

Unitary Trace-Orthogonal Space-Time Block Codes
in Multiple Antenna Wireless Communications

UNITARY TRACE-ORTHOGONAL SPACE-TIME BLOCK CODES
IN MULTIPLE ANTENNA WIRELESS COMMUNICATIONS

By

Jing Liu, M.A.Sc.

SEPTEMBER 2008

A Thesis

Submitted to the Department of Electrical & Computer Engineering

and the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree of

Doctor of Philosophy

McMaster University

© Copyright by Jing Liu, September 2008

DOCTOR OF PHILOSOPHY (2008)
(Electrical & Computer Engineering)

McMaster University
Hamilton, Ontario

TITLE: Unitary Trace-Orthogonal Space-Time Block Codes in
Multiple Antenna Wireless Communications

AUTHOR: Jing Liu
M.A.Sc., McMaster University, Canada

SUPERVISOR: James P. Reilly, Professor

NUMBER OF PAGES: xvii, 203

Dedications

*To my parents,
my husband.*

Abstract

A multiple-input multiple-output (MIMO) communication system has the potential to provide reliable transmissions at high data rates. However, the computational cost of achieving this promising performance can be quite substantial. With an emphasis on practical implementations, the MIMO systems employing the low cost linear receivers are studied in this thesis. The optimum space-time block codes (STBC) that enable a linear receiver to achieve its best possible performance are proposed for various MIMO systems. These codes satisfy an intra and inter orthogonality property, and are called unitary trace-orthogonal codes. In addition, several novel transmission schemes are specially designed for linear receivers with the use of the proposed code structure. The applications of the unitary trace-orthogonal code are not restricted to systems employing linear receivers. The proposed code structure can be also applied to the systems employing other types of receivers where several originally intractable code design problems are successfully solved.

The communication schemes presented in this thesis are outlined as follows:

- For a MIMO system with $N \geq M$, where M and N are the number of transmitter and receiver antennas, respectively, the optimal full rate linear STBC for linear receivers is proposed and named unitary trace-orthogonal code. The proposed code structure is proved to be necessary and sufficient to achieve the minimum detection error probability for the system.

- When applied to a multiple input single output (MISO) communication system, a special linear unitary trace-orthogonal code, named the Toeplitz STBC, is proposed. The code enables a linear receiver to provide full diversity and to achieve the optimal tradeoff between the detection error and the data transmission rate. This is, thus far, the first code that possesses such properties for an arbitrary MISO system that employs a linear receiver.
- In MIMO systems in which $N \geq M$ and the signals are transmitted at full symbol rate, the highest diversity gain achievable by linear receivers is analyzed and shown to be $N - M + 1$. To improve the performance of a linear receiver, a multi-block transmission scheme is proposed, in which signals are coded so that they span multiple independent channel realizations. An optimal full rate linear STBC for this system that minimizes the detection error probability is presented. The code is named multi-block unitary trace-orthogonal code. The resulting system has an improved diversity gain. Furthermore, by relaxing the code from the full symbol rate constraint, a special multi-block transmission scheme is proposed. This scheme achieves a much improved diversity gain than those with full symbol rate.
- The unitary trace-orthogonal code can also be applied to a system that employs a maximum-likelihood (ML) receiver rather than the simple linear receiver. For such a system, a systematic design of full diversity unitary trace-orthogonal code is presented for an arbitrary data transmission rate.

In summary, when a simple linear receiver is employed, unitary trace-orthogonal codes and their optimality properties are exploited for various multiple antenna communication systems. Some members from this code family can also enable an optimal performance of ML detection.

Acknowledgements

I wish to express my gratitude to all those people who gave me the possibility to complete this thesis. I am deeply indebted to my supervisor Prof. J. P. Reilly for his enthusiasm supports, valuable guidance and encouragement. I would also thank Dr. T. N. Davidson, whose continuous feedbacks and insightful discussions are essential in completing this dissertation. I would like to give my special thanks to Dr. J.-K. Zhang, whose help, stimulating suggestions and encouragement are important in accomplishing the related research results. In addition, I would like to thank my supervisory committee members, Dr. T. R. Field for his fruitful discussions throughout my study and research at McMaster, and Prof. N. Balakrishnan for his precious comments and suggestions.

I would like to give thanks to all my colleagues and staff members at McMaster University who have contributed greatly to provide a supportive and exciting research atmosphere.

I am very grateful to my husband for his patient love and support, which enables me to complete this work.

Last but not least, I would like to thank my parents, who are the reason for everything I have achieved.

List of Acronyms

SISO	Single input single output
SNR	Signal to noise ratio
SINR	Signal to interference plus noise ratio
SIMO	Single input multiple output
MISO	Multiple input single output
MIMO	Multiple input multiple output
IID	Independent identical distribution
STBC	Space – time block code
ML	Maximum – likelihood
FRFD	Full rate full diversity
BEP	Bit error Probability
LD	Linear Dispersion
pcu	per channel use
SVD	Singular Value Decomposition
PDF	Probability Density Function
PEP	Pairwise Error Probability
ZF	Zero – Forcing
MMSE	Minimum Mean Square Error
PD	Positive Definite

PSD Positive Semi Definite
SEP Symbol Error Probability

List of Notations

Boldface lowercase letters are used to denote column vectors.

Boldface uppercase letters are used to denote matrices.

$(\cdot)^*$	the conjugate operator
$(\cdot)^T$	the transpose of a vector or a matrix
$(\cdot)^H$	the Hermitian transpose of a vector or a matrix
$(\cdot)^{-1}$	the inversion of a matrix
$(\hat{\cdot})$	the estimate of a parameter
$(\cdot)_j$	the j th element of a vector
$[\cdot]_{ij}$	the ij th element of a matrix
$ \cdot $	the magnitude of a complex quantity
$\ \cdot\ $	the Euclidean norm of a vector
\otimes	the Kronecker matrix product
$E\{\cdot\}$	the statistical expectation operator
$\text{diag}\{\mathbf{a}\}$	the diagonal matrix constructed from elements of \mathbf{a}
$\text{tr}\{\cdot\}$	the trace of a matrix
$\det\{\cdot\}$	the determinant of a matrix
$\text{gcd}(a, b)$	the greatest common divider of a and b
j	$\sqrt{-1}$

$\text{vec}\{\cdot\}$	the operator stacking the columns of a matrix on top of each other
M	number of transmitter antennas
N	number of receiver antennas
H	the channel matrix of a MIMO system
x	the input signal vector
y	the received signal vector
w	the noise vector
ρ	the signal to noise ratio per receiver antenna
X	the input signal matrix
Y	the output signal matrix
W	the noise matrix
s	the original constellation signal
C	the channel capacity
\mathcal{I}	the mutual information
H	the entropy
λ	eigenvalue
R	the data rate
P	Probability
I	Identity matrix
\hat{s}	the equalized signal
\mathcal{E}	error covariance matrix
μ	cardinality of signal constellation
C_k	linear STBC associated with s_k
A_k, B_k	linear dispersion STBC associated with s_k and s_k^*
C	vectorized linear STBC $\left(\text{vec}(C_1) \ \cdots \ \text{vec}(C_K) \right)$

\mathcal{A}	vectorized linear dispersion STBC for \mathbf{s} , $\left(\text{vec}(\mathbf{A}_1) \cdots \text{vec}(\mathbf{A}_K)\right)$
\mathcal{B}	vectorized linear dispersion STBC for \mathbf{s}^* , $\left(\text{vec}(\mathbf{B}_1) \cdots \text{vec}(\mathbf{B}_K)\right)$
γ_i	SNR for i th bit
\mathcal{H}	coded channel matrix
E_s	average symbol energy
ϕ_{\min}, ϕ_{\max}	lower and upper bounds for the determinant of a STBC
$\mathcal{T}(\boldsymbol{\alpha}, K, J)$	Toeplitz STBC
d	diversity gain
L	the number blocks over which a multi – block code spans
\mathbb{Z}	the ring of integers
\mathbb{Q}	the field of rational numbers
\mathbb{R}	the field of real numbers
\mathbb{C}	the field of complex numbers
$\psi(m)$	the Euler function over \mathbb{Z}
ζ_m	the m th root of unity
$\mathbb{Z}[\zeta_m]$	the cyclotomic ring generated by \mathbb{Z} and cyclotomic numbers ζ_m
$\mathbb{Q}(\zeta_m)$	the field generated by \mathbb{Q} and ζ_m
\mathbf{T}	transformation matrix $\frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{I} & j\mathbf{I} \\ \mathbf{I} & -j\mathbf{I} \end{pmatrix}$
\mathbf{D}	DFT matrix
\mathbf{P}	permutation matrix $\begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{I}_{T-1} & \mathbf{0} \end{pmatrix}_{T \times T}$

Contents

Abstract	iv
Acknowledgements	vi
List of Acronyms	vii
List of Notations	ix
1 Introduction	1
1.1 MIMO Communication System	1
1.2 Motivation	6
1.3 Contribution and Thesis Outline	9
1.4 Publication	13
2 Theoretical Background	14
2.1 MIMO Communication System and STBC	14
2.2 Channel Capacity for Fading MIMO Systems	16
2.3 Performance Measures	18
2.4 Detection	25
2.5 Classes of STBC	31

3	Toeplitz STBC: Properties and Application to MISO Communication Systems	38
3.1	Design Criterion for Full-Diversity STBC with Linear Receivers	39
3.2	Toeplitz Space-Time Block Codes and Their Properties	48
3.3	Toeplitz STBC Applied to a MISO System with a Linear Receiver . .	58
3.4	Optimal Toeplitz STBC Design for MISO System with ML Detector .	61
3.5	Numerical Experiments	69
3.6	Conclusion	83
4	Diversity Gain Analysis of a Linear Receiver and Multiple Block Transmission	85
4.1	Introduction	86
4.2	Optimal Linear STBC for Single-Block Transmission and Performance Analysis	89
4.3	Optimal Linear STBC for Multi-Block Transmission with Linear MMSE Receivers	93
4.4	Performance Analysis: Multi-Block Transmission Using Optimum Code	99
4.5	Generation and Detection of the Optimum Multi-Block Code	107
4.6	Numerical Experiments	111
4.7	Conclusion	115
5	Multi-Block Transmission for Linear Receivers	117
5.1	Transmission Scheme and System Performance Analysis	118
5.2	Design Example and Simulations	123
5.3	Alternative Choice of the Precoder	129
5.4	Complexity Analysis	131
5.5	Conclusion and Discussion	132

6	Cyclotomic STBC – a linear dispersion FRFD unitary trace-orthogonal code	133
6.1	Introduction	134
6.2	Channel model with linear dispersion codes	136
6.3	Good structures for LD codes: Unitary trace-orthogonality	137
6.4	Construction of trace orthonormal LD codes	145
6.5	Design of full diversity LD codes	150
6.6	Design of full diversity linear space-time block codes for $N < M$	155
6.7	Design examples and simulations	158
6.8	Conclusion	163
7	Conclusions and Future Work	165
7.1	Conclusions	165
7.2	Future Work	168
A	Proof of Theorem 4.2	170
B	Proof of Theorem 4.3	178
C	Proof of Theorem 6.3	184
D	Proof of Lemma 6.5	186
E	Proof of Theorem 6.4	188

List of Figures

1.1	A SISO communication system with multipath propagation.	2
1.2	A SIMO communication system having N receiver antennas.	2
1.3	A MISO communication system having M receiver antennas.	3
1.4	A MIMO communication system having multiple antennas at both ends.	3
1.5	A MIMO communication system.	5
2.1	Probability for a SIMO channel in deep fading.	20
2.2	Diversity v.s. multiplexing gain tradeoff curve	25
2.3	Diversity v.s. multiplexing gain tradeoff curve for a MISO system.	26
3.1	The average BER performance of the proposed Toeplitz STBC when signals are selected from different constellations.	70
3.2	The average BER performance of the proposed Toeplitz STBC for different K	70
3.3	The average BER performance of the proposed Toeplitz STBC for the MISO system with different number of transmitter antennas.	72
3.4	The average bit error rate comparison of the proposed Toeplitz STBC with i) $\mathbf{B} = [\mathbf{I}_J, \mathbf{0}_{M-J}]$, ii) \mathbf{B}_{op} and iii) $\tilde{\mathbf{B}}_{\text{op}}$. The performances are shown for both linear ZF detectors and ML detectors.	75

3.5	The average bit error rate comparison of the proposed Toeplitz STBC with $\mathbf{B} = \mathbf{I}_M$ and $\tilde{\mathbf{B}}_{\text{op}}$. The performances are shown for both linear ZF detectors and ML detectors.	76
3.6	The average bit error rate comparison of the proposed Toeplitz STBC with other STBC of unit rate.	78
3.7	The average bit error rate comparison of the proposed Toeplitz STBC with the orthogonal STBC.	80
3.8	The average symbol error rate comparison of the proposed Toeplitz STBC with the orthogonal STBC.	80
3.9	Diversity-multiplexing gain tradeoff curves for the proposed Toeplitz STBC and the orthogonal STBCs.	82
4.1	$\mathcal{P}_{\text{mmin}}$ vs SNR for different L for the case of $M = N = 1$, and $R = 2$ bits per channel use.	102
4.2	$\mathcal{P}_{\text{mmin}}$ vs SNR for different L for the case of $M = N = 2$, and $R = 4$ bits per channel use.	102
4.3	$\mathcal{P}_{\text{mmin}}$ vs L	103
4.4	Diversity vs L	103
4.5	Theoretical evaluation of $\mathcal{P}_{\text{mmin}}$ w.r.t. SNR for $L \rightarrow \infty$	106
4.6	Single-block transmission. $M = 2, N = 2, 3, 4$	112
4.7	Single-block transmission. $M = N = 2, 4, 8$	112
4.8	Multi-block transmission performance, $M = N = 1, R = 2$ bits pcu.	114
4.9	Performance of different STBC, $M = N = 2, R = 4$ bits pcu.	114
5.1	Combined precoders/detectors for a multi-block MIMO system.	119
5.2	Plotting of $\bar{s}_\gamma - \bar{s}_k$	125
5.3	The constellation set generated from two 4-PSK.	127

5.4	Performance comparisons between single block, multi-block, and the proposed schemes.	127
5.5	16-QAM	130
5.6	16-PSK	130
5.7	SER comparison of the three constellations.	131
6.1	The error performance comparison of our new code with the current available codes in [1,37,17,107].	160
6.2	The error performance comparison of our new code with the current available codes in [39,22,64].	160
6.3	Comparison of our code with the code in [22].	162
6.4	Comparison of our code with the code in [64].	162

Chapter 1

Introduction

1.1 MIMO Communication System

In conventional wireless communication systems, the transmitter and receiver each have a single antenna, and such systems are often referred to as single input single output (SISO) systems. In a typical wireless environment, the presence of reflectors creates multiple propagation paths from the transmitter to the receiver, as depicted in Fig. 1.1. As a result, the receiver sees multiple copies of the transmitted signal, each experiencing different attenuation, delay and phase shift while travelling from transmitter. This can result in destructive interference at the receiver. Strong destructive interference is frequently referred to as a deep fade and may result in the failure of communication due to poor channel signal-to-noise ratio (SNR).

One way in which more reliable communication can be achieved is by introducing multiple antennas at the receiver end. The resulting system is called a single input multiple output (SIMO) system, as shown in Fig. 1.2. In such a system, multiple copies of the transmitted signal are gathered at the receiver using different antennas. Usually, it is desirable to separate the multiple receiver antennas with sufficient

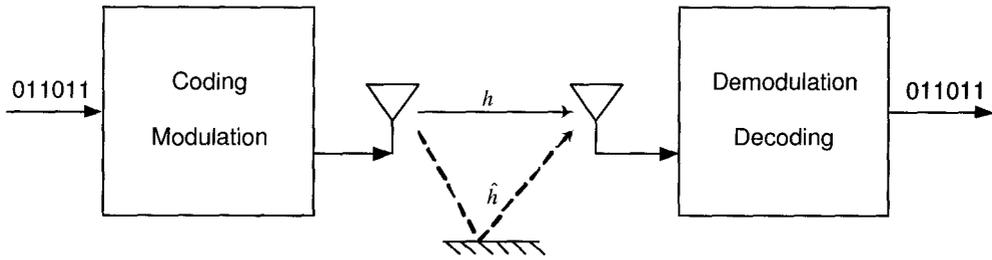


Figure 1.1: A SISO communication system with multipath propagation.

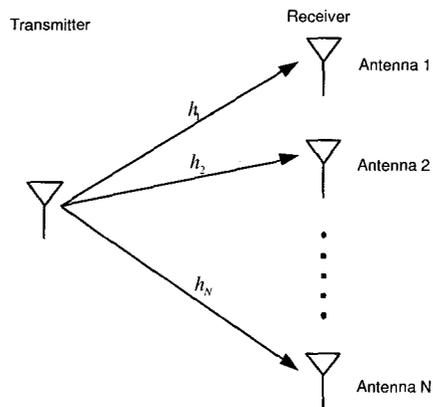


Figure 1.2: A SIMO communication system having N receiver antennas.

distance so that the fading paths seen by different antennas are approximately independent. In this case, the probability that all the paths are in a deep fade is small, and this facilitates a reliable transmission of the original signal. The gain obtained here is called *diversity* gain. Compared to the single antenna system, a lower error probability can be achieved with a SIMO system at the same SNR.

Similarly, a multiple input single output (MISO) system has multiple antennas at the transmitter side, which is shown in Fig. 1.3. When the transmitter antennas are located sufficiently far apart, diversity can be obtained by transmitting signals over the multiple independent fading paths between the transmitter and receiver antennas.

One of the most significant recent developments in wireless communications is the

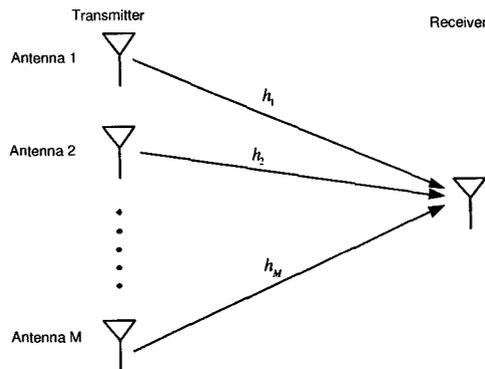


Figure 1.3: A MISO communication system having M receiver antennas.

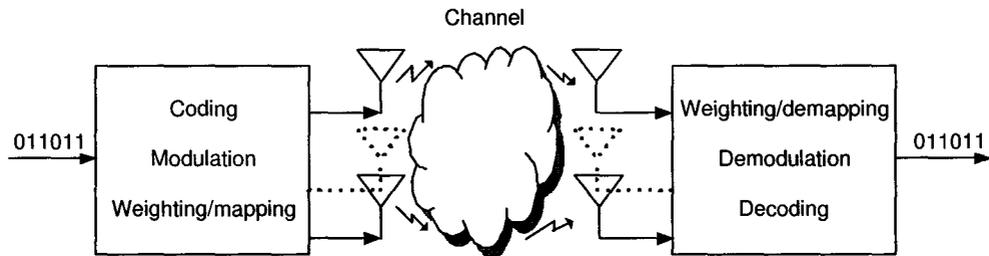


Figure 1.4: A MIMO communication system having multiple antennas at both ends.

multiple input multiple output (MIMO) systems, in which there are multiple antennas at both ends of the link as shown in Fig. 1.4. The multiple antennas provide the potential to achieve both high rate and reliable transmission for information signals. It has been shown [27, 92] that in a rich scattering environment, a MIMO system enables substantially increased capacity over the other systems mentioned above.

Due to the advantages a MIMO system possesses, the analysis and designs for such communication systems is the main focus of this thesis. In the following, the channel models and related assumptions that will be used in this thesis will be introduced. They have been widely employed in the literature of MIMO communications.

- Channel knowledge: The fading channels are modelled by random variables.

The receiver is assumed to know the realizations perfectly, while the transmitter knows only the statistical characteristics of the random channel. Channel state information can be obtained at the receiver by sending known training signals, and its accuracy depends on the received SNR and coherence time. Here, the channel is assumed to vary slowly enough to enable perfect estimation of the channel to be assumed. The characteristics of the class of fading channels that will be considered are discussed as follows.

- Flat fading: When the bandwidth of the transmitted signal is sufficiently narrow, or the symbol period is sufficiently longer than the essential duration of the channel impulse response, all frequency components of the signal will experience the same fading. In that case, the wireless link between the pair transceiver antennas can be simply modelled by a complex scalar, called a fading coefficient.
- Rayleigh fading: When there are many scatters in the environment, the signals experience rich scattering. Based on central limit theory [74], each fading coefficient can be well-modelled by a Gaussian random variable. If there is no specular component, then each such a variable will have zero mean and phase uniformly distributed between 0 and 2π radians. The envelope of each channel fading coefficient will therefore be Rayleigh distributed.
- Spatially IID: In a rich scattering environment, the transmitter and receiver antennas are assumed to be located with sufficient distances for the channel fading coefficients to be modelled as being independent and identically distributed (IID) in space.
- Block fading: The time varying nature of the channel is modelled as block fading. That is, the channel is assumed to remain fixed for a certain duration, called the *coherence time*, and then may change to another independent state.

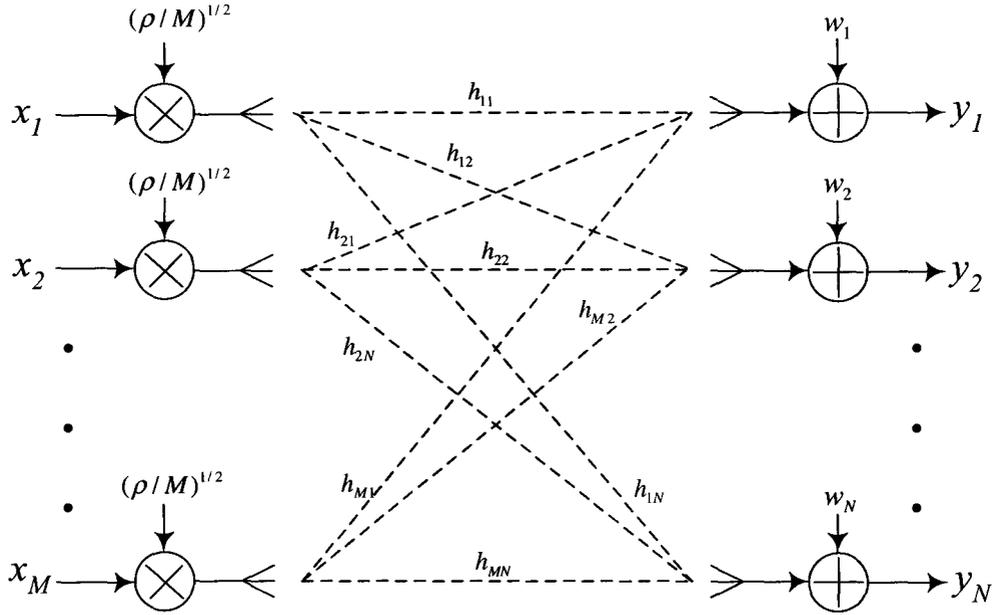


Figure 1.5: A MIMO communication system.

In reality, independence of the fading blocks can be realized by interleaving, where the symbols to be transmitted in each block are interleaved and spaced sufficiently far apart in time.

