 - \|Q_2^*\mathbf{x}\|_F^2$ and $\mathbf{y} = Q_1^*\mathbf{x}$. The sphere becomes

$$\left\| \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,m} \\ 0 & r_{1,1} & \cdots & r_{2,n} \\ \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{n,n} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} \right\|_F^2 \leq d_n^2, \tag{2.16}$$

where $y_i$ indicates the $i$-th entry of $\mathbf{y}$. Note that the $n$-th row of the vector in the left hand side depends only on $s_n$, the $(n-1)$-th row depends only on $s_n$ and $s_{n-1}$, and so on. Looking at only the $n$-th row of (2.16), a necessary condition for (2.16) to hold is $(x_n - r_{n,n}s_n)^2 \leq d_n^2$ which is equivalent to

$$\left\lceil \frac{-d_n + x_n}{r_{n,n}} \right\rceil \leq s_n \leq \left\lfloor \frac{d_n + x_n}{r_{n,n}} \right\rfloor. \tag{2.17}$$

Therefore, the interval for $s_n$ is obtained. For each $s_n$ in the interval, define $d_{n-1}^2 = d_n^2 - (x_n - r_{n,n}s_n)^2$. A stronger necessary condition can be found by looking at the

[2]Here we only discuss the $n \leq m$ case. The $n > m$ case can be seen in [HVb].

$(n-1)$-th row of (2.16): $|x_{n-1} - r_{n-1,n-1}s_{n-1} - r_{n-1,n}s_n|^2 \leq d_{n-1}^2$. Therefore, for each $s_n$ in (2.17), we get an interval for $s_{n-1}$:

$$\left\lceil \frac{-d_{n-1} + x_{n-1} - r_{n-1,n}s_n}{r_{n-1,n-1}} \right\rceil \leq s_{n-1} \leq \left\lfloor \frac{d_{n-1} + x_{n-1} - r_{n-1,n}s_n}{r_{n-1,n-1}} \right\rfloor . \tag{2.18}$$

Continue with this procedure till the interval of $s_1$ for every possible values of $s_n, \cdots, s_2$ is obtained. Thus, all possible points in the sphere (2.15) are found. A flow chart of sphere decoding can be found in [DAML00] and pseudo code can be found in [HVa].

The selection of the search radius in sphere decoding is crucial to the complexity. If the radius is too large, there are too many points in the sphere, and the complexity is high. If the selected radius is too small, it is very probable that there exists no point in the sphere. In [VB93], it is proposed to use the *covering radius* of the lattice. The covering radius is defined as the radius of the spheres centered at the lattice points that cover the whole space in the most economical way. However, the calculation of the covering radius is normally difficult. In [HVa], the authors proposed to choose the initial radius such that the probability of having the correct point in the sphere is 0.9, then increase the radius gradually if there is no point in the sphere. In our simulations, we use this method. Other radius-choosing methods can be found in [DCB00, DAML00]. There are also publications on methods that can further reduce the complexity of sphere decoding, interested readers can refer to [GH03, AVZ02, Art04b, Art04a] .

The sphere decoding algorithm described above applies to real systems when $\mathbf{s}$ is chosen from a real lattice. Therefore, the algorithm can be applied to complex systems when the system equation can be rewritten as linear equations of unknowns with twice the dimension by separating the real and imaginary parts of $\mathbf{x}, H$ and $\mathbf{s}$. Fortunately, this is true for many space-time coding systems ([HH02b, HH02a]). In particular, real sphere decoding is used in the decoding of our Cayley unitary space-

time codes in Chapter 3, the $Sp(2)$ differential unitary space-time codes in Chapter 5, and also the distributed space-time codes in Chapter 7.

Based on real sphere decoding, Hochwald generalized it to the complex case which is more convenient to be applied in wireless communication systems using PSK signals [HtB03]. The main idea is as follows.

The procedure follows all the steps of real sphere decoding. First, use the Cholesky factorization $H = QR$ where $Q$ is an $m \times m$ unitary matrix and $R$ is an upper triangle matrix with positive diagonal entries. Note that generally the off-diagonal entries of $R$ are complex, and $\mathbf{x}, H, \mathbf{s}$ are all complex. The search sphere is the same as in (2.16). As mentioned before, by looking at the last entry of the vector in the left hand side of (2.16), a necessary condition is $|y_n - r_{n,n}s_n|^2 \leq r^2$ or equivalently,

$$|s_n - y_n/r_{n,n}|^2 \leq r^2/r_{n,n}^2.$$

This inequality limits the search to points of the constellation contained in a complex disk of radius $r/r_{n,n}$ centered at $y_n/r_{n,n}$. These points are easily found when the constellation forms a complex circle (as in PSK).

Let $s_n = r_c e^{2\pi j\theta_n}$, where $r_c$ is a positive constant and $\theta_n \in \{0, 2\pi/P, \cdots, 2\pi(P-1)/P\}$. That is, $s_n$ is a $P$-PSK signal. Denote $y_n/r_{n,n}$ as $\hat{r}_c e^{2\pi j\hat{\theta}_n}$ and define $d_n^2 = r^2/r_{n,n}^2$. Then the condition becomes

$$r_c^2 + \hat{r}_c^2 - 2r_c\hat{r}_c\cos(\theta_n - \hat{\theta}_n) \leq r^2/r_{n,n}^2, \tag{2.19}$$

which yields

$$\cos(\theta_n - \hat{\theta}_n) \geq \frac{1}{2r_c\hat{r}_c}(r_c^2 + \hat{r}_c^2 - r^2/r_{n,n}^2).$$

If the right-hand side of the above is greater than 1, the search disk does not contain any point of the PSK constellation. If the value is less than $-1$, then the search disk

includes the entire constellation. Otherwise, the range of the possible angle for $s_n$ is

$$\left\lceil \hat{\theta}_n - \frac{P}{2\pi}\cos^{-1}\frac{(r_c^2+\hat{r}_c^2-r^2/r_{n,n}^2)}{2r_c\hat{r}_c} \right\rceil \le \theta_n \le \left\lfloor \hat{\theta}_n + \frac{P}{2\pi}\cos^{-1}\frac{(r_c^2+\hat{r}_c^2-r^2/r_{n,n}^2)}{2r_c\hat{r}_c} \right\rfloor \quad (2.20)$$

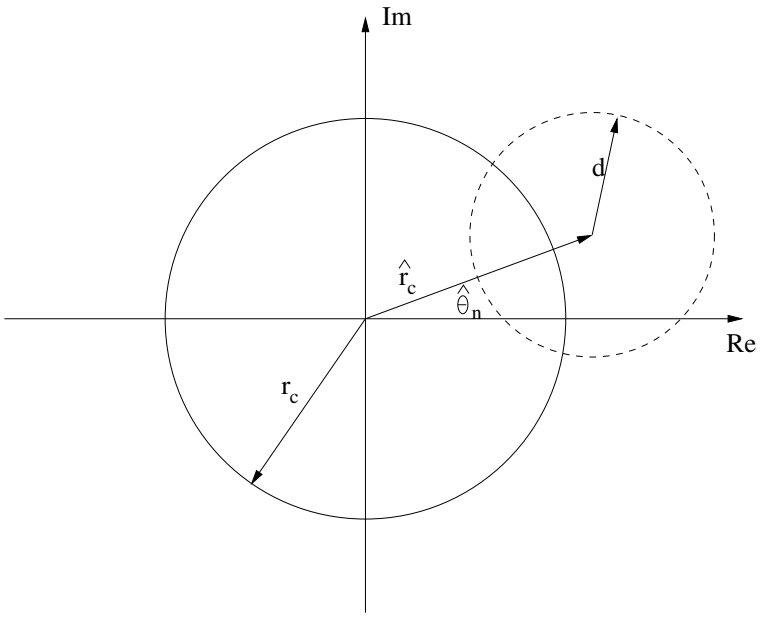


Figure 2.3: Interval searching in complex sphere decoding

This can be easily seen in Figure 2.3. The sphere given in (2.19) is the area bounded by the dashed circle. The values that $s_n$ can take spread on the solid circle uniformally. Note that

$$\frac{1}{2r_c\hat{r}_c}(r_c^2+\hat{r}_c^2-r^2/r_{n,n}^2) > 1 \Leftrightarrow |r_c-\hat{r}_c| > d \Leftrightarrow r_c > \hat{r}_c + d \text{ or } \hat{r}_c > r_c + d.$$

If $r_c > \hat{r}_c + d$, then the dashed circle is inside the solid circle. If $\hat{r}_c > r_c + d$, the two circles are disjoint. Therefore, if either happens, there is no possible $s_n$ in the sphere given in (2.19). If $\frac{1}{2r_c\hat{r}_c}(r_c^2+\hat{r}_c^2-r^2/r_{n,n}^2) < -1$, then the solid circle is contained in the dashed circle, which means that all the PSK signals are in the sphere. Otherwise, the solid circle has an arc that is contained in the sphere, and possible angles are given by (2.20).

Therefore, an interval for $s_n$'s angle, or equivalently, the set of values that $s_n$ can take on is obtained. For any chosen $s_n$ in the set, the set of possible values of $s_{n-1}$ can be found by similar analysis. By continuing with this procedure till the set of possible values of $s_1$ is found, all points in the complex disk are obtained.

## 2.9 Discussion

The results in Sections 2.2-2.6 are based on the assumption that the fading coefficients between pairs of transmit and receiver antennas are frequency non-selective and independent of each other. In this section, situations in which these assumptions are not valid are discussed.

In practice, channels may be correlated especially when the antennas are not sufficiently separated. The correlated fading models are proposed in [ECS+98, SFGK00]. The effects of fading correlation and channel degeneration (known as the keyhole effect) on the MIMO channel capacity have been addressed in [SFGK00, CTK02, CFG02], in which it is shown that channel correlation and degeneration actually degrade the capacity of multiple-antenna systems. Channel correlation can be mitigated using precoding, equalization and other schemes. For more on these issues, refer to [ZG03, KS04, SS03, HS02a, PL03].

In wideband systems,[3] transmitted signals experience frequency-selective fadings, which causes *inter-symbol interference (ISI)*. It is proved in [GL00] that the coding gain of the system is reduced, and it is reported that at high SNR, there exists an irreducible error rate floor. A conventional way to mitigate ISI is to use an equalizer at the receiver ([CC99, AD01]). Equalizers mitigate ISI and convert frequency-selective channels to flat-fading channels. Then, space-time codes designed for flat-fading channels can be applied ([LGZM01]). However, this approach results

[3] If the transmitted signal bandwidth is greater than the channel coherence bandwidth, the communication system is called a wideband system.

in high complexity at the receiver. An alternative approach is to use *orthogonal frequency division multiplexing (OFDM)* modulation. The idea of OFDM can be found in [BS99]. In OFDM, the entire channel is divided into many narrow parallel sub-channels with orthogonal frequencies. In every sub-channel, the fading can be regarded as frequency non-selective. There are many papers on space-time coded OFDM, for example [ATNS98, LW00, LSA98, BGP00].

Space-time coding is also combined with error-correcting schemes to improve coding gain. Space-time trellis codes were first proposed by Tarokh in [TSC98], and after that, they have been widely exploited ([TC01, CYV01, JS03]). The combination of space-time coding with trellis-coded modulation has also been widely analyzed ([BBH00, FVY01, BD01, GL02, TC01, JS03]). Other research investigated the combinations of space-time coding with convolutional codes and turbo codes ([Ari00, SG01, SD01, LFT01, LLC02]). The combinations of these schemes increases the performance of the system, however, the decoding complexity is very high and the performance analysis is very difficult.

## 2.10 Contributions of This Thesis

Contributions of this thesis are mainly on the design of space-time codes for multiple-antenna systems and their implementation in wireless networks. It can be divided into three parts.

In part one, unitary space-time codes are designed for systems with no channel information at both the transmitter and the receiver using Cayley transform. Cayley transform provides an one-to-one mapping from the space of (skew) Hermitian matrices to the space of unitary matrices. Based on the linearity of the space of Hermitian matrices, the transmitted data is first broken into sub-streams $\alpha_1, \cdots, \alpha_Q$, then linearly encoded into the $T \times T$ Hermitian-matrix space. Then a set of $T \times T$

unitary matrices is obtained by applying Cayley transform to the encoded Hermitian matrices. We show that by appropriate constraints on the Hermitian matrices and ignoring the data dependence of the additive noises, $\alpha_1, \cdots, \alpha_Q$ appears linearly at the receiver. Therefore, linear decoding algorithms such as sphere decoding and nulling-and-canceling can be used with polynomial complexity. Our Cayley codes have a similar structure as training-based schemes under transformations.

Cayley codes do not require channel knowledge at either the transmitter or the receiver, are simple to encode and decode, and can be applied to any combination of transmit and receive antennas. They are designed with a probabilistic criterion: they maximize the expected log-determinant of differences between matrix pairs.

The recipe for designing Cayley unitary space-time codes for any combination of transmit and receive antennas and coherence intervals is given, and also simulation examples are presented, which compare our Cayley codes with optimized training-based space-time codes and uncoded training-based schemes for different system settings. Our simulation results are preliminary. They indicate that Cayley codes generated with this recipe only slightly underperform optimized training-based schemes using orthogonal designs and/or linear dispersion codes. However, they are clearly superior to uncoded training-based space-time schemes. Further optimization on basis matrices of Cayley codes is necessary for a complete comparison of Cayley codes with training-based schemes.

The second part of our contributions is the design of unitary space-time codes based on Lie groups for the differential transmission scheme. The work can be regarded as extensions of [HK00]. In Chapter 5, we work on the symplectic group $Sp(n)$ which has dimension $n(2n+1)$ and rank $n$. We first give a parameterization of $Sp(n)$ and then design differential unitary space-time codes which are subsets of $Sp(2)$ by sampling the parameters appropriately. Necessary and sufficient conditions for full diversity of the codes are given. The designed constellations are suitable for

systems with four transmit antennas and any number of receive antennas. The special symplectic structure of the codes lends themselves to linear-algebraic decoding, such as sphere decoding. Simulation results show that they have better performance than the $2 \times 2$ and $4 \times 4$ complex orthogonal designs, group-based diagonal codes, and differential Cayley codes at high SNR. Although they slightly underperform the $k_{1,1,-1}$ finite-group code and the carefully designed non-group code, they do not need exhaustive search (of exponentially growing size) required by such code and therefore are far superior in term of decoding complexity.

In Chapter 6, we keep working on the idea of differential unitary space-time code design based on Lie groups with rank 2 and analyze the special unitary Lie group $SU(3)$, which has dimension 8 and rank 2. The group is not fixed-point-free, but we describe a method to design fully-diverse codes which are subsets of the group. Furthermore, motivated by the structure of the $SU(3)$ codes, we propose a simpler code called the AB code. Both codes are suitable for systems with three transmit antennas. Necessary conditions for full diversity of both codes are given and our conjecture is that they are also sufficient conditions. The codes have simple formulas from which their diversity products can be calculated in a fast way. A fast maximum-likelihood decoding algorithm for AB codes based on complex sphere decoding is given, by which decoding can be done with a complexity that is polynomial in the rate and dimension. Simulation results show that $SU(3)$ codes and AB codes perform as well as finite group-based codes at low rates. At high rates, performance of $SU(3)$ and AB codes is much better than that of finite group-based codes and about the same as that of the carefully designed non-group codes. The AB codes are, in addition, far superior in terms of decoding complexity as exhaustive search (of exponentially growing size) is required in decoding finite group-based and non-group codes. Our work on $Sp(2)$ and $SU(3)$ show the promise of studying constellations inspired by group-theoretical considerations.

The last contribution is on the application of space-time codes in wireless networks, or what we call the *distributed space-time coding*. We propose the use of linear dispersion space-time codes in wireless relay networks with a two-step strategy. We assume that the transmitter and relay nodes do not know the channel realizations but only their statistical distribution. ML decoding and average PEP at the receiver are analyzed. The main result is that diversity, $\min\{T, R\}\left(1 - \frac{\log\log P}{\log P}\right)$, is obtained, where $R$ is the number of relay nodes and $P$ is the average total power consumed in the network. This result indicates that when $T \geq R$ and the average total transmit power is high, relay networks achieve almost the same diversity as multiple-antenna systems with $R$ transmit antennas. This result is also supported by simulations. We further show that with $R = T$, the leading order term in the PEP of wireless relay networks behaves as $\frac{1}{2}\frac{1}{|\det(S_i - S_j)|^2}\left(\frac{8\log P}{P}\right)^R$, which compared to $\frac{1}{2}\frac{1}{|\det(S_i - S_j)|^2}\left(\frac{4}{P}\right)^R$, the PEP of multiple-antenna systems, shows the loss of performance due to the facts that space-time codes relay networks are implemented distributively and the relay nodes have no knowledge of the transmitted symbols. We also observe that the high SNR coding gain, $|\det(S_i - S_j)|^{-2}$, of relay networks is the same as what arises in multiple-antenna systems. The same is true at low SNR where a trace condition comes up.

# Chapter 3 Cayley Unitary Space-Time Codes

## 3.1 Introduction

As discussed in Chapters 1 and 2, multiple transmit and/or receive antennas promise high data rates on wireless channels with multi-path fading [Fos96, Tel99]. Many proposed schemes that achieve these high rates require the propagation environment or channel to be known to the receiver (see, e.g., [Fos96, Ala98, TSC98, HH02b] and the references therein). In practice, knowledge of the channel is often obtained via training: known signals are periodically transmitted for the receiver to learn the channel, and the channel parameters are tracked in between the transmission of the training signals. However, it is not always feasible or advantageous to use training-based schemes, especially when many antennas are used or either end of the link is moving so fast that the channel is changing very rapidly [Mar99, HH03].

Hence, there is much interest in space-time transmission schemes that do not require either the transmitter or receiver to know the channel. Information-theoretic calculations with a multi-antenna channel that changes in a block-fading manner first appeared in [MH99]. Based on these calculations, USTM was proposed in [HM00], in which the transmitted signals form a unitary matrix. Further information-theoretic calculations in [ZT02] and [HM02] show that, at high SNR, USTM schemes are capable of achieving full channel capacity. Furthermore, in [HMH01], it is shown that all these can be done over a single coherence interval, provided the coherence interval and number of transmit antennas are sufficiently large—a phenomenon referred to as

*autocoding*.

While all these are well recognized, it is not clear how to design a constellation of non-square USTM matrices, that deliver on the above information-theoretic results and lend themselves to efficient encoding and decoding. The first technique to design USTM constellations was proposed in [HMR$^+$00], which, while allowing for efficient decoding, was later shown in [MHH02] to have poor performance, especially at high rates. The constellation proposed in [MHH02], on the other hand, while, theoretically having good performance, has to date no tractable decoding algorithm. Recently, a USTM design method based on the exponential map has been proposed in [GKB02].

In USTM, the first $M$ columns of the $T \times T$ unitary matrices are chosen to be the transmitted signal. Therefore, let us first look at the space of $T \times T$ unitary matrices which is referred as the *Stiefel* manifold. It is well-known that this manifold is highly non-linear and non-convex. Note that an arbitrary complex $T \times T$ matrix has $2T^2$ real parameters, but for a unitary one, there are $T$ constraints to force each column to have unit norm and another $2 \times \frac{T(T-1)}{2}$ constraints to make the $T$ columns pairwise orthogonal. Therefore, the Stiefel manifold has dimension $2T^2 - T - 2 \times \frac{T(T-1)}{2} = T^2$. Similarly, the space of $T \times M$ unitary matrices has dimension $2TM - M - 2 \times \frac{M(M-1)}{2} = 2TM - M^2$.

To design codes of unitary matrices, we need first a parameterization of the space of unitary matrices. There are some parameterization methods in existence but all of them suffer from disadvantages for use in unitary space-time code design. We now briefly discuss these. The discussion is based on [HH02a].

The first parameterization method is by Givens rotations. A unitary matrix $\Phi$ can be written as the product

$$\Phi = G_1 G_2 \cdots G_{T(T-1)/2} D G_{T(T+1)/2} \cdots G_{T(T-1)},$$

where $D$ is a diagonal unitary matrix and $G_i$s are Givens (or planar) rotations, one for each of the $\frac{T(T-1)}{2}$ two-dimensional hyperplanes [HJ91]. It is conceivable that one can encode the data onto the angles of rotations and also the diagonal phases of $D$. But it is not a practical method since neither is the parameterization one-to-one (for example, one can re-order Givens rotations) nor does systematic decoding appears to be possible.

Another method is to parameterize with Householder reflections. A unitary matrix $\Phi$ can be written as the product $\Phi = DH_1H_2\cdots H_T$, where $D$ is a diagonal matrix and

$$H_i = I_M - 2\frac{h^{(i)}h^{(i)*}}{\|h^{(i)}\|_F^2}, \quad h^{(i)} = \begin{bmatrix} 0 & \cdots & 0 & 1 & h_{i+1}^{(i)} & \cdots & h_M^{(i)} \end{bmatrix}$$

are Householder matrices [GL96]. This method is also not encouraging to us because we do not know how to encode and decode data onto Householder matrices in any efficient manner.

And also, unitary matrices can be parameterized with the matrix exponential $\Phi = e^{iA}$. When $A$ is $T \times T$ Hermitian, $\Phi$ is unitary. The exponential map also has the difficulty of not being one-to-one. This can be overcome by imposing constraints $0 \leqslant A < 2\pi I$, but the constraints are not linear although convex. We do not know how to sample the space of $A$ to obtain a constellation of $\Phi$. Moreover, the map cannot easily be inverted at the receiver for $T > 1$. Nonetheless, a method based on the exponential map has been proposed in [GKB02].

In this chapter, design of USTM constellations using Cayley transform is proposed. This can be regarded as an extension, to the non-square case, of the earlier work on Cayley codes for *differential* USTM [HH02a]. As will be shown later, this extension is far from trivial. Nonetheless, the codes designed here inherit many of the properties of Cayley differential codes. In particular, they:

1. are very simple to encode: the data is broken into substreams used to parame-

terize the unitary matrices,

2. can be used for systems with any number of transmit and receive antennas,

3. can be decoded in a variety of ways including simple polynomial-time linear-algebraic techniques such as successive nulling-and-cancelling (V-BLAST [GFVW99, Has99]) or sphere decoding [FP85, DCB00],

4. satisfy a probabilistic criterion: they maximize an expected distance between matrix pairs,

The work in this chapter has been published in *IEEE Transactions on Signal Processing Special Issue on MIMO Communications* [JH03e], *the Proceeding of 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)* [HJ02a], and *the Proceeding of 2002 IEEE International Symposium on Information Theory (ISIT'02)* [HJ02b].

## 3.2 Cayley Transform

Cayley transform was proposed in [HH02a] to design codes for differential unitary space-time modulation whereby both good performance and simple encoding and decoding are obtained.

The Cayley transform of a complex $T \times T$ matrix $Y$ is defined to be

$$\Phi = (I + Y)^{-1}(I - Y),$$

where $Y$ is assumed to have no eigenvalue at $-1$ so that the inverse exists. Let $A$ be a $T \times T$ Hermitian matrix and consider the Cayley transform of the skew-Hermitian matrix $Y = iA$:

$$\Phi = (I + iA)^{-1}(I - iA). \tag{3.1}$$

First note that since $iA$ is skew-Hermitian, it has no eigenvalue at $-1$ because all its eigenvalues are strictly imaginary. That means that $(I + iA)^{-1}$ always exists. From definition, Cayley transform is the generalization of the scalar transform

$$v = \frac{1 - ia}{1 + ia},$$

which maps the real line to the unit circle. Notice that no finite point on the real line can be mapped to the point,$-1$, on the unit circle.

The most prominent advantage of Cayley transform is that it maps the complicated space of unitary matrices to the space of Hermitian matrices, which is linear. It can be easily proved that

$$\begin{aligned} \Phi\Phi^* &= (I + iA)^{-1}(I - iA)[(I + iA)^{-1}(I - iA)]^* \\ &= (I + iA)^{-1}(I - iA)(I + iA)(I - iA)^{-1} \\ &= I. \end{aligned}$$

The second equation is true because $I - iA, I + iA, (I - iA)^{-1}$ and $(I + iA)^{-1}$ all commute. Similarly, $\Phi^*\Phi = I$ can also be proved. Therefore, similar to the matrix exponential, Cayley transform maps the complicated Stiefel manifold of unitary matrices to the space of Hermitian matrices. Hermitian matrices are easy to characterize since they form a linear vector space over the reals. Therefore, easy encoding and decoding can be obtained.

From (3.1) it can be proved easily that

$$iA = (I + \Phi)^{-1}(I - \Phi)$$

provided that $(I + \Phi)^{-1}$ exists. This shows that Cayley transform and its inverse transform coincide. Thus, Cayley transform is one-to-one. It is not an onto map

because those unitary matrices with eigenvalues at $-1$ have no inverse images. Recall that the space of Hermitian or skew-Hermitian matrices has dimension $T^2$ which matches that of Stiefel manifold.

We have shown that a matrix with no eigenvalues at $-1$ is unitary if and only if its Cayley transform is skew-Hermitian. Compared with other parameterizations of unitary matrices, the parameterization with Cayley transform is one-to-one and easily invertible.

And also, it is proved in [HH02a] that a set of unitary matrices is fully diverse if and only if the set of their Hermitian inverse Cayley transforms is fully diverse. This suggests that a set of unitary matrices with promising performance can be obtained from a well-designed set of Hermitian matrices by Cayley transform.

In [HH02a], Cayley transform has been used in the design of $M \times M$ unitary space-time codes for differential modulation. The idea is to design a good set of Hermitian matrices and use their Cayley transform as the signal matrices. In this chapter we generalize this idea to the non-square case. It can be seen in the following sections that the generalization is far from trivial since the non-squareness of the matrices causes a lot of problems in the code design.

## 3.3 The Idea of Cayley Unitary Space-Time Codes

Because Cayley transform maps the nonlinear Stiefel manifold to the linear space (over the reals) of Hermitian matrices (and vice-versa), it is convenient and most straightforward to encode data linearly onto Hermitian matrices and then apply Cayley transform to get unitary matrices.

We call a set of $T \times M$ unitary matrices a Cayley unitary space-time code if any

element in the set can be written as

$$S = (I_T + iA)^{-1}(I_T - iA) \begin{bmatrix} I_M \\ 0 \end{bmatrix} \tag{3.2}$$

with the Hermitian matrix $A$ given by

$$A = \sum_{q=1}^{Q} \alpha_q A_q, \tag{3.3}$$

where $\alpha_1, \alpha_2, ..., \alpha_Q$ are real scalars (chosen from a set $\mathcal{A}$ with $r$ possible values) and $A_1, A_2, ..., A_Q$ are fixed $T \times T$ complex Hermitian matrices.

The code is completely determined by the set of matrices $\{A_1, A_2, ..., A_Q\}$, which can be thought of as Hermitian basis matrices. Each individual codeword, on the other hand, is determined by our choice of the scalars $\alpha_1, \alpha_2, ..., \alpha_Q$ whose values are in the set $\mathcal{A}_r$ (the subscript '$r$' represents the cardinality of the set). Since each of the $Q$ real coefficients may take on $r$ possible values and the code occupies $T$ channel uses, the transmission rate is $R = (Q/T)\log_2 r$. We defer the discussions on how to design $A_q$'s, $Q$, and the set $\mathcal{A}_r$ to the later part of this chapter and concentrate on how to decode $\alpha_1, \alpha_2, ..., \alpha_Q$ at the receiver first.

## 3.4 A Fast Decoding Algorithm

Similar to differential Cayley codes, our Cayley unitary space-time codes also have the good property of linear decoding, which means that the receiver can be made to form a system of linear equations in the real scalars $\alpha_1, \alpha_2, ..., \alpha_Q$. First, it is useful to see what our codes and their ML decoding look like.

Partition the $T \times T$ matrix $A$ as

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $A_{11}$ is an $M \times M$ matrix and $A_{22}$ is a $(T-M) \times (T-M)$ matrix. For $A$ being Hermitian, $A_{11}$ and $A_{22}$ must both be Hermitian and also $A_{21} = A_{12}^*$.

Observe that

$$\begin{aligned}
& \Phi \\
= \; & (I + iA)^{-1}(I - iA) \\
= \; & (I + iA)^{-1}[2I - (I + iA)] \\
= \; & 2(I + iA)^{-1} - I \\
= \; & 2 \begin{bmatrix} I_M + iA_{11} & iA_{12} \\ iA_{12}^* & I_{T-M} + iA_{22} \end{bmatrix}^{-1} - I \\
= \; & \begin{bmatrix} 2[I - (I + iA_{11})^{-1}A_{12}\Delta_2^{-1}A_{12}^*](I + iA_{11})^{-1} - I & -2i(I + iA_{11})^{-1}A_{12}\Delta_2^{-1} \\ -2i\Delta_2^{-1}A_{12}^*(I + iA_{11})^{-1} & 2\Delta_2^{-1} - I \end{bmatrix}
\end{aligned}$$

where $\Delta_2 = I + iA_{22} + A_{12}^*(I + iA_{11})^{-1}A_{12}$ is the Schur complement of $I + iA_{11}$ in $I + A$.

Therefore, from (3.2),

$$S = \begin{bmatrix} 2[I - (I + iA_{11})^{-1}A_{12}\Delta_2^{-1}A_{12}^*](I + iA_{11})^{-1} - I \\ -2i\Delta_2^{-1}A_{12}^*(I + iA_{11})^{-1} \end{bmatrix}, \tag{3.4}$$

which is composed by the first $M$ columns of $\Phi$, and

$$S^{\perp} = \begin{bmatrix} -2i(I + iA_{11})^{-1}A_{12}\Delta_2^{-1} \\ 2\Delta_2^{-1} - I \end{bmatrix}$$

is the unitary complement of $S$. In fact, it can be algebraically verified that both $S$ and $S^{\perp}$ are unitary.

By partitioning the received signal matrix $X$ into an $M \times N$ block $X_1$ and a $(T-M) \times N$ block $X_2$ as $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, the second form of the ML decoder in (2.7) reduces to

$$\arg\min_{\{\alpha_q\}} \left\| [-2iX_1^*(I + iA_{11})^{-1}A_{12} + X_2^*(2 - \Delta_2)]\Delta_2^{-1} \right\|_F^2. \tag{3.5}$$

The reason for choosing the second form of the ML decoding, as opposed to the first one, is that we prefer to minimize, rather than maximize the Frobenius norm. In fact, we shall presently see in the following that a simple approximation leads us to a quadratic minimization problem, which can be solved conveniently via sphere decoding.

It is easy to see that the decoding formula given in (3.5) is not quadratic in entries of $A$, which indicates that it is not quadratic in $\alpha_q$s since $A$ is linear in $\alpha_q$s. Therefore, the system equation at the receiver is not linear. The formula looks intractable because it has matrix inverses as well as the Schur complement $\Delta_2$. Adopting the approach of [HH02a] by ignoring the covariance of the additive noise term $\Delta_2^{-1}$, we obtain

$$\arg\min_{\{\alpha_q\}} \left\| 2X_2^* - X_2^*\Delta_2 - 2iX_1^*(I + iA_{11})^{-1}A_{12} \right\|_F^2, \tag{3.6}$$

which, however, is still not quadratic in entries of $A$. Therefore, to simplify the

formula, more constraints should be imposed on the Hermitian matrix $A$. That is, our $A$ matrix should have a more handy structure. Fortunately, observe that the number of degrees of freedom in a $T \times T$ Hermitian matrix is $T^2$, but the number of degrees of freedom in a $T \times M$ unitary matrix $S$ is only $2TM - M^2 = T^2 - (T-M)^2$. There are $(T-M)^2$ more degrees of freedom in $A$ than needed. So let us exploit this. Indeed, if we let

$$(I + iA_{11})^{-1} A_{12} = B \tag{3.7}$$

for some *fixed* $M \times (T-M)$ matrix $B$, by which $2M(T-M)$ degrees of freedom are lost [1]. Then,

$$A_{12} = (I + iA_{11})B \tag{3.8}$$

and

$$\Delta_2 = I + B^*B - iB^*A_{11}B + iA_{22}. \tag{3.9}$$

Some algebra shows that the above decoding formula (3.6) reduces to

$$\hat{\alpha}_{\text{lin}} = \arg\min_{\{\alpha_q\}} \|X_2^* - X_2^*B^*B - 2iX_1^*B + iX_2^*B^*A_{11}B - iX_2^*A_{22}\|_F^2 , \tag{3.10}$$

which is now quadratic in entries of $A$. Fast decoding methods such as sphere decoding and nulling-and-canceling can be used which have polynomial complexity as in BLAST [Fos96].

We call (3.10) the "linearized" decoding because it is equivalent to the decoding

[1] With this condition, the number of degrees of freedom in $A$ is $T^2 - 2TM + 2M^2$, which is greater than $2TM - M^2$, the number of degrees of freedom in an arbitrary $T \times M$ unitary matrix, when $T \geq 3M$.

of a system whose system equation is linear in the unknowns $\alpha_q$s. For a wide range of rates and SNR, (3.10) can be solved exactly in roughly $O(Q^3)$ computations using sphere decoding [FP85, DCB00]. Furthermore, simulation results show that the penalty for using (3.10) instead of the exact ML decoding is small, especially when weighed against the complexity of the exact ML decoding. To facilitate the presentation of the sphere decoding algorithm, the equivalent channel model in matrices are shown in the following subsection.

### 3.4.1 Equivalent Model

From (3.8), $A_{12} = A_{21}^*$ is fully determined by $A_{11}$. Therefore, the degrees of freedoms in $A$ are all in matrices $A_{11}$ and $A_{22}$. The encoding formula (3.3) of $A$ can thus be modified to the following encoding formulas of $A_{11}$ and $A_{22}$:

$$A_{11} = \sum_{q=1}^{Q} \alpha_q A_{11,q} \quad \text{and} \quad A_{22} = \sum_{q=1}^{Q} \alpha_q A_{22,q}, \tag{3.11}$$

where $Q$ is the number of possible $A_{11,q}$s and $A_{22,q}$s, $\alpha_1, \alpha_2, ...\alpha_Q$ are real scalars chosen from the set $\mathcal{A}_r$, and $A_{11,1}, A_{11,2}, ..., A_{11,Q}$ and $A_{22,1}, A_{22,2}, ..., A_{22,Q}$ are fixed $M \times M$ and $(T-M) \times (T-M)$ complex Hermitian matrices.[2] The matrix $A$ is therefore constructed as

$$\begin{aligned} A &= \begin{bmatrix} A_{11} & (I + iA_{11})B \\ B^*(I - iA_{11}) & A_{22} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{q=1}^{Q} \alpha_q A_{11,q} & (I + i\sum_{q=1}^{Q} \alpha_q A_{11,q})B \\ B^*(I - i\sum_{q=1}^{Q} \alpha_q A_{11,q}) & \sum_{q=1}^{Q} \alpha_q A_{22,q} \end{bmatrix} \end{aligned}$$

[2] Actually, in our design , $A_{11}$ and $A_{22}$ can have different numbers of degrees of freedom, $Q_1$ and $Q_2$, and the coefficients of the two basis sets can have non-identical sample spaces. That is, we can have $A_{11} = \sum_{p=1}^{Q_1} \alpha_p A_{11,p}, A_{22} = \sum_{q=1}^{Q_2} \beta_q A_{22,q}$ where $\alpha_i \in \mathcal{A}_{r_1}$ and $\beta_i \in \mathcal{A}_{r_2}$. However, to simplify the design problem, here we just set $Q_1 = Q_2$ and $r_1 = r_2$.

$$= \sum_{q=1}^{Q} \alpha_q \begin{bmatrix} A_{11,q} & iA_{11,q}B \\ -iB^*A_{11,q} & A_{22,q} \end{bmatrix} + \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix}.$$

Therefore, the linearized ML decoding (3.10) can be written as

$$\arg\min_{\{\alpha_q\}} \left\| X_2^* - X_2^*B^*B - 2iX_1^*B + i\sum_{q=1}^{Q} \alpha_q X_2^*B^*A_{11,q}B - i\sum_{q=1}^{Q} \alpha_q X_2^*A_{22,q} \right\|_F^2. \quad (3.12)$$

Define

$$C = X_2^* - X_2^*B^*B - 2iX_1^*B, \quad \text{and} \quad J_q = -iX_2^*B^*A_{11,q}B + iX_2^*A_{22,q} \quad (3.13)$$

for $q = 1, 2, ..., Q$. By decomposing the complex matrices $C$ and $J_q$ into their real and imaginary parts, the decoding formula (3.12) can be further rewritten as

$$\arg\min_{\{\alpha_q\}} \left\| \begin{bmatrix} C_R \\ C_I \end{bmatrix} - \begin{bmatrix} J_{1,R} & \cdots & J_{Q,R} \\ J_{1,I} & \cdots & J_{Q,I} \end{bmatrix} \begin{bmatrix} \alpha_1 I_{T-M} \\ \vdots \\ \alpha_Q I_{T-M} \end{bmatrix} \right\|_F^2,$$

where $C_R, C_I$ are the real and imaginary parts of $C$ and $J_{i,R}, J_{i,I}$ are the real and imaginary parts of $J_i$. Also, denoting by $C_{R,j}, C_{I,j}, J_{i,R,j}, J_{i,I,j}$ the $j$-th columns of $C_R$, $C_I$, $J_{i,R}$, $J_{i,I}$ for $j = 1, 2, ..., (T-M)$, and writing matrices in the above formula column by column, the formula can be further simplified to

$$\arg\min_{\{\alpha_q\}} \|\mathcal{R} - \mathcal{H}\underline{\alpha}\|_F^2, \quad (3.14)$$

where $\mathcal{R}$ is the $2N(T-M)$-dimensional column vector $\begin{bmatrix} C_{R,1}^t & C_{I,1}^t & \cdots & C_{R,T-M}^t & C_{I,T-M}^t \end{bmatrix}^t$ and $\mathcal{H}$ is the $2N(T-M) \times Q)$ matrix

$$\begin{bmatrix} J_{1,R,1} & J_{2,R,1} & \cdots & J_{Q,R,1} \\ J_{1,I,1} & J_{2,I,1} & \cdots & J_{Q,I,1} \\ \vdots & \vdots & \ddots & \vdots \\ J_{1,R,T-M} & J_{2,R,T-M} & \cdots & J_{Q,R,T-M} \\ J_{1,I,T-M} & J_{2,I,T-M} & \cdots & J_{Q,I,T-M} \end{bmatrix}. \tag{3.15}$$

$\underline{\alpha} = [\alpha_1, \cdots, \alpha_Q]^t$ is the vector of unknowns. Therefore, we obtain the equivalent channel model

$$\mathcal{R} = \mathcal{H}\underline{\alpha} + \mathcal{W}, \tag{3.16}$$

where $\mathcal{W}$ is the equivalent noise matrix. $\underline{\alpha}$ appears to pass through an equivalent channel $\mathcal{H}$ and is corrupted by additive noise.[3] The equivalent channel, $\mathcal{H}$ given in (3.15), is known to the receiver because it is a function of $A_{11,1}$, $A_{11,2}$, $\cdots$, $A_{11,Q}$, $A_{22,1}$, $A_{22,2}$, $\cdots$, $A_{22,Q}$, $X_1$, and $X_2$.

Therefore, the decoding is equivalent to decoding of a simple linear system, which can be done using known techniques such as successive nulling-and-canceling, efficient square-root implementation, and sphere decoding. Efficient implementations of nulling-and-canceling generally require $O(Q^3)$ computations. Sphere decoding can be regarded as a generalization of nulling-and-canceling where at each step, rather than making a hard decision on the corresponding $\alpha_q$s, one considers all $\alpha_q$s that lie within a sphere of a certain radius. Sphere decoding has the important advantage over nulling-and-canceling that it computes the *exact* solution. Its worst case behavior is exponential in $Q$, but its average behavior is comparable to nulling-and-canceling. When the number of transmit antennas and the rate are small, exact ML decoding

[3] In general, the covariance of the noise is dependent on the transmitted signal. However, in ignoring $\Delta_2^{-1}$ in (3.6), we have ignored this signal dependence.

using exhaustive search is possible. However, a search over all possible $\alpha_1, ..., \alpha_Q$ may be impractical for large $T$ and $R$. Fortunately, the performance penalty for the linearized Ml decoding given in (3.10) is small, especially weighed against the complexity of exact ML decoding using exhaustive search.

### 3.4.2 Number of Independent Equations

Nulling-and-canceling explicitly requires that the number of equations be at least as large as the number of unknowns. Sphere decoding does not have this hard constraint, but it benefits from more equations because the computational complexity grows exponentially in the difference between the number of unknowns and the number of independent equations. To keep the complexity of sphere decoding algorithm polynomial, it is important that the number of linear equations resulting from (3.10) be at least as large as the number of unknowns. (3.16) suggests that there are $2N(T-M)$ real equations and $Q$ real unknowns. Hence we may impose the constraint

$$Q \leq 2N(T-M).$$

This argument assumes that the matrix $H$ has full column rank. There is, at first glance, no reason to assume otherwise but it turns out to be false. Due to the Hermitian constraint on $A$, not all the $2M(T-M)$ equations are independent. A careful analysis yields the following result.

**Theorem 3.1 (Rank of $H$).** *The matrix given in (3.15) generally has rank*

$$rank(\mathcal{H}) = \begin{cases} \min\left(2N(T-M) - N^2, Q\right) & \text{if } T-M \geq N \\ \min\left((T-M)^2, Q\right) & \text{if } T-M < N \end{cases}. \tag{3.17}$$

**Proof:** First assume that $T - M \geq N$. The rank of $\mathcal{H}$ is the dimension of the

range space of $\underline{c}$ in the equation $\underline{c} = \mathcal{H}\underline{a}$ as $\underline{a}$ varies. Equivalently, the rank of $\mathcal{H}$ is the dimension of the range space of the $N \times (T-M)$ complex matrix $C$ in the equation $C = iX_2^*(A_{22} - B^*A_{11}B)$ when $A_{11}$ and $A_{22}$ vary. Because $A_{11}$ and $A_{22}$ are not arbitrary matrices, the range space of $C$ cannot have all the $2(T-M)N$ dimensions as it appears. Now let's study the number of constraints added on the range space of $C$ as $A_{11}$ and $A_{22}$ can only be Hermitian matrices. Since

$$\begin{aligned}
[C(iX_2)]^* &= -iX_2^*(A_{22} - B^*A_{11}B)(-i)X_2 \\
&= iX_2^*(A_{22} - B^*A_{11}B)(iX_2) \\
&= C(iX_2),
\end{aligned}$$

the $N \times N$ matrix $C(iX_2)$ is Hermitian. This enforces $N^2$ linear constraints on entries of $C$. Therefore, only at most $2(T-M)N - N^2$ entries of all the $2(T-M)N$ entries are free. Since $\mathcal{H}$ is $2(T-M)N \times Q$, the rank of $\mathcal{H}$ is at most $\min\left(2(T-M)N - N^2, Q\right)$.

Now assume that $T - M < N$. We know that the $N \times N$ matrix $C(iX_2)$ is Hermitian but has rank $T - M < N$ now instead of full rank. Therefore, entries of the lower right $[N-(T-M)] \times [N-(T-M)]$ Hermitian sub-matrix of $C(iX_2)$ are uniquely determined by its other entries. Therefore, the number of constraints yielded by equation $C(iX_2) = (C(iX_2))^*$ is $N^2 - (N-(T-M))^2 = 2N(T-M) - (T-M)^2$. Thus, there are at most $2N(T-M) - (2N(T-M) - (T-M)^2) = (T-M)^2$ degrees of freedom in $C$. The rank of $\mathcal{H}$ is at most $\min((T-M)^2, Q)$.

We have essentially proved an upper bound on the rank. Our argument so far has not relied on any specific sets for $A_{11}$ and $A_{22}$. When $A_{11} = 0$, we are reduced to studying $iX_2^*A_{22}$, which is the same setting as that of differential USTM [HH02a]. In Theorem 1 of [HH02a], it is argued that for a generic choice of the basis matrices $A_{22,1}, \cdots, A_{22,Q}$, the rank of $\mathcal{H}$ attains the upper bound. Therefore the same holds here, and $\mathcal{H}$ attains the upper bound. $\square$

Theorem 3.1 shows that even though there are $2N(T-M)$ equations in (3.16), not all of them are independent. To have at least as many equations as unknowns, The following constraint

$$Q \leq \begin{cases} 2N(T-M)-N^2 & \text{if } T-M \geq N \\ (T-M)^2 & \text{if } T-M < N \end{cases}$$

is needed, or equivalently,

$$Q \leq \min(T-M, N)\max(2(T-M)-N, T-M). \tag{3.18}$$

## 3.5 A Geometric Property

With the choice (3.7) or equivalently (3.8), the first block of the transmitted matrix $S$ in (3.4) can be simplified as

$$\begin{aligned} & 2[I-(I+iA_{11})^{-1}A_{12}\Delta_2^{-1}A_{12}^*](I+iA_{11})^{-1}-I \\ = \;& [2I-2B\Delta_2^{-1}B^*(I-iA_{11})-(I+iA_{11})](I+iA_{11})^{-1} \\ = \;& [(I-iA_{11})-2B\Delta_2^{-1}B^*(I-iA_{11})](I+iA_{11})^{-1} \\ = \;& [I-2B\Delta_2^{-1}B^*](I-iA_{11})(I+iA_{11})^{-1}. \end{aligned}$$

The second block of $S$ equals $-2i\Delta_2^{-1}B^*(I-iA_{11})(I+iA_{11})^{-1}$. Since $(I-iA_{11})$ and $(I+iA_{11})^{-1}$ commute,

$$S = \begin{bmatrix} I-2B\Delta_2^{-1}B^* \\ -2i\Delta_2^{-1}B^* \end{bmatrix}(I+iA_{11})^{-1}(I-iA_{11}).$$

Our Cayley unitary space-time code and its unitary complement can be written as

$$S = \begin{bmatrix} I & -iB \\ 0 & I \end{bmatrix} \begin{bmatrix} I_M \\ -2i\Delta_2^{-1}B^* \end{bmatrix} U_1 \quad \text{and} \quad S^{\perp} = \begin{bmatrix} -2iB\Delta_2^{-1} \\ 2\Delta_2^{-1} - I_{T-M} \end{bmatrix}, \quad (3.19)$$

where

$$\Delta_2 = I + B^*B - i\sum_{q=1}^{Q} \alpha_q B^* A_{11,q} B + i\sum_{q=1}^{Q} \alpha_q A_{22,q} \quad (3.20)$$

and $U_1 = (I + iA_{11})^{-1}(I - iA_{11})$ is an $M \times M$ unitary matrix since it is the Cayley transform of the Hermitian matrix $A_{11}$.

The code in (3.19) is completely determined by matrices $A_{11,1}, A_{11,2}, ..., A_{11,Q}$ and $A_{22,1}, A_{22,2}, ..., A_{22,Q}$, which can be thought of as Hermitian basis matrices. Each individual codeword, on the other hand, is determined by our choice of the scalars $\alpha_1, \alpha_2, ..., \alpha_Q$ chosen from the set $\mathcal{A}_r$. Since there are $Q$ basis matrices for $A_{11}$ and $A_{22}$, and the code occupies $T$ channel uses, the transmission rate is

$$R = \frac{Q}{T} \log_2 r. \quad (3.21)$$

Since the channel matrix $H$ is unknown and if it is left multiplied by an $M \times M$ unitary matrix, its distribution remains unchanged, we can combine $U_1$ with the channel matrix $H$ to get $H' = U_1 H$. If we left multiply $X, S$ and $V$ by $\begin{bmatrix} I_M & -iB \\ 0 & I_{T-M} \end{bmatrix}^{-1} = \begin{bmatrix} I_M & iB \\ 0 & I_{T-M} \end{bmatrix}$ to get $X', S'$ and $V'$, the system equation (2.1) can be rewritten as

$$X' = \sqrt{\frac{\rho T}{M}} \begin{bmatrix} I_M \\ -2i\Delta_2^{-1}B^* \end{bmatrix} H' + V'.$$

We can see that this is very similar to the equation of training-based schemes (2.5). The only difference is in the noises. In (2.5), entries of the noise are independent white Gaussian noise with zero-mean and unit-variance. Here, entries of $V'$ are no longer independent with unit-variance, although they still have zero-mean. The dependence of the noises is beneficial to the performance since more information can be obtained.

The following theorem about the structure of $S^{\perp}$ is needed later in the optimization of the basis matrices.

**Theorem 3.2 (Difference of unitary complements of the transmitted signal).** *The difference of the unitary complements $S^{\perp}$ and $\hat{S}^{\perp}$ of the transmitted signals $S$ and $\hat{S}$ can be written as*

$$S^{\perp} - \hat{S}^{\perp} = 2 \begin{bmatrix} -iB \\ I \end{bmatrix} \Delta_2^{-1} (\hat{\Delta}_2 - \Delta_2) \hat{\Delta}_2^{-1}, \tag{3.22}$$

where $\Delta_2$ and $\hat{\Delta}_2$ are the corresponding Schur complements.

**Proof:** From (3.19),

$$S^{\perp} = \begin{bmatrix} -2iB\Delta_2^{-1} \\ 2\Delta_2^{-1} - I \end{bmatrix} = \begin{bmatrix} -2iB \\ 2I - \Delta_2 \end{bmatrix} \Delta_2^{-1}$$

and

$$S^{\perp} = \begin{bmatrix} -2i\Delta_1^{-1} A_{12} (I + iA_{22})^{-1} \\ 2\Delta_2^{-1} - I \end{bmatrix} = \begin{bmatrix} \Delta_1^{-1} & 0 \\ 0 & \Delta_2^{-1} \end{bmatrix} \begin{bmatrix} -2iA_{12}(I + iA_{22})^{-1} \\ 2I - \Delta_2 \end{bmatrix}.$$

From algebra, it is easy to get $\Delta_1^{-1} A_{12} (I + jA_{22})^{-1} = (I + jA_{11})^{-1} A_{12} \Delta_2^{-1}$. From

(3.7), $\Delta_1^{-1}A_{12}(I+jA_{22})^{-1} = B\Delta_2^{-1}$, and thus, $\Delta_1 B = A_{12}(I+jA_{22})^{-1}\Delta_2$. Therefore,

$$
\begin{aligned}
& S^{\perp} - \hat{S}^{\perp} \\
= & \begin{bmatrix} \Delta_1^{-1} & 0 \\ 0 & \Delta_2^{-1} \end{bmatrix} \left( \begin{bmatrix} -2iA_{12}(I+iA_{22})^{-1} \\ 2I - \Delta_2 \end{bmatrix} \hat{\Delta}_2 - \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix} \begin{bmatrix} -2iB \\ 2I - \hat{\Delta}_2 \end{bmatrix} \right) \hat{\Delta}_2^{-1} \\
= & \begin{bmatrix} \Delta_1^{-1} & 0 \\ 0 & \Delta_2^{-1} \end{bmatrix} \begin{bmatrix} -2iA_{12}(I+iA_{22})^{-1}\hat{\Delta}_2 + 2i\Delta_1 B \\ 2\hat{\Delta}_2 - \Delta_2\hat{\Delta}_2 - 2\Delta_2 + \Delta_2\hat{\Delta}_2 \end{bmatrix} \hat{\Delta}_2^{-1} \\
= & \begin{bmatrix} \Delta_1^{-1} & 0 \\ 0 & \Delta_2^{-1} \end{bmatrix} \begin{bmatrix} 2iA_{12}(I+iA_{22})^{-1}\Delta_2 - 2iA_{12}(I+iA_{22})^{-1}\hat{\Delta}_2 \\ 2(\hat{\Delta}_2 - \Delta_2) \end{bmatrix} \hat{\Delta}_2^{-1} \\
= & \begin{bmatrix} -2i\Delta_1^{-1}A_{12}(I+iA_{22})^{-1} & 0 \\ 0 & 2\Delta_2^{-1} \end{bmatrix} \begin{bmatrix} \hat{\Delta}_2 - \Delta_2 \\ \hat{\Delta}_2 - \Delta_2 \end{bmatrix} \hat{\Delta}_2^{-1} \\
= & \begin{bmatrix} -2iB\Delta_2^{-1} \\ 2\Delta_2^{-1} \end{bmatrix} (\hat{\Delta}_2 - \Delta_2)\hat{\Delta}_2^{-1} \\
= & 2\begin{bmatrix} -iB \\ I \end{bmatrix} \Delta_2^{-1}(\hat{\Delta}_2 - \Delta_2)\hat{\Delta}_2^{-1}
\end{aligned}
$$

□

Another way to look at Theorem 3.2 is to note that

$$S^{\perp} = \begin{bmatrix} 0 \\ -I \end{bmatrix} + 2\begin{bmatrix} -iB \\ I \end{bmatrix} \Delta_2^{-1}. \tag{3.23}$$

Without the unitary constraint, this is an affine space since all the data is encoded in $\Delta_2^{-1}$. So, in general, the space of $S^{\perp}$ is the intersection of the linear affine space in (3.23) and the Stiefel manifold $S^{\perp *}S^{\perp} = I$. We can see from (3.22) or (3.23) that the dimension of the range space of $S^{\perp} - S'^{\perp}$ (equivalently the dimension of the affine space) is $T-M$. It is interesting to compare this with that of training-based schemes,

which from (2.6), gives

$$S^{\perp} - \hat{S}^{\perp} = \frac{1}{\sqrt{2}} \begin{bmatrix} -(S_d^* - \hat{S}_d^*) \\ 0 \end{bmatrix}. \tag{3.24}$$

Note now that the dimension of the affine space is $\min(M, T-M)$ which is smaller than $T-M$ when $T > 2M$. So, the affine space of $S^{\perp}$ of Cayley codes has a higher dimension than that of training-based schemes when $T > 2M$.

## 3.6 Design of Cayley Unitary Space-Time Codes

Although the idea of Cayley unitary space-time codes has been introduced in (3.19), we have not yet specified $Q$, nor have we explained how to choose the discrete set $\mathcal{A}_r$ from which $\alpha_q$s are drawn, or the design of the Hermitian basis matrices $\{A_{11,1}, A_{11,2}, ..., A_{11,Q}\}$ and $\{A_{22,1}, A_{22,2}, ..., A_{22,Q}\}$. We now discuss these issues.

### 3.6.1 Design of $Q$

To make the constellation as rich as possible, we should make the number of degrees of freedom $Q$ as large as possible. Therefore, as a general practice, we find it useful to take $Q$ as its upper bound in (3.18). That is,

$$Q = \min(T-M, N)\max(2(T-M) - N, T-M). \tag{3.25}$$

We are left with how to design the discrete set $\mathcal{A}_r$ and how to choose $\{A_{11,1}, A_{11,2}, \cdots, A_{11,Q}\}$ and $\{A_{22,1}, A_{22,2}, \cdots, A_{22,Q}\}$.

### 3.6.2 Design of $\mathcal{A}_r$

As mentioned in Section 2.5, at high SNR, to achieve capacity in the sense of maximizing mutual information between $X$ and $S$, $\Phi = (I+iA)^{-1}(I-iA)$ should assemble samples from an isotropic random distribution. Since our data modulates the $A$ matrix (or equivalently $A_{11}$ and $A_{22}$), we need to find the distribution on $A$ that yields an isotropically distributed $\Phi$.

As proved in [HH02a], the unitary matrix $\Phi$ is isotropically distributed if and only if the Hermitian matrix $A$ has the matrix Cauchy distribution

$$p(A) = \frac{2^{T^2-T}(T-1)!\cdots 1!}{\pi^{T(T+1)/2}} \frac{1}{\det(I+A^2)^T},$$

which is the matrix generalization of the familiar scalar Cauchy distribution

$$p(a) = \frac{1}{\pi(1+a^2)}.$$

For the one-dimensional case, an isotropic-distributed scalar $v$ can be written as $v = e^{i\theta}$, where $\theta$ is uniform over $[0, 2\pi)$. So, $a = -i\frac{1-e^{i\theta}}{1+e^{i\theta}} = -\tan(\theta/2)$ is Cauchy. When there is only one transmit antenna ($M = 1$) and the coherence interval is just one ($T = 1$), the transmitted signals are scalars. There is no need to partition the matrix $A$. Therefore (3.3) is used instead of (3.11). We want our code constellation $A = \sum_{q=1}^{Q} \alpha_q A_q$ to resemble samples from a Cauchy random matrix distribution. Since there is only one degree of freedom in a scalar, it is obvious that $Q = 1$. Without loss of generality, setting $A_1 = 1$, we get

$$v = \frac{1-i\alpha_1}{1+i\alpha_1}, \quad \text{and} \quad \alpha_1 = -i\frac{1-v}{1+v}.$$

To have a code with rate $R = (Q/T)\log_2 r$ at $T = M = 1$, $\mathcal{A}$ should have $r = 2^R$ points. Standard DPSK puts these points uniformly around the unit circle at angular

intervals of $2\pi/r$ with the first point at $\pi/r$. For a point of angle $\theta$ on the unit circle, the corresponding value for $\alpha_1$ is

$$\alpha_1 = -i\frac{1-v}{1+v} = -\tan(\theta/2). \tag{3.26}$$

For example, for $r = 2$, we have the set of points on unit circle $\mathcal{V} = \{e^{i\pi/2}, e^{-i\pi/2}\}$. From (3.26), the set of values for $\alpha_1$ is $\mathcal{A}_2 = \{-1, 1\}$. For the case of $r = 4$, we can get by simple calculation that $\mathcal{A}_4 = \{-2.4142, -0.4142, 0.4142, 2.4142\}$. It can be seen that the points rapidly spread themselves out as $r$ increases, which reflects the heavy tail of the Cauchy distribution.

We denote $\mathcal{A}_r$ to be the image of (3.26) applied to the set $\{\pi/r, 3\pi/r, 5\pi/r, ..., (2r-1)\pi/r\}$. When $r \to \infty$, the fraction of points in the set less than some value $x$ is given by the cumulative Cauchy distribution. Therefore, the set $\mathcal{A}_r$ can be regarded as an $r$-point discretization of a scalar Cauchy random variable.

For the systems with multiple transmit antennas and higher coherence intervals, no direct method is shown about how to choose $\mathcal{A}$. In that case, we also choose our set $\mathcal{A}$ to be the set given above. Thus, $\alpha_q$s are chosen as discretized scalar Cauchy random variables for any $T$ and $M$. But to get rate $R$, from (3.21), we need

$$r^Q = 2^{RT}. \tag{3.27}$$

### 3.6.3 Design of $A_{11,1}, A_{11,2}, ...A_{11,Q}, A_{22,1}, A_{22,2}, ...A_{22,Q}$

To complete the code construction, it is crucial that the two sets of bases $\{A_{11,1}, A_{11,2}, \cdots, A_{11,Q}\}$ and $\{A_{22,1}, A_{22,2}, \cdots, A_{22,Q}\}$ are chosen appropriately, and we present a criterion in this subsection.

If the rates being considered are reasonably small, the diversity product criterion $\min_{l \neq l'} |\det(\Phi_l - \Phi_{l'})^*(\Phi_l - \Phi_{l'})|$ is tractable. At high rates, however, it is not practical to

pursue the full diversity criterion. There are two reasons for this: first, the criterion becomes intractable because of the number of matrices involved and second, the performance of the constellation may not be governed so much by its worst-case pairwise $|\det(\Phi_l - \Phi_{l'})^*(\Phi_l - \Phi_{l'})|$, but rather by how well the matrices are distributed throughout the space of unitary matrices.

Similar to the differential Cayley code design in [HH02a], for given $\mathcal{A}_r$ and the sets of basis matrices $\{A_{11,1}, A_{11,2}, ...A_{11,Q}\}$ and $\{A_{22,1}, A_{22,2}, ...A_{22,Q}\}$, we define a distance criterion for the resulting constellation of matrices $\mathcal{V}$ to be

$$\xi(\mathcal{V}) = \frac{1}{T-M} E \log\det(S^{\perp} - S'^{\perp})^*(S^{\perp} - S'^{\perp}), \tag{3.28}$$

where $S$ is given by (3.19) and (3.20) and $S'$ is given by the same formulas except that the $\alpha_q$s in (3.20) are replace by $\alpha'_q$s. The expectation is over all possible $\alpha_q s$ and $\alpha'_q$s chosen uniformly from $\mathcal{A}_r$ such that $(\alpha_1, ..., \alpha_Q) \neq (\alpha'_1, ..., \alpha'_Q)$. Remember that $S^{\perp}$ denotes the $T \times (T-M)$ unitary complement matrix of the $T \times M$ matrix $S$.

Let us first look at the difference between this criterion with that in [HH02a]. Here, we use $S^{\perp}$ and $S'^{\perp}$ instead of $S$ and $S'$ themselves because the unitary complement instead of the transmitted signal itself is used in the linearized ML decoding. This criterion cannot be directly related to the diversity product as in the case of [HH02a], but still, from the structure, it is a measure of the expected "distance" between matrices $S^{\perp}$ and $S'^{\perp}$. Thus, maximizing $\xi(\mathcal{V})$ should be connected with lowering average pairwise error probability. Hopefully, optimizing the expected "distance" between the unitary complements $S^{\perp}$ and $S'^{\perp}$ instead of that between the unitary signals $S$ and $S'$ themselves will obtain a better performance. And also, since the constraints (3.7) is imposed to simplify $\Delta_2$, which turns out to simplify $S^{\perp}$ as well, the calculation of our criterion is much easier than the calculation of the one used in [HH02a], which maximizes the expected "distance" between the unitary matrices $\Phi$

and $\Phi'$. Therefore, the optimization problem is proposed to be

$$\arg \max_{\{A_{11,q}, A_{22,q}\}, B} \xi(\mathcal{V}). \tag{3.29}$$

By (3.22), we can rewrite the optimization as a function of $A_{11}, A_{22}$ and get the simplified formula,

$$\max_{\{A_{11,q_1}\},\{A_{22,q_2}\},B} \mathrm{E} \log\det[B^*(A_{11} - A'_{11})B - (A_{22} - A'_{22})]^2 \\ -\mathrm{E} \log\det \Delta_2^2 - \mathrm{E} \log\det \Delta_2'^2, \tag{3.30}$$

where

$$\begin{aligned} \Delta_2 &= I + B^*B - iB^*A_{11}B + iA_{22}, \\ \Delta'_2 &= I + B^*B - iB^*A'_{11}B + iA'_{22} \end{aligned}$$

and

$$A_{11} = \sum_{q=1}^{Q} \alpha_q A_{11,q}, \qquad A_{22} = \sum_{q=1}^{Q} \alpha_q A_{22,q}, \\ A'_{11} = \sum_{q=1}^{Q} \alpha'_q A_{11,q}, \qquad A'_{22} = \sum_{q=1}^{Q} \alpha'_q A_{22,q}.$$

When $r$ is large, the discrete sets from which $\alpha_q$s, $\alpha'_q$s are chosen from $(\mathcal{A}_r)$ can be replaced with independent scalar Cauchy distributions. And by noticing that the sum of two independent Cauchy random variables is scaled-Cauchy, our criterion can be simplified to

$$\max_{\{A_{11,q}, A_{22,q}\}, B} \mathrm{E} \log\det(B^*A_{11}B - A_{22})^2 - 2\mathrm{E} \log\det \Delta_2^2. \tag{3.31}$$

### 3.6.4 Frobenius Norm of the Basis Matrices

Entries of $A_{11,q}$s and $A_{22,q}$s in (3.30) are unconstrained other than that they must be Hermitian matrices. However, we found that it is beneficial to constrain the Frobenius norm of all the matrices in $\{A_{11,q}\}$ to be the same, which we denote by $\gamma_1$. This is similarly the case for the matrices $\{A_{22,q}\}$, whose Frobenius norm we denote by $\gamma_2$. In fact, in our experience it is very important, for both the criterion function (3.30) and the ultimate constellation performance, that the correct Frobenius norms of the basis matrices be chosen. The gradients of the Frobenius norms $\gamma_1$ and $\gamma_2$ are given in Section 3.9.2 and the gradient-ascent method is used for the optimization. The matrix $B$ is choosen as $\gamma_3[I_M, 0_{M\times(T-2M)}]$ with $\gamma_3$ close to 1 for the following two reasons. Firstly, the optimization of $B$ is too complicated to be done by the gradient-ascent method. Secondly, as long as $B$ is full rank, simulation shows that the Frobenius norm of $B$ and $B$ itself do not have significant effects on the performance. This has been shown to perform well.

### 3.6.5 Design Summary

We now summarize the design method for Cayley unitary space-time codes with $M$ transmit antennas and $N$ receive antennas, and target rate $R$.

1. Choose $Q \leq \min(T-M, N)\max(2(T-M)-N, T-M)$. Although this inequality is a soft limit for sphere decoding, we choose our $Q$ that obeys the inequality to keep the decoding complexity polynomial.

2. Choose $r$ that satisfies $r^Q = 2^{RT}$. We always choose $r$ to be a power of 2 to simplify the bit allocation and use a standard Gray-code assignment of bits to the symbols in the set $\mathcal{A}_r$.

3. Let $\mathcal{A}_r$ be the $r$-point discretization of the scalar Cauchy distribution obtained

as the image of the function $\alpha = -\tan(\theta/2)$ applied to the set $\{\pi/r,\ 3\pi/r,\ 5\pi/r,\ \cdots,\ (2r-1)\pi/r\}$.

4. Choose $\{A_{11,q}\}$ and $\{A_{22,q}\}$ that solves the optimization problem (3.30). A gradient-ascent method can be used. The computation of the gradients of the criterion in (3.30) is presented in Section 3.9.1. At the end of each iteration, gradient-ascent is used to optimize the Frobenius norms of the basis matrices $A_{11,1}, A_{11,2}, \cdots, A_{11,Q}$ and $A_{22,1}, A_{22,2}, \cdots, A_{22,Q}$. The computation of the gradients is given in Section 3.9.2. Note first that the solution to (3.30) is highly non-unique. Another solution can be obtained by simply reordering $A_{11,q}$s and $A_{22,q}$s. And also, since the criterion function is neither linear nor convex in the design variables $A_{11,q}$ and $A_{22,q}$, there is no guarantee of obtaining a global maximum. However, since the code design is performed off-line and only once, we can use more sophisticated optimization techniques to get a better solution. Simulation results show that the codes obtained by this method have good performance. The number of receive antennas $N$ does not appear explicitly in the criterion (3.30), but it depends on $N$ through the choice of $Q$. Hence, the optimal codes, for a given $M$, are different for different $N$.