Based on the above assumptions, a mathematical description of a MIMO communication system will now be established. Consider a typical MIMO system with M transmitter and N receiver antennas, cf. Fig. 1.5. Based on the flat fading assumption, the fading coefficient between the m th transmitter and the n th receiver antennas is denoted by a complex scalar, h_{nm} . Altogether, MN fading coefficients constitute the $N \times M$ channel matrix, denoted by \mathbf{H} , with h_{nm} being the nm th entry. At a given time slot, M signal symbols (x_1, \dots, x_M) with constrained (total) power are transmitted, one from each of the M transmitter antennas. A linear combination of the M symbols will be received by each of the N receiver antennas, along with

additive noise. The input-output relation is, therefore,

$$\mathbf{y} = \sqrt{\frac{\rho}{M}} \mathbf{H} \mathbf{x} + \mathbf{w} \quad (1.1)$$

where \mathbf{x} is the $M \times 1$ input signal vector, \mathbf{y} is the corresponding $N \times 1$ received signal vector, \mathbf{w} is the additive noise which is assumed to be of IID Gaussian distribution with zero mean and unit variance, and ρ is the SNR per receiver antenna.

Under the block fading assumption, the channel condition is fixed for certain time period, say T time slots. Then, T signal vectors can be transmitted under the same channel conditions. We stack the transmitted T signal vectors as the columns of a $M \times T$ matrix \mathbf{X} , and correspondingly define a $N \times T$ received signal matrix \mathbf{Y} . This results in the following block transmission system model

$$\mathbf{Y} = \sqrt{\frac{\rho}{M}} \mathbf{H} \mathbf{X} + \mathbf{W} \quad (1.2)$$

Each entry of \mathbf{X} can be an original message symbol, but more generally, it can be thought of as a (matrix) codeword. For notation clarity, we write $\mathbf{X} = \mathbf{X}(\mathbf{s})$, where \mathbf{s} is the vector of original message symbols to be transmitted in the block, and is assumed to have zero mean and normalized (unit) variance. The codeword $\mathbf{X}(\mathbf{s})$ is distributed among M transmitter antennas in space and across T slots in time, and is transmitted block by block. Hence, the set of codewords $\{\mathbf{X}(\mathbf{s})\}_{\mathbf{s}}$ is said to be a Space-Time Block Code (STBC).

1.2 Motivation

In the past decade, intensive research on STBC design has been carried out with the aim of improving the system error performance under different types of detection algorithms. Most of this effort has focused on systems in which the maximum-likelihood

(ML) detector is employed, due to its superior performance to other available detectors. In designing a reliable transmission scheme for a given rate, the code should be chosen to minimize the error probability in the detection of the signals. However, when an ML detector is employed, the detection error probability is difficult to formulate in general. Hence, the design efforts have mainly focused on the high SNR performance, where the error probability is dominated by diversity gain. Mathematically, diversity gain describes how fast the error probability decays with SNR at high SNR. As a result, maximizing diversity gain has become one of the most important criteria [91, 20] in designing STBC for ML detectors.

The simplest STBC is Alamouti's code [1], which was invented in 1998 for a two transmitter antenna system. This 2×2 code matrix has an orthogonal structure and belongs to the family of orthogonal STBCs [90, 88, 52, 101, 51]. In a system employing an orthogonal STBC, the maximum diversity gain can be achieved. In addition, the orthogonality property reduces the computational cost of ML detection to that of linear detection. This advantage holds for any type of input signal. Unfortunately, orthogonality also restrains the highest achievable symbol transmission rate to be less than one. As an example [51], for a MIMO system having $M \geq 3$, the highest achievable symbol rate by orthogonal codes is only $\frac{3}{4}$ symbols per channel use. (Actually, there is a fundamental tradeoff [114] between the diversity gain and a measure of the rate that can be achieved by a STBC.)

Efforts to achieve both high rate and good error performance were taken in [37], where the objective of the STBC design is to maximize the Gaussian mutual information. However, the presented codes do not guarantee the maximum diversity gain. Another approach is to design a STBC that has full symbol rate and achieves full diversity (FRFD) [39, 17, 22, 64] when ML detection is employed. Full symbol rate is achieved when the average symbol transmission rate equals $\min\{M, N\}$ symbols

per channel use. Full diversity gain can be achieved when the total spatial degrees of freedom offered by a MIMO system are utilized for diversity purposes. In general, the design of the FRFD codes depends on the properties of the input signals and the available FRFD codes only render full diversity gain for certain structured input signals. Further research in STBC design has led to the FRFD codes that have non-vanishing determinant [107, 28, 5, 79, 48, 24, 73], where the minimum distance between two different code words has been maximized, which has further improved the system performance. However, the superior performance is only applicable to certain types of signals, and is only achieved by using ML detection, which has a high computational cost.

As has been discussed above, numerous STBCs have been proposed to achieve reliable communication with the optimality in certain aspect. However, in general, the structure that a good STBC should possess has not yet been resolved.

It should be noted that detection of a MIMO system demands much higher computational cost compared to a single antenna system in general. Despite the advantages that the optimal ML detector possesses in the system performance perspective, its application is restricted to systems having a small number of antennas. The sub-optimal detection architecture based on decision feedback equalization (DFE) is less costly, but suffers from error propagation. On the other hand, a linear receiver is practically attractive for its simplicity. However, the desirable properties of a STBC for a system with a linear receiver have yet to be established, possibly due to the complicated mathematical formulations and moderate achievable performance compared to the system that employs ML detection. The problems of determining the optimal code for a linear receiver and determining the performance limits for such a system have been left open.

1.3 Contribution and Thesis Outline

In this thesis, the design of STBCs for the MIMO systems that employ linear receivers is studied. A family of STBCs, named *unitary trace-orthogonal* codes, is presented. These codes are optimal in the sense that they minimize the detection error probability for a linear receiver [61, 55] when the MIMO system satisfies $N \geq M$ and the signals are transmitted at full symbol rate. Applying the optimal code structure to such a system, the maximum diversity gain achievable by a linear receiver is analyzed and proved. Several transmission schemes are proposed for improving the diversity gain for linear receivers. Specifically, when a MISO communication system is considered, a special unitary trace-orthogonal code is presented as the first non-orthogonal full diversity STBC for linear receivers. Further analysis reveals that the code can enable a linear receiver to approach the optimal diversity v.s. multiplexing gain tradeoff that was previously only achievable with the use of more sophisticated detectors. The code structure is also successfully applied in solving several previously intractable problems for MIMO systems that employ an ML detector.

Some theoretical background on the MIMO communication systems considered in this thesis is provided in Chapter 2. The unitary trace-orthogonal code [61, 55] is also introduced in that chapter, along with its optimality properties when applied to the MIMO ($N \geq M$) systems that employ linear receivers.

The contributions of this thesis are provided through Chapters 3–6, and are outlined below:

- Chapter 3: Unitary Trace-orthogonal code in a MISO system [60]

In this chapter, the problem of designing full diversity STBCs for a MISO communication system that has a linear receiver is examined. The chapter begins by establishing a mathematical system model for a MISO channel. Based on

that, a condition on STBCs that ensures full diversity when the code is detected by linear receivers is proposed and proved. This is, thus far, the first condition for achieving full diversity by a linear receiver.

Applying the proposed design condition, a Toeplitz STBC is proposed which is a member of unitary trace-orthogonal family. This code is the first non-orthogonal code that enables a linear receiver to achieve full diversity gain. Analysis shows that the symbol rate associated with this code asymptotically approaches unity (i.e., full symbol rate for a MISO system), which is higher than the rates achievable by orthogonal codes when $M \geq 3$. In addition, the code is proved to achieve the optimal Diversity v.s. Multiplexing gain tradeoff for certain signalling schemes. This is the first code that we are aware of to enable a linear receiver to achieve the optimal tradeoff for a MISO system having arbitrary number of transmitter antennas.

In the rest of the chapter, the Toeplitz STBC is applied to a MISO system in which the channel fading coefficients are spatially correlated and ML detection is employed. The beamforming matrix that minimizes the pairwise error probability is designed for such a system. It is shown that by exploiting the special properties of the Toeplitz code, the complicated optimization design problem can be transformed into a convex problem. Several numerical examples are provided in the end of the chapter to support the theoretical analysis.

- Chapter 4: Diversity analysis and multi-block transmission [56]

In this chapter, the spacial diversity gain achievable by a linear receiver is analyzed. This is followed by the introduction of a multi-block transmission scheme that improves the diversity gain of a linear receiver.

The inferior error performance of a linear receiver to that of an ML receiver

has previously been shown by numerical experiments. Theoretically, the performance gap between these two receivers is still unknown. One way to carry out a theoretical comparison is to analyze the maximum diversity gain achievable by each receiver. When an ML receiver is employed, the full diversity gain (MN) of a MIMO system can be exploited. On the other hand, the performance limit for a linear receiver has yet not been analyzed. In [61], unitary trace-orthogonal code is proposed which provides the best performance of a linear receiver. Therefore, the diversity gain of a system with a unitary trace-orthogonal code and a linear receiver must be the maximum gain achievable by a linear receiver. In this chapter, by examining the minimum bit error probability (BEP) associated with the unitary trace-orthogonal code, the maximum diversity for a linear receiver is analyzed and shown to be equal to $N - M + 1$ for $N \geq M$. This result provides a theoretical measure of the performance gap between a linear and an ML receivers.

In the second part of this chapter, a multi-block transmission scheme is presented for improving the system performance of a linear receiver. In this scheme, the signals are coded over several independent channel realizations. By taking advantage of time diversity, each signal symbol may experience an increased number of path gains. As a result, the impact of a particular deep fading path can be reduced. For such a system, the optimal code that minimizes the detection error probability of a linear receiver is proposed, and this code structure is proved to be both necessary and sufficient in achieving the optimality. The code is a generalized version of unitary trace-orthogonal code and is named multi-block unitary trace-orthogonal code. Through both theoretical analysis and numerical experiments, it is shown that the system performance can be significantly improved. The combination of multi-block transmission and a linear

receiver can thus enjoy a high diversity gain with a relatively low computational cost. This system may outperform those systems employing an ML receiver and single block STBCs in the cases when a large number of transmitter antennas is involved, or when the signals are picked from higher order constellations.

- Chapter 5: A multi-block transmission for linear receivers [58]

A transmission scheme specially designed for a linear receiver is proposed in this chapter. This scheme takes greater advantage of time diversity. In particular, the achievable diversity gain increases linearly with the number independent blocks over which a code spans. The idea originates from the observation that the diversity gain of a linear receiver is determined by the difference between the number of transceiver antennas, $(N - M)$. This implies that a taller channel matrix renders higher diversity gain. In the proposed multi-block transmission scheme, the channel matrices corresponding to different realizations are stacked up, resulting in a tall equivalent channel. Simulations show a significantly improved performance for the proposed system. The complexity involved with the linear equalizer remains the same as that for the single block transmission.

- Chapter 6: Unitary trace-orthogonality and ML detection [113]

The importance of unitary trace-orthogonal codes is not constrained to their applications to linear receivers. In this chapter, a system employing an ML receiver is investigated and the problem of designing full diversity STBCs is examined. This problem is generally difficult due to the existence of a large number of unknown parameters. In this chapter, by constraining the design parameters within the family of unitary trace-orthogonal codes, the design problem can be simplified. As a result, a systematic construction of full diversity STBCs is presented based on number theory.

The work presented in this thesis is concluded in Chapter 7, along with a discussion on future work.

1.4 Publication

The contributions of this thesis have been published or submitted in four journal papers [61, 113, 60, 56], one book chapter [112], and eight conference papers [54, 109, 110, 53, 111, 57, 58, 59]. Two full journal papers [61, 113] have been published in *IEEE Transactions on Signal Processing*, and [61] received the *2007 Young Author Best Paper Award* from IEEE Signal Processing Society. The work on the Toeplitz STBC [60] has been accepted for publication as a full paper in the journal of *IEEE Transactions on Information Theory*. The work on diversity analysis and multi-block transmission has been submitted [56] to the same journal with a length of 35 pages, and it is currently under the second round of revision. Parts of the results have also been presented at several international conferences, including *2008 International Conference on Acoustics, Speech, and Signal Processing* at Las Vegas, *2006 IEEE International Symposium on Information Theory* at Seattle, and *2005 IEEE International Symposium on Information Theory* at Adelaide, Australia. The results on linear dispersion codes for ML detection has also appeared in Chapter 5 of the book “Space-Time Processing for MIMO Communications”, jointly edited by A. B. Gershman and N. D. Sidiropoulos, 2005.

Chapter 2

Theoretical Background

The design of Space-Time block codes is the main focus of this thesis. In order to simplify the exposition of the proposed design techniques, some theoretical background will be introduced in this chapter. The chapter begins with a discussion of the MIMO communication system model upon which the designs are performed. This discussion includes a brief review of the channel capacity, followed by a discussion of several performance measures for STBC designs, including diversity gain and a fundamental tradeoff in the system performance. A brief review of some detection architectures for MIMO systems is also provided. Finally, two commonly employed classes of STBC are discussed, and this leads to the introduction of the unitary trace-orthogonal code and a discussion of its properties.

2.1 MIMO Communication System and STBC

As discussed in Chapter 1, the output of a MIMO communication system that has a flat fading channel can be expressed as follows

$$\mathbf{y} = \sqrt{\frac{\rho}{M}} \mathbf{H} \mathbf{x} + \mathbf{w} \quad (2.1)$$

Based on the assumptions provided in Chapter 1, the $N \times M$ complex channel matrix \mathbf{H} is random and has IID circularly symmetric complex Gaussian entries with zero mean and unit variance. It is also assumed that perfect knowledge of the channel state information is available at the receiver side and that the transmitter only knows the distribution of the channel. At each signally instant, an $M \times 1$ signal vector \mathbf{x} is transmitted through the channel, and received as an $N \times 1$ vector \mathbf{y} . The signal vector \mathbf{x} is assumed to have zero mean and a normalized (unit) variance. The receiver also sees additive white Gaussian noise \mathbf{w} , of which the elements are circularly symmetric complex Gaussian having zero mean and unit variance. The SNR at each receiver is denoted by ρ .

The signal vector \mathbf{x} may be the original signal \mathbf{s} , or may be part of a codeword representing \mathbf{s} . Hence, it is more precise to write $\mathbf{x} = \mathbf{x}(\mathbf{s})$. Suppose the channel remains constant for a period of time, say T time slots. Then during this time, there are T vectors of $\mathbf{x}(\mathbf{s})$ that can be transmitted, and these vectors can be represented by the columns of an $M \times T$ matrix $\mathbf{X}(\mathbf{s})$. Thus, the following block transmission model can be obtained

$$\mathbf{Y} = \sqrt{\frac{\rho}{M}} \mathbf{H} \mathbf{X}(\mathbf{s}) + \mathbf{W} \quad (2.2)$$

The matrix $\mathbf{X}(\mathbf{s})$ is a codeword for a STBC and the design of such codes is the main focus of this thesis. Before a discussion is provided on that topic, some related background about MIMO communication systems will be introduced in the following.

2.2 Channel Capacity for Fading MIMO Systems

The channel capacity for a MIMO system that has a random flat fading channel will be discussed in this section. We examine the following MIMO communication model

$$\mathbf{y} = \sqrt{\frac{\rho}{M}} \mathbf{H} \mathbf{x} + \mathbf{w} \quad (2.3)$$

where the channel realization is available at the receiver, but the transmitter only knows its distribution. Then, the mutual information between the input and output is [11, 92]

$$\mathcal{I}(\mathbf{x}; \mathbf{y}, \mathbf{H}) = H(\mathbf{y}|\mathbf{H}) - H(\mathbf{y}|\mathbf{x}, \mathbf{H}) \quad (2.4)$$

where $\mathcal{I}(\cdot)$ and $H(\cdot)$ stand for “mutual information” and “entropy”, respectively.

The mutual information in Eq. (2.4) is a function of the random channel, and hence it is also a random variable. Depending on the latency requirement, different measures on the distribution of the mutual information yield appropriate notions of reliable communication. For the MIMO systems that we will consider, there are two relevant capacity definitions, namely, ergodic capacity and capacity versus outage.

2.2.1 Ergodic Capacity

When the transmission time is sufficient long and the channel fading process is assumed to be an ergodic process, the fundamental limit on the rate of reliable communication is the ergodic capacity,

$$\max_{p(\mathbf{x})} : \mathbb{E}_{\mathbf{H}} \{ \mathcal{I}(\mathbf{x}; \mathbf{y}, \mathbf{H}) \} \quad (2.5)$$

where $\mathbb{E}_{\mathbf{H}}$ denotes the expectation over the random channel \mathbf{H} , and $p(\mathbf{x})$ is the distribution of the input \mathbf{x} . The input distribution that achieves this maximum is a

circularly symmetric complex Gaussian distribution with zero mean and covariance \mathbf{Q} . This leads to the following expression of the ergodic capacity [92]

$$\max_{\mathbf{Q}} : E_{\mathbf{H}} \left\{ \log_2 \det \left(\mathbf{I} + \frac{\rho}{M} \mathbf{H} \mathbf{Q} \mathbf{H}^H \right) \right\} \quad (2.6)$$

where \mathbf{I} denotes the identity matrix. In the case of IID fading described in Chapter 1, the optimal transmit covariance \mathbf{Q} is a scaled identity matrix, and the capacity can be written as [92],

$$C = E_{\mathbf{H}} \left\{ \log_2 \det \left(\mathbf{I} + \frac{\rho}{M} \mathbf{H} \mathbf{H}^H \right) \right\} \quad (2.7)$$

The quantity in Eq. (2.7) is a constant positive number. Hence, the ergodic capacity is a capacity with classical meaning, i.e., when the transmission data rate is lower than the capacity, there exists a code such that the error probability is exponentially decaying with the transmission length, and at rates above the capacity, it is impossible to have reliable data transmission.

2.2.2 Outage Capacity

In applications in which low latency is required, such as interactive communications, the ergodicity requirement that the transmission time is sufficiently long will be violated. The instantaneous channel capacity is a random variable and it depends on the particular realization of the channel parameters. When the channel realization is not available at the transmitter, it is possible that the transmission data rate exceeds the instantaneous mutual information no matter how low the rate may be. Therefore, the error probability does not exponentially decay with the increase of the transmission length. The wireless link in this case is said to be in an outage state. In this case, one can talk about the relation between outage probability and the supportable rate.

Given a transmission data rate R , the outage probability is defined as [92]

$$P_{\text{outage}}(R) = \inf_{\mathbf{Q}} : \text{Prob} \left\{ \log_2 \det \left(\mathbf{I} + \frac{\rho}{M} \mathbf{H} \mathbf{Q} \mathbf{H}^H \right) < R \right\} \quad (2.8)$$

The capacity in this case does not have the classical meaning. Accordingly, an outage capacity is defined as the maximum data rate that can be supported by a MIMO channel for a prescribed outage probability.

Since the channel capacity describes the amount of information that can be reliably transmitted through a communication system, capacity based criteria are widely employed in the design of STBCs [92, 37, 113].

2.3 Performance Measures

The discussion in the previous section was focused on the achievable information rate of a MIMO communication system. In the design of STBCs, systems having a fixed transmission data rate are frequently considered. For such systems, the design of STBCs is usually aimed at achieving low probability of error. In this section, two commonly employed criteria for the design of STBCs which are based on error probability will be discussed.

2.3.1 Error Performance: Diversity Gain

In Chapter 1, it has been discussed that the reliability of wireless communications very much depends on the path SNR. When a transmission path suffers from a deep fade, errors are likely to occur. A simple way to combat this is to transmit information signals through multiple independent paths so that the probability that these paths are simultaneously in deep fades is exponentially decreased. The performance of the transmissions can therefore be dramatically improved. This technique is called

diversity.

Diversity can be realized in a variety of ways. For multiple antenna systems, *space* diversity can be realized by spacing the antennas sufficiently so that the fading paths between the transmitter and receiver are statistically independent. When the channel is of the fast fading type, signals can be coded or interleaved to achieve *time* diversity. The information symbols are dispersed over several coherence intervals, and they experience independent channel states. Similarly, *frequency* diversity can be exploited in frequency selective channels.

To obtain further insights on how diversity influences on the reliability of a communication system, we take a simple example here. Consider a flat fading SIMO channel that has N receiver antennas. The single transmitter antenna is fed with a symbol x at an instant of time. Correspondingly, the signals at the N receiver antennas can be written in the following vector form

$$\mathbf{y} = \sqrt{\rho} \mathbf{h}x + \mathbf{w} \quad (2.9)$$

where \mathbf{h} and \mathbf{w} are statistically independent, and they have IID Gaussian elements distributed as $\mathcal{CN}(0, 1)$. In detecting x , a sufficient statistic of \mathbf{y} can be obtained in a scalar form as

$$\frac{\mathbf{h}^H}{\|\mathbf{h}\|} \mathbf{y} = \sqrt{\rho} \|\mathbf{h}\|x + \frac{\mathbf{h}^H}{\|\mathbf{h}\|} \mathbf{w} \quad (2.10)$$

where the noise $\frac{\mathbf{h}^H}{\|\mathbf{h}\|} \mathbf{w}$ is $\mathcal{CN}(0, 1)$. The accuracy of detecting x is determined by the received SNR, i.e., $\rho\|\mathbf{h}\|^2$. A detection error is likely to occur when $\rho\|\mathbf{h}\|^2$ is a small quantity. The probability of this event is

$$P\left(\|\mathbf{h}\|^2 < \frac{\epsilon}{\rho}\right) \quad (2.11)$$

where ϵ is a small positive constant. Under the assumption that the channel coefficient \mathbf{h} is circularly symmetric complex IID Gaussian, the random variable $\|\mathbf{h}\|^2$ is thus the

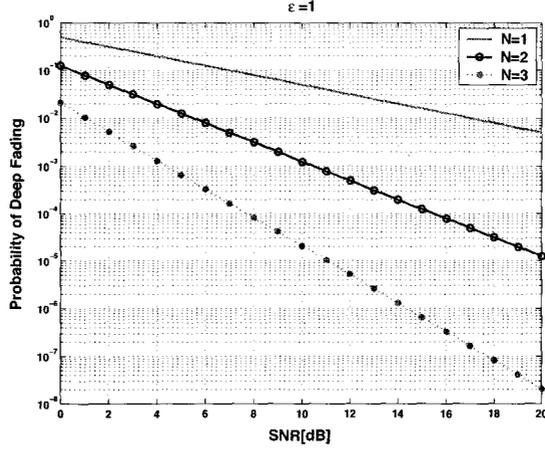


Figure 2.1: Probability for a SIMO channel in deep fading.

sum of $2N$ independent real Gaussian random variables. Hence, it has a Chi-square distribution [72] with $2N$ degrees of freedom. The probability density function (PDF) of $z_h \triangleq \|\mathbf{h}\|^2$ is [72]

$$p(z_h) = \frac{1}{2^N(N-1)!} z_h^{N-1} e^{-z_h/2} \quad (2.12)$$

When z_h is small,

$$p(z_h) \approx \frac{1}{2^N(N-1)!} z_h^{N-1} \quad (2.13)$$

Now if we apply Eq. (2.13) to Eq. (2.11), the probability in Eq. (2.11) can be written as

$$P\left(\|\mathbf{h}\|^2 < \frac{\epsilon}{\rho}\right) \approx \int_0^{\frac{\epsilon}{\rho}} \frac{1}{2^N(N-1)!} x^{N-1} dx = \frac{\epsilon^N}{2^N N!} \rho^{-N} \quad (2.14)$$

The probability in Eq. (2.14) is plotted in Fig. 2.1 for various values of N . We observe that the probability of the channel being in a deep fade is exponentially decaying with N , the number of receiver antennas. This is shown in the figure by the slope of the probability curve. This gain is obtained by diversity, and is called diversity gain.

2.3.1.1 Definition

In general, diversity gain describes how fast the error probability P_e decays with the increase of SNR. Mathematically, it is defined as the negative power of SNR in the expression of P_e at high SNRs, i.e., [96],

$$d = - \lim_{\rho \rightarrow \infty} \frac{\log P_e}{\log \rho} \quad (2.15)$$

We note that diversity gain is defined at high SNR. This is because only when SNR goes to infinity, the remaining factors in the error probability become negligible. Hence, the diversity gain is the gain that dominates the high SNR performance of a communication system. The definition provided in Eq.(2.15) is applicable to any communication system. For a given system, the value of diversity gain d depends on the transmission strategy, the channel model, and the receiver architecture. In the following, the diversity gain for a system having an ML receiver will be derived.

2.3.1.2 Diversity gain for ML detectors

MIMO communication systems with ML detection have been widely studied, due to their optimum performance. However, for such a system, an expression for the exact error probability is hard to obtain, due to the mathematical complexity. For that reason, the pairwise error probability (PEP) is commonly considered in analyzing the performance of an ML detector. The pairwise error probability refers to the probability of transmitting a symbol vector \mathbf{s} that is detected as \mathbf{s}' . When white Gaussian noise is encountered, this probability is controlled by the Euclidean distance between the two symbol vectors $\{\mathbf{s}, \mathbf{s}'\}$. For a MIMO system transmitting a coded signal $\mathbf{X}(\mathbf{s})$, the distance between the two points is measured by the Frobenius norm of the matrix $\mathbf{Z}(\mathbf{e}) = \mathbf{X}(\mathbf{s}) - \mathbf{X}(\mathbf{s}')$, with $\mathbf{e} = \mathbf{s} - \mathbf{s}'$. Since Gaussian noise is assumed, the PEP for a given channel realization can be mathematically described using the

Q -function [77]. Note that the channel is random with IID Gaussian entries, and we are interested in the PEP averaged over all the realizations of the random channel \mathbf{H} . This results in the following expression of the PEP as a function of the STBC [91]:

$$P(\mathbf{s} \rightarrow \mathbf{s}') = \frac{1}{\pi} \int_0^{\pi/2} \frac{d\theta}{\det\left(\mathbf{I} + \frac{\rho}{8M \sin^2 \theta} \mathbf{Z}(\mathbf{e}) \mathbf{Z}^H(\mathbf{e})\right)^N} \quad (2.16)$$

Using the fact that $\sin^2 \theta \leq 1$, an upper bound on Eq. (2.16) can be obtained as

$$\begin{aligned} P(\mathbf{s} \rightarrow \mathbf{s}') &\leq \frac{1}{2} \det\left(\mathbf{I}_M + \frac{\rho}{8M} \mathbf{Z}(\mathbf{e}) \mathbf{Z}^H(\mathbf{e})\right)^{-N} \\ &= \frac{1}{2} \prod_{m=1}^M \left(1 + \frac{\rho}{8M} \cdot \lambda_m\right)^{-N} \\ &\leq \frac{1}{2} \prod_{m=1}^r \left(\frac{\rho}{8M} \cdot \lambda_m\right)^{-N} \\ &= \frac{1}{2} \left(\frac{\rho}{8M}\right)^{-rN} \left(\prod_{m=1}^r \lambda_m\right)^{-N} \end{aligned} \quad (2.17)$$

where $r (\leq M)$ is the rank of $\mathbf{Z}(\mathbf{e}) \mathbf{Z}^H(\mathbf{e})$, and $\{\lambda_m\}, m = 1, \dots, r$ are its non-zero eigenvalues. To realize low error rate transmission, it is desirable to keep the upper bound in Eq. (2.17) low. In the following, we examine the two factors in Eq. (2.17):

- The Rank Factor r : At high SNR, the first term in Eq. (2.17) dominates, and therefore we should make the exponent, rN , known as the *diversity gain*, as large as possible. For *full diversity*, the minimum rank of $\mathbf{Z}(\mathbf{e})$ taken over *all* distinct pairs $\{\mathbf{s}, \mathbf{s}'\}$ should be maximized, i.e., the matrix $\mathbf{Z}(\mathbf{e}) \mathbf{Z}^H(\mathbf{e})$ should be of full rank, resulting in full diversity being equal to MN .
- The Determinant Factor ($\prod_{m=1}^r \lambda_m$): The second term consists of the product of the non-zero eigenvalues of $\mathbf{Z}(\mathbf{e}) \mathbf{Z}^H(\mathbf{e})$. The minimum value of $(\prod_{m=1}^r \lambda_m)^{1/r}$, taken over all distinct symbol vector pairs $\{\mathbf{s}, \mathbf{s}'\}$, is called the *coding gain* and should be made large.

The rank factor and the determinant factor are frequently employed as design criteria for STBCs that are to be used in systems equipped with an ML detector. From the above derivations, it is clear that maximizing these two factors are necessary for the minimization of PEP.

From the discussions in this subsection, we observe that by applying the general definition of diversity gain in Eq. (2.15) to a system with a STBC, an IID Rayleigh channel and an ML receiver, an expression that is specialized for such a system is obtained, and this expression is dependent on the rank of the code matrix. This illustrates the fact that for different communication systems, the expression for the diversity gain may appear in special forms.

2.3.2 Diversity v.s. Multiplexing tradeoff

2.3.2.1 Multiplexing Gain

For a communication system operating at high SNRs, diversity gain ensures that the error probability decreases rapidly with SNR. However, achieving diversity gain occupies some of the degrees of freedom offered by the system, such as those provided by the presence of multiple antennas. These degrees of freedom could also be used to provide higher rate of transmission. For example, in a MIMO system having M transmitter antennas, the symbol transmission data rate could be M symbols per channel use (pcu) if an independent symbol is transmitted from each antenna. This suggests that there may be a quantified trade-off between transmission data rate and diversity gain. To explore that possibility, we first define a notion of multiplexing gain that is compatible with the notion of diversity gain.

Firstly, we observe that channel capacity is an increasing function of SNR. Hence we will consider a transmission scheme with supportable rate that increases as SNR.

We consider a family of codes with one at each SNR level. The achievable rate by this scheme is denoted by $R(\rho)$. The multiplexing gain g is defined to reflect the change of this rate at high SNRs, i.e. [96],

$$g = \lim_{\rho \rightarrow \infty} \frac{\log R(\rho)}{\log \rho} \quad (2.18)$$

2.3.2.2 $M \times N$ MIMO channel

Diversity gain and multiplexing gain are two benefits offered by a MIMO system. These dual benefits are captured by a fundamental tradeoff of the two gains — the diversity v.s. multiplexing gain tradeoff [114]. Consider an IID Rayleigh fading MIMO channel having $T \geq M + N - 1$. Then the optimal tradeoff between (g, d) is a piece-wise linear curve joining the points [114]:

$$(k, (M - k)(N - k)), \quad k = 0, \dots, \min\{M, N\} \quad (2.19)$$

as shown in Fig. 2.2. When $g \rightarrow 0$, e.g., the transmission rate is fixed and finite, the maximum diversity gain MN is achievable as shown in Fig. 2.2. When $g = \min\{M, N\}$, the full degrees of freedom are exploited for rate gain and we have no spacial diversity gain. The diversity v.s. multiplexing gain tradeoff bridges the two extreme cases and provides a more general picture of the performance capability of a MIMO system.

2.3.2.3 $M \times 1$ MISO channel

As a special case of Section 2.3.2.2, the diversity v.s. multiplexing gain tradeoff for a MISO communication system is provided in the following. This tradeoff curve will be employed later in Chapter 3 in designing STBC for such a system.

Consider an IID Rayleigh fading MISO system having M transmitter antennas,

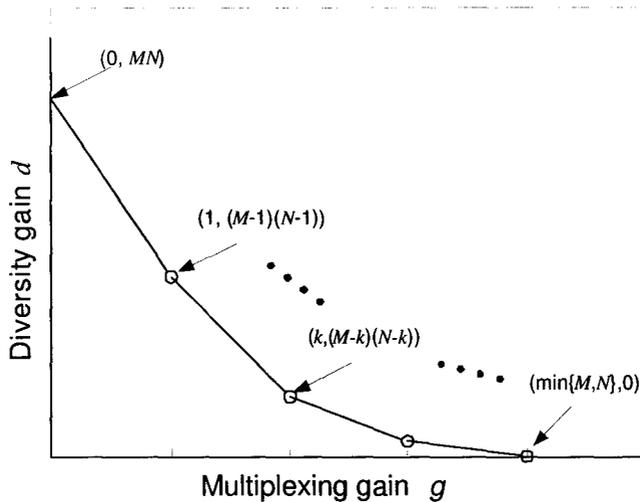


Figure 2.2: Diversity v.s. multiplexing gain tradeoff curve

and the relation of the optimal diversity and multiplexing gain is described by [96]

$$d = M(1 - g) \quad 0 \leq g \leq 1 \quad (2.20)$$

as shown in Fig. 2.3. Eq. (2.20) indicates that a MISO system provides an M -fold increase of diversity gain for all multiplexing gain.

2.4 Detection

For a MIMO communication system with perfect channel knowledge at the receiver, there are several detection architectures, including maximum likelihood detection, detection involving decision feedback equalization, and detection including linear pre-processing. In this section, we focus on describing the simple linear receiver, including its Zero-Forcing (ZF) and Minimum Mean Square Error (MMSE) versions, and the maximum likelihood receiver. These receivers will be employed throughout the thesis.

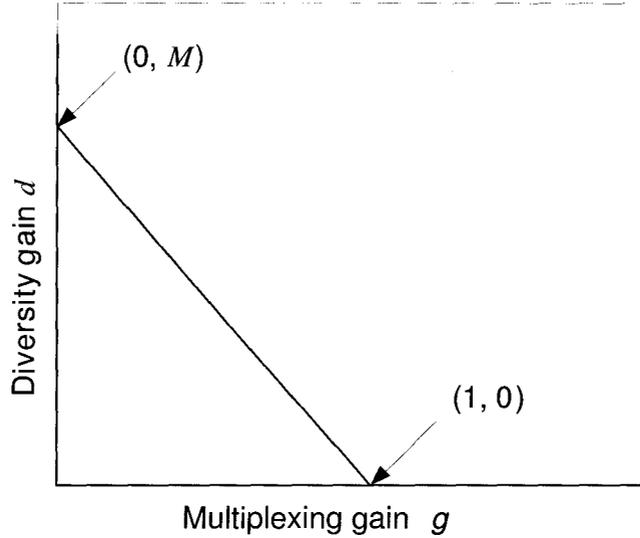


Figure 2.3: Diversity v.s. multiplexing gain tradeoff curve for a MISO system.

2.4.1 Linear ZF Receiver

Consider the signal transmitted through a MIMO channel at a particular time instant,

$$\mathbf{y} = \sqrt{\frac{\rho}{M}} \mathbf{H} \mathbf{x} + \mathbf{w} \quad (2.21)$$

where we have the following assumptions: i) \mathbf{x} is the input signal vector that has independently elements and satisfies $E\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}$; ii) \mathbf{w} has zero mean and unit variance; iii) \mathbf{x} and \mathbf{w} are independent; and iv) \mathbf{H} has full column rank (which implies $N \geq M$). The M independent elements contained in \mathbf{x} are each fed to one of the M transmitter antennas. Now, let us consider the detection of the m th element, x_m , which passes through N fading paths between the m th transmitter and all the receiver antennas. The receiver therefore sees N copies of x_m , which can be mathematically described as $\mathbf{H}(:, m)x_m$, where $\mathbf{H}(:, m)$ denotes the m th column of \mathbf{H} . Additionally, the received signal \mathbf{y} also contains the signals transmitted through other transmitter antennas, $\sum_{k=1, k \neq m}^M \mathbf{H}(:, k)x_k$, which constitute interference in the

detection of the desired m th symbol x_m .

One way of detecting x_m is to completely remove the interference from \mathbf{y} , normalize the resulting symbol, and detect it independently from the other symbols. This can be realized by projecting the received signal \mathbf{y} onto the direction orthogonal to interference space spanned by $\{\mathbf{H}(:, k), k = 1, \dots, M, k \neq m\}$. As a result, when the noise is circularly symmetric white Gaussian with zero mean, the detection error will be orthogonal to the desired signal. A projector is a linear processor, and can be denoted by a matrix \mathbf{G} . Hence, the received signal \mathbf{y} in Eq. (2.21) is processed as

$$\hat{\mathbf{x}} \triangleq \mathbf{G}\mathbf{y} = \sqrt{\frac{\rho}{M}} \mathbf{G}\mathbf{H}\mathbf{x} + \mathbf{G}\mathbf{w} \quad (2.22)$$

where $\hat{\mathbf{x}}$ denotes the equalized signal. To cancel the interferences for all x_m , $m = 1, \dots, M$, \mathbf{G} should be chosen in the way that the detection error is uncorrelated to the input signal \mathbf{s} , i.e.,

$$\mathbf{E}\{\mathbf{e}\mathbf{x}^H\} = \mathbf{0} \quad (2.23)$$

where $\mathbf{e} = \hat{\mathbf{x}} - \mathbf{x}$. Now, substituting Eq. (2.22) into Eq. (2.23) results in

$$\begin{aligned} \mathbf{E}\{\mathbf{e}\mathbf{x}^H\} &= \mathbf{E}\left\{\left(\sqrt{\frac{\rho}{M}} \mathbf{G}\mathbf{H} - \mathbf{I}\right) \mathbf{x}\mathbf{x}^H + \mathbf{G}\mathbf{w}\mathbf{x}^H\right\} \\ &= \left(\sqrt{\frac{\rho}{M}} \mathbf{G}\mathbf{H} - \mathbf{I}\right) \mathbf{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{0} \end{aligned} \quad (2.24)$$

The last step is obtained under the assumption that \mathbf{x} is independent with the noise that has a zero mean. Eq. (2.24) gives rise to

$$\sqrt{\frac{\rho}{M}} \mathbf{G}\mathbf{H} = \mathbf{I} \quad (2.25)$$

The solution to Eq. (2.25) is

$$\mathbf{G}_{\text{ZF}} = \sqrt{\frac{M}{\rho}} \mathbf{H}^\dagger \quad (2.26)$$

where \mathbf{H}^\dagger is the *Pseudo Inverse* [34] of \mathbf{H} . This processor removes all the interferences and is called a Zero-Forcing equalizer. A subscript ZF is added to \mathbf{G}_{ZF} for notational clarity. As a result, the equalized signals have the following form

$$\hat{\mathbf{x}} = \mathbf{x} + \mathbf{G}_{\text{ZF}}\mathbf{w} \quad (2.27)$$

The transmission model in Eq. (2.27) is now equivalent to M parallel scalar channels. Detection for that model can be simply carried out by employing a scalar symbol by symbol threshold detector on each channel. The receiver having a linear ZF equalizer followed by a threshold detector is called a linear ZF receiver.

A linear ZF receiver performs well for a system having either at high SNRs or with a well conditioned channel matrix. The performance will deteriorate when neither of these conditions are satisfied.

2.4.2 Linear MMSE Receiver

A linear MMSE equalizer is a linear processor that minimizes the averaged detection error, i.e., $\min: E\{\|\mathbf{e}\|^2 = \|\hat{\mathbf{x}} - \mathbf{x}\|^2\}$. A linear MMSE equalizer satisfies [38]:

$$E\{\mathbf{e}\mathbf{y}^H\} = \mathbf{0} \quad (2.28)$$

Substituting Eq. (2.22) into Eq. (2.28), we have

$$\begin{aligned} E\{\mathbf{e}\mathbf{y}^H\} &= E\left\{\left(\sqrt{\frac{\rho}{M}}\mathbf{G}\mathbf{H}\mathbf{x} - \mathbf{x} + \mathbf{G}\mathbf{w}\right)\left(\sqrt{\frac{\rho}{M}}\mathbf{H}\mathbf{x} + \mathbf{w}\right)\right\} \\ &= \frac{\rho}{M}\mathbf{G}\mathbf{H}\mathbf{H}^H - \sqrt{\frac{\rho}{M}}\mathbf{H}^H + \mathbf{G} = \mathbf{0} \end{aligned} \quad (2.29)$$

where we have used the assumption that $E\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}_M$ and $E\{\mathbf{w}\mathbf{w}^H\} = \mathbf{I}_N$. The expression of an MMSE equalizer \mathbf{G}_{MMSE} can be obtained from Eq. (2.29) as

$$\mathbf{G}_{\text{MMSE}} = \sqrt{\frac{\rho}{M}}\mathbf{H}^H\left(\mathbf{I} + \frac{\rho}{M}\mathbf{H}\mathbf{H}^H\right)^{-1} \quad (2.30)$$

or

$$\mathbf{G}_{\text{MMSE}} = \sqrt{\frac{M}{\rho}} \left(\frac{M}{\rho} \mathbf{I} + \mathbf{H}^H \mathbf{H} \right)^{-1} \mathbf{H}^H \quad (2.31)$$

where the equivalency of the two expressions of \mathbf{G}_{MMSE} can be simply proved by employing the matrix inversion lemma. In analyzing the system performance when a linear equalizer is employed, the error covariance matrix $\mathcal{E} \triangleq \text{E}\{\mathbf{e}\mathbf{e}^H\}$ is frequently considered. Here, when a linear MMSE equalizer is considered, the error covariance of the system can be found as [55]

$$\mathcal{E}_{\text{MMSE}} = \text{E}\{\mathbf{e}\mathbf{e}^H\} = \left(\mathbf{I} + \frac{\rho}{M} \mathbf{H}^H \mathbf{H} \right)^{-1} \quad (2.32)$$

The m th diagonal element of \mathcal{E} corresponds to the variance of the detection error for the m th symbol x_m . Therefore, taking trace of Eq. (2.32) results in the mean square error (MSE) of the system. After the linear equalization, the interference to the desired symbol is treated as noise. Then, the equalized signal $\hat{\mathbf{x}}$ with an appropriate scaling can be detected by a symbol-by-symbol threshold detector. A receiver having such architecture is called a linear MMSE receiver.

It should be noted that the MMSE equalizer approaches the ZF equalizer at high SNR. Refer to Eq.(2.31), when $\rho \rightarrow \infty$, the identity matrix is negligible, and we have

$$\mathbf{G}_{\text{MMSE}} \approx \sqrt{\frac{M}{\rho}} (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H, \quad \rho \rightarrow \infty \quad (2.33)$$

The matrix part in Eq.(2.33) is the pseudo inverse of \mathbf{H} that is assumed to have full column rank, and therefore it is a ZF equalizer according to Eq. (2.26). Now we calculate the error covariance for a ZF equalizer,

$$\mathcal{E}_{\text{ZF}} = \text{E}\{\mathbf{G}_{\text{ZF}} \mathbf{w} \mathbf{w}^H \mathbf{G}_{\text{ZF}}^H\} = \frac{M}{\rho} (\mathbf{H}^H \mathbf{H})^{-1} \quad (2.34)$$

where we have used the assumption $\text{E}(\mathbf{w}\mathbf{w}^H) = \mathbf{I}$, and have employed $(\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H$ as the pseudo inverse of \mathbf{H} . From Eqs. (2.34) and (2.32), it is clear that both equalizers have the same error covariance at high SNR.

However, at low SNR, an MMSE equalizer is more robust than a ZF equalizer. Consider the error covariance associated with a ZF equalizer in Eq. (2.34). If the channel matrix is ill-conditioned, taking the inverse such a matrix will cause a large error to the system. However for an MMSE equalizer, the identity matrix in Eq. (2.32) bounds the diagonal entries of the error covariance matrix to be no greater than one.

The linear ZF/MMSE receiver has a reasonably low computational cost. The cost is dominated by the matrix inversion involved in the equalization, which is of order M^3 for an $M \times M$ matrix. It should be noted that the complexity of the linear processing component of a linear receiver is independent of the type of signalling. Therefore, a linear receiver is practically attractive especially for higher order constellations. In the case of block transmission where MT symbols are jointly transmitted and detected, the computational cost in detecting these symbols by employing a linear receiver has terms that are $O(M^3)$, $O(M^2T)$, and $O(MT\mu)$, where μ is the cardinality of the signal constellation.

2.4.3 ML Detector

The maximum likelihood detector is a non-linear detector. It detects the signals jointly by choosing the symbol vector that is most likely to have been transmitted. Therefore, we consider the conditional probability density function $p(\mathbf{y}|\mathbf{s})$ with \mathbf{s} and \mathbf{y} being the transmitted and received signal vectors respectively. The method of maximum likelihood is to estimate the parameter \mathbf{s} such that the associated conditional PDF $p(\mathbf{y}|\mathbf{s})$ is at the maximum. Hence we write $p(\mathbf{y}|\mathbf{s})$ as a function of \mathbf{s} , called the likelihood function, i.e.,

$$l(\mathbf{s}) = p(\mathbf{y}|\mathbf{s})$$

which is to be maximized in detecting \mathbf{s} . When the signal vector \mathbf{s} is equally likely picked from its constellation, the maximum likelihood detection is theoretically optimal in minimizing the probability of vector decision error.

Despite the optimality of an ML detector, its computational complexity makes it infeasible in MIMO systems having a large number of transmitter antennas. For example, consider the block fading MIMO system where the input data sequences are transmitted and received in blocks. When MT signal symbols are processed per block, the computational complexity of detecting these symbols by employing ML detection is $O(\mu^{MT})$ [77]. In comparison to that of a linear receiver, ML detection incurs higher costs when either a high order signal constellation is employed, or a large number of symbols are processed per block.

2.5 Classes of STBC

We now turn our attention onto the design of STBC. We begin by describing several classes of STBC considered in this thesis.

2.5.1 Linear STBC

A linear STBC is a code in which each codeword $\mathbf{X}(\mathbf{s})$ is constructed as the linear combination of a set of code matrices,

$$\mathbf{X}(\mathbf{s}) = \sum_{k=1}^K s_k \mathbf{C}_k \quad (2.35)$$

where \mathbf{C}_k is the $M \times T$ code matrix associated with the symbol s_k . It satisfies the power constraint that

$$\text{tr} \left(\sum_{k=1}^K \mathbf{C}_k \mathbf{C}_k^H \right) = K \quad (2.36)$$

i.e., the codes maintain the same transmission power as that of the original signal \mathbf{s} . Here, K is the total number of symbols transmitted per block (in T time slots). For a full symbol rate transmission, $K = \min\{M, N\}T$. For a linear code defined in Eq. (2.35), the design target is on the structure of \mathbf{C}_k . We observe that the code matrix \mathbf{C}_k may span all the transmitter antennas in space (column dimension of the code is M) and all the time slots in time (row dimension of the code is T). Therefore in transmitting each symbol of \mathbf{s} , such a code provides the possibility for the system to take advantage of all the degrees of freedom offered by a flat fading MIMO channel.

2.5.2 Linear Dispersion STBC

When signals are picked from a complex constellation, e.g., a Quadrature Amplitude Modulation (QAM) [77] constellation, the real and imaginary parts in each symbol typically contain different information. A linear dispersion (LD) code is a STBC specially designed for complex signal symbols where for each signal symbol and its conjugate, different code matrices are assigned as shown in the following,

$$\mathbf{X}(\mathbf{s}) = \sum_{k=1}^K s_k \mathbf{A}_k + \sum_{k=1}^K s_k^* \mathbf{B}_k \quad (2.37)$$

where \mathbf{A}_k and \mathbf{B}_k are the $M \times T$ code matrices associated with the k th signal symbol s_k and its conjugate s_k^* respectively. They are constrained so that the total transmission power remains a constant, i.e.,

$$\text{tr} \left(\sum_{k=1}^K (\mathbf{A}_k \mathbf{A}_k^H + \mathbf{B}_k \mathbf{B}_k^H) \right) = K \quad (2.38)$$

The class of linear dispersion codes is a generalization of the class of linear codes. By setting $\mathbf{B}_k = \mathbf{0}$, Eq. (2.37) reduces to Eq. (2.35).

2.5.3 Unitary Trace-Orthogonal Code

As a subset of linear and linear dispersion STBC, unitary trace-orthogonal code has some special optimality properties. In the following, the definition and properties of the unitary trace-orthogonal code will be provided.

2.5.3.1 Definition

Definition 2.1 *A linear STBC is said to be unitary trace-orthogonal if [61, 55]*

$$\mathbf{C}_k \mathbf{C}_k^H = \frac{1}{M} \mathbf{I}; \quad (2.39a)$$

$$\text{tr}(\mathbf{C}_k \mathbf{C}_\ell^H) = \delta(k - \ell), \quad k, \ell = 1, \dots, K \quad (2.39b)$$

where $\delta(\cdot)$ is the Kronecker delta function.