## 3.7 Simulation Results

In this section, we give examples of Cayley unitary space-time codes and the simulated performance of the codes for various number of antennas and rates. The fading coefficient between every transmit-and-receive antenna pair is modeled independently as a complex Gaussian variable with zero-mean and unit-variance and is kept constant for $T$ channel uses. At each time, a zero-mean, unit-variance complex Gaussian noise is added to the received signal at every receive antenna. Two error events are demonstrated including block errors, which correspond to errors in decoding the

$T \times M$ matrices $S_1, ..., S_L$, and bit errors, which correspond to errors in decoding $\alpha_1, ..., \alpha_Q$. The bits are allocated to each $\alpha_q$ by a Gray code and therefore, a block error may correspond to only a few bit errors. We first give an example to compare the performance of the linearized ML, which is given by (3.10) with that of the true ML, then performance comparisons of our codes with training-based methods are given.

### 3.7.1 Linearized ML vs. ML

In communications and code designs, the decoding complexity is an important issue. In our problem, when the transmission rate is high, for example, $R = 3$ and $T = 6, M = 3$, for one coherence interval, the true ML decoding involves a search over $2^{RT} = 2^{18} = 262,144$ $6 \times 3$ matrices, which is not practical. This is why we linearize the ML decoding to use the sphere decoding algorithm.

But we need to know what is the penalty of using (3.10) instead of the true ML. Here an example is given for the case of a two transmit, one receive antenna system with coherence interval of four channel uses operating at rate $R = 1.5$ with $Q = 3$ and $r = 2$. The number of signal matrices is $2^{RT} = 64$, for which the true ML is feasible. The resulting bit error rate and block error rate curves for the linearized ML are the line with circles and line with stars in Figure 3.1. The resulting bit error rate and block error rate curves for the the true ML are the solid line and the dashed line in the figure. We can see from Figure 3.1 that the performance loss for the linearized ML decoding is almost neglectable but the computational complexity is saved greatly by using the linearized ML decoding which is implemented by sphere decoding.

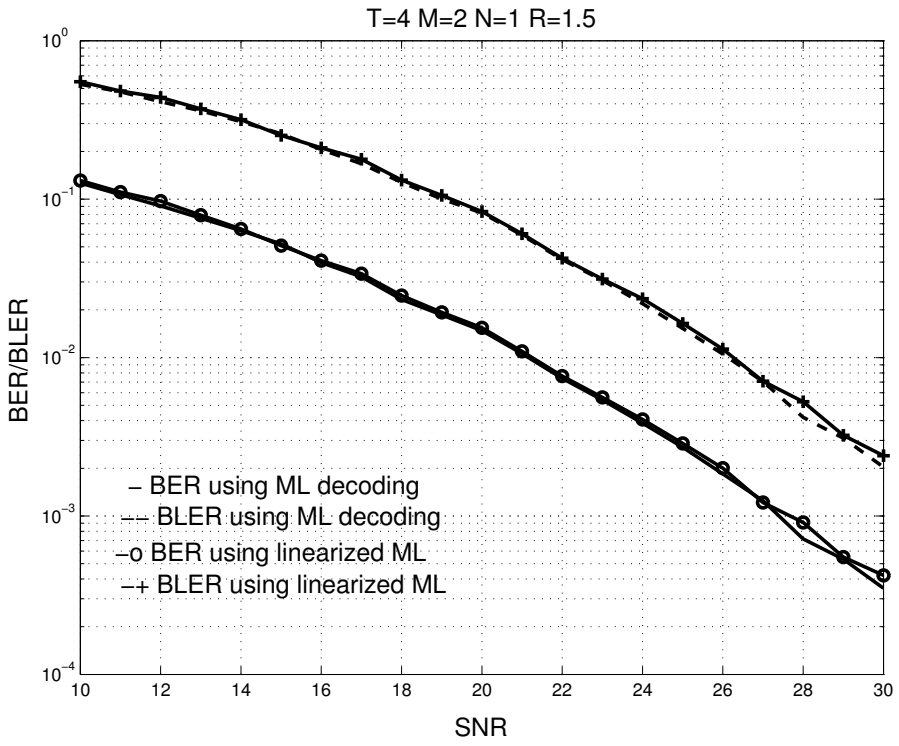


Figure 3.1: $T = 4, M = 2, N = 1, R = 1.5$: BER and BLER of the linearized ML given by (3.10) compared with the true ML

### 3.7.2 Cayley Unitary Space-Time Codes vs. Training-Based Codes

In this subsection a few examples of Cayley codes for various multiple-antenna communication systems are given and their performance are compared with that of training-based codes.

As discussed in Chapter 2, a commonly used scheme for unknown channel multiple-antenna communication systems is to obtain the channel information via training. It is important and meaningful to compare our code with that of training-based codes. Training-based schemes and the optimal way to do training are discussed in Section 2.4 . In our simulations of training-based schemes, the LMMSE estimation is used. We set the training period $T_\tau$ as $M$ and the training signal matrix $S_\tau$ as $\rho_\tau I_M$, which are optimal. For simplicity, we use equal-training-and-data-power by setting $\rho_d = \rho_\tau = \sqrt{M}$, which is optimal if $T = 2M$. In most of the following simulations, different space-time codes are used in the data transmission phase for different system settings. Sphere decoding is used in decoding all the Cayley codes and the decoding of the training-based codes is always ML, but the algorithm varies according to the codes used.

**Example of $T = 4, M = 2, N = 2$**

The first example is for the case of two transmit and two receive antennas with coherence interval $T = 4$. For training-based schemes, half of the coherence interval is used for training. For the data transmission phase, we consider two different space-time codes. The first one is the well-known orthogonal design in which the transmitted data matrix has the following structure:

$$S_d = \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}.$$

By choosing $a$ and $b$ from the signal set of 16-QAM equally likely, the rate of the training-based code is 2 bits per channel use. The same as Cayley codes, bits are allocated to each entry by Gray code. The second one is the LD code proposed in [HH02b]:

$$S_d = \sum_{q=1}^{4}(\alpha_q A_q + i\beta_q B_q), \quad \text{with } \alpha_q, \beta_q \in \{-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\},$$

where

$$A_1 = B_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = B_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$
$$A_3 = B_3 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad A_4 = B_4 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Clearly, the rate of the training-based LD code is also 2. For the Cayley code, from (3.25), we choose $Q = 4$. To attain rate 2, $r = 4$ from (3.27). The Cayley code was obtained by finding a local maximum to (3.31).

The performance curves are shown in Figure 3.2. The dashed line and dashed line with plus signs indicate the BER and BLER of the Cayley code at rate 2, respectively. The solid line and solid line with plus signs indicate the BER and BLER of the training-based orthogonal design at rate 2 respectively and the dash-dotted line and dash-dotted line with plus sighs show the BER and BLER of the training-based LD code at rate 2 respectively. We can see from the figure that the Cayley code underperforms the optimal training-based codes by $3-4$dB. However, our results are preliminary and it is conceivable that better performance may be obtained by further optimization of (3.30) or (3.31).

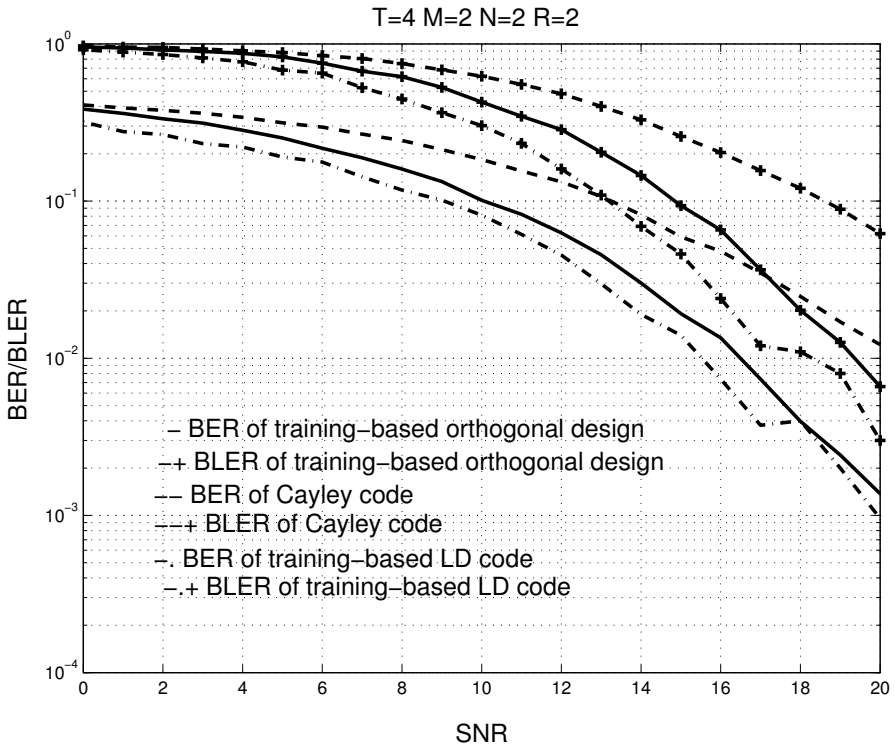


Figure 3.2: $T = 4, M = 2, N = 2, R = 2$: BER and BLER of the Cayley code compared with the training-based orthogonal design and the training-based LD code

**Example of** $T = 5, M = 2, N = 1$

For the training-based scheme of this setting, 2 channel uses of each coherence interval are allocated to training. Therefore, in the data transmission phase, bits are encoded into a $3 \times 2$ data matrix $S_d$. Since we are not aware of any $3 \times 2$ space-time code, we employ an uncoded transmission scheme, where each element of $S_d$ is chosen independently from a BPSK constellation, resulting in rate 6/5. This allows us to compare the Cayley codes with the the uncoded training-based scheme. Two Cayley codes are analyzed here: the Cayley code at rate 1 with $Q = 5, r = 2$ and the Cayley code at rate 2 with $Q = 5, r = 4$.

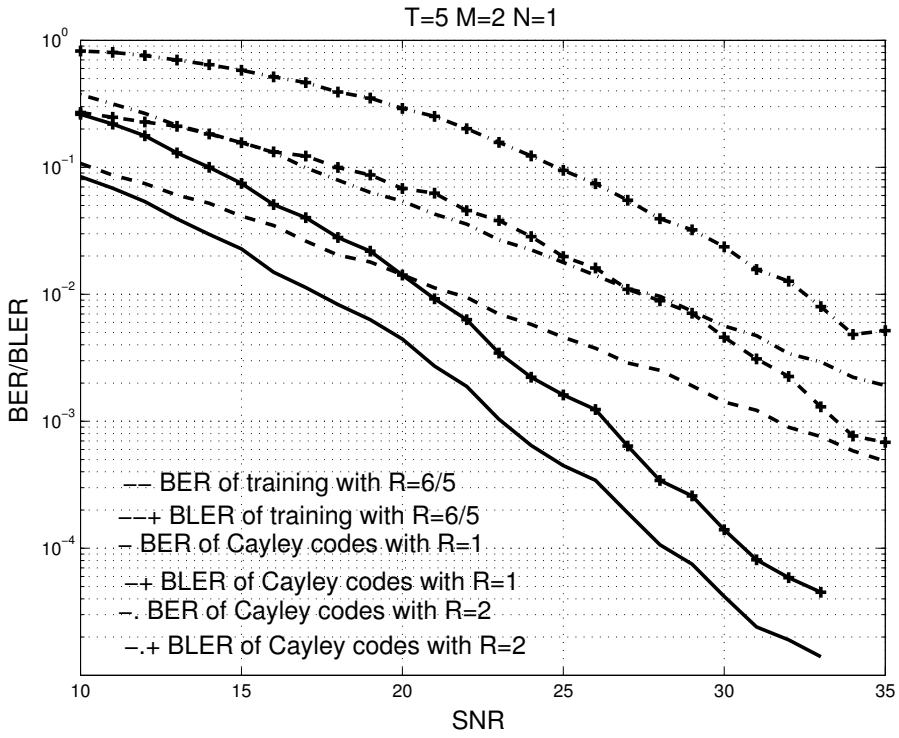


Figure 3.3: $T = 5, M = 2, N = 1$: BER and BLER of the Cayley codes compared with the uncoded training-based scheme

The performance curves are shown in Figure 3.3. The solid line and solid line with plus signs indicate the BER and BLER of the Cayley code at rate 1, respectively, the dash-dotted line and dash-dotted line with plus signs show the BER and BLER

of the Cayley code at rate 2, respectively, and the dashed line and dashed line with plus signs shows the BER and BLER of the training-based scheme, which has a rate of 6/5. Exhaustive search is used in decoding the training-based scheme and sphere decoding is applied to decode the Cayley codes.

We can see that our Cayley code at rate 1 has lower BER and BLER than the training-based scheme at rate 6/5 at any SNR. And, even at a rate which is 4/5 higher, 2 compared with 6/5, the performance of the Cayley code is comparable to that of the training-based scheme when the SNR is as high as 35dB.

**Example of** $T = 7, M = 3, N = 1$

For this system setting, three channel uses of each coherence interval are allocated to training. In the data transmission phase of the training-based scheme, we use the optimized LD code given in [HH02b]:

$$S_d = \begin{bmatrix} \alpha_1 + \alpha_3 + i\left[\frac{\beta_2+\beta_3}{\sqrt{2}} + \beta_4\right] & \frac{\alpha_2-\alpha_4}{\sqrt{2}} - i\left[\frac{\beta_1}{\sqrt{2}} + \frac{\beta_2-\beta_3}{2}\right] & 0 \\ \frac{-\alpha_2+\alpha_4}{\sqrt{2}} - i\left[\frac{\beta_1}{\sqrt{2}} + \frac{\beta_2-\beta_3}{2}\right] & \alpha_1 - i\frac{\beta_2+\beta_3}{\sqrt{2}} & -\frac{\alpha_2+\alpha_4}{\sqrt{2}} + i\left[\frac{\beta_1}{\sqrt{2}} - \frac{\beta_2-\beta_3}{2}\right] \\ 0 & \frac{\alpha_2+\alpha_4}{\sqrt{2}} + i\left[\frac{\beta_1}{\sqrt{2}} - \frac{\beta_2-\beta_3}{2}\right] & \alpha_1 - \alpha_3 + i\left[\frac{\beta_2+\beta_3}{\sqrt{2}} - \beta_4\right] \\ \frac{\alpha_2-\alpha_4}{\sqrt{2}} + i\left[\frac{\beta_1}{\sqrt{2}} + \frac{\beta_2-\beta_3}{2}\right] & -\alpha_3 + i\beta_4 & -\frac{\alpha_2+\alpha_4}{\sqrt{2}} + i\left[\frac{\beta_1}{\sqrt{2}} - \frac{\beta_2-\beta_3}{2}\right] \end{bmatrix}.$$

By setting $\alpha_i, \beta_i$ as BPSK, we obtain a LD code at rate 8/7. For the Cayley code, we choose $Q = 7$ and $r = 2$ and the rate of the code is 1.

The performance curves are shown in Figure 3.4. The solid line and solid line with plus signs indicate the BER and BLER of the Cayley code at rate 1 respectively and the dashed line and dashed line with plus signs show the BER and BLER of the training-based LD code, which has a rate of 8/7. Sphere decoding is applied in the decoding of both codes. From the figure we can see that the performance of the Cayley code is close to the performance of the training-based LD code. Therefore, at a rate 1/7 lower, the Cayley code is comparable with the training-based LD code.

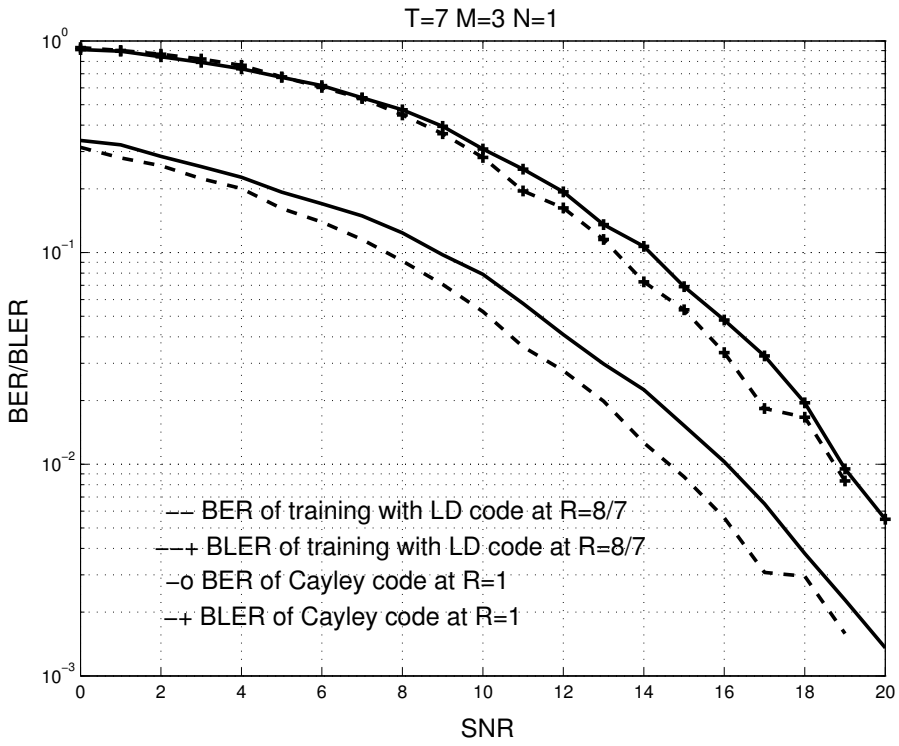


Figure 3.4: $T = 7, M = 3, N = 1$: BER and BLER of the Cayley code compared with the training-based LD code

Again, our results are preliminary and further optimization of (3.30) or (3.31) may yield improved performance.

## 3.8 Conclusion

Cayley unitary space-time codes are proposed in this chapter. The codes do not require channel knowledge at either the transmitter or the receiver, are simple to encode and decode, and apply to systems with any combination of transmit and receive antennas. They are designed with a probabilistic criterion: they maximize the expected log-determinant of the difference between matrix pairs. Cayley transform is used to construct the codes because it maps the nonlinear Stiefel manifold of unitary matrices to the linear space of Hermitian matrices. The transmitted data is broken into sub-streams $\alpha_1, ..., \alpha_Q$, then linearly encoded in the Cayley transform domain. We showed that by constraining $A_{12} = (I + iA_{11})B$ and ignoring the data dependence of the additive noise, $\alpha_1, ..., \alpha_Q$ appear linearly at the receiver. Therefore, linear decoding algorithms such as sphere decoding and nulling-and-canceling can be used whose complexity is polynomial in the rate and dimension. Our code has a similar structure to training-based schemes after transformations.

The recipe for designing Cayley unitary space-time codes for any combination of transmit/receive antennas and coherence intervals is given and also, simulation examples are shown to compare our Cayley codes with optimized training-based space-time codes and uncoded training-based schemes for different system settings. Our simulation results are preliminary, but indicate that the Cayley codes generated with this recipe only slightly underperform optimized training-based schemes using orthogonal designs and LD codes. However, they are clearly superior to uncoded training-based space-time schemes. Further optimization of the Cayley code basis matrices (in (3.30) or (3.31)) is necessary for a complete comparison of the performance with training-

based schemes.

## 3.9 Appendices

### 3.9.1 Gradient of Criterion (3.30)

In the simulations, the maximization of the design criterion function (3.30) is performed using a simple gradient-ascent method. In this section, we compute the gradient of (3.30) that is required in this method.

We are interested in the gradient with respect to the matrices $A_{11,1}, ..., A_{11,Q}$ and $A_{22,1}, ..., A_{22,Q}$ of the design function (3.30), which is equivalent to

$$\max_{\{A_{11,q},A_{22,q}\},B} \mathrm{E}\,\log\det[B^*(A_{11}-A'_{11})B-(A_{22}-A'_{22})]^2 - 2\mathrm{E}\,\log\det\Delta_2^2. \quad (3.32)$$

To compute the gradient of a real function $f(A_q)$ with respect to the entries of the Hermitian matrix $A_q$, we use the formulas

$$\left[\frac{\partial f(A_q)}{\partial \Re A_q}\right]_{j,k} = \min_{\delta\to 0}\frac{1}{\delta}[f(A_q+\delta(e_je_k^t+e_ke_j^t))-f(A_q)], j\neq k, \quad (3.33)$$

$$\left[\frac{\partial f(A_q)}{\partial \Im A_q}\right]_{j,k} = \min_{\delta\to 0}\frac{1}{\delta}[f(A_q+i\delta(e_je_k^t-e_ke_j^t))-f(A_q)], j\neq k, \quad (3.34)$$

$$\left[\frac{\partial f(A_q)}{\partial A_q}\right]_{j,j} = \min_{\delta\to 0}\frac{1}{\delta}[f(A_q+\delta e_je_j^t)-f(A_q)], \quad (3.35)$$

where $e_j$ is the unit column vector of the same dimension of columns of $A_q$ which has a one in the $j$-th entry and zeros elsewhere. That is, while calculating the gradient with respect to $A_{11,q}$, $e_j$ should have dimension $M$ and for the gradient with respect to $A_{22,q}$, the dimension should be $T-M$ instead.

First, note that $A_{11}-A'_{11}=\sum_{q=1}^{Q}A_{11,q}a_q$ where $a_q=\alpha_q-\alpha'_q$ and similarly, $A_{22}-A'_{22}=\sum_{q=1}^{Q}A_{22,q}a_q$. Therefore, to apply (3.33) to the first term of (3.32) with respect

to $A_{11,q}$, let $H = B^*(A_{11} - A'_{11})B - (A_{22} - A'_{22})$. Therefore,

$$\begin{aligned}
& \log\det[B^*(A_{11} - A'_{11})B - (A_{22} - A'_{22}) + B^*(e_j e_k^t + e_k e_j^t)B\delta a_q]^2 \\
= \;& \log\det\{H^2 + [HB^*(e_j e_k^t + e_k e_j^t)B + B^*(e_j e_k^t + e_k e_j^t)BH]\delta a_q + o(\delta^2)I\} \\
= \;& \log\det H^2 + \log\det\{I + H^{-2}[HB^*(e_j e_k^t + e_k e_j^t)B + B^*(e_j e_k^t + e_k e_j^t)BH]\delta a_q + o(\delta^2)I\} \\
= \;& \log\det H^2 + \mathrm{tr}\,\{H^{-2}[HB^*(e_j e_k^t + e_k e_j^t)B + B^*(e_j e_k^t + e_k e_j^t)BH]\delta a_q\} + o(\delta^2) \\
= \;& \log\det H^2 + \mathrm{tr}\,\{H^{-1}B^*(e_j e_k^t + e_k e_j^t)B + H^{-2}B^*(e_j e_k^t + e_k e_j^t)BH\delta a_q\} + o(\delta^2) \\
= \;& \log\det H^2 + \mathrm{tr}\,\{BH^{-1}B^*(e_j e_k^t + e_k e_j^t) + BH^{-1}B^*(e_j e_k^t + e_k e_j^t)\delta a_q\} + o(\delta^2) \\
= \;& \log\det H^2 + (2\{BH^{-1}B^*\}_{k,j} + 2\{BH^{-1}B^*\}_{j,k})a_q + o(\delta^2) \\
= \;& \log\det H^2 + 4\Re\{BH^{-1}B^*\}_{j,k}a_q + o(\delta^2),
\end{aligned}$$

where $\{A\}_{i,j}$ indicates the $(i,j)$-th entry of matrix $A$ and $\Re\{A\}_{i,j}$ indicates the real part of the $(i,j)$-th entry of matrix $A$. We use $\mathrm{tr}\,AB = \mathrm{tr}\,BA$ and the last equality follows because $BH^{-1}B^*$ is Hermitian. We may now apply (3.33) to obtain

$$\left[\frac{\partial \log\det[B^*(A_{11} - A'_{11})B - (A_{22} - A'_{22})]^2}{\partial \Re A_{11,q}}\right]_{j,k} = 4\mathrm{E}\,\Re\{BH^{-1}B^*\}_{j,k}a_q, \quad j \neq k.$$

The gradient with respect to the imaginary components of $A_{11,q}$ can be obtained in a similar way as the following

$$\left[\frac{\partial \log\det[B^*(A_{11} - A'_{11})B - (A_{22} - A'_{22})]^2}{\partial \Im A_{11,q}}\right]_{j,k} = 4\mathrm{E}\,\Im\{BH^{-1}B^*\}_{j,k}a_q, \quad j \neq k,$$

where $\Im\{A\}_{i,j}$ indicates the imaginary part of the $(i,j)$-th entry of matrix $A$. And the gradient with respect to the diagonal elements is

$$\left[\frac{\partial \log\det[B^*(A_{11} - A'_{11})B - (A_{22} - A'_{22})]^2}{\partial A_{11,q}}\right]_{j,j} = 2\mathrm{E}\,\{BH^{-1}B^*\}_{j,j}a_q.$$

Similarly, we get the gradient with respect to $A_{22,q}$,

$$\begin{aligned}
\left[\frac{\partial \log\det[B^*(A_{11}-A'_{11})B-(A_{22}-A'_{22})]^2}{\partial \Re A_{22,q}}\right]_{j,k} &= -4\mathrm{E}\,\Re H^{-1}_{j,k}a_q, \quad j\neq k,\\
\left[\frac{\partial \log\det[B^*(A_{11}-A'_{11})B-(A_{22}-A'_{22})]^2}{\partial \Im A_{22,q}}\right]_{j,k} &= -4\mathrm{E}\,\Im H^{-1}_{j,k}a_q, \quad j\neq k,\\
\left[\frac{\partial \log\det[B^*(A_{11}-A'_{11})B-(A_{22}-A'_{22})]^2}{\partial A_{22,q}}\right]_{j,j} &= -2\mathrm{E}\,H^{-1}_{j,j}a_q.
\end{aligned}$$

For the second term, by using the same method, the following results are obtained

$$\begin{aligned}
\left[\frac{\partial \log\det\Delta_2^2}{\partial \Re A_{11,q}}\right]_{j,k} &= 2\mathrm{E}\,\Re(D+D^*+E+E^*)_{j,k}\alpha_q, \quad j\neq k,\\
\left[\frac{\partial \log\det\Delta_2^2}{\partial \Im A_{11,q}}\right]_{j,k} &= 2\mathrm{E}\,\Im(D+D^*+E+E^*)_{j,k}\alpha_q, \quad j\neq k,\\
\left[\frac{\partial \log\det\Delta_2^2}{\partial A_{11,q}}\right]_{j,j} &= 2\mathrm{E}\,(D+E)_{j,j}\alpha_q,\\
\left[\frac{\partial \log\det\Delta_2^2}{\partial \Re A_{22,q}}\right]_{j,k} &= 2\mathrm{E}\,\Re(F+F^*+G+G^*)_{j,k}\alpha_q, \quad j\neq k,\\
\left[\frac{\partial \log\det\Delta_2^2}{\partial \Im A_{22,q}}\right]_{j,k} &= 2\mathrm{E}\,\Im(F+F^*+G+G^*)_{j,k}\alpha_q, \quad j\neq k,\\
\left[\frac{\partial \log\det\Delta_2^2}{\partial A_{22,q}}\right]_{j,j} &= 2\mathrm{E}\,(F+G)_{j,j}\alpha_a,
\end{aligned}$$

where

$$\begin{aligned}
D &= iB\Delta_2^{-2}(I+iA_{22})B^*,\\
E &= iB\Delta_2^{-2}B^*(I-iA_{11})BB^*,\\
F &= \Delta_2^{-2}A_{22},\\
G &= iB^*(I+iA_{11})B\Delta_2^{-2},
\end{aligned}$$

and all the expectations are over all possible $\alpha_1, ..., \alpha_Q$.

### 3.9.2 Gradient of Frobenius Norms of the Basis Sets

Let $\gamma_1$ be a multiplicative factor that we use to multiple every $A_{11,q}$ and $\gamma_2$ a multiplicative factor that we use to multiple every $A_{22,q}$. Thus, $\gamma_1$ and $\gamma_2$ are the Frobenius norms of matrices in $\{A_{11,q}\}$ and $\{A_{22,q}\}$. We solve for the optimal $\gamma_1, \gamma_2 > 0$ by maximizing the criterion function in (3.30), which is

$$\xi(\gamma_1, \gamma_2) = \mathrm{E}\,\log\det[\gamma_1 B^*(A_{11} - A'_{11})B - \gamma_2(A_{22} - A'_{22})]^2 - 2\mathrm{E}\,\log\det\Delta_2^2,$$

where

$$\Delta_2 = I + B^*B - i\gamma_1 B^* \sum_{q=1}^{Q} \alpha_q A_{11,q} B + i\gamma_2 \sum_{q=1}^{Q_1} \alpha_q A_{22,q}.$$

As in the optimization of $A_{11,q}, A_{22,q}$, gradient-ascent method is used. To compute the gradient of a real function $f(x_1, x_2)$ with respect to $x_1$ and $x_2$, we use the formulas

$$\begin{aligned} \frac{\partial f(x_1, x_2)}{\partial x_1} &= \lim_{\delta \to 0} \frac{1}{\delta}[f(x_1 + \delta, x_2) - f(x_1, x_2)], \\ \frac{\partial f(x_1, x_2)}{\partial x_2} &= \lim_{\delta \to 0} \frac{1}{\delta}[f(x_1, x_2 + \delta) - f(x_1, x_2)]. \end{aligned}$$

And the results are

$$\begin{aligned} &\frac{\partial \xi(\gamma_1, \gamma_2)}{\partial \gamma_1} \\ = \; & -2\mathrm{E}\,\mathrm{tr}\,\{f^{-1}[2\gamma_1 B^* A_{11} B B^* A_{11} B + iB^*(BB^* A_{11} - A_{11} BB^*)B \\ & -\gamma_2(A_{22} B^* A_{11} B + A_{11} B A_{22} B^*)]\} + \mathrm{E}\,\mathrm{tr}\,[g^{-1}(2\gamma_1 B^*(A_{11} - A'_{11})BB^*(A_{11} - A'_{11})B \\ & -\gamma_2((A_{22} - A'_{22})B^*(A_{11} - A'_{11})B + (A_{11} - A'_{11})B(A_{22} - A'_{22})B^*))] \end{aligned}$$

and

$$\begin{aligned}
&\frac{\partial \xi(\gamma_1,\gamma_2)}{\partial \gamma_2} \\
= \;& -2\mathrm{E}\,\mathrm{tr}\,[f^{-1}(2\gamma_1 A_{22}^2 - i(B^*BA_{22} + A_{22}BB^*) - \gamma_1(A_{22}B^*A_{11}B + A_{11}BA_{22}B^*))] \\
& \mathrm{E}\,\mathrm{tr}\,[g^{-1}(2\gamma_2 A_{22}^2 - \gamma_1((A_{22}-A'_{22})B^*(A_{11}-A'_{11})B + (A_{11}-A'_{11})B(A_{22}-A'_{22})B^*))].
\end{aligned}$$

Simulation shows that good performance is obtained when $\gamma_1$ and $\gamma_2$ are not too far away from unity.

# Chapter 4 Groups and Representation Theory

## 4.1 Advantages of Group Structure

Another interesting space-time codes design scheme is the group codes proposed originally in [SHHS01, Hug00b, HK00], in which the set of matrices, which are space-time codes, forms a group. The motivation of group-based codes is as the following.

As discussed in Section 2.6, the space-time code design problem for differential unitary space-time modulation, which is well-tailored for the unknown-channel case, is: given the number of transmitter antennas $M$ and the transmission rate $R$, find a set $\mathcal{C}$ of $L = 2^{MR}$ $M \times M$ unitary matrices, such that the diversity product

$$\xi_{\mathcal{C}} = \frac{1}{2} \min_{S_i \neq S_{i'} \in \mathcal{C}} |\det(S_i - S_{i'})|^{\frac{1}{M}} \tag{4.1}$$

is as large as possible. This design problem is very difficult to solve because of the following reasons. First, it is easy to see that the objective function of the code design problem, given in (4.1), is not convex. Second, the constraint space, which is the space of unitary matrices, is also non-convex. Furthermore, when our desired rate of transmission $R$ is not very low, the constellation size $L = 2^{MR}$ can be huge, which make the problem even more difficult according to computation complexity. For example, if there are four transmit antennas and we want to transmit at a rate of four bits per channel use, we need to find a set of $L = 2^{16} = 65,536$ unitary matrices whose minimum value of the determinants of the pairwise difference matrices is maximized. Therefore, it appears that there is no efficient algorithm with tractable computational

complexity to find the exact solution to this problem. To simplify the design problem, it is necessary to introduce some structure to the constellation set $\mathcal{C}$. Group structure turns our to be a very good one.

**Definition 4.1 (Group).** *[DF99] A* group *is an ordered pair $(G, \star)$, where $G$ is a set and $\star$ is a binary operation on $G$ satisfying the following axioms.*

1. *$(a \star b) \star c = a \star (b \star c)$, for all $a, b, c \in G$, i.e., $\star$ is associative.*
2. *There exists an element $e$ in $G$, called the* identity *of $G$, such that for all $a \in G$, $a \star e = e \star a = a$.*
3. *For each $a \in G$, there exists an element $a^{-1}$ of $G$, called an* inverse *of $a$, such that $a \star a^{-1} = a^{-1} \star a = e$.*

Normally, the binary operation is called *multiplication*. In brief, a group is a set of elements that is closed under both multiplication and inversion.

The *order* of a finite set is simply the number of elements in it. A subset $H \subset G$ is a *subgroup* if it is closed under the group multiplication and $h^{-1} \in H$ for all $h \in H$.

**Example 1.** The set of integers $\mathbb{Z}$, rational numbers $\mathbb{Q}$, real numbers $\mathbb{R}$, and complex numbers $\mathbb{C}$ are all groups under the operation $+$ with $e = 0$ and $a^{-1} = -a$.

**Example 2.** $GL_n(\mathbb{C})$, the set of all invertible $n \times n$ matrices with entries in $\mathbb{C}$, is a group under the operation of matrix multiplication with $e = I_n$ and $a^{-1}$ the inverse matrix of $a$. Note that the matrix multiplication operation is not commutative, that is, $ab = ba$ is not true in general.

**Example 3.** $U(n)$, the set of all unitary $n \times n$ matrices, is a group under the operation of matrix multiplication with $e = I_n$ and $a^{-1}$ the inverse matrix of $a$. It is a *subgroup* of the group $GL_n(\mathbb{C})$.

Now we are going to discuss the advantages of group structure in the space-time code design problem. In general, for two arbitrary elements $A$ and $B$ in a set, $\mathcal{C}$, with

cardinality $L$, $|\det(A-B)|$ takes on $L(L-1)/2$ distinct values if $A \neq B$. Therefore, when $L$ is large, diversity product of the set may be quite small. If $\mathcal{C}$ forms a group, for any two matrices $A$ and $B$ in the set, there exists a matrix $C \in \mathcal{C}$ such that $C = A^{-1}B$. Therefore,

$$|\det(A-B)| = |\det A||\det(I_M - A^{-1}B)| = |\det(I_M - A^{-1}B)| = |\det(I_M - C)|,$$

which takes on at most $L-1$ distinct values as $I_M \neq C$ if $A \neq B$. Therefore, the chance of having a large diversity product is greatly increased. Another advantage is that the calculations of both the diversity product and the transmission matrix are greatly simplified. From the above formula, it can be easily seen that the complexity of calculating the diversity product is reduced dramatically. In general, to calculate the diversity product of a set with $L$ elements, $L(L-1)/2$ calculations of matrix determinants are needed, which is quadratic in $L$. However, if the set forms a group, it has been shown that only $L-1$ calculations of matrix determinants are needed, which is linear in $L$. For the calculation of the transmission signal matrix, generally, the multiplication of two $M \times M$ matrices is needed. If the set $\mathcal{C}$ is a group, the product is also in the group, which means that every transmission matrix is an element of $\mathcal{C}$. Therefore, explicit matrix multiplication can be replaced by simpler group table-lookup. Finally, note that $\det(I-C) = 0$ if and only if $C$ has a unit eigenvalue. Another important advantage of group structure is that if the set forms a group, its full diversity has a practical meaning: its non-identity elements have no unit eigenvalue, which gives a possible method to design fully diverse codes.

There are a lot of well-known group-based codes. Two examples are given in the following.

**Example 1. Cyclic group-based codes**. [Hug00b] [SHHS01]

$$\mathcal{S} = \{I_M, S, S^2, \cdots, S^{L-1}\},$$

where

$$S = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & e^{2\pi j \frac{l_1}{L}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{2\pi j \frac{l_{M-1}}{L}} \end{bmatrix}$$

with $l_1, \cdots, l_{M-1} \in [0, L-1]$. The code is called cyclic code since any element in the set is a power of $S$ and $S^L = I_M$. It is easy to check that $\mathcal{S}$ forms a group.

**Example 2. Alamouti's orthogonal design**. [Ala98]

$$\mathcal{S} = \left\{ \frac{1}{\sqrt{|x|^2 + |y|^2}} \begin{bmatrix} x & y \\ -y^* & x^* \end{bmatrix} \middle| x, y \in \mathbb{C} \right\}.$$

It is easy to check that $\mathcal{S}$ forms a group. Actually it is the Lie group $SU(2)$: the set of $2 \times 2$ unitary matrix with unit determinant. This group has infinitly many elements. To get a finite constellation, $x$ and $y$ can be chosen from some finite sets $\mathcal{S}_1$ and $\mathcal{S}_2$, for example PSK or QAM signals[1].

## 4.2 Introduction to Groups and Representations

Before getting into more details about the design of space-time group codes, some group theory and representation theory are reviewed in this section that are needed later.

[1]By these choices, the resulted sets might not form a group anymore.

**Definition 4.2.** *A subgroup $H$ of $G$ is a* **normal subgroup** *if*

$$ghg^{-1} \in H \quad \text{for all } g \in G, h \in H.$$

**Definition 4.3.** *If $G$ is a group, the* **center** *of $G$ is the set of elements in $G$ that commutes with all other elements of $G$.*

Since group is an abstract concept, normally representation theory is used to map abstract groups to subsets of matrices.

**Definition 4.4.** *Let $(G, \star)$ and $(H, \diamond)$ be groups. A map $\phi : G \to H$ is a group* **homomorphism** *if*

$$\phi(x \star y) = \phi(x) \diamond \phi(y) \quad \text{for all } x, y \in G.$$

**Definition 4.5.** *Let $G$ be a group. $F$ be a field and $V$ a vector space over $F$.*[2]

1. *A* **linear representation** *of $G$ is any homomorphism from $G$ to $GL(V)$. The* **degree** *of the representation is the dimension of $V$.*
2. *A* **matrix representation** *of $G$ is any homomorphism from $G$ into $GL_n(\mathbb{C})$.*
3. *A linear or matrix representation is* **faithful** *if it is injective.*

**Definition 4.6.** *Two matrix representations $\phi$ and $\psi$ of $G$ are* **equivalent** *if there is a fixed invertible matrix $P$ such that*

$$P\phi(g)P^{-1} = \psi(g) \quad \text{for all } g \in G.$$

**Definition 4.7.** *The* **direct sum** *$\phi \oplus \psi$ of two representations $\phi, \psi$ of $G$ with degrees*

[2]For definitions of field and vector space, repfer to [DF99].

*$d_1$ and $d_2$ respectively is the degree $d_1 + d_2$ representation of $G$ such that*

$$(\phi \oplus \psi)(g) = \begin{bmatrix} \phi(g) & 0 \\ 0 & \psi(g) \end{bmatrix} \quad \textit{for all } g \in G.$$

**Definition 4.8.** *A matrix representation is called* **reducible** *if it is equivalent to a direct sum of two (or more) representations. Otherwise, it is called* **irreducible**.

As discussed in the previous section, if our signal set has a group structure under matrix multiplication, its diversity product can be simplified to

$$\xi_{\mathcal{C}} = \frac{1}{2} \min_{I \neq V \in \mathcal{C}} |\det(I - V)|^{\frac{1}{M}}.$$

If we insist on a fully diverse constellation, which means that $\xi_{\mathcal{C}} \neq 0$, then from the above equalities, the eigenvalues of all non-identity elements in the constellation must be different from one. This leads to the following definition.

**Definition 4.9 (Fix-point-free group).** *[HK00] A group $G$ is called* **fixed-point-free (fpf)** *if and only if it has a faithful representation as unitary matrices with the property that the representation of each non-unit element of the group has no eigenvalue at unity.*

Note that the above definition does not require that in *every* representation of the group, non-unit elements have no eigenvalue at unity, but rather that there exists one representation with this propriety. This is because any non-faithful representation of a group cannot be fpf. The reason that fpf groups have been defined as those for which the representation of each non-unit element in the group, rather than each non-identity matrix in the representation, has no eigenvalue at unity is that had we not done so, all groups would have been fpf if all elements in the groups are represented as the identity matrix. For more information on groups, see [DF99] and [Hun74].

## 4.3 Constellations Based on Finite Fixed-Point-Free Groups

To get space-time codes with good performance, fpf groups have been widely investigated. Shokrollahi, Hassibi, Hochwald and Sweldens classified all finite fpf groups in their magnificent paper [SHHS01] based on Zassenhaus's work [Zas36] in 1936. There are only six types of finite fpf groups as given in the following. Before giving the big theorem, we first introduce a definition.

**Definition 4.10.** *Given a pair of integers $(m, r)$, define $n$ as the smallest positive integer such that $r^n = 1 \mod m$. Define $t = m/\gcd(r-1, m)$. The pair $(m, r)$ is called* **admissible** *if $\gcd(n, t) = 1$.*

**Theorem 4.1 (Classification of finite fixed-point-free groups).** *[SHHS01] A finite group is fixed-point-free if and only if it is isomorphic to one of the following six types of abstract groups.*

1. $$G_{m,r} = \left\langle \sigma, \tau | \sigma^m = 1, \tau^n = \sigma^t, \sigma^\tau = \sigma^r \right\rangle,$$

   *where $(m, r)$ is admissible. The order of $G_{m,r}$ is $mn$.*

2. $$D_{m,r,l} = \left\langle \sigma, \tau, \gamma | \sigma^m = 1, \tau^n = \sigma^t, \sigma^\tau = \sigma^r, \sigma^\gamma = \sigma^l, \tau^\gamma = \tau^l, \gamma^2 = \tau^{nr_0/2} \right\rangle,$$

   *where $nr_0$ is even, $(m, r)$ is admissible, $l^2 = 1 \mod m$, $l = 1 \mod n$, and $l = -1 \mod s$ with $s$ the highest power of 2 divising $mn$. The order of $D_{m,r,l}$ is $2mn$.*

*3.*

$$E_{m,r} = \Big\langle \sigma, \tau, \mu, \gamma | \sigma^m = 1, \tau^n = \sigma^t, \sigma^\tau = \sigma^r, \mu^{\sigma^{m/t}} = \mu, \gamma^{\sigma^{m/t}} = \gamma, \mu^4 = 1, \mu^2 = \gamma^2, \mu^\gamma = \mu^{-1}, \mu^\tau = \gamma, \gamma^\tau = \mu\gamma \Big\rangle,$$

*where* $(m, r)$ *is admissible,* $mn$ *is odd, and* $nr_0$ *is divisible by* 3*. The order of* $E_{m,r}$ *is* $8mn$*.*

*4.*

$$F_{m,r,l} = \Big\langle \sigma, \tau, \mu, \gamma, \nu | \sigma^m = 1, \tau^n = \sigma^t, \sigma^\tau = \sigma^r, \mu^{\sigma^{m/t}} = \mu, \gamma^{\sigma^{m/t}} = \gamma, \mu^\tau = \gamma, \gamma^\tau = \mu\gamma, \mu^4 = 1, \mu^2 = \gamma^2, \mu^\gamma = \mu^{-1}, \nu^2 = \mu^2, \sigma^\nu = \sigma^l, \tau^\nu = \tau^l, \mu^\nu = \gamma^{-1}, \gamma^\nu = \mu^{-1} \Big\rangle,$$

*where* $(m, r)$ *is admissible,* $mn$ *is odd,* $r$ *is divisible by* 3*,* $n$ *is not divisible by* 3*,* $l^2 = 1 \mod m$*,* $l = 1 \mod n$*, and* $l = -1 \mod 3$*. The order of* $F_{m,r,l}$ *is* $16mn$*.*

*5.*

$$J_{m,r} = SL_2(\mathbf{F}_5) \times G_{m,r}$$

*where* $(m, r)$ *is admissible,* $\gcd(mn, 120) = 1$*.* $SL_2(\mathbf{F}_5)$ *is the group of* $2 \times 2$ *matrices with elements in the finite Galois field [GS79]* $\mathbf{F}_5$ *and determinant* 1*. It can also be defined as the following abstract group,*

$$SL_2(\mathbf{F}_5) = \left\langle \mu, \gamma | \mu^2 = \gamma^3 = (\mu\gamma)^5, \mu^4 = 1 \right\rangle.$$

*The order of* $J_{m,r}$ *is* $120mn$*.*

*6.*

$$K_{m,n,l} = \langle J_{m,r}, \nu \rangle$$

*with relations*

$$\nu^2 = \mu^2, \mu^\nu = (\mu\gamma)^7(\gamma\mu)^2\gamma(\gamma\mu)^2, \gamma^\nu = \gamma, \sigma^\nu = \sigma^l, \tau^\nu = \tau^l,$$

*where* $\mu, \gamma$ *are as in* $J_{m,r}$, $l^2 = 1 \mod m$, *and* $l = 1 \mod n$. *The order of* $K_{m,r,l}$ *is* $240mn$.

Some unitary representations of these abstract groups are also given in [SHHS01]. Constellations based on representations of these finite fpf groups give space-time codes with amazingly good performances at low to moderate rates, for example, $SL_2(\mathbf{F}_5)$ , which works for systems with two transmit antennas at rate 3.45. In that paper, the authors also give non-group constellations which are generalizations of some finite groups and products of group representations.

## 4.4 Introduction to Lie Groups and Lie Algebras

As shown in [SHHS01], these finite fpf groups are few and far between. There are only six types of them and unitary representations of them have dimension and rate constraints. Although very good constellations are obtained for low to moderate rates, no good constellations are obtained for very high rates from these finite groups. This motivates the search for *infinite* fpf groups, in particular, their most interesting case, Lie groups.

**Definition 4.11 (Lie group).** *[BtD95] A* **Lie group** *is a differential manifold which is also a group such that the group multiplication and inversion maps are differential.*

The above definition gives us the main reason for studying Lie groups. Since Lie groups have an underlying manifold strcuture, finite constellations, which are subsets of infinite Lie groups, can be obtained by sampling the underlying continuous manifold appropriately.

**Definition 4.12 (Lie algebra).** *[SW86] A* **Lie algebra** $\mathfrak{g}$ *is a vector space over a field* $F$ *on which a product* $[,]$*, called the Lie bracket, is defined, which satisfies*

1. $X, Y \in \mathfrak{g}$ *implies* $[X, Y] \in \mathfrak{g}$,
2. $[X, \alpha Y + \beta Z] = \alpha[X, Y] + \beta[X, Z]$ *for* $\alpha, \beta \in F$ *and* $X, Y, Z \in \mathfrak{g}$,
3. $[X, Y] = -[Y, X]$,
4. $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

It turns out that there is a close connection between Lie groups and Lie algebras.

**Theorem 4.2 (Connection between Lie group and Lie algebra).** *[SW86] Let* $G$ *be a Lie group of matrices. Then* $\mathfrak{g}$*, the set of tangent vectors to all curves in* $G$ *at the identity, is a Lie algebra. Let* $\mathfrak{g}$ *be a linear algebra generated by the basis* $g_1, \cdots, g_n$*, then* $g(\theta) = e^{\theta_1 g_1 + \cdots + \theta_n g_n}$ *is a local Lie group for small enough* $\theta$*.*

Therefore, to obtain many, if not most, of its properties, one can study the Lie algebra, rather than the Lie group itself. Lie algebras are easier to be analyzed because they are vector spaces with good properties.

**Example 1.** $GL(n, \mathbb{C})$ is the Lie group of non-singular $n \times n$ complex matrices. Its Lie algebra is the space of $n \times n$ complex matrices.

**Example 2.** $SL(n, \mathbb{C})$ is the Lie group of unit-determinant non-singular $n \times n$ complex matrices. Its Lie algebra is the space of $n \times n$ traceless matrices.

**Example 3.** $U(n)$ is the Lie group of $n \times n$ complex unitary matrices. Its Lie algebra is the space of $n \times n$ skew-Hermitian matrices.

**Example 4.** $SU(n)$ is the Lie group of unit-determinant $n \times n$ unitary matrices. Its Lie algebra is the space of $n \times n$ traceless, skew-Hermitian matrices.

**Definition 4.13.** *A Lie sub-algebra $\mathfrak{h} \subset \mathfrak{g}$ is an* **ideal** *if it satisfies the condition*

$$[X, Y] \in \mathfrak{h} \quad \textit{for all } X \in \mathfrak{h}, Y \in \mathfrak{g}.$$

**Definition 4.14.** *A Lie algebra $\mathfrak{g}$ is* **simple** *if $\dim \mathfrak{g} > 1$ and it contains no nontrivial ideals. Or equivalently, the Lie group $G$ of the Lie algebra has no nontrivial normal Lie subgroups.*

**Definition 4.15.** *Define the series $\{\mathcal{D}^k\mathfrak{g}\}$ inductively by*

$$\mathcal{D}^1\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$$

*and*

$$\mathcal{D}^k\mathfrak{g} = [\mathcal{D}^{k-1}\mathfrak{g}, \mathcal{D}^{k-1}\mathfrak{g}].$$

*$\mathfrak{g}$ is* **solvable** *if $\mathcal{D}^k\mathfrak{g} = 0$ for some $k$.*

**Definition 4.16.** *A Lie algebra $\mathfrak{g}$ is* **semi-simple** *if $\mathfrak{g}$ has no nonzero solvable ideals.*

The *rank* of a Lie algebra $\mathfrak{g}$ equals the maximum number of commuting basis elements it has or the dimension of a maximum Abelian subgroup[3] of the correspondence Lie group $G$.

It is proved in [SHHS01] that any representation of a finite group is equivalent to a representation using only unitary matrices. However, this is not true for infinite groups and Lie groups.

**Theorem 4.3 (Lie groups with unitary representations).** *[HK00] A Lie group has a representation as unitary matrices if and only of its Lie algebra is a compact*

[3]If for any two elements $f$ and $g$ in a group $G$, $f \star g = g \star f$, $G$ is an Abelian group.

*semi-simple Lie algebra or the direct sum of* $u(1)$ *and a compact semi-simple Lie algebra.*

For more on the definition of semi-simple and simple Lie algebras, see [HK00, BtD95, Ser92].

## 4.5 Rank 2 Compact Simple Lie Groups

To design differential unitary space-time codes with good performance, two conditions must be satisfied: the unitarity and the full diversity. For unitarity, from Theorem 4.3, to get unitary constellations, we should look at compact, semi-simple Lie groups. Since any semi-simple Lie group can be decomposed as a direct sum of simple Lie groups, for simplicity, we look at compact, simple Lie groups.

For full diversity, we want to design constellations with positive diversity product, or in other words, constellations whose non-identity elements have no unit eigenvalues. It is proved in [HK00] that the only fpf infinite Lie groups are $U(1)$, the group of unit-modulus scalars, and $SU(2)$, the group of unit-determinant $2 \times 2$ unitary matrices. As discussed at the end of [HK00], due to their dimensions, constellations based on these two Lie groups are constrained to systems with one and two transmit antennas. (Codes constructed based on higher-dimensional representations of $SU(2)$ can be found in [Sho00].) To obtain constellations that work for systems with more than two transmit antennas, we relax the fpf condition, which is equivalent to non-identity elements have no unit eigenvalues, and consider Lie groups whose non-identity elements have no more than $k > 0$ unit eigenvalues ($k = 0$ corresponds to fpf groups). Since constellations of finite size are obtained by sampling the Lie group's underlying manifold. When $k$ is small, there is a good chance that, by sampling appropriately, fully diverse subsets can be obtained.

It follows from the exponential map relating Lie groups with Lie algebras that

a matrix element of a Lie group has unit-eigenvalues if and only if the corresponding matrix element (the logarithm of the element) in the corresponding Lie algebra (tangent space at identity) has zero-eigenvalues and vice versa. Thus classifying Lie groups whose non-identity elements have no more than $k$ eigenvalues at 1 is the same as classifying Lie algebras whose non-zero elements have no more than $k$ eigenvalues at zero. Unfortunately, there does not appear to be a straightforward way of analyzing the number of zero-eigenvalues of a matrix element of a Lie algebra. However, it turns out that the number of zero-eigenvalues of a matrix element of a Lie algebra can be related to the rank of its Lie algebra.

**Lemma 4.1.** *If a Lie algebra $\mathfrak{g}$ of $M \times M$ matrices has rank $r$, it has at least one non-zero element with $r-1$ eigenvalues at zero.*

**Proof:** Assume that $b_1, \cdots, b_r$ are the commuting basis elements of the Lie algebra $\mathfrak{g}$. Since they commute, there exists a matrix $T$ such that $b_1, \cdots, b_r$ can be diagonalized simultaneously. That is, there exists an $M \times M$ invertible matrix $T$, such that $b_i = TD_iT^{-1}$ where $D_i = \text{diag}\,(b_1^i, \cdots, b_M^i)$. Therefore, it is possible to design scalars $\alpha_1, \cdots, \alpha_r$ such that $\sum_{i=1}^r \alpha_i D_i$ is a diagonal matrix with the first $r-1$ diagonal elements being zero. The matrix $b = \sum_{i=1}^r \alpha_i b_i = T(\sum_{i=1}^r \alpha_i D_i)T^{-1}$, which is also an element in the Lie algebra, therefore, has $r-1$ eigenvalues at zero. □

Therefore, instead of exploring Lie groups whose non-identity elements have no more than $k$ unit eigenvalues, we work on the compact simple Lie groups whose rank is no more than $k+1$, and obtain a finite subset of it as our constellation by sampling its underlying manifold. As discussed before, for full diversity, we should begin with Lie groups with rank 2.

Combining the two conditions, unitarity and full diversity, a beginning point is to look at simple, compact Lie groups of rank 2. The following table is a complete list of simple, simply connected, compact Lie groups [Sim94]. In the table, $\mathbf{Z}(G)$ indicates the center of the group $G$. There are three groups with rank 2: the Lie group of unit-

| Group | Dimension | $\mathbf{Z}(G)$ | Cartan Name | Rank |
|---|---|---|---|---|
| $SU(n); n \geq 2$ | $n^2 - 1$ | $\mathbf{Z}_n$ | $A_{n-1}$ | $n-1$ |
| $Sp(n); n \geq 2$ | $n(2n+1)$ | $\mathbf{Z}_2$ | $C_n$ | $n$ |
| $Spin(2n+1); n \geq 3$ | $n(2n+1)$ | $\mathbf{Z}_2$ | $B_n$ | $n$ |
| $Spin(2n); n \geq 4$ | $n(2n-1)$ | $\mathbf{Z}_4$ ($n$ odd)<br>$\mathbf{Z}_2 \times \mathbf{Z}_2$ ($n$ even) | $D_n$ | $n$ |
| $E_6$ | 78 | $\mathbf{Z}_3$ | $E_6$ | 6 |
| $E_7$ | 133 | $\mathbf{Z}_2$ | $E_7$ | 7 |
| $E_8$ | 248 | 0 | $E_8$ | 8 |
| $F_4$ | 52 | 0 | $F_4$ | 4 |
| $G_2$ | 14 | 0 | $G_2$ | 2 |

Table 4.1: The simple, simply-connected, compact Lie groups

determinant $3 \times 3$ unitary matrices, $SU(3)$, the Lie group of $4 \times 4$ unitary, symplectic matrices, $Sp(2)$, and one of the five celebrated exceptional Lie groups of E. Cartan, $G_2$. $Sp(2)$ is analyzed in the next chapter and $SU(3)$ is analyzed in the chapter after. $G_2$ has dimension 14, and its simplest matrix representation is 7-dimensional, which is very difficult to be parameterized.

# Chapter 5 Differential Unitary Space-Time Codes Based on $Sp(2)$

## 5.1 Abstract

As discussed in Section 4.5, $Sp(n)$ is a Lie group with rank $n$. In this chapter, the focus is on $Sp(2)$, which, as will be seen later, can be regarded as a generalization of $SU(2)$, which results in Alamouti's scheme. Differential unitary space-time codes suitable for systems with four transmit antennas are designed based on this Lie group. The codes are fully diverse and their structure lends themselves to polynomial-time ML decoding via sphere decoding. Simulation results show that they have better performance than $2 \times 2$ and $4 \times 4$ orthogonal designs, Cayley differential codes, and some finite-group-based codes at high SNR. It is also shown that they are comparable to the carefully designed product-of-groups code.

This chapter is organized as follows. In Section 5.2, the Lie group $Sp(n)$ is discussed and a parameterization method of it is given. Based on the parameterization, in Section 5.3, differential unitary space-time codes that are subsets of $Sp(2)$ are designed. The full diversity of the codes is proved in Section 5.4. In Section 5.5, $Sp(2)$ codes with higher rates are proposed. It is shown in Section 5.6 that the codes have a fast ML decoding algorithm using sphere decoding. Finally, in Section 5.7 the performance of the $Sp(2)$ code is compared with that of other existing codes including Alamouti's orthogonal design, the $4 \times 4$ complex orthogonal design, Cayley differential unitary space-time codes, and finite-group-based codes. Section 5.8 provides the conclusion. Section 5.9 includes some of the technical proofs.

The work in this chapter has been published in *the Proceeding of the Thirty-Sixth Asilomar Conference on Signals, Systems, and Computers (Asilomar'02)* [JH02], *the Proceeding of 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)* [JH03a], and *the Proceeding of 2003 IEEE International Symposium on Information Theory (ISIT'03)* [JH03c]. The journal paper [JH03b] is accepted in *IEEE Transactions on Information Theory*.

## 5.2 The Symplectic Group and Its Parameterization

**Definition 5.1 (Symplectic group).** *[Sim94] $Sp(n)$, the $n$-th order symptectic group, is the set of complex $2n \times 2n$ matrices $S$ obeying*

1. *unitary condition: $S^*S = SS^* = I_{2n}$,*
2. *symplectic condition: $S^t J_{2n} S = J_{2n}$,*

*where* $J_{2n} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$.

$Sp(n)$ has dimension $n(2n+1)$ and rank $n$. As mentioned before, we are most interested in the lowest rank case, which is also the simplest case of $n = 2$. Also note that $Sp(1) = SU(2)$, and $SU(2)$ constitutes the orthogonal design of Alamouti [Ala98]. The symplectic group $Sp(2)$ can be regarded as a generalization of the orthogonal design. Even though Lemma 4.1 claims that there exists an element of $Sp(2)$ with at least one unit eigenvalue, it can be shown that a non-identity element of $Sp(2)$ can have up to 2 unit eigenvalues.

**Lemma 5.1.** *The multiplicity of the unit eigenvalue of any matrix in $Sp(2)$ is even.*

**Proof:** Assume $S$ is a matrix in $Sp(2)$ and $x$ is an eigenvector of $S$ with eigenvalue 1. Then we have $Sx = x$. From the symplectic condition $S^t JS = J$, $S^* J\bar{S} = J$. Since $S$ is unitary, $J\bar{S} = SJ$. Therefore, $SJ\bar{x} = J\bar{S}\bar{x} = J\overline{Sx} = J\bar{x}$, which means that $J\bar{x}$ is also an eigenvector of $S$ with eigenvalue 1. We now argue that $x \neq J\bar{x}$. Assume $x = J\bar{x}$. Partition $x$ as $[x_1, x_2]^t$, where $x_1$ and $x_2$ are 2-dimensional vectors. We have

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & I_2 \\ -I_2 & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} \Rightarrow \begin{cases} x_1 = \bar{x}_2 \\ x_2 = -\bar{x}_1 \end{cases} \Rightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases},$$

from which we get $x = 0$. This contradicts the assumption that $x$ is an eigenvector. Therefore, $x \neq J\bar{x}$, which means that the number of eigenvectors for any unit eigenvalue is even. Thus, the multiplicity of the unit eigenvalue is even. □

From Lemma 5.1, if a matrix in $Sp(2)$ has a unit eigenvalue then its multiplicity must be 2 or 4. 4 unit eigenvalues means that the matrix is $I_4$. Therefore, a non-identity element of $Sp(2)$ can have 0 or 2 unit eigenvalues.

From Condition 1 of Definition 5.1, $(S^t)^{-1} = \bar{S}$, where $\bar{S}$ is the conjugate of $S$, so condition 2 becomes

$$JS = \bar{S}J. \tag{5.1}$$

If the matrix $S$ is further partitioned into a $2 \times 2$ block matrix of $n \times n$ sub-matrices $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$, from (5.1), $C = \bar{B}$ and $D = \bar{A}$. Therefore, any matrices in $Sp(n)$ have the form

$$\begin{bmatrix} A & B \\ -\bar{B} & \bar{A} \end{bmatrix}, \tag{5.2}$$

which is similar to Alamouti's two-dimensional orthogonal design [Ala98], but here

instead of complex scalars, $A$ and $B$ are $n$ by $n$ complex matrices.[1] The group, $Sp(n)$, can thus be identified as the subgroup of unitary matrices with generalized orthogonal design form. To get a more detailed structure of the Lie group, let us look at the conditions imposed on $A$ and $B$ for $S$ to be unitary. From $SS^* = I_{2n}$ or $S^*S = I_{2n}$,

$$\begin{cases} AA^* + BB^* = I_n \\ BA^t = AB^t \end{cases} \quad \text{or} \quad \begin{cases} A^*A + B^t\bar{B} = I_n \\ A^*B = B^t\bar{A} \end{cases} . \tag{5.3}$$

**Lemma 5.2.** *For any $n \times n$ complex matrices $A$ and $B$ that satisfy (5.3), there exist unitary matrices $U$ and $V$ such that $A = U\Sigma_A V$ and $B = U\Sigma_B \bar{V}$, where $\Sigma_A$ and $\Sigma_B$ are diagonal matrices whose diagonal elements are singular values of $A$ and $B$.*

To proof this lemma, two intermediate lemmas are needed.

**Lemma 5.3.** *Let $D_1$ and $D_2$ be $n \times n$ diagonal matrices with non-increasing diagonal entries. If $UD_1U^* = D_2$ for some unitary matrix $U$, then $D_1 = D_2 = D$ and $U$ is a block diagonal matrix whose block sizes equal the number of times that the diagonal elements of $D$ are repeated.*

**Proof:** Denote the $i$-th diagonal element of $D_1$ and $D_2$ as $d_{ii}^{(1)}$ and $d_{ii}^{(2)}$, respectively. Since $UD_1U^*$ is a similarity transformation, which preserves the eigenvalues, the set of eigenvalues of $D_1$ is the same as the set of eigenvalues of $D_2$, or in other words, $d_{ii}^{(1)} = d_{jj}^{(2)}$ for some $j$. Notice that $d_{ii}^{(1)}s$ and $d_{ii}^{(2)}s$ are ordered non-increasingly. Therefore, $d_{ii}^{(1)} = d_{ii}^{(2)}$ for $i = 1, 2, ..., n$, that is, $D_1 = D_2 = D$. Now write $D$ as $\text{diag}\{d_1 I_{P_1}, \cdots, d_q I_{P_q}\}$, where $P_i$ is the number of times the element $d_i$ appears in $D$ for $i = 1, \cdots, q$. It is obvious that $U$ can be written as $\text{diag}\{U_1, \cdots, U_q\}$ where the size of $U_i$ is $P_q$ for $i = 1, \cdots, q$. □

[1] Note that the structure in (5.2) is akin to the quasi-orthogonal space-time block codes in [TBH00, Jaf01]. The crucial difference, however, is that in our paper, we shall insist that (5.2) be a unitary matrix. This leads to further conditions on $A$ and $B$, which are described below, and do not appear in quasi-orthogonal space-time block codes.

**Lemma 5.4.** *If $UD^2 = D^2U$ for any $n \times n$ positive semi-definite diagonal matrix $D$ and any $n \times n$ matrix $U$, then $UD = DU$.*

**Proof:** By looking at the $(i, j)$-th entries of $UD^2$ and $D^2U$, $u_{ij}d_{jj}^2 = d_{ii}^2 u_{ij}$. If $d_{ii}^2 \neq d_{jj}^2$, $u_{ij} = 0$ is obtained, and therefore $u_{ij}d_{jj} = d_{ii}u_{ij}$. If $d_{ii}^2 = d_{jj}^2$, since $D$ is a positive semi-definite matrix, $d_{ii}$ is non-negative, therefore, $d_{ii} = d_{jj}$ and so $u_{ij}d_{jj} = d_{ii}u_{ij}$ is obtained. Therefore, $UD = DU$. □

Now Lemma 5.2 is ready to be proved.