A linear dispersion STBC is said to be unitary trace-orthogonal if [113]

$$\mathbf{A}_k \mathbf{A}_k^H + \mathbf{B}_k \mathbf{B}_k^H = \frac{1}{M} \mathbf{I} \quad (2.40a)$$

$$\mathbf{A}_k \mathbf{B}_k^H + \mathbf{B}_k \mathbf{A}_k^H = \mathbf{0} \quad (2.40b)$$

$$\text{tr}(\mathbf{A}_k \mathbf{A}_\ell^H + \mathbf{B}_k \mathbf{B}_\ell^H) = \delta(k - \ell) \quad (2.40c)$$

$$\text{tr}(\mathbf{A}_k \mathbf{B}_\ell^H + \mathbf{B}_k \mathbf{A}_\ell^H) = 0 \quad (2.40d)$$

The definition provided here is under the power assumption that

$$\mathbb{E} \{ \text{tr}(\mathbf{X}^H(\mathbf{s}) \mathbf{X}(\mathbf{s})) \} = \mathbb{E} \{ \mathbf{s}^H \mathbf{s} \} = K$$

For a system having different power constraint, the code will be different from the definition provided here by a constant scalar.

The conditions in Eqs. (2.39a), (2.40a), and (2.40b) show the intra-unitary structure for the code associated with each individual symbol, and lead to the term *unitary*. The rest unveil the inter-unitary properties of different code matrices, and lead to the

term *trace-orthogonal*. For a better understanding of the unitary trace-orthogonal structure, we vectorize the linear STBC defined in Eq. (2.35) and obtain

$$\text{vec}(\mathbf{X}(\mathbf{s})) = \text{vec}\left(\sum_{k=1}^K s_k \mathbf{C}_k\right) = \mathbf{C}\mathbf{s} \quad (2.41)$$

where

$$\mathbf{C} \triangleq \left(\text{vec}(\mathbf{C}_1) \quad \cdots \quad \text{vec}(\mathbf{C}_K)\right) \quad (2.42)$$

is the equivalent code matrix having dimension $MT \times K$. In Eq. (2.41), the originally mixed code and signals are now separated and represented by individual variables. This expression will be frequently employed later on in solving code design problems, as it simplifies the analysis in most cases. Now applying Eq. (2.39b) to Eq. (2.41), and utilizing the fact that $\text{tr}(\mathbf{C}_k \mathbf{C}_\ell^H) = \text{vec}(\mathbf{C}_\ell)^H \text{vec}(\mathbf{C}_k)$, we obtain that the trace-orthogonal structure defined in Eq. (2.39b) is equivalent to that the matrix \mathbf{C} has *orthonormal column vectors*.

Similarly, consider the linear dispersion code defined in Eq. (2.37) and we have

$$\begin{pmatrix} \text{vec}(\mathbf{X}(\mathbf{s})) \\ \text{vec}(\mathbf{X}^*(\mathbf{s})) \end{pmatrix} = \begin{pmatrix} \mathbf{A}\mathbf{s} + \mathbf{B}\mathbf{s}^* \\ \mathbf{B}^*\mathbf{s} + \mathbf{A}^*\mathbf{s}^* \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{s} \\ \mathbf{s}^* \end{pmatrix} \quad (2.43)$$

where

$$\mathbf{A} \triangleq \left(\text{vec}(\mathbf{A}_1) \quad \cdots \quad \text{vec}(\mathbf{A}_K)\right) \quad (2.44a)$$

$$\mathbf{B} \triangleq \left(\text{vec}(\mathbf{B}_1) \quad \cdots \quad \text{vec}(\mathbf{B}_K)\right) \quad (2.44b)$$

If we define

$$\mathcal{F} \triangleq \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \quad (2.45)$$

and apply Eqs. (2.40c) and (2.40d), and then we obtain that a linear dispersion STBC is trace-orthogonal when \mathcal{F} has orthonormal column vectors.

The vectorized code gives rise to a physical interpretation of the unitary trace-orthogonal structure. Here we focus on the linear STBC only and the linear dispersion STBC can be dealt with in a similar way. From Eq.(2.41) and (2.42), we observe that each column vector of \mathbf{C} is associated with one of the signal symbols. That \mathbf{C} has orthonormal columns (i.e., trace-orthogonal) evenly spreads the K signal symbols over the whole space, with each of them having the same transmission power. The unitary structure of each \mathbf{C}_k further evenly spreads the associated symbol s_k over the whole space and time slots that a block code spans.

2.5.3.2 Properties

We consider the properties of the unitary trace-orthogonal code when it is applied to a flat block fading MIMO channel that has IID fading coefficients and white noise. We will show its optimality from the view points of mutual information and detection errors. Here we consider linear STBC only. Similar properties also hold for linear dispersion code, and the details will be provided in Chapter 6.

Information Lossless

Consider a MIMO communication system in which the signals are transmitted using a linear STBC, i.e.,

$$\mathbf{Y} = \sqrt{\frac{\rho}{M}} \mathbf{H} \left(\sum_{m=1}^M \mathbf{C}_m s_m \right) + \mathbf{W} \quad (2.46)$$

Here we assume $N \geq M$ and $K = MT$, i.e., the transmission is of full symbol rate, and the signals are IID circularly symmetric complex Gaussian with zero mean and unit variance. If we vectorize both sides of Eq. (2.46), the following expression can be obtained

$$\mathbf{y} = \sqrt{\frac{\rho}{M}} (\mathbf{I}_T \otimes \mathbf{H}) \mathbf{C} \mathbf{s} + \mathbf{w} \quad (2.47)$$

where \mathbf{y} and \mathbf{w} are the vectorized forms of the received signals and noise, and \mathbf{C} is a square code matrix defined in Eq. (2.42) with dimension $MT \times MT$. In Eq. (2.47), if we define a coded channel as $\mathcal{H}_{\mathbf{C}} = (\mathbf{I}_T \otimes \mathbf{H}) \mathbf{C}$, then the system model can be written as

$$\mathbf{y} = \sqrt{\frac{\rho}{M}} \mathcal{H}_{\mathbf{C}} \mathbf{s} + \mathbf{w} \quad (2.48)$$

For the system in Eq. (2.48), the Gaussian mutual information between \mathbf{s} and \mathbf{y} for a given coded channel $\mathcal{H}_{\mathbf{C}}$ is (e.g. [92])

$$\mathcal{I}(\mathbf{s}; \mathbf{y} | \mathcal{H}_{\mathbf{C}}) = \log \det \left(\mathbf{I} + \frac{\rho}{M} \mathcal{H}_{\mathbf{C}}^H \mathcal{H}_{\mathbf{C}} \right) \quad (2.49)$$

Suppose the channel is of the fast fading type, and we consider the averaged behavior. By taking the expectation over Eq. (2.49), $\mathbb{E}\{\mathcal{I}(\mathbf{y}; \mathbf{s} | \mathcal{H}_{\mathbf{C}})\}$ is obtained as a function with respect to the parameter \mathbf{C} . We observe that in Eq. (2.49), the matrix inside the parentheses is symmetric positive definite (PD). On the other hand, it is known $\log \det(\cdot)$ is a concave function [9] over a PD matrix. Hence, for all the possible choices of \mathbf{C} , there exists at least one that achieves the maximum mutual information. Such a code is said to be *information lossless*.

In [55], it is shown that $\mathbb{E}\{\mathcal{I}(\mathbf{y}; \mathbf{s} | \mathcal{H}_{\mathbf{C}})\}$ is maximized if and only if \mathbf{C} is a unitary matrix, i.e., the code is trace-orthogonal as defined in Eq. (2.39b).

Detection Error Minimization

We now show the optimality of the unitary trace-orthogonal code in its error performance. Here we assume that the channel is flat fading with $N \geq M$, and a linear MMSE receiver is employed. It is also assumed that the signals are transmitted at full symbol rate, i.e., $K = M$ (which implies \mathbf{C} is a square matrix). If the signals are picked from a 4-QAM constellation, then each signal symbol contains two bits. For a block code transmitting MT symbols, there are $2MT$ bits in total. Let γ_i denote the signal to interference plus noise ratio (SINR) associated with the i th bit

at the receiver end, then the averaged BEP over the $2MT$ bits transmitted per block is [77, 61, 55]

$$P_e = \frac{1}{2MT} \sum_{i=1}^{2MT} \mathbf{E}_{\mathbf{H}} \{Q(\sqrt{\gamma_i})\} \quad (2.50)$$

where [77] $Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} \exp(-x^2/2) dx$, and γ_i is a function of \mathbf{H} and \mathbf{C} . It can be shown [61, 55] that the minimum of P_e in Eq. (2.50) is

$$P_e \geq \mathbf{E}_{\mathbf{H}} \left\{ Q \left(\sqrt{\frac{M}{\text{tr}(\mathbf{I} + \frac{\rho}{M} \mathbf{H}^H \mathbf{H})^{-1}} - 1}} \right) \right\} \quad (2.51)$$

This lower bound is achieved if and only if the linear STBC has unitary and trace-orthogonal structure. This holds true for any range of SNR.

When the signals are picked from an arbitrary QAM constellation, this optimality property holds at high SNR [61]. More generally, when signals are picked from an arbitrary constellation, it is difficult to formulate the system BEP. For this reason, the performance is analyzed by considering the MSE in detecting the signals, which can be obtained by taking the trace of the error covariance matrix. Following this idea, by replacing \mathbf{H} with $\mathcal{H}_{\mathbf{C}}$ in the general expression of the error covariance matrix provided in Eq. (2.32), the MSE can be found as follows

$$\text{MSE} = \mathbf{E}_{\mathbf{H}} \left\{ \text{tr} \left(\mathbf{I} + \frac{\rho}{M} \mathcal{H}_{\mathbf{C}}^H \mathcal{H}_{\mathbf{C}} \right)^{-1} \right\} \quad (2.52)$$

The MSE in Eq. (2.52) is a function of the code matrix \mathbf{C} . In fact, $\text{tr}(\cdot)^{-1}$ is a convex function [9] over a PD matrix. Under the power constraint, the minimum of MSE can be achieved if and only if the code is of trace-orthogonal structure [61, 55].

Chapter 3

Toeplitz STBC: Properties and Application to MISO Communication Systems

We start by discussing the application of unitary trace-orthogonal codes in a MISO communication system. We will present the first non-orthogonal STBC that enables a linear receiver to achieve full diversity gain for the system.

In this chapter ¹, we consider a system in which signals are detected by a *linear* receiver. We propose a general design criterion for full-diversity STBCs. By applying Toeplitz structure into the design of STBC, we obtain a simple full-diversity STBC that satisfies this criterion and converts an original MISO flat fading channel into a Toeplitz virtual MIMO channel. The code is therefore named Toeplitz STBC. We also propose the following important properties of the Toeplitz code: a) The symbol transmission rate can approach unity. b) For any signalling scheme, employing the

¹ The work related to this chapter has been accepted as full paper [60] to be published in the journal of *IEEE Transaction on Information Theory*. Part of the work has also been presented in *ISIT2005* [111], Adelaide, Australia.

Toeplitz code results in a non-vanishing determinant. We show that the Toeplitz STBC is a member of unitary trace-orthogonal code.

We examine the performance of the Toeplitz STBC in a MISO system equipped with different receivers. When applied in conjunction with a linear ZF or MMSE receiver in a MISO system, the Toeplitz STBC can achieve full diversity. In particular, if QAM is used as the signalling scheme, then for independent MISO flat fading channels, we prove that the Toeplitz codes can approach the *optimal diversity v.s. multiplexing gain tradeoff* with a ZF receiver when the number of channel uses is large. This, so far, is the first STBC proved to achieve the optimal tradeoff curve for a *linear* receiver in an arbitrary MISO communication system.

On the other hand, when applied to a MISO system fitted with an ML receiver and the channel coefficients are independent, the Toeplitz STBC minimizes the worst case average pair-wise error probability resulting in the coding gain being maximized. If the channel coefficients are correlated, the inherent transmission matrix in the Toeplitz STBC can be designed to minimize the average worst case pair-wise error probability. In particular, when pair-wise error probability is approximated by the Chernoff bound, a closed-form optimal solution is obtained.

3.1 Design Criterion for Full-Diversity STBC with Linear Receivers

Consider a MISO wireless communication system having M transmitter antennas and a single receiver antenna. At any time slot (channel use) t , each of the M transmitter antennas is fed a coded symbol for transmission. Each of these transmitter antennas is linked to the receiver antenna through a channel h_m , $m = 1, \dots, M$. At the receiver of such a system, at the end of T time slots ($t = 1, \dots, T$), we receive an

T -dimensional signal vector $\mathbf{y} = [y_1 \ y_2 \ \cdots \ y_T]^T$ which, according to the input-output model of the system, can then be written as

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{w} \quad (3.1)$$

where \mathbf{X} is the $T \times M$ coding matrix, each row of which consists of coded symbols fed to the M transmitter antennas during a particular time slot, \mathbf{h} is an $M \times 1$ channel vector, and \mathbf{w} is an $T \times 1$ complex noise vector. Throughout this chapter, we adopt the following assumptions:

- a) The channel \mathbf{h} is circularly-symmetric complex Gaussian distributed, with zero-mean and positive definite covariance matrix $\mathbf{\Sigma}$;
- b) \mathbf{w} is a circularly-symmetric complex Gaussian noise vector with covariance $\sigma^2 \mathbf{I}_N$.

Through the MISO system, we transmit signals coded with a linear STBC

$$\mathbf{X} = \sum_{k=1}^K \mathbf{C}_k s_k \quad (3.2)$$

where \mathbf{C}_k is the $T \times M$ code matrix associated with the k th transmitted symbol s_k , and K is the total number of symbols to be transmitted per block ($K = T$ for a rate-one system). Write the symbols to be transmitted as a vector and define a code-channel as

$$\mathcal{H} = \left(\mathbf{C}_1 \mathbf{h} \ \mathbf{C}_2 \mathbf{h} \ \cdots \ \mathbf{C}_K \mathbf{h} \right) \quad \text{and} \quad \mathbf{s} = [s_1 \ s_2 \ \cdots \ s_K]^T \quad (3.3)$$

and then the received signal vector can be written as

$$\mathbf{y} = \mathcal{H}\mathbf{s} + \mathbf{w} \quad (3.4)$$

In the following, we will derive a condition on the equivalent channel \mathcal{H} that renders full-diversity when the signals are received by a linear receiver. First, we present the following properties of the equivalent channel matrix \mathcal{H} :

Property 3.1 Suppose the equivalent channel \mathcal{H} in Eq. (3.3) satisfies that the matrix $\mathcal{H}^H \mathcal{H}$ is non-singular for any nonzero \mathbf{h} . Then we have the following inequality:

$$\phi_{\min} \|\mathbf{h}\|^{2K} \leq \det(\mathcal{H}^H \mathcal{H}) \leq \phi_{\max} \|\mathbf{h}\|^{2K} \quad (3.5)$$

where ϕ_{\min} and ϕ_{\max} are positive constants independent of \mathbf{h} . ■

Proof. Since \mathbf{h} is nonzero, we normalize the $K \times K$ matrix $\mathcal{H}^H \mathcal{H}$ by dividing each of its elements with $\|\mathbf{h}\|^2$, i.e., $\mathcal{H}^H \mathcal{H} = \|\mathbf{h}\|^2 \mathbb{H}$, where \mathbb{H} is the normalized matrix with the ij th element being equal to

$$[\mathbb{H}]_{ij} = \frac{\mathbf{h}^H \mathbf{C}_i^H \mathbf{C}_j \mathbf{h}}{\|\mathbf{h}\| \|\mathbf{h}\|} \quad i, j = 1, 2, \dots, K$$

The determinant of positive semi-definite (PSD) matrix \mathbb{H} is continuous in a closed bounded feasible set $\{\bar{\mathbf{h}} : \|\bar{\mathbf{h}}\|^2 = 1\}$ where $\bar{\mathbf{h}} \triangleq \frac{\mathbf{h}}{\|\mathbf{h}\|}$. It has the maximum and minimum values that are denoted by ϕ_{\max} and ϕ_{\min} respectively. Now, since $\mathcal{H}^H \mathcal{H}$ is non-singular for any nonzero \mathbf{h} , its determinant is positive. Therefore, $0 < \phi_{\min} \leq \phi_{\max}$ and Eq. (3.5) holds. □

We provide a specific example in the following to illustrate the logic in developing the proof.

Example 3.1 Consider the following channel matrix

$$\mathcal{H} = \begin{pmatrix} h_1 & 0 \\ h_2 & h_1 \\ 0 & h_2 \end{pmatrix} \quad (3.6)$$

The determinant of matrix $\mathcal{H}^H \mathcal{H}$ can be written as

$$\det(\mathcal{H}^H \mathcal{H}) = \|\mathbf{h}\|^4 \left(1 - \frac{|h_1|^2 |h_2|^2}{\|\mathbf{h}\|^2 \|\mathbf{h}\|^2} \right) \quad (3.7)$$

Observe that $\frac{|h_1|^2}{\|\mathbf{h}\|^2} + \frac{|h_2|^2}{\|\mathbf{h}\|^2} = 1$, and hence, we can define $\frac{|h_1|}{\|\mathbf{h}\|} = \cos \theta$, and $\frac{|h_2|}{\|\mathbf{h}\|} = \sin \theta$, and Eq. (3.7) becomes

$$\det(\mathcal{H}^H \mathcal{H}) = \|\mathbf{h}\|^4 (1 - \sin^2 \theta \cos^2 \theta) = \|\mathbf{h}\|^4 \left(1 - \frac{1}{4} \sin^2(2\theta)\right)$$

It is obvious that the function $f(\theta) = 1 - \frac{1}{4} \sin^2(2\theta)$ is continuous in a closed bounded set. The minimum and maximum of it can be easily obtained, i.e.,

$$\phi_{\min} = \frac{3}{4}; \quad \phi_{\max} = 1$$

Both values are constants, independent of the random channel. Thus, the determinant of the channel matrix is bounded as

$$\frac{3}{4} \|\mathbf{h}\|^4 \leq \det(\mathcal{H}^H \mathcal{H}) \leq \|\mathbf{h}\|^4 \quad (3.8)$$

Property 3.2 If $\mathcal{H}^H \mathcal{H}$ is non-singular for any nonzero \mathbf{h} , then we have

$$\left[(\mathcal{H}^H \mathcal{H})^{-1} \right]_{kk}^{-1} \geq \phi_0 \|\mathbf{h}\|^2 \quad (3.9)$$

for $k = 1, 2, \dots, K$ where ϕ_0 is a constant independent of \mathbf{h} . ■

Proof. From the matrix inversion algorithm, we have

$$\left[(\mathcal{H}^H \mathcal{H})^{-1} \right]_{kk}^{-1} = \frac{\det(\mathcal{H}^H \mathcal{H})}{\det(\bar{\mathcal{H}}_k^H \bar{\mathcal{H}}_k)} \quad (3.10)$$

where $\bar{\mathcal{H}}_k$ is the matrix obtained by deleting the k the column vector, $\mathbf{C}_k \mathbf{h}$, from \mathcal{H} . We notice that the matrix $\bar{\mathcal{H}}_k^H \bar{\mathcal{H}}_k$ is still PSD and therefore satisfies the right side inequality of Eq. (3.5) having an upper bound denoted by $\phi_{k \max} \|\mathbf{h}\|^{2(K-1)}$. Applying the lower bound of Eq. (3.5) to the numerator and the upper bound to the denominator of Eq. (3.10), we obtain

$$\left[(\mathcal{H}^H \mathcal{H})^{-1} \right]_{kk}^{-1} \geq \frac{\phi_{\min} \|\mathbf{h}\|^{2K}}{\phi_{k \max} \|\mathbf{h}\|^{2(K-1)}} = \frac{\phi_{\min}}{\phi_{k \max}} \|\mathbf{h}\|^2 \geq \phi_0 \|\mathbf{h}\|^2$$

where $\phi_0 = \phi_{\min}/\bar{\phi}_{k\max}$, with $\bar{\phi}_{k\max} = \max\{\phi_{k\max}, k = 1, 2, \dots, K\}$. \square

Now, we are ready to discuss the condition on \mathcal{H} for which full-diversity is achieved by a linear receiver for a MISO system. To do this, we need only to consider the system equipped with a linear ZF receiver. The same condition automatically extends to systems with linear MMSE receivers or other more sophisticated receivers such as decision feedback (DFE) or ML receivers. To analyze the diversity gain achieved by the system, we consider the symbol error probability when signals are of: 1) square QAM, 2) PAM and 3) PSK constellations respectively and let μ denote the cardinality. Firstly, we summarize the definition of some common parameters which govern the performance of the linear (ZF and MMSE) detectors under these three signalling schemes. We use the index $i = 1, 2, 3$ to denote parameters associated with the three signalling schemes as ordered above. Let E_{si} , $i = 1, 2, 3$, denote the respective average symbol energy in each of the above signalling schemes, and let σ^2 be the noise variance at the receiver antenna. Therefore, the SNR for each symbol at the receiver is given by

$$\rho_i = E_{si}/\sigma^2; \quad i = 1, 2, 3 \quad (3.11)$$

We note that $\sigma^2[\mathcal{H}^H\mathcal{H}]_{kk}^{-1}$ is in fact the noise power at the output of the ZF equalizer for the k th symbol.

3.1.1 Symbol Error Probability of Various Signalling Schemes

We now summarize the symbol error probability (SEP) when a MISO communication system transmits signals using the above three signalling schemes and a ZF receiver.

1. *PAM signals*: The SEP of the ZF receiver for a μ -ary PAM signal s_k is given by [87]

$$P_2(\mathbf{h}, s_k) = \frac{2(\mu-1)}{\mu} Q \left(\sqrt{\frac{3E_{s2}}{(\mu^2-1)\sigma^2 [\mathcal{H}^H\mathcal{H}]_{kk}^{-1}}} \right) \quad (3.12)$$

We use the following alternative expressions for the Q - and Q^2 -functions [12, 75, 102, 87]

$$Q(z) = \frac{1}{\pi} \int_0^{\pi/2} \exp\left(-\frac{z^2}{2 \sin^2 \theta}\right) d\theta \quad z \geq 0 \quad (3.13)$$

$$Q^2(z) = \frac{1}{\pi} \int_0^{\pi/4} \exp\left(-\frac{z^2}{2 \sin^2 \theta}\right) d\theta \quad z \geq 0 \quad (3.14)$$

Now, using the expression of the Q -function in Eq. (3.13) and noting that by putting $\sin \theta = 1$ in the integral, we have $Q(z) \leq \exp(-z^2)$ for $z \geq 0$. Hence, we arrive at an upper bound of Eq. (3.12), i.e.,

$$P_2(\mathbf{h}, s_k) \leq \frac{\mu - 1}{\mu} \exp\left(-\frac{3E_{s2}}{2(\mu^2 - 1)\sigma^2 \left[(\mathcal{H}^H \mathcal{H})^{-1}\right]_{kk}}\right) \quad (3.15)$$

2. *Square QAM signals*: The SEP of a ZF receiver for the square QAM signal s_k is [87]

$$\begin{aligned} P_1(\mathbf{h}, s_k) &= 4 \left(1 - \frac{1}{\sqrt{\mu}}\right) Q\left(\sqrt{\frac{3E_{s1}}{2(\mu - 1)\sigma^2 \left[(\mathcal{H}^H \mathcal{H})^{-1}\right]_{kk}}}\right) \\ &\quad - 4 \left(1 - \frac{1}{\sqrt{\mu}}\right)^2 Q^2\left(\sqrt{\frac{3E_{s1}}{2(\mu - 1)\sigma^2 \left[(\mathcal{H}^H \mathcal{H})^{-1}\right]_{kk}}}\right) \end{aligned} \quad (3.16)$$

Substituting Eqs. (3.13) and (3.14) into Eq. (3.16) and after a little manipulation, we obtain

$$\begin{aligned} P_1(\mathbf{h}, s_k) &= \frac{4}{\pi} \left(1 - \frac{1}{\sqrt{\mu}}\right) \int_0^{\pi/4} \exp\left(-\frac{3E_{s1}}{4(\mu - 1)\sigma^2 \left[(\mathcal{H}^H \mathcal{H})^{-1}\right]_{kk} \sin^2 \theta}\right) d\theta + \\ &\quad \frac{4}{\pi\sqrt{\mu}} \left(1 - \frac{1}{\sqrt{\mu}}\right) \int_{\pi/4}^{\pi/2} \exp\left(-\frac{3E_{s1}}{4(\mu - 1)\sigma^2 \left[(\mathcal{H}^H \mathcal{H})^{-1}\right]_{kk} \sin^2 \theta}\right) d\theta \end{aligned} \quad (3.17)$$

We can obtain an upper bound for Eq. (3.17) by putting $\sin \theta = 1$ in the two integrals, and this easily simplifies to

$$P_1(\mathbf{h}, s_k) \leq \frac{\mu - 1}{\mu} \exp \left(-\frac{3E_{s1}}{4(\mu - 1)\sigma^2 \left[(\mathcal{H}^H \mathcal{H})^{-1} \right]_{kk}} \right) \quad (3.18)$$

3. *PSK signals*: The SEP of a ZF receiver for the PSK signal s_k is given by [87]

$$P_3(\mathbf{h}, s_k) = \frac{1}{\pi} \int_0^{(\mu-1)\pi/\mu} \exp \left(-\frac{E_{s3} \sin^2(\pi/\mu)}{2\sigma^2 \left[(\mathcal{H}^H \mathcal{H})^{-1} \right]_{kk} \sin^2 \theta} \right) d\theta \quad (3.19)$$

which, similar to the PAM signal, can be upper bounded by

$$P_3(\mathbf{h}, s_k) \leq \frac{(\mu - 1)}{\mu} \exp \left(-\frac{E_{s3} \sin^2(\pi/\mu)}{2\sigma^2 \left[(\mathcal{H}^H \mathcal{H})^{-1} \right]_{kk}} \right) \quad (3.20)$$

3.1.2 Design Criterion for Full-Diversity STBC Applied to a MISO System with Linear Receivers

We now examine the *diversity gain* [91, 114] achieved by a MISO system. From the expression of SEP, we examine the diversity gain for a MISO system having a linear receiver. The result is provided by the following theorem:

Theorem 3.1 *For a MISO system employing a square QAM, PAM, or PSK signalling scheme of cardinality μ in the transmission, a linear (ZF/MMSE) receiver achieves full diversity when $\mathcal{H}^H \mathcal{H}$ is non-singular for any nonzero \mathbf{h} . ■*

Proof. From Eqs. (3.18), (3.15), and (3.20), we can arrive at a generalized upper bound on the symbol error probability for the μ -ary QAM, PAM, and PSK signals such that

$$P_i(\mathbf{h}, s_k) \leq \frac{\mu - 1}{\mu} \exp \left(-\frac{a_i \rho_i}{\left[(\mathcal{H}^H \mathcal{H})^{-1} \right]_{kk}} \right), \quad i = 1, 2, 3 \quad (3.21)$$

where $\rho_i = E_{s_i}/\sigma^2$, and

$$a_1 = 3/[4(\mu - 1)], \quad a_2 = 3/[2(\mu^2 - 1)], \quad \text{and} \quad a_3 = \sin^2(\pi/\mu)/2 \quad (3.22)$$

Since $\mathcal{H}^H\mathcal{H}$ is non-singular for any nonzero \mathbf{h} , we can apply Eq. (3.9) to Eq. (3.21). Here, we see that the SEP of all the three signalling schemes have a general upper bound given by

$$P_i(\mathbf{h}) \leq \frac{\mu - 1}{\mu} \exp(-a_i \rho_i \phi_0 \|\mathbf{h}\|^2) = \frac{\mu - 1}{\mu} \exp(-a_i \rho_i \phi_0 \mathbf{h}^H \mathbf{h}) \quad (3.23)$$

Now, \mathbf{h} is assumed to be Gaussian with zero mean and covariance matrix Σ . Therefore, averaging the exponential part of the right side of Eq. (3.23) over the density function of \mathbf{h} yields

$$\begin{aligned} & \frac{1}{\pi^M \det \Sigma} \int \exp(-a_i \rho_i \phi_0 \mathbf{h}^H \mathbf{h}) \exp(-\mathbf{h}^H \Sigma^{-1} \mathbf{h}) d\mathbf{h} \\ &= \left[\frac{\det((a_i \rho_i \phi_0 \mathbf{I} + \Sigma^{-1})^{-1})}{\det \Sigma} \right] = \det(\mathbf{I} + a_i \rho_i \phi_0 \Sigma)^{-1} \end{aligned} \quad (3.24)$$

Substituting Eq. (3.24) into Eq. (3.23), we establish the following inequalities:

$$\begin{aligned} \mathbb{E}[P_i(\mathbf{h})] &\leq \frac{\mu - 1}{\mu} \det(\mathbf{I} + a_i \rho_i \phi_0 \Sigma)^{-1} \\ &\leq \left(\frac{\mu - 1}{\mu} \det(\phi_0 \Sigma)^{-1} a_i^{-M} \right) \rho_i^{-M}, \quad i = 1, 2, 3 \end{aligned} \quad (3.25)$$

The power of ρ_i in Eq. (3.25) indicates that the upper bound of the SEP using the three signalling schemes and a ZF receiver in a MISO system indeed achieves full diversity for non-singular $\mathcal{H}^H\mathcal{H}$. \square

Remarks on Theorem 1:

- a) Although the proof provided here is for square QAM, PAM and PSK signalings, Theorem 3.1 can be shown to be valid for any signal constellation.

b) Since the condition provided here is sufficient for a linear receiver to achieve full diversity, the same condition naturally yields full diversity for more sophisticated receivers such as MMSE/ZF-DFE or ML receivers.

Corollary 3.1 *In a MISO system having a linear receiver, a STBC $\mathbf{X}(\mathbf{s})$ achieves full diversity if $\mathbf{X}^H(\mathbf{s})\mathbf{X}(\mathbf{s})$ is non-singular for any non-zero \mathbf{s} .*

Proof: Corollary 3.1 can be proved by showing the equivalency of the following two statements:

- 1) $\mathbf{X}^H(\mathbf{s})\mathbf{X}(\mathbf{s})$ is non-singular for any nonzero \mathbf{s} ; and
- 2) $\mathbf{H}^H\mathbf{H}$ is non-singular for any nonzero \mathbf{h} .

We will show 2) \Rightarrow 1), and the reverse can be similarly proved. From the development of Eq. (3.4), we have $\mathbf{X}(\mathbf{s})\mathbf{h} = \mathcal{H}\mathbf{s}$. Now, if $\mathcal{H}^H\mathcal{H}$ is non-singular for any nonzero \mathbf{h} , then \mathcal{H} has full column rank, and hence $\mathcal{H}\mathbf{s} \neq \mathbf{0}$ for any $\mathbf{s} \neq \mathbf{0}$. Therefore $\mathbf{X}(\mathbf{s})\mathbf{h} \neq \mathbf{0}$ for any $\mathbf{s} \neq \mathbf{0}$, $\mathbf{h} \neq \mathbf{0}$. This implies full column rank of matrix $\mathbf{X}(\mathbf{s})$, and hence $\mathbf{X}^H(\mathbf{s})\mathbf{X}(\mathbf{s})$ is non-singular for any nonzero \mathbf{s} . \square

Remark on Corollary 3.1:

The condition provided here enables full diversity for a linear receiver, and hence it is also valid for more sophisticated types of receivers. Corollary 3.1 provides a criterion in designing full diversity STBC in a MISO system having a linear receiver. Thus, the design of a STBC should aim at making $\mathbf{X}^H(\mathbf{s})\mathbf{X}(\mathbf{s})$ full-rank for any nonzero \mathbf{s} .

In the following, we present a special code, *the Toeplitz STBC*, which has a simple structure and satisfies the full-rank condition in Theorem 3.1, and is therefore a full-diversity STBC. We will first introduce the definition and properties of the Toeplitz STBC. This is then followed by its performance analysis when the signals are detected

by both linear receivers and ML receivers. The symbol transmission rate of Toeplitz STBC is shown to be less than, but asymptotically approaching, rate one. The design of a STBC which simultaneously achieves rate one and full diversity for a linear receiver is still an open problem.

3.2 Toeplitz Space-Time Block Codes and Their Properties

3.2.1 Toeplitz STBC for a MISO System

Let us now introduce the Toeplitz space-time block code [111] to the MISO system. Let $\boldsymbol{\alpha} = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_K]^T$. Then, a $(K + J - 1) \times J$ Toeplitz matrix generated by $\boldsymbol{\alpha}$ and a positive integer J , denoted by $\mathcal{T}(\boldsymbol{\alpha}, K, J)$, is defined as

$$[\mathcal{T}(\boldsymbol{\alpha}, K, J)]_{ij} = \begin{cases} \alpha_{i-j+1}, & \text{if } i \geq j \text{ and } i - j < K \\ 0, & \text{otherwise} \end{cases} \quad (3.26)$$

which can be explicitly written as

$$\mathcal{T}(\boldsymbol{\alpha}, K, J) = \begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ \alpha_2 & \alpha_1 & \cdots & 0 \\ \vdots & \alpha_2 & \ddots & \vdots \\ \alpha_K & \ddots & \ddots & \alpha_1 \\ 0 & \ddots & \ddots & \alpha_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & \alpha_K \end{pmatrix}_{(K+J-1) \times J} \quad (3.27)$$

If we replace $\boldsymbol{\alpha}$ by \mathbf{s} , the information symbols to be transmitted, then a *Toeplitz* STBC matrix $\mathbf{X}_B(\mathbf{s})$ is defined as

$$\mathbf{X}_B(\mathbf{s}) = \mathcal{T}(\mathbf{s}, K, J) \cdot \mathbf{B} \quad (3.28)$$

where, for $J \leq M$, \mathbf{B} is a $J \times M$ matrix of rank J , and is placed in the coding matrix to allow for the flexibility of facilitating the transmitter antennas with beamforming capability.

The Toeplitz STBC $\mathcal{T}(\mathbf{s}, K, J)$ is a member of of unitary trace-orthogonal code. This can be observed by writing $\mathcal{T}(\mathbf{s}, K, J)$ in the form of a linear STBC,

$$\mathcal{T}(\mathbf{s}, K, J) = \sum_{k=1}^K s_k \mathbf{C}_k \quad (3.29)$$

with

$$\mathbf{C}_k = \mathbf{P}^{k-1} \mathbf{C}_0, \quad k = 1, \dots, K \quad (3.30)$$

where

$$\mathbf{P} = \begin{pmatrix} \mathbf{0}_{(T-1) \times 1} & 1 \\ \mathbf{I}_{(T-1)} & \mathbf{0}_{1 \times (T-1)} \end{pmatrix}$$

is a row permutation matrix and is unitary, and

$$\mathbf{C}_0 = \begin{pmatrix} \mathbf{I}_K \\ \mathbf{0}_{(J-1) \times J} \end{pmatrix}$$

is a tall matrix. For example, when $K = 2, J = 2$, the two linear codes are

$$\mathbf{C}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \mathbf{C}_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The unitary property of the code can be shown as follows

$$\mathbf{C}_k^H \mathbf{C}_k = \mathbf{C}_0^H (\mathbf{P}^H \mathbf{P})^{k-1} \mathbf{C}_0 = \mathbf{C}_0^H \mathbf{C}_0 = \mathbf{I}$$

To show the trace-orthogonality, we consider

$$\text{tr}(\mathbf{C}_k^H \mathbf{C}_\ell) = \text{tr}(\mathbf{C}_0^H \mathbf{P}^{(k-1)H} \mathbf{P}^{\ell-1} \mathbf{C}_0) \quad (3.31)$$

When $k = \ell$, there is

$$\text{tr}(\mathbf{C}_k^H \mathbf{C}_\ell) = \text{tr}(\mathbf{C}_0^H \mathbf{C}_0) = K \quad (3.32)$$

which is a constant. If $k \neq \ell$, then

$$\text{tr}(\mathbf{C}_k^H \mathbf{C}_\ell) = \text{tr}(\mathbf{P}^{\ell-k} \mathbf{C}_0 \mathbf{C}_0^H) \quad (3.33)$$

Now

$$\mathbf{C}_0 \mathbf{C}_0^H = \begin{pmatrix} \mathbf{I}_K & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \quad (3.34)$$

of which the non-zero elements are located only on the diagonal. The row permutation operated by $\mathbf{P}^{\ell-k}$ will move the non-zero entries off the diagonal. This results in a zero trace. Hence we conclude $\text{tr}(\mathbf{C}_k^H \mathbf{C}_\ell) = K\delta(k, \ell)$.

At time slot t , the t th row of the $T \times M$ matrix $\mathbf{X}_B(\mathbf{s})$ is fed to the M transmitter antennas for transmission. Apply the Toeplitz space-time block coding matrix to the MISO system described in Eq. (3.1), and we have

$$\mathbf{y} = \mathcal{T}(\tilde{\mathbf{h}}, K, J) \mathbf{s} + \mathbf{w} \quad (3.35)$$

where $\tilde{\mathbf{h}} = \mathbf{B}\mathbf{h}$, and $K = T - J + 1$. $\mathcal{T}(\tilde{\mathbf{h}}, K, J)$ can be viewed as the overall channel matrix of the MISO system.

Example 3.2 For $K = M = J = 2, T = K + J - 1 = 3$, and $\mathbf{B} = \mathbf{I}_2$, the codeword matrix and channel matrix are, respectively,

$$\mathbf{X}_{\mathbf{I}_2}(\mathbf{s}) = \begin{pmatrix} s_1 & 0 \\ s_2 & s_1 \\ 0 & s_2 \end{pmatrix}, \quad \mathcal{T}(\tilde{\mathbf{h}}, 2, 2) = \begin{pmatrix} h_1 & 0 \\ h_2 & h_1 \\ 0 & h_2 \end{pmatrix}$$

For this code, there are $K = 2$ symbols to be transmitted in $T = 3$ channel uses. Therefore, the symbol transmission rate of this system is $R_s = \frac{2}{3}$ symbols per channel use.

Remarks on Toeplitz STBC:

- a) It can be seen that in Eq. (3.35), the original MISO channel is transformed into a Toeplitz virtual MIMO channel. Such a channel is a special intersymbol interference channel for block transmission with zero-padding. Thus, we can utilize the efficient Viterbi algorithm [98] to detect the signal \mathbf{s} if perfect channel knowledge is available at the receiver.
- b) We can make use of the second order statistics of the received signal to blindly identify the channel when channel coefficients are not known at the receiver [95,82].
- c) When the Toeplitz STBC is applied to a MISO system, the space diversity has been transformed into delay (time) diversity. This is realized by transforming the flat fading channel into a frequency selective channel with zero-padding. This technique is parallel to that employed in [103]. In [103], only a ML detector is considered whereas here, achieving full diversity by the detection of Toeplitz STBC using a linear receiver is proved, as well as an additional design parameter \mathbf{B} for ML detection is incorporated. Also in [103], performance analysis is carried out under the assumption that for independent channels, only one bit error occurs with the signals being restricted to a BPSK constellation, whereas here, all the possible error events for different signal constellations with the channel being either independent or dependent are considered.
- d) When $\mathbf{B} = \mathbf{I}$, a Toeplitz STBC is reminiscent of a delay diversity code (DDC). In general, a DDC [84, 36] is applied with the use of outer channel coding and ML detectors to achieve the full diversity gain. However, a Toeplitz STBC owns special properties that enable full space diversity even with the use of the simple linear receiver and the signals can be of any type.

e) Toeplitz STBC is the first non-orthogonal STBC that has non-vanishing determinant regardless of the signalling. Hence according to Corollary 3.1, the code achieves full diversity with the use of a linear receiver. For the full diversity ML STBC (e.g., [18, 86, 89, 25]), since they only preserve non-vanishing determinant for certain types of signalling, the full diversity gain is not necessarily guaranteed with the use of a linear receiver.

3.2.2 Properties of Toeplitz STBC

We now examine some important properties of the Toeplitz space-time block codes introduced in the previous subsection. These properties will be useful in performance analysis and code designs in the ensuing sections.

Property 3.3 *The definition of the Toeplitz space-time code shows that the symbol transmission rate is $R_s = \frac{K}{T} = \frac{T-J+1}{T}$ symbols per channel use when $J \leq M$. Therefore, for a fixed M , the transmission rate R can approach unity if the number of channel uses is sufficiently large.*

Property 3.4 *For any nonzero vector α , there exists $0 < \phi_{T \min} \leq \phi_{T \max} \leq 1$, and the matrix $(\mathcal{T}^H(\alpha, K, J)\mathcal{T}(\alpha, K, J))$ satisfies the following inequality,*

$$\phi_{T \min} \|\alpha\|^{2J} \leq \det(\mathcal{T}^H(\alpha, K, J)\mathcal{T}(\alpha, K, J)) \leq \phi_{T \max} \|\alpha\|^{2J} \quad (3.36)$$

Proof. By letting $\alpha = \mathbf{h}$ in Eq. (3.3) and choosing

$$\mathbf{C}_k = \mathbf{P}^{k-1} \mathbf{C}_0, \quad k = 1, \dots, K$$

as shown in Eq. (3.30), we obtain an equivalent channel \mathcal{H} of the same structure as $\mathcal{T}(\alpha, K, J)$. Hence, $\mathcal{T}(\alpha, K, J)$ is a special case of \mathcal{H} . Thus, from Property 3.1, there exist $\phi_{T \min}$ and $\phi_{T \max}$ for which Eq. (3.36) holds. Now, we note that

the diagonal entries of the matrix $\mathcal{T}^H(\boldsymbol{\alpha}, K, J)\mathcal{T}(\boldsymbol{\alpha}, K, J)$ are all the same and are equal to $[\mathcal{T}^H(\boldsymbol{\alpha}, K, J)\mathcal{T}(\boldsymbol{\alpha}, K, J)]_{jj} = \|\boldsymbol{\alpha}\|^2$, $j = 1, \dots, J$. By applying Hadamard's inequality [43], we arrive at:

$$\det(\mathcal{T}^H(\boldsymbol{\alpha}, K, J)\mathcal{T}(\boldsymbol{\alpha}, K, J)) \leq \|\boldsymbol{\alpha}\|^{2J} \quad \Rightarrow \quad \phi_{T_{\max}} \leq 1 \quad (3.37)$$

Furthermore, since $\boldsymbol{\alpha}$ is nonzero, we can assume, without loss of generality, that the first element $\alpha_1 \neq 0$. (Otherwise, we can always permute the nonzero element to the first position.) The $T \times J$ "tall" matrix, $\mathcal{T}(\boldsymbol{\alpha}, K, J)$, can be partitioned into a top $J \times J$ matrix $\boldsymbol{\Omega}_1$, and a bottom matrix $\boldsymbol{\Omega}_2$ containing the rest of $\mathcal{T}(\boldsymbol{\alpha}, K, J)$, i.e.,

$$\mathcal{T}(\boldsymbol{\alpha}, K, J) = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ \alpha_2 & \alpha_1 & \ddots & \vdots \\ \vdots & \alpha_2 & \ddots & 0 \\ \alpha_J & \alpha_{J-1} & \dots & \alpha_1 \\ \text{---} & \text{---} & \text{---} & \text{---} \\ \alpha_{J+1} & \alpha_J & \dots & \alpha_2 \\ \vdots & \vdots & \dots & \vdots \\ \alpha_K & \alpha_{K-1} & \dots & \alpha_{K-J+1} \\ 0 & \alpha_K & \dots & \alpha_{K-J+2} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_K \end{pmatrix} \begin{matrix} \boldsymbol{\Omega}_1 \\ \boldsymbol{\Omega}_2 \end{matrix}$$

We note that $\boldsymbol{\Omega}_1$ is a lower triangular matrix having equal diagonal elements $\alpha_1 \neq 0$. (Here, we assume that $J \leq K$. The proof is equally valid if $K \leq J$ by exchanging the roles of K and J). Since $\mathcal{T}^H(\boldsymbol{\alpha}, K, J)\mathcal{T}(\boldsymbol{\alpha}, K, J) = \boldsymbol{\Omega}_1^H\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2^H\boldsymbol{\Omega}_2$, using a standard result [43, p. 484] on the determinant of the sum of a positive definite and

a positive semi-definite matrix we have

$$\begin{aligned} \det(\mathcal{T}^H(\boldsymbol{\alpha}, K, J)\mathcal{T}(\boldsymbol{\alpha}, K, J)) &\geq \det(\boldsymbol{\Omega}_1^H\boldsymbol{\Omega}_1) + \det(\boldsymbol{\Omega}_2^H\boldsymbol{\Omega}_2) \\ &\geq \det(\boldsymbol{\Omega}_1^H\boldsymbol{\Omega}_1) = |\alpha_1|^{2J} \end{aligned} \quad (3.38)$$

Since $\alpha_1 \neq 0$, $\det(\mathcal{T}^H(\boldsymbol{\alpha}, K, J)\mathcal{T}(\boldsymbol{\alpha}, K, J))$ is nonzero for any nonzero $\boldsymbol{\alpha}$ and thus we have

$$\phi_{\mathcal{T}\min} > 0 \quad (3.39)$$

The conclusions in Eqs. (3.37) and (3.39) complete the proof of the property. \square

We see that \mathcal{T} is non-singular by Property 3.4, then using the fact that \mathcal{T} as a special case of \mathcal{H} in Property 3.2, we have:

Property 3.5

$$[(\mathcal{T}^H(\boldsymbol{\alpha}, K, J)\mathcal{T}(\boldsymbol{\alpha}, K, J))^{-1}]_{jj}^{-1} \geq \phi_{\mathcal{T}\min}\|\boldsymbol{\alpha}\|^2 \quad \text{for } j = 1, 2, \dots, J \quad (3.40)$$

where the existence of $\phi_{\mathcal{T}\min} > 0$ is shown in Property 3.4, holds. \blacksquare

We now introduce the following definition related to the measure in a signal constellation \mathcal{S} :

Definition 3.1 For $s, s' \in \mathcal{S}$, the minimum distance of the signal constellation is defined as

$$d_{\min}(\mathcal{S}) = \min_{s \neq s'} \|s - s'\| \quad (3.41)$$

If $\|s - s'\| = d_{\min}(\mathcal{S})$, we say that s and s' are *neighbours*.