**Proof of Lemma 5.2:** Suppose $A = U_A\Sigma_A V_A^*$ and $B = U_B\Sigma_B V_B^*$ are the singular value decompositions of $A$ and $B$ with the non-negative diagonal elements of $\Sigma_A$ *non-decreasingly* ordered and the non-negative diagonal elements of $\Sigma_B$ *non-increasingly* ordered. From the equation $AA^* + BB^* = I_n$ in (5.3), the following series of equations can be obtained.

$$\begin{aligned}
& U_A\Sigma_A V_A^* V_A \Sigma_A U_A^* + U_B\Sigma_B V_B^* V_B \Sigma_B U_B^* = I_n \\
\Rightarrow\ & U_A\Sigma_A^2 U_A^* + U_B\Sigma_B^2 U_B^* = I_n \\
\Rightarrow\ & \Sigma_A^2 + (U_A^*U_B)\Sigma_B^2(U_A^*U_B)^* = I_n \\
\Rightarrow\ & (U_A^*U_B)\Sigma_B^2(U_A^*U_B)^* = I_n - \Sigma_A^2.
\end{aligned}$$

Since $U_A$ and $U_B$ are unitary, $U_A^*U_B$ is also unitary. Now since the diagonal entries of $\Sigma_B^2$ and $I_n - \Sigma_A^2$ are non-increasingly ordered, from Lemma 5.3,

$$\Sigma_B^2 = I_n - \Sigma_A^2. \tag{5.4}$$

Define $U_D = U_A^*U_B$. Therefore $U_B = U_AU_D$ and

$$U_D(I_n - \Sigma_A^2) = (I_n - \Sigma_A^2)U_D \Leftrightarrow U_D\Sigma_A^2 = \Sigma_A^2 U_D.$$

Since $\Sigma_A$ is a positive semi-definite matrix, from Lemma 5.4, $U_D\Sigma_A = \Sigma_A U_D$ and therefore $U_D^*\Sigma_A = \Sigma_A U_D^*$. Further define $U_1 = U_B = U_A U_D$. Then,

$$A = U_A\Sigma_A V_A^* = U_A U_D U_D^*\Sigma_A V_A^* = U_A U_D \Sigma_A (V_A U_D)^* = U_1\Sigma_A V_A'^*,$$

where $V_A'$ is defined as $V_A' = V_A U_D$. Since $U_1 = U_B$, $B = U_1\Sigma_B V_B^*$, Thus $A$ and $B$ have the same left singular vectors.

We now focus on the right singular vectors. From the equation $A^*A + B^t\bar{B} = I_n$ in (5.3), the following series of equations can be obtained.

$$\begin{aligned}
& V_A'\Sigma_A U_1^* U_1 \Sigma_A V_A'^* + \bar{V}_B\Sigma_B U_1^t \bar{U}_1 \Sigma_B V_B^t = I_n \\
\Rightarrow\ & V_A'\Sigma_A^2 V_A'^* + \bar{V}_B\Sigma_B^2 V_B^t = I_n \\
\Rightarrow\ & \Sigma_A^2 + (V_A'^*\bar{V}_B)\Sigma_B^2(V_A'^*\bar{V}_B)^* = I_n \\
\Rightarrow\ & (V_A'^*\bar{V}_B)\Sigma_B^2(V_A'^*\bar{V}_B)^* = I_n - \Sigma_A^2.
\end{aligned}$$

Therefore, $I_n - \Sigma_A^2 = \Sigma_B^2$ and $(V_A'^*\bar{V}_B)\Sigma_B^2 = \Sigma_B^2(V_A'^*\bar{V}_B)$ by using Lemma 5.3. Define $V_D = V_A'^*\bar{V}_B$, which is obviously a unitary matrix. Therefore $\bar{V}_B = V_A' V_D$, $V_D\Sigma_B^2 = \Sigma_B^2 V_D$, and $V_D\Sigma_A^2 = \Sigma_A^2 V_D$, from which $V_D\Sigma_B = \Sigma_B V_D$ and $V_D\Sigma_A = \Sigma_A V_D$ can be obtained by using Lemma 5.4. Now according to Lemma 5.3, $V_D$ can be written as $\mathrm{diag}\{V_1, \cdots, V_q\}$ with each $V_i$ being a unitary matrix. Since $V_i$ is unitary, there exists a Hermitian matrix $G_i$ such that $V_i = e^{jG_i}$. Because $G_i$ is Hermitian, so is $G_i/2$. Therefore, the matrix $e^{jG_i/2}$, which is the square root of $V_i$, is also a unitary matrix and $e^{jG/2}\Sigma_A = \Sigma_A e^{jG/2}, e^{jG/2}\Sigma_B = \Sigma_B e^{jG/2}$ can be obtained, where $G/2 = \mathrm{diag}\{G_1/2, \cdots, G_q/2\}$. Therefore $e^{jG/2}$ is the square root of $V_D$. Thus, $A$ and $B$ can

be written as

$$\begin{aligned} A &= U_1\Sigma_A V_D V_1^* = U_1\Sigma_A e^{jG} V_1^* = U_1 e^{jG/2}\Sigma_A e^{jG/2} V_1^* = (U_1 e^{jG/2})\Sigma_A (V_1 e^{-jG/2})^*, \\ B &= U_1\Sigma_B V_1^t = (U_1 e^{jG/2})\Sigma_B (V_1 e^{-jG/2})^t, \end{aligned}$$

where $V_1$ is defined as $V_1 = \bar{V}_B = V_A' V_D$. Therefore, $A = U\Sigma_A V^*$ and $Y = U\Sigma_B V^t$ for some unitary matrices $U = U_1 e^{jG/2}$ and $V = V_1 e^{-jG/2}$ or equivalently, $A = U\Sigma_A V$ and $Y = U\Sigma_B \bar{V}$ if $U$ and $V$ are defined as $U = U_1 e^{jG/2}$and $V = (V_1 e^{-jG/2})^*$. □

Lemma 5.2 indicates that $A$ and $B$ can be diagonalized by the same pair of unitary matrices. This leads to the following parameterization theorem of $Sp(n)$.

**Theorem 5.1 (Parameterization of $Sp(n)$).** *A matrix $S$ belongs to $Sp(n)$ if and only if it can be written as*

$$S = \begin{bmatrix} U\Sigma_A V & U\Sigma_B \bar{V} \\ -\bar{U}\Sigma_B V & \bar{U}\Sigma_A \bar{V} \end{bmatrix}, \tag{5.5}$$

*where $\Sigma_A = diag(\cos\theta_1, \cdots \cos\theta_n), \Sigma_B = diag(\sin\theta_1, \cdots, \sin\theta_n)$ for some $\theta_1, \cdots, \theta_n \in [0, \pi/2]$ and $U$ and $V$ are $n \times n$ unitary matrices.*

**Proof:** Lemma 5.2 and formula (5.2) imply that any matrix in $Sp(2)$ can be written as the form in (5.5). Conversely, for any matrix $S$ with the form of (5.5), it is easy to verify the unitary and symplectic conditions in Definition 5.1. □

Now let us look at the dimension of $S$. It is known that an $n \times n$ unitary matrix has dimension $2n^2 - n - 2 \times \frac{n(n-1)}{2} = n^2$. Therefore, there are all together $2n^2$ degrees of freedom in the unitary matrices $U$ and $V$. Together with the $n$ real angles $\theta_i$, the dimension of $S$ is therefore $n(2n+1)$, which is exactly the same as that of $Sp(n)$. But from the discussion above, an extra condition is imposed on the matrix $S$: the diagonal elements of $\Sigma_A$ and $\Sigma_B$ are non-negative and non-increasingly/non-decreasingly ordered. This might cause the dimension of $S$ to be less than matrices

in $Sp(2)$ at first glance. However, the order and signs of the diagonal elements of $\Sigma_A$ and $\Sigma_B$ can be changed by right multiplying $U$ and left multiplying $V$ with two types of matrices: permutation matrices and diagonal matrices with diagonal elements 1 and $-1$.[2] Therefore, the constraint does not result in dimension reduction. Based on Theorem 5.1, matrices in $Sp(n)$ can be parameterized by elements of $U, V$ and the real angles $\theta_i$s. Therefore, we can obtain finite subsets (samplings) of the infinite Lie group $Sp(n)$ by sampling these parameters.

## 5.3 Design of $Sp(2)$ Codes

Let us now focus on the case of $n = 2$. The goal is to find fully diverse subsets of $Sp(2)$. For simplicity, first let $\Sigma_A = \Sigma_B = \frac{1}{\sqrt{2}} I_2$, by which 2 degrees of freedom are neglected. To get a finite subset of unitary matrices from the infinite Lie group, further choose $U$ and $V$ as orthogonal designs with entries of $U$ chosen from the set of $P$-PSK signals: $\{1, e^{j\frac{2\pi}{P}}, \cdots, e^{j\frac{2\pi(P-1)}{P}}\}$ and entries of $V$ chosen from the set of $Q$-PSK signals shifted by an angle $\theta$: $\{e^{j\theta}, e^{j(\frac{2\pi}{Q}+\theta)}, \cdots, e^{j(\frac{2\pi(Q-1)}{Q}+\theta)}\}$ [Hug00b, Ala98]. The following code is obtained.

$$\mathcal{C}_{P,Q,\theta} = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} UV & U\bar{V} \\ -\bar{U}V & \bar{U}\bar{V} \end{bmatrix} \middle| \begin{array}{c} U = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j\frac{2\pi k}{P}} & e^{j\frac{2\pi l}{P}} \\ -e^{-j\frac{2\pi l}{P}} & e^{-j\frac{2\pi k}{P}} \end{bmatrix}, \\ V = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j(\frac{2\pi m}{Q}+\theta)} & e^{j(\frac{2\pi n}{Q}+\theta)} \\ -e^{-j(\frac{2\pi n}{Q}+\theta)} & e^{-j(\frac{2\pi m}{Q}+\theta)} \end{bmatrix}, \\ 0 \le k, l < P, 0 \le m, n < Q \end{array} \right\}, \quad (5.6)$$

where $P$ and $Q$ are positive integers. $\theta \in [0, 2\pi)$ is an angle to be chosen later. There are $P^2$ possible $U$ matrices and $Q^2$ possible $V$ matrices. Since the channel is used in

[2]Definition of permutation matrices can be found in [Art99]. It is easy to see that both types of matrices are unitary, therefore, the unitarity of $U$ and $V$ keeps unchanged.

blocks of four transmissions, the rate of the code is therefore

$$\frac{1}{2}(\log_2 P + \log_2 Q). \tag{5.7}$$

It is easy to see that any transmission matrix in the code can be identified by the 4-tuple $(k, l, m, n)$. The angle $\theta$, an extra degree of freedom added to increase the diversity product, is used in the proof of the full diversity of the code. However, simulation results indicate that the code is always fully-diverse no matter what value $\theta$ takes.

A similar code as follows can also be considered.

$$\mathcal{C}'_{P,Q,\theta} = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} UV & U\bar{V} \\ -\bar{U}V & \bar{U}\bar{V} \end{bmatrix} \middle| \begin{array}{c} U = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j\frac{2\pi k}{P}} & e^{j\frac{2\pi l}{Q}} \\ -e^{-j\frac{2\pi l}{Q}} & e^{-j\frac{2\pi k}{P}} \end{bmatrix}, \\ V = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j(\frac{2\pi m}{P}+\theta)} & e^{j(\frac{2\pi n}{Q}+\theta)} \\ -e^{-j(\frac{2\pi n}{Q}+\theta)} & e^{-j(\frac{2\pi m}{P}+\theta)} \end{bmatrix}, \\ 0 \le k, m < P, 0 \le l, n < Q \end{array} \right\}.$$

The rate of the code is the same as that of $\mathcal{C}_{P,Q,\theta}$ and its full diversity can be proved similarly. Here, however, the focus is on the code given in (5.6).

## 5.4 Full Diversity of $Sp(2)$ Codes

In differential unitary space-time code design for multiple-antenna systems, the most widely used criterion is the full diversity of the code since as discussed in Section 2.6, the diversity product is directly related to the pairwise error probability of the systems. This issue is discussed in this section.

To calculate the diversity product of the code $\mathcal{C}_{P,Q,\theta}$ given in (5.6), first choose

any 2 transmission matrices $S_1$ and $S_2$ in $\mathcal{C}_{P,Q,\theta}$ where

$$S_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} U_1V_1 & U_1\bar{V}_1 \\ -\bar{U}_1V_1 & \bar{U}_1\bar{V}_1 \end{bmatrix} \quad \text{and} \quad S_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} U_2V_2 & U_2\bar{V}_2 \\ -\bar{U}_2V_2 & \bar{U}_2\bar{V}_2 \end{bmatrix} \tag{5.8}$$

and

$$U_i = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j\frac{2\pi k_i}{P}} & e^{j\frac{2\pi l_i}{P}} \\ -e^{-j\frac{2\pi l_i}{P}} & e^{-j\frac{2\pi k_i}{P}} \end{bmatrix} \quad \text{and} \quad V_i = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j(\frac{2\pi m_i}{Q}+\theta)} & e^{j(\frac{2\pi n_i}{Q}+\theta)} \\ -e^{-j(\frac{2\pi n_i}{Q}+\theta)} & e^{-j(\frac{2\pi m_i}{Q}+\theta)} \end{bmatrix}. \tag{5.9}$$

$k_i, l_i \in [0, P)$ and $m_i, n_i \in [0, Q)$ are integers for $i = 1, 2$. Before calculating the determinant of the difference of the two matrices, some well-known facts about $2 \times 2$ orthogonal design [Ala98] are first stated as follows.

**Lemma 5.5.** *For any non-zero* $2 \times 2$ *matrix*

$$M = \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix},$$

1. $\det M = |a|^2 + |b|^2$,
2. $M^*M = MM^* = (\det M)I_2$,
3. $\det M = 0$ *if and only if* $M = \mathbf{0}_{22}$,
4. $M^{-1} = \frac{M^*}{\det M}$ *and* $\bar{M}^{-1} = \frac{M^t}{\det M}$.

**Proof:** Straightforward algebra. □

This lemma shows that the determinant of any non-zero matrix with an orthogonal design structure is real and positive.

Define

$$O_1 = U_1V_1 - U_2V_2 \quad \text{and} \quad O_2 = U_1\bar{V}_1 - U_2\bar{V}_2. \tag{5.10}$$

Therefore,

$$\begin{aligned}
& \det(S_1 - S_2) \\
= & \frac{1}{\sqrt{2}} \det \begin{bmatrix} U_1V_1 - U_2V_2 & U_1\bar{V}_1 - U_2\bar{V}_2 \\ -(\bar{U}_1V_1 - \bar{U}_2V_2) & \bar{U}_1\bar{V}_1 - \bar{U}_2\bar{V}_2 \end{bmatrix} \\
= & \frac{1}{\sqrt{2}} \det \begin{bmatrix} O_1 & O_2 \\ -\bar{O}_2 & \bar{O}_1 \end{bmatrix} \\
= & \frac{1}{\sqrt{2}} \det O_1 \det(\bar{O}_1 + \bar{O}_2 O_1^{-1} O_2) \qquad (5.11)
\end{aligned}$$

if $O_1$ is invertible. Since $U_1$, $U_2$, $V_1$ and $V_2$ are all orthogonal designs and it is easy to prove that the addition, multiplication, and conjugate operations preserve this property, $O_1$ and $O_2$ are also have the orthogonal design structure. By taking advantage of this, when $\det O_1 \neq 0$ and $\det O_2 \neq 0$, the determinant of the difference can be calculated to be

$$\begin{aligned}
& \det(S_1 - S_2) \\
= & \frac{1}{\sqrt{2}} \det O_1 \det\left(\bar{O}_1 + \frac{\bar{O}_2 O_1^* O_2}{\det O_1}\right) \\
= & \frac{1}{\sqrt{2}} \det O_1 \det\left(\frac{\bar{O}_1 O_2^* O_2}{\det O_2} + \frac{\bar{O}_2 O_1^* O_2}{\det O_1}\right) \\
= & \frac{1}{\sqrt{2}} \det O_1 \det\left(\frac{\bar{O}_1 O_2^*}{\det O_2} + \frac{\bar{O}_2 O_1^*}{\det O_1}\right) \det O_2 \\
= & \frac{1}{\sqrt{2}} \det\left(\frac{\sqrt{\det O_1 \det O_2}}{\det O_2} \bar{O}_1 O_2^* + \frac{\sqrt{\det O_1 \det O_2}}{\det O_1} \bar{O}_2 O_1^*\right) \\
= & \frac{1}{\sqrt{2}} \det\left(a\bar{O}_1 O_2^* + \frac{1}{a}(\bar{O}_1 O_2^*)^t\right) \\
= & \frac{1}{\sqrt{2}} \det\left(a \begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix} + \frac{1}{a} \begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix}^t\right) \\
= & \frac{1}{\sqrt{2}} \left(\left|a\alpha + \frac{1}{a}\alpha\right|^2 + \left|a\beta - \frac{1}{a}\bar{\beta}\right|^2\right)
\end{aligned}$$

$$= \frac{1}{\sqrt{2}}|\alpha|^2\left(a+\frac{1}{a}\right)^2+\frac{1}{2}\left|a\beta-\frac{\bar{\beta}}{a}\right|^2, \tag{5.12}$$

where $a=\sqrt{\frac{\det O_1}{\det O_2}}$ is a positive number and $(\alpha,\beta)$ is the first row of $\bar{O}_1O_2^*$.

**Lemma 5.6.** *For any $S_1$ and $S_2$ given in (5.8) and (5.9) where $S_1\neq S_2$, $\det(S_1-S_2)=0$ if and only if $O_1=\pm J\bar{O}_2$, or equivalently*

$$\begin{cases} e^{2j\theta}w^+=x^+ \\ e^{2j\theta}y^+=z^+ \end{cases}, \qquad or \qquad \begin{cases} e^{2j\theta}w^-=x^- \\ e^{2j\theta}y^-=z^- \end{cases}, \tag{5.13}$$

*where*

$$\begin{cases} w^+=e^{2\pi j(\frac{k_1}{P}+\frac{m_1}{Q})}-e^{2\pi j(\frac{k_2}{P}+\frac{m_2}{Q})}+e^{2\pi j(\frac{l_1}{P}+\frac{m_1}{Q})}-e^{2\pi j(\frac{l_2}{P}+\frac{m_2}{Q})} \\ x^+=-e^{2\pi j(\frac{k_1}{P}-\frac{n_1}{Q})}+e^{2\pi j(\frac{k_2}{P}-\frac{n_2}{Q})}+e^{2\pi j(\frac{l_1}{P}-\frac{n_1}{Q})}-e^{2\pi j(\frac{l_2}{P}-\frac{n_2}{Q})} \\ y^+=e^{2\pi j(\frac{k_1}{P}+\frac{n_1}{Q})}-e^{2\pi j(\frac{k_2}{P}+\frac{n_2}{Q})}+e^{2\pi j(\frac{l_1}{P}+\frac{n_1}{Q})}-e^{2\pi j(\frac{l_2}{P}+\frac{n_2}{Q})} \\ z^+=e^{2\pi j(\frac{k_1}{P}-\frac{m_1}{Q})}-e^{2\pi j(\frac{k_2}{P}-\frac{m_2}{Q})}-e^{2\pi j(\frac{l_1}{P}-\frac{m_1}{Q})}+e^{2\pi j(\frac{l_2}{P}-\frac{m_2}{Q})} \end{cases} \tag{5.14}$$

*and*

$$\begin{cases} w^-=e^{2\pi j(\frac{k_1}{P}+\frac{m_1}{Q})}-e^{2\pi j(\frac{k_2}{P}+\frac{m_2}{Q})}-e^{2\pi j(\frac{l_1}{P}+\frac{m_1}{Q})}+e^{2\pi j(\frac{l_2}{P}+\frac{m_2}{Q})} \\ x^-=e^{2\pi j(\frac{k_1}{P}-\frac{n_1}{Q})}-e^{2\pi j(\frac{k_2}{P}-\frac{n_2}{Q})}+e^{2\pi j(\frac{l_1}{P}-\frac{n_1}{Q})}-e^{2\pi j(\frac{l_2}{P}-\frac{n_2}{Q})} \\ y^-=e^{2\pi j(\frac{k_1}{P}+\frac{n_1}{Q})}-e^{2\pi j(\frac{k_2}{P}+\frac{n_2}{Q})}-e^{2\pi j(\frac{l_1}{P}+\frac{n_1}{Q})}+e^{2\pi j(\frac{l_2}{P}+\frac{n_2}{Q})} \\ z^-=-e^{2\pi j(\frac{k_1}{P}-\frac{m_1}{Q})}+e^{2\pi j(\frac{k_2}{P}-\frac{m_2}{Q})}-e^{2\pi j(\frac{l_1}{P}-\frac{m_1}{Q})}+e^{2\pi j(\frac{l_2}{P}-\frac{m_2}{Q})} \end{cases}. \tag{5.15}$$

**Proof:** See Section 5.9.1. □

Here is the main theorem.

**Theorem 5.2 (Condition for full diversity).** *There exists a $\theta$ such that the code in (5.6) is fully diverse if and only if $P$ and $Q$ are relatively prime.*

This theorem provides both the sufficient and the necessary condition for the code set to be fully diverse. Before proving this theorem, a few lemmas are given first which are needed in the proof of the theorem.

**Lemma 5.7.** *For any four points on the unit circle that add up (as complex numbers) to zero, it is always true that two of them add up to zero. (Clearly, the other two must also have a summation of zero.)*

**Proof:** See Section 5.9.2. □

**Lemma 5.8.** *If $P$ and $Q$ are relatively prime, then for any non-identical pairs, $(k_1, l_1, m_1, n_1)$ and $(k_2, l_2, m_2, n_2)$, where $k_1, l_1, k_2, l_2 \in [0, P)$ and $m_1, n_1, m_2, n_2 \in [0, Q)$ are integers, $w^+$, $x^+$, $y^+$, $z^+$, as defined in (5.14), cannot be zero simultaneously. Also $w^-$, $x^-$, $y^-$, $z^-$, as defined in (5.15), cannot be zero simultaneously.*

**Proof:** See Section 5.9.3. □

Now it is ready to prove Theorem 5.2.

**Proof of Theorem 5.2:** The proof has two steps. First, we prove the sufficiency of the condition, that is, assuming $P$ and $Q$ are relatively prime, we show that there exists a $\theta$ such that the code is fully-diverse. If $P$ and $Q$ are relatively prime, by Lemma 5.8, for any non-identical pair of signal matrices $(k_1, l_1, m_1, n_1)$ and $(k_2, l_2, m_2, n_2)$, $w^+$, $x^+$, $y^+$, $z^+$ cannot be zero simultaneously. (For definitions of $w^+$, $x^+$, $y^+$, $z^+$, see (5.14)). Define

$$\theta^+_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\mathrm{Arg}\left(\frac{x^+}{w^+}\right) & \mod 2\pi \ \text{ if } \ w^+ \neq 0 \\ -\frac{1}{2}\mathrm{Arg}\left(\frac{w^+}{x^+}\right) & \mod 2\pi \ \text{ if } \ w^+ = 0, x^+ \neq 0 \\ \frac{1}{2}\mathrm{Arg}\left(\frac{z^+}{y^+}\right) & \mod 2\pi \ \text{ if } \ w^+ = x^+ = 0, y^+ \neq 0 \\ -\frac{1}{2}\mathrm{Arg}\left(\frac{y^+}{z^+}\right) & \mod 2\pi \ \text{ if } \ w^+ = x^+ = y^+ = 0, z^+ \neq 0 \end{cases},$$

which is the same as

$$\theta^{+}_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\mathrm{Arg}\,(\frac{x^+}{w^+}) \mod 2\pi & \text{if } w^+ \neq 0 \\ 0 & \text{if } w^+ = 0, x^+ \neq 0 \\ \frac{1}{2}\mathrm{Arg}\,(\frac{z^+}{y^+}) \mod 2\pi & \text{if } w^+ = x^+ = 0, y^+ \neq 0 \\ 0 & \text{if } w^+ = x^+ = y^+ = 0, z^+ \neq 0 \end{cases} \tag{5.16}$$

$\mathrm{Arg}\,c$ indicates the argument of the complex number $c$. Also, by Lemma 5.8, $w^-, x^-, y^-, z^-$ cannot be zero simultaneously. (For definitions of $w^-, x^-, y^-, z^-$, see (5.15)). Define

$$\theta^{-}_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\mathrm{Arg}\,(\frac{x^-}{w^-}) \mod 2\pi & \text{if } w^- \neq 0 \\ -\frac{1}{2}\mathrm{Arg}\,(\frac{w^-}{x^-}) \mod 2\pi & \text{if } w^- = 0, x^- \neq 0 \\ \frac{1}{2}\mathrm{Arg}\,(\frac{z^-}{y^-}) \mod 2\pi & \text{if } w^- = x^- = 0, y^- \neq 0 \\ -\frac{1}{2}\mathrm{Arg}\,(\frac{y^-}{z^-}) \mod 2\pi & \text{if } w^- = x^- = y^- = 0, z^- \neq 0 \end{cases},$$

which is the same as

$$\theta^{-}_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\mathrm{Arg}\,(\frac{x^-}{w^-}) \mod 2\pi & \text{if } w^- \neq 0 \\ 0 & \text{if } w^- = 0, x^- \neq 0 \\ \frac{1}{2}\mathrm{Arg}\,(\frac{z^-}{y^-}) \mod 2\pi & \text{if } w^- = x^- = 0, y^- \neq 0 \\ 0 & \text{if } w^- = x^- = y^- = 0, z^- \neq 0 \end{cases} \tag{5.17}$$

By choosing

$$\theta \notin \{\theta^{+}_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2}, \\ \Big| \; |w^+| = |x^+|, |y^+| = |z^+|, 0 \leq k_1, l_1, k_2, l_2 < P, 0 \leq m_1, n_1, m_2, n_2 < Q\} \tag{5.18}$$

and

$$\theta \notin \left\{\theta^{-}_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2}\ , \right. \\ \left. \Big| \ ||w^-| = |x^-|, |y^-| = |z^-|, 0 \le k_1, l_1, k_2, l_2 < P, 0 \le m_1, n_1, m_2, n_2 < Q\right\}, \tag{5.19}$$

(5.13) cannot be true at the same time. Therefore, by Lemma 5.6, $\det(S_1 - S_2) \neq 0$, which means that the code is fully diverse. An angle in $[0, 2\pi)$ that satisfies (5.18) can always be found since the two sets at the right-hand side of (5.18) and (5.19) are finite. This proves the sufficiency of the condition ($P$ and $Q$ are relatively prime) in Theorem 5.2.

In the second step, we prove the necessity of the condition, that is, assuming that $P$ and $Q$ are not relatively prime, we show that there exist two signals in the code such that the determinant of the difference of the two is zero for any $\theta$. Assume that the greatest common divisor of $P$ and $Q$ is $G > 1$, then there exist positive integers $P'$ and $Q'$ such that $P = P'G$ and $Q = Q'G$. Consider the following two signal matrices $S_1$ and $S_2$ as given in (5.8) and (5.9) with $k_2 = k_1 - P'$, $l_2 = l_1 - P'$, $m_2 = m_1 + Q'$, $n_2 = n_1 + Q'$, $k_1 = l_1$, and $k_2 = l_2$. Assume that $k_1, l_1 \in [P', P)$, and $m_1, n_1 \in [0, Q - Q')$. Since $P > P'$and $Q > Q'$, we can always choose $k_1, l_1, m_1, n_1$ that satisfy the conditions. It is easy to verify that $w^+ = x^+ = y^+ = z^+ = 0$, which means that the first set of equations in (5.13) is true for any angle $\theta$. Therefore, $O_1 = J\bar{O}_2$, and from Lemma 5.6, we have $\det(S_1 - S_2) = 0$, that is, the signal set in (5.6) is not fully-diverse. □

**Remark:** Note that we have actually proved that the codes in (5.6) are fully diverse for almost any $\theta$ except for a measure zero set. However, this is a sufficient condition and may be not necessary. The diversity products of many codes for $\theta$ from 0 to $2\pi$ with step size 0.001 are calculated by simulation. Two of these are shown in the following. Simulation results show that the codes are fully diverse for all $\theta$.

The following two plots show the diversity products of two $Sp(2)$ codes at different

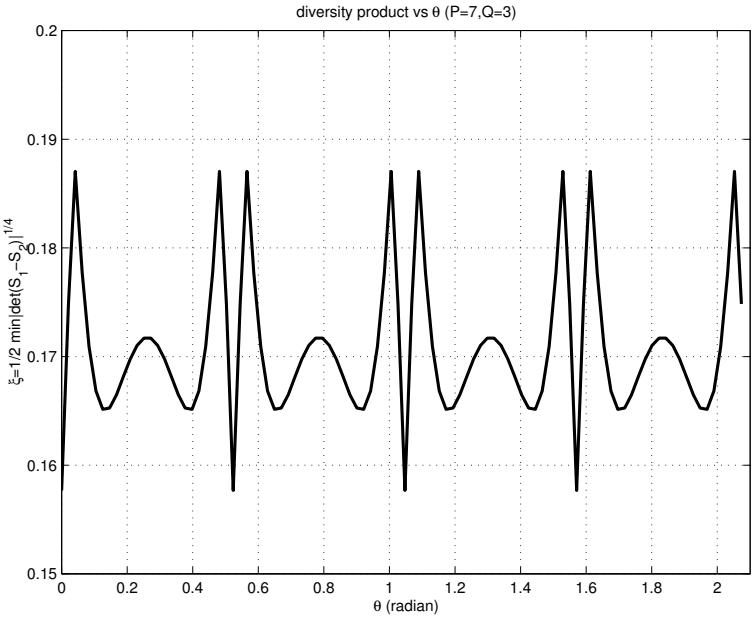


Figure 5.1: Diversity product of the $P = 7, Q = 3$ $Sp(2)$ code

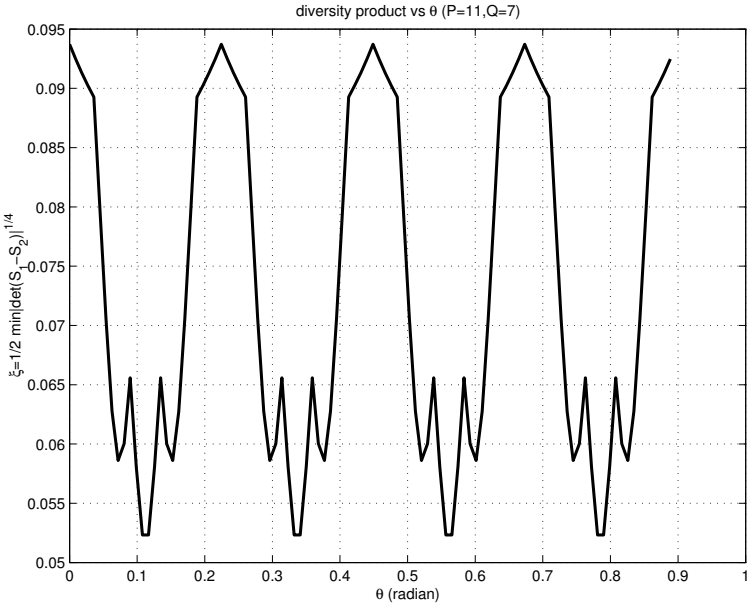


Figure 5.2: Diversity product of the $P = 11, Q = 7$ $Sp(2)$ code

$\theta$. Figure 5.1 shows the diversity product of the $Sp(2)$ code with $P = 7$ and $Q = 3$ and Figure 5.2 shows the diversity product of the $Sp(2)$ code with $P = 11$ and $Q = 7$. Since the angles of the elements in the $V$ matrix of (5.6) are chosen from $Q$-PSK signals shifted by an angle $\theta$, it is enough to set the changing region of $\theta$ as $[0, 2\pi/Q)$ instead of $[0, 2\pi)$. It can be seen from the two plots that the $Sp(2)$ code with $P = 7$ and $Q = 3$ gets its highest diversity product, 0.1870, at $\theta = 0.0419$ and the $Sp(2)$ code with $P = 7$ and $Q = 3$ gets its highest diversity product, 0.0937, at $\theta = 0$. Although the codes are fully diverse at any $\theta$.

## 5.5 $Sp(2)$ Codes of Higher Rates

In section 5.4, $Sp(2)$ codes are designed with the 2 degrees of freedom in $\Sigma_A$ and $\Sigma_B$ unused. For higher rate code design, one of the two degrees of freedom can be added in by letting

$$\Sigma_A = \cos\gamma_i I_2 \quad \text{and} \quad \Sigma_B = \sin\gamma_i I_2,$$

where $\gamma_i \in \Gamma$ for some real set $\Gamma$. The code can be constructed as follows.

$$\mathcal{C}_{P,Q,\theta,\Gamma} = \left\{ \left[ \begin{array}{cc} \cos\gamma_i UV & \sin\gamma_i U\bar{V} \\ -\sin\gamma_i \bar{U}V & \cos\gamma_i \bar{U}\bar{V} \end{array} \right] \middle| \begin{array}{c} U = \frac{1}{\sqrt{2}} \left[ \begin{array}{cc} e^{j\frac{2\pi k}{P}} & e^{j\frac{2\pi l}{P}} \\ -e^{-j\frac{2\pi l}{P}} & e^{-j\frac{2\pi k}{P}} \end{array} \right], \\ V = \frac{1}{\sqrt{2}} \left[ \begin{array}{cc} e^{j(\frac{2\pi m}{Q}+\theta)} & e^{j(\frac{2\pi n}{Q}+\theta)} \\ -e^{-j(\frac{2\pi n}{Q}+\theta)} & e^{-j(\frac{2\pi m}{Q}+\theta)} \end{array} \right], \\ 0 \le k, l < P, 0 \le m, n < Q, \gamma_i \in \Gamma \end{array} \right\}, \quad (5.20)$$

where $P$ and $Q$ are positive integers and $\theta \in [0, 2\pi)$ is a constant to be chosen later. It can be easily seen that any signal matrix in the code can be identified by the 5-tuple $(k, l, m, n, \gamma_i)$. The code proposed in (5.6) is a special case of this code, which can be obtained by setting $\Gamma = \{\frac{\pi}{4}\}$. Since the set has altogether $P^2Q^2|\Gamma|$ matrices and the

channel is used in blocks of four transmissions, the rate of the code is

$$\frac{1}{2}(\log_2 P + \log_2 Q) + \frac{1}{4}\log_2 |\Gamma|, \tag{5.21}$$

where $|\Gamma|$ indicates the cardinality of the set $\Gamma$. A conclusion similar to Theorem 5.2 about the full diversity of the code can be stated as follows.

**Theorem 5.3 (Condition for full diversity).** *If $P$ and $Q$ are relatively prime with $P$ odd and the set $\Gamma$ satisfies the following conditions,*

1. $\Gamma \subset (0, \frac{\pi}{2})$,

2. *For any $\gamma \in \Gamma$,*

$$\tan\gamma \neq \pm\frac{\sin\pi(\frac{l}{P}+\frac{m}{Q})}{\sin\pi(\frac{k}{P}+\frac{m}{Q})}, \tag{5.22}$$

   *where $k, l \in (-P, P)$, $m \in (-Q, Q)$ are integers and $(k, m) \neq (0, 0)$,*

3. *For any two different $\gamma_i, \gamma_j \in \Gamma$,*

$$\frac{\sin 2\gamma_i}{\sin 2\gamma_j} \neq \frac{\cos 2\pi\frac{s}{P}}{\cos 2\pi\frac{l}{P}}, \tag{5.23}$$

   *where $s, l \in [0, P)$ are integers,*

*then there exists $\theta$ such that the signal set in (5.20) is fully-diverse.*

**Proof:** First we need to show that the right-hand side formula of (5.22) is well defined, that is, $\sin\pi(\frac{k}{P}+\frac{m}{Q}) \neq 0$ for any integers $k \in (-P, P)$, $m \in (-Q, Q)$, and $(k, m) \neq (0, 0)$. This can be proved by contradiction. Assume that $\sin\pi(\frac{k}{P}+\frac{m}{Q}) = 0$. Therefore, $\pi(\frac{k}{P}+\frac{m}{Q}) = l\pi$, which is equivalent to $\frac{k}{P}+\frac{m}{Q} = l$, for some integer $l$. Since $P$ and $Q$ are relatively prime, $P|k$ and $Q|m$. Since $k \in (-P, P)$ and $m \in (-Q, Q)$, $k = 0$ and $m = 0$. This contradicts the condition that $(k, m) \neq (0, 0)$. We now prove

that the right-hand side formula of (5.23) is well defined, that is, $\cos 2\pi \frac{l}{P} \neq 0$ for any $l \in [0, P)$. Again, this can be proved by contradiction. Assume that $\cos 2\pi \frac{l}{P} = 0$. Therefore, $2\pi \frac{l}{P} = r\pi + \frac{\pi}{2}$, that is, $\frac{l}{P} = \frac{2r+1}{4}$ for some integer $r$. Since $2r + 1$ is odd, $4|P$ and this contradicts the condition that $P$ is odd.

Now we prove this theorem. Assume that $P$ and $Q$ are relatively prime, $P$ is odd and the set $\Gamma \subset (0, \frac{\pi}{2})$ satisfies conditions (5.22) and (5.23). We want to show that there exists a $\theta$ such that the code is fully diverse. It is equivalent to show that for any non-identical pair of signals $S_i$ and $S_j$ of the code, $\det(S_i - S_j) \neq 0$. Without loss of generality, assume

$$S_i = \begin{bmatrix} \cos\gamma_i U_i V_i & \sin\gamma_i U_i \bar{V}_i \\ -\sin\gamma_i \bar{U}_i V_i & \cos\gamma_i \bar{U}_i \bar{V}_i \end{bmatrix} \text{ and } S_j = \begin{bmatrix} \cos\gamma_i U_j V_j & \sin\gamma_i U_j \bar{V}_j \\ -\sin\gamma_i \bar{U}_j V_j & \cos\gamma_i \bar{U}_j \bar{V}_j \end{bmatrix}, \quad (5.24)$$

where $U_i, V_i$ are $2 \times 2$ unitary matrices given by (5.9) and $U_j, V_j$ are the two $2 \times 2$ matrices in (5.9) by replacing $i$ by $j$. The two signals being different indicates that the two 5-tuples, $(k_i, l_i, m_i, n_i, \gamma_i)$ and $(k_j, l_j, m_j, n_j, \gamma_j)$, are not identical. From the proof of Lemma 5.6, $\det(S_i - S_j)$ is zero if and only if $O'_1 = \pm J\bar{O}'_2$, where $O'_1$ and $O'_2$ are defined as

$$O'_1 = \cos\gamma_i U_i V_i - \cos\gamma_j U_j V_j \text{ and } O'_2 = \sin\gamma_i U_i \bar{V}_i - \sin\gamma_j U_j \bar{V}_j. \quad (5.25)$$

By using (5.25) and (5.9), similar to the argument in the proof of Lemma 5.6, $O'_1 = \pm J\bar{O}'_2$ can be equivalently written as

$$\begin{cases} e^{2j\theta}(\cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{m_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{m_j}{Q})} \pm \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{m_i}{Q})} \mp \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{m_j}{Q})}) \\ \qquad = \mp\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{n_i}{Q})} \pm \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{n_j}{Q})} + \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{n_j}{Q})} \\ \text{or} \\ e^{2j\theta}(\cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{n_j}{Q})} \pm \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{n_i}{Q})} \mp \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{n_j}{Q})}) \\ \qquad = \pm\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{m_i}{Q})} \mp \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{m_j}{Q})} - \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{m_i}{Q})} + \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{m_j}{Q})} \end{cases} \quad (5.26)$$

Define

$$\begin{cases} \tilde{w}^+ = \cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{m_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{m_j}{Q})} + \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{m_i}{Q})} - \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{m_j}{Q})} \\ \tilde{x}^+ = -\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{n_i}{Q})} + \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{n_j}{Q})} + \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{n_j}{Q})} \\ \tilde{y}^+ = \cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{n_j}{Q})} + \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{n_i}{Q})} - \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{n_j}{Q})} \\ \tilde{z}^+ = +\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{m_i}{Q})} - \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{m_j}{Q})} - \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{m_i}{Q})} + \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{m_j}{Q})} \end{cases} \tag{5.27}$$

and

$$\begin{cases} \tilde{w}^- = \cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{m_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{m_j}{Q})} - \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{m_i}{Q})} + \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{m_j}{Q})} \\ \tilde{x}^- = +\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{n_i}{Q})} - \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{n_j}{Q})} + \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{n_j}{Q})} \\ \tilde{y}^- = \cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{n_j}{Q})} - \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{n_i}{Q})} + \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{n_j}{Q})} \\ \tilde{z}^- = -\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{m_i}{Q})} + \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{m_j}{Q})} - \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{m_i}{Q})} + \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{m_j}{Q})} \end{cases} \tag{5.28}$$

(5.26) is thus equivalent to,

$$\begin{cases} e^{2j\theta}\tilde{w}^+ = \tilde{x}^+ \\ e^{2j\theta}\tilde{y}^+ = \tilde{z}^+ \end{cases} \quad \text{or} \quad \begin{cases} e^{2j\theta}\tilde{w}^- = \tilde{x}^- \\ e^{2j\theta}\tilde{y}^- = \tilde{z}^- \end{cases} . \tag{5.29}$$

Now we need the following lemma.

**Lemma 5.9.** *For any non-identical pairs $(k_i, l_i, m_i, n_i, \gamma_i)$ and $(k_j, l_j, m_j, n_j, \gamma_j)$, where $k_i, l_i, k_j, l_j \in [0, P)$, $m_i, n_i, m_j, n_j \in [0, Q)$ are integers and $\gamma_i, \gamma_j \in \Gamma$, if $P$ and $Q$ are relatively prime with $P$ odd and if the set $\Gamma \subset (0, \frac{\pi}{2})$ satisfies conditions (5.22) and (5.23), then $\tilde{w}^+$, $\tilde{x}^+$, $\tilde{y}^+$, $\tilde{z}^+$ (as defined in (5.27)) cannot be zero simultaneously. Also $\tilde{w}^-$, $\tilde{x}^-$, $\tilde{y}^-$, $\tilde{z}^-$ (as defined in (5.28)) cannot be zero simultaneously.*

**Proof:** See Section 5.9.4. □

The proof is very similar to the proof of Theorem 5.2. By Lemma 5.9, $\tilde{w}^+$, $\tilde{x}^+$,

$\tilde{y}^+$, $\tilde{z}^+$ cannot be zero simultaneously. Define

$$\tilde{\theta}^+_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{x}^+}{\tilde{w}^+}) & \mod 2\pi \ \text{ if } \ \tilde{w}^+ \neq 0 \\ -\frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{w}^+}{\tilde{x}^+}) & \mod 2\pi \ \text{ if } \ \tilde{w}^+ = 0, \tilde{x}^+ \neq 0 \\ \frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{z}^+}{\tilde{y}^+}) & \mod 2\pi \ \text{ if } \ \tilde{w}^+ = \tilde{x}^+ = 0, \tilde{y}^+ \neq 0 \\ -\frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{y}^+}{\tilde{z}^+}) & \mod 2\pi \ \text{ if } \ \tilde{w}^+ = \tilde{x}^+ = \tilde{y}^+ = 0, \tilde{z}^+ \neq 0 \end{cases},$$

which is the same as

$$\tilde{\theta}^+_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{x}^+}{\tilde{w}^+}) & \mod 2\pi \ \text{ if } \ \tilde{w}^+ \neq 0 \\ 0 & \text{ if } \ \tilde{w}^+ = 0, \tilde{x}^+ \neq 0 \\ \frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{z}^+}{\tilde{y}^+}) & \mod 2\pi \ \text{ if } \ \tilde{w}^+ = \tilde{x}^+ = 0, \tilde{y}^+ \neq 0 \\ 0 & \text{ if } \ \tilde{w}^+ = \tilde{x}^+ = \tilde{y}^+ = 0, \tilde{z}^+ \neq 0 \end{cases}.$$

Also from Lemma 5.9, $\tilde{w}^-$, $\tilde{x}^-$, $\tilde{y}^-$, $\tilde{z}^-$ cannot be zeros simultaneously. Define

$$\tilde{\theta}^-_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{x}^-}{\tilde{w}^-}) & \mod 2\pi \ \text{ if } \ \tilde{w}^- \neq 0 \\ -\frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{w}^-}{\tilde{x}^-}) & \mod 2\pi \ \text{ if } \ \tilde{w}^- = 0, \tilde{x}^- \neq 0 \\ \frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{z}^-}{\tilde{y}^-}) & \mod 2\pi \ \text{ if } \ \tilde{w}^- = \tilde{x}^- = 0, \tilde{y}^- \neq 0 \\ -\frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{y}^-}{\tilde{z}^-}) & \mod 2\pi \ \text{ if } \ \tilde{w}^- = \tilde{x}^- = \tilde{y}^- = 0, \tilde{z}^- \neq 0 \end{cases},$$

which is the same as

$$\tilde{\theta}^-_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} -\frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{x}^-}{\tilde{w}^-}) & \mod 2\pi \ \text{ if } \ \tilde{w}^- \neq 0 \\ 0 & \text{ if } \ \tilde{w}^- = 0, \tilde{x}^- \neq 0 \\ -\frac{1}{2}\mathrm{Arg}\,(\frac{\tilde{z}^-}{\tilde{y}^-}) & \mod 2\pi \ \text{ if } \ \tilde{w}^- = \tilde{x}^- = 0, \tilde{y}^- \neq 0 \\ 0 & \text{ if } \ \tilde{w}^- = \tilde{x}^- = \tilde{y}^- = 0, \tilde{z}^- \neq 0 \end{cases}.$$

By choosing

$$\theta \notin \left\{ \tilde{\theta}^{+}_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} \;\middle|\; |\tilde{w}^{+}| = |\tilde{x}^{+}|, |\tilde{y}^{+}| = |\tilde{z}^{+}|, 0 \leq k_1, l_1, k_2, l_2 < P, 0 \leq m_1, n_1, m_2, n_2 < Q \right\} \tag{5.30}$$

and

$$\theta \notin \left\{ \tilde{\theta}^{-}_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} \;\middle|\; |\tilde{w}^{-}| = |\tilde{x}^{-}|, |\tilde{y}^{-}| = |\tilde{z}^{-}|, 0 \leq k_1, l_1, k_2, l_2 < P, 0 \leq m_1, n_1, m_2, n_2 < Q \right\} \tag{5.31}$$

(5.29) cannot be true. Therefore $\det(S_i - S_j) \neq 0$, which means that the code is fully-diverse. An angle in $[0, 2\pi)$ that satisfies both (5.30) and (5.31) can always be found since the two sets at the right-hand side of (5.30) and (5.31) are finite. □

## 5.6 Decoding of $Sp(2)$ Codes

One of the most prominent properties of our $Sp(2)$ codes is that it is a generalization of the orthogonal design. In this section, it is shown how this property can be used to get linear-algebraic decoding, which means that the receiver can be made to form a system of linear equations in the unknowns.

### 5.6.1 Formulation

The ML decoding for differential USTM is given in (2.12), which, in our system, can be written as

$$\begin{aligned} & \arg \max_{l=0,\ldots,L-1} \|X_\tau - V_l X_{\tau-1}\|_F^2 \\ = \; & \arg \max_{i,j} \left\| X_\tau - \frac{1}{\sqrt{2}} \begin{bmatrix} U_i & 0 \\ 0 & \bar{U}_i \end{bmatrix} \begin{bmatrix} I_2 & I_2 \\ -I_2 & I_2 \end{bmatrix} \begin{bmatrix} V_j & 0 \\ 0 & \bar{V}_j \end{bmatrix} X_{\tau-1} \right\|_F^2 \end{aligned}$$

$$= \arg\max_{i,j} \left\| \begin{bmatrix} U_i^* & 0 \\ 0 & U_i^t \end{bmatrix} X_\tau - \frac{1}{\sqrt{2}} \begin{bmatrix} V_j & \bar{V}_j \\ -V_j & \bar{V}_j \end{bmatrix} X_{\tau-1} \right\|_F^2 .$$

By writing the matrices in the norm column by column, we get

$$\arg\max_{i,j} \left\| \begin{bmatrix} \begin{bmatrix} U_i^* & 0 \\ 0 & U_i^t \end{bmatrix} X_1^{(\tau)} \\ \vdots \\ \begin{bmatrix} U_i^* & 0 \\ 0 & U_i^t \end{bmatrix} X_N^{(\tau)} \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} \begin{bmatrix} V_j & \bar{V}_j \\ -V_j & \bar{V}_j \end{bmatrix} X_1^{(\tau-1)} \\ \vdots \\ \begin{bmatrix} V_j & \bar{V}_j \\ -V_j & \bar{V}_j \end{bmatrix} X_N^{(\tau-1)} \end{bmatrix} \right\|_F^2 ,$$

where $X_i^{(\tau)}$ denotes the $i$-th column of $X_\tau$ and $X_i^{(\tau-1)}$ denotes the $i$-th column of $X_{\tau-1}$. It is obvious that $X_i^{(\tau)}$ and $X_i^{(\tau-1)}$ are $4 \times 1$ column vectors. We further denote the $(i,j)$-th entry of $X_\tau$ as $x_{ij}^{(\tau)}$ and denote the $(i,j)$-th entry of $X_\tau$ as $x_{ij}^{(\tau-1)}$ for $i = 1, 2, 3, 4$ and $j = 1, 2, \cdots, N$. The ML decoding is equivalent to

$$\arg\max_{i,j} \left\| \begin{bmatrix} \begin{bmatrix} U_i^* \left[x_{11}^{(\tau)}, x_{21}^{(\tau)}\right]^t \\ U_j^t \left[x_{31}^{(\tau)}, x_{41}^{(\tau)}\right]^t \end{bmatrix} \\ \vdots \\ \begin{bmatrix} U_i^* \left[x_{N1}^{(\tau)}, x_{N2}^{(\tau)}\right]^t \\ U_j^t \left[x_{N3}^{(\tau)}, x_{N4}^{(\tau)}\right]^t \end{bmatrix} \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} \begin{bmatrix} V_j \left[x_{11}^{(\tau-1)}, x_{21}^{(\tau-1)}\right]^t + \bar{V}_j \left[x_{31}^{(\tau-1)}, x_{41}^{(\tau-1)}\right]^t \\ -V_j \left[x_{11}^{(\tau-1)}, x_{21}^{(\tau-1)}\right]^t + \bar{V}_j \left[x_{31}^{(\tau-1)}, x_{41}^{(\tau-1)}\right]^t \end{bmatrix} \\ \vdots \\ \begin{bmatrix} V_j \left[x_{1N}^{(\tau-1)}, x_{2N}^{(\tau-1)}\right]^t + \bar{V}_j \left[x_{3N}^{(\tau-1)}, x_{4N}^{(\tau-1)}\right]^t \\ -V_j \left[x_{1N}^{(\tau-1)}, x_{2N}^{(\tau-1)}\right]^t + \bar{V}_j \left[x_{3N}^{(\tau-1)}, x_{4N}^{(\tau-1)}\right]^t \end{bmatrix} \end{bmatrix} \right\|_F^2 .$$

From the design of the code, we know that the matrices $U_i, V_j$ and their conjugates and transposes are all orthogonal designs. For any orthogonal design $M = \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$

and any two-dimensional vector $X = [x_1, x_2]^t$, $MX$ can be written equivalently as

$$\begin{bmatrix} \Re x_1 & \Im x_1 & -\Re x_2 & \Im x_2 \\ \Re x_2 & -\Im x_2 & \Re x_1 & \Im x_1 \end{bmatrix} \begin{bmatrix} \Re a \\ \Im a \\ \Re b \\ \Im b \end{bmatrix} + i \begin{bmatrix} \Im x_1 & -\Re x_1 & -\Im x_2 & -\Re x_2 \\ \Im x_2 & \Re x_2 & \Im x_1 & -\Re x_1 \end{bmatrix} \begin{bmatrix} \Re a \\ \Im a \\ \Re b \\ \Im b \end{bmatrix},$$

where $\Re x$ indicates the real part of $x$ and $\Im x$ indicates the imaginary part of $x$. It can be seen that the roles of $M$ and $X$ are interchanged. Therefore, by careful calculation, the ML decoding of $Sp(2)$ codes can be shown to be equivalent to

$$\arg \max_{0 \le k,l < P, 0 \le m,n < Q} \left\| \begin{bmatrix} \begin{bmatrix} \mathcal{A}_1(X_1^{(\tau)}) & -\mathcal{C}_1(X_1^{(\tau-1)}) \\ \mathcal{B}_1(X_1^{(\tau)}) & -\mathcal{D}_1(X_1^{(\tau-1)}) \end{bmatrix} \\ \cdots \\ \begin{bmatrix} \mathcal{A}_N(X_N^{(\tau)}) & -\mathcal{C}_N(X_N^{(\tau-1)}) \\ \mathcal{B}_N(X_N^{(\tau)}) & -\mathcal{D}_N(X_N^{(\tau-1)}) \end{bmatrix} \end{bmatrix} \begin{bmatrix} \cos\frac{2\pi k}{P} \\ \sin\frac{2\pi k}{P} \\ \cos\frac{2\pi l}{P} \\ \sin\frac{2\pi l}{P} \\ \cos(\frac{2\pi m}{Q} + \theta) \\ \sin(\frac{2\pi m}{Q} + \theta) \\ \cos(\frac{2\pi n}{Q} + \theta) \\ \sin(\frac{2\pi n}{Q} + \theta) \end{bmatrix} \right\|_F^2, \quad (5.32)$$

where

$$\mathcal{A}_i(X_i^{(\tau)}) = \frac{1}{\sqrt{2}} \begin{bmatrix} \Re x_{i1}^{(\tau)} & \Im x_{i1}^{(\tau)} & -\Re x_{i2}^{(\tau)} & \Im x_{i2}^{(\tau)} \\ \Re x_{i2}^{(\tau)} & -\Im x_{i2}^{(\tau)} & \Re x_{i1}^{(\tau)} & \Im x_{i1}^{(\tau)} \\ \Re x_{i3}^{(\tau)} & -\Im x_{i3}^{(\tau)} & -\Re x_{i4}^{(\tau)} & -\Im x_{i4}^{(\tau)} \\ \Re x_{i4}^{(\tau)} & \Im x_{i4}^{(\tau)} & \Re x_{i3}^{(\tau)} & -\Im x_{i3}^{(\tau)} \end{bmatrix}, \quad (5.33)$$

$$\mathcal{B}_i(X_i^{(\tau)}) = \frac{1}{\sqrt{2}} \begin{bmatrix} \Im x_{i1}^{(\tau)} & -\Re x_{i1}^{(\tau)} & -\Im x_{i2}^{(\tau)} & -\Re x_{i2}^{(\tau)} \\ \Im x_{i2}^{(\tau)} & \Re x_{i2}^{(\tau)} & \Im x_{i1}^{(\tau)} & -\Re x_{i1}^{(\tau)} \\ \Im x_{i3}^{(\tau)} & \Re x_{i3}^{(\tau)} & -\Im x_{i4}^{(\tau)} & \Re x_{i4}^{(\tau)} \\ \Im x_{i4}^{(\tau)} & -\Re x_{i4}^{(\tau)} & \Im x_{i3}^{(\tau)} & \Re x_{i3}^{(\tau)} \end{bmatrix}, \tag{5.34}$$

$$\mathcal{C}_i(X_i^{(\tau-1)}) = \tfrac{1}{2} \begin{bmatrix} \Re(x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) & \Im(-x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) \\ \Re(x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Im(x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) \\ \Re(-x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) & \Im(x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) \\ \Re(-x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Im(-x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) \end{bmatrix} \begin{bmatrix} \Re(x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Im(-x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) \\ \Re(-x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) & \Im(-x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) \\ \Re(-x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Im(x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) \\ \Re(x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) & \Im(x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) \end{bmatrix}, \tag{5.35}$$

$$\mathcal{D}_i(X_i^{(\tau-1)}) = \tfrac{1}{2} \begin{bmatrix} \Im(x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) & \Re(x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) \\ \Im(x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Re(-x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) \\ \Im(-x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) & \Re(-x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) \\ \Im(-x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Re(x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) \end{bmatrix} \begin{bmatrix} \Im(x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Re(x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) \\ \Im(-x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) & \Re(x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) \\ \Im(-x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Re(-x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) \\ \Im(x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) & \Re(-x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) \end{bmatrix}, \tag{5.36}$$

and $\left[\cos\frac{2\pi k}{P}, \sin\frac{2\pi k}{P}, \cos\frac{2\pi l}{P}, \sin\frac{2\pi l}{P}, \cos(\frac{2\pi m}{Q}+\theta), \sin(\frac{2\pi m}{Q}+\theta), \cos(\frac{2\pi n}{Q}+\theta), \sin(\frac{2\pi n}{Q}+\theta)\right]^t$ is the vector of unknowns. Notice that $\mathcal{B}_i(X_i^{(\tau)})$ can also be constructed as $\mathcal{A}_i(-jX_i^{(\tau)})$. It can be seen that formula (5.32) is quadratic in sines and cosines of the unknowns.

Thus, it is possible to use fast decoding algorithms such as sphere decoding to achieve the exact ML solution in polynomial time.

In this paragraph, the sphere decoding for codes given in (5.20) is discussed. For each of the angle $\gamma_i$, a sphere decoding is applied and one signal is gained. In doing the sphere decoding for each $\gamma_i$, matrices $\mathcal{A}_i, \mathcal{B}_i$ for $i = 1, 2, \cdots, N$ are the same as those in (5.33) and (5.34), but the $\mathcal{C}_i$ and $\mathcal{D}_i$ matrices should be modified to

$$\mathcal{C}_i = \begin{bmatrix} \Re(\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) & \Im(-\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) \\ \Re(\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Im(\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Re(-\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) & \Im(\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) \\ \Re(-\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Im(-\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Re(\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Im(-\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Re(-\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) & \Im(-\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) \\ \Re(-\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Im(\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Re(\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) & \Im(\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) \end{bmatrix} \tag{5.37}$$

and

$$\mathcal{D}_i = \begin{bmatrix} \Im(\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) & \Re(\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) \\ \Im(\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Re(-\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Im(-\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) & \Re(-\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) \\ \Im(-\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Re(\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Im(\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Re(\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Im(-\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) & \Re(\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) \\ \Im(-\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Re(-\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Im(\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) & \Re(-\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) \end{bmatrix}. \tag{5.38}$$

For each transmission, $|\Gamma|$ sphere decoding is used and therefore $|\Gamma|$ signal matrices are obtained in total. ML decoding given in (5.32) is then used to get the optimal

one. The complexity of this decoding algorithm is $|\Gamma|$ times the original one, but it is still cubic polynomial in the transmission rate and dimension.

### 5.6.2 Remarks on Sphere Decoding

Below are some remarks on the implementation of sphere decoding in our systems.

1. The main idea of sphere decoding is disucssed in Section 2.8. The choice of the searching radius is very crucial to the speed of the algorithm. Here, the radius is initialized as a small value and then increase it gradually based on the noise level [HV02]. The searching radius $\sqrt{C}$ is initialized in such a way that the probability that the correct signal is in the sphere is 0.9, that is,

$$\mathrm{P}\left(\|v\|_F < \sqrt{C}\right) = 0.9. \tag{5.39}$$

   If no point is found in this sphere, the searching radius is then raised such that the probability is increased to 0.99 and so on. Using this algorithm, the probability that a point can be found during the first search is high. The noise of the system is given in (2.11). Since $W_\tau$, $W_{\tau-1}$ and the transmitted unitary matrix $V_{z_\tau}$ are independent, it is easy to prove that the noise has mean zero and variance $2NI_4$. Each component of the $4 \times N$-dimensional noise vector has mean zero and variance 2. Therefore the random variable $v = \|W'_\tau\|_F^2$ has Gamma distribution with mean $4N$. The value of $C$ that satisfies (5.39) can be easily calculated.

2. From (5.32), it can be seen that the unknowns are in forms of sines and cosines. Notice that for any $\alpha \neq \beta \in [0, 2\pi)$, $\sin\alpha = \sin\beta$ if and only if $\beta = (2k+1)\pi - \alpha$ for some integer $k$. When $P, Q$ are odd, it can be seen that $\pi$ cannot be in the set $\Theta_P = \{\frac{2\pi k}{P} | k = 0, 1, 2, ..., P-1\}$, which is the set of all possible angles of $U_i$'s entries, and $\pi + \theta$ cannot be in the set

$\Theta_Q = \{\frac{2\pi k}{Q} + \theta | k = 0, 1, 2, ..., P-1\}$, which is the set of all possible angles of $V_i$'s entries. Therefore, the map $f_P : \Theta_P \to \{\sin x | x \in \Theta_P\}$ by $f_P(\theta) = \sin\theta$ and the map $f_Q : \Theta_Q \to \{\sin x | x \in \Theta_Q\}$ by $f_Q(\theta) = \sin\theta$ are one-to-one and onto. The independent unknowns $k, l, m, n$ can thus be replaced equivalently by their sines: $\sin\frac{2\pi k}{P}, \sin\frac{2\pi l}{P}, \sin(\frac{2\pi m}{Q} + \theta), \sin(\frac{2\pi n}{Q} + \theta)$.

3. Notice that there are only four independent unknowns but eight components in the unknown vector in (5.32). We combine the $2i$-th component (with the form of $\cos x$) and the $(2i+1)$-th component (with the form of $\sin x$) for $i = 1, 2, 3, 4$. From previous discussions we know that for any value in the set $\{\sin x | x \in \Theta_P\}$ or $\{\sin x | x \in \Theta_Q\}$, there is only one possible value in $\Theta_P$ or $\Theta_Q$ whose sine equals the value. Therefore, there is only one possible value of the cosine. In other words, for any possible value of the $2i$-th component, there is one unique value for the $(2i+1)$-th component. Therefore, it is natural to combine the $2i$-th and the $(2i+1)$-th components. To simplify the programming, while considering the searching range of each unknown variable, we skip the $2i$-th component and only consider the $(2i+1)$-th one. For example, instead of analyzing all the possible values of $\sin\frac{2\pi n}{Q}$ (the 8th component) satisfying $q_{88}\sin^2\frac{2\pi n}{Q} + q_{77}(\cos\frac{2\pi n}{Q} + S_8)^2 \leq C$, all the possible values of $\sin\frac{2\pi n}{Q}$ satisfying $q_{88}\sin^2\frac{2\pi n}{Q} \leq C$ are considered [DAML00]. It may seem that more points than needed is searched, but actually the extra points will be eliminated in the next step of the sphere decoding algorithm.

4. Complex sphere decoding can also be used here to obtain ML results, which actually is simplier than the real sphere decoding. However, in all the simulations, real sphere decoding is used.

## 5.7 Simulation Results

In this section, examples of $Sp(2)$ codes and also the simulated performance of the codes at different rates are given. The fading coefficient between every transmit-and-receive-antenna pair is modeled independently as a complex Gaussian variable with zero-mean and unit-variance and keeps constant for $2M$ channel uses. At each channel use, zero-mean, unit-variance complex Gaussian noise is added to each received signal. The block error rate (BLER), which corresponds to errors in decoding the $4 \times 4$ transmitted matrices, is demonstrated as the error event of interest. The performance of the $Sp(2)$ codes is also compared with that of some group-based codes [SHHS01], the differential Cayley codes [HH02a], the $2\times2$ Alamouti's complex orthogonal designs [Ala98] of the form

$$\mathcal{C}(z_1, z_2) = \begin{bmatrix} z_1 & z_2 \\ -z_2^* & z_1^* \end{bmatrix}, \tag{5.40}$$

and the $4 \times 4$ complex orthogonal design

$$\mathcal{C}(z_1, z_2, z_3) = \begin{bmatrix} z_1 & z_2 & z_3 & 0 \\ -z_2^* & z_1^* & 0 & -z_3 \\ -z_3^* & 0 & z_1^* & z_2 \\ 0 & z_3^* & -z_2^* & z_2 \end{bmatrix} \tag{5.41}$$

proposed in [TH02].

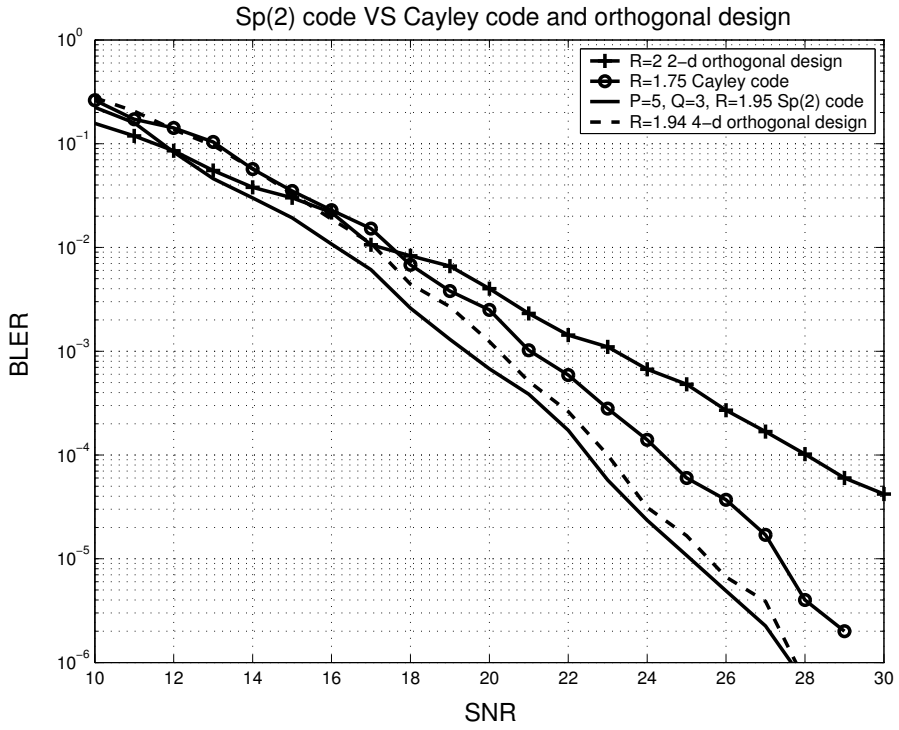


Figure 5.3: Comparison of the rate 1.95 $Sp(2)$ code with the rate 1.75 differential Cayley code, the rate 2, $2 \times 2$ complex orthogonal design, and the rate 1.94, $4 \times 4$ complex orthogonal design with $N = 1$ receive antennas

### 5.7.1 $Sp(2)$ Code vs. Cayley Code and Complex Orthogonal Designs

The first example is the $Sp(2)$ code with $P = 5, Q = 3, \theta = 0$, that is, entries of the $U$ matrix of the code are chosen from the 5-PSK signal set $\{1, e^{j\frac{2\pi}{5}}, e^{j\frac{4\pi}{5}}, e^{j\frac{6\pi}{5}}, e^{j\frac{8\pi}{5}}\}$, and entries of the $V$ matrix are chosen from the 3-PSK signal set $\{1, e^{j\frac{2\pi}{3}}, e^{j\frac{4\pi}{3}}\}$. Therefore by (5.7), the rate of the code is 1.95. It is compared with 3 code: the rate 2, $2 \times 2$ complex orthogonal design given by (5.40), where $z_1, z_2$ are chosen from the 4-PSK signal set $\{1, e^{j\frac{2\pi}{4}}, e^{j\frac{4\pi}{4}}, e^{j\frac{6\pi}{4}}\}$; a rate 1.75 differential Cayley code with parameters $Q = 7, r = 2$ [HH02a]; and also the rate 1.94, $4 \times 4$ complex orthogonal design given by (5.41), where $z_1, z_2, z_3$ are chosen from the 6-PSK signal set $\{1, e^{j\frac{2\pi}{6}}, e^{j\frac{4\pi}{6}}, e^{j\frac{6\pi}{6}}, e^{j\frac{8\pi}{6}}, e^{j\frac{10\pi}{6}}\}$. The number of receive antennas is 1. The performance curves are shown in Figure 5.3. The solid line indicates the BLER of the $Sp(2)$ code. The lines with circles indicates the BLER of the differential Cayley code. The line with plus signs and the dashed line show the BLER of the $2 \times 2$ and $4 \times 4$ complex orthogonal designs, respectively. From the plot, it can be seen that the $Sp(2)$ code has the lowest BLER at high SNR. For example, at a BLER of $10^{-3}$, the $Sp(2)$ code is 2dB better than the differential Cayley code, even though the Cayley code has a lower rate, 1dB better than the $4 \times 4$ complex orthogonal design, and 4dB better than the $2 \times 2$ complex orthogonal design.