From the definition of the coding matrix \mathbf{X} in Eq. (3.28) with $\mathbf{B} = \mathbf{I}_M$, we can now establish a lower bound for a metric between $\mathbf{X}_{I_M}(s)$ and $\mathbf{X}_{I_M}(s')$ in $d_{\min}(\mathcal{S})$. Let $\mathbf{e} = (s - s')$. For notational convenience, we let $\mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\})$ denote the matrix consisting of m of the columns of $\mathbf{X}_{I_M}(\mathbf{e})$ indexed by $\{i_1, i_2, \dots, i_m\}$ where $i_1 < i_2 < \dots < i_m$ and these columns are not necessarily consecutively chosen. Then, we have

Property 3.6 For $\mathbf{s} \neq \mathbf{s}' \in \mathcal{S}^K$, where $\mathcal{S}^K = \mathcal{S} \times \mathcal{S} \cdots \times \mathcal{S}$, and any nonzero vector $\mathbf{e} = (\mathbf{s} - \mathbf{s}')$, we have

$$\det [\mathbf{X}_{I_M}^H(\mathbf{e}, \{i_1, i_2, \dots, i_m\}) \mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\})] \geq d_{\min}^{2m}(\mathcal{S}) \quad (3.42)$$

for $m = 0, 1, \dots, M-1$, where the equality holds if and only if \mathbf{s} and \mathbf{s}' are neighbours, i.e., iff $\|\mathbf{e}\| = d_{\min}(\mathcal{S})$. \blacksquare

Proof. The proof of this property is similar to that of Property 3.4. As in Property 3.4, without loss of generality, we can always assume that $e_1 \neq 0$ with e_1 being the first element of \mathbf{e} . Thus, $\mathbf{X}_{I_M}(\mathbf{e})$ can be written as

$$\mathbf{X}_{I_M}(\mathbf{e}) = \begin{pmatrix} e_1 & 0 & \dots & 0 \\ e_2 & e_1 & \dots & 0 \\ \vdots & e_2 & \ddots & \vdots \\ e_M & \ddots & \ddots & e_1 \\ \vdots & e_M & \ddots & e_2 \\ e_K & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & e_M \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & e_K \end{pmatrix}_{T \times M} \quad (3.43)$$

An important observation in Eq. (3.43) is that the top submatrix consisting of the first M rows of $\mathbf{X}_{I_M}(\mathbf{e})$ is a $M \times M$ lower triangular matrix with nonzero diagonal entries and therefore, nonzero determinant. We can also see that the submatrix $\mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\})$ preserves the same property because by permuting its rows and columns, an $m \times m$ lower triangular matrix can always be formed as its top part, i.e., $\mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\})$ can be expressed as

$$\mathbf{\Pi}_1 \mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\}) \mathbf{\Pi}_2 = \begin{pmatrix} \Omega_1 \\ \hline \Omega_2 \end{pmatrix}$$

where $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$ denote the $T \times T$ and $m \times m$ permutation matrices, respectively, $\mathbf{\Omega}_1$ contains the first m rows of $\mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\})$ and hence is lower triangular, and $\mathbf{\Omega}_2$ denotes the remaining submatrix of $\mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\})$. Since the permutation of the rows and columns of $\mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\})$ does not change the determinant of its autocorrelation matrix $\mathbf{X}_{I_M}^H(\mathbf{e}, \{i_1, i_2, \dots, i_m\})\mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\})$, therefore, as in Property 3.4, we arrive at

$$\begin{aligned} \det [\mathbf{X}_{I_M}^H(\mathbf{e}, \{i_1, i_2, \dots, i_m\})\mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\})] \\ \geq \det(\mathbf{\Omega}_1^H \mathbf{\Omega}_1) + \det(\mathbf{\Omega}_2^H \mathbf{\Omega}_2) \geq d_{\min}^{2m}(\mathcal{S}) \end{aligned}$$

where the equality holds iff $\mathbf{\Omega}_2$ is a zero matrix, i.e., iff \mathbf{s} and \mathbf{s}' are neighbouring points. \square

We can establish another useful property on the metric between $\mathbf{X}_{I_M}(\mathbf{s})$ and $\mathbf{X}_{I_M}(\mathbf{s}')$ by first recalling an important property in matrix algebra [49]:

The characteristic polynomial of an $M \times M$ matrix \mathbf{A} is the polynomial whose roots are the eigenvalues of \mathbf{A} . Mathematically, it can be re-written as

$$h(\nu) \triangleq \det(\mathbf{I} + \nu \mathbf{A}) = \nu^M + c_1 \nu^{M-1} + \dots + c_{M-1} \nu + c_M \quad (3.44)$$

$$\text{such that} \quad c_m = \sum_{\vartheta} \det(\mathbf{A})_{i_1, \dots, i_m} \quad (3.45)$$

where $\mathbf{A}_{i_1, \dots, i_m}$ denotes the principal submatrix obtained by deleting the rows and columns of \mathbf{A} except the i_1 th, the i_2 th, \dots , and the i_m th ones, and ϑ denotes the combination set of i_1, \dots, i_m . We note, in particular, $c_1 = \text{tr}(\mathbf{A})$ and $c_M = \det(\mathbf{A})$. Now, the following property provides us with another lower bound on the metric between $\mathbf{X}_{I_M}(\mathbf{s})$ and $\mathbf{X}_{I_M}(\mathbf{s}')$ in relation to $d_{\min}(\mathcal{S})$.

Property 3.7 *Let $\mathbf{\Delta} = \text{diag}(\delta_1, \delta_2, \dots, \delta_M)$ with $\delta_m > 0$ for $m = 1, 2, \dots, M$. Then, for any nonzero vector \mathbf{e} , the following inequality holds*

$$\det(\mathbf{\Delta} + \mathbf{X}_{I_M}^H(\mathbf{e})\mathbf{X}_{I_M}(\mathbf{e})) \geq \prod_{m=1}^M (\delta_m + d_{\min}^2(\mathcal{S})) \quad (3.46)$$

with equality holding if and only if \mathbf{s} and \mathbf{s}' are neighbours. ■

Proof. Let us first rewrite the left side of Eq. (3.46) as

$$\begin{aligned} \det(\Delta + \mathbf{X}_{I_M}^H(\mathbf{e})\mathbf{X}_{I_M}(\mathbf{e})) \\ = \det(\Delta) \det\left(\mathbf{I} + (\mathbf{X}_{I_M}(\mathbf{e})\Delta^{-1/2})^H \mathbf{X}_{I_M}(\mathbf{e})\Delta^{-1/2}\right) \end{aligned} \quad (3.47)$$

Now, let $\mathbf{A} = (\mathbf{X}_{I_M}(\mathbf{e})\Delta^{-1/2})^H \mathbf{X}_{I_M}(\mathbf{e})\Delta^{-1/2}$, then Eqs. (3.44) and (3.45) becomes

$$\det\left[\mathbf{I} + (\mathbf{X}_{I_M}(\mathbf{e})\Delta^{-1/2})^H \mathbf{X}_{I_M}(\mathbf{e})\Delta^{-1/2}\right] = 1 + \sum_{m=1}^M c_m, \quad (3.48)$$

where

$$c_m = \sum_{\vartheta} \det\left[\mathbf{X}_{I_M}^H(\mathbf{e}, \{i_1, i_2, \dots, i_m\})\mathbf{X}_{I_M}(\mathbf{e}, \{i_1, i_2, \dots, i_m\})\right] \prod_{\ell=1}^m \delta_{i_\ell}^{-1} \quad (3.49)$$

Using Eq. (3.42) in Property 3.6 on the right side of Eq. (3.49), we have

$$c_m \geq d_{\min}^{2m}(\mathcal{S}) \sum_{\vartheta} \prod_{\ell=1}^m \delta_{i_\ell}^{-1} \quad (3.50)$$

where $\sum_{\vartheta} \prod_{\ell=1}^m \delta_{i_\ell}^{-1}$ denotes the sum of the combination of the product of δ_{i_ℓ} taken m at a time. Equality in Eq. (3.50) holds if and only if \mathbf{s} and \mathbf{s}' are neighbours.

Combining Eqs. (3.48) and (3.50) results in

$$\det\left[\mathbf{I} + (\mathbf{X}_{I_M}(\mathbf{e})\Delta^{-1/2})^H \mathbf{X}_{I_M}(\mathbf{e})\Delta^{-1/2}\right] \geq 1 + \sum_{m=1}^M d_{\min}^{2m}(\mathcal{S}) \sum_{\vartheta} \prod_{\ell=1}^m \delta_{i_\ell}^{-1} \quad (3.51a)$$

$$= \prod_{m=1}^M (1 + \delta_m^{-1} d_{\min}^2(\mathcal{S})) \quad (3.51b)$$

where in the second step, we recognize that the right side of Eq. (3.51a) is the eigen polynomial of the matrix $d_{\min}^2 \Delta$ which, in turn, equals $\det(\mathbf{I} + d_{\min}^2 \Delta)$ and hence Eq. (3.51b). Combining Eqs. (3.51b) and (3.47), we complete the proof of Property 3.7. □

3.3 Toeplitz STBC Applied to a MISO System with a Linear Receiver

We now apply the Toeplitz STBC to the MISO communication system using the properties presented in Section 3.2.2. From Property 3.3 and 3.4 in Section 3.2.2 and Corollary 3.1 in Section 3.1.2, we can see that the Toeplitz STBC can approach *unit-rate* as well as *full diversity* even if only a linear ZF or linear MMSE receiver is used in a MISO system. In the following, we examine the optimal tradeoff between diversity gain and multiplexing gain [114] when the Toeplitz STBC is employed in a MISO system equipped with a linear receiver. We first make the assumption that the channel coefficients are independent, i.e., $\Sigma = \mathbf{I}$.

Now, our MISO system has M transmitter antennas transmitting a signal vector \mathbf{s} of length K in $T = K + M - 1$ time slots. Also, all the above three signalling schemes have constellation cardinality μ . Thus, employing any of the three schemes described in Section 3.1 in our MISO system will result in a bit transmission data rate r given by

$$r = \frac{K}{T} \log_2 \mu \quad (3.52)$$

Note that r is the *bit* rate of transmission. The multiplexing gain g , on the other hand, is dependent on the scheme and in general, is defined as [114]

$$g_i = \frac{r}{\log_2 \text{SNR}_i} \quad (3.53)$$

where “ SNR_i ” refers to the general SNR in the received data. Here in our analysis, we use the SNR of the received *data block* for the SNR_i and denote this by ρ_{bli} , $i = 1, 2, 3$ when the i th signalling scheme is employed. Notice that in the MISO system, we always have $0 \leq g_i \leq 1 \forall i$, since the system has only one receiver antenna. From Eqs. (3.52) and (3.53), we can write $\mu = \rho_{bli}^{Tg_i/K}$, which implies that the cardinality

of the constellation is increasing with SNR. It has been shown [114] that at high SNR, we can trade-off the multiplexing gain for diversity gain and *vice versa*, and the optimum trade-off for our MISO system with M transmitter antennas is given by

$$d_{\text{op}} = M(1 - g) \quad (3.54)$$

where d_{op} is the optimal diversity gain. Let us examine trade-off of the multiplexing gain for diversity gain in the three signalling schemes:

1. *QAM signals*: The averaged symbol energy E_s for square QAM signal is [77]

$$E_{s1} = \frac{2}{3}(\mu - 1) \quad (3.55)$$

We note that E_{s1} increases with the constellation cardinality μ . From Eq. (3.55), the averaged transmission energy *per block* can be calculated as $E_{\text{bl1}} = \frac{2}{3}(\mu - 1)MK$. Given σ^2 being the noise variance at the receiver antenna, the averaged noise power per block is $\sigma_{\text{bl1}}^2 = \sigma^2 T$. Therefore, the block SNR is

$$\rho_{\text{bl1}} = \frac{2(\mu - 1)MK}{3T\sigma^2}$$

leading to

$$\sigma^2 \approx \frac{2\mu MK}{3T\rho_{\text{bl1}}} = \frac{2MK}{3T}\rho_{\text{bl1}}^{\frac{Ng_1}{K}-1} \quad (3.56)$$

where the approximation is under the assumption of large μ , and Eq. (3.53) has been used. Therefore, from Eqs. (3.11), (3.22), (3.55) and (3.56), we obtain

$$a_1\rho_1 = \frac{3E_s}{4(\mu - 1)\sigma^2} = \frac{1}{2\sigma^2} = \frac{3T}{4MK}\rho_{\text{bl1}}^{1-\frac{Tg}{K}} \quad (3.57)$$

Now, consider Eq. (3.25) on the upper bound of the SEP for a ZF receiver, i.e.,

$$\text{E}[P_1(\mathbf{h})] \leq C_{T\text{min}}^{-M} (a_1\rho_1)^{-M} \quad (3.58)$$

Substituting Eq. (3.57) into Eq. (3.58), we obtain

$$E[P_1(\mathbf{h})] \leq C_{T\min}^{-M} \left(\frac{3T}{4MK} \right)^{-M} \rho_{\text{bl}}^{\frac{MT}{K}g-M} \quad (3.59)$$

Hence, the diversity gain for this scheme, D_1 , is given by

$$d_1(g) = M \left(1 - \frac{T}{K}g \right) = M(1 - g) - \varepsilon Mg = d_{\text{op}}(g) - \varepsilon Mg \quad (3.60)$$

where $\varepsilon = \frac{M-1}{K} \geq 0$. From Eq. (3.60), we can see that $d_1(g) \leq d_{\text{op}}(g)$. However, we can make ε small by choosing K sufficiently large so that $d_1(g) \approx d_{\text{op}}(g)$. Hence, $D_1(g)$ is the ε -approximation of $D_{\text{op}}(g)$. We can always choose $K = \lceil \frac{M-1}{\varepsilon} \rceil + 1$, where $\lceil \cdot \rceil$ denotes the integer part of a quantity, and therefore, we can say that *the ZF receiver is able to approach the optimal diversity-multiplexing tradeoff if the proposed Toeplitz code is used with a square QAM signalling scheme and the cardinality is increasing with SNR at a rate $\rho_{\text{bl}}^{Tg_1/K}$.*

If an MMSE receiver is employed, utilizing an expression parallel to Eq. (3.59) for the MMSE receiver, we can also show that the optimal diversity-multiplexing tradeoff can be asymptotically achieved if the Toeplitz STBC is applied to the MISO system in which a square QAM signalling scheme is employed.

2. *PAM signals*: We note that the averaged transmission energy E_s for μ -ary PAM signal is given by [77] $E_s = \frac{1}{6}(\mu^2 - 1)$. Hence the averaged transmission energy per block is $E_{\text{sbl}} = \frac{1}{6}(\mu^2 - 1)MK$. Following similar arguments resulting in Eq. (3.56), for PAM signals we have,

$$\sigma^2 \approx \frac{\mu^2 MK}{6T \rho_{\text{bl}2}} = \frac{MK}{6T} \rho_{\text{bl}2}^{\frac{2Tg}{K}-1}$$

Also, similarly to Eq. (3.57), we have

$$a_2 \rho_2 = \frac{3E_s}{2(\mu^2 - 1)\sigma^2} = \frac{1}{4\sigma^2} = \frac{3T}{2MK} \rho_{\text{bl}2}^{1-\frac{2Tg}{K}}$$

Therefore, from Eq. (3.25), the upper bound on SEP for PAM signal is

$$\mathbb{E}[P_2(\mathbf{h})] \leq C_{T \min}^{-M} (a_2 \rho_2)^{-M} = \left(\frac{3TC_{T \min}}{2MK} \right)^{-M} \rho_{\text{bl}2}^{\frac{2MT}{K}g-M}$$

Hence, the diversity order is

$$\begin{aligned} d_2(g) &= M \left(1 - \frac{2T}{K}g \right) = M(1 - 2g) - 2\epsilon M g \\ &\leq M(1 - 2g) \leq M(1 - g) = d_{\text{op}}(g) \end{aligned} \quad (3.61)$$

Equality in (3.61) holds iff $g = 0$. Therefore, for *finite multiplexing gain*, $d_2(g)$ cannot approach the optimal tradeoff $d_{\text{op}}(g)$.

3. *PSK signals*: The averaged transmission energy E_s for μ -ary PSK signal is [77] $E_s = 1$. Hence the averaged transmission energy per block is $E_{\text{sbl}} = MK$. Therefore, we have

$$\sigma^2 = \frac{MK}{T\rho_{\text{bl}3}}$$

We also have

$$a_3 \rho_3 = \frac{E_s \sin^2(\pi/\mu)}{2\sigma^2} \approx \frac{\pi^2}{2\sigma^2 \mu^2} = \frac{T\pi^2}{2MK} \rho_{\text{bl}3}^{1-\frac{2Tg}{K}} \quad (3.62)$$

where, the second step comes from the assumption of large μ . Following similar arguments as PAM scheme, it can be shown that *PSK signalling cannot achieve the optimal tradeoff of diversity-multiplexing gains*.

3.4 Optimal Toeplitz STBC Design for MISO System with ML Detector

The previous section shows what could be achieved when the Toeplitz STBC is applied to a MISO system equipped with a *linear* receiver. In this section, we will examine

the application of the Toeplitz STBC to a MISO system equipped with a ML detector. In particular, we seek for the optimal design of the matrix \mathbf{B} inherent in a Toeplitz space-time block code Eq. (3.28) so that the worst case pair-wise error probability is minimized when a maximum likelihood detector is employed.

Given a channel realization \mathbf{h} and a transmission matrix \mathbf{B} , the probability $P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{h}, \mathbf{B})$ of transmitting \mathbf{s} and deciding in favor of $\mathbf{s}' \neq \mathbf{s}$ with the ML detector is given by [26]

$$P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{h}, \mathbf{B}) = Q\left(\frac{d(\mathbf{s}, \mathbf{s}')}{2\sigma}\right) \quad (3.63)$$

where $d(\mathbf{s}, \mathbf{s}')$ is the Euclidean distance between the Toeplitz coded signals \mathbf{s} and \mathbf{s}' after being transmitted through the channel, i.e., it is the Euclidean distance between $\mathbf{X}_B(\mathbf{s})\mathbf{h}$ and $\mathbf{X}_B(\mathbf{s}')\mathbf{h}$. Because of the relation of Eq. (3.28), we can write:

$$d^2(\mathbf{s}, \mathbf{s}') = (\mathbf{s} - \mathbf{s}')^H \mathbf{T}^H(\tilde{\mathbf{h}}, M, N) \mathbf{T}(\tilde{\mathbf{h}}, M, N) (\mathbf{s} - \mathbf{s}') = \mathbf{h}^H \mathbf{X}_B^H(\mathbf{e}) \mathbf{X}_B(\mathbf{e}) \mathbf{h} \quad (3.64)$$

where $\mathbf{e} = \mathbf{s} - \mathbf{s}'$. By employing the alternative expression of the Q -function in Eq. (3.13) and taking the average of Eq. (3.63) over the Gaussian random vector \mathbf{h} , the average pair-wise error probability can be written as

$$P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}) = \frac{1}{\pi} \int_0^{\pi/2} \frac{d\theta}{\det(\mathbf{I} + (8\sigma^2 \sin^2 \theta)^{-1} \Sigma \mathcal{X}_B^H(\mathbf{e}) \mathbf{X}_B(\mathbf{e}))} \quad (3.65)$$

with Σ being the covariance matrix of \mathbf{h} . The design problem can now be stated as:

Design Problem: For a fixed number M of transmitter antennas, find a $J \times M$, ($J \leq M$), matrix \mathbf{B} such that the worst-case average pair-wise error probability $P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B})$ is minimized, subject to the transmission power constraint, $\text{tr}(\mathbf{B}^H \mathbf{B}) \leq 1$, i.e.,

$$\mathbf{B}_{\text{op}} = \arg \min_{\text{tr}(\mathbf{B}^H \mathbf{B}) \leq 1} \max_{\substack{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^L \\ \mathbf{s}' \neq \mathbf{s}}} P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}) \quad (3.66)$$

where $\mathcal{S}^K = \mathcal{S} \times \mathcal{S} \cdots \times \mathcal{S}$ with \mathcal{S} denoting the signal constellation of each element of \mathbf{s} .

To solve the above design problem, we not only have to find the optimum \mathbf{B} , but also have to determine its dimension J . Let us first examine the $M \times M$ covariance matrix, $\mathbf{\Sigma} = \mathbb{E}[\mathbf{h}\mathbf{h}^H]$, of the transmission channels \mathbf{h} . Suppose we perform an eigen decomposition such that $\mathbf{\Sigma} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$ where \mathbf{V} is an $M \times M$ unitary matrix and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_M)$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M > 0$. The following theorem provides us with an optimum design of \mathbf{B} :

Theorem 3.2 *Let $\mathbf{\Gamma} = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_J)$, $J \leq M$, be the singular values of \mathbf{B} and let $G(\mathbf{\Lambda}_J\mathbf{\Gamma}, \varepsilon)$ denote the integral*

$$G(\mathbf{\Lambda}_J\mathbf{\Gamma}, \varepsilon) = \frac{1}{\pi} \int_0^{\pi/2} \prod_{j=1}^J \left(1 + \frac{\varepsilon \lambda_j \gamma_j^2}{\sin^2 \theta} \right)^{-1} d\theta \quad \text{for } \varepsilon > 0 \quad (3.67)$$

An optimal $\mathbf{\Gamma}$ can be obtained by solving the following convex optimization problem.²

$$\mathbf{\Gamma}_{\text{op}} = \arg \min_{\text{tr}(\mathbf{\Gamma}) \leq 1} G \left(\mathbf{\Lambda}_J\mathbf{\Gamma}, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2} \right) \quad (3.68)$$

where $\mathbf{\Gamma}_{\text{op}}$ is a $J \times J$ diagonal matrix given by $\mathbf{\Gamma}_{\text{op}} = \text{diag}(\gamma_{\text{op}1}, \gamma_{\text{op}2}, \dots, \gamma_{\text{op}J})$ with J being the largest integer for which $[\mathbf{\Gamma}_{\text{op}}]_{jj} = \gamma_{\text{op}j} > 0$, $j = 1, 2, \dots, J$. Then the optimum transmission matrix is given by

$$\mathbf{B}_{\text{op}} = \mathbf{\Gamma}_{\text{op}}\mathbf{V}_J^H \quad (3.69)$$

where \mathbf{V}_J is the $M \times J$ matrix containing the J eigenvectors corresponding to the J largest eigenvalues in the eigen decomposition of $\mathbf{\Sigma}$. Furthermore, the worst case

² Note that the work presented here is different from that in [40] in which a precoder matrix is designed for a frequency-selective fading channel even though both involve Toeplitz structured matrices. Here, the Toeplitz matrix containing \mathbf{B} and \mathbf{h} is separated from the signal vector \mathbf{s} . This, by the properties of the Toeplitz STBC shown, transforms the design of \mathbf{B} into a convex optimization problem. In [40], however, the design parameter and the signal vector are all parts of the Toeplitz structure resulting in a non-convex design problem that can only be solved by numerical method with no guarantee for global optimality.

pair-wise error probability is lower bounded by

$$\max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^L, \mathbf{s} \neq \mathbf{s}'} P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}) \geq G \left(\Lambda_J \Gamma_{\text{opt}}, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2} \right). \quad (3.70)$$

Equality in Eq. (3.70) holds if and only if

i) $\|\mathbf{s} - \mathbf{s}'\| = d_{\min}(\mathcal{S})$, and

ii) $\mathbf{B} = \mathbf{B}_{\text{op}}$. ■

Proof. We first establish a lower bound on the worst case average pair-wise error probability. Let \mathbf{s}_k and \mathbf{s}'_k be neighbour symbol vectors differing in only the k th symbol, i.e., $\mathbf{s}_k - \mathbf{s}'_k = \mathbf{e}_k = [0 \ \cdots \ 0 \ e_k \ 0 \ \cdots \ 0]^T$ where $|e_k| = d_{\min}(\mathcal{S})$. Then, we can write

$$\begin{aligned} \det(\mathbf{I}_M + \frac{1}{8\sigma^2 \sin^2 \theta} \Sigma \mathcal{X}_B^H(\mathbf{e}_k) \mathbf{X}_B(\mathbf{e}_k)) &= \det(\mathbf{I}_M + \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2 \sin^2 \theta} \Sigma \mathbf{B}^H \mathbf{B}) \\ &\leq \prod_{j=1}^J \left(1 + \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2 \sin^2 \theta} \gamma_j^2 \lambda_j \right) \end{aligned} \quad (3.71)$$

where the first step is a result of the structure of \mathbf{e}_k on the Toeplitz code, and the second step is the result of an inequality for the determinant of a matrix [105, 11]. Equality in Eq. (3.71) holds if and only if $\mathbf{B} = \mathbf{B}_o = \Gamma \mathbf{V}_J^H$, i.e., the singular vectors of \mathbf{B} are the eigenvectors of Σ . Substituting the inequality of (3.71) in Eq. (3.65), we have $P(\mathbf{s}_k \rightarrow \mathbf{s}'_k | \mathbf{B}) \geq G \left(\Lambda_J \Gamma, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2} \right)$. Since $\left(\max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K, \mathbf{s} \neq \mathbf{s}'} P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}) \right) \geq P(\mathbf{s}_k \rightarrow \mathbf{s}'_k | \mathbf{B})$, the worst case average pair-wise error probability is lower bounded by

$$\max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K, \mathbf{s} \neq \mathbf{s}'} P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}) \geq G \left(\Lambda_J \Gamma, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2} \right) \quad (3.72)$$

If we minimize both sides of Eq. (3.72), we can write

$$\min_{\mathbf{B}} \left(\max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K, \mathbf{s} \neq \mathbf{s}'} P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}) \right) \geq G \left(\Lambda_J \Gamma_{\text{op}}, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2} \right) \quad (3.73)$$

where $\mathbf{\Gamma}_{\text{op}}$ is obtained according to Eq. (3.68).

Let us now establish an upper bound for the worst case average pair-wise error probability for the specially structured transmission matrix \mathbf{B}_o above. For any error vector \mathbf{e} , we have

$$\begin{aligned} \det\left(\mathbf{I}_M + \frac{1}{8\sigma^2 \sin^2 \theta} \mathbf{\Sigma} \mathcal{X}_{\mathbf{B}_o}^H(\mathbf{e}) \mathbf{X}_{\mathbf{B}_o}(\mathbf{e})\right) \\ = \left(\frac{1}{8\sigma^2 \sin^2 \theta}\right)^M \det(\mathbf{\Lambda}_J \mathbf{\Gamma}^2) \det(\mathbf{\Delta} + \mathbf{X}_{\mathbf{I}_M}^H(\mathbf{e}) \mathbf{X}_{\mathbf{I}_M}(\mathbf{e})) \end{aligned} \quad (3.74)$$

where the special structure of \mathbf{B}_o has been utilized, $\mathbf{\Lambda}_J$ denotes the diagonal matrix containing the largest J positive eigenvalues of $\mathbf{\Sigma}$ and $\mathbf{\Delta} = (8\sigma^2 \sin^2 \theta) \mathbf{\Lambda}_J^{-1} \mathbf{\Gamma}^{-2}$. Using Eq. (3.74) in Property 3.7, for any nonzero vector \mathbf{e} and nonzero θ in the interval $[0, \pi/2]$, we have

$$\det(\mathbf{\Delta} + \mathbf{X}_{\mathbf{I}_M}^H(\mathbf{e}) \mathbf{X}_{\mathbf{I}_M}(\mathbf{e})) \geq \prod_{k=1}^M \left(\frac{8\sigma^2 \sin^2 \theta}{\gamma_k^2 \lambda_k} + d_{\min}^2(\mathcal{S}) \right) \quad (3.75)$$

where, according to Property 3.7, equality holds if and only if \mathbf{s} and \mathbf{s}' are neighbour vectors. Eq. (3.74) and Eq. (3.75) together yield

$$\det\left(\mathbf{I}_M + \frac{1}{8\sigma^2 \sin^2 \theta} \mathbf{\Sigma} \mathbf{X}_{\mathbf{B}_o}^H(\mathbf{e}) \mathbf{X}_{\mathbf{B}_o}(\mathbf{e})\right) \geq \prod_{k=1}^M \left(1 + \frac{d_{\min}^2(\mathcal{S}) \gamma_k^2 \lambda_k}{8\sigma^2 \sin^2 \theta} \right) \quad (3.76)$$

Again, substituting Eq. (3.76) in Eq. (3.65) and using the optimum $\mathbf{\Gamma}_{\text{op}}$ yields

$$\max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K, \mathbf{s} \neq \mathbf{s}'} P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}_{\text{op}}) \leq G \left(\mathbf{\Lambda}_J \mathbf{\Gamma}_{\text{op}}, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2} \right)$$

where equality holds if and only if $\|\mathbf{s} - \mathbf{s}'\| = d_{\min}(\mathcal{S})$. This results in

$$\begin{aligned} \min_{\mathbf{B}} \left(\max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K, \mathbf{s} \neq \mathbf{s}'} P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}) \right) &\leq \max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K, \mathbf{s} \neq \mathbf{s}'} P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}_{\text{op}}) \\ &\leq G \left(\mathbf{\Lambda}_J \mathbf{\Gamma}_{\text{op}}, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2} \right) \end{aligned} \quad (3.77)$$

Combining Eq. (3.77) with Eq. (3.73) yields

$$\min_{\mathbf{B}} \left(\max_{\substack{s, s' \in \mathcal{S}^K \\ s \neq s'}} P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}) \right) = G \left(\mathbf{\Lambda}_J \mathbf{\Gamma}_{\text{op}}, \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2} \right) \quad (3.78)$$

Eq. (3.78) holds iff $\mathbf{B} = \mathbf{\Gamma}_{\text{op}} \mathbf{V}_J^H$ and $\|\mathbf{s} - \mathbf{s}'\| = d_{\min}(\mathcal{S})$. Thus, the proof of Theorem 3.2 is complete. \square

Remarks on Theorem 3.2:

- a) Theorem 3.2 shows us that the lower bound of the worst case pair-wise error probability can be reached by having $\mathbf{B} = \mathbf{\Gamma}_{\text{op}} \mathbf{V}_J^H$. Thus, the design problem in Eq. (3.66) becomes finding the optimum $\mathbf{\Gamma}_{\text{op}}$ of Eq. (3.68).
- b) The original non-convex optimization problem has been transformed into the convex problem in Eq. (3.68) and can be solved efficiently by interior point methods. The convexity of the objective function can be verified by re-writing $G(\mathbf{\Lambda}_J \mathbf{\Gamma}, \varepsilon)$ in Eq. (3.67) as

$$G(\mathbf{\Lambda}_J \mathbf{\Gamma}, \varepsilon) = \frac{1}{\pi} \int_0^{\pi/2} \exp \left(- \sum_{j=1}^J \ln \left(1 + \frac{\varepsilon \lambda_j \gamma_j^2}{\sin^2 \theta} \right) \right) d\theta, \quad \varepsilon > 0 \quad (3.79)$$

We notice that $-\ln(\cdot)$ is a convex function over γ_j^2 , and hence their sum is also convex over $[\gamma_1^2, \dots, \gamma_J^2]$. Now, $\exp(x)$ is monotonically increasing with x . By *composition rule* [9] (Page 84), the integrand in Eq. (3.79) is a convex function implying that $G(\mathbf{\Lambda}_J \mathbf{\Gamma}, \varepsilon)$ is convex.

- c) The solution of Eq. (3.68) yields the values of the diagonal elements $\{\gamma_{\text{op}1}, \gamma_{\text{op}2}, \dots, \gamma_{\text{op}M}\}$. Some of these values may not be positive. We choose all the K positive ones to form the singular values of \mathbf{B}_{op} .

Theorem 3.2 provides us with an efficient scheme to obtain the optimal matrix \mathbf{B}_{op} by numerically minimizing $G(\mathbf{\Lambda}_J \mathbf{\Gamma}, \varepsilon)$. However, if the Chernoff bound [87] of the

pairwise error probability is employed as the objective function for minimization instead, a closed-form optimal \mathbf{B} can be obtained. This can be shown by setting $\sin^2 \theta = 1$ in the pairwise error probability of Eq. (3.65), so that we obtain the Chernoff bound as

$$P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}) \leq \frac{1}{2 \det(\mathbf{I} + (8\sigma^2)^{-1} \Sigma \mathcal{X}_{\mathbf{B}}^H(\mathbf{e}) \mathbf{X}_{\mathbf{B}}(\mathbf{e}))} \quad (3.80)$$

Seeking to minimize the worst case Chernoff bound, and following similar arguments which establish the the optimization problem in Eq. (3.68), we arrive at the following problem,

$$\tilde{\Gamma}_{\text{op}} = \arg \min_{\text{tr}(\tilde{\Gamma}) \leq 1} \frac{1}{2} \prod_{j=1}^J \left(1 + \frac{d_{\min}^2(\mathcal{S})}{8\sigma^2} \lambda_j \tilde{\gamma}_j^2 \right)^{-1} \quad (3.81)$$

where $\tilde{\Gamma}_{\text{op}}$ is a diagonal matrix with diagonal elements $\tilde{\gamma}_{\text{op}j}$. This problem is a relaxed form of that in Eq. (3.68) and its solution is provided by the following corollary:

Corollary 3.2 *The solution, $\tilde{\Gamma}_{\text{op}}$, for the optimization problem of Eq. (3.81) can be obtained by employing the water-filling strategy [11]. The diagonal elements of $\tilde{\Gamma}_{\text{op}}$ are given by*

$$\tilde{\gamma}_{\text{op}j} = \sqrt{\left[\frac{1}{M_0} \left(1 + \frac{8\sigma^2}{d_{\min}^2(\mathcal{S})} \sum_{\ell=1}^{M_0} \frac{1}{\lambda_{\ell}} \right) - \frac{1}{\lambda_n} \right]_+}, \quad (3.82)$$

$$j = 1, \dots, J \quad (3.83)$$

where notation $[x]_+$ denotes $\max(x, 0)$. The optimal choice of J is $J = M_0$, where M_0 is the maximum positive integer satisfying

$$\frac{1}{M_0} \left(1 + \frac{8\sigma^2}{d_{\min}^2(\mathcal{S})} \sum_{\ell=1}^{M_0} \frac{1}{\lambda_{\ell}} \right) - \frac{1}{\lambda_m} > 0, \quad m = 1, 2, \dots, M_0$$

The optimum transmission matrix is thus $\tilde{\mathbf{B}}_{\text{op}} = \tilde{\Gamma}_{\text{op}} \mathbf{V}_J^H$.

Proof. The minimization of the type of problems of Eq. (3.81) has been studied by several researchers and the water-filling solution is shown in [11]. \square

Remarks on Corollary 3.2:

- a) For the particular case in which the channel coefficients are mutually independent, i.e., $\mathbf{\Sigma} = \mathbf{I}_M$, then any $M \times M$ unitary matrix scaled by a factor $1/\sqrt{M}$ is a suitable choice for $\tilde{\mathbf{B}}_{\text{op}}$.
- b) Since $\tilde{\mathbf{B}}_{\text{op}}$ minimizes the Chernoff bound on the worst case PEP, it implicitly maximizes the coding gain [91], which is defined to be the normalized minimum determinant of $\mathbf{X}(\mathbf{e})^H \mathbf{X}(\mathbf{e})$ for all nonzero \mathbf{e} ; see Eq. (2.17).

Remarks on the Optimum Transmission Matrix Design for Linear Receivers:

- The derivations of Theorem 3.2 and Corollary 3.2 are based on the consideration of using an ML receiver in the MISO system. For the case when the system is equipped with a linear ZF (or MMSE) receiver under the environment of correlated channels, the problem of obtaining an optimum \mathbf{B} becomes very complicated. This is because we seek for a matrix \mathbf{B} to minimize the respective average error probability obtained by averaging the expressions of error probability in Eqs. (3.16), (3.12) and (3.19) for the respective signalling schemes over the random channel matrix. This requires the knowledge of the PDF of the equivalent channel matrix. However, the equivalent channel matrix in these cases is of a Toeplitz structure for which the PDF is unavailable, and therefore, the expressions for the average error probability cannot be obtained. (For the case of linear MMSE receivers, a similar problem exists).
- An alternative way to attack the problem in the case when a linear receiver is employed is to consider the upper bound on the averaged error probability given

in Eq. (3.25). We can minimize this upper bound with respect to \mathbf{B} . However, this necessitates the knowledge of the value of C_0 . From Property 3.4, C_0 is the minimum value of the determinant of the Toeplitz matrix having its column vector belonging to unit ball and this renders the solving of C_0 difficult. Thus, in this thesis, we are unable to come up with any true optimal \mathbf{B} for MISO systems equipped with linear receivers.

3.5 Numerical Experiments

In this section, we examine the performance of the Toeplitz code in a MISO system. We first evaluate the performance of the system equipped with a linear receiver under the condition that different parameters are varied. We then evaluate the performance of the system employing different beamformers as well as a linear or an ML receiver in an environment in which the channels are correlated. We note that for a linear receiver, the major computation occurs in the inverse of the Toeplitz matrix for which the complexity is of order $\mathcal{O}(KM)$ [34], where K is the length of signal vector and M is the number of transmitter antennas for the MISO system. On the other hand, the complexity using an ML detector for this MISO system transmitting the Toeplitz code is of order μ^M where μ is the constellation cardinality. Thus, for a reasonably large constellation and/or a comparatively large number of transmitter antennas, the ML receiver is substantially more complex than a linear receiver. Finally, we compare the performance of the Toeplitz code with some other known efficient codes.

Example 1: In this example, we examine the performance of Toeplitz STBC for a MISO communication system with independent channel fading, i.e., $\mathbf{\Sigma} = \mathbf{I}$. The system is equipped with a linear ZF detector at the receiver end. For the Toeplitz STBC, we choose $\mathbf{B} = \mathbf{I}$ in Eq. (3.28). The following three experiments are performed:

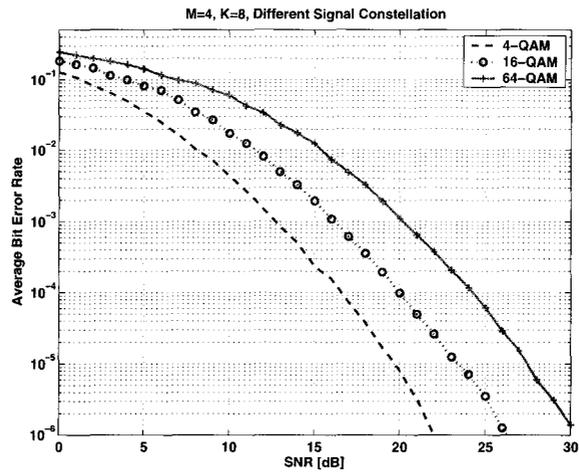


Figure 3.1: The average BER performance of the proposed Toeplitz STBC when signals are selected from different constellations.

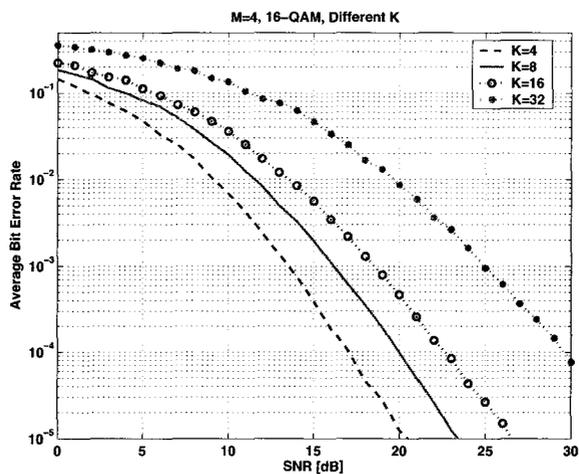


Figure 3.2: The average BER performance of the proposed Toeplitz STBC for different K .

1. We fix the number of transmitter antennas to be $M = 4$ and the length of signal vector \mathbf{s} to be $K = 8$, and the symbol transmission data rate is therefore $R_s = K/T = 0.7273$ symbols pcu. The signals are randomly selected respectively from the constellations of 4-QAM, 16-QAM and 64-QAM. The signals are transmitted through the MISO system having zero-mean unit-variance i.i.d. Gaussian channels and additive white Gaussian noise as described in Section 3.1 and the BER curves are plotted in Fig.3.1 in which the three different curves correspond to the performance of the system using the three signal constellations respectively. It should be noted that different constellation size results in different transmission bit rates. For the system we examine, the bit rates are $R_b = 1.4545$, 2.9091 , and 5.8182 bits pcu corresponding to 4-QAM, 16-QAM and 64-QAM respectively. Therefore, for larger constellation size, worse BER performance is expected. This is indeed the case as shown by the three BER curves plotted in Fig. 3.1 from which it is observed that, for a BER of 10^{-5} , the difference in SNR between 4-QAM and 16-QAM is approximately 3dB and that between 16-QAM and 64-QAM is approximately 5 dB.

2. In this experiment, we fix the signal constellation to be 16-QAM for a four transmitter antenna MISO system and perform simulations for different signal lengths, $K = 4, 8, 16, 32$. For different choices of L , the system has different transmission symbol rates, which are $R_s = 0.5714, 0.7273, 0.8421, 0.9143$ symbols pcu, respectively. The channel and noise assumptions are the same as those in the previous experiment. The BER curves at different SNR are plotted in Fig. 3.2. It can be observed that the longer is the transmitted signal, the worse is the system BER performance. Again, this is due to the fact that larger K corresponds to higher transmission data rate resulting in worse performance.

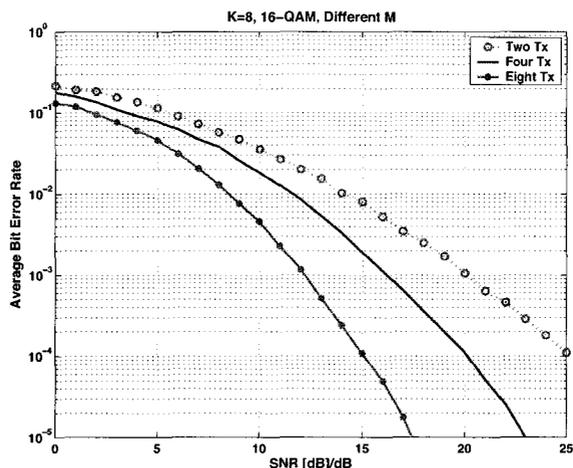


Figure 3.3: The average BER performance of the proposed Toeplitz STBC for the MISO system with different number of transmitter antennas.

3. In this experiment, we vary the number of the transmitter antennas M . An increase in M increases the diversity and decreases the transmission symbol rate. Therefore, it is expected that the performance of the system will be enhanced with the increase of the number of transmitter antennas. This is indeed the case as illustrated in Fig. 3.3 where we compare the performance of the MISO systems having $M = 2, 4$, and 8 antennas with $K = 8$ in the Toeplitz codes and the signals selected from a 16-QAM constellation.

Example 2: In this example, we test the performance of Toeplitz STBC for correlated channels in a MISO system equipped with four transmitter antennas in a linear array and the one receiver antenna on the normal to the axis of the transmitter antenna array (“broadside”). For small angle spread, the correlation coefficient between the

m_1 th and m_2 th transmitter antennas is [13, 115]

$$[\Sigma]_{m_1 m_2} \approx \frac{1}{2\pi} \int_0^{2\pi} \exp\left(-j2\pi(m_1 - m_2)\Delta \frac{d_t}{\varsigma} \sin\theta\right) d\theta \quad (3.84)$$

where d_t is the antenna spacing, ς is the wavelength of the (narrowband) signal, and Δ is the angle spread. Here in our simulations, we choose $d_t = 0.5\varsigma$ and $\Delta = 5^\circ$. We examine the performance of the MISO system transmitting 4-QAM signals in the following three cases using Toeplitz STBC having the structures:

- i) $\mathbf{B} = \frac{1}{\sqrt{M}}\mathbf{I}$. This is an approximately optimal transmission matrix at sufficiently high SNR in the minimization of the Chernoff bound under the assumption that \mathbf{B} is a square matrix. The approximate optimality can be shown as follows: For Eq. (3.80), under high SNR, we ignore the identity matrix \mathbf{I} in the denominator and obtain

$$P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{B}) < \frac{(8\sigma^2)^M}{2 \det(\Sigma \mathcal{X}_{\mathbf{B}}^H(\mathbf{e}) \mathbf{X}_{\mathbf{B}}(\mathbf{e}))} = \frac{(8\sigma^2)^M}{2 \det(\Sigma) \det(\mathbf{B}^H \mathbf{T}^H(\mathbf{e}) \mathbf{T}(\mathbf{e}) \mathbf{B})} \quad (3.85)$$

To minimize the right side of Eq. (3.85), we maximize the second determinant in the denominator. We note that

$$\begin{aligned} \det(\mathbf{B}^H \mathbf{T}^H(\mathbf{e}) \mathbf{T}(\mathbf{e}) \mathbf{B}) &\leq \det(\mathbf{T}^H(\mathbf{e}) \mathbf{T}(\mathbf{e})) \prod_{m=1}^M [\mathbf{B} \mathbf{B}^H]_{mm} \\ &\leq \det(\mathbf{T}^H(\mathbf{e}) \mathbf{T}(\mathbf{e})) \left(\frac{1}{M} \text{tr}(\mathbf{B} \mathbf{B}^H)\right)^M \\ &= \frac{1}{M^M} \det(\mathbf{T}^H(\mathbf{e}) \mathbf{T}(\mathbf{e})) \end{aligned} \quad (3.86)$$

The first inequality is due to Hadamard Inequality [43] and equality holds iff $\mathbf{B} \mathbf{B}^H$ is diagonal. The second inequality is due to geometric mean being no larger than arithmetic mean, with equality holding iff $\mathbf{B} \mathbf{B}^H$ has equal diagonal elements. Finally, the trace of $\mathbf{B} \mathbf{B}^H$ is equal to unity due to the power constraint. Hence, the condition for maximum in Eq. (3.86) is that \mathbf{B} is a scaled unitary matrix of which $\mathbf{B} = \mathbf{I}$ is one choice.

- ii) $\mathbf{B} = \mathbf{B}_{\text{op}}$. This is the optimal solution to minimize the worst case PEP derived in Theorem 3.2 and can be obtained numerically by solving the convex optimization problem of Eq. (3.68) and then using the result in Eq. (3.69).
- iii) $\mathbf{B} = \tilde{\mathbf{B}}_{\text{op}}$. This minimizes Chernoff bound on PEP as described in Corollary 3.2.

In the simulations, the transmitted signal vector is of length $K = 10$ and each of the symbols is randomly selected from the 4-QAM constellation. At the receiver, the signals in each of the three cases are detected separately by a linear ZF detector and an ML detector and the respective performances of the two detectors are examined. (As mentioned in the end of the last section, we cannot obtain an exact optimum \mathbf{B} for the ZF receiver. Nonetheless, we will employ the two optimum transmission matrices derived for the ML receiver \mathbf{B}_{op} and $\tilde{\mathbf{B}}_{\text{op}}$ to the case of ZF receiver to see how the performance is improved). We note that due to the variation of the channel fading, the dimension J of the optimum transmission matrices in Cases ii) and iii) change with SNR. For the specific correlated channel described in Eq. (3.84), for Case ii), we found that $J = 1$ when SNR ≤ 8 dB and $J = 2$ at higher SNR, whereas for Case iii), we found that $J = 1$ when SNR ≤ 10 dB and $J = 2$ at higher SNR. Therefore, the transmission data rate for these two cases is $R_1 = \frac{K}{K+J-1}$ symbols pcu, which is higher than $R_2 = \frac{K}{K+M-1}$ for the case of $\mathbf{B} = \mathbf{I}_M$. For a fair comparison, we choose two different structures for \mathbf{B} in Case i) in the following experiments:

1. We maintain the transmission data rate in Case i) the same as that in case iii). This is realized by setting $\mathbf{B} = \left[\frac{1}{\sqrt{K}} \mathbf{I}_J, \mathbf{0}_{(J, M-J)} \right]$ in Case i) with J being the dimension of $\tilde{\mathbf{B}}_{\text{op}}^H \tilde{\mathbf{B}}_{\text{op}}$. We evaluated the error performances of the systems equipped with different \mathbf{B} in all three cases and the results are shown in Fig. 3.4 from which the following observations can be made:

- For the system employing an ML detector, performance of Case ii) and

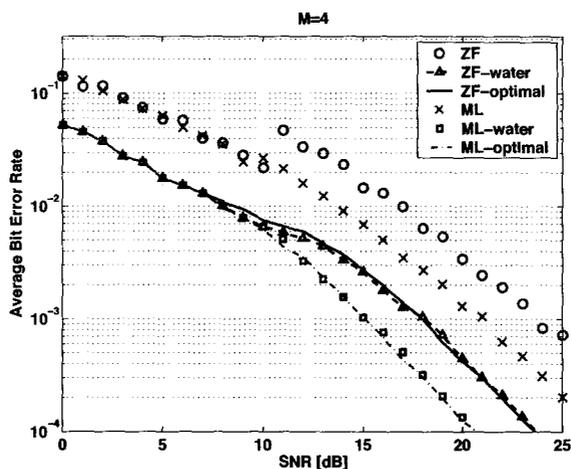


Figure 3.4: The average bit error rate comparison of the proposed Toeplitz STBC with i) $\mathbf{B} = [\mathbf{I}_J, \mathbf{0}_{M-J}]$, ii) \mathbf{B}_{op} and iii) $\tilde{\mathbf{B}}_{\text{op}}$. The performances are shown for both linear ZF detectors and ML detectors.