### 5.7.2 $Sp(2)$ Code vs. Finite-Group Constellations

In this subsection, the same $Sp(2)$ code is compared with a group-based diagonal code and the $K_{1,1,-1}$ code both at rate 1.98 [SHHS01]. The $K_{1,1,-1}$ code is in one of the 6 types of the finite fixed-point-free groups given in [SHHS01]. The number of receive antennas is 1. In Figure 5.4, the solid line indicates the BLER of the $Sp(2)$ code and the line with circles and plus signs show the BLER of the $K_{1,1,-1}$ code and

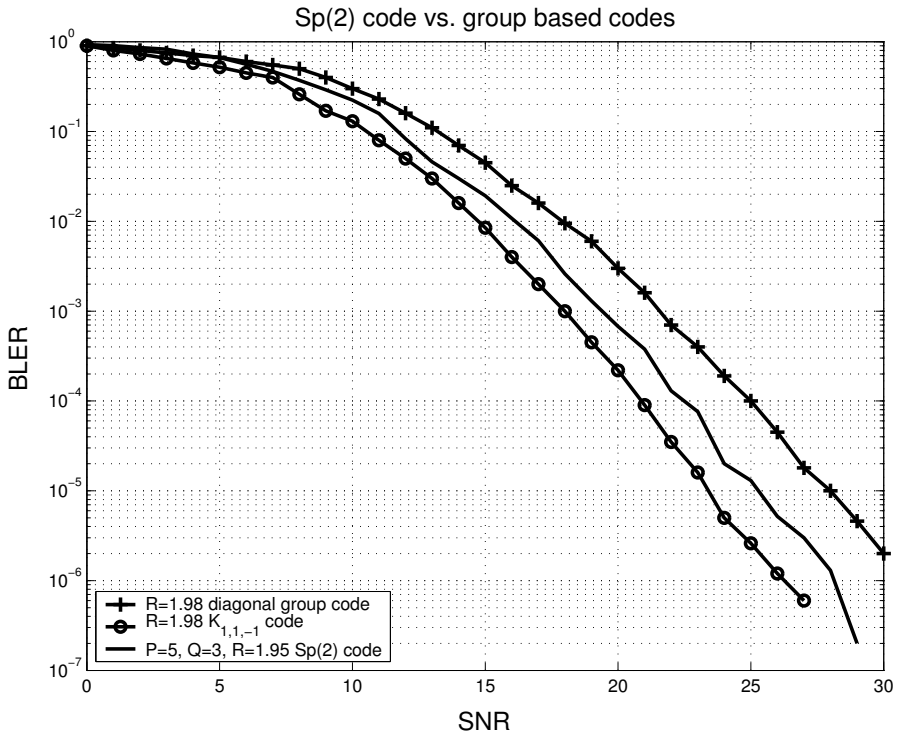


Figure 5.4: Comparison of the rate 1.95 $Sp(2)$ code with the rate 1.98 group-based $K_{1,1,-1}$ code and a rate 1.98 group-based diagonal code with $N = 1$ receive antennas

the diagonal code, respectively. The plot indicates that the $Sp(2)$ code is better than the diagonal code but worse than $K_{1,1,-1}$ code according to the BLER. For example, at a BLER of $10^{-3}$, the $Sp(2)$ code is 2dB better than the diagonal code, but 1.5dB worse than $K_{1,1,-1}$ group code. However, decoding $K_{1,1,-1}$ code requires an exhaustive search over the entire constellation.

### 5.7.3 $Sp(2)$ Codes vs. Complex Orthogonal Designs

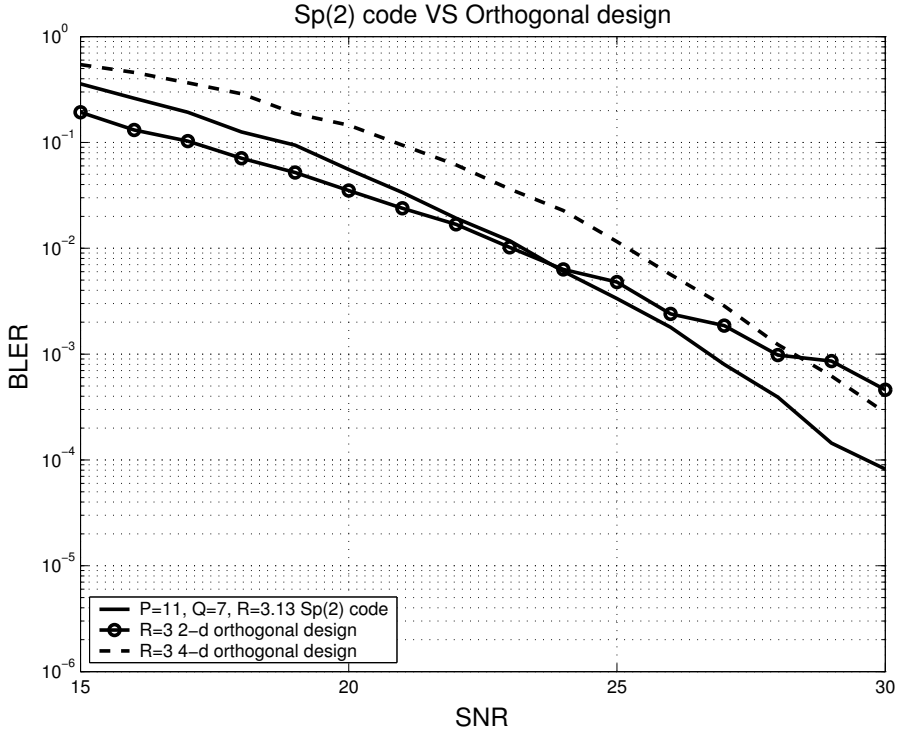


Figure 5.5: Comparison of the rate 3.13 $Sp(2)$ code with the rate 3, $2 \times 2$ and $4 \times 4$ complex orthogonal designs with $N = 1$ receive antenna

The comparison of the $Sp(2)$ codes with complex orthogonal designs at rate approximately 3 and 4 is shown in Figures 5.5 and 5.6. In Figure 5.5, the solid line indicates the BLER of the $Sp(2)$ code of $P = 11, Q = 7, \theta = 0$. The line with circles shows the BLER of the $2 \times 2$ complex orthogonal design (5.40) with $z_1, z_2$ chosen from 8-PSK. The dashed line indicates the BLER of the rate 3, $4 \times 4$ complex orthogonal

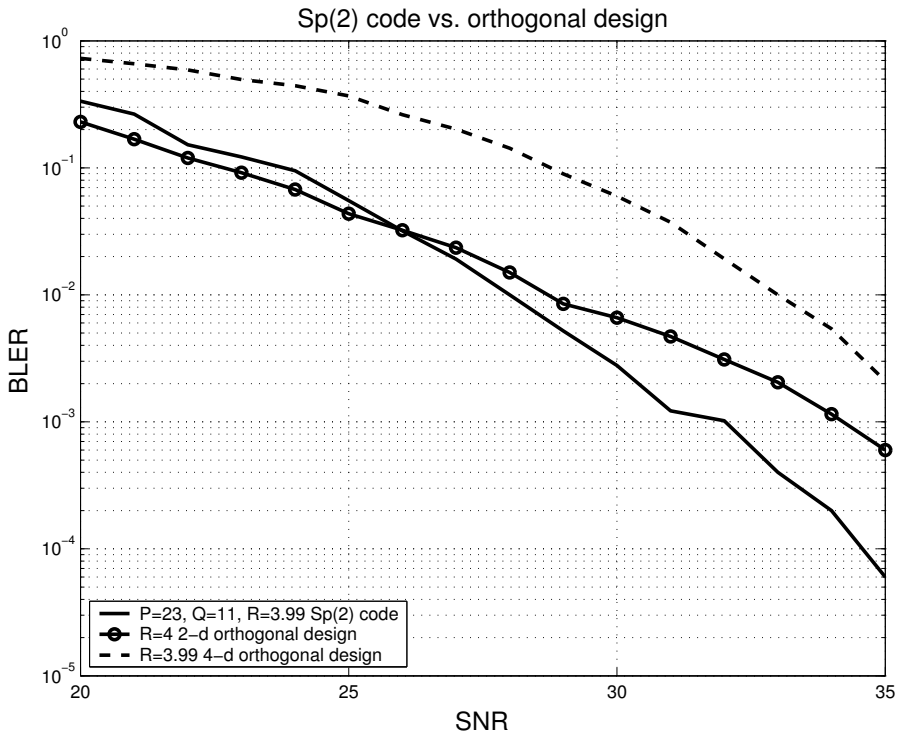


Figure 5.6: Comparison of the rate 3.99 $Sp(2)$ code with the rate 4, $2 \times 2$ and rate 3.99, $4 \times 4$ complex orthogonal designs with $N = 1$ receive antenna

design (5.41) with $z_1, z_2, z_3$ chosen from 16-PSK. Therefore, the rate of the $Sp(2)$ code is 3.13 and the rate of the $2 \times 2$ and $4 \times 4$ orthogonal designs is 3. Similarly, in Figure 5.6, the solid line indicates the BLER of the $Sp(2)$ code of $P = 23, Q = 11, \theta = 0$. The line with circles shows the BLER of the $2 \times 2$ complex orthogonal design (5.40) with $z_1, z_2$ chosen from 16-PSK and the dashed line indicates the BLER of the $4 \times 4$ complex orthogonal design (5.41) with $z_1, z_2, z_3$ chosen from 40-PSK. Therefore, the rate of the $Sp(2)$ code is 3.99 and the rates of the $2 \times 2$ and $4 \times 4$ complex orthogonal designs are 4 and 3.99. The number of receive antennas is 1. It can be seen from the two figures that the $Sp(2)$ codes are better than the $4 \times 4$ complex orthogonal designs for all the SNRs and are better than the $2 \times 2$ complex orthogonal designs at high SNR.

### 5.7.4 Performance of $Sp(2)$ Codes at Higher Rates

In this subsection, simulated performances of the $Sp(2)$ codes at higher rates, as given in (5.20), are shown for different $\Gamma$ and are compared with the corresponding original codes given in (5.6), whose $\Gamma$ is $\{\frac{\pi}{4}\}$.

The first example is the $Sp(2)$ code with $P = 11, Q = 7, \theta = 0, \Gamma = \{\frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}\} + 0.012$ and $\Gamma = \{\frac{\pi}{12}, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{5\pi}{12}\} + 0.02$. A small value is added to the set that is uniform on $(0, \frac{\pi}{2})$ to make the resulting $\Gamma$ set satisfy conditions (5.22) and (5.23), that is, to guarantee the full diversity of the code. According to (5.21), the rates of the codes are 3.5296 and 3.7139. In Figure 5.7, the dashed line and the line with plus signs indicate the BLER of the $Sp(2)$ codes that are just mentioned, which we call the new codes, and the solid line shows the BLER of the $P = 11, Q = 7, \theta = 0$ $Sp(2)$ code with $\Gamma = \{\frac{\pi}{4}\}$ and rate 3.1334, which we call the original code. The figure shows that the new codes are about only 1dB and 2dB worse than the original one with rates 0.3962 and 0.5805 higher. The BLER of the rate 4 non-group code given in [SHHS01], which has the structure of product-of-groups, is also shown in the figure by the line with circles. It can be seen that performance of the new code is very close to that of the non-group code with rate 0.4704 lower. The result is actually encouraging since the design of the non-group is very difficult and its decoding needs exhaustive search over $2^{16} = 65,536$ possible signal matrices.

The second example is the $P = 9, Q = 5, \theta = 0.0377, \Gamma = \{\frac{\pi}{12}, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{5\pi}{12}\} + 0.016$ $Sp(2)$ code. The rate of the code is therefore 3.3264 by formula (5.21). In Figure 5.8, the dashed line indicates the BLER of the rate 3.3264 $Sp(2)$ code we just mention, which we call the new code, and the solid line shows the BLER of the $P = 9, Q = 5, \theta = 0.0377$ $Sp(2)$ with $\Gamma = \{\frac{\pi}{4}\}$ and rate 2.7459, which we call the original code. The figure shows that the new code is only about 2dB worse than the original one with rate 0.5805 higher. Also, the BLER of the rate 4 non-group code given in [HH02a] is shown in the figure by the line with circles. It can be seen that the new code is

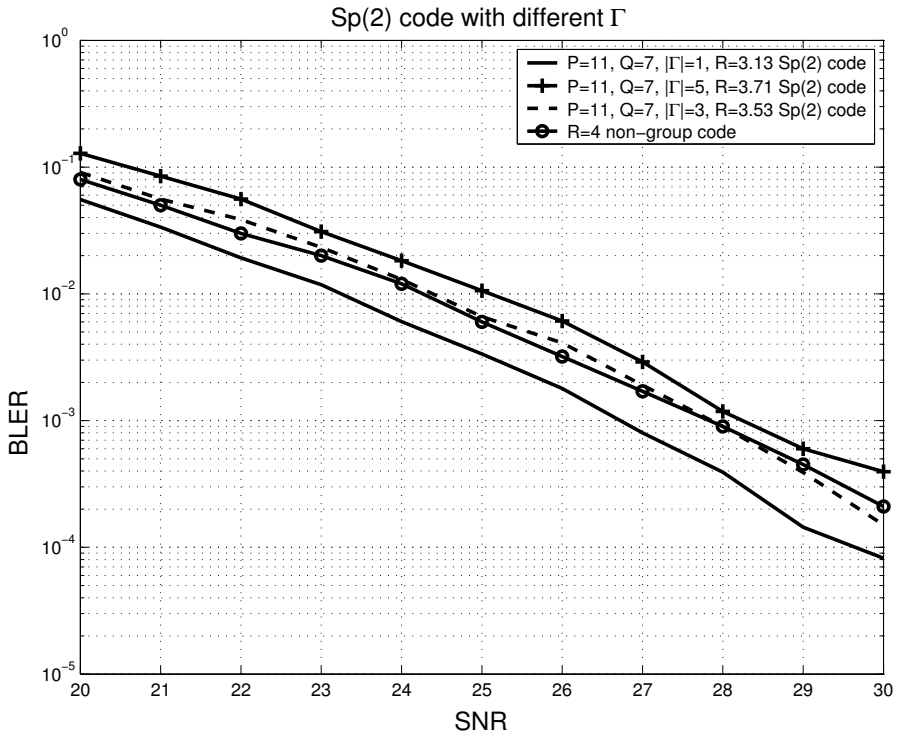


Figure 5.7: Comparison of $P = 11, Q = 7, \theta = 0$ $Sp(2)$ codes of $\Gamma = \{\frac{\pi}{4}\}$, $R = 3.1334$, $\Gamma = \{\frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}\} + 0.012$, $R = 3.5296$, and $\Gamma = \{\frac{\pi}{12}, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{5\pi}{12}\} + 0.02$, $R = 3.7139$ with the non-group code

1dB better than the non-group code with rate 0.6736 lower. As mentioned before, the result is actually encouraging since the design of the non-group is very difficult and its decoding needs exhaustive search over $2^{16} = 65,536$ possible signal matrices.

## 5.8 Conclusion

In this chapter, the symplectic group $Sp(n)$, which has dimension $n(2n + 1)$ and rank $n$, are analyzed and differential USTM codes based on $Sp(2)$ are designed. The group, $Sp(2)$, is not fpf, but a method to design fully-diverse codes which are subsets of the group are proposed. The constellations designed are suitable for systems with four transmit antennas and any number of receive antennas. The special symplectic structure of the codes lend themselves to decoding by linear-algebraic techniques,

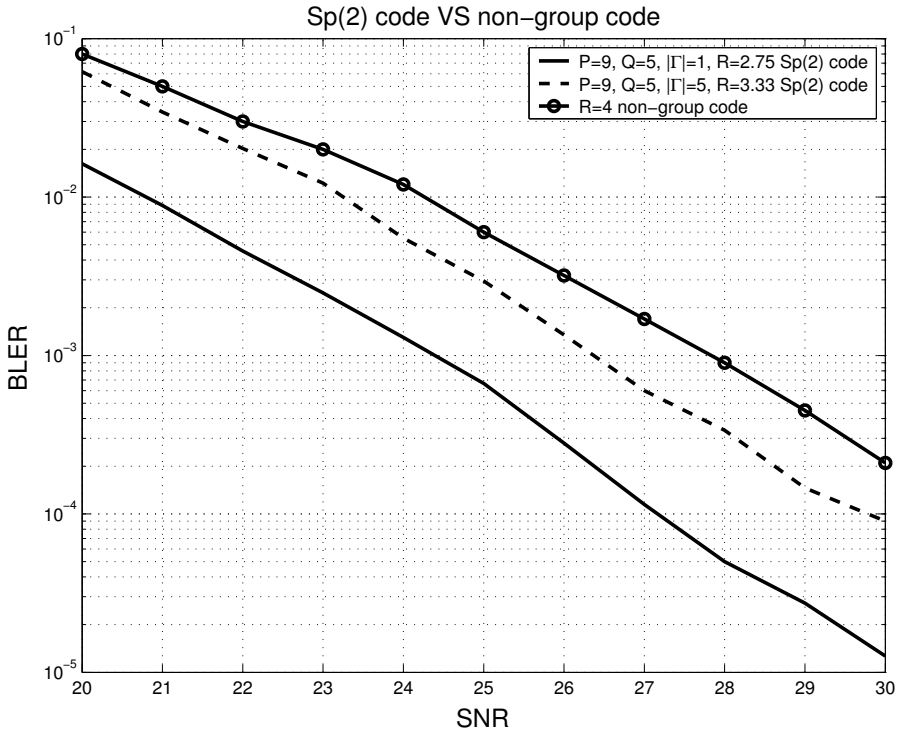


Figure 5.8: Comparison of $P = 9, Q = 5, \theta = 0.0377$ $Sp(2)$ code of $\Gamma = \{\frac{\pi}{4}\}, R = 2.7459$ and $\Gamma = \{\frac{\pi}{12}, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{5\pi}{12}\} + 0.016, R = 3.3264$ with the non-group code

such as sphere decoding. Simulation results show that they have better performance than the $2 \times 2$ and $4 \times 4$ complex orthogonal designs, a group-based diagonal code as well as differential Cayley codes at high SNR. Although they slightly underperform the $k_{1,1,-1}$ finite-group code and the carefully designed non-group code, they do not need the exhaustive search (of exponentially growing size) required for such codes and therefore are far superior in term of decoding complexity. Our work shows the promise of studying constellations inspired by group-theoretic considerations.

# 5.9 Appendices

## 5.9.1 Proof of Lemma 5.6

**Proof:** First assume that the determinant is zero and prove that $O_1 = \pm J\bar{O}_2$. Assume $\det(S_1 - S_2) = 0$. If $\det O_1 = 0$, from Lemma 5.5, $O_1 = \mathbf{0}_{22}$. Therefore

$$\det(S_1 - S_2) = \frac{1}{\sqrt{2}} \det \begin{bmatrix} 0 & O_2 \\ -\bar{O}_2 & 0 \end{bmatrix} = \det{}^2 O_2 = 0.$$

Thus, $O_2 = \mathbf{0}_{22}$ by Lemma 5.5. This indicates that $S_1 = S_2$, which contradicts $S_1 \neq S_2$. And the same is true for the case of $\det O_2 = \mathbf{0}_{22}$. Therefore, $\det O_1 \neq 0$ and $\det O_2 \neq 0$. From (5.12), $\det(S_1 - S_2)$ is always non-negative and equals 0 if and only if $|\alpha|^2(a + \frac{1}{a})^2 = 0$ and $|a\beta - \frac{\bar{\beta}}{a}|^2 = 0$. Since $(a + \frac{1}{a})^2 \geq 2$, the determinant equals zero if and only if $\alpha = 0$ and $a\beta = \frac{\bar{\beta}}{a}$, which can be written as $a^2\beta = \bar{\beta}$. By looking at the norm of each side of the equation, we get $a^2 = 1$. Since $a$ is positive, $a = 1$ and thus $\beta = \bar{\beta}$, which means that $\beta$ is real. Therefore, $\bar{O}_1 O_2^* = \begin{bmatrix} 0 & \beta \\ -\beta & 0 \end{bmatrix}$ with real $\beta$, which indicates that

$$\bar{O}_1 = \frac{\beta}{\det O_2} J O_2,$$

or equivalently,

$$O_1 = \frac{\beta}{\det O_2} J \bar{O}_2$$

since $O_2^{-1} = \frac{O_2^*}{\det O_2}$ by Lemma 5.5. Since $a = 1$ we have $\det O_1 = \det O_2$. Therefore, the following equations can be obtained.

$$\det O_2 = \det O_1 = \det(\frac{\beta}{\det O_2} J\bar{O}_2) = \left(\frac{\beta}{\det O_2}\right)^2 \overline{\det O_2} = \left(\frac{\beta}{\det O_2}\right)^2 \det O_2.$$

Therefore,

$$\frac{\beta}{\det O_2} = \pm 1.$$

Thus, $O_1 = \pm J\bar{O}_2$.

Now assume that $O_1 = \pm J\bar{O}_2$ and prove that $\det(S_1 - S_2) = 0$. First, assume that $O_1$ is invertible. If $O_1 = \pm J\bar{O}_2$, we have $O_1^{-1} = (\pm J\bar{O}_2)^{-1} = \pm\bar{O}_2^{-1}\bar{J}^{-1} = \mp\bar{O}_2^{-1}J$. From (5.11),

$$\begin{aligned} & \det(S_1 - S_2) \\ = & \frac{1}{\sqrt{2}} \det O_1 \det(\pm\overline{JO_2} \mp \bar{O}_2\bar{O}_2^{-1}JO_2) \\ = & \frac{1}{\sqrt{2}} \det O_1 \det(\pm\overline{JO_2} \mp JO_2) \\ = & 0. \end{aligned}$$

Secondly, assume that $O_1$ is not invertible, that is $\det O_1 = 0$. From Lemma 5.5, $O_1 = \mathbf{0}_{22}$. Therefore, from $O_1 = \pm J\bar{O}_2$, $O_2 = \mathbf{0}_{22}$. Thus, $S_1 - S_2 = \mathbf{0}_{44}$ and $\det(S_1 - S_2) = 0$.

Now what left to be proved is that $O_1 = \pm J\bar{O}_2$ is equivalent to (5.13). By (5.10), it is equivalent to

$$U_1V_1 - U_2V_2 = \pm J(\bar{U}_1V_1 - \bar{U}_2V_2),$$

and thus,

$$(U_1 \mp J\bar{U}_1)V_1 = (U_2 \mp J\bar{U}_2)V_2.$$

Using (5.9), the following series of equations can be obtained.

$$
\begin{aligned}
&\frac{1}{2}\left(\begin{bmatrix} e^{2\pi j\frac{k_1}{P}} & e^{2\pi j\frac{l_1}{P}} \\ -e^{-2\pi j\frac{l_1}{P}} & e^{-2\pi j\frac{k_1}{P}} \end{bmatrix} \mp \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} e^{-2\pi j\frac{k_1}{P}} & e^{-2\pi j\frac{l_1}{P}} \\ -e^{2\pi j\frac{l_1}{P}} & e^{2\pi j\frac{k_1}{P}} \end{bmatrix}\right)\begin{bmatrix} e^{j\left(2\pi\frac{m_1}{Q}+\theta\right)} & e^{j\left(2\pi\frac{n_1}{P}+\theta\right)} \\ -e^{-j\left(2\pi\frac{n_1}{Q}+\theta\right)} & e^{-j\left(2\pi\frac{m_1}{Q}+\theta\right)} \end{bmatrix} \\
=&\frac{1}{2}\left(\begin{bmatrix} e^{2\pi j\frac{k_2}{P}} & e^{2\pi j\frac{l_2}{P}} \\ -e^{-2\pi j\frac{l_2}{P}} & e^{-2\pi j\frac{k_2}{P}} \end{bmatrix} \mp \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} e^{-2\pi j\frac{k_2}{P}} & e^{-2\pi j\frac{l_2}{P}} \\ -e^{2\pi j\frac{l_2}{P}} & e^{2\pi j\frac{k_2}{P}} \end{bmatrix}\right)\begin{bmatrix} e^{j\left(2\pi\frac{m_2}{Q}+\theta\right)} & e^{j\left(2\pi\frac{n_2}{P}+\theta\right)} \\ -e^{-j\left(2\pi\frac{n_2}{Q}+\theta\right)} & e^{-j\left(2\pi\frac{m_2}{Q}+\theta\right)} \end{bmatrix} \\
\Leftrightarrow&\begin{bmatrix} e^{j\left[2\pi\left(\frac{k_1}{P}+\frac{m_1}{Q}\right)+\theta\right]} \pm e^{j\left[2\pi\left(\frac{k_1}{P}-\frac{n_1}{Q}\right)-\theta\right]} - e^{j\left[2\pi\left(\frac{l_1}{P}+\frac{m_1}{Q}\right)+\theta\right]} \pm e^{j\left[2\pi\left(\frac{l_1}{P}-\frac{n_1}{Q}\right)-\theta\right]} \\ \pm e^{-j\left[2\pi\left(\frac{k_1}{P}-\frac{m_1}{Q}\right)-\theta\right]} - e^{-j\left[2\pi\left(\frac{k_1}{P}+\frac{n_1}{Q}\right)+\theta\right]} - e^{-j\left[2\pi\left(\frac{l_1}{P}-\frac{m_1}{Q}\right)-\theta\right]} \mp e^{-j\left[2\pi\left(\frac{l_1}{P}+\frac{n_1}{Q}\right)+\theta\right]} \end{bmatrix} \\
&\qquad\begin{matrix} \mp e^{j\left[2\pi\left(\frac{k_1}{P}-\frac{m_1}{Q}\right)-\theta\right]} + e^{j\left[2\pi\left(\frac{k_1}{P}+\frac{n_1}{Q}\right)+\theta\right]} + e^{j\left[2\pi\left(\frac{l_1}{P}-\frac{m_1}{Q}\right)-\theta\right]} \pm e^{j\left[2\pi\left(\frac{l_1}{P}+\frac{n_1}{Q}\right)+\theta\right]} \\ e^{-j\left[2\pi\left(\frac{k_1}{P}+\frac{m_1}{Q}\right)+\theta\right]} \pm e^{-j\left[2\pi\left(\frac{k_1}{P}-\frac{n_1}{Q}\right)-\theta\right]} - e^{-j\left[2\pi\left(\frac{l_1}{P}+\frac{m_1}{Q}\right)+\theta\right]} \pm e^{-j\left[2\pi\left(\frac{l_1}{P}-\frac{n_1}{Q}\right)-\theta\right]} \end{matrix} \\
=&\begin{bmatrix} e^{j\left[2\pi\left(\frac{k_2}{P}+\frac{m_2}{Q}\right)+\theta\right]} \pm e^{j\left[2\pi\left(\frac{k_2}{P}-\frac{n_2}{Q}\right)-\theta\right]} - e^{j\left[2\pi\left(\frac{l_2}{P}+\frac{m_2}{Q}\right)+\theta\right]} \pm e^{j\left[2\pi\left(\frac{l_2}{P}-\frac{n_2}{Q}\right)-\theta\right]} \\ \pm e^{-j\left[2\pi\left(\frac{k_2}{P}-\frac{m_2}{Q}\right)-\theta\right]} - e^{-j\left[2\pi\left(\frac{k_2}{P}+\frac{n_2}{Q}\right)+\theta\right]} - e^{-j\left[2\pi\left(\frac{l_2}{P}-\frac{m_2}{Q}\right)-\theta\right]} \mp e^{-j\left[2\pi\left(\frac{l_2}{P}+\frac{n_2}{Q}\right)+\theta\right]} \end{bmatrix} \\
&\qquad\begin{matrix} \mp e^{j\left[2\pi\left(\frac{k_2}{P}-\frac{m_2}{Q}\right)-\theta\right]} + e^{j\left[2\pi\left(\frac{k_2}{P}+\frac{n_2}{Q}\right)+\theta\right]} + e^{j\left[2\pi\left(\frac{l_2}{P}-\frac{m_2}{Q}\right)-\theta\right]} \pm e^{j\left[2\pi\left(\frac{l_2}{P}+\frac{n_2}{Q}\right)+\theta\right]} \\ e^{-j\left[2\pi\left(\frac{k_2}{P}+\frac{m_2}{Q}\right)+\theta\right]} \pm e^{-j\left[2\pi\left(\frac{k_2}{P}-\frac{n_2}{Q}\right)-\theta\right]} \pm e^{-j\left[2\pi\left(\frac{l_2}{P}+\frac{m_2}{Q}\right)+\theta\right]} - e^{-j\left[2\pi\left(\frac{l_2}{P}-\frac{n_2}{Q}\right)-\theta\right]} \end{matrix} \\
\Leftrightarrow&\begin{cases} e^{2\theta j}\left(e^{2\pi j\left(\frac{k_1}{P}+\frac{m_1}{Q}\right)} - e^{2\pi j\left(\frac{k_2}{P}+\frac{m_2}{Q}\right)} \pm e^{2\pi j\left(\frac{l_1}{P}+\frac{m_1}{Q}\right)} \mp e^{2\pi j\left(\frac{l_2}{P}+\frac{m_2}{Q}\right)}\right) \\ \qquad = \mp e^{2\pi j\left(\frac{k_1}{P}-\frac{n_1}{Q}\right)} \pm e^{2\pi j\left(\frac{k_2}{P}-\frac{n_2}{Q}\right)} + e^{2\pi j\left(\frac{l_1}{P}-\frac{n_1}{Q}\right)} - e^{2\pi j\left(\frac{l_2}{P}-\frac{n_2}{Q}\right)} \\ e^{2\theta j}\left(e^{2\pi j\left(\frac{k_1}{P}+\frac{n_1}{Q}\right)} - e^{2\pi j\left(\frac{k_2}{P}+\frac{n_2}{Q}\right)} \pm e^{2\pi j\left(\frac{l_1}{P}+\frac{n_1}{Q}\right)} \mp e^{2\pi j\left(\frac{l_2}{P}+\frac{n_2}{Q}\right)}\right) \\ \qquad = \pm e^{2\pi j\left(\frac{k_1}{P}-\frac{m_1}{Q}\right)} \mp e^{2\pi j\left(\frac{k_2}{P}-\frac{m_2}{Q}\right)} - e^{2\pi j\left(\frac{l_1}{P}-\frac{m_1}{Q}\right)} + e^{2\pi j\left(\frac{l_2}{P}-\frac{m_2}{Q}\right)} \end{cases}.
\end{aligned}
$$

It is easy to see that the equation is equivalent to (5.13). □

### 5.9.2 Proof of Lemma 5.7

**Proof:** Assume $e^{j\theta_1} + e^{j\theta_2} + e^{j\theta_3} + e^{j\theta_4} = 0$, then the following series of equations are

true.

$$
\begin{aligned}
& e^{j\frac{\theta_1+\theta_2}{2}}\left(e^{j\frac{\theta_1-\theta_2}{2}}+e^{-j\frac{\theta_1-\theta_2}{2}}\right)+e^{j\frac{\theta_3+\theta_4}{2}}\left(e^{j\frac{\theta_3-\theta_4}{2}}+e^{-j\frac{\theta_3-\theta_4}{2}}\right)=0 \\
\Rightarrow\quad & 2\cos\left(\frac{\theta_1-\theta_2}{2}\right)e^{j\frac{\theta_1+\theta_2}{2}}=-2\cos\left(\frac{\theta_3-\theta_4}{2}\right)e^{j\frac{\theta_3+\theta_4}{2}} \\
\Rightarrow\quad & \begin{cases} \frac{\theta_1-\theta_2}{2}=\pm\frac{\theta_3-\theta_4}{2}+2k\pi \\ \frac{\theta_1+\theta_2}{2}=\frac{\theta_3+\theta_4}{2}+(2l+1)\pi \end{cases} \text{ or } \begin{cases} \frac{\theta_1-\theta_2}{2}=\pm\frac{\theta_3-\theta_4}{2}+(2k+1)\pi \\ \frac{\theta_1+\theta_2}{2}=\frac{\theta_3+\theta_4}{2}+2l\pi \end{cases},
\end{aligned}
$$

for some integers $k$ and $l$. Without loss of generality, only the first case is considered here. By adding the two equations,

$$
\begin{cases} \theta_1=\theta_3+(2k+2l+1)\pi \\ \theta_2=\theta_4+(-2k+2l+1)\pi \end{cases} \Rightarrow \begin{cases} e^{j\theta_1}+e^{j\theta_3}=0 \\ e^{j\theta_2}+e^{j\theta_4}=0 \end{cases}
$$

when plus sign is applied, or

$$
\begin{cases} \theta_1=\theta_4+(2k+2l+1)\pi \\ \theta_2=\theta_3+(-2k+2l+1)\pi \end{cases} \Rightarrow \begin{cases} e^{j\theta_1}+e^{j\theta_4}=0 \\ e^{j\theta_2}+e^{j\theta_3}=0 \end{cases}
$$

when minus sign is applied.

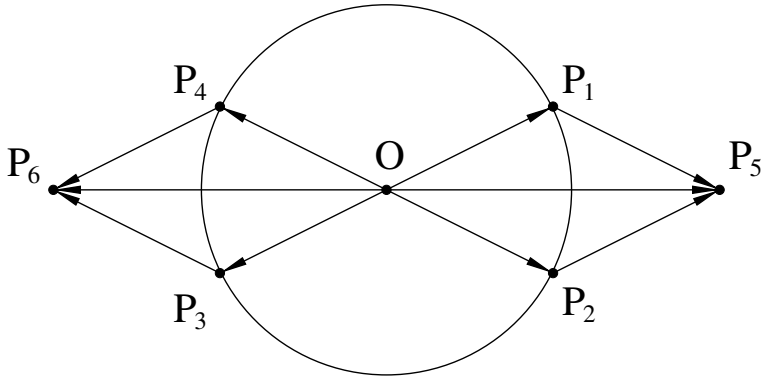


Figure 5.9: Figure for Lemma 5.7

This lemma can also be proved easily in a geometric way. As in Figure 5.9,

$P_1, P_2, P_3, P_4$ are the four points on the unit circle that add up to zero, and $O$ is the center of the unit circle, which is the origin. Without loss of generality, assume $P_2$ is the point that is closest to $P_1$. Since the four points add up to zero, $\overrightarrow{OP}_5$, which is the summation of $\overrightarrow{OP_1}$ and $\overrightarrow{OP}_2$, and $\overrightarrow{OP}_6$, which is the summation of $\overrightarrow{OP}_3$ and $\overrightarrow{OP}_4$, are on the same line with inverse directions and have the same length. Since $\overline{OP_1} = \overline{OP_3}$, $\overline{P_1P_5} = \overline{OP_2} = \overline{OP_4} = \overline{P_3P_6}$, $\triangle OP_1P_5 = \triangle OP_3P_6$. Therefore, $\angle P_1OP_5 = \angle P_3OP_6$. Thus, $\overrightarrow{OP}_1$ and $\overrightarrow{OP}_3$ are on the same line with inverse directions and have the same length, which means that $\overrightarrow{OP}_1 + \overrightarrow{OP}_3 = 0$. □

### 5.9.3 Proof of Lemma 5.8

**Proof:** Here, we only prove that $w^+, x^+, y^+, z^+$ as defined in (5.14) cannot be zero simultaneously. The proof of the other part ($w^-, x^-, y^-, z^-$ as defined in (5.15) cannot be zero simultaneously) is very similar to it. It can be proved by contradiction. Assume that $P$ and $Q$ are relatively prime and there exist integers $k_1, l_1, k_2, l_2$ in $[0, P-1]$ and $m_1, n_1, m_2, n_2$ in $[0, Q-1]$ such that $w^+ = x^+ = y^+ = z^+ = 0$. Since $x^+ = 0$, by Lemma 5.7,

$$\begin{cases} e^{2\pi j(\frac{k_1}{P} - \frac{n_1}{Q})} = e^{2\pi j(\frac{l_1}{P} - \frac{n_1}{Q})} \\ e^{2\pi j(\frac{k_2}{P} - \frac{n_2}{Q})} = e^{2\pi j(\frac{l_2}{P} - \frac{n_2}{Q})} \end{cases}, \quad \begin{cases} e^{2\pi j(\frac{k_1}{P} - \frac{n_1}{Q})} = e^{2\pi j(\frac{k_2}{P} - \frac{n_2}{Q})} \\ e^{2\pi j(\frac{l_1}{P} - \frac{n_1}{Q})} = e^{2\pi j(\frac{l_2}{P} - \frac{n_2}{Q})} \end{cases},$$

or

$$\begin{cases} e^{2\pi j(\frac{k_1}{P} - \frac{n_1}{Q})} = -e^{2\pi j(\frac{l_2}{P} - \frac{n_2}{Q})} \\ e^{2\pi j(\frac{k_2}{P} - \frac{n_2}{Q})} = -e^{2\pi j(\frac{l_1}{P} - \frac{n_1}{Q})} \end{cases}.$$

Without loss of generality and to simplify the proof, only the first case is discussed

here. From the first set of equations, there exist integers $i_1$ and $i_2$ such that

$$\begin{cases} 2\pi(\frac{k_1}{P} - \frac{n_1}{Q}) = 2\pi(\frac{l_1}{P} - \frac{n_1}{Q}) + 2\pi i_1 \\ 2\pi(\frac{k_2}{P} - \frac{n_2}{Q}) = 2\pi(\frac{l_2}{P} - \frac{n_2}{Q}) + 2\pi i_2 \end{cases},$$

and therefore

$$\begin{cases} \frac{k_1 - l_1}{P} = i_1 \\ \frac{k_2 - l_2}{P} = i_2 \end{cases}. \tag{5.42}$$

Since $k_1, l_1, l_1, l_2 \in [0, P)$, $k_1 - l_1, k_2 - l_2 \in (-P, P)$. Therefore, $i_1 = i_2 = 0$, from which $k_1 = l_1$ and $k_2 = l_2$ are obtained. Using this result and $w^+ = 0$, $2e^{2\pi j(\frac{k_1}{P} + \frac{m_1}{Q})} = 2e^{2\pi j(\frac{k_2}{P} + \frac{m_2}{Q})}$ can be obtained, from which $\frac{k_1}{P} + \frac{m_1}{Q} = \frac{k_2}{P} + \frac{m_2}{Q} + i$ is true for some integer $i$. The equation is equivalent to $\frac{k_1 - k_2}{P} = i - \frac{m_1 - m_2}{Q}$. Since $P$ and $Q$ are relatively prime, $P|(k_1 - k_2)$ and $Q|(m_1 - m_2)$. Since $k_1, k_2 \in [0, P)$ and $m_1, m_2 \in [0, Q)$, $k_1 - k_2 \in (-P, P)$ and $m_1 - m_2 \in (-Q, Q)$. Therefore, $k_1 - k_2 = 0$ and $m_1 - m_2 = 0$ which, combined with (5.42), means that $k_1 = k_2 = l_1 = l_2$ and $m_1 = m_2$. From $y^+ = 0$ and $e^{2\pi j \frac{k_1}{P}} \neq 0$, $2e^{2\pi j \frac{n_1}{Q}} = 2e^{2\pi j \frac{n_2}{Q}}$ can be achieved, which, similarly, leads to $\frac{n_1}{Q} = \frac{n_1}{Q} + i$ and therefore $n_1 - n_2 = iQ$, for some integer $i$. Since $n_1, n_2 \in [0, Q)$, $n_1 - n_2 \in (-Q, Q)$. Therefore, $i = 0$, that is $n_1 = n_2$. Therefore, $(k_1, l_1, m_1, n_1) = (k_2, l_2, m_2, n_2)$, which contradicts the condition of the lemma. □

### 5.9.4 Proof of Lemma 5.9

**Proof:** This lemma is porved by contradiction. We only prove that $\tilde{w}^+, \tilde{x}^+, \tilde{y}^+, \tilde{z}^+$ cannot be zeros simultaneously here since proving that $\tilde{w}^-, \tilde{x}^-, \tilde{y}^-, \tilde{z}^-$ cannot be zeros simultaneously is very similar.

Assume $\tilde{w}^+ = \tilde{x}^+ = \tilde{y}^+ = \tilde{z}^+ = 0$. From the definition of $\tilde{w}^+, \tilde{x}^+, \tilde{y}^+, \tilde{z}^+$ in (5.27),

$$e^{2\pi i\frac{m_i}{Q}}\left(\cos\gamma_i e^{2\pi i\frac{k_i}{P}} + \sin\gamma_i e^{2\pi i\frac{l_i}{P}}\right) = e^{2\pi i\frac{m_j}{Q}}\left(\cos\gamma_j e^{2\pi i\frac{k_j}{P}} + \sin\gamma_j e^{2\pi i\frac{l_j}{P}}\right) \quad (5.43)$$

$$e^{-2\pi i\frac{n_i}{Q}}\left(\sin\gamma_i e^{2\pi i\frac{k_i}{P}} - \cos\gamma_i e^{2\pi i\frac{l_i}{P}}\right) = e^{-2\pi i\frac{n_j}{Q}}\left(\sin\gamma_j e^{2\pi i\frac{k_j}{P}} - \cos\gamma_j e^{2\pi i\frac{l_j}{P}}\right) \quad (5.44)$$

$$e^{2\pi i\frac{n_i}{Q}}\left(\cos\gamma_i e^{2\pi i\frac{k_i}{P}} + \sin\gamma_i e^{2\pi i\frac{l_i}{P}}\right) = e^{2\pi i\frac{n_j}{Q}}\left(\cos\gamma_j e^{2\pi i\frac{k_j}{P}} + \sin\gamma_j e^{2\pi i\frac{l_j}{P}}\right) \quad (5.45)$$

$$e^{-2\pi i\frac{m_i}{Q}}\left(\sin\gamma_i e^{2\pi i\frac{k_i}{P}} - \cos\gamma_i e^{2\pi i\frac{l_i}{P}}\right) = e^{-2\pi i\frac{m_j}{Q}}\left(\sin\gamma_j e^{2\pi i\frac{k_j}{P}} - \cos\gamma_j e^{2\pi i\frac{l_j}{P}}\right) \quad (5.46)$$

The square of the norm of the left hand side of (5.43) equals

$$\begin{aligned}
&\left|\cos\gamma_i e^{2\pi i\frac{k_i}{P}} + \sin\gamma_i e^{2\pi i\frac{l_i}{P}}\right|^2 \\
=\ & (\cos\gamma_i\cos 2\pi\frac{k_i}{P} + \sin\gamma_i\cos 2\pi\frac{l_i}{P})^2 + (\cos\gamma_i\sin 2\pi\frac{k_i}{P} + \sin\gamma_i\sin 2\pi\frac{l_i}{P})^2 \\
=\ & \cos^2\gamma_i\cos^2 2\pi\frac{k_i}{P} + \sin^2\gamma_i\cos^2 2\pi\frac{l_i}{P} + 2\cos\gamma_i\cos 2\pi\frac{k_i}{P}\sin\gamma_i\cos 2\pi\frac{l_i}{P} + \\
& \cos^2\gamma_i\sin^2 2\pi\frac{k_i}{P} + \sin^2\gamma_i\sin^2 2\pi\frac{l_i}{P} + 2\cos\gamma_i\sin 2\pi\frac{k_i}{P}\sin\gamma_i\sin 2\pi\frac{l_i}{P} \\
=\ & \cos^2 2\pi\frac{k_i}{P} + \sin^2 2\pi\frac{k_i}{P} + 2\cos\gamma_i\sin\gamma_i(\cos 2\pi\frac{k_i}{P}\cos 2\pi\frac{l_i}{P} + \sin 2\pi\frac{k_i}{P}\sin 2\pi\frac{l_i}{P}) \\
=\ & 1 + \sin 2\gamma_i\cos 2\pi\frac{k_i - l_i}{P}.
\end{aligned}$$

Similarly, the square of the norm of the right hand side of (5.43) is equal to

$$1 + \sin 2\gamma_j\cos 2\pi\frac{k_j - l_j}{P}.$$

Comparing the norms of both side of (5.43),

$$1 + \sin 2\gamma_i\cos 2\pi\frac{k_i - l_i}{P} = 1 + \sin 2\gamma_j\cos 2\pi\frac{k_j - l_j}{P},$$

which is equivalent to

$$\sin 2\gamma_i\cos 2\pi\frac{k_i - l_i}{P} = \sin 2\gamma_j\cos 2\pi\frac{k_j - l_j}{P}. \quad (5.47)$$

Define $r = |k_i - l_i|$ and $s = |k_j - l_j|$. Since $k_i, l_i, k_j, l_j \in [0, P)$, $k_i - l_i, k_j - l_j \in (-P, P)$. Therefore, $r, s \in [0, P)$, $\cos 2\pi \frac{r}{P} = \cos 2\pi \frac{k_i - l_i}{P}$ and $\cos 2\pi \frac{s}{P} = \cos 2\pi \frac{k_j - l_j}{P}$. Thus,

$$\sin 2\gamma_i \cos 2\pi \frac{r}{P} = \sin 2\gamma_j \cos 2\pi \frac{s}{P}.$$

Since $\Gamma \subset (0, \frac{\pi}{0})$, $\sin\gamma \neq 0$ and $\cos\gamma \neq 0$ for any $\gamma \in \Gamma$. Therefore, when $\gamma_i \neq \gamma_j$, this contradicts (5.22). Therefore, when $\gamma_i \neq \gamma_j$ and (5.22) is satisfied, $\tilde{w}^+, \tilde{x}^+, \tilde{y}^+, \tilde{z}^+$ cannot be zero simultaneously.

Now look at the case of $\gamma_i = \gamma_j$. From (5.43),

$$\cos\gamma_i \left( e^{2\pi j \left(\frac{k_j}{P} + \frac{m_j}{Q}\right)} - e^{2\pi j \left(\frac{k_j}{P} + \frac{m_j}{Q}\right)} \right) = -\sin\gamma_i \left( e^{2\pi j \left(\frac{l_j}{P} + \frac{m_i}{Q}\right)} - e^{2\pi j \left(\frac{l_j}{P} + \frac{m_j}{Q}\right)} \right)$$

and

$$\begin{aligned} & 2j \cos\gamma_i \sin 2\pi \left( \frac{k_i - k_j}{2P} + \frac{m_i - m_j}{2Q} \right) e^{2\pi j \left(\frac{k_i + k_j}{2P} + \frac{m_i + m_j}{Q}\right)} \\ = \quad & -2 \sin\gamma_i \sin 2\pi \left( \frac{l_i - l_j}{2P} + \frac{m_i - m_j}{2Q} \right) e^{2\pi j \left(\frac{l_i + l_j}{2P} + \frac{m_i + m_j}{2Q}\right)}. \end{aligned}$$

Therefore,

$$\cos\gamma_i \sin 2\pi \left( \frac{k_i - k_j}{2P} + \frac{m_i - m_j}{2Q} \right) = \pm \sin\gamma_i \sin 2\pi \left( \frac{l_i - l_j}{2P} + \frac{m_i - m_j}{2Q} \right). \quad (5.48)$$

From (5.23),

$$\cos\gamma_i \sin 2\pi \left( \frac{k_i - k_j}{2P} + \frac{m_i - m_j}{2Q} \right) \neq \pm \sin\gamma_i \sin 2\pi \left( \frac{l_i - l_j}{2P} + \frac{m_i - m_j}{2Q} \right),$$

where $k_i - k_j, l_i - l_j \in (-P, P), m_i - m_j \in (-Q, Q)$, and $(k_i - k_j, m_i - m_j) \neq (0, 0)$. Therefore, for (5.48) to be true, $k_i - k_j = 0$ and $m_i - m_j = 0$. Thus, $\sin 2\pi(\frac{k_i - k_j}{2P} + \frac{m_i - m_j}{2Q}) = 0$. Since $\sin\gamma_i \neq 0$, $\sin 2\pi(\frac{l_i - l_j}{2P} + \frac{m_i - m_j}{2Q}) = \sin\pi\frac{l_i - l_j}{2P} = 0$.

Therefore, $l_i - l_j = 0$. So we get

$$\begin{cases} k_i = k_j \\ l_i = l_j \\ m_i = m_j \end{cases}.$$

Similarly, from (5.44), we have

$$\sin\gamma_i \left( e^{2\pi j\left(\frac{k_i}{P}-\frac{n_i}{Q}\right)} - e^{2\pi j\left(\frac{k_j}{P}-\frac{n_j}{Q}\right)} \right) = \cos\gamma_i \left( e^{2\pi j\left(\frac{l_i}{P}-\frac{n_i}{Q}\right)} - e^{2\pi j\left(\frac{l_j}{P}-\frac{n_j}{Q}\right)} \right)$$

and

$$\begin{aligned} & 2i\sin\gamma_i \sin 2\pi \left( \frac{k_i - k_j}{2P} - \frac{n_i - n_j}{2Q} \right) e^{2\pi j\left(\frac{k_i+k_j}{2P}-\frac{n_i+n_j}{Q}\right)} \\ = \;& 2j\cos\gamma_i \sin 2\pi \left( \frac{l_i - l_j}{2P} - \frac{n_i - n_j}{2Q} \right) e^{2\pi j\left(\frac{l_i+l_j}{2P}-\frac{n_i+n_j}{2Q}\right)}. \end{aligned}$$

Therefore,

$$\sin\gamma_i \sin 2\pi \left( \frac{k_i - k_j}{2P} - \frac{n_i - n_j}{2Q} \right) = \pm\cos\gamma_i \sin 2\pi \left( \frac{l_i - l_j}{2P} - \frac{n_i - n_j}{2Q} \right). \quad (5.49)$$

By a similar argument,

$$\begin{cases} k_i = k_j \\ l_i = l_j \\ n_i = n_j \end{cases}.$$

Therefore, $(k_i, l_i, m_i, n_i, \gamma_i) = (k_j, l_j, m_j, n_j, \gamma_j)$, and this contradicts the condition that they are different. □

# Chapter 6 Differential Unitary Space-Time Codes Based on $SU(3)$

## 6.1 Abstract

In this chapter, the special unitary Lie group $SU(3)$ is discussed. Based on the structure of matrices in $SU(3)$, two methods to design constellations of $3 \times 3$ unitary matrices are proposed. One of the methods gives codes that are subsets of $SU(3)$. The other codes are derived from the $SU(3)$ code by a simple modification, which are called AB codes. Simple formulas are derived by which diversity products of the codes can be calculated in a fast way. Necessary conditions for full-diversity of the codes are also proved. Our conjecture is that they are also sufficient conditions. Simulation results given in Section 6.6 show that the codes have better performances than the group-based codes [SHHS01] especially at high rates and are as good as the elaborately-designed non-group codes [SHHS01]. Another exceptional feature of AB codes is that they have a fast maximum-likelihood decoding algorithm based on complex sphere decoding.

The work in this chapter has been published in *the Proceeding of the Thirty-Seventh Asilomar Conference on Signals, Systems, and Computers (Asilomar'03)* [JH03d] and *the Proceeding of 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'04) [JH04d].* The journal paper is submitted to *IEEE Transactions on Signal Processing* [JH04e].

## 6.2 The Special Unitary Lie Group and Its Parameterization

**Definition 6.1 (Special unitary group).** *[Sim94] $SU(n)$ is the group of complex $n \times n$ matrices obeying $U^*U = UU^* = I_n$ and* $\det U = 1$.

From the definition, $SU(n)$ is the group of complex $n \times n$ unitary matrices with determinant 1. It is called the special unitary group. It is also known that $SU(n)$ is a compact, simple, simply-connected Lie group of dimension $n^2 - 1$ and rank $n - 1$. Since we are most interested in the case of rank 2, here the focus is on $SU(3)$, which has dimension 8. The following theorem on the parameterization of $SU(3)$ is proved.

**Theorem 6.1 (Parameterization of $SU(3)$).** *Any matrix $U$ belongs to $SU(3)$ if and only if it can be written as*

$$U = \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Phi \end{bmatrix} \begin{bmatrix} a & 0 & -\sqrt{1-|a|^2} \\ 0 & 1 & 0 \\ \sqrt{1-|a|^2} & 0 & \bar{a} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Psi \end{bmatrix}, \tag{6.1}$$

*where $\Phi, \Psi \in SU(2)$ and $a$ is a complex number with $|a| < 1$.*

**Proof:** See Section 6.8.1. □

From the proof of the theorem, $\bar{a}$ actually equals the determinant of the submatrix $\begin{bmatrix} u_{22} & u_{23} \\ u_{32} & u_{33} \end{bmatrix}$ of $U$. This theorem indicates that any matrix in $SU(3)$ can be written as a product of three $3 \times 3$ unitary matrices which are basically $SU(2)$ since they are actually reducible $3 \times 3$ unitary representations of $SU(2)$ by adding an identity block. Now let's look at the number of degrees of freedom in $U$. Since $\Phi, \Psi \in SU(2)$, there are 6 degrees of freedom in them. Together with the 2 degrees of freedom in the complex scalar $\alpha$, the dimension of $U$ is 8, which is exactly the same

as that of $SU(3)$. Based on (6.1), matrices in $SU(3)$ can be parameterized by entries of $\Phi, \Psi$ and $\alpha$, that is, any matrix in $SU(3)$ can be identified with a 3-tuple $(\Phi, \Psi, \alpha)$. From (6.1), it can also be seen that all the three matrices are block-diagonal with a unit block. The first and third matrices have the unit element at the $(1,1)$ entry and the second matrix has the unit element at the $(2,2)$ entry. To get a more symmetric parameterization method, the following corollary is proved.

**Corollary 6.1.** *Any matrix $U$ belongs to $SU(3)$ if and only if it can be written as*

$$U = \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Phi \end{bmatrix} \begin{bmatrix} \alpha e^{j\omega} & 0 & \sqrt{1-|\alpha|^2}e^{j\beta} \\ 0 & 1 & 0 \\ -\sqrt{1-|\alpha|^2}e^{-j\beta} & 0 & \alpha e^{-j\omega} \end{bmatrix} \begin{bmatrix} \Psi & \mathbf{0}_{21} \\ \mathbf{0}_{12} & 1 \end{bmatrix}, \quad (6.2)$$

*where $\Phi, \Psi \in SU(2)$, $\alpha \in [0,1]$. $\omega$ and $\beta$ are arbitrary angles.*

**Proof:** First, it is easy to prove that any matrix with the structure in (6.2) is in $SU(3)$ by checking the unitary and determinant conditions. What is left is to prove that any matrix in $SU(3)$ can be written as the formula in (6.2).

For any matrix $U \in SU(3)$, define $U' = U \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$. It is easy to check that $U'$ is also a matrix in $SU(3)$. Therefore, from Theorem 6.1, there exist matrices $\Phi', \Psi'' \in SU(2)$ and a complex scalar $a$, such that

$$U' = \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Phi' \end{bmatrix} \begin{bmatrix} a & 0 & -\sqrt{1-|a|^2} \\ 0 & 1 & 0 \\ \sqrt{1-|a|^2} & 0 & \bar{a} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Psi'' \end{bmatrix}.$$

Let $\Psi'' = \begin{bmatrix} \psi'_{11} & \psi'_{12} \\ -\bar{\psi}'_{12} & \bar{\psi}'_{11} \end{bmatrix}$, where $|\psi'_{11}|^2 + |\psi'_{12}|^2 = 1$. Note that

$$\begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Psi'' \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Psi' & \mathbf{0}_{21} \\ \mathbf{0}_{12} & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

where we have defined $\Psi' = \begin{bmatrix} \bar{\psi}'_{11} & -\bar{\psi}'_{12} \\ \psi'_{12} & \psi'_{11} \end{bmatrix}$ (it is easy to see that $\Psi' \in SU(2)$). Therefore,

$$\begin{aligned} & U' \\ = & \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Phi' \end{bmatrix} \begin{bmatrix} a & 0 & -\sqrt{1-|a|^2} \\ 0 & 1 & 0 \\ \sqrt{1-|a|^2} & 0 & \bar{a} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Psi' & \mathbf{0}_{21} \\ \mathbf{0}_{12} & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ = & \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Phi' \end{bmatrix} \begin{bmatrix} \sqrt{1-|a|^2} & 0 & a \\ 0 & 1 & 0 \\ -\bar{a} & 0 & \sqrt{1-|a|^2} \end{bmatrix} \begin{bmatrix} \Psi' & \mathbf{0}_{21} \\ \mathbf{0}_{12} & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{aligned}$$

and

$$U = \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Phi' \end{bmatrix} \begin{bmatrix} \sqrt{1-|a|^2} & 0 & a \\ 0 & 1 & 0 \\ -\bar{a} & 0 & \sqrt{1-|a|^2} \end{bmatrix} \begin{bmatrix} \Psi' & \mathbf{0}_{21} \\ \mathbf{0}_{12} & 1 \end{bmatrix}.$$

It is easy to check that

$$\begin{aligned} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{j\omega} & 0 \\ 0 & 0 & e^{-j\omega} \end{bmatrix} \begin{bmatrix} \sqrt{1-|a|^2} & 0 & a \\ 0 & 1 & 0 \\ -\bar{a} & 0 & \sqrt{1-|a|^2} \end{bmatrix} \begin{bmatrix} e^{j\omega} & 0 & 0 \\ 0 & e^{-j\omega} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ = & \begin{bmatrix} \sqrt{1-|a|^2}e^{j\omega} & 0 & a \\ 0 & 1 & 0 \\ -\bar{a} & 0 & \sqrt{1-|a|^2}e^{-j\omega} \end{bmatrix} \end{aligned}$$

for any angle $\omega$. Define

$$\Phi = \Phi' \begin{bmatrix} e^{-j\omega} & 0 \\ 0 & e^{j\omega} \end{bmatrix} \quad \text{and} \quad \Psi = \begin{bmatrix} e^{-j\omega} & 0 \\ 0 & e^{j\omega} \end{bmatrix} \Psi'.$$

It is easy to see that $\Phi, \Psi \in SU(2)$. (6.2) is obtained by letting $\alpha = \sqrt{1-|a|^2}$ and $\beta = \angle a$, □

The parameter $\omega$ does not add any degrees of freedom as can be seen in the proof of the corollary. However, as will be seen later that it is important to our code design. From formula (6.2), any matrix in $SU(3)$ can be written as a product of three basically $SU(2)$ matrices. The first matrix is a three-dimensional representation of $SU(2)$ with an identity block at the $(1,1)$ entry. The second matrix is a a three-dimensional representation of $SU(2)$ with an identity block at the $(2,2)$ entry and the third matrix is a a three-dimensional representation of $SU(2)$ with an identity block at the $(3,3)$ entry.

## 6.3 $SU(3)$ Code Design

To get finite constellations of unitary matrices from the infinite Lie group $SU(3)$, the parameters, $\Phi, \Psi, \alpha, \beta, \omega$, need to be sampled appropriately. We first sample $\Phi$ and $\Psi$. As discussed in Chapter 2, Alamouti's orthogonal design

$$\left\{ \frac{1}{\sqrt{|x|^2+|y|^2}} \begin{bmatrix} x & y \\ -\bar{x} & \bar{y} \end{bmatrix} \middle| \, x, y \in \mathbb{C} \right\}$$

is a faithful representation of the group $SU(2)$. To get a discrete set, $x$ and $y$ must belong to discrete sets. As is well known, the PSK signal is a very good and simple

modulation. Therefore, $\Phi$ and $\Psi$ are chosen as follows.[1]

$$\Phi = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{2\pi j \frac{p}{P}} & e^{2\pi j \frac{q}{Q}} \\ -e^{-2\pi j \frac{q}{Q}} & e^{-2\pi j \frac{p}{P}} \end{bmatrix} \quad \text{and} \quad \Psi = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{2\pi j \frac{r}{R}} & e^{2\pi j \frac{s}{S}} \\ -e^{-2\pi j \frac{s}{S}} & e^{-2\pi j \frac{r}{R}} \end{bmatrix},$$

where $P, Q, R, S$ are positive integers.

Since the group is not fpf, we cannot use all the eight degrees of freedom in it to get fully-diverse codes. To simplify the structure, let $\alpha = 1$, by which the middle matrix in (6.2) is a diagonal matrix. Also, fully-diverse subsets are desired. Therefore the angle $\omega$ should depend on $\Phi$ and $\Psi$, or equivalently, it is a function of $p, q, r, s$. To see this, assume that $\omega$ is independent of $\Phi$. Then, the determinant of $U_1(\Phi_1, \Psi, 1, \omega) - U_1(\Phi_2, \Psi, 1, \omega)$ is zero since $\Phi$ has a unit block at its $(1,1)$ entry. The same is true for $\Psi$. Therefore, let $\omega = 2\pi \left( \frac{p}{P} - \frac{q}{Q} + \frac{r}{R} + \frac{s}{S} \right)$. The reason for this will be illuminated later. Define

$$\theta_{p,q} = 2\pi \left( \frac{p}{P} - \frac{q}{Q} \right) \text{ and } \xi_{r,s} = 2\pi \left( \frac{r}{R} + \frac{s}{S} \right). \tag{6.3}$$

(6.2) becomes

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} e^{2\pi j \frac{p}{P}} & \frac{1}{\sqrt{2}} e^{2\pi j \frac{q}{Q}} \\ 0 & -\frac{1}{\sqrt{2}} e^{-2\pi j \frac{q}{Q}} & \frac{1}{\sqrt{2}} e^{-2\pi j \frac{p}{P}} \end{bmatrix} \text{diag} \left\{ e^{\theta_{p,q}}, 1, e^{-\theta_{p,q}} \right\} \text{diag} \left\{ e^{\xi_{r,s}}, 1, e^{-\xi_{r,s}} \right\} \begin{bmatrix} \frac{1}{\sqrt{2}} e^{2\pi j \frac{r}{R}} & \frac{1}{\sqrt{2}} e^{2\pi j \frac{s}{S}} & 0 \\ -\frac{1}{\sqrt{2}} e^{-2\pi j \frac{s}{S}} & \frac{1}{\sqrt{2}} e^{-2\pi j \frac{r}{R}} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Define

$$A^{(1)}_{(p,q)} = \begin{bmatrix} e^{j\theta_{p,q}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} e^{2\pi j \frac{p}{P}} & \frac{1}{\sqrt{2}} e^{2\pi j \frac{q}{Q}} e^{-j\theta_{p,q}} \\ 0 & -\frac{1}{\sqrt{2}} e^{-2\pi j \frac{q}{Q}} & \frac{1}{\sqrt{2}} e^{-2\pi j \frac{p}{P}} e^{-j\theta_{p,q}} \end{bmatrix}, \tag{6.4}$$

[1] PSK symbols have been analyzed in [SWWX04], where it is shown that having a full parameterization of $SU(2)$, that is, parameterizing $x$ and $y$ fully (both the norms and the arguments) gives about 1-2 dB improvement but with a much more complicated decoding. Here, to make our main idea clear, $x$ and $y$ are chosen as simple $PSK$ signals.

which is the product of the first two matrices in the above formula, and

$$B^{(1)}_{(r,s)} = \begin{bmatrix} \frac{1}{\sqrt{2}} e^{2\pi j \frac{r}{R}} e^{j\xi_{r,s}} & \frac{1}{\sqrt{2}} e^{2\pi j \frac{s}{S}} e^{j\xi_{r,s}} & 0 \\ -\frac{1}{\sqrt{2}} e^{-2\pi j \frac{s}{S}} & \frac{1}{\sqrt{2}} e^{-2\pi j \frac{r}{R}} & 0 \\ 0 & 0 & e^{-j\xi_{r,s}} \end{bmatrix}, \tag{6.5}$$

which is the product of the last two matrices. The following codes are obtained.

$$\mathcal{C}^{(1)}_{(P,Q,R,S)} = \left\{ A^{(1)}_{(p,q)} B^{(1)}_{(r,s)} \,|p \in [0,P), q \in [0,Q), r \in [0,R), s \in [0,S) \right\} \tag{6.6}$$

The set is a subset of $SU(3)$. We call it the $SU(3)$ code. There are all together $PQRS$ elements in the code (6.6). Since the channel is used in blocks of three transmissions, the rate of the code is

$$R = \frac{1}{3} \log_2(PQRS). \tag{6.7}$$

**Theorem 6.2 (Calculation of the diversity product).** *Define*

$$\begin{aligned} x &= e^{2\pi j \left( \frac{p_1 - p_2}{2P} - \frac{q_1 - q_2}{2Q} \right)} \cos 2\pi \left( \frac{p_1 - p_2}{2P} + \frac{q_1 - q_2}{2Q} \right) \\ w &= e^{2\pi j \left( -\frac{r_1 - r_2}{2R} - \frac{s_1 - s_2}{2S} \right)} \cos 2\pi \left( \frac{r_1 - r_2}{2R} - \frac{s_1 - s_2}{2S} \right) \end{aligned}. \tag{6.8}$$

*For* $U_1(p_1, q_1, r_1, s_1), U_2(p_2, q_2, r_2, s_2) \in \mathcal{C}^{(1)}$,

$$\det(U_1 - U_2) = 2j \Im[(1 - \bar{\Theta} x)(1 - \Theta w)], \tag{6.9}$$

*where* $\Theta = e^{-2\pi j \left( \frac{p_1 - p_2}{P} - \frac{q_1 - q_2}{Q} + \frac{r_1 - r_2}{R} + \frac{s_1 - s_2}{S} \right)}$. $\Im[c]$ *indicates the imaginary part of the complex scalar* $c$.

**Proof:** See Section 6.8.2. □

The diversity product of the code is therefore

$$\zeta_{\mathcal{C}^{(1)}}(P,Q,R,S) = \frac{1}{2} \min_{\substack{\delta_p \in (-P,P),\, \delta_q \in (-Q,Q),\\ \delta_r \in (-R,R),\, \delta_s \in (-S,S)\\ (\delta_p,\delta_q,\delta_r,\delta_s) \neq (0,0,0,0)}} |2\Im[(1-\bar{\Theta}x)(1-\Theta w)]|^{1/3}. \tag{6.10}$$

In general, to obtain the diversity product, determinants of $\frac{1}{2}L(L-1) = \frac{1}{2}(L^2-L)$ difference matrices, which is quadratic in $L$, need to be calculated. However, from Theorem 6.2, $x$, $w$, $\Theta$, $\Theta_1$, and $\Theta_2$ only depend on the differences $\delta_p = p_1 - p_2$, $\delta_q = q_1 - q_2$, $\delta_r = r_1 - r_2$, and $\delta_s = s_1 - s_2$ instead of $p_1, p_2, q_1, q_2, r_1, r_2, s_1, s_2$. That is, $\det(U_1 - U_2)$ can be written as $\Delta(\delta_p, \delta_q, \delta_r, \delta_s)$, a function of $\delta_p, \delta_q, \delta_r, \delta_s$. Since $\delta_p, \delta_q, \delta_r, \delta_s$ can take on $2P-1, 2Q-1, 2R-1, 2S-1$ possible values, respectively, to get the diversity products, calculation of determinants of $(2P-1)(2Q-1)(2R-1)(2S-1)-1 < 16PQRS = 16L$ difference matrices is required, which is linear in $L$. Actually, instead of $16L$, less than $8L$ calculations is enough because of the symmetry in (6.10). Note that

$$\begin{aligned} & |\Delta(\delta_p,\delta_q,\delta_r,\delta_s)| \\ = \; & |2\Im[(1-\bar{\Theta}x)(1-\Theta w)]|^{1/3} \\ = \; & |2\Im[(1-\Theta\bar{x})(1-\overline{\Theta w})]|^{1/3} \\ = \; & |\Delta(-\delta_p,-\delta_q,-\delta_r,-\delta_s)|. \end{aligned}$$

Therefore, only half of the determinants are needed to be calculated. The computational complexity is greatly reduced especially for codes of high rates, that is, when $PQRS$ is large.

From the symmetries of $\delta_p, \delta_q, \delta_r, \delta_s$ in (6.10), it is easy to prove that

$$\zeta_{\mathcal{C}^{(1)}}(P,Q,R,S) = \zeta_{\mathcal{C}^{(1)}}(Q,P,R,S) = \zeta_{\mathcal{C}^{(1)}}(P,Q,S,R) = \zeta_{\mathcal{C}^{(1)}}(R,S,P,Q),$$

| $(P, Q, R, S)$ | Rate | Diversity Product |
|---|---|---|
| $(5, 7, 3, 1)$ | 2.2381 | 0.2133 |
| $(5, 7, 9, 1)$ | 2.7664 | 0.1714 |
| $(7, 9, 11, 1)$ | 3.1456 | 0.0961 |
| $(3, 7, 5, 11)$ | 3.3912 | 0.0803 |
| $(5, 9, 7, 11)$ | 3.9195 | 0.0510 |
| $(7, 11, 9, 13)$ | 4.3791 | 0.0316 |

Table 6.1: Diversity products of $SU(3)$ codes

| Group | Rate | Diversity Product |
|---|---|---|
| cyclic group with $u = (1, 17, 26)$ | 1.99 | 0.3301 |
| cyclic group with $u = (1, 11, 27)$ | 2 | 0.2765 |
| $G_{21,4}$ | 1.99 | 0.3851 |
| $G_{171,64}$ | 3 | 0.1353 |
| $G_{1365,16}$ | 4 | 0.0361 |
| $G_{10815,46}$ | 5 | 0.0131 |
| non-group code with $L_A = 65, u = (1, 30, 114)$ | 4.01 | 0.0933 |

Table 6.2: Diversity products of some group-based codes and a non-group code

which indicates that switching the positions of $P$ and $Q$, $R$ and $S$, or $(P, Q)$ and $(R, S)$ does not affect the diversity product. But generally, $\zeta_{\mathcal{C}^{(1)}}(P, Q, R, S) \neq \zeta_{\mathcal{C}^{(1)}}(P, R, Q, S)$.

Diversity products of some of the $SU(3)$ codes are given in Table 6.1. Diversity products of some of the group-based codes and non-group codes in [SHHS01] are also given in Table 6.2 for comparison. It can be seen from the tables that diversity products of $SU(3)$ codes are about the same as those of the group-based codes at low rates, but when the rates are high, diversity products of $SU(3)$ codes are much greater than those of the group-based codes at about the same rates. However, diversity products of the $SU(3)$ code at rate 3.9195, which is 0.0510, is smaller than that of the non-group code at rate 4.01, which is 0.0933. But simulated performance shows that the code performs as well as the non-group code, which will be seen in Section 6.6.

**Theorem 6.3 (Necessary conditions for full diversity).** *Necessary conditions*

*for code $\mathcal{C}^{(1)}$ to be fully diverse are that any two of the integers $(P, Q, R, S)$ are relatively prime and none of them are even.*

**Proof:** First, we prove that $\gcd(P, Q) = 1$ is a necessary condition for the set $\{A^{(1)}_{(p,q)}\}$ to be fully diverse and thus a necessary condition for full diversity of the code. Let

$$A^{(1)}_{(p_1,q_1)} = \begin{bmatrix} e^{j\theta_1} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} e^{2\pi j \frac{p_1}{P}} & \frac{1}{\sqrt{2}} e^{2\pi j \frac{q_1}{Q}} e^{-j\theta_1} \\ 0 & -\frac{1}{\sqrt{2}} e^{-2\pi j \frac{q_1}{Q}} & \frac{1}{\sqrt{2}} e^{-2\pi j \frac{p_1}{P}} e^{-j\theta_1} \end{bmatrix}$$

and

$$A^{(1)}_{(p_2,q_2)} = \begin{bmatrix} e^{j\theta_2} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} e^{2\pi j \frac{p_2}{P}} & \frac{1}{\sqrt{2}} e^{2\pi j \frac{q_2}{Q}} e^{-j\theta_2} \\ 0 & -\frac{1}{\sqrt{2}} e^{-2\pi j \frac{q_2}{Q}} & \frac{1}{\sqrt{2}} e^{-2\pi j \frac{p_2}{P}} e^{-j\theta_2} \end{bmatrix},$$

where $\theta_1 = 2\pi(\frac{p_1}{P} - \frac{q_1}{Q})$ and $\theta_2 = 2\pi(-\frac{p_2}{P} - \frac{q_2}{Q})$. Therefore,

$$\det\left(A^{(1)}_{(p_1,q_1)} - A^{(1)}_{(p_2,q_2)}\right) = \left(e^{j\theta_1} - e^{j\theta_2}\right) X$$

for some $X$. If $\gcd(P, Q) = G > 1$, let $0 \leq p_1 = p_2 + P/G < P$ and $0 \leq q_1 = q_2 + Q/G < Q$. Then

$$e^{j\theta_1} - e^{j\theta_2} = e^{2\pi j\left(\frac{p_2}{P} - \frac{q_2}{Q}\right)} - e^{2\pi j\left(\frac{p_2}{P} - \frac{q_2}{Q}\right)} = 0.$$

Therefore, $\gcd(P, Q) = 1$ is a necessary condition for the set $\{A^{(1)}_{(p,q)}\}$ to be fully diverse. By a similar argument, $\gcd(R, S) = 1$ is also a necessary condition.

Now assume $\gcd(P, R) = G > 1$. Let $q_1 = q_2, s_1 = s_2$, $0 \leq p_1 = p_2 + P/G$, and $0 \leq r_1 = r_2 + R/G$. From (6.8), $x = e^{j\frac{\pi}{G}} \cos\frac{\pi}{G}$, $w = e^{-j\frac{\pi}{G}} \cos\frac{\pi}{G}$, and $\Theta = e^{j\frac{4\pi}{G}}$.

Therefore, for the two matrices $U_1(p_1, q_1, r_1, s_1)$ and $U_2(p_2, q_2, r_2, s_2)$ in $\mathcal{C}^{(1)}$,

$$\begin{aligned}
& \det\left(U_1(p_1, q_1, r_1, s_1) - U_2(p_2, q_2, r_2, s_2)\right) \\
= \;& 2j(\mathcal{I}m(wx) - \mathcal{I}m\bar{\Theta}x - \mathcal{I}m\Theta w) \\
= \;& 2j\left(-\cos 2\pi\frac{1}{2G}\sin 2\pi\frac{5}{2G} - \cos 2\pi\frac{1}{2G}\sin 2\pi\left(-\frac{5}{2G}\right)\right) \\
= \;& 0.
\end{aligned}$$

So, $\mathcal{C}^{(1)}$ is not fully-diverse, which means $\gcd(P, R) = 1$ is a necessary condition. From the symmetries of $P$ and $Q$, $R$ and $S$, $\gcd(P, S) = \gcd(Q, R) = \gcd(Q, S) = 1$ are also necessary. Therefore, any two of the four integers $P, Q, R, S$ are relatively prime is necessary for the code $\mathcal{C}^{(1)}$ to be fully-diverse.

Now assume that $P$ is even. Let's look at the two matrices $U_1$ and $U_2$ with $(q_1, r_1, s_1) = (q_2, r_2, s_2)$ and $p_1 - p_2 = P/2$. (Since $P$ is even, this is achievable.) Therefore,

$$\begin{aligned}
& \det\left(U_1(p_1, q_1, r_1, s_1) - U_2(p_2, q_2, r_2, s_2)\right) \\
= \;& 2j\mathcal{I}\left(\cos 2\pi\frac{p_1 - p_2}{2P}e^{2\pi j\frac{p_1-p_2}{2P}} - \cos 2\pi\frac{p_1 - p_2}{2P}e^{2\pi j\frac{p_1-p_2}{2P}}e^{2\pi j\frac{p_1-p_2}{P}} - e^{2\pi j(-\frac{p_1-p_2}{P})}\right) \\
= \;& 2j\left(\cos\frac{\pi}{2}\sin\frac{\pi}{2} - \cos\frac{\pi}{2}\sin\frac{3\pi}{2} - \sin(-\pi)\right) \\
= \;& 0,
\end{aligned}$$

which means that $\mathcal{C}^{(1)}$ is not fully-diverse. By similar argument, if $Q$, $R$, or $S$ is even, $\mathcal{C}^{(1)}$ is not fully diverse. □

We are not able to give sufficient conditions for full diversity of the $SU(3)$ codes. Here is our conjecture.

**Conjecture 6.1 (Sufficient conditions for full diversity).** *The conditions, that any two of the integers $(P, Q, R, S)$ are relatively prime and none of them are even,*

*are sufficient for code* $\mathcal{C}^{(1)}_{(P,Q,R,S)}$ *to be fully diverse.*

## 6.4 AB Code Design

Note from (6.4) and (6.5) that the $e^{-j\theta_{p,q}}$ in the last column of $A^{(1)}_{p,q}$ and the $e^{j\xi}$ in the first row of $B^{(1)}_{p,q}$ are used to make the matrices have determinant 1. However, in differential unitary space-time code design, the signal matrices are only needed to be unitary. Therefore, the structure of the code can be further simplified by abandoning the restriction that both matrices have unit determinant. Define

$$A^{(2)}_{(p,q)} = \begin{bmatrix} e^{j\theta'_{p,q}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}e^{2\pi j\frac{p}{P}} & \frac{1}{\sqrt{2}}e^{2\pi j\frac{q}{Q}} \\ 0 & -\frac{1}{\sqrt{2}}e^{-2\pi j\frac{q}{Q}} & \frac{1}{\sqrt{2}}e^{-2\pi j\frac{p}{P}} \end{bmatrix} \quad B^{(2)}_{(r,s)} = \begin{bmatrix} \frac{1}{\sqrt{2}}e^{2\pi j\frac{r}{R}} & \frac{1}{\sqrt{2}}e^{2\pi j\frac{s}{S}} & 0 \\ -\frac{1}{\sqrt{2}}e^{-2\pi j\frac{s}{S}} & \frac{1}{\sqrt{2}}e^{-2\pi j\frac{r}{R}} & 0 \\ 0 & 0 & e^{-j\xi'_{r,s}} \end{bmatrix} \tag{6.11}$$

and[2]

$$\theta'_{p,q} = 2\pi\left(\pm\frac{p}{P} \pm \frac{q}{Q}\right) \qquad \xi'_{r,s} = 2\pi\left(\pm\frac{r}{R} \pm \frac{s}{S}\right). \tag{6.12}$$

The following codes with a simpler structure are obtained.

$$\mathcal{C}^{(2)}_{(P,Q,R,S)} = \left\{A^{(2)}_{(p,q)}B^{(2)}_{(r,s)} \,|p \in [0,P), q \in [0,Q), r \in [0,R), s \in [0,S)\right\}. \tag{6.13}$$

They are not subsets of the Lie group $SU(3)$ any more since the determinant of the matrices is now $e^{j(\theta'-\xi')}$ which is not 1 in general. However, the matrices in the codes are still unitary. Since any matrix in the code is a product of two unitary matrices (they are not representations of $SU(2)$ anymore because their determinants are no longer 1), we call it AB code. Simulations show that they have the same and sometimes slightly better diversity products than the codes in (6.6), which is

[2]There are actually 16 possibilities in (6.12). Different codes are obtained by different choices of signs. Two of them are used in this chapter.

not surprising since we now get rid of the constraint of unit determinant. In the next section, it will be seen that the handy structure of AB codes results in a fast maximum-likelihood decoding algorithm. The code has the same rate as the code in (6.6). It is easy to see that any matrix $U$ in the two codes can be identified by the 4-tuple $(p, q, r, s)$.

**Theorem 6.4 (Calculation of the diversity product).** *For any two matrices* $U_1(p_1, q_1, r_1, s_1)$ *and* $U_2(p_2, q_2, r_2, s_2)$ *in the code* $\mathcal{C}^{(2)}$,

$$\det(U_1 - U_2) = 2e^{j\theta'_1} e^{-j\xi'_2} \bar{\Theta}_1 \bar{\Theta}_2 \Im[(\Theta_1 - \bar{\Theta}_1 w)(\bar{\Theta}_2 - \Theta_2 x)], \tag{6.14}$$

*where* $\Theta_1 = e^{2\pi j\left(\pm\frac{p_1-p_2}{2P} \pm \frac{q_1-q_2}{2Q}\right)}$ *and* $\Theta_2 = e^{2\pi j\left(\pm\frac{r_1-r_2}{2R} \pm \frac{s_1-s_2}{2S}\right)}$.

**Proof:** See Section 6.8.3. □

Therefore, diversity product of the AB code is

$$\zeta_{\mathcal{C}^{(2)}}(P, Q, R, S) = \frac{1}{2} \min_{\substack{\delta_p \in (-P, P), \delta_q \in (-Q, Q), \\ \delta_r \in (-R, R), \delta_s \in (-S, S) \\ (\delta_p, \delta_q, \delta_r, \delta_s) \neq (0,0,0,0)}} |2\Im[(\Theta_1 - \bar{\Theta}_1 w)(\bar{\Theta}_2 - \Theta_2 x)]|^{1/3}. \tag{6.15}$$

Similar to the argument in the previous section, less than $8L$ calculations of the determinants of difference matrices are enough to obtain the diversity product. AB codes also have the symmetry that $\zeta_{\mathcal{C}^{(2)}}(P, Q, R, S) = \zeta_{\mathcal{C}^{(2)}}(Q, P, R, S) = \zeta_{\mathcal{C}^{(2)}}(P, Q, S, R) = \zeta_{\mathcal{C}^{(2)}}(R, S, P, Q)$. But generally, $\zeta_{\mathcal{C}^{(1)}}(P, Q, R, S) \neq \zeta_{\mathcal{C}^{(1)}}(P, R, Q, S)$.

As mentioned before, for AB codes, the choices for the angles $\theta$ and $\xi$ are not unique. Based on (6.12), there are actually 16 possible choices. Two of them are used here,

$$\theta' = 2\pi\left(-\frac{p}{P} + \frac{q}{Q}\right), \xi' = 2\pi\left(-\frac{r}{R} - \frac{s}{S}\right)$$

and

$$\theta' = 2\pi\left(\frac{p}{P} - \frac{q}{Q}\right), \xi' = 2\pi\left(-\frac{r}{R} - \frac{s}{S}\right),$$

| $(P,Q,R,S)$ | Rate | Type | Diversity Product |
|---|---|---|---|
| $(1,3,4,5)$ | 1.9690 | I | 0.2977 |
| $(4,5,3,7)$ | 2.9045 | I | 0.1413 |
| $(3,7,5,11)$ | 3.3912 | II | 0.0899 |
| $(4,7,5,11)$ | 3.5296 | I | 0.0731 |
| $(5,9,7,11)$ | 3.9195 | I | 0.0510 |
| $(5,8,9,11)$ | 3.9838 | II | 0.0611 |
| $(9,10,11,13)$ | 4.5506 | II | 0.0336 |
| $(11,13,14,15)$ | 4.9580 | II | 0.0276 |

Table 6.3: Diversity products of AB codes

which we call type I AB code and type II AB code, respectively.

Diversity products of some of the two types of AB codes are given in Table 6.3. By comparing with Table 6.2, it can be seen that AB codes have about the same diversity products as those of group-based codes at low rates, but when the rates are high, diversity products of AB codes are much greater than those of group-based codes at about the same rates. However, diversity product of the AB code at rate 3.9838, which is 0.0661, is smaller than that of the non-group code at rate 4.01, which is 0.0933. However, simulated performances in Section 6.6 show that the code performs as well as the non-group code.

**Theorem 6.5.** *The set* $\left\{A^{(2)}_{(p,q)}, p \in [0,P), q \in [0,Q)\right\}$ *is fully diverse if and only if* $\gcd(P,Q) = 1$. *The set* $\left\{B^{(2)}_{(r,s)}, r \in [0,R), s \in [0,S)\right\}$ *is fully diverse if and only if* $\gcd(R,S) = 1$.

**Proof:** We first prove that the set $\left\{A^{(2)}_{(p,q)}, p \in [0,P), q \in [0,Q)\right\}$ is fully diverse if and only if $P$ and $Q$ are relatively prime. For any two different matrices $A^{(2)}_{(p_1,q_1)}$ and $A^{(2)}_{(p_2,q_2)}$ in the set, denote

$$A^{(2)}_{(p_1,q_1)} = \begin{bmatrix} e^{j\theta_1} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}e^{2\pi j\frac{p_1}{P}} & \frac{1}{\sqrt{2}}e^{2\pi j\frac{q_1}{Q}} \\ 0 & -\frac{1}{\sqrt{2}}e^{-2\pi j\frac{q_1}{Q}} & \frac{1}{\sqrt{2}}e^{-2\pi j\frac{p_1}{P}} \end{bmatrix} \quad A^{(2)}_{(p_2,q_2)} = \begin{bmatrix} e^{j\theta_2} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}e^{2\pi j\frac{p_2}{P}} & \frac{1}{\sqrt{2}}e^{2\pi j\frac{q_2}{Q}} \\ 0 & -\frac{1}{\sqrt{2}}e^{-2\pi j\frac{q_2}{Q}} & \frac{1}{\sqrt{2}}e^{-2\pi j\frac{p_2}{P}} \end{bmatrix},$$

where $\theta_1 = 2\pi(\pm\frac{p_1}{P} \pm \frac{q_1}{Q})$ and $\theta_2 = 2\pi(\pm\frac{p_2}{P} \pm \frac{q_2}{Q})$. Therefore,

$$\begin{aligned}
& \det\left(A^{(2)}_{(p_1,q_1)} - A^{(2)}_{(p_2,q_2)}\right) \\
= & \left(e^{j\theta_1} - e^{j\theta_2}\right) \det \begin{bmatrix} \frac{1}{\sqrt{2}}e^{2\pi j\frac{p_1}{P}} - \frac{1}{\sqrt{2}}e^{2\pi j\frac{p_2}{P}} & \frac{1}{\sqrt{2}}e^{2\pi j\frac{q_1}{Q}} - \frac{1}{\sqrt{2}}e^{2\pi j\frac{q_2}{Q}} \\ -\frac{1}{\sqrt{2}}e^{-2\pi j\frac{q_1}{Q}} + \frac{1}{\sqrt{2}}e^{-2\pi j\frac{q_2}{Q}} & \frac{1}{\sqrt{2}}e^{-2\pi j\frac{p_1}{P}} - \frac{1}{\sqrt{2}}e^{-2\pi j\frac{p_2}{P}} \end{bmatrix} \\
= & \frac{1}{2}\left(e^{j\theta_1} - e^{j\theta_2}\right)\left(\left|e^{2\pi j\frac{p_1}{P}} - e^{2\pi j\frac{p_2}{P}}\right|^2 + \left|e^{2\pi j\frac{q_1}{Q}} - e^{2\pi j\frac{q_2}{Q}}\right|^2\right).
\end{aligned}$$

The second factor equals zero if and only if $e^{2\pi j\frac{p_1}{P}} = e^{2\pi j\frac{p_2}{P}}$ and $e^{2\pi j\frac{q_1}{Q}} = e^{2\pi j\frac{q_2}{Q}}$. Since $p_1, p_2 \in [0, P)$ and $q_1, q_2 \in [0, Q)$, this is equivalent to $(p_1, q_1) = (p_2, q_2)$, which cannot be true since the two matrices are different. Therefore, the determinant equals zero if and only if $e^{j\theta_1} = e^{j\theta_2}$.

Now assume that $\gcd(P, Q) = G > 1$, that is, $P$ and $Q$ are not relatively prime. When $p_1 - p_2 = \frac{P}{G}$ and $q_1 - q_2 = -\frac{Q}{G}$ (since $G$ divides both $P$ and $Q$, this is achievable,)

$$\theta_1 - \theta_2 = 2\pi\left(\pm\frac{p_1 - p_2}{P} \pm \frac{q_1 - q_2}{Q}\right) = 2\pi\left(\pm\frac{1}{G} \pm \left(-\frac{1}{G}\right)\right) = 0,$$

which means that $e^{j\theta_1} = e^{j\theta_2}$. Therefore, the set $\left\{A^{(2)}_{(p,q)}, p \in [0, P), q \in [0, Q)\right\}$ is not fully diverse.

Now assume that $\gcd(P, Q) = 1$. If $e^{j\theta_1} = e^{j\theta_2}$, $\theta_1 - \theta_2 = 2k\pi$ for some integer $k$, which means that $\pm\frac{p_1-p_2}{P} \pm \frac{q_1-q_2}{Q} = k$. Therefore, $\frac{p_1-p_2}{P} = \frac{\pm kQ \mp (q_1-q_2)}{Q}$. Since $\gcd(P, Q) = 1$, $P|(p_1 - p_2)$. However, because $p_1 - p_2 \in (-P + 1, P - 1)$, the only possibility is that $p_1 - p_2 = 0$. Therefore, $\frac{q_1-q_2}{Q} = \pm k$. From $q_1 - q_2 \in (-Q + 1, Q - 1)$, $q_1 - q_2 = 0$ and $k = 0$. So, $(p_1, q_1) = (p_2, q_2)$ is obtained, and this is a contradiction since the two matrices are different. Therefore, $e^{j\theta_1} \neq e^{j\theta_2}$. So $\det\left(A^{(2)}_{(p_1,q_1)} - A^{(2)}_{(p_2,q_2)}\right) \neq 0$. Therefore, $\gcd(P, Q) = 1$ is a sufficient condition for the set to be fully diverse.

By similar argument, we can prove that the set $\left\{B^{(2)}_{(r,s)}, r \in [0, R), s \in [0, S)\right\}$ is

fully-diverse if and only if $R$ and $S$ are relatively prime. □

**Theorem 6.6 (Necessary conditions for fully diversity).**

1. *Necessary conditions for full diversity of the type* I *AB code are that any two of* $P, Q, R, S$ *are relatively prime and at most one of the four integers* $P, Q, R, S$ *is even.*

2. *Necessary conditions for the full diversity of the type* II *AB code are* $\gcd(P, Q) = \gcd(R, S) = 1$ *and at most one of the four integers* $P, Q, R, S$ *is even.*

**Proof:** Since for code $\mathcal{C}^{(2)}$ to be fully diverse, $\left\{A^{(2)}_{(p,q)}\right\}$ and $\left\{B^{(2)}_{(r,s)}\right\}$ must be fully-diverse. Therefore, from Theorem 6.5, $\gcd(P, Q) = \gcd(R, S) = 1$ are necessary conditions for $\mathcal{C}^{(2)}$ to be fully diverse.

Assume that both $P$ and $R$ are even. For any two matrices $U_1(p_1, q_1, r_1, s_1)$ and $U_2(p_2, q_2, r_2, s_2)$ in $\mathcal{C}^{(2)}$, choose $q_1 = q_2$, $s_1 = s_2$, $p_1 - p_2 = \frac{P}{2}$, and $r_1 - r_2 = \frac{R}{2}$. Since both $P$ and $R$ are even, this is achievable and the two matrices are different. Therefore, from the proof of Theorem 6.2,

$$|\det(U_1(p_1, q_1, r_1, s_1) - U_2(p_2, q_2, r_2, s_2)| = 2|\Im(\Theta_1\bar{\Theta}_2 - \Theta_1\Theta_2 x - \bar{\Theta}_1\bar{\Theta}_2 w + \bar{\Theta}_1\Theta_2 wx)|,$$

where

$$\begin{aligned}
x &= e^{2\pi j\left(\frac{p_1-p_2}{2P} - \frac{q_1-q_2}{2Q}\right)} \cos 2\pi \left(\frac{p_1 - p_2}{2P} + \frac{q_1 - q_2}{2Q}\right) = e^{j\frac{\pi}{2}} \cos\frac{\pi}{2} = 0, \\
w &= e^{2\pi j\left(-\frac{r_1-r_2}{2R} - \frac{s_1-s_2}{2S}\right)} \cos 2\pi \left(\frac{r_1 - r_2}{2R} - \frac{s_1 - s_2}{2S}\right) = e^{-j\frac{\pi}{2}} \cos\frac{\pi}{2} = 0, \\
\Theta_1 &= e^{2\pi j\left(\pm\frac{p_1-p_2}{2P} \pm \frac{q_1-q_2}{2Q}\right)} = e^{\pm j\frac{\pi}{2}}, \quad \text{and} \quad \Theta_2 = e^{2\pi j\left(\pm\frac{r_1-r_2}{2R} \pm \frac{s_1-s_2}{2S}\right)} = e^{\pm j\frac{\pi}{2}}.
\end{aligned}$$

Therefore,

$$|\det(U_1(p_1, q_1, r_1, s_1) - U_2(p_2, q_2, r_2, s_2)| = 2|\Im(\Theta_1\bar{\Theta}_2)| = 0,$$

which indicates that the code is not fully diverse. By a similar argument, it can be proved that when any two of the integers $P, Q, R, S$ are even, code $\mathcal{C}^{(2)}$ is not fully diverse. Therefore, necessary conditions for code $\mathcal{C}^{(2)}$ to be fully-diverse are $\gcd(P, Q) = \gcd(R, S) = 1$ and among the four integers $P, Q, R, S$, at most one is even.

What is left is to prove that for type I AB code to be fully diverse, $\gcd(P, R) = \gcd(P, R) = \gcd(Q, R) = \gcd(Q, S) = 1$ are necessary. This is proved by contradiction. Assume that $\gcd(P, Q) = G > 1$. Let $p_1 = p_2 + G$, $q_1 = q + 2$, $r_1 = r_2 + G$, and $s_1 = s_2$. Therefore, $\Theta = \Theta_2 = e^{-j\frac{2\pi}{G}}$ and $x = \bar{w} = e^{j\frac{\pi}{G}} \cos\frac{\pi}{G}$. It is easy to check by formula (6.15) that $\det(U_1 - U_2) = 0$. Therefore, $\gcd(P, Q) = 1$ is necessary. The proofs are similar for other pairs. □

Note that, for type II AB code, any two of $(P, Q, R, S)$ being relatively prime is not a necessary condition. By computer simulation, $\zeta_{\mathcal{C}^{(2)}}(3, 5, 3, 5) = 0.1706 > 0$ and $\zeta_{\mathcal{C}^{(2)}}(7, 9, 7, 9) = 0.0219 > 0$, which indicates that $\gcd(P, R) = 1$, $\gcd(P, S) = 1$, $\gcd(Q, R) = 1$, or $\gcd(Q, S) = 1$ are not necessary for the code to be fully-diverse. We are not able to give sufficient conditions for the full diversity of the AB codes. Our conjectures are that the necessary conditions are also sufficient.

**Conjecture 6.2 (Sufficient conditions for full diversity).** *The necessary conditions given in Theorem 6.6 are also sufficient conditions for the two types of AB codes to be fully diverse.*

## 6.5 A Fast Decoding Algorithm for AB Codes

From (6.13), it can be seen that any matrix in code $\mathcal{C}^{(2)}_{(P,Q,R,S)}$ is a product of two basically $U(2)$ matrices $A^{(2)}_{(p,q)}$ and $B^{(2)}_{(r,s)}$.[3] It is easy to see from formula (6.11) that the two matrices have orthogonal design structure. This handy property can be used

[3]Although, $A^{(2)}_{(p,q)}$ and $B^{(2)}_{(r,s)}$ are not in $U(2)$, they are 3-dimensional representations of two $U(2)$ matrices by a reducible homomorphism.

to get linear-algebraic decoding, which means that the receiver can be made to form a system of linear equations in the unknowns.

The ML decoder for differential USTM given in (2.12) is equivalent to

$$\arg\min_{p,q,r,s} \left\| X_\tau - A^{(2)}_{(p,q)} B^{(2)}_{(r,s)} X_{\tau-1} \right\|_F^2 = \arg\min_{p,q,r,s} \left\| A^{(2)*}_{(p,q)X_\tau} - B^{(2)}_{(r,s)} X_{\tau-1} \right\|_F^2 .$$

Therefore, the decoding formula for code $\mathcal{C}^{(2)}_{(P,Q,R,S)}$ can be written as

$$\begin{aligned} &\arg\max_{p,q,r,s} \left\| \begin{bmatrix} e^{-j\theta_{p,q}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} e^{-2\pi j \frac{p}{P}} & -\frac{1}{\sqrt{2}} e^{2\pi j \frac{q}{Q}} \\ 0 & \frac{1}{\sqrt{2}} e^{-2\pi j \frac{q}{Q}} & \frac{1}{\sqrt{2}} e^{2\pi j \frac{p}{P}} \end{bmatrix} \begin{bmatrix} x_{\tau,11} & \cdots & x_{\tau,1N} \\ x_{\tau,21} & \cdots & x_{\tau,2N} \\ x_{\tau,31} & \cdots & x_{\tau,3N} \end{bmatrix} \right. \\ &\qquad \left. - \begin{bmatrix} \frac{1}{\sqrt{2}} e^{2\pi j \frac{r}{R}} & \frac{1}{\sqrt{2}} e^{2\pi j \frac{s}{S}} & 0 \\ -\frac{1}{\sqrt{2}} e^{-2\pi j \frac{s}{S}} & \frac{1}{\sqrt{2}} e^{-2\pi j \frac{r}{R}} & 0 \\ 0 & 0 & e^{-j\xi_{r,s}} \end{bmatrix} \begin{bmatrix} x_{\tau-1,11} & \cdots & x_{\tau-1,1N} \\ \tau-1,21 & \cdots & x_{\tau-1,2N} \\ x_{\tau-1,31} & \cdots & x_{\tau-1,3N} \end{bmatrix} \right\|_F^2 \\ = \quad &\arg\max_{p,q,r,s} \left\| \begin{bmatrix} x_{\tau,11} & 0 & 0 & 0 & -\frac{x_{\tau-1,11}}{\sqrt{2}} & -\frac{x_{\tau-1,21}}{\sqrt{2}} \\ 0 & \frac{\bar{x}_{\tau,21}}{\sqrt{2}} & -\frac{\bar{x}_{\tau,31}}{\sqrt{2}} & 0 & -\frac{\bar{x}_{\tau-1,21}}{\sqrt{2}} & \frac{\bar{x}_{\tau-1,11}}{\sqrt{2}} \\ 0 & \frac{x_{\tau,31}}{\sqrt{2}} & \frac{x_{\tau,21}}{\sqrt{2}} & -x_{\tau-1,31} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{\tau,1N} & 0 & 0 & 0 & -\frac{x_{\tau-1,1N}}{\sqrt{2}} & -\frac{x_{\tau-1,2N}}{\sqrt{2}} \\ 0 & \frac{\bar{x}_{\tau,2N}}{\sqrt{2}} & -\frac{\bar{x}_{\tau,3N}}{\sqrt{2}} & 0 & -\frac{\bar{x}_{\tau-1,2N}}{\sqrt{2}} & \frac{\bar{x}_{\tau-1,1N}}{\sqrt{2}} \\ 0 & \frac{x_{\tau,3N}}{\sqrt{2}} & \frac{x_{\tau,2N}}{\sqrt{2}} & -x_{\tau-1,3N} & 0 & 0 \end{bmatrix} \begin{bmatrix} e^{-j\theta'_{p,q}} \\ e^{2\pi j \frac{p}{P}} \\ e^{-2\pi j \frac{q}{Q}} \\ e^{-j\xi'_{r,s}} \\ e^{2\pi j \frac{r}{R}} \\ e^{2\pi j \frac{s}{S}} \end{bmatrix} \right\|_F^2 , \end{aligned}$$

where $x_{t,ij}$ indicates the $(i,j)$-th entry of the $M \times N$ matrix $X_t$ for $t = \tau, \tau-1$. The equality is obtained since the matrices $A^{(2)}_{(p,q)}$ and $B^{(2)}_{(r,s)}$ have orthogonal structure. It is easy to see that the formula inside the norm is linear in the PSK unknown signals. Therefore, sphere decoding for complex channels proposed in [HtB03] can be used with slight modification. The only difference here is that the unknowns $e^{-j\theta'_{p,q}}$ and $e^{-j\xi'_{r,s}}$ are not independent unknown PSK signals but are determined by $e^{2\pi j \frac{p}{P}}, e^{-2\pi j \frac{q}{Q}}$ and

$e^{2\pi j\frac{r}{R}}, e^{2\pi j\frac{s}{S}}$. Therefore, in the sphere decoding, instead of searching over the intervals for the unknowns $e^{-j\theta'_{p,q}}$ and $e^{-j\xi'_{r,s}}$, their values are calculated by values of $p, q$ and $r, s$ respectively based on formulas in (6.12) depending on the choices of the angles $\theta'_{p,q}$ and $\xi'_{r,s}$ used in the code. Since sphere decoding has an average complexity that is cubic in rate and dimension of the system, and at the same time achieves the exact ML results, a fast decoding algorithm for AB codes is found.

In digital communication, choice of the searching radius is crucial to the speed of the algorithm. If the initial radius is chosen to be very large, then actually most of the points are being searched, by which not too much improvement on computational complexity can be gained over exhaustive search. On the other hand, if the radius is chosen to be too small, then there may be no point in the sphere being searched. It is better to start with a small value then increase it gradually. In [DAML00], the authors proposed to choose the packing radius or the estimated packing radius to be the initial searching radius. Here, another initialization for the searching radius based on the noise level as in [HV02] and [JH03b] is used. The noise of the system is given in (2.11). Since $W_\tau, W_{\tau-1}$ and the transmitted unitary matrix, $V_{z_\tau}$, are independent, it is easy to prove that the noise matrix has mean zero and variance $2NI_3$. Each component of the $3 \times N$-dimensional noise vector has mean zero and variance 2. Therefore the random variable $v = \|W'_\tau\|_F^2$ has Gamma distribution with mean $3N$. The searching radius $\sqrt{C}$ is initialized in such a way that the probability that the correct signal is in the sphere is 0.9, that is, $\mathrm{P}\left(\|v\|_F < \sqrt{C}\right) = 0.9$. If no point is found in the sphere, then the searching radius is raised such that the probability is increased to 0.99 and so on. Using this algorithm, the probability that a point can be found during the first search is high. For more details of sphere decoding and sphere decoding for complex channels, please refer to [DAML00] and [HtB03].

Although $SU(3)$ codes also have the structure of products of two unitary matrices $A^{(1)}_{(p,q)}$ and $B^{(1)}_{(r,s)}$, since the two unitary matrices do not have the orthogonal design

structure, we cannot find a way to simplify the decoding. Therefore, for $SU(3)$ codes, exhaustive search is used to obtain the ML results.

## 6.6 Simulation Results

In this section, examples of both $SU(3)$ codes and the two types of AB codes are shown and also the simulated performance of the codes at different rates. The number of transmit antennas is three. The fading coefficient from each transmit antenna to each receive antenna is modeled independently as a complex Gaussian variable with zero-mean and unit-variance and keeps constant for $2M = 6$ channel uses. At each channel use, zero-mean, unit-variance complex Gaussian noise is added to each receive antenna. The block error rate (BLER), which corresponds to errors in decoding the $3 \times 3$ transmitted matrices, is demonstrated as the error event of interest. The comparison of the proposed codes with some of the group-based codes and the non-group code in [SHHS01] is also shown.

### 6.6.1 AB Code vs. Group-Based Codes at $R \approx 2$

The first example is the $P = 1, Q = 3, R = 4, S = 5$ AB code with $\theta_{p,q} = 2\pi\left(-\frac{p}{P} + \frac{q}{Q}\right)$ and $\xi_{r,s} = 2\pi\left(-\frac{r}{R} - \frac{s}{S}\right)$. In brief, we call it the $(1, 3, 4, 5)$ type I AB code. From (6.7), the rate of the code is 1.9690. From Table 6.3, diversity product of the code is 0.2977. Its BLER is compared with the $G_{21,4}$ group code at rate $R = 1.99$ with diversity product 0.3851 and also the best cyclic group code at rate 1.99, whose diversity product is 0.3301, with $u = (1, 17, 26)$. The number of receive antennas is one. The performance curves are shown in Figure 6.1. The solid line indicates the BLER of the type I AB code. The solid line with circles indicates the BLER of the $G_{21,4}$ code and the dashed line indicates the BLER of the cyclic group code. It can be seen from the plot that the performance of the three codes is close to each other. The AB code is a little

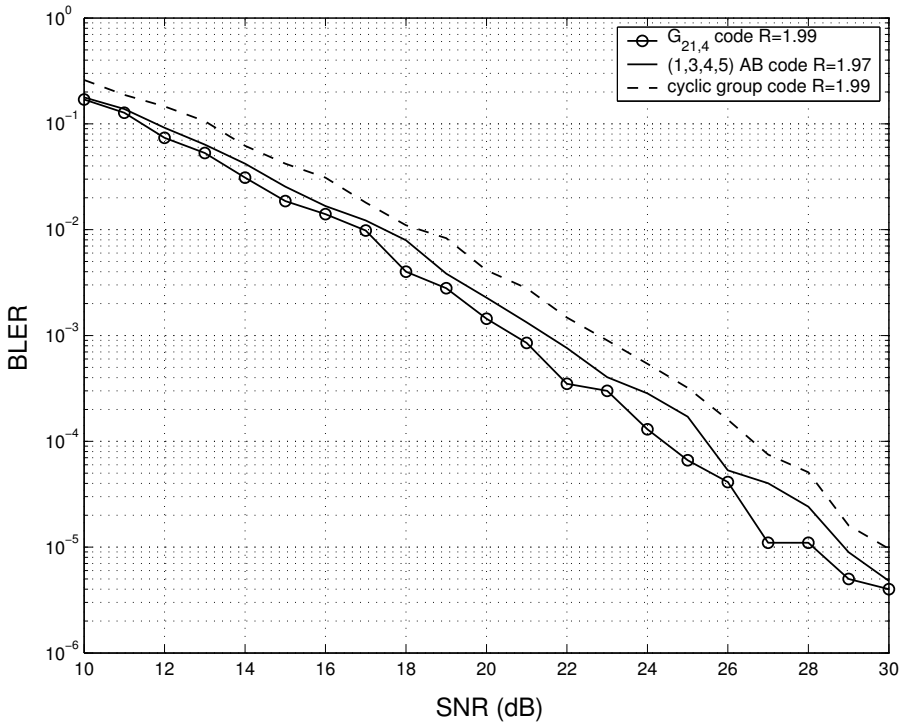


Figure 6.1: Comparison of the rate 1.9690, $(1, 3, 4, 5)$ type I AB code with the rate 1.99 $G_{21,4}$ code and the best rate 1.99 cyclic group code

(0.5dB-1dB) worse than the $G_{21,4}$ code and better (0.5dB-1dB) than the cyclic group code. Notice that the decoding of both group-based codes needs exhaustive search but the AB code has a fast decoding method. Therefore, at rate approximately 2, the AB code is as good as the group-based codes with far superior decoding complexity.

### 6.6.2 $SU(3)$ Codes and AB Codes vs. Group-Based Codes at $R \approx 3$

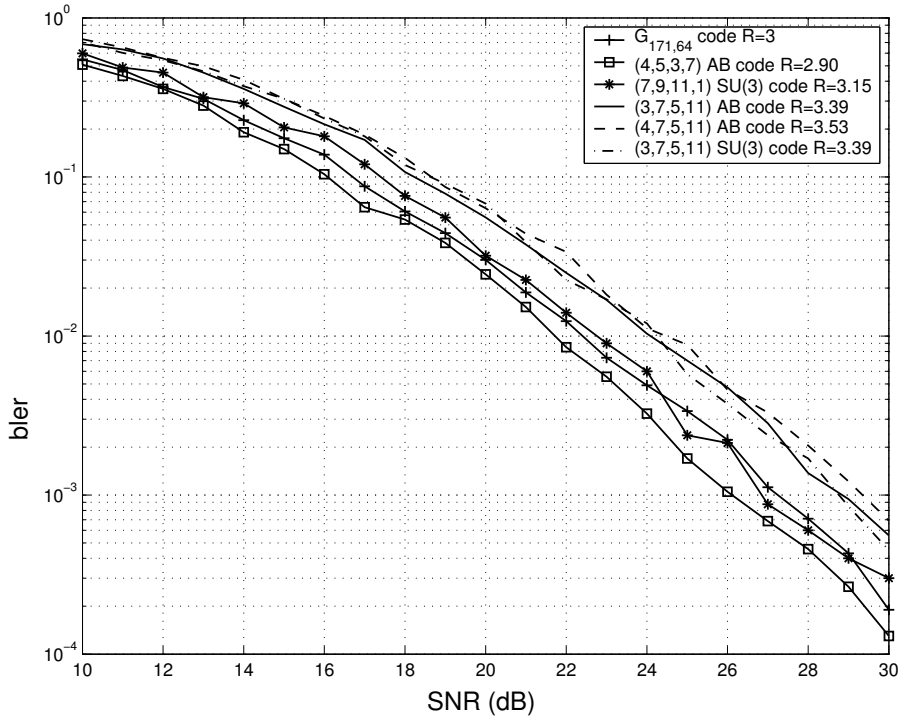


Figure 6.2: Comparison of the **1**) rate $2.9045, (4, 5, 3, 7)$ type I AB code, **2**) rate $3.15, (7, 9, 11, 1)$, $SU(3)$ code, **3**) rate $3.3912, (3, 7, 5, 11)$ type II AB code, **4**) rate $3.5296, (4, 7, 5, 11)$ type I AB code, and **5**) rate 3.3912, $(3, 7, 5, 11)$, $SU(3)$ code with **6**) the rate $3, G_{171,64}$ code

In this subsection, two sets of codes are compared. The first set includes the $(4, 5, 3, 7)$ type I AB code with rate $R = 2.9045$, the $G_{171,64}$ group-based code at rate 3, and the $SU(3)$ with $(P, Q, R, S) = (7, 9, 11, 1)$ and rate 3.1456. The number of receive antennas is one. The simulated BLERs are shown in Figure 6.2. The line

with squares indicates the bler of the type I AB code at rate 2.9045. The line with plus signs indicates the BLER of the $G_{171,64}$ code and the line with stars shows the BLER of the $SU(3)$ code at rate 3.1456. It can be seen from the plot that the rate 2.9045 AB code is about 1dB better than the $G_{171,64}$ code whose rate is 3. The $SU(3)$ code has about the same performance as the group-based code with a rate 0.1456 higher.

The second set of codes includes the $(3, 7, 5, 11)$ type II AB code at rate $R = 3.3912$, the $(4, 7, 5, 11)$ type I AB code with rate $R = 3.5296$, and the $(3, 7, 5, 11)$ $SU(3)$ code with rate $R = 3.3912$. The number of receive antennas is one. The simulated BLERs are also shown in Figure 6.2. The solid and dashed lines show the BLER of the rate 3.3912 and rate 3.5296 AB code, respectively. BLER of the $SU(3)$ code is shown by the dash-dotted line. The three codes have very close performance. Compared with the performance of the $G_{171,64}$ code, which is shown by the line with circles, the three codes, the two AB codes and the $SU(3)$ codes with rates 0.3912,0.5296, and 0.3912 higher, perform about 1.5dB worse than that of the group-based code. Note that the AB codes can be decoded much faster than the $G_{171,64}$ code and the $SU(3)$ codes.

### 6.6.3 $SU(3)$ Codes and AB Codes vs. Group-Based Codes and the Non-Group Code at $R \approx 4$

The comparison of the $(5, 8, 9, 11)$ type II AB code at rate 3.9838 and the $(9, 10, 11, 13)$ type II AB code at rate 4.5506, the $(5, 9, 7, 11)$ $SU(3)$ code at rate 3.9195, and the $(7, 11, 9, 13)$ $SU(3)$ code at rate 4.3791, with the rate 4 group-based $G_{1365,16}$ code is given in Figure 6.3. As can be seen in Figure 6.3, the line with circles indicates the BLER of the $G_{1365,16}$ code. The line with plus signs and the solid line show the BLER of the rate 3.9838 and 4.5506 AB code, respectively. The dashed and the dash-dotted

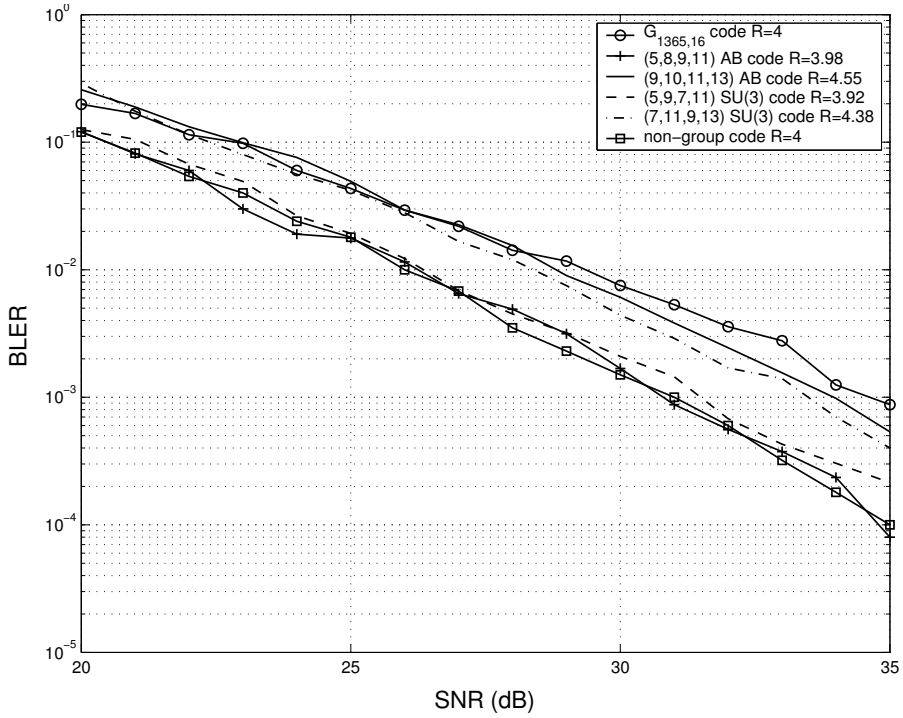


Figure 6.3: Comparison of the **1**) rate 3.9838, $(5, 8, 9, 11)$ type II AB code, **2**) rate 4.5506, $(9, 10, 11, 13)$ type II AB code, **3**) rate 3.9195, $(5, 9, 7, 11)$, $SU(3)$ code, and **4**) rate 4.3791, $(7, 11, 9, 13)$, $SU(3)$ code with the **5**) rate 4 $G_{1365,16}$ code and **6**) rate 4 non-group code

line show the BLER of the rate 3.9195 and 4.3791 $SU(3)$ code, respectively. The number of receive antennas is one. It can been seen from the plot that at about the same rate, the $(5, 8, 9, 11)$ type II AB code and the $(5, 9, 7, 11)$ $SU(3)$ code perform a lot better than the $G_{1365,16}$ code. For example, at the BLER of $10^{-3}$, the AB code has an advantage of about 4dB and the $SU(3)$ code has an advantage of about 3.5dB. Also, at rate 0.3791 higher, the $(7, 11, 9, 13)$, $SU(3)$ code is more than 1dB better than the $G_{1365,16}$ code does at high SNRs. The BLER of the $(9, 10, 11, 13)$ type II AB code is slightly lower than that of the $G_{1365,16}$ code even with a rate 0.5506 higher. The performance of the non-group code is also shown, which is indicated by the line with squares. It can be seen from the plot that the $(5, 8, 9, 11)$ type II AB code and the $SU(3)$ code at rates 3.9838 and 3.9195 are as good as the non-group code given in [SHHS01] at rate 4 according to BLER, although diversity product of the non-group code is much higher than those of the AB and $SU(3)$ codes from Tables 6.1-6.3. The reason might be that although in the AB and $SU(3)$ codes, the minimum of the determinants of the differences of two matrices is much smaller than that of the non-group code, the overall distribution of elements in the AB and $SU(3)$ codes are as good as the overall distribution of the non-group code. Or in other words, in the AB and $SU(3)$ codes, pairs of matrices that have very small difference determinant is scarce. The expected difference determinant, $\mathrm{E}_{1\le i<j\le L} \det|U_i - U_j|$, of the AB and $SU(3)$ codes may be as large as that of the non-group code. When the rate is high, the probability that matrices that are close to others are transmitted is small. It is the expected "distance"[4] instead of the worse-case "distance" that dominants the BLER.

This plot shows that both the AB codes and the $SU(3)$ codes have much better performance than the group-based code. They even have the same good performance as the elaborately designed non-group codes. Another advantage is that the AB

[4]Here, the distance of two matrices, $A$ and $B$, is $|\det(A-B)|$. It is quoted since it is not a metric by definition. For definition of metric, see [Yos78].

codes have a fast decoding algorithm while the decoding of both the group-based and non-group codes needs exhaustive search.

### 6.6.4 AB Code vs. Group-Based Code at Higher Rates

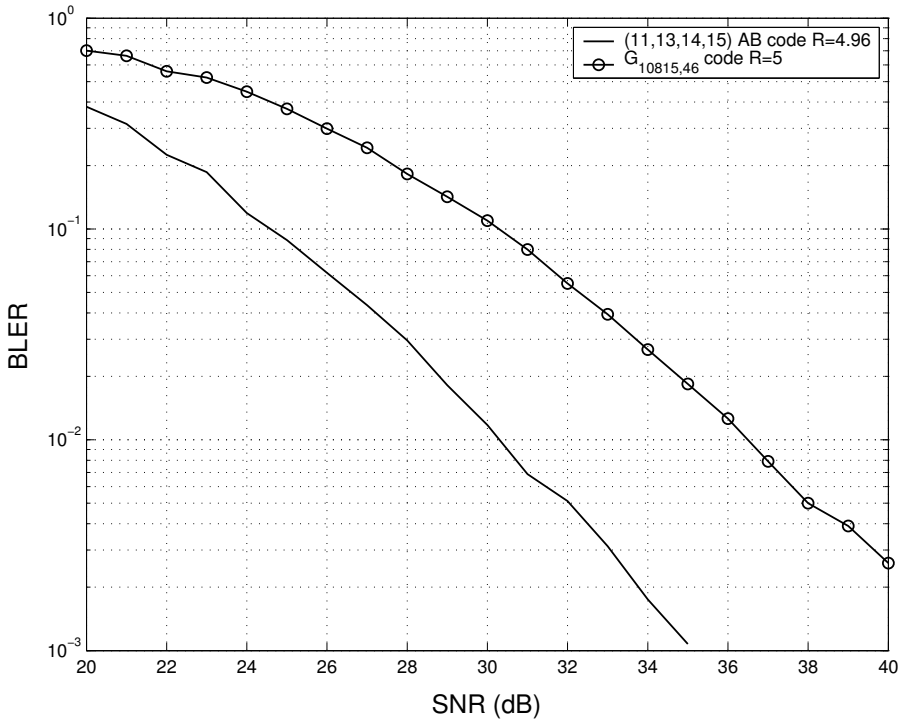


Figure 6.4: Comparison of the rate $4.9580, (11, 13, 14, 15)$ type II AB code with the rate 5 $G_{10815,46}$ code

In this subsection, the $(11, 13, 14, 15)$ type II AB code is compared with the $G_{10815,46}$ group-based code. The rate of the AB code is 4.9580 and the rate of the group-based code is 5. The performance is shown in Figure 6.4. The line with circles indicates BLER of the $G_{10815,46}$ code and the solid line shows BLER of the AB code. The plot shows that the AB code has a much better performance. For example, at the BLER of $10^{-3}$, the AB code is 6dB better and the performance gap is even higher for lower BLERs or higher SNRs. As mentioned before, the AB code has a fast decoding algorithm while decoding the group-based codes needs exhaustive search. Therefore,

at high rates, AB codes have great advantages over group-based codes in both the performance and decoding complexity.

## 6.7 Conclusion

In this chapter, the research on the idea of differential unitary space-time code designs based on Lie groups with rank 2, which is first discussed in Chapter 4, is continued. The special unitary Lie group $SU(3)$ is analyzed, which has dimension 8 and rank 2. The group is not fixed-point-free, but a method to design fully-diverse codes, which are subsets of the group, is described. Furthermore, motivated by the structure of the $SU(3)$ codes proposed, a simpler code, called the AB code, is proposed. Both codes are suitable for systems with three transmit antennas. Necessary conditions for the full diversity of both codes are given and our conjecture is that they are also sufficient conditions. The codes have simple formulas from which their diversity products can be calculated in a fast way. A fast maximum-likelihood decoding algorithm for AB codes based on complex sphere decoding is given by which the codes can be decoded in a complexity that is polynomial in the rate and dimension. Simulation results show that both $SU(3)$ codes and AB codes perform as well as the finite group-based codes at low rates. But they do not need the exhaustive search (of exponentially growing size) required of group-based codes and therefore are far superior in terms of decoding complexity. Both $SU(3)$ and AB codes have great advantages over the finite group-based codes at high rates and perform as well as the carefully designed non-group code at rate 4.

# 6.8 Appendices

## 6.8.1 Proof of Theorem 6.1

**Proof:** It is easy to check that any matrix $U$ that satisfies (6.1) is in $SU(3)$ by checking that $UU^* = I_3$ and $\det U = 1$. Now what is left to prove is that any matrix $U \in SU(3)$ can be written as (6.1). Partition $U$ into $\begin{bmatrix} a & U_{12} \\ U_{21} & U_{22} \end{bmatrix}$ where $a$ is a complex number, $U_{12}$ is $1 \times 2$, $U_{21}$ is $2 \times 1$, and $U_{22}$ is $2 \times 2$. Since $UU^* = I_3$,

$$\begin{bmatrix} a & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} \bar{a} & U_{21}^* \\ U_{12}^* & U_{22}^* \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & I_2 \end{bmatrix}.$$

Comparing the $(1,1)$ entries, $U_{12}U_{12}^* = 1 - |a|^2$ can be obtained. Therefore $|a|^2 < 1$. Comparing the $(1,2)$ entries,

$$aU_{21}^* + U_{12}U_{22}^* = \mathbf{0}_{12} \Rightarrow U_{21}^* = -a^{-1}U_{12}U_{22}^*.$$

Comparing the $(2,2)$ entries and using the above equality, we have

$$U_{21}U_{21}^* + U_{22}U_{22}^* = I_2 \Rightarrow -a^{-1}U_{21}U_{12}U_{22}^* + U_{22}U_{22}^* = I_2 \Rightarrow a^{-1}U_{21}U_{12} = U_{22} - U_{22}^{-*}.$$

Now let's look at the unit determinant-constraint, which gives us

$$1 = \det U = a \cdot \det(U_{22} - U_{21}a^{-1}U_{12}) = a \cdot \det U_{22}^{-*} = a(\overline{\det U_{22}})^{-1}.$$

Therefore, $\det U_{22} = \bar{a}$. From $U_{21}U_{21}^* + U_{22}U_{22}^* = I_2$, it is obvious that $I_2 - U_{22}U_{22}^*$ has rank 1. So, 1 is an eigenvalue of $U_{22}U_{22}^*$. The other eigenvalue must be $|a|^2$ since $\det U_{22}U_{22}^* = |a|^2$. Thus, the Hermitian and positive matrix $U_{22}U_{22}^*$ can be

decomposed as $U_{22}U_{22}^* = \Phi \begin{bmatrix} 1 & 0 \\ 0 & |\alpha|^2 \end{bmatrix} \Phi^*$, for some unitary matrix $\Phi$ with determinant 1. Therefore, there exists a unitary matrix $\Psi$ with determinant 1 such that $U_{22} = \Phi \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \Psi$.

Again, from $U_{21}U_{21}^* + U_{22}U_{22}^* = I_2$,

$$\begin{aligned} & U_{21}U_{21}^* \\ = \; & I_2 - U_{22}U_{22}^* \\ = \; & \Phi \begin{bmatrix} 0 & 0 \\ 0 & 1-|a|^2 \end{bmatrix} \Phi^* \\ = \; & \Phi \begin{bmatrix} 0 \\ \sqrt{1-|a|^2}e^{j\zeta} \end{bmatrix} \begin{bmatrix} 0 & \sqrt{1-|a|^2}e^{-j\zeta} \end{bmatrix} \Phi^*. \end{aligned}$$

A general solution for $U_{21}$ is

$$U_{21} = \Phi \begin{bmatrix} 0 \\ \sqrt{1-|a|^2}e^{j\zeta} \end{bmatrix},$$

where $\zeta$ is an arbitrary angle. By similar argument, a general solution for $U_{12}$ is,

$$U_{12} = \begin{bmatrix} 0 & \sqrt{1-|\alpha|^2}e^{j\eta} \end{bmatrix} \Psi$$

where $\eta$ is an arbitrary angle. Also,

$$\begin{aligned} & aU_{21}^* + U_{12}U_{22}^* = \mathbf{0}_{12} \\ \Rightarrow \; & a \begin{bmatrix} 0 & \sqrt{1-|a|^2}e^{-j\zeta} \end{bmatrix} \Phi^* + \begin{bmatrix} 0 & \sqrt{1-|a|^2}e^{j\eta} \end{bmatrix} \Psi\Psi^* \begin{bmatrix} 1 & 0 \\ 0 & a \end{bmatrix} \Phi^* = 0 \\ \Rightarrow \; & e^{j\eta} = -e^{-j\zeta}. \end{aligned}$$

Therefore, we have proved that matrices in $SU(3)$ can be written as

$$\begin{bmatrix} a & \begin{bmatrix} 0 & -\sqrt{1-|a|^2}e^{-j\zeta} \end{bmatrix}\Psi \\ \Phi\begin{bmatrix} 0 \\ \sqrt{1-|a|^2}e^{j\zeta} \end{bmatrix} & \Phi\begin{bmatrix} 1 & 0 \\ 0 & \bar{a} \end{bmatrix}\Psi \end{bmatrix}$$
$$= \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Phi \end{bmatrix} \begin{bmatrix} \bar{\alpha} & 0 & -\sqrt{1-|\alpha|^2}e^{-j\zeta} \\ 0 & 1 & 0 \\ \sqrt{1-|\alpha|^2}e^{j\zeta} & 0 & \alpha \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}_{12} \\ \mathbf{0}_{21} & \Psi \end{bmatrix}.$$

Since

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-j\beta} & 0 \\ 0 & 0 & e^{j\beta} \end{bmatrix} \begin{bmatrix} a & 0 & -\sqrt{1-|a|^2}e^{-j\zeta} \\ 0 & 1 & 0 \\ \sqrt{1-|a|^2}e^{j\zeta} & 0 & \bar{a} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{j\beta} & 0 \\ 0 & 0 & e^{-j\beta} \end{bmatrix}$$
$$= \begin{bmatrix} a & 0 & -\sqrt{1-|a|^2}e^{-j(\zeta+\beta)} \\ 0 & 1 & 0 \\ \sqrt{1-|a|^2}e^{j(\zeta+\beta)} & 0 & \bar{a} \end{bmatrix}$$

for any real angle $\beta$, the angle $\zeta$ is a redundant degree of freedom. Therefore, we can set $\zeta = 0$. Thus, (6.1) is obtained. □

### 6.8.2 Proof of Theorem 6.2

**Proof:** Define $\theta_i = 2\pi\left(\frac{p_i}{P} - \frac{q_i}{Q}\right)$ and $\xi_i = 2\pi\left(\frac{r_i}{R} + \frac{s_i}{S}\right)$ for $i = 1, 2$. Furthermore, define $\gamma_1 = \frac{p_1-p_2}{2P} + \frac{q_1-q_2}{2Q}$ and $\gamma_2 = \frac{r_1-r_2}{2P} - \frac{s_1-s_2}{2Q}$. For any $U_1(p_1, q_1, r_1, s_1)$ and $U_2(p_2, q_2, r_2, s_2)$ in $\mathcal{C}^{(1)}$, we can write $U_1 = A^{(1)}_{(p_1,q_1)}B^{(1)}_{(r_1,s_1)}$ and $U_2 = A^{(1)}_{(p_2,q_2)}B^{(1)}_{(r_2,s_2)}$

where $A^{(1)}$ and $B^{(1)}$ are as defined in (6.4) and (6.5). Therefore,

$$
\begin{aligned}
&\left(A^{(1)}_{(p_2,q_2)}\right)^{-1} A^{(1)}_{(p_1,q_1)} \\
&= \begin{bmatrix} e^{-j\theta_2} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}e^{-2\pi\frac{p_2}{P}} & -\frac{1}{\sqrt{2}}e^{2\pi\frac{q_2}{Q}} \\ 0 & \frac{1}{\sqrt{2}}e^{-2\pi\frac{q_2}{Q}}e^{j\theta_2} & \frac{1}{\sqrt{2}}e^{2\pi\frac{p_2}{P}}e^{j\theta_2} \end{bmatrix} \begin{bmatrix} e^{j\theta_1} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}e^{2\pi\frac{p_1}{P}} & \frac{1}{\sqrt{2}}e^{2\pi\frac{q_1}{Q}}e^{-j\theta_1} \\ 0 & -\frac{1}{\sqrt{2}}e^{-2\pi\frac{q_1}{Q}} & \frac{1}{\sqrt{2}}e^{-2\pi\frac{p_1}{P}}e^{-j\theta_1} \end{bmatrix} \\
&= \begin{bmatrix} e^{j(\theta_1-\theta_2)} & & \\ 0 & \frac{1}{2}\left(e^{2\pi j\frac{p_1-p_2}{P}} + e^{-2\pi j\frac{q_1-q_2}{Q}}\right) & \frac{1}{2}\left(e^{2\pi j(-\frac{p_2}{P}+\frac{q_1}{Q})} - e^{2\pi j(-\frac{p_1}{P}+\frac{q_2}{Q})}\right)e^{-j\theta_1} \\ 0 & \frac{1}{2}\left(e^{2\pi j(\frac{p_1}{P}-\frac{q_2}{Q})} - e^{2\pi j(\frac{p_2}{P}-\frac{q_1}{Q})}\right)e^{j\theta_2} & \frac{1}{2}\left(e^{-2\pi j\frac{p_1-p_2}{P}} + e^{2\pi j\frac{q_1-q_2}{Q}}\right)e^{-j(\theta_1-\theta_2)} \end{bmatrix} \\
&= \begin{bmatrix} e^{j(\theta_1-\theta_2)} & 0 & 0 \\ 0 & e^{2\pi j\left(\frac{p_1-p_2}{2P}-\frac{q_1-q_2}{2Q}\right)}\cos\gamma_1 & e^{-j\theta_1}je^{2\pi j\left(-\frac{p_1+p_2}{2P}+\frac{q_1+q_2}{2Q}\right)}\sin\gamma_1 \\ 0 & e^{j\theta_2}je^{2\pi j\left(\frac{p_1+p_2}{2P}-\frac{q_1+q_2}{2Q}\right)}\sin\gamma_1 & e^{-j(\theta_1-\theta_2)}e^{2\pi j\left(-\frac{p_1-p_2}{2P}+\frac{q_1-q_2}{2Q}\right)}\cos\gamma_1 \end{bmatrix} \\
&= \begin{bmatrix} e^{-j\theta_2} & 0 & 0 \\ 0 & e^{2\pi j(-\frac{p_2}{2P}+\frac{q_2}{2Q})} & 0 \\ 0 & 0 & e^{j\theta_2}e^{2\pi j(\frac{p_2}{2P}-\frac{q_2}{2Q})} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma_1 & j\sin\gamma_1 \\ 0 & j\sin\gamma_1 & \cos\gamma_1 \end{bmatrix} \\
&\qquad \begin{bmatrix} e^{j\theta_1} & 0 & 0 \\ 0 & e^{2\pi j(\frac{p_1}{2P}-\frac{q_1}{2Q})} & 0 \\ 0 & 0 & e^{-j\theta_1}e^{2\pi j(-\frac{p_1}{2P}+\frac{q_1}{2Q})} \end{bmatrix},
\end{aligned}
$$

and

$$
\begin{aligned}
&B^{(1)}_{(r_2,s_2)} \left(B^{(1)}_{(r_1,s_1)}\right)^{-1} \\
&= \begin{bmatrix} \frac{1}{\sqrt{2}}e^{2\pi\frac{r_2}{R}}e^{j\xi_2} & \frac{1}{\sqrt{2}}e^{2\pi\frac{s_2}{S}}e^{j\xi_2} & 0 \\ -\frac{1}{\sqrt{2}}e^{-2\pi\frac{s_2}{S}} & \frac{1}{\sqrt{2}}e^{-2\pi\frac{r_2}{R}} & 0 \\ 0 & 0 & e^{-j\xi_2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}}e^{-2\pi\frac{r_1}{R}}e^{-j\xi_1} & -\frac{1}{\sqrt{2}}e^{2\pi\frac{s_1}{S}} & 0 \\ \frac{1}{\sqrt{2}}e^{-2\pi\frac{s_1}{S}}e^{-j\xi_1} & \frac{1}{\sqrt{2}}e^{2\pi\frac{r_1}{R}} & 0 \\ 0 & 0 & e^{j\xi_1} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&= \begin{bmatrix} \frac{1}{2}\left(e^{-2\pi j\frac{r_1-r_2}{R}}+e^{-2\pi j\frac{s_1-s_2}{S}}\right)e^{-j(\xi_1-\xi_2)} & \frac{1}{2}\left(e^{2\pi j(\frac{r_1}{R}+\frac{s_2}{S})}-e^{2\pi j(\frac{r_2}{R}+\frac{s_1}{S})}\right)e^{j\xi_2} & 0 \\ \frac{1}{2}\left(e^{2\pi j(-\frac{r_2}{R}-\frac{s_1}{S})}-e^{2\pi j(-\frac{r_1}{R}-\frac{s_2}{S})}\right)e^{-j\xi_1} & \frac{1}{2}\left(e^{2\pi j\frac{r_1-r_2}{R}}+e^{2\pi j\frac{s_1-s_2}{S}}\right) & 0 \\ 0 & 0 & e^{j(\xi_1-\xi_2)} \end{bmatrix} \\
&= \begin{bmatrix} e^{-j(\xi_1-\xi_2)}e^{2\pi j\left(-\frac{r_1-r_2}{2R}-\frac{s_1-s_2}{2S}\right)}\cos\gamma_2 & je^{j\xi_2}e^{2\pi j\left(\frac{r_1+r_2}{2R}+\frac{s_1+s_2}{2S}\right)}\sin\gamma_2 & 0 \\ je^{-j\xi_1}e^{2\pi j\left(-\frac{r_1+r_2}{2R}-\frac{s_1+s_2}{2S}\right)}\sin\gamma_2 & e^{2\pi j(\frac{r_1-r_2}{2R}+\frac{s_1-s_2}{2S})}\cos\gamma_2 & 0 \\ 0 & 0 & e^{j(\xi_1-\xi_2)} \end{bmatrix} \\
&= \begin{bmatrix} e^{j\xi_2}e^{2\pi j\left(\frac{r_2}{2R}+\frac{s_2}{2S}\right)} & 0 & 0 \\ 0 & e^{2\pi j\left(-\frac{r_2}{2R}-\frac{s_2}{2S}\right)} & 0 \\ 0 & 0 & e^{-j\xi_2} \end{bmatrix} \begin{bmatrix} \cos\gamma_2 & j\sin\gamma_2 & 0 \\ j\sin\gamma_2 & \cos\gamma_2 & \\ 0 & 0 & 1 \end{bmatrix} \\
&\qquad \begin{bmatrix} e^{-j\xi_1}e^{2\pi j\left(-\frac{r_1}{2R}-\frac{s_1}{2S}\right)} & 0 & 0 \\ 0 & e^{2\pi j\left(\frac{r_1}{2R}+\frac{s_1}{2S}\right)} & 0 \\ 0 & 0 & e^{j\xi_1} \end{bmatrix}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\det(U_1(p_1,q_1,r_1,s_1)-U_2(p_2,q_2,r_2,s_2) \\
&= \det A^{(1)}_{(p_2,q_2)} \det B^{(1)}_{(r_1,s_1)} \det\left(\left(A^{(1)}_{(p_2,q_2)}\right)^{-1} A^{(1)}_{(p_1,q_1)} - B^{(1)}_{(r_2,s_2)}\left(B^{(1)}_{(r_1,s_1)}\right)^{-1}\right) \\
&= \det\left(\begin{bmatrix} e^{-j\theta_2} & 0 & 0 \\ 0 & e^{2\pi j(-\frac{p_2}{2P}+\frac{q_2}{2Q})} & 0 \\ 0 & 0 & e^{j\theta_2}e^{2\pi j(\frac{p_2}{2P}-\frac{q_2}{2Q})} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma_1 & j\sin\gamma_1 \\ 0 & j\sin\gamma_1 & \cos\gamma_2 \end{bmatrix}\right. \\
&\qquad \left.\begin{bmatrix} e^{j\theta_1} & 0 & 0 \\ 0 & e^{2\pi j(\frac{p_1}{2P}-\frac{q_1}{2Q})} & 0 \\ 0 & 0 & e^{-j\theta_1}e^{2\pi j(-\frac{p_1}{2P}+\frac{q_1}{2Q})} \end{bmatrix}\right)
\end{aligned}
$$

$$
\left.
-\begin{bmatrix} e^{j\xi_2}e^{2\pi j\left(\frac{r_2}{2R}+\frac{s_2}{2S}\right)} & 0 & 0 \\ 0 & e^{2\pi j\left(-\frac{r_2}{2R}-\frac{s_2}{2S}\right)} & 0 \\ 0 & 0 & e^{-j\xi_2} \end{bmatrix}
\begin{bmatrix} \cos\gamma_2 & j\sin\gamma_2 & 0 \\ j\sin\gamma_2 & \cos\gamma_2 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\right.
$$

$$
\left.
\begin{bmatrix} e^{-j\xi_1}e^{2\pi j\left(-\frac{r_1}{2R}-\frac{s_1}{2S}\right)} & 0 & 0 \\ 0 & e^{2\pi j\left(\frac{r_1}{2R}+\frac{s_1}{2S}\right)} & 0 \\ 0 & 0 & e^{j\xi_1} \end{bmatrix}
\right)
$$

$$
\begin{aligned}
= \quad & e^{j(\theta_i-\theta_2+\xi_i-\xi_2)} \\
& \det\left(
\begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{2\pi j(-\frac{p_2}{2P}+\frac{q_2}{2Q})} & 0 \\ 0 & 0 & e^{j(\theta_2+\xi_2)}e^{2\pi j(\frac{p_2}{2P}-\frac{q_2}{2Q})} \end{bmatrix}
\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma_1 & j\sin\gamma_1 \\ 0 & j\sin\gamma_1 & \cos\gamma_1 \end{bmatrix}
\right. \\
& \qquad \left.
\begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{2\pi j(\frac{p_1}{2P}-\frac{q_1}{2Q})} & 0 \\ 0 & 0 & e^{-j(\theta_1+\xi_1)}e^{2\pi j(-\frac{p_1}{2P}+\frac{q_1}{2Q})} \end{bmatrix}
\right. \\
& \qquad -\begin{bmatrix} e^{j(\theta_2+\xi_2)}e^{2\pi j\left(\frac{r_2}{2R}+\frac{s_2}{2S}\right)} & 0 & 0 \\ 0 & e^{2\pi j\left(-\frac{r_2}{2R}-\frac{s_2}{2S}\right)} & 0 \\ 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} \cos\gamma_2 & j\sin\gamma_2 & 0 \\ j\sin\gamma_2 & \cos\gamma_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
& \qquad \left.
\begin{bmatrix} e^{-j(\theta_1+\xi_1)}e^{2\pi j\left(-\frac{r_1}{2R}-\frac{s_1}{2S}\right)} & 0 & 0 \\ 0 & e^{2\pi j\left(\frac{r_1}{2R}+\frac{s_1}{2S}\right)} & 0 \\ 0 & 0 & 1 \end{bmatrix}
\right) \\
= \quad & e^{j(\theta_i-\theta_2+\xi_i-\xi_2)} \\
& \det\left(
\begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{2\pi j\left(\frac{p_1-p_2}{2P}-\frac{q_1-q_2}{2Q}\right)}\cos\gamma_1 & je^{-j(\theta_1+\xi_1)}e^{2\pi j\left(-\frac{p_1+p_2}{2P}+\frac{q_1+q_2}{2Q}\right)}\sin\gamma_1 \\ 0 & je^{j(\theta_2+\xi_2)}e^{2\pi j\left(\frac{p_1+p_2}{2P}-\frac{q_1+q_2}{2Q}\right)}\sin\gamma_1 & e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}e^{2\pi j\left(-\frac{p_1-p_2}{2P}+\frac{q_1-q_2}{2Q}\right)}\cos\gamma_1 \end{bmatrix}
\right. \\
& \qquad \left.
-\begin{bmatrix} e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}e^{2\pi j\left(-\frac{r_1-r_2}{2R}-\frac{s_1-s_2}{2S}\right)}\cos\gamma_2 & je^{j(\theta_2+\xi_2)}e^{2\pi j\left(\frac{r_1+r_2}{2R}+\frac{s_1+s_2}{2S}\right)}\sin\gamma_2 & 0 \\ je^{-j(\theta_1+\xi_1)}e^{2\pi j\left(-\frac{r_1+r_2}{2R}-\frac{s_1+s_2}{2S}\right)}\sin\gamma_2 & e^{2\pi j\left(\frac{r_1-r_2}{2R}+\frac{s_1-s_2}{2S}\right)}\cos\gamma_2 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\right).
\end{aligned}
$$

Define

$$\begin{aligned} y &= je^{2\pi j\left(-\frac{p_1+p_2}{2P}+\frac{q_1+q_2}{2Q}\right)}\sin 2\pi\left(\frac{p_1-p_2}{2P}+\frac{q_1-q_2}{2Q}\right), \\ z &= je^{2\pi j\left(\frac{r_1+r_2}{2R}+\frac{s_1+s_2}{2S}\right)}\sin 2\pi\left(\frac{r_1-r_2}{2R}-\frac{s_1-s_2}{2S}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} & \det(U_1-U_2) \\ = \ & \bar{\Theta}\det\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & x & e^{-j(\theta_1+\xi_1)}y \\ 0 & -e^{j(\theta_2+\xi_2)}\bar{y} & e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}\bar{x} \end{bmatrix} - \begin{bmatrix} e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}w & e^{j(\theta_2+\xi_2)}z & 0 \\ -e^{-j(\theta_1+\xi_1)}\bar{z} & \bar{w} & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \\ = \ & \bar{\Theta}\det\begin{bmatrix} 1-e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}w & -e^{j(\theta_2+\xi_2)}z & 0 \\ e^{-j(\theta_1+\xi_1)}\bar{z} & x-\bar{w} & e^{-j(\theta_1+\xi_1)}y \\ 0 & -e^{j(\theta_2+\xi_2)}\bar{y} & e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}\bar{x}-1 \end{bmatrix} \\ = \ & \bar{\Theta}\big[(1-e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}w)(x-\bar{w})(e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}\bar{x}-1)+ \\ & \quad e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}|z|^2(e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}\bar{x}-1)+e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}|y|^2(1-e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}w)\big] \\ = \ & \bar{\Theta}\big[(1-\Theta w)(x-\bar{w})(\Theta\bar{x}-1)+\Theta|z|^2(\Theta\bar{x}-1)+\Theta|y|^2(1-\Theta w)\big] \\ = \ & \bar{\Theta}\big[(x-\Theta wx-\bar{w}+\Theta|w|^2)(\Theta\bar{x}-1)+\Theta|z|^2(\Theta\bar{x}-1)+\Theta|y|^2(1-\Theta w)\big] \\ = \ & \bar{\Theta}\big[(\Theta\bar{x}-1)(x-\Theta wx-\bar{w}+\Theta)+\Theta|y|^2(1-\Theta w)\big] \\ = \ & \bar{\Theta}\big[\Theta|x|^2-x-\Theta^2|x|^2w+\Theta wx-\Theta\bar{x}\bar{w}+\bar{w}+\Theta^2\bar{x}-\Theta+\Theta|y|^2-\Theta^2|y|^2w'\big] \\ = \ & \bar{\Theta}\big[\Theta wx-\Theta\bar{x}\bar{w}-(x-\Theta^2\bar{x})-(\Theta^2w-\bar{w})\big] \\ = \ & \bar{\Theta}[wx-\bar{x}\bar{w}-(\bar{\Theta}x-\Theta\bar{x})-(\Theta w-\bar{\Theta}\bar{w})] \\ = \ & 2j(\mathcal{I}m(wx)-\mathcal{I}m\bar{\Theta}x-\mathcal{I}m\Theta w) \\ = \ & 2j\mathcal{I}m[(1-\bar{\Theta}x)(1-\Theta w)]. \end{aligned}$$

□

### 6.8.3 Proof of Theorem 6.4

**Proof:** Define $\theta_i = 2\pi j\left(\pm\frac{p_i}{P} \pm \frac{q_i}{Q}\right)$ and $\xi_i = 2\pi j\left(\pm\frac{r_i}{R} \pm \frac{s_i}{S}\right)$ for $i = 1, 2$. For any $U_1(p_1, q_1, r_1, s_1)$ and $U_2(p_2, q_2, r_2, s_2)$ in $\mathcal{C}^{(2)}$, we write $U_1 = A^{(2)}_{(p_1,q_1)} B^{(2)}_{(r_1,s_1)}$ and $U_2 = A^{(2)}_{(p_2,q_2)} B^{(2)}_{(r_2,s_2)}$, where $A^{(2)}$ and $B^{(2)}$ are as defined in (6.11). By a similar argument to the proof of Theorem 6.2,

$$
\begin{aligned}
&\left(A^{(2)}_{(p_2,q_2)}\right)^{-1} A^{(2)}_{(p_1,q_1)} \\
&= \begin{bmatrix} e^{-j\theta_2} & 0 & 0 \\ 0 & e^{2\pi j\left(-\frac{p_2}{2P}+\frac{q_2}{2Q}\right)} & 0 \\ 0 & 0 & e^{2\pi j\left(\frac{p_2}{2P}-\frac{q_2}{2Q}\right)} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 2\pi\left(\frac{p_1-p_2}{2P}+\frac{q_1-q_2}{2Q}\right) & j\sin 2\pi\left(\frac{p_1-p_2}{2P}+\frac{q_1-q_2}{2Q}\right) \\ 0 & j\sin 2\pi\left(\frac{p_1-p_2}{2P}+\frac{q_1-q_2}{2Q}\right) & \cos 2\pi\left(\frac{p_1-p_2}{2P}+\frac{q_1-q_2}{2Q}\right) \end{bmatrix} \\
&\qquad \begin{bmatrix} e^{j\theta_1} & 0 & 0 \\ 0 & e^{2\pi j\left(\frac{p_1}{2P}-\frac{q_1}{2Q}\right)} & 0 \\ 0 & 0 & e^{2\pi j\left(-\frac{p_1}{2P}+\frac{q_1}{2Q}\right)} \end{bmatrix},
\end{aligned}
$$

and

$$
\begin{aligned}
&B^{(2)}_{(r_2,s_2)} \left(B^{(2)}_{(r_1,s_1)}\right)^{-1} \\
&= \begin{bmatrix} e^{2\pi j\left(\frac{r_2}{2R}+\frac{s_2}{2S}\right)} & 0 & 0 \\ 0 & e^{2\pi j\left(-\frac{r_2}{2R}-\frac{s_2}{2S}\right)} & 0 \\ 0 & 0 & e^{-j\xi_2} \end{bmatrix} \begin{bmatrix} \cos 2\pi\left(\frac{r_1-r_2}{2R}-\frac{s_1-s_2}{2S}\right) & j\sin 2\pi\left(\frac{r_1-r_2}{2R}-\frac{s_1-s_2}{2S}\right) & 0 \\ j\sin 2\pi\left(\frac{r_1-r_2}{2R}-\frac{s_1-s_2}{2S}\right) & \cos 2\pi\left(\frac{r_1-r_2}{2R}-\frac{s_1-s_2}{2S}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&\qquad \begin{bmatrix} e^{2\pi j\left(-\frac{r_1}{2R}-\frac{s_1}{2S}\right)} & 0 & 0 \\ 0 & e^{2\pi j\left(\frac{r_1}{2R}+\frac{s_1}{2S}\right)} & 0 \\ 0 & 0 & e^{j\xi_1} \end{bmatrix}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\det(U_1(p_1, q_1, r_1, s_1) - U_2(p_2, q_2, r_2, s_2) \\
&= \det A^{(2)}_{(p_2,q_2)} \det B^{(2)}_{(r_1,s_1)} \det\left(\left(A^{(2)}_{(p_2,q_2)}\right)^{-1} A^{(2)}_{(p_1,q_1)} - B^{(2)}_{(r_2,s_2)} \left(B^{(2)}_{(r_1,s_1)}\right)^{-1}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= e^{j\theta_1}e^{-j\xi_2}\det\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & x & e^{-j\xi_1}y \\ 0 & -e^{j\xi_2}\bar{y} & e^{-j(\xi_1-\xi_2)}\bar{x} \end{bmatrix} - \begin{bmatrix} e^{-j(\theta_1-\theta_2)}w & e^{j\theta_2}z & 0 \\ -e^{-j\theta_1}\bar{z} & \bar{w} & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) \\
&= e^{j\theta_1}e^{-j\xi_2}\det\begin{bmatrix} 1-e^{-j(\theta_1-\theta_2)}w & -e^{j\theta_2}z & 0 \\ e^{-j\theta_1}\bar{z} & x-\bar{w} & e^{-j\xi_1}y \\ 0 & -e^{j\xi_2}\bar{y} & e^{-j(\xi_1-\xi_2)}\bar{x}-1 \end{bmatrix} \\
&= e^{j\theta_1}e^{-j\xi_2}\left[(e^{-j(\xi_1-\xi_2)}-e^{-j(\theta_1-\theta_2)}) - (x-e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}\bar{x}) + (\bar{w}-e^{-j(\theta_1-\theta_2+\xi_1-\xi_2)}w)\right. \\
&\qquad \left. +(e^{-j(\theta_1-\theta_2)}wx - e^{-j(\xi_1-\xi_2)}\bar{x}\bar{w})\right] \\
&= e^{j\theta_1}e^{-j\xi_2}\left[(\bar{\Theta}_2^2-\bar{\Theta}_1^2) - (x-\bar{\Theta}_1^2\bar{\Theta}_2^2\bar{x}) + (\bar{w}-\bar{\Theta}_1^2\bar{\Theta}_2^2 w) + (\bar{\Theta}_1^2 wx - \bar{\Theta}_2^2\bar{x}\bar{w})\right] \\
&= e^{j\theta_1}e^{-j\xi_2}\bar{\Theta}_1\bar{\Theta}_2\left[(\Theta_1\bar{\Theta}_2 - \overline{\Theta_1\bar{\Theta}_2}) - (\Theta_1\Theta_2 x - \overline{\Theta_1\Theta_2 x}) - (\bar{\Theta}_1\bar{\Theta}_2 w - \overline{\bar{\Theta}_1\bar{\Theta}_2 w})\right. \\
&\qquad \left. +(\bar{\Theta}_1\Theta_2 wx - \overline{\bar{\Theta}_1\Theta_2 wx})\right] \\
&= 2je^{j\theta_1}e^{-j\xi_2}\bar{\Theta}_1\bar{\Theta}_2\mathcal{I}m(\Theta_1\bar{\Theta}_2 - \Theta_1\Theta_2 x - \bar{\Theta}_1\bar{\Theta}_2 w + \bar{\Theta}_1\Theta_2 wx) \\
&= 2je^{j\theta_1}e^{-j\xi_2}\bar{\Theta}_1\bar{\Theta}_2\mathcal{I}m[(\Theta_1-\bar{\Theta}_1 w)(\bar{\Theta}_2-\Theta_2 x)] \\
&= 2je^{j\theta_1}e^{-j\xi_2}\bar{\Theta}_1\bar{\Theta}_2\mathcal{I}m[\Theta_1\bar{\Theta}_2(1-\bar{\Theta}_1^2 w)(1-\Theta_2^2 x)].
\end{aligned}
$$

□

# Chapter 7 Using Space-Time Codes in Wireless Networks

## 7.1 Abstract

In this chapter, the idea of space-time coding devised for multiple-antenna systems is applied to communications over wireless relay networks. A two-stage protocol is used in the network communications, where in one stage the transmitter transmits information and in the other, the relay nodes encode their received signals into a "distributed" linear dispersion space-time code, and then transmit the coded signals to the receive node. It is shown that at high SNR, the PEP behaves as $\left(\frac{\log P}{P}\right)^{\min\{T,R\}}$, with $T$ the coherence interval, $R$ the number of relay nodes, and $P$ the total transmitted power. Thus, apart from the $\log P$ factor and assuming $T \geq R$, the network has the same diversity as a multiple-antenna system with $R$ transmit antennas, which is the same as assuming that the $R$ relay nodes can fully cooperate and have full knowledge of the transmitted signal. It is further shown that for a fixed total transmit power across the entire network, the optimal power allocation is for the transmitter to expend half the power and for the relays to collectively expend the other half. It is also proved that at low and high SNR, the coding gain is the same as that of a multiple-antenna system with $R$ transmit antennas. However, at intermediate SNR, it can be quite different.

## 7.2 Introduction

The communication systems that have been discussed or worked with in previous chapters are point-to-point communication systems, which only have two users: one is the transmitter and the other is the receiver. Recently, communications in wireless networks are of great interest because of their diverse applications. Wireless networks consist of a number of nodes or users communicating over wireless channels. Roughly, there are two types of wireless networks according to the structure. One type is networks that have a master node or base station. All nodes communicate with the base station directly and the base station is in control of all transmissions and forwarding data to the intended users. A cellular phone system is the most popular example of this kind of wireless networks. Another example is satellite communication systems. The other kind of wireless networks is ad hoc or sensory networks, which is the type of networks that are going to be dealt with in this chapter.

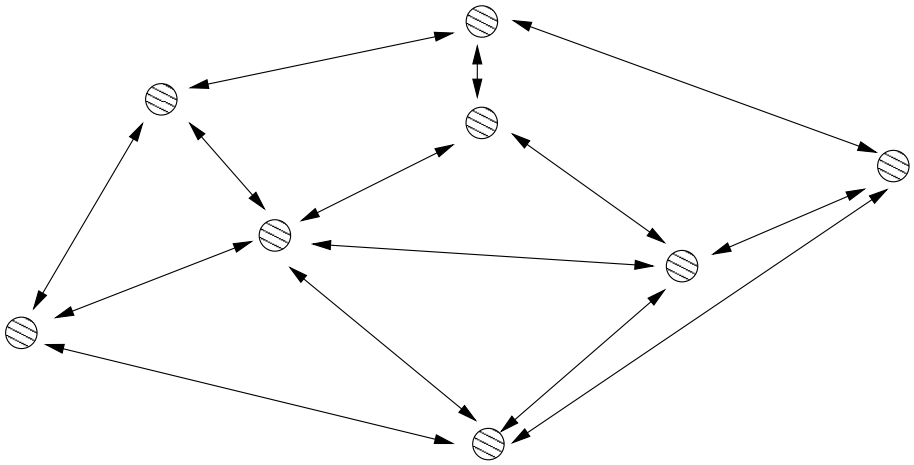

Figure 7.1: Ad hoc network

An ad hoc wireless network is a collection of wireless mobile nodes that self-configure to form a network without the aid of any established infrastructure [GW02]. Figure 7.1 is a simple diagram of wireless ad hoc networks. A wireless links exists between each pair of nodes. In the figure, only some of the links are represented. In ad

hoc wireless networks, there is no master node or base station. All communications are peer to peer. As every node may not be in the direct communication range[1] of every other node, nodes can cooperate in routing each other's data. Therefore, transmissions may be completed by one-hop routing or even multiple-hop routing. In addition, nodes in an ad-hoc network may be mobile. The difference between an ad-hoc network and sensory network is that in the former, nodes may be mobile and there can be more than one pair of nodes communicating at the same time, while, for sensory networks, the nodes are normally static and there is only one pair of nodes communicating at a time.

According to the features mentioned above, ad-hoc and sensory networks can be rapidly deployed and reconfigured, can be easily tailored to specific applications, and are robust due to the distributed nature and redundancy of the nodes. Because of these advantages of ad-hoc and sensory networks, they have many applications, for example, the data network, the home network, the wireless network of mobile laptops, PDAs and smart phones, the automated transportation systems, sensor dust, Bluetooth [Har00], etc.[2] However, because of exactly the same unique features, the analysis on wireless ad hoc networks is very difficult in networking, signal processing, and especially information theoretical aspects.

There are many preliminary results in ad hoc wireless networks. In 2000, the capacity of wireless ad-hoc networks was first analyzed in the landmark paper [GK00]. It is proved that the optimal bit-distance product can be transported by a network placed in a disk of unit area scales as $O(\sqrt{n})$ bit-meters per second, where $n$ is the number of nodes in the network. In [GT02], it is proved that the mobility of nodes can

[1]There are many ways to define communication range of a wireless network according to the transmit power, interference, distance, and other factors in the network. For example, in the protocol model in [GK00], it is defined that one node located at $X_i$ can transmit to another node located at $X_j$ successfully if $|X_k - X_j| \geq (1+\Delta)|X_i - X_j|$ for every other node located at $X_k$ simultaneously transmitting over the same sub-channel. $\Delta$ is a positive number that models situations where a guard zone is specified by the protocol to prevent a neighboring node from transmitting on the same sub-channel at the same time.

[2]For more applications and introduction, refer to [GW02] and [Per01].

increase the per-session throughput greatly. Results on the network layer designing, interference, and energy management can be found in [PGH00, BMJ+98, RkT99, RM99, DBT03]. Although these work illuminate issues in ad hoc networks with specific network models and under specific conditions, most of the questions about ad hoc networks are still open. For example, what is the Shannon capacity region, how to do scheduling and coding to achieve capacity, and how to allocate power among the nodes? In this chapter, we use the space-time coding idea, which is widely used in multiple-antenna systems, in wireless networks to improve the performance of network communications.

As has been mentioned in Chapter 1, multiple antennas can greatly increase the capacity and reliability of a wireless communication link in a fading environment using space-time codes [Tel99, MH99, Fos96, TSC98]. Recently, with the increasing interests in ad hoc networks, researchers have been looking for methods to exploit spatial diversity using the antennas of different users in the network [SEA03a, SEA03b, TV01, LW03, NBK04]. In [LW03], the authors exploit spatial diversity using the repetition and space-time algorithms. The mutual information and outage probability of the network are analyzed. However, in their model, the relay nodes need to decode their received signals, which causes extra consumption in both time and energy and also may cause error propagation. In [NBK04], a network with a single relay under different protocols is analyzed and second order spatial diversity is achieved. In [HMC03], the authors use space-time codes based on Hurwitz-Radon matrices and conjecture a diversity factor around $R/2$ from their simulations. Also, their simulations in [CH03] show that the use of Khatri-Rao codes lowers the average bit error rate. In this chapter, relay networks with fading are considered and linear dispersion space-time codes [HH02b] are applied among the relays. The problem we are interested in is: can we increase the reliability of a wireless network by using space-time codes among the relay nodes?

A key feature of this work is that no decoding is required at the relay nodes. This has two main benefits: first, the computation at the relay nodes is considerably simplified, and second, we can avoid imposing bottlenecks on the rate by requiring some relay nodes to decode (See e.g., [DSG+03]).

The wireless relay network model used here is similar to those in [GV02, DH03]. In [GV02], the authors show that the capacity of the wireless relay network with $n$ nodes behaves like $\log n$. In [DH03], a power efficiency that behaves like $\sqrt{n}$ is obtained. Both results are based on the assumption that each relay knows its local channels so that they can work coherently. Therefore, the system should be synchronized at the carrier level. Here, it is assumed that the relay nodes do not know the channel information. All we need is the much more reasonable assumption that the system is synchronized at the symbol level.

The work in this chapter shows that the use of space-time codes among the relay nodes, with linear dispersion structure, can achieve a diversity, $\min\{T, R\}\left(1 - \frac{\log\log P}{\log P}\right)$. When $T \geq R$, the transmit diversity is linear in the number of relays (size of the network) and is a function of the total transmit power. When $P$ is very large, the diversity is approximately $R$. The coding gain for large $R$ and very large $P$ is $\det{}^{-1}(S_i - S_j)^*(S_i - S_j)$, where $S_i$ is the distributed space-time code. Therefore, with very large transmit power and a big network, the same transmit diversity and coding gain are obtained as in the multiple-antenna case, which means that the systems works as if the relays can fully cooperate and have full knowledge of the transmitted signal.

This chapter is organized as follows. In the following section, the network model and the two-step protocol is introduced. The distributed space-time coding scheme is explained in Section 7.4 and the pairwise error probability (PEP) is calculated in Section 7.5. In Section 7.6, the optimum power allocation based on the PEP is derived. Sections 7.7 and 7.8 contain the main results. The transmit diversity and the

coding gain are derived. To motivate the main results, simple approximate derivations are given first in Section 7.7, and then in Section 7.8 the more involved rigorous derivation is shown. In Section 7.9, the transmit diversity obtained in Sections 7.7 and 7.8 is improved slightly, and the optimality of the new diversity is proved. A more general distributed linear dispersion space-time coding is discussed in Section 7.10, and in Section 7.11 the transmit diversity and coding gain for a special case are obtained, which coincide with those in Sections 7.7 and 7.8. The performance of relay networks with randomly chosen distributed linear dispersion space-time codes is simulated and compared with the performance of the same space-time codes used in multiple-antenna systems with $R$ transmit antennas and one receive antenna. The details of the simulations and the BER and BLER figures are given in Section 7.12. Section 7.13 provides the conclusion and future work. Section 7.14 contains some of the technical proofs.

The work in this chapter has been published in *the Proceeding of the Third Sensory Array and Multi-Channel Signal Processing Workshop (SAM'04)* [JH04f] and is accepted in *the Forty-Second Annual Allerton Conference on Communication, Control, and Computing (Allerton'04)* [JH04a]. The journal papers, [JH04b] and [JH04c], are submitted to *IEEE Transactions on Wireless Communications*.

## 7.3 System Model

Consider a wireless network with $R + 2$ nodes which are placed randomly and independently according to some distribution. There is one transmit node and one receive node. All the other $R$ nodes work as relays. Every node has one antenna. Antennas at the relay nodes can be used for both transmission and reception. Denote the channel from the transmitter to the $i$-th relay as $f_i$, and the channel from the $i$-th relay to the receiver as $g_i$. Assume that $f_i$ and $g_i$ are independent complex Gaussian

with zero-mean and unit-variance. If the fading coefficients $f_i$ and $g_i$ are known to relay $i$, it is proved in [GV02] and [DH03] that the capacity behaves like $\log R$ and a power efficiency that behaves like $\sqrt{R}$ can be obtained. However, these results rely on the assumption that the relay nodes know their local connections, which requires the system to be synchronized at the carrier level. However, for ad hoc networks with a lot of nodes which can also be mobile, this is not a realistic assumption. In our work, a much more practical assumption, that the relay nodes are only coherent at the symbol level, is made. In the relay network, it is assumed that the relay nodes, know only the statistical distribution of the channels. However, we make the assumption that the receiver knows all the fading coefficients $f_i$ and $g_i$, which needs the network to be synchronized at the symbol level. Its knowledge of the channels can be obtained by sending training signals from the relays and the transmitter. The main question is what gains can be obtained? There are two types of gains: improvement in the outage capacity and improvement in the PEP. In this chapter, the focus is on the latter.

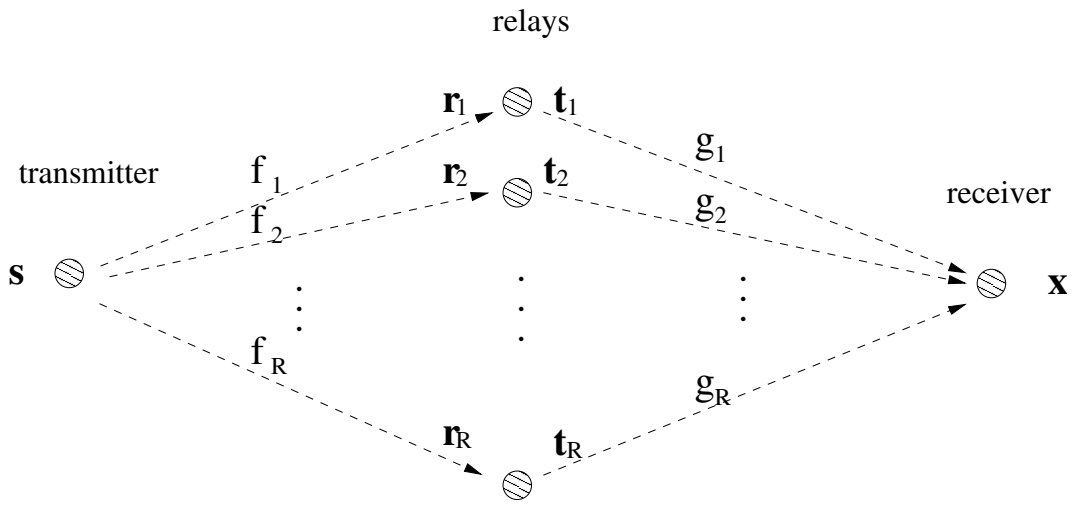


Figure 7.2: Wireless relay network

Assume that the transmitter wants to send the signal $\mathbf{s} = [s_1, \cdots, s_T]^t$ in the codebook $\{\mathbf{s}_1, \cdots, \mathbf{s}_L\}$ to the receiver, where $L$ is the cardinality of the codebook. $\mathbf{s}$

is normalized as

$$\mathrm{E}\,\mathbf{s}^*\mathbf{s} = 1. \tag{7.1}$$

The transmission is accomplished by the following two-step strategy, which is also shown in Figure 7.2.[3] From time 1 to $T$, the transmitter sends signals $\sqrt{P_1 T}s_1, \cdots,$ $\sqrt{P_1 T}s_T$ to each relay. Based on the normalization of $\mathbf{s}$ in (7.1), the average total transmit power of the $T$ transmissions is $P_1 T$. The received signal at the $i$-th relay at time $\tau$ is denoted as $r_{i,\tau}$, which is corrupted by the noise $v_{i,\tau}$. From time $T+1$ to $2T$, the $i$-th relay node transmits $t_{i,1}, \cdots, t_{i,T}$ to the receiver based on its received signals. Denote the received signal at the receiver at time $\tau + T$ by $x_\tau$, and the noise at the receiver at time $\tau + T$ by $w_\tau$. Assume that the noises are complex Gaussian with zero-mean and unit-variance, that is, the distribution of $v_{i,\tau}$ and $w_\tau$ are $\mathcal{CN}(0,1)$.

The following notations are used:

$$\mathbf{v}_i = \begin{bmatrix} v_{i,1} \\ v_{i,2} \\ \vdots \\ v_{i,T} \end{bmatrix}, \quad \mathbf{r}_i = \begin{bmatrix} r_{i,1} \\ r_{i,2} \\ \vdots \\ r_{i,T} \end{bmatrix}, \quad \mathbf{t}_i = \begin{bmatrix} t_{i,1} \\ t_{i,2} \\ \vdots \\ t_{i,T} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_T \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix}.$$

Note that $\mathbf{v}_i$, $\mathbf{r}_i$, $\mathbf{t}_i$, $\mathbf{w}$, and $\mathbf{x}$ are all $T$-dimensional vectors. Clearly

$$\mathbf{r}_i = \sqrt{P_1 T} f_i \mathbf{s} + \mathbf{v}_i \tag{7.2}$$

and

$$\mathbf{x} = \sum_{i=1}^{R} g_i \mathbf{t}_i + \mathbf{w}. \tag{7.3}$$

[3] Although in the figure, all the relay nodes sit on a line in the middle of the transmitter and the receiver, this does not means that they must be in the middle of the two communicating nodes to relay the information. The positions of the relay nodes are arbitrary. For simplicity and clearness of the figure, we draw it this way.

## 7.4 Distributed Space-Time Coding

The key question is what the relay nodes should do. There are two widely used cooperative strategies for the relay nodes. The first one is called amplify-and-forward, in which the relays just amplify their received signals according to power constraints and forward to the receiver. The other is called decode-and-forward, in which the relay nodes do fully decoding and then send their decoded information to the receiver. If the relay nodes know their local connections, beamforming can be done by amplify-and-forward. However, it is obvious that if the relay nodes do not know the channels, amplify-and-forward is not optimal. For decode-and-forward, if the relays can decode the signal correctly, which happens when the transmit power is very high or the transmission rate is very low, the system is equivalent to a multiple-antenna system with $R$ transmit antennas and one receive antenna, and the best diversity $R$ can be obtained. However if some relay nodes decode incorrectly, whether because of bad channel, low transmit power, or high transmission rate, they will forward incorrect signals to the receiver, which will harm the decoding at the receiver greatly. Therefore, for ad hoc networks whose nodes have limited power, decode-and-forward puts a heavy restriction on the transmission rate. Another disadvantage of decode-and-forward is that because of the decoding complexity, it causes both extra power consumption and time delay.

We will instead focus on the diversity achievable without requiring the relay nodes to decode. The strategy we use is called distributed space-time coding, in which simple signal processing is done at relay nodes. No decoding is need at relay nodes, which saves both time and energy, and more importantly, there is no rate constraint on transmissions. As will be seen later, this strategy leads to the optimal diversity, $R$, with asymptotically high transmit power.[4]

[4] A combination of requiring some relay nodes to decode and others to not, may also considered. However, in the interest of space, we shall not do so here.

In our approach, we use the idea of the linear dispersion space-time codes [HH02b] for multi-antenna systems by designing the transmitted signal at every relay as a linear function of its received signal:[5]

$$t_{i,\tau} = \sqrt{\frac{P_2}{P_1+1}} \sum_{t=1}^{T} a_{i,\tau t} r_{i,t} = \sqrt{\frac{P_2}{P_1+1}} [a_{i,\tau 1}, a_{i,\tau 2}, \cdots, a_{i,\tau T}] \mathbf{r}_i,$$

or in other words,

$$\mathbf{t}_i = \sqrt{\frac{P_2}{P_1+1}} A_i \mathbf{r}_i, \tag{7.4}$$

where

$$A_i = \begin{bmatrix} a_{i,11} & a_{i,12} & \cdots & a_{i,1T} \\ a_{i,21} & a_{i,22} & \cdots & a_{i,2T} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i,T1} & a_{i,T2} & \cdots & a_{i,TT} \end{bmatrix}, \quad \text{for } i = 1, 2, \cdots, R.$$

While within the framework of linear dispersion codes, the $T \times T$ matrices $A_i$ can be quite arbitrary (apart from a Frobenius norm constraint). In the network, since the relay nodes have no knowledge of the channels, there is no reason to put more weight on anyone of the relay nodes or any time instant. To have a protocol that is equitable among different users and among different time instants, we shall henceforth assume that $A_i$ are unitary matrices. As we shall presently see, this also simplifies the analysis considerably since it keeps the noises forwarded by the relay nodes to the receiver white.

Now let's discuss the transmit power at each relay node. Because $\operatorname{tr} \mathbf{s}\mathbf{s}^* = 1$, $f_i, v_{i,j}$ are $\mathcal{CN}(0,1)$, and $f_i, s_i, v_{i,j}$ are independent,

$$\mathrm{E}\, \mathbf{r}_i^* \mathbf{r}_i = \mathrm{E}\, (\sqrt{P_1 T} f_i \mathbf{s} + \mathbf{v}_i)^* (\sqrt{P_1 T} f_i \mathbf{s} + \mathbf{v}_i) = \mathrm{E}\, \left( P_1 T |f_i|^2 \mathbf{s}^* \mathbf{s} + \mathbf{v}_i^* \mathbf{v}_i \right) = (P_1 + 1) T.$$

[5]Note that the conjugate of $r_i$ does not appear in (7.4). The case with $\overline{r_i}$ is discussed in Section 7.10.

Therefore the average transmit power at relay node $i$ is

$$\mathrm{E}\,\mathbf{t}_i^*\mathbf{t}_i = \frac{P_2}{P_1+1}\mathrm{E}\,(A_i\mathbf{r}_i)^*(A_i\mathbf{r}_i) = \frac{P_2}{P_1+1}\mathrm{E}\,\mathbf{r}_i^*\mathbf{r}_i = P_2 T,$$

which explains our normalization in (7.4). The expected transmit power for one transmission at each relay is $P_2$.

Let us now focus on the received signal. Clearly from (7.2) and (7.3),

$$\begin{aligned}
\mathbf{x} &= \sqrt{\frac{P_2}{P_1+1}}\sum_{i=1}^{R} g_i A_i \mathbf{r}_i + \mathbf{w} \\
&= \sqrt{\frac{P_2}{P_1+1}}\sum_{i=1}^{R} g_i A_i (\sqrt{P_1 T} f_i \mathbf{s} + \mathbf{v}_i) + \mathbf{w} \\
&= \sqrt{\frac{P_1 P_2 T}{P_1+1}}[A_1\mathbf{s}, \cdots, A_R\mathbf{s}] \begin{bmatrix} f_1 g_1 \\ \vdots \\ f_R g_R \end{bmatrix} + \sqrt{\frac{P_2}{P_1+1}}\sum_{i=1}^{R} g_i A_i \mathbf{v}_i + \mathbf{w}.
\end{aligned}$$

Define

$$S = [A_1\mathbf{s}, A_2\mathbf{s}, \cdots, A_R\mathbf{s}], \quad H = \begin{bmatrix} f_1 g_1 \\ f_2 g_2 \\ \vdots \\ f_R g_R \end{bmatrix}, \quad \text{and} \quad W = \sqrt{\frac{P_2}{P_1+1}}\sum_{i=1}^{R} g_i A_i \mathbf{v}_i + \mathbf{w}.$$

The received signal can therefore be written as

$$\mathbf{x} = \sqrt{\frac{P_1 P_2 T}{P_1+1}} SH + W. \tag{7.5}$$

**Remark:** From equation (7.5), it can be seen that the $T \times R$ matrix $S$ works like the space-time code in the multiple-antenna case. We call it the *distributed space-time code* to emphasize that it has been generated in a distributed way by the relay nodes,

without having access to $\mathbf{s}$. $H$, which is $R \times 1$, is the equivalent channel matrix and $W$, which is $T \times 1$, is the equivalent noise, $W$ is clearly influenced by the choice of the space-time code. Using the unitarity of the $A_i$, it is easy to get the normalization of $S$:

$$\operatorname{tr} S^* S = \sum_{i=1}^{R} \mathbf{s}^* A_i^* A_i \mathbf{s} = \sum_{i=1}^{R} \mathbf{s}^* \mathbf{s} = R.$$

## 7.5 Pairwise Error Probability

Since $A_i$s are unitary and $w_j$, $v_{i,j}$ are independent Gaussian, $W$ is also Gaussian when $g_i$s are known. It is easy to see that $\mathrm{E}\, W = \mathbf{0}_{T1}$ and

$$\begin{aligned}
\operatorname{Var}(W|g_i) &= \mathrm{E}\, WW^* \\
&= \mathrm{E}\left(\sqrt{\frac{P_2}{P_1+1}}\sum_{i=1}^{R} g_i A_i \mathbf{v}_i + \mathbf{w}\right)\left(\sqrt{\frac{P_2}{P_1+1}}\sum_{i=1}^{R} g_i A_i \mathbf{v}_i + \mathbf{w}\right)^* \\
&= \frac{P_2}{P_1+1}\mathrm{E}\sum_{i=1,j=1}^{R,R} g_i \bar{g}_j A_i \mathbf{v}_i \mathbf{v}_j^* A_j^* + I_T \\
&= \frac{P_2}{P_1+1}\sum_{i=1}^{R} |g_i|^2 A_i A_i^* + I_T \\
&= \left(1 + \frac{P_2}{P_1+1}\sum_{i=1}^{R} |g_i|^2\right) I_T.
\end{aligned}$$

Thus, $W$ is both spatially and temporally white. This implies that, when both $f_i$ and $g_i$ are known, $\mathbf{x}|\mathbf{s}_i$ is also Gaussian with the following mean and variance.

$$\mathrm{E}(\mathbf{x}|\mathbf{s}_i) = \sqrt{\frac{P_1 P_2 T}{P_1+1}} S_i H$$

and

$$\mathrm{Var}\,(\mathbf{x}|\mathbf{s}_i) = \mathrm{Var}\,W \quad = \quad \left(1 + \frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2\right) I_T.$$

Thus,

$$\mathrm{P}\,(\mathbf{x}|\mathbf{s}_i) = \frac{1}{\left[2\pi\left(1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2\right)\right]^T} e^{-\frac{(\mathbf{x}-\sqrt{\frac{P_1P_2T}{P_1+1}}S_iH)^*(\mathbf{x}-\sqrt{\frac{P_1P_2T}{P_1+1}}S_iH)}{1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2}}.$$

The ML decoder of the system can be easily calculated to be

$$\arg\max_{\mathbf{s}_i}\mathrm{P}\,(\mathbf{x}|\mathbf{s}_i) = \arg\min_{\mathbf{s}_i}\left\|\mathbf{x} - \sqrt{\frac{P_1P_2T}{P_1+1}}S_iH\right\|_F^2. \tag{7.6}$$

Recall that $S_i = [A_1\mathbf{s}_i, \cdots, A_R\mathbf{s}_i]$, with $\mathbf{s}_i$ in the code book $\{\mathbf{s}_1, \cdots, \mathbf{s}_L\}$. By splitting the real and imaginary parts, the ML decoding in (7.6) is equivalent to

$$\arg\min_{\mathbf{s}_i}\left\|\begin{bmatrix}\mathbf{x}_{Re}\\ \mathbf{x}_{Im}\end{bmatrix} - \sqrt{\frac{P_1P_2T}{P_1+1}}\begin{bmatrix}\left(\sum_{i=1}^{R}f_ig_iA_i\right)_{Re} & -\left(\sum_{i=1}^{R}f_ig_iA_i\right)_{Im}\\ \left(\sum_{i=1}^{R}f_ig_iA_i\right)_{Im} & \left(\sum_{i=1}^{R}f_ig_iA_i\right)_{Re}\end{bmatrix}\begin{bmatrix}\mathbf{s}_{Re}\\ \mathbf{s}_{Im}\end{bmatrix}\right\|_F^2 \tag{7.7}$$

Since (7.7) is equivalent to the decoding of a real linear system, sphere decoding can be used whose complexity is polynomial in the transmission rate and dimension at almost any practical SNR [DAML00, HV02].

**Theorem 7.1 (Chernoff bound on PEP).** *With the ML decoding in (7.6), the PEP, averaged over the channel coefficients, of mistaking* $\mathbf{s}_i$ *by* $\mathbf{s}_j$ *has the following Chernoff bound.*

$$Pe \leq \mathop{\mathrm{E}}_{f_i,g_i} e^{-\frac{P_1P_2T}{4(1+P_1+P_2\sum_{i=1}^{R}|g_i|^2)}H^*(S_i-S_j)^*(S_i-S_j)H}.$$

*By integrating over* $f_i$*s in the above formula, we can get the following inequality on*

*PEP.*

$$Pe \leq \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left[I_R + \frac{P_1 P_2 T}{4\left(1+P_1+P_2\sum_{i=1}^{R}|g_i|^2\right)}(S_i-S_j)^*(S_i-S_j)\mathrm{diag}\left\{|g_1|^2,\cdots,|g_R|^2\right\}\right] \quad (7.8)$$

**Proof:** The PEP of mistaking $S_1$ by $S_i$ has the following Chernoff upper bound [SOSL85].

$$Pe \leq \mathrm{E}\, e^{\lambda(\ln \mathrm{P}(\mathbf{x}|S_i) - \ln \mathrm{P}(\mathbf{x}|S_1))}.$$

Since $S_i$ is transmitted, $\mathbf{x} = \sqrt{\frac{P_1P_2T}{P_1+1}}S_iH + W$. Therefore, from (7.6),

$$\begin{aligned} & \ln \mathrm{P}(\mathbf{x}|S_j) - \ln \mathrm{P}(\mathbf{x}|S_i) \\ = & -\frac{\left[\frac{P_1P_2T}{P_1+1}H^*(S_i-S_j)^*(S_i-S_j)H + \sqrt{\frac{P_1P_2T}{P_1+1}}H^*(S_i-S_j)^*W + \sqrt{\frac{P_1P_2T}{P_1+1}}W^*(S_i-S_j)H\right]}{1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2}. \end{aligned}$$

Thus,

$$\begin{aligned} & Pe \\ \leq & \mathop{\mathrm{E}}_{f_i,g_i,W} e^{-\frac{\lambda}{1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2}\left[\frac{P_2}{P_1+1}P_1TH^*(S_i-S_j)^*(S_i-S_j)H+\sqrt{\frac{P_1P_2T}{P_1+1}}H^*(S_i-S_j)^*W+\sqrt{\frac{P_1P_2T}{P_1+1}}W^*(S_i-S_j)H\right]} \\ = & \mathop{\mathrm{E}}_{f_i,g_i}\int \frac{e^{-\frac{\lambda\left[\frac{P_2}{P_1+1}P_1TH^*(S_i-S_j)^*(S_i-S_j)H+\sqrt{\frac{P_1P_2T}{P_1+1}}H^*(S_i-S_j)^*W+\sqrt{\frac{P_1P_2T}{P_1+1}}W^*(S_i-S_j)H\right]+WW^*}{1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2}}}{\left[2\pi\left(1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2\right)\right]^T}dW \\ = & \mathop{\mathrm{E}}_{f_i,g_i} e^{-\frac{\lambda(1-\lambda)\frac{P_1P_2T}{1+P_1}}{1+\alpha\sum_{i=1}^{R}|g_i|^2}H^*(S_i-S_j)^*(S_i-S_j)H}\int \frac{e^{-\frac{\left(\lambda\sqrt{\frac{P_1P_2T}{P_1+1}}(S_i-S_j)H+W\right)^*\left(\lambda\sqrt{\frac{P_1P_2T}{P_1+1}}(S_i-S_j)H+W\right)}{1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2}}}{\left[2\pi\left(1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2\right)\right]^T}dW \\ = & \mathop{\mathrm{E}}_{f_i,g_i} e^{-\frac{\lambda(1-\lambda)\frac{P_1P_2T}{P_1+1}}{1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2}H^*(S_i-S_j)^*(S_i-S_j)H} \\ = & \mathop{\mathrm{E}}_{f_i,g_i} e^{-\frac{\lambda(1-\lambda)P_1P_2T}{1+P_1+P_2\sum_{i=1}^{R}|g_i|^2}H^*(S_i-S_j)^*(S_i-S_j)H}. \end{aligned}$$

Choose $\lambda = \frac{1}{2}$ which maximizes $\lambda(1-\lambda) = \frac{1}{4}$ and therefore minimizes the right-hand side of the above formula. Thus,

$$Pe \leq \mathop{\mathrm{E}}_{f_i,g_i} e^{-\frac{P_1P_2T}{4(1+P_1+P_2\sum_{i=1}^{R}|g_i|^2)}H^*(S_i-S_j)^*(S_i-S_j)H}. \tag{7.9}$$

This is the first upper bound in Theorem 7.1. To get the second upper bound, the expectation over $f_i$s must be calculated. Notice that

$$H = \begin{bmatrix} f_1g_1 \\ \vdots \\ f_Rg_R \end{bmatrix} = \begin{bmatrix} g_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & g_R \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_R \end{bmatrix}.$$

Denote $\mathrm{diag}\{g_1, \cdots, g_R\}$ as $G$, $[f_1, \cdots, f_R]^t$ as $\mathbf{f}$. (7.9) becomes,

$$\begin{aligned}
Pe &\leq \mathop{\mathrm{E}}_{f_i,g_i} e^{-\frac{P_1P_2T}{4(1+P_1+P_2\sum_{i=1}^{R}|g_i|^2)}\mathbf{f}^*\overline{G}(S_i-S_j)^*(S_i-S_j)G\mathbf{f}} \\
&= \frac{1}{2}\mathop{\mathrm{E}}_{g_i}\int \frac{1}{(2\pi)^R} e^{-\frac{P_1P_2T}{4(1+P_1+P_2\sum_{i=1}^{R}|g_i|^2)}\mathbf{f}^*\overline{G}(S_i-S_j)^*(S_i-S_j)G\mathbf{f}} e^{-\mathbf{f}^*\mathbf{f}} d\mathbf{f} \\
&= \mathop{\mathrm{E}}_{g_i}\int \frac{1}{(2\pi)^R} e^{-\mathbf{f}^*\left(I_R+\frac{P_1P_2T}{4(1+P_1+P_2\sum_{i=1}^{R}|g_i|^2)}\overline{G}(S_i-S_j)^*(S_i-S_j)G\right)\mathbf{f}} d\mathbf{f} \\
&= \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left[I_R + \frac{P_1P_2T}{4\left(1+P_1+P_2\sum_{i=1}^{R}|g_i|^2\right)}\overline{G}(S_i-S_j)^*(S_i-S_j)G\right] \\
&= \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left[I_R + \frac{P_1P_2T}{4\left(1+P_1+P_2\sum_{i=1}^{R}|g_i|^2\right)}(S_i-S_j)^*(S_i-S_j)\mathrm{diag}\{|g_1|^2, \cdots, |g_R|^2\}\right]
\end{aligned}$$

as desired. □

Let's compare (7.8) with the Chernoff bound on the PEP of a multiple-antenna system with $R$ transmit antennas and 1 receive antenna (the receiver knows the channel) [TSC98, HM00]:

$$Pe \leq \det{}^{-1}\left[I_R + \frac{PT}{4R}(S_i-S_j)^*(S_i-S_j)\right].$$

The difference is that now the expectations over the $g_i$ must be calculated. Similar to the multiple-antenna case, the "full diversity" condition can be obtained from (7.8). It is easy to see that if $S_i - S_j$ drops rank, the upper bound in (7.8) increases. Therefore, the Chernoff bound is minimized when $S_i - S_j$ is full-rank, or equivalently, $\det(S_i - S_j)^*(S_i - S_j) \neq 0$ for any $1 \leq i \neq j \leq L$.

## 7.6 Optimum Power Allocation

In this section, the optimum power allocation between the transmit node and relay nodes, that minimize the PEP, is discussed. Because of the expectations over $g_i$, this is easier said than done. Therefore, a heuristic argument is used. Note that $g = \sum_{i=1}^{R} |g_i|^2$ has the gamma distribution [EHP93],

$$p(g) = \frac{g^{R-1}e^{-g}}{(R-1)!},$$

whose mean and variance are both $R$. By the law of large numbers, almost surely $\frac{1}{R}g \to 1$ when $R \to \infty$. It is therefore reasonable to approximate $g$ by its mean, i.e., $\sum_{i=1}^{R} |g_i|^2 \approx R$, especially for large $R$. Therefore, (7.8) becomes

$$Pe \lesssim \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left[I_T + \frac{P_1 P_2 T}{4\,(1 + P_1 + P_2 R)}(S_i - S_j)^*(S_i - S_j)\mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right]. \quad (7.10)$$

It can be seen that the upper bound in (7.10) is minimized when $\frac{P_1 P_2 T}{4(1+P_1+P_2 R)}$ is maximized, which can be easily done.

Assume that the total power consumed in the whole network is $PT$ for transmissions of $T$ symbols. Since the power used at the transmitter and each relay are $P_1$ and $P_2$ respectively for each transmission, $P = P_1 + RP_2$. Therefore,

$$\frac{P_1 P_2 T}{4\,(1 + P_1 + P_2 R)} = \frac{P_1 \frac{P - P_1}{R} T}{4(1 + P_1 + P - P_1)} = \frac{P_1(P - P_1)T}{R(1+P)} \leq \frac{P^2 T}{16R(1+P)}$$

with equality when

$$P_1 = \frac{P}{2} \quad \text{and} \quad P_2 = \frac{P}{2R}. \tag{7.11}$$

Therefore, the optimum power allocation is such that the transmitter uses half the total power and the relay nodes share the other half fairly. So, for large $R$, the relay nodes spend only a very small amount of power to help the transmitter.

With this optimum power allocation, for high total transmit power ($P \gg 1$),

$$\begin{aligned} & \frac{P_1 P_2 T}{4\left(1 + P_1 + P_2 \sum_{i=1}^{R} |g_i|^2\right)} \\ = & \frac{\frac{P}{2}\frac{P}{2R} T}{4\left(1 + \frac{P}{2} + \frac{P}{2R} \sum_{i=1}^{R} |g_i|^2\right)} \\ \approx & \frac{\frac{P}{2}\frac{P}{2R} T}{4\left(\frac{P}{2} + \frac{P}{2R} \sum_{i=1}^{R} |g_i|^2\right)} \\ = & \frac{PT}{8(R + \sum_{i=1}^{R} |g_i|^2)}. \end{aligned}$$

(7.8) becomes

$$Pe \lesssim \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left[I_T + \frac{PT}{8(R + \sum_{i=1}^{R} |g_i|^2)} (S_i - S_j)^* (S_i - S_j) \mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right]. \tag{7.12}$$

## 7.7 Approximate Derivations of the Diversity

As mentioned earlier, to obtain the diversity, the expectation in (7.8) must be calculated. This will be done rigorously in Section 7.8. However, since the calculations are detailed and give little insight, a simple approximate derivation, which leads to the same diversity result, is given here. As discussed in the previous section, when $R$

is large, $\sum_{i=1}^{R} |g_i|^2 \approx R$ with high probability. In this section, this approximation is used to simplify the derivation.

Define

$$M = (S_i - S_j)^*(S_i - S_j). \tag{7.13}$$

To highlight the transmit diversity result, we first upper bound the PEP using the minimum nonzero singular values of $M$, which is denoted as $\sigma_{min}^2$. Therefore, from (7.12),

$$\begin{aligned} Pe &\lesssim \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left[I_T + \frac{PT\sigma_{min}^2}{16R}\mathrm{diag}\left\{I_{\mathrm{rank}\,M}, 0\right\}\mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right] \\ &= \mathop{\mathrm{E}}_{g_i} \prod_{i=1}^{\mathrm{rank}\,M}\left(1 + \frac{PT\sigma_{min}^2}{16R}|g_i|^2\right)^{-1} \\ &= \left[\int_0^\infty \left(1 + \frac{PT\sigma_{min}^2}{16R}x\right)^{-1} e^{-x}dx\right]^{\mathrm{rank}\,M} \\ &= \left(\frac{PT\sigma_{min}^2}{16R}\right)^{-\mathrm{rank}\,M}\left[-e^{-\frac{16R}{PT\sigma_{min}^2}}\mathbf{Ei}\left(-\frac{16R}{PT\sigma_{min}^2}\right)\right]^{\mathrm{rank}\,M}, \end{aligned}$$

where

$$\mathbf{Ei}(\chi) = \int_{-\infty}^{\chi} \frac{e^t}{t}dt, \quad \chi < 0$$

is the exponential integral function [GR00]. For $\chi < 0$,

$$\mathbf{Ei}(\chi) = c + \log(-\chi) + \sum_{k=1}^{\infty} \frac{(-1)^k \chi^k}{k \cdot k!}, \tag{7.14}$$

where $c$ is the Euler constant.[6] For $\log P \gg 1$,

$$e^{-\frac{16R}{PT\sigma_{min}^2}} = 1 + O\left(\frac{1}{P}\right) \approx 1$$

[6]The Euler-Mascheroni constant is defined by $c = \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \log n\right)$. $c$ has the numerical value $0.57721566\cdots$.

and

$$-\mathbf{Ei}\left(-\frac{16R}{PT\sigma_{min}^2}\right) = \log P + O(1) \approx \log P.$$

Therefore,

$$Pe \lesssim \left(\frac{16R}{T\sigma_{min}^2}\right)^{\text{rank}\, M}\left(\frac{\log P}{P}\right)^{\text{rank}\, M} = \left(\frac{16R}{T\sigma_{min}^2}\right)^{\text{rank}\, M} P^{\text{rank}\, M\left(1-\frac{\log\log P}{\log P}\right)} \quad (7.15)$$

When $M$ is full rank, the transmit diversity is $\min\{T,R\}\left(1-\frac{\log\log P}{\log P}\right)$. Therefore, similar to the multiple-antenna case, there is no point in having more relays than the coherence interval according to the diversity. Thus, we will henceforth always assume $T \geq R$. The transmit diversity is therefore $R\left(1-\frac{\log\log P}{\log P}\right)$. (7.15) also shows that the PEP is smaller for bigger coherence interval $T$. A tighter upper bound is given in the following theorem.

**Theorem 7.2.** *Design the transmit signal at the i-th relay node as in (7.4) and use the power allocation in (7.11). For full diversity of the space-time code, assume $T \geq R$. If $P \gg 1$, for any positive $x$, the PEP has the following upper bound*

$$Pe \lesssim \sum_{k=0}^{R}\left(\frac{16R}{PT}\right)^k \sum_{1\leq i_1<\cdots<i_k\leq R} \det{}^{-1}[M]_{i_1,\cdots,i_k}\left(1-e^{-x}\right)^{R-k}\left[-\mathbf{Ei}(-x)\right]^k, \quad (7.16)$$

*where $[M]_{i_1,\cdots,i_k}$ denotes the $k\times k$ matrix composed by choosing the $i_1,\cdots,i_k$-th rows and columns of $M$.*

**Proof:** From (7.10),

$$Pe \lesssim \int_0^\infty \cdots \int_0^\infty \det{}^{-1}\left[I_R + \frac{PT}{16R} M \text{diag}\left\{\lambda_1,\cdots,\lambda_R\right\}\right] e^{-\lambda_1}\cdots e^{-\lambda_R} d\lambda_1 \cdots d\lambda_R,$$

where $\lambda_i$ is defined as $\lambda_i = |g_i|^2$. Therefore, $\lambda_i$ is a random variable with exponential distribution $p_{\lambda_i}(x) = e^{-x}$. We upper bound this by breaking each integral into two parts: the integration from 0 to an arbitrary positive number $x$ and from $x$ to $\infty$,

and then upper bound every one of the resulting $2^R$ terms. That is,

$$\begin{aligned} & Pe \\ \lesssim\ & \left(\int_0^x + \int_x^\infty\right)\cdots\left(\int_0^x + \int_x^\infty\right) \det\left(I_R + \frac{PT}{16R} M \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\}\right)^{-1} \\ & \qquad\qquad e^{-\lambda_1}\cdots e^{-\lambda_R} d\lambda_1 \cdots d\lambda_R \\ =\ & \frac{1}{2}\sum_{k=0}^{R} \sum_{1\le i_1<\cdots<i_k\le R} T_{i_1,\cdots,i_k}, \end{aligned}$$

where

$$T_{i_1,\cdots,i_k} = \underset{\substack{\text{the } i_1,\cdots i_k\text{-th integrals} \\ \text{are from } x \text{ to } \infty, \\ \text{all others are from 0 to } x}}{\int\cdots\int} \det\left(I_R + \frac{PT}{16R} M \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\}\right)^{-1} e^{-\lambda_1}\cdots e^{-\lambda_R} d\lambda_1 \cdots d\lambda_R.$$

Without loss of generality, $T_{1,\cdots,k}$ is calculated.

$$T_{1,\cdots,k} = \underbrace{\int_x^\infty \cdots \int_x^\infty}_{k} \underbrace{\int_0^x \cdots \int_0^x}_{R-k} \det\left(I_R + \frac{PT}{16R} M \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\}\right)^{-1} e^{-\lambda_1}\cdots e^{-\lambda_R} d\lambda_1 \cdots d\lambda_R.$$

Note that since $M > 0$, for any $\lambda_k + 1, \cdots, \lambda_R > 0$,

$$\begin{aligned} & \det\left(I_R + \frac{PT}{16R} M \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\}\right) \\ >\ & \det\left(I_R + \frac{PT}{16R} M \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_k, 0, \cdots, 0\right\}\right) \\ =\ & \det\left(I_R + \frac{PT}{16R} M \begin{bmatrix} \det[M]_{1,\cdots,k}\mathrm{diag}\left\{\lambda_1, \cdots, \lambda_k\right\} & \mathbf{0_{k,R-k}} \\ * & \mathbf{0_{R-k,R-k}} \end{bmatrix}\right) \\ =\ & \det\left(I_k + \frac{PT}{16R}[M]_{1,\cdots,k}\mathrm{diag}\left\{\lambda_1, \cdots, \lambda_k\right\}\right) \\ >\ & \det\left(\frac{PT}{16R}[M]_{1,\cdots,k}\mathrm{diag}\left\{\lambda_1, \cdots, \lambda_k\right\}\right) \\ =\ & \left(\frac{PT}{16R}\right)^k \det[M]_{1,\cdots,k}\lambda_1\cdots\lambda_k, \end{aligned}$$

where $[M]_{i_1,\cdots,i_k}$ is defined in Theorem 7.2. Therefore,

$$\begin{aligned} T_{1,\cdots,k} &< \left(\frac{16R}{PT}\right)^k \det{}^{-1}[M]_{1,\cdots,k} \int_0^x \cdots \int_0^x e^{-\lambda_{k+1}} \cdots e^{-\lambda_R} d\lambda_{k+1} \cdots d\lambda_R \\ &\qquad \int_x^\infty \cdots \int_x^\infty \frac{e^{-\lambda_1}}{\lambda_1} \cdots \frac{e^{-\lambda_k}}{\lambda_k} d\lambda_1 \cdots d\lambda_k \\ &= \left(\frac{16R}{PT}\right)^k \det{}^{-1}[M]_{1,\cdots,k} \left(\int_0^x e^{-\lambda} d\lambda\right)^{R-k} \left(\int_x^\infty \frac{e^{-\lambda}}{\lambda} d\lambda\right)^k \\ &= \left(\frac{16R}{PT}\right)^k \det{}^{-1}[M]_{1,\cdots,k} \left(1-e^{-x}\right)^{R-k} \left[\mathbf{Ei}(-x)\right]^k. \end{aligned}$$

In general,

$$T_{i_1,\cdots,i_k} < \left(\frac{16R}{PT}\right)^k \det{}^{-1}[M]_{1,\cdots,k} \left(1-e^{-x}\right)^{R-k} \left[\mathbf{Ei}(-x)\right]^k.$$

The upper bound in (7.16) is obtained. □

It is easy to see that for any $1 \le i_1 < \cdots < i_k \le R$, $[M]_{i_1,\cdots,i_k}$ is a positive definite matrix since $M$ is positive definite. Therefore, all terms in (7.16) are positive.

**Corollary 7.1.** *If* $\log P \gg 1$,

$$Pe \lesssim \frac{1}{P^R} \sum_{k=0}^{R} \left(\frac{16R}{T}\right)^k \sum_{1 \le i_1 < \cdots < i_k \le R} \det{}^{-1}[M]_{i_1,\cdots,i_k} \log^k P. \tag{7.17}$$

**Proof:** Set $x = \frac{1}{P}$.[7] Therefore,

$$\left(1-e^{-x}\right)^{R-k} = \left(1-e^{-\frac{1}{P}}\right)^{R-k} = \left(\frac{1}{P} + o\left(\frac{1}{P}\right)\right)^{R-k} = \frac{1}{P^{R-k}} + o\left(\frac{1}{P^{R-k}}\right).$$

[7] Actually, this is not the optimum choice based on the transmit diversity. The transmit diversity can be improved slightly by choosing a optimum $x$. However, the coding gain of that case is smaller than the coding gain in (7.17). The details will be discussed in Section 7.9.

From (7.14),

$$-\mathbf{Ei}(-x) = \log P + O(1).$$

Therefore, (7.17) is obtained from (7.16) by omitting the higher order terms of $\frac{1}{P}$. □

The same as the PEP Chernoff upper bound of multiple-antenna systems with $R$ transmit antennas and one receive antenna at high SNR, which is

$$Pe \leq \frac{1}{P^R} \det{}^{-1}(S_i - S_j)^*(S_i - S_j) \left(\frac{4R}{T}\right)^R,$$

the factor $\frac{1}{P^R}$ is also obtained in the network case. However, instead of a constant that is independent of $P$, the coefficient of the factor in (7.17) is a polynomial in $\log P$, which actually changes the diversity result.

To get the diversity, we should look at the the term with the highest order of $P$ in (7.17), which is the $k = R$ term: $\det{}^{-1} M \left(\frac{16R}{T}\right)^R \frac{\log^R P}{P^R}$. By simple rewriting, it is equivalent to

$$\det{}^{-1} M \left(\frac{16R}{T}\right)^R P^{-R\left(1-\frac{\log\log P}{\log P}\right)}. \tag{7.18}$$

Therefore, as in (7.15), transmit diversity of the distributed space-time code is, again, $R\left(1 - \frac{\log\log P}{\log P}\right)$, which is linear in the number of relays. When $P$ is very large ($P \gg \log P$), $\frac{\log\log P}{\log P} \ll 1$, and a transmit diversity about $R$ is obtained which is the same as the transmit diversity of a multiple-antenna system with $R$ transmit antennas and one receive antenna. That is, the system works as if the $R$ relay nodes can fully cooperate and have full knowledge of the transmitted signal as in the multiple-antenna case. However, for any general average total transmit power, the transmit diversity depends on the average total transmit power $P$.

## 7.8 Rigorous Derivation of the Diversity

In the previous section, we use the approximation $\sum_{i=1}^{R} |g_i|^2 \approx R$. In this section, a rigorous derivation of the Chernoff upper bound on the PEP is given. The same transmit diversity is obtained but the coding gain becomes more complicated. Here is the main result.

**Theorem 7.3.** *Design the transmit signal at the i-th relay node as in (7.4) and use the power allocation in (7.11). For full diversity of the space-time code, assume* $T \geq R$. *If* $\log P \gg 1$, *the PEP has the following Chernoff bound.*

$$Pe \lesssim \frac{1}{P^R} \sum_{k=0}^{R} \left(\frac{8}{T}\right)^k \sum_{1 \leq i_1 < \cdots < i_k \leq R} \det{}^{-1}[M]_{i_1,\cdots,i_k} \sum_{l=0}^{k} B_R(k-l,k) \log^l P, \quad (7.19)$$

*where*

$$B_R(j,k) = \begin{pmatrix} k \\ j \end{pmatrix} \sum_{i_1=1}^{k} \sum_{i_2=1}^{k-i_1} \cdots \sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}} \begin{pmatrix} k \\ i_1 \end{pmatrix} \cdots \begin{pmatrix} k-i_1-\cdots-i_{j-1} \\ i_j \end{pmatrix} (i_1-1)! \cdots (i_j-1)! R^{k-i_1-\cdots-i_j}. \quad (7.20)$$

**Proof:** Before proving the theorem, we first give a lemma that is needed.

**Lemma 7.1.** *If* $A$ *is a constant,*

$$\int_x^\infty \cdots \int_x^\infty \left(A + \sum_{i=1}^{k} \lambda_i\right)^k \frac{e^{-\lambda_1} \cdots e^{-\lambda_k}}{\lambda_1 \cdots \lambda_k} d\lambda_1 \cdots d\lambda_k = \sum_{j=0}^{k} B_{A,x}(j,k) \left[-\mathbf{Ei}(-x)\right]^{k-j}, \quad (7.21)$$

*where*

$$B_{A,x}(j,k) = \begin{pmatrix} k \\ j \end{pmatrix} \sum_{i_1=1}^{k} \sum_{i_2=1}^{k-i_1} \cdots \sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}} \begin{pmatrix} k \\ i_1 \end{pmatrix} \cdots \begin{pmatrix} k-i_1-\cdots-i_{j-1} \\ i_j \end{pmatrix}$$
$$\Gamma(i_1,x)\cdots\Gamma(i_j,x)A^{k-i_1-\cdots-i_j}$$

*and*

$$\Gamma(i,x) = \int_x^\infty e^{-t}t^{i-1}dt$$

*is the incomplete Gamma function [GR00].*

**Proof:** See Section 7.14.1. □

Now we prove Theorem 7.3. From (7.12), we need to upper bound

$$\int_0^\infty \cdots \int_0^\infty \det{}^{-1}\left[I_R + \frac{PT}{8\left(R+\sum_{i=1}^R \lambda_i\right)} M \mathrm{diag}\left\{\lambda_1,\cdots,\lambda_R\right\}\right] e^{-\lambda_1}\cdots e^{-\lambda_R} d\lambda_1\cdots d\lambda_R.$$

We use the same method as in the previous section: breaking every integral into two parts. Therefore,

$$Pe \;\leq\; \frac{1}{2}\sum_{k=0}^{R} \sum_{1\leq i_1<\cdots<i_k\leq R} T'_{i_1,\cdots,i_k},$$

while

$$T'_{i_1,\cdots,i_k}$$
$$= \int\cdots\int_{\substack{\text{the } i_1,\cdots i_k\text{-th integrals} \\ \text{are from } x \text{ to } \infty, \\ \text{all others are from 0 to } x}} \det{}^{-1}\left[I_R + \frac{PT}{8\left(R+\sum_{i=1}^R \lambda_i\right)} M \mathrm{diag}\left\{\lambda_1,\cdots,\lambda_R\right\}\right]$$
$$e^{-\lambda_1}\cdots e^{-\lambda_R} d\lambda_1\cdots d\lambda_R.$$

Without loss of generality, $T'_{1,\cdots,k}$ is calculated.

$$
\begin{aligned}
& T'_{1,\cdots,k} \\
&= \underbrace{\int_x^\infty \cdots \int_x^\infty}_{k} \underbrace{\int_0^x \cdots \int_0^x}_{R-k} \det{}^{-1}\left[ I_R + \frac{PT}{8\left(R+\sum_{i=1}^R \lambda_i\right)} M \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\} \right] \\
&\qquad\qquad e^{-\lambda_1} \cdots e^{-\lambda_R} d\lambda_1 \cdots d\lambda_R.
\end{aligned}
$$

For any $0 < \lambda_{k+1} < x, \cdots, 0 < \lambda_R < x$,

$$
\begin{aligned}
& \det\left[ I_R + \frac{PT}{8\left(R+\sum_{i=1}^R \lambda_i\right)} M \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\} \right] \\
> \; & \det\left[ I_R + \frac{PT}{8\left(R+(R-k)x+\sum_{i=1}^k \lambda_i\right)} M \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_k, 0, \cdots, 0\right\} \right] \\
> \; & \det\left[ \frac{PT}{8\left(R+(R-k)x+\sum_{i=1}^k \lambda_i\right)} [M]_{1,\cdots,k} \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_k\right\} \right] \\
= \; & \left[ \frac{PT}{8\left(R+(R-k)x+\sum_{i=1}^k \lambda_i\right)} \right]^k \det [M]_{1,\cdots,k} \lambda_1 \cdots \lambda_k.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
T'_{1,\cdots,k} \; < \; & \left(\frac{8}{PT}\right)^k \det{}^{-1}[M]_{1,\cdots,k} \int_0^x \cdots \int_0^x e^{-\lambda_{k+1}} \cdots e^{-\lambda_R} d\lambda_{k+1} \cdots d\lambda_R \\
& \int_x^\infty \cdots \int_x^\infty \left[ R + (R-k)x + \sum_{i=1}^k \lambda_i \right]^k \frac{e^{-\lambda_1} \cdots e^{-\lambda_k}}{\lambda_1 \cdots \lambda_k} d\lambda_1 \cdots d\lambda_k.
\end{aligned}
$$

Using Lemma 7.1,

$$
T'_{1,\cdots,k} < \left(\frac{8}{PT}\right)^k \det{}^{-1}[M]_{1,\cdots,k} \left(1-e^{-x}\right)^{R-k} \sum_{j=0}^k B_{R+(R-k)x,x}(j,k) \left[-\mathbf{Ei}(-x)\right]^{k-j}.
$$

Choose $x = \frac{1}{P}$. Similarly, for large $P$,

$$\left[R + (R-k)\frac{1}{P}\right]^k \approx R^k, \qquad -\mathbf{Ei}\left(-\frac{1}{P}\right) \approx \log P,$$
$$1 - e^{-\frac{1}{P}} \approx \frac{1}{P}, \qquad \Gamma(i,x) \approx (i-1)!.$$

Therefore,

$$\begin{aligned} T'_{1,\cdots,k} &< \left(\frac{8}{PT}\right)^k \det{}^{-1}[M]_{1,\cdots,k} \frac{1}{P^{R-k}} \sum_{j=0}^{k} B_{R,x}(j,k) \log^{k-j} P \\ &= \left(\frac{8}{PT}\right)^k \det{}^{-1}[M]_{1,\cdots,k} \frac{1}{P^{R-k}} \sum_{l=0}^{k} B_{R,x}(k-l,k) \log^{l} P. \end{aligned}$$

In general,

$$T'_{i_1,\cdots,i_k} < \frac{1}{P^R} \left(\frac{8}{T}\right)^k \det{}^{-1}[M]_{i_1,\cdots,i_k} \sum_{l=0}^{k} B_{R,x}(k-l,k) \log^{l} P.$$

Thus, (7.19) is obtained. □

**Corollary 7.2.** *If* $R \gg 1$,

$$Pe \lesssim \frac{1}{P^R} \sum_{k=0}^{R} \left(\frac{8R}{T}\right)^k \sum_{1 \le i_1 < \cdots < i_k \le R} \det{}^{-1}[M]_{i_1,\cdots,i_k} \log^k P. \tag{7.22}$$

**Proof:** When $R \gg 1$, $B_R(0,k) >> B_R(l,k)$ for all $l > 0$ since $B_R(0,k) = R^k$ is the term with the highest order of $R$. Therefore, (7.22) is obtained from (7.19). □

**Remarks:**

1. The $k = l = R$ term,

$$\det{}^{-1} M \left(\frac{8R \log P}{TP}\right)^R, \tag{7.23}$$

in (7.19) has the highest order of $P$. By simple rewriting, it is equivalent to

$$\det{}^{-1} M \left(\frac{8R}{T}\right)^R P^{-R\left(1-\frac{\log\log P}{\log P}\right)}, \tag{7.24}$$

which is the same as (7.18) except for a coefficient of $2^R$. Therefore, the same transmit diversity, $R\left(1-\frac{\log\log P}{\log P}\right)$, is obtained.

2. In a multiple-antenna system with $R$ transmit antennas and one receive antenna, at high transmit power (or SNR), the PEP has the following upper bound, (which is given in (2.4) in Section 2.3,)

$$\det{}^{-1} M \left(\frac{4R}{PT}\right)^R.$$

Comparing this with the highest order term given in (7.23), we can see the relay network has a performance that is

$$(3 + 10\log_{10}\log P)\, dB \tag{7.25}$$

worse. This analysis is also verified by simulations in Section 7.12.

3. Corollary 7.2 also gives the coding gain for networks with large number of relay nodes. When $P$ is very large ($\log P \gg 1$), the dominant term in (7.22) is (7.24). The coding gain is therefore $\det{}^{-1} M$, which is the same as the multiple-antenna case. When $P$ is not very large, the second term in (7.22),

$$\left(\frac{8R}{T}\right)^{R-1} \sum_{i=1}^{R} \det{}^{-1}[M]_{1,\cdots,i-1,i+1,\cdots,R} \frac{\log^{R-1} P}{P^R},$$

cannot be ignored and even the $k = 3, 4, \cdots$ terms have non-neglectable contributions. Therefore, we want not only $\det M$ to be large but also $\det[M]_{i_1,\cdots,i_k}$

to be large for all $0 \leq k \leq R, 1 \leq i_1 < \cdots < i_k \leq R$. Note that

$$[M]_{i_1,\cdots,i_k} = ([S_i]_{i_1,\cdots,i_k} - [S_j]_{i_1,\cdots,i_k})^*([S_i]_{i_1,\cdots,i_k} - [S_j]_{i_1,\cdots,i_k}),$$

where $[S_i]_{i_1,\cdots,i_k} = [A_{i_1}\mathbf{s}_i \ \cdots \ A_{i_k}\mathbf{s}_i]$ is the distributed space-time code when only the $i_1, \cdots, i_k$-th relay nodes are working. To have a good performance for not very large transmit power, Corollary 7.2 indicates that the distributed space-time code should have the property that it is "scale-free" in the sense that it is still a good distributed space-time code when some of the relays are not working. In general, for networks with any number of relay nodes, the same conclusion can be obtained from (7.19).

4. Now we look at the low average total transmit power case, that is the $P \ll 1$ case. With the same approximation $\sum_{i=1}^{R} |g_i|^2 \approx R$, using the power allocation given in (7.11),

$$\frac{P_1 P_2 T}{4\left(1 + P_1 + P_2 \sum_{i=1}^{R} |g_i|^2\right)} \approx \frac{\frac{P}{2}\frac{P}{2R}T}{4\,(1+P)} = \frac{P^2 T}{16R}.$$

Therefore, (7.8) becomes

$$\begin{aligned}
Pe &\lesssim \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left(I_R + \frac{P^2T}{16R} M \mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right) \\
&= \mathop{\mathrm{E}}_{g_i}\left[1 + \frac{P^2T}{16R}\mathrm{tr}\,\left(M\mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right) + o(P^2)\right]^{-1} \\
&= \mathop{\mathrm{E}}_{g_i}\left(1 + \frac{P^2T}{16R}\sum_{i=1}^{R} m_{ii}|g_i|^2 + o(P^2)\right)^{-1} \\
&= \mathop{\mathrm{E}}_{g_i}\left(1 - \frac{P^2T}{16R}\sum_{i=1}^{R} m_{ii}|g_i|^2\right) + o(P^2) \\
&= \left(1 - \frac{P^2T}{16R}\sum_{i=1}^{R} m_{ii}\right) + o(P^2)
\end{aligned}$$

$$= \left(1 - \frac{P^2 T}{16R} \mathrm{tr}\, M\right) + o(P^2),$$

where $m_{ii}$ is the $(i,i)$ entry of $M$. Therefore, the same as in the multiple-antenna case, the coding gain at low total transmit power is $\mathrm{tr}\, M$. The design criterion is to maximize $\mathrm{tr}\, M$.

5. Corollary 7.2 also shows that the results obtained by the rigorous derivation in this section is consistent with the approximate derivation in the previous section except for a coefficient $2^k$. Actually the upper bound in (7.22) is tighter than the one in (7.17). This is reasonable since in (7.22) all the terms except the one with the highest order of $R$ are omitted, however in the derivation of (7.17), we approximate $\sum_{i=1}^{R} |g_i|^2$ by its expected value $R$.

## 7.9 Improvement in the Diversity

In Corollary 7.1 and Theorem 7.3, $x = \frac{1}{P}$ is used, which actually is not the optimal choice according to transmit diversity. The transmit diversity can be improved slightly by choosing the positive number $x$ optimally.

**Theorem 7.4.** *The best transmit diversity can be obtained using the distributed space-time codes is $\alpha_0 R$, where $\alpha_0$ is the solution of*

$$\alpha + \frac{\log \alpha}{\log P} = 1 - \frac{\log \log P}{\log P}. \tag{7.26}$$

*For $P \gg \log P$, the PEP has the following upper bound,*

$$Pe \lesssim \sum_{k=0}^{R} \left(\frac{8}{T}\right)^k \sum_{1 \le i_1 < \cdots < i_k \le R} \det{}^{-1}[M]_{i_1,\cdots,i_k} \sum_{l=0}^{k} B_R(k-l,k) P^{-[\alpha_0 R + (1-\alpha_0)(k-l)]}. \tag{7.27}$$

*If* $R \gg 1$,

$$Pe \lesssim \left[\sum_{k=0}^{R}\left(\frac{8R}{T}\right)^{k}\sum_{1\le i_1<\cdots<i_k\le R}\det{}^{-1}[M]_{i_1,\cdots,i_k}\right]P^{-\alpha_0 R}. \tag{7.28}$$

**Proof:** According to the proof of Theorem 7.3,

$$\begin{aligned}&Pe\\ &\le \sum_{k=0}^{R}\sum_{1\le i_1<\cdots<i_k\le R}\det{}^{-1}[M]_{i_1,\cdots,i_k}\left(\tfrac{8}{PT}\right)^{k}(1-e^{x})^{R-k}\textstyle\sum_{l=0}^{k}B_{R+(R-k)x}(k-l,k)[-\mathbf{Ei}(-x)]^{l}\end{aligned} \tag{7.29}$$

Set $x=\frac{1}{P^{\alpha}}$ with $\alpha>0$. Therefore,

$$\left(1-e^{-x}\right)^{R-k}=\left(1-e^{-\frac{1}{P^{\alpha}}}\right)^{R-k}=\left(\frac{1}{P^{\alpha}}+o\left(\frac{1}{P^{\alpha}}\right)\right)^{R-k}=\frac{1}{P^{\alpha(R-k)}}+o\left(\frac{1}{P^{\alpha(R-k)}}\right).$$

and

$$\left(R+(R-k)\frac{1}{P^{\alpha}}\right)^{k}=R^{k}+O\left(\frac{1}{P^{\alpha}}\right).$$

From (7.14),

$$-\mathbf{Ei}(-x)=\log P^{\alpha}+O(1)=\alpha\log P+O(1).$$

To omit high order terms of $P$, we need $P^{\alpha}\gg 1$ and $\log P^{\alpha}\gg 1$. Combining the two, $\log P\gg\frac{1}{\alpha}$ is needed.

Assume $\log P\gg\frac{1}{\alpha}$. (7.29) becomes

$$\begin{aligned}Pe &\lesssim \sum_{k=0}^{R}\sum_{1\le i_1<\cdots<i_k\le R}\det{}^{-1}[M]_{i_1,\cdots,i_k}\left(\frac{8}{PT}\right)^{k}\sum_{l=0}^{k}B_R(k-l,k)\frac{\alpha^{l}\log^{l}P}{P^{\alpha(R-k)}}\\ &= \sum_{k=0}^{R}\sum_{1\le i_1<\cdots<i_k\le R}\det{}^{-1}[M]_{i_1,\cdots,i_k}\left(\frac{8}{T}\right)^{k}\sum_{l=0}^{k}B_R(k-l,k)P^{-[k+\alpha(R-k)]}\alpha^{l}\log^{l}P.\end{aligned}$$

Note that $\alpha^l = P^{l\frac{\log\alpha}{\log P}}$ and $\log^l P = P^{l\frac{\log\log P}{\log P}}$. Therefore,

$$\lesssim \sum_{k=0}^{R} \sum_{1\le i_1<\cdots<i_k\le R} \det{}^{-1}[M]_{i_1,\cdots,i_k} \left(\frac{8}{PT}\right)^k \sum_{l=0}^{k} B_R(k-l,k) P^{-\left[\alpha R+(1-\alpha)k-l\frac{\log\alpha}{\log P}-l\frac{\log\log P}{\log P}\right]}.$$

When $\alpha = \alpha_0$, (7.27) is obtained. As in the proof of Corollary 7.2, for $R \gg 1$, we only keep the terms with the highest order of $R$, that is, the $l = k$ terms. (7.28) is then obtained from (7.27).

What left to prove is the optimality of the choice $x = \frac{1}{P^{\alpha_0}}$. For any $k$, let's look at the term with the highest order of $P$, which is the $l = k$ term. Define

$$\beta(\alpha,k) = \alpha R + (1-\alpha)k - k\frac{\log\alpha}{\log P} - k\frac{\log\log P}{\log P},$$

which is the negative of the highest order of $P$. To obtain the best transmit diversity, $\alpha$ should be chosen to maximize $\min_{k\in[1,R]} \beta(\alpha,k)$. Note that

$$\frac{\partial\beta}{\partial k} = (1-\alpha) - \frac{\log\alpha}{\log P} - \frac{\log\log P}{\log P}.$$

From the definition of $\alpha_0$, we have

$$\frac{\partial\beta}{\partial k}\Big|_{\alpha=\alpha_0} = 0.$$

Also note

$$\frac{\partial^2\beta}{\partial\alpha\partial k} = -1 - \frac{1}{\alpha\log P} < 0.$$

Therefore,

$$\begin{cases} \frac{\partial\beta}{\partial k} > 0 \text{ if } \alpha < \alpha_0 \\ \frac{\partial\beta}{\partial k} < 0 \text{ if } \alpha > \alpha_0 \end{cases}.$$

Thus,

$$\min_{k\in[1,R]} \beta(\alpha,k) = \begin{cases} \beta(\alpha,0) = \alpha R & \text{if } \alpha \le \alpha_0 \\ \beta(\alpha,R) = R\left(1 - \frac{\log\alpha}{\log P} - \frac{\log\log P}{\log P}\right) & \text{if } \alpha \ge \alpha_0 \end{cases}.$$

It is easy to see that for $\alpha \le \alpha_0$, $\max_{\alpha\le\alpha_0} \min_{k\in[1,R]} \beta(\alpha,k) = \max_{\alpha\le\alpha_0} \alpha R = \alpha_0 R$, which is obtained when $\alpha = \alpha_0$. Now let's look at the case of $\alpha \ge \alpha_0$. Note that

$$\frac{d\min_{k\in[1,R]}\beta(\alpha,k)}{d\alpha} = \frac{\beta(\alpha,R)}{d\alpha} = -\frac{R}{\alpha\log P} < 0.$$

Therefore,

$$\max_{\alpha\ge\alpha_0}\min_{k\in[1,R]} \beta(\alpha,k) = \beta(\alpha_0,R) = R\left(1 - \frac{\log\alpha_0}{\log P} - \frac{\log\log P}{\log P}\right) = \alpha_0 R,$$

which is obtained when $\alpha = \alpha_0$. Therefore, $\alpha_0$ is the optimum and

$$\alpha_0 R + (1-\alpha_0)k - l\frac{\log\alpha_0}{\log P} - l\frac{\log\log P}{\log P} = \alpha_0 R + (1-\alpha_0)(k-l).$$

Still we need to check the condition $\log P \gg \frac{1}{\alpha_0}$. Define $\gamma(\alpha) = \alpha + \frac{\log\alpha}{\log P}$. Then $\frac{d\gamma(\alpha)}{d\alpha} = 1 + \frac{1}{\alpha\log P} > 0$. Since

$$\gamma(1) = 1 > 1 - \frac{\log\log P}{\log P}$$

and

$$\gamma\left(1 - \frac{\log\log P}{\log P}\right) = 1 - \frac{\log\log P}{\log P} + \frac{\log\left(1 - \frac{\log\log P}{\log P}\right)}{\log P} < 1 - \frac{\log\log P}{\log P},$$

therefore,

$$1 - \frac{\log\log P}{\log P} < \alpha_0 < 1. \tag{7.30}$$

Therefore, $\alpha_0 \log P > \log P - \log\log P$. If $\log P \gg \log\log P$, $\log P \gg 1$ is true and thus $\log P - \log\log P \gg 1$. The condition is satisfied. □

There is no closed form for the solution of equation (7.26). The following theorem gives a region of $\alpha_0$ and also gives some ideas about how much improvement in transmit diversity is obatined.

**Theorem 7.5.** *For $P > e$,*

$$1 - \frac{\log\log P}{\log P} < \alpha_0 < 1 - \frac{\log\log P}{\log P} + \frac{\log\log P}{\log P(\log P - \log\log P)}.$$

**Proof:** From the proof of Theorem 7.4, we know that $1 - \frac{\log\log P}{\log P} < \alpha_0$. We only need to prove the other part. Let

$$\alpha_1 = 1 - \frac{\log\log P}{\log P} + \frac{\log\log P}{(\log P - \log\log P)\log P}.$$

Since as in the proof of Theorem 7.4, $\gamma'(\alpha) > 0$. We just need to prove that $\gamma(\alpha_1) - 1 - \frac{\log\log P}{\log P} > 0$.

Let's first prove

$$\log\alpha_1 > -\frac{\log P}{\log P - \log\log P}\left(\frac{\log\log P}{\log P} - \frac{\log\log P}{(\log P - \log\log P)\log P}\right).$$

Define $g(x) = \log(1-x) + cx$. Since $g'(x) = c - \frac{1}{1-x}$, $g'(x) > 0$ if $x < 1 - \frac{1}{c}$. Note that $g(0) = 0$, therefore, $g(x) > 0$ or equivalently $\log(1-x) > -cx$ when $0 < x < 1 - \frac{1}{c}$. Let $x_0 = \frac{\log\log P}{\log P} - \frac{\log\log P}{\log P(\log P - \log\log P)}$ and $c_0 = \frac{\log P}{\log P - \log\log P}$. $1 - \frac{1}{c_0} = \frac{\log\log P}{\log P} > x_0$ for $P > e$.

It is also easy to check that $x_0 > 0$ for $P > e$. Therefore, $\log \alpha_1 = \log(1-x_0) > -c_0 x_0$ and

$$\begin{aligned}
& \gamma(\alpha_1) - 1 - \frac{\log\log P}{\log P} \\
= & \frac{\log\log P}{\log P(\log P - \log\log P)} + \frac{1}{\log P}\log\alpha_1 \\
> & \frac{\log\log P}{\log P(\log P - \log\log P)} - \frac{1}{\log P}\frac{\log P}{\log P - \log\log P}\left(\frac{\log\log P}{\log P} - \frac{\log\log P}{\log P(\log P - \log\log P)}\right) \\
= & \frac{\log\log P}{\log P(\log P - \log\log P)^2} \\
> & 0.
\end{aligned}$$

□

Theorem 7.5 indicates that the PEP Chernoff bound of the distributed space-time codes decreases faster than

$$\sum_{k=0}^{R}\left(\frac{8R}{T}\right)^k \sum_{1\le i_1<\cdots<i_k\le R} \det{}^{-1}[M]_{i_1,\cdots,i_k}\left(\frac{\log P}{P}\right)^R$$

and slower than

$$\sum_{k=0}^{R}\left(\frac{8R}{T}\right)^k \sum_{1\le i_1<\cdots<i_k\le R} \det{}^{-1}[M]_{i_1,\cdots,i_k}\left(\frac{(\log P)^{1-\frac{1}{\log P-\log\log P}}}{P}\right)^R.$$

When $P$ is large ($\log P \gg 1$), $1 - \frac{\log\log P}{\log P}$ is a very accurate approximation of $\alpha_0$. The improvement in transmit diversity is small.

Now let's comapre the new upper bound in (7.28) with the one in (7.22). As discussed above, a slightly better transmit diversity is obtained. However, the coding gain in (7.28) smaller. The coding gain of (7.28) is

$$\left[\sum_{k=0}^{R}\left(\frac{8R}{T}\right)^k \sum_{1\le i_1<\cdots<i_k\le R} \det{}^{-1}[M]_{i_1,\cdots,i_k}\right]^{-1},$$

and the coding gain of (7.22) for very high SNR $\log P \gg 1$ is $\det M$. To compare the two, we assume that the singular values of $M$ take their maximum value, $\sqrt{2}$, and $R = T$. Therefore the coding gain of (7.28) is $\left[\sum_{k=0}^{R} \begin{pmatrix} R \\ k \end{pmatrix} 4^k\right]^{-1} = 5^{-R}$. The coding gain of (7.22) is $4^{-R}$. The upper bound in (7.22) is 0.97dB better according to coding gain.

Therefore, when $P$ is extremely large, the new upper bound is tighter than the previous one since it has a lager diversity. Otherwise, the previous bound is tighter since it has a larger coding gain.

## 7.10 A More General Case

In this section, a more general type of distributed linear dispersion space-time codes [HH02b] is discussed. The transmitted signal at the $i$-th relay node is designed as

$$\mathbf{t}_i = \sqrt{\frac{P_2}{P_1 + 1}}(A_i \mathbf{r}_i + B_i \overline{\mathbf{r}_i}) \quad i = 1, 2, \cdots, R, \tag{7.31}$$

where $A_i$ and $B_i$ are $T \times T$ real matrices. Similar to before, we assume that $A_i + B_i$ and $A_i - B_i$ are orthogonal, which is equivalent to

$$\begin{cases} A_i A_i^T + B_i B_i^T = I_T \\ A_i B_i^T = -(A_i B_i^T)^T \end{cases}.$$

By separating the real and imaginary parts, (7.31) can be written equivalently as

$$\begin{bmatrix} \mathbf{t}_{i,Re} \\ \mathbf{t}_{i,Im} \end{bmatrix} = \sqrt{\frac{P_2}{P_1 + 1}} \begin{bmatrix} A_i + B_i & \mathbf{0} \\ \mathbf{0} & A_i - B_i \end{bmatrix} \begin{bmatrix} \mathbf{r}_{i,Re} \\ \mathbf{r}_{i,Im} \end{bmatrix}. \tag{7.32}$$

The expected total transmit power at the $i$-th relay can therefore be calculated to be $P_2T$.

Now let's look at the received signal. Similar to the rewriting of (7.31), (7.2) can be equivalently written as

$$\begin{bmatrix} \mathbf{r}_{i,Re} \\ \mathbf{r}_{i,Im} \end{bmatrix} = \sqrt{P_1 T} \begin{bmatrix} f_{i,Re} I_T & -f_{i,Im} I_T \\ f_{i,Im} I_T & f_{i,Re} I_T \end{bmatrix} \begin{bmatrix} \mathbf{s}_{Re} \\ \mathbf{s}_{Im} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{i,Re} \\ \mathbf{v}_{i,Im} \end{bmatrix}.$$

Therefore,

$$\begin{bmatrix} \mathbf{t}_{i,Re} \\ \mathbf{t}_{i,Im} \end{bmatrix} = \sqrt{\frac{P_1 P_2 T}{P_1 + 1}} \begin{bmatrix} A_i + B_i & \mathbf{0} \\ \mathbf{0} & A_i - B_i \end{bmatrix} \begin{bmatrix} f_{i,Re} I_T & -f_{i,Im} I_T \\ f_{i,Im} I_T & f_{i,Re} I_T \end{bmatrix} \begin{bmatrix} \mathbf{s}_{Re} \\ \mathbf{s}_{Im} \end{bmatrix} + \sqrt{\frac{P_2}{P_1 + 1}} \begin{bmatrix} \mathbf{v}_{i,Re} \\ \mathbf{v}_{i,Im} \end{bmatrix}$$

For the $T \times 1$ complex vector $\mathbf{x}$, define the $2T \times 1$ real vector $\hat{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_{Re} \\ \mathbf{x}_{Im} \end{bmatrix}$. Further define the $2T \times 2T$ real matrix

$$\mathcal{H} = \sum_{i=1}^{R} \begin{bmatrix} g_{i,Re} I_T & -g_{i,Im} I_T \\ g_{i,Im} I_T & g_{i,Re} I_T \end{bmatrix} \begin{bmatrix} A_i + B_i & \mathbf{0} \\ \mathbf{0} & A_i - B_i \end{bmatrix} \begin{bmatrix} f_{i,Re} I_T & -f_{i,Im} I_T \\ f_{i,Im} I_T & f_{i,Re} I_T \end{bmatrix},$$

and the $2T \times 1$ real vector

$$\mathcal{W} = \begin{bmatrix} \mathbf{w}_{Re} \\ \mathbf{w}_{Im} \end{bmatrix} + \sqrt{\frac{P_2}{P_1 + 1}} \sum_{i=1}^{R} \begin{bmatrix} g_{i,Re} I_T & -g_{i,Im} I_T \\ g_{i,Im} I_T & g_{i,Re} I_T \end{bmatrix} \begin{bmatrix} A_i + B_i & \mathbf{0} \\ \mathbf{0} & A_i - B_i \end{bmatrix} \begin{bmatrix} \mathbf{v}_{i,Re} \\ \mathbf{v}_{i,Im} \end{bmatrix}.$$

The following equivalent system equation is obtained.

$$\hat{\mathbf{x}} = \sqrt{\frac{P_1 P_2 T}{P_1 + 1}} \mathcal{H} \hat{\mathbf{s}} + \mathcal{W},$$

where $\mathcal{H}$ is the equivalent channel matrix and $\mathcal{W}$ is the equivalent noise.

**Theorem 7.6 (ML decoding and PEP).** *Design the transmit signal at the i-th relay node as in (7.31). Then*

$$\mathrm{P}\left(\mathbf{x}|\mathbf{s}_i\right) = \frac{1}{\left[2\pi\left(1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2\right)\right]^T} e^{-\frac{\left(\hat{\mathbf{x}}-\sqrt{\frac{P_1P_2T}{P_1+1}}\mathcal{H}\hat{\mathbf{s}}_i\right)^*\left(\hat{\mathbf{x}}-\sqrt{\frac{P_1P_2T}{P_1+1}}\mathcal{H}\hat{\mathbf{s}}_i\right)}{2\left(1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2\right)}}, \tag{7.33}$$

*and the ML decoding is*

$$\arg\max_{\mathbf{s}_i}\mathrm{P}\left(\mathbf{x}|\mathbf{s}_i\right) = \arg\min_{\mathbf{s}_i}\left\|\hat{\mathbf{x}}-\sqrt{\frac{P_1P_2T}{P_1+1}}\mathcal{H}\hat{\mathbf{s}}_i\right\|_F^2.$$

*Using the optimum power allocation given in (7.11), the PEP of mistaking $\mathbf{s}_i$ with $\mathbf{s}_j$ has the following Chernoff upper bound for large $P$.*

$$Pe \leq \mathop{\mathrm{E}}_{g_i}\det{}^{-1/2}\left[I_{2R}+\frac{PT}{8\left(R+\sum_{k=1}^{R}|g_k|^2\right)}\sum_{k=1}^{R}\mathcal{G}_k\mathcal{G}_k^t\right], \tag{7.34}$$

*where*

$$\mathcal{G}_k = \begin{bmatrix} g_{k,Re}I_T & -g_{k,Im}I_T \\ g_{k,Im}I_T & g_{k,Re}I_T \end{bmatrix}\begin{bmatrix} A_k+B_k & \mathbf{0} \\ \mathbf{0} & A_k-B_k \end{bmatrix}\begin{bmatrix} (\mathbf{s}_i-\mathbf{s}_j)_{Re} & -(\mathbf{s}_i-\mathbf{s}_j)_{Im} \\ (\mathbf{s}_i-\mathbf{s}_j)_{Im} & (\mathbf{s}_i-\mathbf{s}_j)_{Re} \end{bmatrix}.$$

**Proof:** To get the distribution of $\mathbf{x}|\mathbf{s}_i$, let's first discuss the noise part. Since $\mathbf{v}_i$ and $\mathbf{w}$ are independent circularly symmetric Gaussian with mean 0 and variance $I_T$, $\mathcal{W}$ is Gaussian with mean zero and its variance can be calculated to be $\left(1+\frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2\right)I_{2T}$. Therefore, when $f_i$ and $g_i$ are known, $\hat{\mathbf{x}}|\mathbf{s}_i$ is Gaussian with mean $\sqrt{\frac{P_1P_2T}{P_1+1}}\mathcal{H}\hat{\mathbf{s}}_i$ and variance the same as that of $\mathcal{W}$. Thus, $\mathrm{P}\left(\mathbf{x}|\mathbf{s}_i\right) = \mathrm{P}\left(\hat{\mathbf{x}}|\hat{\mathbf{s}}_i\right)$ is as given in (7.33). It is straightforward to get the ML decoding from the distribution.

Now let's look at the PEP of mistaking $\mathbf{s}_i$ by $\mathbf{s}_j$. By the same argument as in the

proof of Theorem 7.1, the PEP has the following Chernoff upper bound.

$$Pe \leq \underset{f_i,g_i}{\mathrm{E}}\ e^{-\frac{P_1P_2T}{8\left(1+P_1+P_2\sum_{i=1}^{R}|g_i|^2\right)}[\mathcal{H}(\widehat{\mathbf{s}_i-\mathbf{s}_j})]^t\mathcal{H}(\widehat{\mathbf{s}_i-\mathbf{s}_j})}.$$

Note that

$$\begin{bmatrix} f_{i,Re}I_T & -f_{i,Im}I_T \\ f_{i,Im}I_T & f_{i,Re}I_T \end{bmatrix}\begin{bmatrix} (\mathbf{s}_i-\mathbf{s}_j)_{Re} \\ (\mathbf{s}_i-\mathbf{s}_j)_{Im} \end{bmatrix} = \begin{bmatrix} (\mathbf{s}_i-\mathbf{s}_j)_{Re} & -(\mathbf{s}_i-\mathbf{s}_j)_{Im} \\ (\mathbf{s}_i-\mathbf{s}_j)_{Im} & (\mathbf{s}_i-\mathbf{s}_j)_{Re} \end{bmatrix}\begin{bmatrix} f_{i,Re} \\ f_{i,Im} \end{bmatrix}.$$

Therefore,

$$\begin{aligned} & \mathcal{H}(\widehat{\mathbf{s}_i-\mathbf{s}_j}) \\ = & \sum_{i=1}^{R}\begin{bmatrix} g_{i,Re}I_T & -g_{i,Im}I_T \\ g_{i,Im}I_T & g_{i,Re}I_T \end{bmatrix}\begin{bmatrix} A_i+B_i & \mathbf{0} \\ \mathbf{0} & A_i-B_i \end{bmatrix}\begin{bmatrix} f_{i,Re}I_T & -f_{i,Im}I_T \\ f_{i,Im}I_T & f_{i,Re}I_T \end{bmatrix}\begin{bmatrix} (\mathbf{s}_i-\mathbf{s}_j)_{Re} \\ (\mathbf{s}_i-\mathbf{s}_j)_{Im} \end{bmatrix} \\ = & \sum_{i=1}^{R}\begin{bmatrix} g_{i,Re}I_T & -g_{i,Im}I_T \\ g_{i,Im}I_T & g_{i,Re}I_T \end{bmatrix}\begin{bmatrix} A_i+B_i & \mathbf{0} \\ \mathbf{0} & A_i-B_i \end{bmatrix}\begin{bmatrix} (\mathbf{s}_i-\mathbf{s}_j)_{Re} & -(\mathbf{s}_i-\mathbf{s}_j)_{Im} \\ (\mathbf{s}_i-\mathbf{s}_j)_{Im} & (\mathbf{s}_i-\mathbf{s}_j)_{Re} \end{bmatrix}\begin{bmatrix} f_{i,Re} \\ f_{i,Im} \end{bmatrix} \\ = & \hat{\mathcal{H}}\begin{bmatrix} \vdots \\ \begin{bmatrix} f_{i,Re} \\ f_{i,Im} \end{bmatrix} \\ \vdots \end{bmatrix}, \end{aligned}$$

where $\hat{\mathcal{H}} = [\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_R]$ is a $2T \times 2R$ real matrix. Now we can calculate the expectation over $f_{i,Re}$ and $f_{i,Im}$. Similar to the argument in the proof of Theorem 7.1, the following can be proved.

$$Pe \leq \underset{g_i}{\mathrm{E}}\det{}^{-1/2}\left[I_{2R}+\frac{P_1P_2T}{4\left(1+P_1+P_2\sum_{i=1}^{R}|g_i|^2\right)}\hat{\mathcal{H}}\hat{\mathcal{H}}^t\right].$$

The same as before, using the approximation $\sum_{i=1}^{R} |g_i|^2 \approx R$, the optimum power allocation is as given in (7.11). Using this power allocation, (7.34) is obtained. □

## 7.11 Either $A_i = 0$ or $B_i = 0$

We have not yet been able to explicitly evaluate the expectation in (7.34). Our conjecture is that when $T \geq R$, the same transmit diversity $R\left(1 - \frac{\log\log P}{\log P}\right)$ will be obtained. Here we give an analysis of a much simpler, but far from trivial, case: for any $i$, either $A_i = 0$ or $B_i = 0$. That is, each relay node sends a signal that is either linear in its received signal or linear in the conjugate of its received signal. It is clear to see that Alamouti's scheme is included in this case with $R = 2, A_1 = I_2, B_1 = 0, A_2 = 0$, and $B_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. The conditions that $A_i + B_i$ and $A_i - B_i$ are orthogonal become that $A_i$ is orthogonal if $B_i = 0$ and $B_i$ is orthogonal if $A_i = 0$.

**Theorem 7.7.** *Design the transmitted signal at the i-th relay node as in (7.31). Use the optimum power allocation given in (7.11). Further assume that for any $i = 1, \cdots, R$, either $A_i = 0$ or $B_i = 0$. The PEP of mistaking $\mathbf{s}_i$ with $\mathbf{s}_j$ has the following Chernoff upper bound.*

$$Pe \leq \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left[I_R + \frac{PT}{8\left(R + \sum_{i=1}^{R} |g_i|^2\right)}(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right], \quad (7.35)$$

*where*

$$\hat{S}_i = [A_1\mathbf{s}_i + B_1\overline{\mathbf{s}_i}, \cdots, A_R\mathbf{s}_i + B_R\overline{\mathbf{s}_i}] \quad (7.36)$$

*is a $T \times R$ matrix which is the distributed space-time code.*

**Proof:** See Section 7.14.2. □

(7.35) is exactly the same as (7.12) except that now the distributed space-time

code is $\hat{S}$ instead of $S$. Therefore, by the same argument, the following theorem can be obtained.

**Theorem 7.8.** *Design the transmit signal at the i-th relay as in (7.31). Use the optimum power allocation as given in (7.11). For the full diversity of the space-time code, assume $T \geq R$. Define*

$$\hat{M} = (\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j). \tag{7.37}$$

*If $\log P \gg 1$, the PEP has the following Chernoff bound.*

$$Pe \lesssim \sum_{k=0}^{R} \left(\frac{8}{T}\right)^k \sum_{1 \leq i_1 < \cdots < i_k \leq R} \left[\sum_{l=0}^{k} B_R(k-l,k)\right] \det{}^{-1}[\hat{M}]_{i_1,\cdots,i_k} \frac{\log^l P}{P^R}.$$

*The best transmit diversity that can be obtained is $\alpha_0 R$. When $P \gg \log P$,*

$$Pe \lesssim \sum_{k=0}^{R} \left(\frac{8}{T}\right)^k \sum_{1 \leq i_1 < \cdots < i_k \leq R} \left[\sum_{l=0}^{R} B_R(k-l,k)\right] \det{}^{-1}[\hat{M}]_{i_1,\cdots,i_k} P^{-[\alpha_0 R + (1-\alpha_0)(k-l)]}.$$

**Proof:** The same as the proof of Theorem 7.3 and 7.4. □

Therefore, exactly the same transmit diversity is obtained as in Section 7.7 and 7.8. The coding gain for very large $P$ ($P \gg \log P$) is $\det \hat{M}$. When $P$ is not very large, we not only want $\det \hat{M}$ to be large but also want $\det[\hat{M}]_{i_1,\cdots,i_k}$ to be large for all $0 \leq k \leq R, 1 \leq i_1 < \cdots < i_k \leq R$. That is, to have good performance for not very large transmit power, the distributed space-time code should have the property that it is "scale-free" in the sense that it is still a good distributed space-time code when some of the relays are not working.

## 7.12 Simulation Results

In this section, we give the simulated performance of the distributed space-time codes for different values of the coherence interval $T$, number of relay nodes $R$, and total transmit power $P$. The fading coefficients between the transmitter and the relays, $f_i$, and between the receiver and the relays, $g_i$, are modeled as independent complex Gaussian variables with zero-mean and unit-variance. The fading coefficients keep constant for $T$ channel uses. The noises at the relays and the receiver are also modeled as independent zero-mean unit-variance Gaussian additive noise. The block error rate (BLER), which corresponds to errors in decoding the vector of transmitted signals $\mathbf{s}$, and the bit error rate (BER), which corresponds to errors in decoding $s_1, \cdots, s_T$, is demonstrated as the error events of interest. Note that one block error rate may correspond to only a few bit errors.

The transmit signals at each relay are designed as in (7.4). We should remark that our goal here is to compare the performance of linear dispersion (LD) codes implemented distributively over wireless networks with the performance of the same codes in multiple-antenna systems. Therefore the actual design of the LD codes and their optimality is not an issue here: all that matters is that the codes used for simulations in both systems be the same.[8] Therefore, the matrices, $A_i$, are generated randomly based on the isotropic distribution on the space of $T \times T$ unitary matrices. It is certainly conceivable that the performance obtained in the following figures can be improved by several dB if $A_i$s are chosen optimally.

The transmitted signals $s_1, \cdots, s_T$ are designed as independent $N^2$-QAM signals. Both the real and imaginary parts of $s_i$ are equal probably chosen from the $N$-PAM

[8]The question of how to design optimal codes is an interesting one, but is beyond the scope of this thesis.

signal set:

$$\sqrt{\frac{6}{T(N^2-1)}}\{-(N-1)/2, \cdots, -1/2, 1/2, \cdots, (N-1)/2\},$$

where $N$ is a positive integer. The coefficient $\sqrt{\frac{6}{T(N^2-1)}}$ is used for the normalization of $\mathbf{s}$ given in (7.1). The number of possible transmitted signal is therefore $L^{2T}$. Since the channel is used in blocks of $T$ transmissions, the rate of the code is, therefore,

$$\frac{1}{T}\log N^{2T} = 2\log N.$$

In the simulations of the multiple-antenna systems, the number of transmit antennas is $R$ and the number of receive antennas is 1. We also model the fading coefficients between the transmit antennas and the receive antenna as independent zero-mean unit-variance complex Gaussian. The noises at the receive antenna are also modeled as independent zero-mean unit-variance complex Gaussian. As discussed in the chapter, the space-time code used is the $T \times R$ matrix $S = [A_1\mathbf{s}, \cdots, A_R\mathbf{s}]$. The rate of the space-time code is again $2\log N$. In both systems, sphere decoding [DAML00, HV02] is used to obtain the ML results.

### 7.12.1 Performance of Wireless Networks with Different $T$ and $R$

In Figure 7.3, the BER curves of relay networks with different coherence interval $T$ and number of relay nodes $R$ are shown. The solid line indicates the BER of a network with $T = R = 5$, the line with circles indicates the BER of a network with $T = 10$ and $R = 5$, the dash-dotted line indicates the BER of a network with $T = R = 10$, and the line with stars indicates the BER of a network with $T = R = 20$. It can be seen from the plot that the bigger $R$, the faster the BER

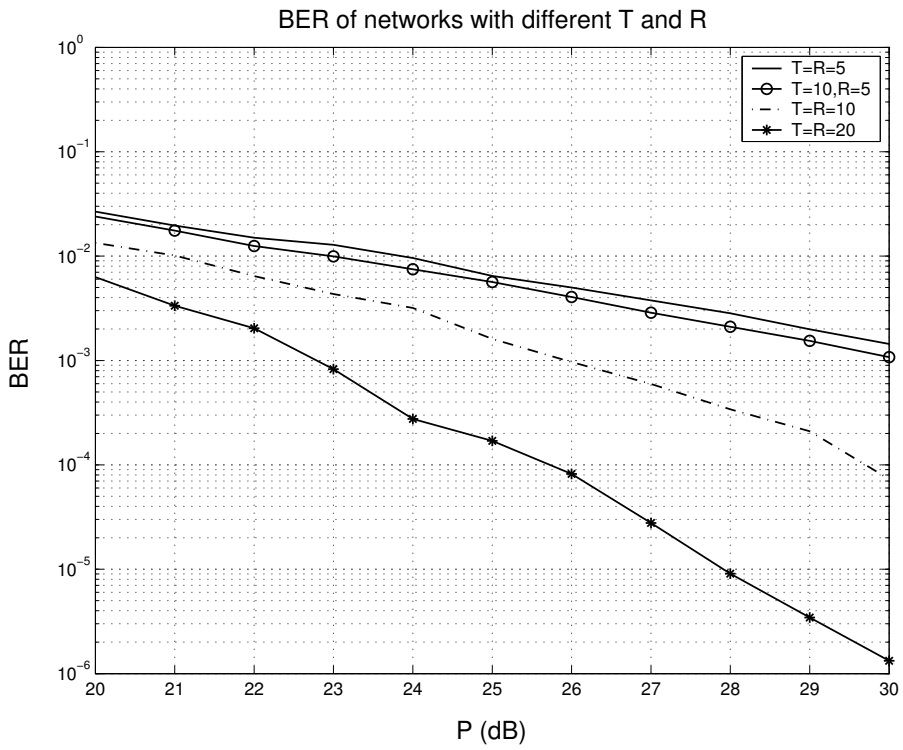


Figure 7.3: BER comparison of wireless networks with different $T$ and $R$

curve decreases, which verifies our analysis that the diversity is linear in $R$ when $T \geq R$. However, the slopes of the BER curves of networks with $T = R = 5$ and $T = 10, R = 5$ is the same. This verifies our result that the transmit diversity only depends on $\min\{T, R\}$, which is always $R$ in our examples. Increasing the coherence interval does not improve the diversity. According to the analysis in Section 6 and 7, increasing $T$ can improvement the coding gain. However, when having a larger coherence interval, not much performance improvement can be seen from the plot by comparing the solid line (the BER curve of network with $T = R = 5$) and the line with circles (the BER curve of network with $T = 10, R = 5$). The reason may be that our code is randomly chosen without any optimization.

### 7.12.2 Perfromance Comparisions of Distributed Space-Time Codes with Space-Time Codes

In this subsection, the performance of relay networks using distributed space-time codes is compared with that of multiple-antenna systems using the same space-time codes. The performance is compared in two ways. In one, we assume that the *total transmit power* for both the systems is the same. This is done since the noise and channel variances are everywhere normalized to unity. In other words, the total transmit power in the networks (summed over the transmitter and $R$ relay nodes) is the same as the transmit power of the multiple-antenna systems. In the other, we assume that the *SNR at the receiver* is the same for the two systems. Assuming that the total transmit power is $P$, in the distributed scheme the SNR can be calculated to be $\frac{P^2}{4(1+P)}$, and in the multiple-antenna setting it is $P$. Thus, roughly a 6 dB increase in power is needed to make the SNR of the relay networks identical to that of the multiple-antenna systems. In the examples below, plots of both comparisons are provided.

In the first example, $T = R = 5$ and $N = 2$. Therefore, the rate of both the distributed space-time code and the space-time code is 2 bits per transmission. The BER and BLER curves are shown in Figure 7.4 and 7.5. Figure 7.4 shows the BER and BLER of the two systems with respect to the total transmit power. Figure 7.5 shows the BER and BLER of the two systems with respect to the receive SNR. In both figures, the solid and dashed curves indicate the BER and BLER of the relay network. The curve with plus signs and curve with circles indicate the BER and BLER of the multiple-antenna system. It can be seen from the figures that the performance of the multiple-antenna system is always better than the relay network at any total transmit power or SNR. This is what we expected because in the multiple-antenna system, the transmit antennas of the transmitter can fully cooperate and

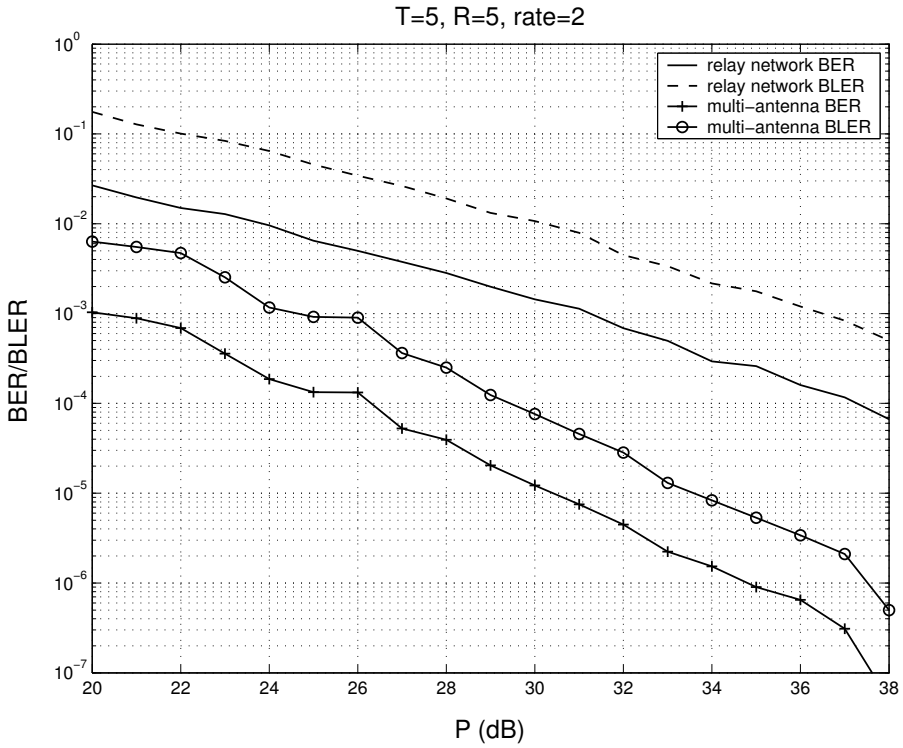


Figure 7.4: BER/BLER comparison of relay network with multiple-antenna system at $T = R = 5, \text{rate} = 2$ and the same total transmit power

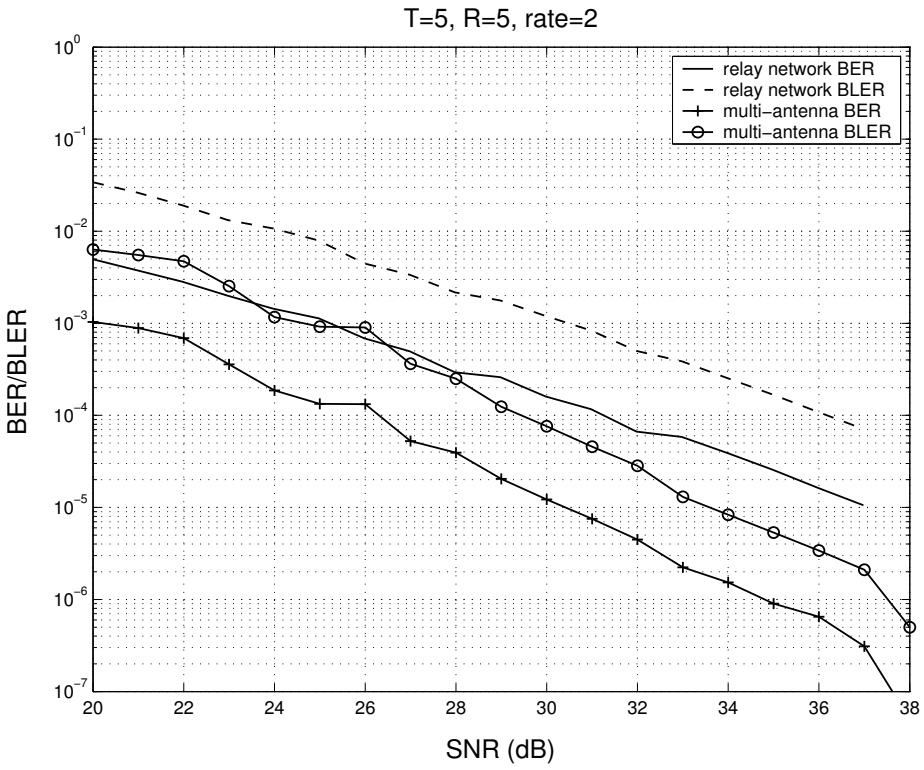


Figure 7.5: BER/BLER comparison of relay network with multiple-antenna system at $T = R = 5, \text{rate} = 2$ and the same receive SNR

have perfect information of the transmit signal. Also, Figure 7.4 shows that the BER and BLER curves of the multiple-antenna system goes done faster than those of the relay network. However, the differences of the slopes of the BER and BLER curves of the two systems are diminishing as the total transmit power goes higher. This can be seen more clearly in Figure 7.5. At low SNR (0-10dB), the BER and BLER curves of the multiple-antenna system go down faster than those of the relay network. As SNR goes higher, the differences of slopes of the BER curves and BLER curves vanishes, which indicates that the two system have about the same diversity at high SNR. This verifies our analysis of the transmit diversity.

Also, in Figure 7.4, at the BER of $10^{-4}$, the total transmit power of the relay network is about 37.5dB. Our analysis of (7.25) indicates that the performance of the relay network should be 12.36dB worse. Reading from the plot, we get a 11.5dB difference. This verifies the correctness and tightness of our upper bound.

In the next example, $T = R = 10$ and $N = 2$. Therefore, the rate is again 2. The simulated performances are as shown in Figure 7.6 and 7.7. Figure 7.6 shows the BER and BLER of the two systems with respect to the total transmit power. Figure 7.7 shows the BER and BLER of the two systems with respect to the receive SNR. The indicators of the curves are the same as before. The BER and BLER of both the relay network and the multiple-antenna system are lower than those in the previous example. This is because there are more relay nodes or transmit antennas present. From Figure 7.6, it can be seen that the multiple-antenna system has a higher diversity at low transmit power. However, as the total transmit power or SNR goes higher, the slope differences of the BER and BLER curves between the two systems diminish. Figure 7.7 shows the same phenomenon. When the receive SNR is low (0-10dB), the performance of the two systems are about the same. However, the BER and BLER curves of the multiple-antenna system goes done faster than those of the relay network. When SNR is high (above 20dB), the BER and BLER curves

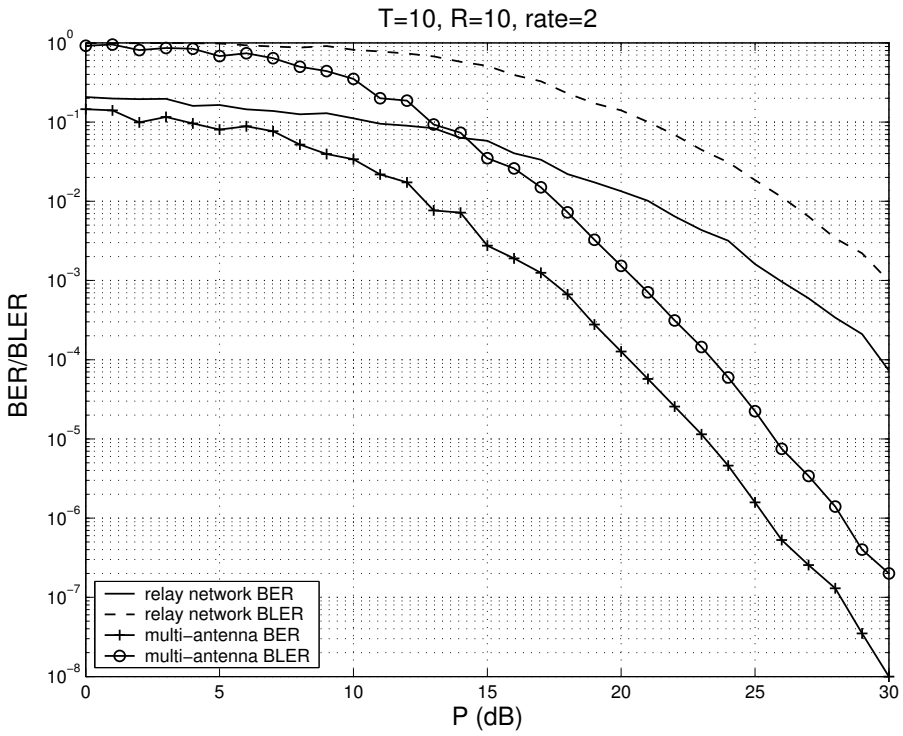


Figure 7.6: BER/BLER comparison of the relay network with the multiple-antenna system at $T = R = 10, \text{rate} = 2$ and the same total transmit power

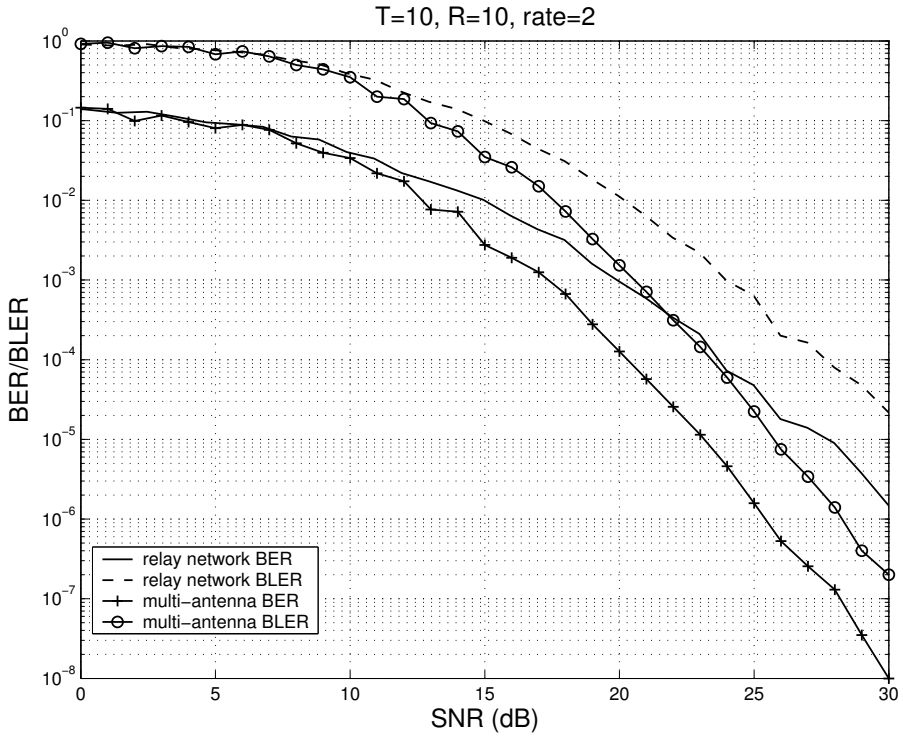


Figure 7.7: BER/BLER comparison of the relay network with the multiple-antenna system at $T = R = 10, \text{rate} = 2$ and the same receive SNR

have about the same slope.

Also, in Figure 7.6, at the BER of $10^{-3}$, the total transmit power of the relay network is about 26dB. Our analysis of (7.25) indicates that the performance of the relay network should be 10.77dB worse than that of the multiple-antenna system. Reading from the plot, we get a 9dB difference. At a BER of $10^{-4}$, the total transmit power of the relay network is about 30dB. Our analysis of (7.25) indicates that the performance of the relay network should be 11.39dB worse. Reading from the plot, we get a 10dB difference.

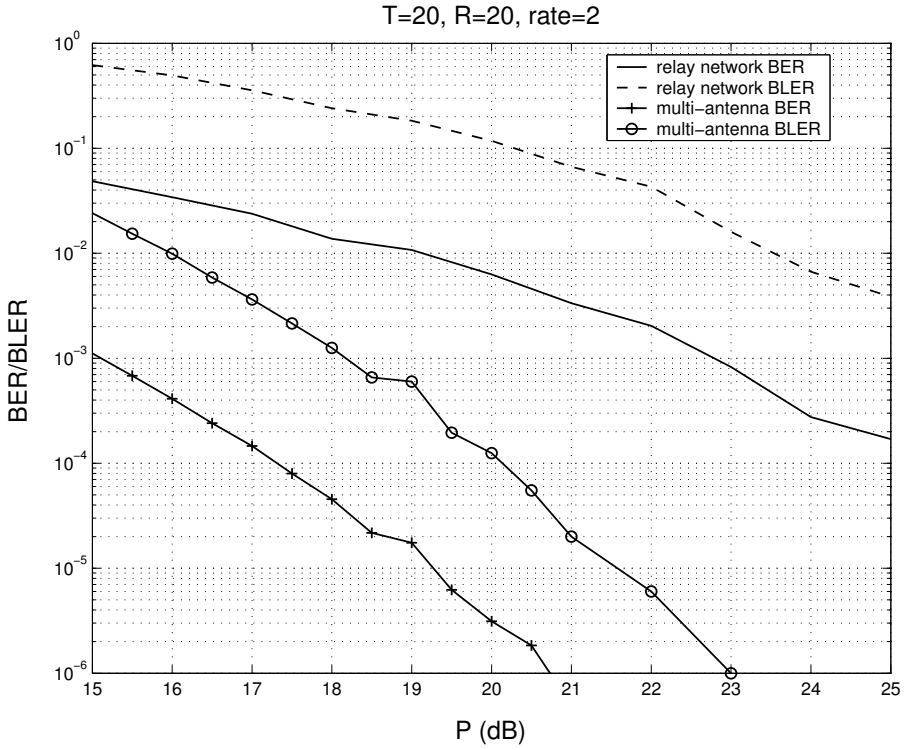


Figure 7.8: BER/BLER comparison of the relay network with the multiple-antenna system at $T = R = 20, \text{rate} = 2$ and the same total transmit power

Figure 7.8 and Figure 7.9 show the performance of systems with $T = R = 20$ and $N = 2$. The rate is again 2. Figure 7.8 shows the BER and BLER of the two systems with respect to the total transmit power. Figure 7.9 shows the BER and BLER of the two systems with respect to the receive SNR. The indicators of the curves are

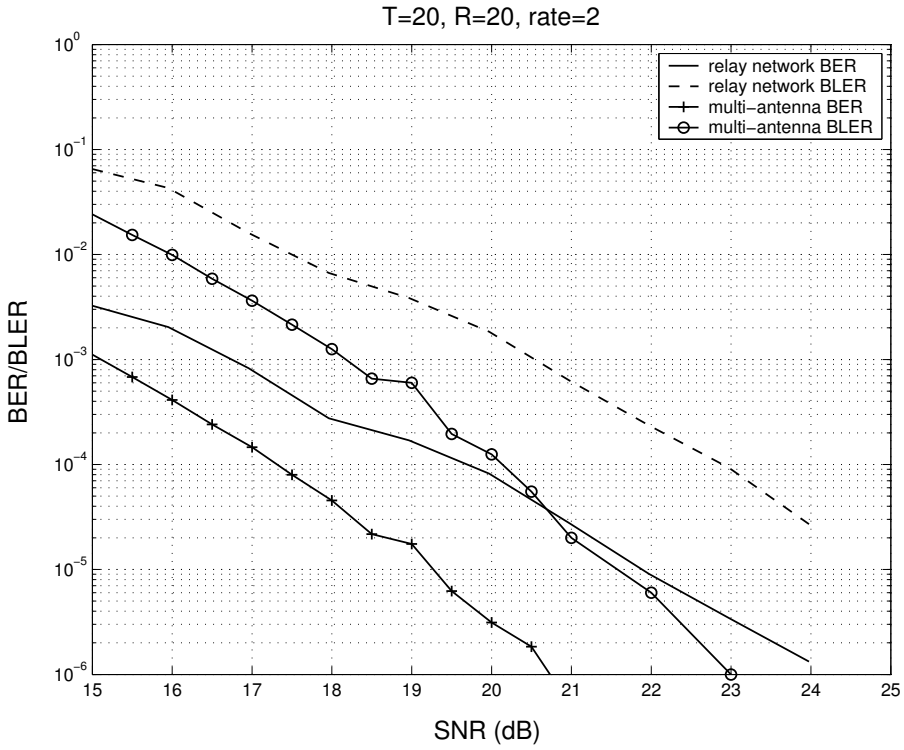


Figure 7.9: BER/BLER comparison of the relay network with the multiple-antenna system with $T = R = 20, \text{rate} = 2$ and the same receive SNR

the same as before. It can be seen from the figures that for total transmit power or SNR higher than 20, the slopes of the BER and BLER curves of the two systems are about the same. Also, from Figure 7.9 we can see that for SNR less than 14dB, the performance of the two systems are about the same. However, the BER and BLER curves of the multiple-antenna system goes down faster than those of the relay network. When SNR is high (above 20dB), the performance difference converges.

Again, in Figure 7.8, at a BER of $10^{-4}$, the total transmit power of the relay network is about 26dB. Our analysis of (7.25) indicates that the performance of the relay network should be 10.77dB worse. Reading from the plot, we get a 9dB difference.

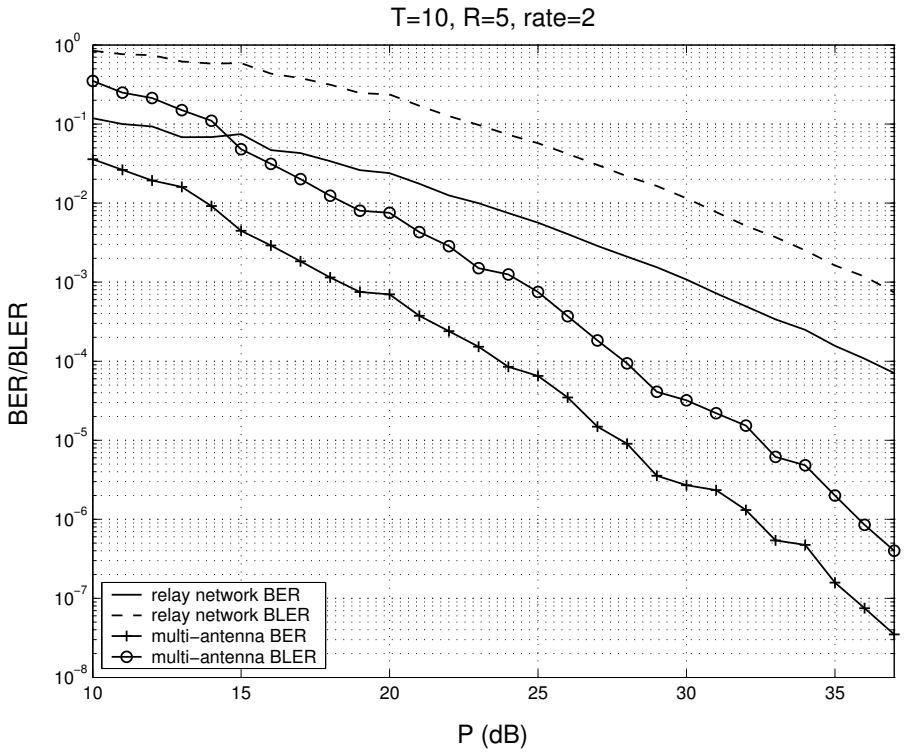


Figure 7.10: BER/BLER comparison of the relay network with the multiple-antenna system at $T = 10, R = 5, \text{rate} = 2$ and the same total transmit power

Finally, we give an example with $T \neq R$. In this example, $T = 10$, $R = 5$ and $N = 2$. The rate of the system is again 2. The BER and BLER curves of both

the relay network and the multiple antenna system with respect to the average total transmit power are shown in Figure 7.10. The indicators of the curves are the same as before. The same phenomenon can be observed.

## 7.13 Conclusion and Future Work

In this chapter, the use of linear dispersion space-time codes in wireless relay networks is proposed. We assume that the transmitter and relay nodes do not know the channel realizations but only their statistical distribution. The ML decoding and pairwise error probability at the receiver is analyzed. The main result is that the diversity of the system behaves as $\min\{T, R\}\left(1 - \frac{\log\log P}{\log P}\right)$, which shows that when $T \geq R$ and the average total transmit power is very high ($P \gg \log P$), the relay network has almost the same diversity as a multiple-antenna system with $R$ transmit antennas and one receive antenna. This result is also supported by simulations. It is further shown that, assuming $R = T$, the leading order term in the PEP behaves as $\frac{1}{|\det(S_i - S_j)|^2}\left(\frac{8\log P}{P}\right)^R$, which compared to $\frac{1}{|\det(S_i - S_j)|^2}\left(\frac{4}{P}\right)^R$, the PEP of a space-time code, shows the loss of performance due to the fact that the code is implemented distributively and the relay nodes have no knowledge of the transmitted symbols. We also observe that the high SNR coding gain, $|\det(S_i - S_j)|^{-2}$, is the same as that arises in space-time coding. The same is true at low SNR where a trace condition comes up.

We then improve the achieved diversity gain slightly (by the order no larger than $O\left(\frac{\log\log P}{\log^2 P}\right)$). Furthermore, a more general type of distributed space-time linear codes is discussed, in which the transmission signal from each relay node to the receive node is designed as a linear combination of both its received signal and the conjugate of its received signal. For a special case, which includes the Alamouti's scheme, exactly the same diversity gain can be obtained. Simulation results on some randomly generated

distributed space-time codes are demonstrated, which verify our theoretical analysis on both the diversity and coding gain.

There are several directions for future work that can be envisioned. One is to study the outage capacity of our scheme. Another is to determine whether the diversity order $\min\{T, R\}\left(1 - \frac{\log\log P}{\log P}\right)$ can be improved by other coding methods that are more complicated and general than linear code used here. We conjecture that it cannot. Another interesting question is to study the design and optimization of distributed space-time codes. For this the PEP expression (7.19) in Theorem 7.3 should be useful. In fact, relay networks provide an opportunity for the design of space-time codes with a large number of transmit antennas, since $R$ can be quite large. Also, in our network model, only single antenna is used at every node. What if there are multiple antennas at the transmit node, the receive node, and/or the relay nodes? For multiple-antenna systems, it has been shown in Chapter 2 that the diversity increases linearly in the number of transmit and receive antennas. Here, in relay networks, can we obtain the same linear increase in the number of antennas nodes in the network?

As discussed in the introduction of this chapter, decode-and-forward can achieve higher diversity however with a strict rate constraint. With the scheme of distributed space-time coding, in which no decoding is need at any relay node, there is no rate constraint. If, in the relay network, we allow some of the relay nodes to decode and then all the relay nodes, those who decode and those who do not, generate a distributed space-time code, it is conceivable that the diversity can be improved with some sacrifice of rate (needed for the decoding relay nodes to decode correctly). Therefore there is a diversity-and-rate trade-off to be analyzed.

Finally, in our network model, it is assumed that the receive node knows all the channel information, which needs the system to be synchronized at the symbol level and needs training symbols to be sent from both the transmit node and the relay

nodes. It is interesting to see whether differential space-time coding techniques can be generalized to the distributed setting that there is no channel information at the receiver as well. The Cayley codes [HH02a] might be a suitable candidate for this.

## 7.14 Appendices

### 7.14.1 Proof of Lemma 7.1

**Proof:** We want to explicitly evaluate

$$I \equiv \int_x^\infty \cdots \int_x^\infty \left( A + \sum_{i=1}^k \lambda_i \right)^k \frac{e^{-\lambda_1} e^{-\lambda_2} \cdots e^{-\lambda_k}}{\lambda_1 \cdots \lambda_k} \, d\lambda_1 \cdots d\lambda_k.$$

Consider the expansion of $\left( A + \sum_{i=1}^k \lambda_i \right)^k$ into monomial terms. We have

$$\left( A + \sum_{i=1}^k \lambda_i \right)^k = \sum_{j=0}^k \left( \sum_{1 \le l_1 < \cdots < l_j \le k} \sum_{i_1=1}^{k} \sum_{i_2=1}^{k-i_1} \cdots \sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}} C(i_1, \ldots, i_j) \lambda_{l_1}^{i_1} \lambda_{l_2}^{i_2} \cdots \lambda_{l_j}^{i_j} A^{k-i_1-\cdots-i_j} \right),$$

where $j$ denotes how many $\lambda$'s are present, $l_1, \ldots, l_j$ are the subscripts of the $j$ $\lambda$'s that appears, $i_m \ge 1$ indicates that $\lambda_{l_m}$ is taken to the $i_m$-th power (the summation should be

$$\sum_{\substack{i_1, \ldots, i_j \ge 1 \\ \sum i_m \le k}},$$

which is equivalent to

$$\sum_{i_1=1}^{k} \sum_{i_2=1}^{k-i_1} \cdots \sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}}.$$

if we sum $i_1$ first, then $i_2$, etc. ), and finally

$$C(i_1, \ldots, i_j) = \binom{k}{i_1} \binom{k-i_1}{i_2} \cdots \binom{k-i_1-\cdots-i_{j-1}}{i_j}$$

counts how many times the term $\lambda_{l_1}^{i_1}\lambda_{l_2}^{i_2}\cdots\lambda_{l_j}^{i_j}A^{k-i_1-\cdots-i_j}$ appears in the expansion.

Thus we have

$$I=\sum_{j=0}^{k}\sum_{1\le l_1<\cdots<l_j\le k}\sum_{i_1=1}^{k}\cdots\sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}}C(i_1,\ldots,i_j)I(j;l_1,\ldots,l_j;i_1,\ldots,i_j),$$

where

$$I(j;l_1,\ldots,l_j;i_1,\ldots,i_j)\equiv\int_x^\infty\cdots\int_x^\infty\lambda_{l_1}^{i_1}\lambda_{l_2}^{i_2}\cdots\lambda_{l_j}^{i_j}A^{k-i_1-\cdots-i_j}\frac{e^{-\lambda_1}\cdots e^{-\lambda_k}}{\lambda_1\cdots\lambda_k}\,d\lambda_1\cdots d\lambda_k.$$

We compute

$$\begin{aligned}
& I(j;l_1,\ldots,l_j;i_1,\ldots,i_j)\\
= \;& A^{k-i_1-\cdots-i_j}\left(\prod_{m=1}^{j}\int_x^\infty\lambda_{l_m}^{i_m-1}e^{-\lambda_{l_m}}\,d\lambda_{l_m}\right)\prod_{i\neq i_1,\ldots i_j}\int_x^\infty\frac{e^{-\lambda_i}}{\lambda_i}\,d\lambda_i\\
= \;& A^{k-i_1-\cdots-i_j}\left(\prod_{m=1}^{j}\Gamma(i_m,x)\right)\left[-\mathbf{Ei}(-x)\right]^{k-j}.
\end{aligned}$$

Note that the result is independent of $l_1,\ldots,l_j$. Finally adding the terms up, we have

$$\begin{aligned}
& I\\
= \;& \sum_{j=0}^{k}\sum_{1\le l_1<\cdots<l_j\le k}\sum_{i_1=1}^{k}\cdots\sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}}C(i_1,\ldots,i_j)A^{k-i_1-\cdots-i_j}\left[-\mathbf{Ei}(-x)\right]^{k-j}\prod_{m=1}^{j}\Gamma(i_m,x)\\
= \;& \sum_{j=0}^{k}\left[\left(\sum_{1\le l_1<\cdots<l_j\le k}1\right)\left(\sum_{i_1=1}^{k}\cdots\sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}}C(i_1,\ldots,i_j)A^{k-i_1-\cdots-i_j}\Gamma(i_1,x)\cdots\Gamma(i_j,x)\right)\right]\\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\left[-\mathbf{Ei}(-x)\right]^{k-j}\\
= \;& \sum_{j=0}^{k}\left[\binom{k}{j}\sum_{i_1=1}^{k}\cdots\sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}}\binom{k}{i_1}\cdots\binom{k-i_1-\cdots-i_{j-1}}{i_j}\Gamma(i_1,x)\cdots\Gamma(i_j,x)A^{k-i_1-\cdots-i_j}\right]\\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\left[-\mathbf{Ei}(-x)\right]^{k-j}\\
\equiv \;& \sum_{j=0}^{k}B_{A,x}(j,k)\left[-\mathbf{Ei}(-x)\right]^{k-j}.
\end{aligned}$$

Thus ends the proof. □

### 7.14.2 Proof of Theorem 7.7

**Proof:** Note that

$$\begin{bmatrix} g_{k,Re}I_T & -g_{k,Im}I_T \\ g_{k,Im}I_T & g_{k,Re}I_T \end{bmatrix} \begin{bmatrix} A_k & 0 \\ 0 & A_k \end{bmatrix} = \begin{bmatrix} A_k & 0 \\ 0 & A_k \end{bmatrix} \begin{bmatrix} g_{k,Re}I_T & -g_{k,Im}I_T \\ g_{k,Im}I_T & g_{k,Re}I_T \end{bmatrix},$$

$$\begin{bmatrix} g_{k,Re}I_T & -g_{k,Im}I_T \\ g_{k,Im}I_T & g_{k,Re}I_T \end{bmatrix} \begin{bmatrix} 0 & B_k \\ B_k & 0 \end{bmatrix} = \begin{bmatrix} 0 & B_k \\ B_k & 0 \end{bmatrix} \begin{bmatrix} g_{k,Re}I_T & g_{k,Im}I_T \\ -g_{k,Im}I_T & g_{k,Re}I_T \end{bmatrix},$$

and

$$\begin{bmatrix} g_{k,Re}I_T & \mp g_{k,Im}I_T \\ \pm g_{k,Im}I_T & g_{k,Re}I_T \end{bmatrix} \begin{bmatrix} (\mathbf{s}_i - \mathbf{s}_j)_{Re} & -(\mathbf{s}_i - \mathbf{s}_j)_{Im} \\ (\mathbf{s}_i - \mathbf{s}_j)_{Im} & (\mathbf{s}_i - \mathbf{s}_j)_{Re} \end{bmatrix}$$

$$= \begin{bmatrix} (\mathbf{s}_i - \mathbf{s}_j)_{Re} & -(\mathbf{s}_i - \mathbf{s}_j)_{Im} \\ (\mathbf{s}_i - \mathbf{s}_j)_{Im} & (\mathbf{s}_i - \mathbf{s}_j)_{Re} \end{bmatrix} \begin{bmatrix} g_{k,Re} & \mp g_{k,Im} \\ \pm g_{k,Im} & g_{k,Re} \end{bmatrix}.$$

Therefore,

$$\begin{bmatrix} g_{k,Re}I_T & -g_{k,Im}I_T \\ g_{k,Im}I_T & g_{k,Re}I_T \end{bmatrix} \begin{bmatrix} A_k + B_k & \mathbf{0} \\ \mathbf{0} & A_k - B_k \end{bmatrix} \begin{bmatrix} (\mathbf{s}_i - \mathbf{s}_j)_{Re} & -(\mathbf{s}_i - \mathbf{s}_j)_{Im} \\ (\mathbf{s}_i - \mathbf{s}_j)_{Im} & (\mathbf{s}_i - \mathbf{s}_j)_{Re} \end{bmatrix}$$

$$= \begin{bmatrix} A_k + B_k & \mathbf{0} \\ \mathbf{0} & A_k - B_k \end{bmatrix} \begin{bmatrix} (\mathbf{s}_i - \mathbf{s}_j)_{Re} & -(\mathbf{s}_i - \mathbf{s}_j)_{Im} \\ (\mathbf{s}_i - \mathbf{s}_j)_{Im} & (\mathbf{s}_i - \mathbf{s}_j)_{Re} \end{bmatrix} \begin{bmatrix} g_{k,Re}I_T & -\mathrm{sgn}_k g_{k,Im}I_T \\ \mathrm{sgn}_k g_{k,Im}I_T & g_{k,Re}I_T \end{bmatrix},$$

where $\mathrm{sgn}_k = 1$ if $B_k = 0$ and $\mathrm{sgn}_k = -1$ if $A_k = 0$. Thus,

$$\mathcal{G}_k\mathcal{G}_k^t = \begin{bmatrix} A_k + B_k & \mathbf{0} \\ \mathbf{0} & A_k - B_k \end{bmatrix} \begin{bmatrix} (\mathbf{s}_i - \mathbf{s}_j)_{Re} & -(\mathbf{s}_i - \mathbf{s}_j)_{Im} \\ (\mathbf{s}_i - \mathbf{s}_j)_{Im} & (\mathbf{s}_i - \mathbf{s}_j)_{Re} \end{bmatrix} \begin{bmatrix} |g_k|^2 & 0 \\ 0 & |g_k|^2 \end{bmatrix} \begin{bmatrix} (\mathbf{s}_i - \mathbf{s}_j)_{Re} & -(\mathbf{s}_i - \mathbf{s}_j)_{Im} \\ (\mathbf{s}_i - \mathbf{s}_j)_{Im} & (\mathbf{s}_i - \mathbf{s}_j)_{Re} \end{bmatrix}^t \begin{bmatrix} A_k + B_k & \mathbf{0} \\ \mathbf{0} & A_k - B_k \end{bmatrix}^t.$$

Define

$$S_i' = \left[ \begin{bmatrix} A_1 - B_1 & \mathbf{0} \\ \mathbf{0} & A_1 - B_1 \end{bmatrix} \begin{bmatrix} \mathbf{s}_{i,Re} & -\mathbf{s}_{i,Im} \\ \mathbf{s}_{i,Im} & \mathbf{s}_{i,Re} \end{bmatrix} \quad \cdots \quad \begin{bmatrix} A_R + B_R & \mathbf{0} \\ \mathbf{0} & A_R - B_R \end{bmatrix} \begin{bmatrix} \mathbf{s}_{i,Re} & -\mathbf{s}_{i,Im} \\ \mathbf{s}_{i,Im} & \mathbf{s}_{i,Re} \end{bmatrix} \right] \quad (7.38)$$

(7.34) thus becomes,

$$\begin{aligned} Pe &\lesssim \mathop{\mathrm{E}}_{g_i} \det{}^{-1} \left[ I_{2R} + \frac{PT}{8\left(R + \sum_{i=1}^R |g_i|^2\right)} (S_i' - S_j') \mathrm{diag}\left\{|g_1|^2, |g_1|^2, \cdots, |g_R|^2, |g_R|^2\right\} (S_i' - S_j')^t \right] \\ &= \mathop{\mathrm{E}}_{g_i} \det{}^{-1} \left[ I_{2R} + \frac{PT}{8\left(R + \sum_{i=1}^R |g_i|^2\right)} (S_i' - S_j')^t (S_i' - S_j') \mathrm{diag}\left\{|g_1|^2, |g_1|^2, \cdots, |g_R|^2, |g_R|^2\right\} \right]. \end{aligned}$$

Define

$$\mathrm{SGN} = \mathrm{diag}\left\{ \cdots \quad \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{sgn}_k \end{bmatrix} \quad \cdots \right\}.$$

Note that $\det \mathrm{SGN} = 1$. From the definition of $S_i'$ in (7.38),

$$\begin{aligned} &(S_i' - S_j')\mathrm{SGN} \\ = &\begin{bmatrix} \cdots & \begin{bmatrix} A_k + B_k & \mathbf{0} \\ \mathbf{0} & A_k - B_k \end{bmatrix} \begin{bmatrix} \mathbf{s}_{i,Re} - \mathbf{s}_{j,Re} & -\mathrm{sgn}_k(\mathbf{s}_{i,Im} - \mathbf{s}_{j,Im}) \\ \mathbf{s}_{i,Im} - \mathbf{s}_{j,Im} & \mathrm{sgn}_k(\mathbf{s}_{i,Re} - \mathbf{s}_{j,Re}) \end{bmatrix} & \cdots \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} \cdots & \begin{bmatrix} [A_k(\mathbf{s}_i - \mathbf{s}_j) + B_k(\overline{\mathbf{s}}_i - \overline{\mathbf{s}}_j)]_{Re} & -[A_k(\mathbf{s}_i - \mathbf{s}_j) + B_k(\overline{\mathbf{s}}_i - \overline{\mathbf{s}}_j)]_{Im} \\ [A_k(\mathbf{s}_i - \mathbf{s}_j) + B_k(\overline{\mathbf{s}}_i - \overline{\mathbf{s}}_j)]_{Im} & [A_k(\mathbf{s}_i - \mathbf{s}_j) + B_k(\overline{\mathbf{s}}_i - \overline{\mathbf{s}}_j)]_{Re} \end{bmatrix} & \cdots \end{bmatrix}.$$

It is easy to see that the matrix $\begin{bmatrix} [\hat{S}_i - \hat{S}_j]_{Re} & -[\hat{S}_i - \hat{S}_j]_{Im} \\ [\hat{S}_i - \hat{S}_j]_{Im} & [\hat{S}_i - \hat{S}_j]_{Re} \end{bmatrix}$ can be obtained by switching the columns of $(S'_i - S'_j)\mathrm{SGN}$. More precisely,

$$\begin{bmatrix} [\hat{S}_i - \hat{S}_j]_{Re} & -[\hat{S}_i - \hat{S}_j]_{Im} \\ [\hat{S}_i - \hat{S}_j]_{Im} & [\hat{S}_i - \hat{S}_j]_{Re} \end{bmatrix} = (S'_i - S'_j)\mathrm{SGN} \prod_{k=2}^{R} E_k,$$

where $E_k = [e_1, \cdots, e_{k-1}, e_{2k-1}, e_k, \cdots, e_{2k-2}, e_{2k}, \cdots, e_{2R}]$ with $\{e_k\}$ the standard basis of $\mathfrak{R}^R$. It is easy to see that $E_k^{-1} = E_k^t$ and $\det E_k = 1$. Right multiplying by $E_k$, we move the $(2k-1)$-th column of a matrix to the $k$-th position and shift the $k$-th to $(2k-2)$-th columns one column right, that is,

$$[c_1, \cdots, c_n]\, E_k = [c_1 \cdots, c_{k-1}, c_{2k-1}, c_k, \cdots, c_{2k-2}, c_{2k}, \cdots, c_{2R}]\,.$$

Therefore,

$$\begin{aligned} & (S'_i - S'_j)^t (S'_i - S'_j)\mathrm{diag}\,\{|g_1|^2, |g_1|^2, \cdots, |g_R|^2, |g_R|^2\} \\ = \ & \mathrm{SGN} \prod_{k=2}^{R} E_k \begin{bmatrix} [\hat{S}_i - \hat{S}_j]_{Re} & [\hat{S}_i - \hat{S}_j]_{Im} \\ -[\hat{S}_i - \hat{S}_j]_{Im} & [\hat{S}_i - \hat{S}_j]_{Re} \end{bmatrix}^t \begin{bmatrix} [\hat{S}_i - \hat{S}_j]_{Re} & [\hat{S}_i - \hat{S}_j]_{Im} \\ -[\hat{S}_i - \hat{S}_j]_{Im} & [\hat{S}_i - \hat{S}_j]_{Re} \end{bmatrix} \\ & \qquad \prod_{k=R}^{2} E_k^t \mathrm{SGN}\,\mathrm{diag}\,\{|g_1|^2, |g_1|^2, \cdots, |g_R|^2, |g_R|^2\} \\ = \ & \mathrm{SGN} \prod_{k=2}^{R} E_k \begin{bmatrix} [\hat{S}_i - \hat{S}_j]_{Re} & [\hat{S}_i - \hat{S}_j]_{Im} \\ -[\hat{S}_i - \hat{S}_j]_{Im} & [\hat{S}_i - \hat{S}_j]_{Re} \end{bmatrix}^t \begin{bmatrix} [\hat{S}_i - \hat{S}_j]_{Re} & [\hat{S}_i - \hat{S}_j]_{Im} \\ -[\hat{S}_i - \hat{S}_j]_{Im} & [\hat{S}_i - \hat{S}_j]_{Re} \end{bmatrix} \\ & \qquad \mathrm{diag}\,\{|g_1|^2, \cdots, |g_R|^2, |g_1|^2, \cdots, |g_R|^2\} \prod_{k=R}^{2} E_k^t \mathrm{SGN} \end{aligned}$$

$$= \text{SGN} \prod_{k=2}^{R} E_k \begin{bmatrix} [(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)]_{Re} & [(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)]_{Im} \\ -[(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)]_{Im} & [(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)]_{Re} \end{bmatrix}$$

$$\text{diag}\,\{\mathbf{G}, \mathbf{G}\} \prod_{k=R}^{2} E_k^t \text{SGN}$$

$$= \text{SGN} \prod_{k=2}^{R} E_k \begin{bmatrix} [(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Re} & [(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Im} \\ -[(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Im} & [(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Re} \end{bmatrix} \prod_{k=R}^{2} E_k^t \text{SGN}\,,$$

where we have defined $\mathbf{G} = \text{diag}\,\{|g_1|^2, \cdots, |g_1|^2\}$. Note that for any complex matrix $A$,

$$\det \begin{bmatrix} A_{Re} & A_{Im} \\ -A_{Im} & A_{Re} \end{bmatrix} = |\det A|^2.$$

Therefore,

$$\det\left(I_{2R} + \frac{PT}{8\left(R + \sum_{i=1}^{R} |g_i|^2\right)} (S_i' - S_j')^t (S_i' - S_j') \text{diag}\,\{|g_1|^2, |g_1|^2, \cdots, |g_R|^2, |g_R|^2\}\right)$$

$$= \det\left(I_{2R} + \frac{PT}{8\left(R + \sum_{i=1}^{R} |g_i|^2\right)} \begin{bmatrix} [(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Re} & [(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Im} \\ -[(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Im} & [(\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Re} \end{bmatrix}\right)$$

$$= \det \begin{bmatrix} [I_R + \frac{PT}{8\left(R + \sum_{i=1}^{R} |g_i|^2\right)} (\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Re} & [I_R + \frac{PT}{8\left(R + \sum_{i=1}^{R} |g_i|^2\right)} (\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Im} \\ -[I_R + \frac{PT}{8\left(R + \sum_{i=1}^{R} |g_i|^2\right)} (\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Im} & [I_R + \frac{PT}{8\left(R + \sum_{i=1}^{R} |g_i|^2\right)} (\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j)\mathbf{G}]_{Re} \end{bmatrix}$$

$$= \det\left[I_R + \frac{PT}{8\left(R + \sum_{i=1}^{R} |g_i|^2\right)} (\hat{S}_i - \hat{S}_j)^*(\hat{S}_i - \hat{S}_j) G G^*\right].$$

□

# Chapter 8 Summary and Discussion

Reaching the end of this thesis, to conclude, a brief summary of contributions of this thesis and discussions on possible future research directions are given in the following. As this thesis can be roughly divided into two big parts: the MIMO/multiple-antenna systems part and the wireless ad hoc network part, a separate summary and discussion are provided for each.

## 8.1 Summary and Discussion on Multiple-Antenna Systems

From Chapter 1 to Chapter 6, multiple-antenna systems are discussed. In the first two chapters, an introduction of multiple-antenna systems is given including the fading model, the systems model, the Shannon capacity, the training-based scheme, the differential and non-differential unitary space-time modulations, and the diversity and coding gain according to the PEP of multiple-antenna systems. An example, which is the well-known Alamouti's scheme, is presented. A brief review of the real and complex sphere decoding algorithms, which are widely used as fast ML decoding algorithms in multiple-antenna communications, is also given. In Chapter 3, non-square unitary space-time codes is designed via Cayley transform. The code can be used to systems with any number of transmit and receive antennas with a fast nearly optimal decoding algorithm. Preliminary simulations show that the code is far better than the uncoded training-based space-time schemes and only slightly underperforms optimized training-based schemes using orthogonal designs and linear dispersion codes. Chapters 4, 5, and 6 are on differential unitary space-time codes based on groups. In

Chapter 4, the idea of group-based differential unitary space-time codes is introduced and its advantages are explained. The research on rank 2 Lie groups is also motivated in Chapter 4. Following this, differential unitary space-time codes based on rank 2 Lie groups, $Sp(2)$ and $SU(3)$, are described, which can be used in systems with four and three transmit antennas and any number of receive antennas, respectively. The codes are fully diverse with high diversity products. Simulations show that they are superior to existing differential Cayley codes, orthogonal designs, finite-group-based codes, and are comparable to the elaborately-designed non-group codes which have a structure of products of groups.

There are still many open questions and unsolved problems in this area. The most prominent one is the capacity. The capacity of multiple-antenna systems is still unknown when neither the transmitter and the receiver has the channel information, which is the most practical case. Although some results are obtained for very high [MH99, HM00, ZT02, LM03] and very low SNR cases [LTV03, PV02, HS02b, RH04], we just scratched the surface of the research on the capacity of multiple-antenna systems. The capacity when partial channel information is available and the capacity for systems with frequency-selective channels are also open. This area of research will remain timely and important for many years.

As most research on multiple-antenna systems focused on exploiting the diversity gain provided by multiple antennas, there is another gain, called the *spatial multiplexing* or the *degrees of freedom*, corresponding to the increase in the data rate provided by multiple antennas. In [ZT03], it is proved that these two types of gains can be obtained simultaneously, however, there is a fundamental trade-off between how much each of the two gains can be extracted. Then comes the question of finding practical codes which can actually achieve the optimal trade-off between diversity and spatial multiplexing with good performance. In [GCD04], a coding scheme called LAST coding is proposed and is proved to achieve the optimal trade-off. Other related work

can be found in [LK04, GT04, TV04].

Another open problem is the error rate of multiple-antenna systems. The analysis on the exact block or bit error rate is very difficult. In all the analysis given in this thesis, the Chernoff upper bound on the pairwise error probability is used. There are also works on the exact pairwise error probability [TB02]. Any improvement in the analysis of exact block/bit error rate or any results of non-trivial lower bound on the PEP will be very interesting.

## 8.2 Summary and Discussion on Wireless Ad Hoc Networks

Chapter 7 is about wireless ad hoc networks. In this chapter, the idea of space-time coding proposed for multiple-antenna systems is applied to wireless relay networks, by which diversity $R$ is achieved when the transmit power is asymptotically high, where $R$ is the number of relay nodes in the network. This result indicates that wireless networks with $R$ relay nodes can achieve the same diversity as multiple-antenna systems with $R$ transmit antennas and one receive antenna at asymptotically high transmit power although space-time codes are used distributively among the $R$ relay nodes.

As discussed in Chapter 7, the straightforward future research are the analysis on the outage capacity with this coding scheme and coding scheme designs when no channel information is known at the receiver. Another possibility is the optimization of the distributed space-time codes. It is mentioned in Chapter 7 that the design criterion is the same as space-time codes for multiple-antenna systems for very high and low transmit power. However, it is different for intermediate transmit power. Also, the design of space-time codes for large $R$ is rare.

In research on wireless relay networks, as mentioned in the introduction of Chap-

ter 7, there are two mainly used cooperative diversity algorithms for transmissions between a pair of nodes through a bunch of relay nodes: amplify-and-forward [DH03] and decoding-and-forward [LW03, NBK04]. Intuitively, when the receive SNR of a relay node is low (for example, if the relay is far from the transmitter), it is not beneficial for the relay to do decoding since with high probability, it will make an error. As discussed in [DH03], wireless networks are most power efficient in the low power regime, in which case, the receive SNR at the relay nodes is low. In this situation, decoding-and-forward is not advantageous. However, if some relay nodes are very near the transmitter, it might be advantageous for them to decode since they have high receive SNRs according to diversity gain, capacity, outage capacity, etc. In our approach, simple signal processing, which is called the distributed space-time coding, is used at the relay nodes. No decoding needs to be done at the relay nodes, which both saves the computational complexity and improves the reliability when the SNR is not very high. This algorithm is superior to amplify-and-forward since the latter is actually a special case of the former. Other work based on this algorithm can be find in [CH03, HMC03, LW03]. A mixed algorithm of decode-and-forward and distributed space-time coding according to the instantaneous SNR and transmission rate at the relay nodes will be interesting. A trade-off between diversity and rate is expected.

In our network model, all the relay nodes have the same power allocation. This might not be applicable for real ad hoc or sensory networks. What is more important is that this might be not optimal if some relay nodes have full or partial knowledge of their local channels. Relay nodes can estimate their instantaneous receive SNR from the transmit node at every time. Therefore, it might be advantageous for those relay nodes who have high SNR to use higher transmit power to relay the signals. Therefore, the optimal power allocation among the relay nodes is another interesting problem.

In the network model in Chapter 7, there is only one transmit-and-receive pair,

which is applicable to most sensory networks but not to general ad hoc wireless networks. When there are multiple pairs of transmit and receive nodes, not only does noise exist but also interference. One most straightforward method to solve this is to use time division or frequency division by assigning a different time instant or frequency interval to every pair. However, it will be interesting to see if there are better and more efficient strategies.

As discussed before, because of their features and related issues (distributivity, interference, routing, power constraint, mobility, etc.), analysis on wireless ad hoc networks is very difficult. Most questions on wireless ad hoc networks are still open. For example, what is the capacity, what is the optimal diversity gain, is multi-hop routing better than single-hop routing, and what is the optimal power allocation? Related work can be found in [GK00, GK01, GT02, TG03]. To get some results in this area, most of the work nowadays focus on one of the two special networks: networks with a small amount of nodes so that theoretical analysis are possible (for example, [TG03, CH04]) and networks with very large number of nodes in which asymptotic results may be obtained [GK00, GK01, GT02]. Understanding wireless ad hoc network is the key to our ultimate goal of wireless communication: to communicate with anybody anywhere at anytime for anything. For a considerable period of time, research on wireless ad hoc networks will keep timely, interesting and significant.

# Bibliography

[AD01] N. Al-Dhahir. Single-carrier frequency-domain equalization for sapce-time block-coded transmission over frequency-selective channels. *IEEE Communications Letters*, 5:304–306, July 2001.

[Ala98] S. M. Alamouti. A simple transmitter diversity scheme for wireless communications. *IEEE Jour. on Selected Areas in Communications*, pages 1451–1458, Oct. 1998.

[Ari00] S. L. Ariyavisitakul. Turbo space-time processing to improve wireless channel capacity. *IEEE Trans. on Communications*, 48:1347–1359, Aug. 2000.

[Art99] M. Artin. *Algebra*. Printice-Hall, 1999.

[Art04a] H. Artés. Channel dependent tree pruning for the sphere decoder. In *Proc. of 2004 IEEE International Symposium on Information Theory (ISIT'04)*, June-July 2004.

[Art04b] H. Artés. Efficient near-optimum decoding of space-time linear dispersion codes using a modified sphere decoder. In *Proc. of 2004 International Symposium on Signals, Systems, and Electronics (ISSSE'04)*, Aug. 2004.

[ARU01] D. Agrawal, T. J. Richardson, and R. L. Urbanke. Multiple-antenna signal constellations for fading channels. *IEEE Trans. on Information Theory*, 47:2618–2626, Sep. 2001.

[ATNS98] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri. Space-time coded OFDM for high data rate wireless communication over wide-band chan-

nels. In *Proc. of 1998 Fall IEEE Vehicular Technology Conference (VTC'98-Fall)*, pages 2232–2236, Sep. 1998.

[AVZ02] E. Agrell, A. Vardy, and K. Zeger. Closest point search in lattices. *IEEE Trans. on Information Theory*, 48:2201–2214, Aug. 2002.

[BBH00] S. Baro, G. Bauch, and A. Hansmann. Improved codes for space-time trellis-coded modulation. *IEEE Trans. on Communications Letters*, 4:20–22, Jan. 2000.

[BD01] I. Bahceci and T. M. Duman. Trellis coded unitary space-time modulation. In *Proc. of 2001 IEEE Global Telecommunications Conference (GLOBECOM'01)*, volume 2, pages 25–29, Nov. 2001.

[BGP00] H. Boleskei, D. Gesbert, and A. J. Paulraj. On the capacity of OFDM-based multi-antenna systems. In *Proc. of 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP'00)*, volume 5, pages 2569–2572, June 2000.

[BGT93] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon limit error-correcting coding: turbo codes. In *Proc. of 1993 International Conference on Communications (ICC'93)*, pages 1064–1070, May 1993.

[BMJ+98] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva. A performance comparison of multi-hop wireless ad hoc network routing protocols. In *Proc. of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'98)*, Oct. 1998.

[BS99] A. R. Bahai and B. R. Saltzberg. *Multicarrier Digital Communications: Theory and Applications of OFDM.* Plenum Publishing Corp., 1999.

[BtD95] T. Brocker and T tom Dieck. *Representations of Compact Lie Groups*. Springer-Verlag, 1995.

[Cal03] G. M. Calhoun. *The Third Generation Wireless Systems, Volume I: Post-Shannon Signal Architectures*. Artech House, 2003.

[CC99] W. J. Choi and J. M. Cioffi. Space-time block codes over frequency selective Rayleigh fading channels. In *Proc. of 1999 Fall IEEE Vehicular Technology Conference (VTC'99-Fall)*, pages 2541–2545, Sep. 1999.

[CFG02] D. Chizhik, G. J. Foschini, and M. J. Gans. Keyholes, correlations, and capacities of multielement transmit and receive antennas. *IEEE Trans. on Wireless Communications*, 1:361–368, Apr. 2002.

[CH03] Y. Chang and Y. Hua. Application of space-time linear block codes to parallel wireless relays in mobile ad hoc networks. In *Proc. of the 37th Asilomar Conference on Signals, Systems and Computers (Asilomar'03)*, volume 1, pages 1002 – 1006, Nov. 2003.

[CH04] Y. Chang and Y. Hua. Diversity analysis orthogonal space-time modulation for distributed wireless relays. In *Proc. of 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP'04)*, volume IV, pages 561–564, May 2004.

[CT91] T. Cover and J. Thomas. *Elements of Information Theory*. John Wiley and Sons Inc., 1991.

[CTK02] C. N. Chuah, D. Tse, and J. M. Kahn. Capacity scaling in MIMO wireless systems under correlated fading. *IEEE Trans. on Information Theory*, 48:637–650, Mar. 2002.

[CYV01] Z. Chen, J. Yuan, and B. Vucetic. Improved space-time trellis-coded modulation scheme on slow Rayleigh fading channels. *IEEE Electronic Letters*, 37:440–441, Mar. 2001.

[DAML00] M. O. Damen, K. Abed-Meraim, and M. S. Lemdani. Further results on the sphere decoder algorithm. *Submitted to IEEE Trans. on Information Theory*, 2000.

[DBT03] O. Dousse, F. Baccelli, and P. Thiran. Impact of interferences on connectivity in ad hoc networks. In *Proc. of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (Infocom'03)*, volume 3, pages 1724 – 1733, Mar.-Apr. 2003.

[DCB00] M. O. Damen, A. Chkeif, and J.-C. Belfiore. Lattice code decoder for space-time codes. *IEEE Communications Letters*, pages 161–163, May 2000.

[DF99] Dummit and Foote. *Abstract Algebra*. John Wiley and Sons Inc., 2nd edition, 1999.

[DH03] A. F. Dana and B. Hassibi. On the power-efficiency of sensory and ad hoc wireless networks. *Submitted to IEEE Trans. on Information Theory*, 2003.

[DSG+03] A. F. Dana, M. Sharif, R. Gowaikar, B. Hassibi, and M. Effros. Is broadcast plus multi-access optimal for Gaussian wireless network? In *Proc. of the 37th Asilomar Conference on Signals, Systems and Computers (Asilomar'03)*, volume 2, pages 1748 – 1752, Nov. 2003.

[DTB02] M. O. Damen, A. Tewfik, and J. C. Belfiore. A construction of space-time code based on number theory. *IEEE Trans. on Information Theory*, 48:753–760, Mar. 2002.

[ECS+98] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed. Overview of spatial channel models for antenna array communication systems. *IEEE Trans. on Personal Communications*, 5:502–512, Feb. 1998.

[Ede89] A. Edelman. *Eigenvalues and Condition Numbers of Random Matrices*. Ph.D. Thesis, MIT, Department of Math., 1989.

[EHP93] M. Evans, N. Hastings, and B. Peacock. *Statistical Distributions*. John Wiley and Sons Inc., 2nd edition, 1993.

[Fos96] 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.

[FP85] U. Fincke and M. Pohst. Improved methods for calculating vectors of short length in a lattice, including a complexity analysis. *Mathematics of Computation*, 44(170):463–471, April 1985.

[FVY01] W. Firmanto, B. S. Vucetic, and J. H. Yuan. Space-time TCM with improved performance on fast fading channels. *IEEE Trans. on Communications Letters*, 5:154–156, Apr. 2001.

[Gal62] R.G. Gallager. Low-density parity-check codes. *IEEE Trans. on Information Theory*, 8:21–28, Jan. 1962.

[GCD04] H. E. Gamal, G. Caire, and M. O. Damen. Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels. *IEEE Trans. on Information Theory*, 50:968–985, June 2004.

[GD03] H. E. Gamal and M. O. Damen. Universal space-time coding. *IEEE Trans. on Information Theory*, 49:1097–1119, May. 2003.

[GFVW99] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky. Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture. *Electronic Letters*, 35:14–16, Jan. 1999.

[GH03] R. Gowaikar and B. Hassibi. Statistical pruning for maximum-likelihood decoding. *submitted to IEEE Trans. on Information Theory*, 2003.

[GK00] P. Guptar and P. R. Kumar. The capacity of wireless network. *IEEE Trans. on Information Theory*, 46:388–404, Mar. 2000.

[GK01] P. Gupta and P. R. Kumar. Internet in the sky: the capacity of three-dimensional wireless networks. *Communications in Information and Systems*, 1:33–50, Jan. 2001.

[GKB02] S. Galliou, I. Kammoun, and J. Belfiore. Space-time codes for the GLRT noncoherent detector. In *Proc. of 2002 International Symposium on Information Theory (ISIT'02)*, page 419, 2002.

[GL96] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 3rd edition, 1996.

[GL00] Y. Gong and K. B. Letaief. Performance evaluation and analysis of space-time coding in unequalized multipath fading links. *IEEE Trans. on Communications*, 48:1778–1782, Nov. 2000.

[GL02] Y. Gong and K. B. Letaief. Concatenated space-time block coding with trellis-coded modulation in fading channels. *IEEE Trans. on Wireless Communications*, 1:580–590, Oct. 2002.

[GR00] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series and Products*. Academic Press, 6th edition, 2000.

[GS79] A.V. Geramita and J. Seberry. *Orthogonal Designs, Quadratic Forms and Hadamard Matrices, Lecture Notes in Pure and Applied Mathematicatics*, volume 43. New York and Basel: Marcel Dekker, 1979.

[GT02] M. Grossglauser and D. Tse. Mobility incereses the capacity of ad hoc wireless networks. *IEEE/ACM Trans. on Networking*, 10:477–486, Aug. 2002.

[GT04] L. Grokop and D. N. C. Tse. Diversity/multiplexing tradeoff in ISI channels. In *Proc. of 2004 International Symposium on Information Theory (ISIT'04)*, June-July 2004.

[GV02] M. Gastpar and M. Vetterli. On the capacity of wireless networks: the relay case. In *Proc. of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (Infocom'02)*, pages 1577 – 1586, June 2002.

[GW02] A. J. Goldsmith and S. B. Wicker. Design challenges for energy-constrained ad hoc wireless networks. *IEEE Wireless Communications*, 9:8–27, Aug. 2002.

[Har00] J. C. Harrtsen. The Bluetooth radio system. *IEEE Personal Communications Magazine*, 7:28–36, Feb. 2000.

[Has99] B. Hassibi. An efficient square-root algorithm for BLAST. *submitted to IEEE Trans. on Signal Processing*, 1999.

[Hay01] S. Haykin. *Communications Systems*. John Wiley and Sons Inc., 4th edition, 2001.

[HH02a] B. Hassibi and B. Hochwald. Cayley differential unitary space-time codes. *IEEE Trans. on Information Theory*, 48:1458–1503, June 2002.

[HH02b] B. Hassibi and B. Hochwald. High-rate codes that are linear in space and time. *IEEE Trans. on Information Theory*, 48(7):1804–1824, July 2002.

[HH03] B. Hassibi and B. Hochwald. How much training is needed in multiple-antenna wireless links? *IEEE Trans. on Information Theory*, 49(4):951–963, Apr. 2003.

[HJ91] R. Horn and C. Johnson. *Topics in Matrix Analysis*. University Press, Cambridge, 1991.

[HJ02a] B. Hassibi and Y. Jing. Unitary space-time codes and the Cayley transform. In *Proc. of 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP'02)*, volume 3, pages 2409–2412, May 2002.

[HJ02b] B. Hassibi and Y. Jing. Unitary space-time modulation via the Cayley transform. In *Proc. of 2002 IEEE International Symposium on Information Theory (ISIT '02)*, page 134, June-July 2002.

[HK00] B. Hassibi and M. Khorrami. Fully-diverse multi-antenna constellations and fixed-point-free Lie groups. *submitted to IEEE Trans. on Information Theory*, 2000.

[HM00] B. M. Hochwald and T. L. Marzetta. Unitary space-time modulation for multiple-antenna communication in Rayleigh flat-fading. *IEEE Trans. on Information Theory*, 46:543–564, Mar. 2000.

[HM02] B. Hassibi and T. L. Marzetta. Multiple-antennas and isotropically-random unitary inputs: the received signal density in closed-form. *IEEE Trans. on Information Theory*, 48:1473–1484, June 2002.

[HMC03] Y. Hua, Y. Mei, and Y. Chang. Wireless antennas - making wireless communications perform like wireline communications. In *Proc. of 2003 IEEE Topical Conference on Wireless Communication Technology*, pages 47 – 73, Oct. 2003.

[HMH01] B. Hochwald, T. Marzetta, and B. Hassibi. Space-time autocoding. *IEEE Trans. on Information Theory*, 47:2761–2781, Nov 2001.

[HMR+00] B. Hochwald, T. Marzetta, T. Richardson, W. Sweldens, and R. Urbanke. Systematic design of unitary space-time constellations. *IEEE Trans. on Information Theory*, 46:1962–1973, Sep. 2000.

[HS00] B. Hochwald and W. Sweldens. Differential unitary space-time modulation. *IEEE Trans. on Communications*, 48:2041–2052, Dec. 2000.

[HS02a] A. Paulraj H. Sampath. Linear precoding for space-time coded systems with known fading correlations. *IEEE Communications Letters*, 6:239–241, June 2002.

[HS02b] B. Hajek and V. Subramaniam. Capacity and reliability function for small peak signal constraints. *IEEE Trans. on Information Theory*, 48:828–839, Apr. 2002.

[HtB03] B.M. Hochwald and S. ten Brink. Achieving near-capacity on a multiple-antenna channel. *IEEE Trans. on Communications*, 51:389–399, Mar. 2003.

[Hug00a] B. Hughes. Differential space-time modulation. *IEEE Trans. on Information Theory*, 46:2567–2578, Nov. 2000.

[Hug00b] B. Hughes. Optimal space-time constellations from groups. *IEEE Trans. on Information Theory*, 49:401 – 410, Feb. 2000.

[Hun74] T. W. Hungerford. *Algebra*. Springer-Verla New York Inc., 1974.

[HVa] B. Hassibi and H. Vikalo. On sphere decoding algorithm: Part I, expected complexity. *submitted to IEEE Trans. on Signal Processing*.

[HVb] B. Hassibi and H. Vikalo. On sphere decoding algorithm: Part II, generalizations, second statistics, and application to communications. *submitted to IEEE Trans. on Signal Processing*.

[HV02] B. Hassibi and H. Vikalo. On the expected complexity of integer least-squares problems. In *Proc. of 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)*, pages 1497 – 1500, April 2002.

[Jaf01] H. Jafarkhani. A quasi-orthogonal space-time code. *IEEE Trans. on Communications*, 49:1–4, Jan. 2001.

[JH02] Y. Jing and B. Hassibi. Fully-diverse multi-antenna space-time codes based on Sp(2). In *Proc. of the 36th Asilomar Conference on Signals, Systems and Computers (Asilomar'02)*, volume 2, pages 1142 – 1146, Nov. 2002.

[JH03a] Y. Jing and B. Hassibi. Design of fully-diverse multi-antenna codes based on Sp(2). In *Proc. of 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing 2003 (ICASSP '03)*, volume 4, Apr. 2003.

[JH03b] Y. Jing and B. Hassibi. Design of fully-diverse multiple-antenna codes based on Sp(2). *to Appear in IEEE Trans. on Information Theory*, 2003.

[JH03c] Y. Jing and B. Hassibi. Fully-diverse Sp(2) code design. In *Proc. of 2003 IEEE International Symposium on Information Theory 2003 (ISIT '03)*, page 299, June-July 2003.

[JH03d] Y. Jing and B. Hassibi. High-rate space-time codes motivated by SU(3). In *Proc. of the 36th Asilomar Conference on Signals, Systems and Computers (Asilomar'03)*, volume 2, pages 1801 – 1805, Nov. 2003.

[JH03e] Y. Jing and B. Hassibi. Unitary space-time modulation via Cayley transform. *IEEE Trans. on Signal Processing Special Issue on MIMO Communications*, 51:2891–2904, Nov. 2003.

[JH04a] Y. Jing and B. Hassibi. Distributed diversity in wireless networks using space-time coding. In *Proc. of the 42nd Annual Allerton Conference on Communication, Control, and Computing*, Sep.-Oct. 2004.

[JH04b] Y. Jing and B. Hassibi. Distributed space-time coding in wireless relay networks - Part I: basic diversity results. *Submitted to IEEE Trans. on Wireless Communications*, 2004.

[JH04c] Y. Jing and B. Hassibi. Distributed space-time coding in wireless relay networks - Part II: tighter bounds and a more general case. *Submitted to IEEE Trans. on Wireless Communications*, 2004.

[JH04d] Y. Jing and B. Hassibi. Space-time code design for three-transmit-antenna systems. In *Proc. of 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP'04)*, May 2004.

[JH04e] Y. Jing and B. Hassibi. Three-transmit-antenna space-time codes based on SU(3). *Submitted to IEEE Trans. on Signal Porcessing*, 2004.

[JH04f] Y. Jing and B. Hassibi. Using space-time codes in wireless relay networks. In *Proc. of the 3rd IEEE Sensory Array and Multichannel Signal Processing Workshop (SAM'04)*, July. 2004.

[JS03] H. Jafarkhani and N. Seshadri. Super-orthogonal space-time trellis codes. *IEEE Trans. on Information Theory*, 49:937–950, Apr. 2003.

[KS04] J. H. Kotecha and A. M. Sayeed. Transmit signal design for optimal estimation of correlated MIMO channels. *IEEE Trans. on Signal Processing*, 52:546–557, Feb. 2004.

[LB02] X. T. Lin and R. S. Blum. Systematic design of space-time codes employing multiple trellis-coded modulation. *IEEE Trans. on Communications*, 50:608–615, Apr. 2002.

[LFT01] Y. J. Liu, M. P. Fitz, and O. Y. Takechita. Full rate space-time turbo codes. *IEEE Jour. on Selected Areas in Communications*, 19:969–980, May 2001.

[LGZM01] Z. Liu, G. B. Giannakis, S. Zhou, and B. Muquet. Space-time coding for broadband wireless communications. *Wireless Communications and Mobile Computing*, 1:35–53, Jan. 2001.

[LK04] H. Lu and P.V. Kumar. Generalized unified construction of space-time codes with optimal rate-diversity tradeoff. In *Proc. of 2004 International Symposium on Information Theory (ISIT'04)*, June-July 2004.

[LLC02] J. Li, K. B. Letaief, and Z. Cao. Adaptive co-channel interference cancellation in space-time coded communication systems. *IEEE Electronic Letters*, 38:129–131, Jan 2002.

[LM03] A. Lapidoth and S. M. Moser. Capacity bounds via duality with applications to multi-antenna systems on flat-fading channels. *IEEE Trans. on Information Theory*, 49:2426–2467, Oct. 2003.

[LSA98] Y. Li, N. Seshadri, and S. Ariyavisitakul. Transmit diversity for OFDM systems with mobile wireless channels. In *Proc. of 1998 IEEE Global Telecommunications Conference (GLOBECOM'98)*, pages 968–973, Nov. 1998.

[LTV03] A. Lozano, A. Tullino, and S. Verdú. Multiple-antenna capacity in the low-power regime. *IEEE Trans. on Information Theory*, 49:2527–2544, 2003.

[LW00] B. Lu and X. Wang. Space-time code design in OFDM systems. In *Proc. of 2000 IEEE Global Telecommunications Conference (GLOBECOM'00)*, pages 1000–1004, Nov. 2000.

[LW03] J. N. Laneman and G. W. Wornell. Distributed space-time-coded protocols for exploiting cooperative diversity in wireless betwork. *IEEE. Trans. on Information Theory*, 49:2415–2425, Oct. 2003.

[Mar99] T. L. Marzetta. BLAST training: estimating channel characteristics for high-capacity space-time wireless. In *Proc. of the 37th Annual Allerton Conference on Communications, Control, and Computing*, Sept. 22-24 1999.

[McE02] Robert J. McEliece. *The Theory of Information and Coding*. Cambridge University Press, 2nd edition, 2002.

[MH99] T. L. Marzetta and B. M. Hochwald. Capacity of a mobile multiple-antenna communication link in Rayleigh flat-fading. *IEEE Trans. on Information Theory*, 45:139–157, Jan. 1999.

[MHH02] T. Marzetta, B. Hassibi, and B. Hochwald. Structured unitary space-time constellations. *IEEE Trans. on Information Theory*, 48:942–950, Apr. 2002.

[MN96] D. J. C. MacKay and R. M. Neal. Near Shannon limit performance of low density parity check codes. *IEE Electronics Letters*, 32:1645–1655, Aug. 1996.

[NBK04] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler. Fading relay channels: performance limits and space-time signal design. *IEEE Jour. on Selected Areas in Communications*, 22:1099 – 1109, Aug. 2004.

[Per01] C. E. Perkins. *Ad Hoc Networking*. Addison-Wesley, 2001.

[PGH00] G. Pei, M. Gerla, and X. Hong. Lanmar: landmark routing for large scale wireless ad hoc networks with group mobility. In *Proc. of the 1st Annual workshop on Mobile and Ad Hoc Networking and Computing*, Aug. 2000.

[PL03] D. P Palomar and M. A. Lagunas. Joint transmit-receive space-time equalization in spatially correlated MIMO channels: a beamforming approach. *IEEE Jour. on Selected Areas in Communications*, 21:730–743, June 2003.

[Poh81] M. Pohst. On the computation of lattice vectors of minimal length, successive minima, and reduced basis with applications. *ACM SIGSAM*, 15:37–44, 1981.

[Pro00] J. G. Proakis. *Digital Communications*. McGraw-Hill, 2000.

[PV02] V. Prelov and S. Verdú. Second-order asymptotics of mutual information. *Center for Discrete Mathematics and Theoretical Computer Science (DIMACS) Technical Report 2002-16*, Apr. 2002.

[Rap02] T. S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall PTR, 2nd edition, 2002.

[RH04] C. Rao and B. Hassibi. Analysis of multiple antenna wireless links at low SNR. *IEEE Trans. on Information Theory*, 50:2123–2130, Sep. 2004.

[RkT99] E. M. Royer and C. k. Toh. A review of current routing protocols for ad hoc mobile wireless networks. *IEEE Personal Communications Magazine*, 6:46–55, Apr. 1999.

[RM99] V. Rudolphu and T. H. Meng. Minimum energy mobile wireless networks. *IEEE Jour. on Selected Areas in Communicarions*, 17:1333–1344, Aug. 1999.

[SD01] A. Stefanov and T. M. Duman. Turbo-coded modulation for systems with transmit and receive antenna diversity over block fading channels: system model, decoding approaches, and practical considerations. *IEEE Jour. on Selected Areas in Communications*, 19:958–968, May 2001.

[SEA03a] A. Sendonaris, E. Erkip, and B. Aazhang. User cooperation diversity - Part I: system description. *IEEE Trans. on Communications*, 51:1927–1938, Nov. 2003.

[SEA03b] A. Sendonaris, E. Erkip, and B. Aazhang. User cooperation diversity - Part II: implementation aspects and performance analysis. *IEEE Trans. on Communications*, 51:1939–1948, Nov. 2003.

[Ser92] Jean-Pierre Serre. *Lie Algebras and Lie Groups*. Springer-Verlag, 1992.

[SFGK00] D. S. Shiu, G. Foschini, M. Gans, and J. Kahn. Fading correlations and effect on the capacity of multielement antenna systems. *IEEE Trans. on Communications*, 48:502–512, Mar. 2000.

[SG01] H. J. Su and E. Geraniotis. Space-time turbo codes with full antenna diversity. *IEEE Trans. on Communications*, 49:47–57, Jan. 2001.

[SHHS01] A. Shokrollahi, B. Hassibi, B. Hochwald, and W. Sweldens. Representation theory for high-rate multiple-antenna code design. *IEEE Trans. on Information Theory*, 47(6):2335–2367, Sep. 2001.

[Sho00] A. Shokrollahi. Design of unitary space-time codes from representations of $su(2)$. *Preprint*, 2000.

[Sim94] B. Simon. *Representations of Finite and Compact Lie Groups*. Providence, R.I. : American Mathematical Society, 1994.

[SOSL85] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt. *Spread Spectrum Communications*, volume 1. Computer Science Press, 1985.

[SS03] O. Simeone and U. Spagnolini. Combined linear pre-equalization and blast equalization with channel correlation feedback. *IEEE Communications Letters*, 7:487–489, Oct. 2003.

[Stu00] G. L. Stuber. *Principles of Mobile Communications*. KAP, 2nd edition, 2000.

[SW86] Sattinger and Weaver. *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*. Springer-Verlag, 1986.

[SWWX04] A. Song, G. Wang, W. Wu, and X.-G. Xia. Unitary space-time codes from Alamouti's scheme with APSK signals. *IEEE Trans. on Wireless Communications*, Sep. 2004.

[TB02] D. Taricco and E. Biglieri. Exact pairwise error probability of space-time codes. *IEEE Trans. on Information Theory*, 48:510–513, Feb. 2002.

[TBH00] O. Tirkkonen, A. Boariu, and A. Hottinen. Minimal nonorthogonality rate 1 space-time block code for 3+ tx antennas. In *Proc. of 2000 IEEE*

*international Symposium on Spread Spectrum Techniques and Applications (ISSSTA'00)*, volume 2, pages 429–432, Sep. 2000.

[TC01] M. X. Tao and R. S. Cheng. Improved design criteria and new trellis codes for space-time coded modulation in slow flat-fading channels. *IEEE Communications Letters*, 5:312–315, July 2001.

[Tel99] I. E. Telatar. Capacity of multi-antenna gaussian channels. *European Trans. on Telecommunications*, 10:585–595, Nov. 1999.

[TG03] S. Toumpis and A.J. Goldsmith. Capacity regions for wireless ad hoc networks. *IEEE Trans. on Wireless Communications*, 2:736–748, July. 2003.

[TH02] O. Tirkkonen and A. Hottinen. Square-matrix embeddable space-time block codes for complex signal constellations. *IEEE Trans. on Information Theory*, 48:384–395, 2002.

[TJC99] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block codes from orthogonal designs. *IEEE Trans. on Information Theory.*, 45:1456–1467, July 1999.

[TK02] V. Tarokh and I. Kim. Existence and construction of non-coherent unitary space-time codes. *IEEE Trans. on Information Theory*, 48:3112–3117, Dec. 2002.

[TSC98] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: performance criterion and code construction. *IEEE Trans. on Information Theory*, 44:744–765, 1998.

[TV01] Y. Tang and M. C. Valenti. Coded transmit macrodiversity: block space-time codes over distributed antennas. In *Proc. of 2001 Spring IEEE*

*Vehicular Technology Conference (VTC'01-Spring)*, volume 2, pages 1435 – 1438, May 2001.

[TV04] S. Tavildar and P. Viswanath. Permutation codes: achieving the diversity-multiplexing tradeoff. In *Proc. of 2004 International Symposium on Information Theory (ISIT'04)*, June-July 2004.

[UG00] M. Uysal and C. N. Georghiades. Error performance analysis of space-time codes over Rayleigh fading channels. In *Proc. of 2000 Fall IEEE Vehicular Technology Conference (VTC'00-Fall)*, volume 5, pages 2285–2290, sep. 2000.

[VB93] E. Viterbo and E. Biglieri. A universal decoding algorithm for lattice codes. In *Quatorzieme colloque GRETSI*, pages 611–614, Sep. 1993.

[VY03] B. Vucetic and J. Yuan. *Space-Time Coding*. Wiley, 2003.

[Win83] J. H. Winters. Switched diversity with feedback for DPSK mobile radio systems. *IEEE Trans. on Vechnology*, 32:134–150, 1983.

[Wit93] A. Wittneben. A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation. In *Proc. of 1993 International Conference on Communications (ICC'93)*, pages 1630–1634, May 1993.

[WX03] H. Wang and X-G Xia. Upper bounds of rates of complex orthogonal space-time block codes. *IEEE Trans. on Information Theory*, 49:2788–2796, Oct. 2003.

[Yos78] K. Yosida. *Functional Analysis*, volume 1. Springer-Verlag, 5th edition, 1978.

[Zas36] H. Zassenhaus. Über endliche Fastkörper. *Abh. Math. Sem. Hamburg*, 11:187–220, 1936.

[ZAS00] S. D. Zummo and S. A. Al-Semari. A tight bound on the error probability of space-time codes for rapid fading channels. In *Proc. of 2000 IEEE Wireless Communications and Networking Conference (WCNC'00)*, pages 1086–1089, 2000.

[ZG03] S. L. Zhou and G. B. Giannakis. Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations. *IEEE Trans. on Information Theory*, 49:1673–1690, Jul. 2003.

[ZT02] L. Zheng and D. Tse. Communication on the Grassman manifold: a geometric approach to the noncoherent multiple-antenna channel. *IEEE Trans. on Information Theory*, 48:359–383, Feb. 2002.

[ZT03] L Zheng and D. Tse. Diversity and multiplexing: A fundemental trade-off in multiple-antenna channels. *IEEE Trans. on Information Theory*, 49:1072–1096, May 2003.