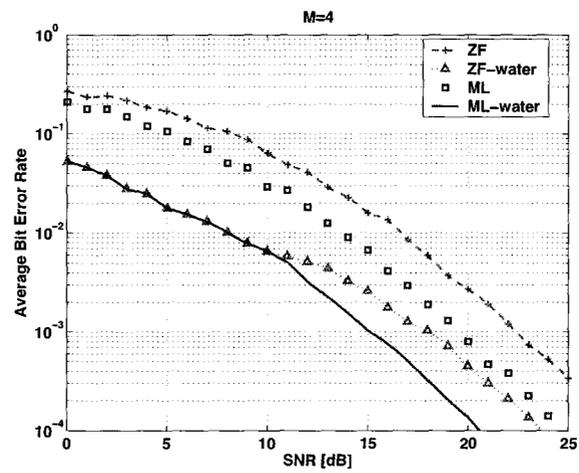


Figure 3.5: The average bit error rate comparison of the proposed Toeplitz STBC with $\mathbf{B} = \mathbf{I}_M$ and $\tilde{\mathbf{B}}_{\text{op}}$. The performances are shown for both linear ZF detectors and ML detectors.

- iii) are superior to that of Case i), confirming the theoretical analyses in Theorem 3.2 and Corollary 3.2.
- For the system employing an ML detector, the BER performance for Cases ii) and iii) employing \mathbf{B}_{op} and $\tilde{\mathbf{B}}_{\text{op}}$ respectively are very close. This shows that Chernoff bound is tight for this system. Close performance in the two cases is also true for the system using a ZF detector.
 - Although \mathbf{B}_{op} and $\tilde{\mathbf{B}}_{\text{op}}$ are optimal transmission matrices developed for the ML detector, they are equally effective in providing substantial performance improvement for the same system employing a linear ZF detector.
 - At lower SNR, we have $J = 1$, i.e., only one transmitter antenna is effective. Therefore, given a coded system, linear ZF and ML detectors provide the same performance.
2. In the second part of the experiment, we put $\mathbf{B} = \mathbf{I}_M$ in Case i) and examine its performance. The error performance for such a choice is shown in Fig. 3.5. Here, the system using $\mathbf{B} = \mathbf{I}_M$ has higher transmission data rate than those in Cases ii) and iii). For the sake of comparison, we have re-plotted in Fig. 3.5 the performance curves from Fig. 3.4 of Case iii) corresponding to the uses of $\tilde{\mathbf{B}}_{\text{op}}$ as a transmission matrix. (Since the performance of Cases ii) is almost the same as that of Case iii), we have omitted here the performance curves corresponding to the use of \mathbf{B}_{op}). It should be noted that when the signals are detected by an ML detector, the system coded with $\mathbf{B} = \mathbf{I}_M$ has higher diversity gain over the system with $\tilde{\mathbf{B}}_{\text{op}}$. This is due to the fact that water-filling strategy may not employ all the available transmitter antennas for correlated channels. Specifically in this example, the effective number of antennas for $\tilde{\mathbf{B}}_{\text{op}}$ is $J \leq 2 < 4$. However, the optimal coding gain achieved by $\tilde{\mathbf{B}}_{\text{op}}$ with

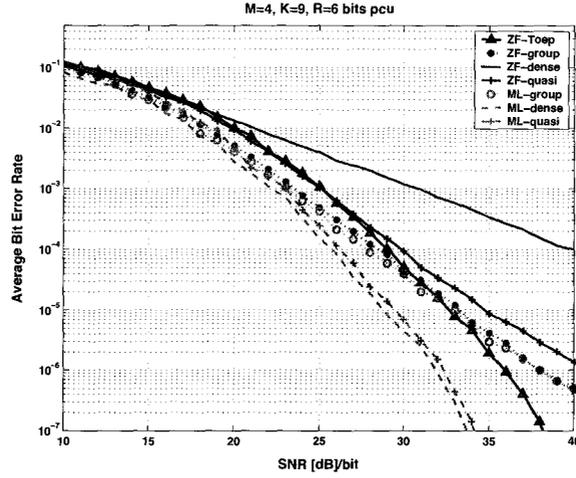


Figure 3.6: The average bit error rate comparison of the proposed Toeplitz STBC with other STBC of unit rate.

ML detectors ensures a better performance. It is also important to note that the employments of $\mathbf{B} = \mathbf{I}_M$ and $\tilde{\mathbf{B}}_{\text{op}}$ result in a relatively large difference in performances, revealing that the upper bound on PEP given in Eq. (3.85) is not tight. Thus, even though this bound is quite commonly employed in STBC designs for independent channels, the results here show that this relaxed bound is a poor design criterion for an environment of highly correlated channel coefficients.

Example 3: In this example, we compare the BER performance of Toeplitz STBC with other STBC for independent MISO channels. Here again, we choose $\mathbf{B} = \mathbf{I}$ for Toeplitz STBC. The experiments are performed for the two cases in which the number of transmitter antennas in the communication system are $M = 4$ and $M = 8$ respectively:

1. $M = 4$ transmitter antennas and a single receiver antenna: We compare BER performance of Toeplitz STBC with other rate one STBC [89, 41, 18, 47]:

- Quasi-orthogonal STBC. The code for four transmitter antennas was presented in [89], and the maximization of its coding gain was subsequently shown in [18].
- Dense full-diversity STBC [41]
- Multi-group decodable STBC [47]

For the Toeplitz STBC, we choose $K = 9$ for which the symbol transmission data rate is $R_s = K/T = 3/4$ symbols pcu. To achieve a fair comparison, the same transmission *bit* rate is imposed on all the codes such that signals are selected from 256-QAM constellation for Toeplitz STBC and from 64-QAM for the other full-rate STBC. Therefore, the same transmission bit rate, $R_b = 6$ bits pcu, is employed for all the systems. At the receiver, the Toeplitz STBC is processed by a linear ZF equalizer followed by a symbol-by-symbol detector. For the other full-rate STBC, we examine the two cases in which the signals are processed by a) an ML detector and b) a linear ZF receiver. The BER curves are plotted in Fig. 3.6. When a linear ZF equalizer and a symbol-by-symbol detector is applied at the receiver, it can be observed that Toeplitz STBC outperforms “quasi-orthogonal” STBC and “dense” STBC, and at higher SNR, its performance is superior to multi-group code. It is also interesting to observe that at higher SNR, for Toeplitz STBC with linear ZF receivers, the performance is also superior to that of the Multi-group STBC using an ML receiver. In fact, for the range of SNR tested, the slope of its BER curve is the same as those of the “dense” STBC and the “quasi-orthogonal” STBC processed by ML detectors, indicating they have the same diversity gain.

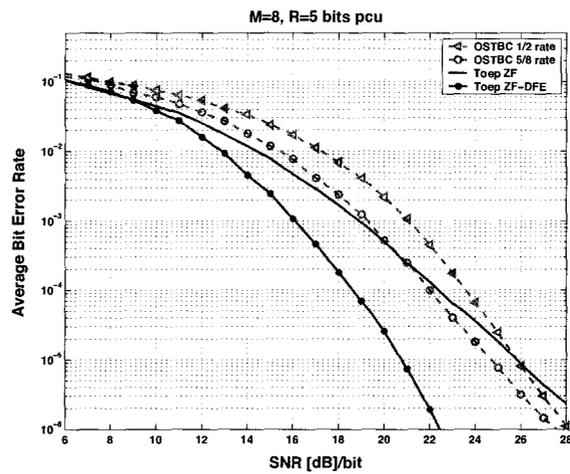


Figure 3.7: The average bit error rate comparison of the proposed Toeplitz STBC with the orthogonal STBC.

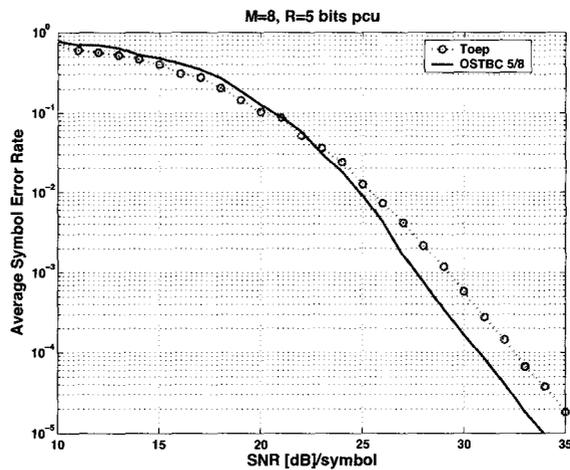


Figure 3.8: The average symbol error rate comparison of the proposed Toeplitz STBC with the orthogonal STBC.

2. We now consider the system having $M = 8$ transmitter antennas. For the Toeplitz code, we choose $K = 35$ and therefore, the symbol transmission data rate is $K/T = 5/6$ symbols pcu. We compare the bit error rate performance of our Toeplitz code with that of the orthogonal STBC having symbol transmission rate of:

- i) $1/2$ symbols pcu [90, 1, 94] and
- ii) $5/8$ symbols pcu [63] (this the highest symbol rate achievable by the orthogonal STBC applied to an eight transmitter antenna system).

To achieve a fair comparison, the transmitted signals are selected from a 64-QAM constellation for our Toeplitz code, a 256-QAM constellation for the $5/8$ rate orthogonal code and a 1024-QAM constellation for the $1/2$ rate orthogonal code. Hence, all of the codes have the same transmission data rate in bits, i.e., $R_b = 5$ bits pcu. At the receiver end, the orthogonal STBC is decoded by a linear ZF detector for which, because of the orthogonality, the performance is the same as that of an ML detector. For Toeplitz STBC, the signals are decoded separately by a linear ZF receiver and a ZF-DFE receiver. The average bit error rate for these codes are plotted Fig. 3.7. It can be observed that the performance of the Toeplitz code detected with a linear ZF receiver is superior to that of the $\frac{1}{2}$ -rate orthogonal STBC when the SNR is less than or equal to 25 dB. When the Toeplitz STBC is received by a ZF-DFE receiver, its performance is significantly better than that of the orthogonal STBC. In Fig. 3.7 at 10^{-5} , the Toeplitz code with a ZF-DFE receiver outperforms the orthogonal code by about 4 dB.

Here it should be noted that for the Toeplitz code, both linear ZF and ZF-DFE receivers can achieve full-diversity. However, this property is not reflected in Fig. 3.7 where the BER curves are not parallel. The reason for this is that

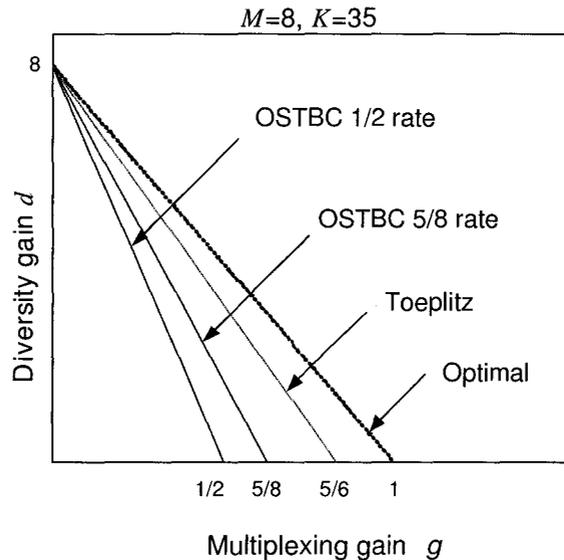


Figure 3.9: Diversity-multiplexing gain tradeoff curves for the proposed Toeplitz STBC and the orthogonal STBCs.

the employed SNR range is not sufficiently high. To simulate the high SNR BER performance in this example demands exhaustive computational costs. Therefore, we choose symbol error rate (SER) here to demonstrate the diversity gain. To show the diversity gain achieved by the Toeplitz code, we only need to compare the SER curve of the Toeplitz code with that of 5/8 orthogonal code when both are detected by a linear ZF receiver. The results are provided in Fig. 3.8. From the plotting, it is obvious that the curves for the two codes are parallel when SNR is above 30dB, indicating the same diversity gain. Hence, the Toeplitz code with a linear ZF receiver (and hence any superior detector) indeed achieves full diversity.

To compare the diversity and multiplexing gains achieved by the codes examined

in this example, in Fig. 3.9 the tradeoff curves associated with each of the three codes are provided. It can be observed that the curve for the Toeplitz STBC is closer to the optimal tradeoff curve than the other two. This curve will approach the optimal one as K increases.

3.6 Conclusion

In this chapter, we have presented a general design criterion for full-diversity linear STBC when the signals are transmitted through a MISO communication system and processed by a linear receiver. This is, to our knowledge, the first design criterion for linear receivers to achieve full diversity. Specifically, we proposed a linear Toeplitz STBC for a MISO channel which satisfies the criterion and achieves full-diversity. We have shown that such a code possesses many interesting properties, two of which recapitulated here are of practical importance:

1. The symbol transmission rate for the code approaches one when the number of channel uses ($T > M$) is large.
2. Employing the Toeplitz code results in a non-vanishing determinant.

When employed in a MISO system equipped with a linear receiver (ZF or MMSE), the Toeplitz code can provide full diversity. Furthermore, when the number of channel uses is large, in an independent MISO flat fading environment, the Toeplitz code can approach the Zheng-Tse optimal diversity v.s. multiplexing gain tradeoff.

When employed in a MISO system equipped with an ML detector, for both independent and correlated channel coefficients, we can design the transmission matrix inherent in the proposed Toeplitz STBC to minimize the exact worst case average pair-wise error probability resulting in full diversity and optimal coding gain being

achieved. In particular, when the design criterion of the worst case average pair-wise error probability is approximated by the Chernoff bound, we obtain a closed-form optimal solution.

The use of the Toeplitz STBC (having an identity transmission matrix) in a MISO system fitted with a ZF receiver has been shown by simulations to have the same slope of the BER curves to other full rate STBC employing an ML detector, whereas even better performance can be achieved by using receivers (such as ZF-DFE) more sophisticated than the linear ones to detect the Toeplitz code. For correlated channels, employing the optimum transmission matrices in the Toeplitz code results in substantial additional improvements in performance to using the identity transmission matrix. This substantial improvement of performance is observed in either case for which an ML or a ZF receiver is used.

Chapter 4

Diversity Gain Analysis of a Linear Receiver and Multiple Block Transmission

In the previous chapter, a special unitary trace-orthogonal code is presented which enables a linear receiver to achieve full diversity for a MISO system. More generally, for a MIMO system having $N \geq M$, the highest diversity gain achievable by a linear receiver is still unknown. In this chapter¹, we examine the maximum diversity gain achievable by a linear receiver for such a MIMO system, and we propose a multi-block transmission scheme to improve the system performance for a linear receiver.

Consider a MIMO ($N \geq M$) communication system where the signals are processed by a full rate linear STBC and detected by a linear MMSE equalizer followed by a symbol-by-symbol threshold detector. From the detection error probability expression, we analyze the maximum diversity gain associated with such a system and

¹The results related to this chapter has been submitted to *IEEE Transaction on Information Theory*, under second round of review. Part of the work has also been presented in *ISIT2006* [57] at Seattle, July 2006.

show that it is given by $(N - M + 1)$ for any square QAM signals. Employing the same criterion of minimizing the detection error probability, we extend the optimum design of STBC for the MIMO system to systems in which the data span L independent realizations of a block-static fading channel, and obtained the necessary and sufficient structures for the optimality of the code. The new system takes advantage of time diversity and the minimum achievable detection error probability is proved to be a decreasing function of L . Furthermore, the quantitative increase in diversity gain with L can be numerically evaluated, and shown to approach infinity as L grows. Thus, by increasing the number of blocks over which the data is transmitted and utilizing the time diversity provided, a simple linear MMSE receiver can achieve similar or higher diversity gain to that for an ML detector with a STBC designed over one block. These codes can be detected with a computational complexity to be little more than that for a single-block. Hence, when signals are from dense constellations, a simple linear receiver is able to outperform ML detectors that demands higher computational costs. A systematic method of generating such optimum codes for multi-block data transmission is also presented. The validity of our analysis is confirmed by simulation results.

4.1 Introduction

In MIMO communications, the design of STBC for high rate transmission with low detection error probability have attracted much of researchers' attention recently (e.g., [1, 91, 92, 37, 39, 94, 88, 51, 52, 22, 93]). *Diversity gain* is a concept central to code designs since it is an indication of the rate of decay of the error probability with SNR in the range of high SNR. The design of full-rate full-diversity STBC over Rayleigh flat fading channels have been proposed in [64, 15] based on the pair-wise

error probability for ML detectors. These codes are designed for the transmission of a *single* space-time block of coded data using M transmitter antennas in T time-slots and achieve full diversity, i.e., $d = MN$. However, when the signal constellation is large, the detection complexity for these codes is generally too high for practical applications.

A linear MMSE receiver, on the other hand, is simple to implement. The full-rate linear STBC that minimize the bit error rate (BER) for a MIMO system employing a linear MMSE receiver have been presented in [54, 61, 3]. The proposed codes have been shown to provide significantly better BER performance than existing STBC with linear receivers, but, due to the simplicity of the linear receivers, the diversity provided by a MIMO system cannot be fully exploited. From the expression of the minimum detection error probability, we will show in this chapter that the *maximum* diversity gain achieved by a linear receiver for a MIMO system is lower than the full diversity MN , resulting in the performance of such systems being inferior to that of codes designed for ML detectors.

Instead of designing the STBC for the transmission of one block of data, by assuming that the channel coefficients remain constant for one block and change independently from block to block, the codeword design has been extended [23] to cover multiple blocks for MIMO systems employing ML receivers. The code construction for this multi-block transmission that achieves the optimal diversity v.s. multiplexing gain tradeoff with the use of ML detection was proposed in [62]. Unfortunately, the high detection complexity renders the resulting code difficult to implement in practice. On the other hand, for such a system employing a linear MMSE receiver, experiments have been performed in [33]. The results show that by interchanging part of the signals between different blocks, the system error performance may be improved.

In this chapter, we consider the design of optimum linear STBC covering L blocks for a MIMO system equipped with a linear MMSE receiver. The new scheme takes advantage of time diversity provided by extra blocks that improves the system performance, and at the same time, involves low computational costs by choosing a linear receiver. As in [23], we assume that the channel coefficients remain constant for one block and may vary independently from block to block. (The independence of the channel changes can be made practically feasible by interleaving the transmission so that the channel states for one data block are separated by sufficiently large time intervals.) For these L -block system, we present necessary and sufficient code structures for the minimization of the detection error probability, and show that this minimum error probability is a decreasing function of L . We also show that for codes with these optimal structures, the normalized detection complexity has the same highest order as that of the code designed for a single-block, i.e., $\mathcal{O}(M^3)$, with the second-order term growing only linearly with L . Furthermore, by evaluating the analytic expression of the detection error probability numerically, we show that the diversity order grows with L . Hence, a proper choice of L will render the diversity gain of such codes comparable to that of FRFD codes designed for a single-block employing an ML detector. For a system in which latency can be accommodated, such a scheme would provide high diversity without the high computational complexity associated with ML detections. This is especially true when signals are selected from denser constellations.

4.2 Optimal Linear STBC for Single-Block Transmission and Performance Analysis

4.2.1 Optimal Linear STBC for Single-Block Transmission

Consider a MIMO communication system through which one block data is transmitted, i.e.,

$$\mathbf{Y} = \sqrt{\frac{\rho}{M}} \mathbf{H} \mathbf{X} + \mathbf{W} \quad (4.1)$$

The transmission code is specified to be a full rate linear STBC

$$\mathbf{X} = \sum_{i=1}^{MT} s_i \mathbf{C}_i \quad (4.2)$$

which satisfied the power constraint

$$\sum_{i=1}^{MT} \text{tr}(\mathbf{C}_i^H \mathbf{C}_i) = MT \quad (4.3)$$

s_i are the transmitted symbols selected from a particular constellation. Statistically, the symbols s_i are assumed to be independent with zero mean and unit variance. In the case of full-rate single-block transmission, the goal is to design optimal \mathbf{C}_i yielding minimum detection error probability for a linear MMSE receiver while satisfying the power constraint. Several researchers have studied this problem in the past few years [37,39,80,3,33,61]. In particular, the following theorem [61] provides necessary and sufficient structures for \mathbf{C}_i to be optimal in BER for a system with 4-QAM signalling and linear MMSE equalization, under the standard approximation [74] [76] that the residual inter symbol interference (ISI) is Gaussian:

Theorem 4.1 *Consider a MIMO system transmitting 4-QAM signals using a full-rate linear STBC and a linear MMSE receiver. Under Gaussian approximation of the*

interference at the detector, the average BER, \mathcal{P}_{es} , has a lower bound given by

$$\mathcal{P}_{\text{es}} \geq E_{\mathbf{H}} \left\{ Q \left(\sqrt{\frac{M}{\text{tr} \left((\mathbf{I} + \frac{\rho}{M} \mathbf{H}^H \mathbf{H})^{-1} \right)} - 1} \right) \right\} \triangleq \mathcal{P}_{\text{min}} \quad (4.4)$$

where $E_{\mathbf{H}}(\cdot)$ denotes expectation over the random channels. Equality in Eq. (4.4) holds if and only if the code matrices \mathbf{C}_i satisfy the following two conditions:

- i) \mathbf{C}_i is trace-orthogonal, i.e., $\text{tr}(\mathbf{C}_i \mathbf{C}_j^H) = \delta_{ij}$ for $i, j = 1, 2, \dots, MT$, where δ_{ij} is the Kronecker delta;
- ii) \mathbf{C}_i is unitary up to a constant, i.e., $\mathbf{C}_i \mathbf{C}_i^H = \frac{1}{M} \mathbf{I}$. ■

Theorem 4.1 states that the jointly unitary and trace-orthogonal code structure is necessary and sufficient for minimizing the average BER for 4-QAM signals under single-block transmission. These optimality conditions also apply to any square QAM constellation at high SNR [61]. They are also optimal in minimizing MSE for other mode of signalling. A code satisfying Theorem 4.1 can be systematically generated from a normalized DFT matrix [61].

4.2.2 Performance Analysis

For codes designed for single-block transmissions, the minimum BER given by Theorem 4.1 is in the form of the expected value of a Q -function over the random channels. To gain insight into the role of each parameter in the performance of a system employing such codes, we would like to express the minimum BER in a form independent of the random channels. Specifically, we examine the minimum BER at high SNR in terms of diversity gain d defined in Eq. (2.15). Therefore, we concentrate on the dominant term (the term containing the lowest order of ρ^{-1}) in the asymptotic performance at high SNR, i.e.,

$$\mathcal{P}_{\text{min}} = \mathcal{K}^{-1} \rho^{-d} + (\text{terms involving higher order of } \rho^{-1})$$

where \mathcal{K} is the coefficient often referred to as the *coding gain* [91].

Since the channels h_{nm} are assumed to be IID zero-mean Gaussian random variables, the matrix product $(\mathbf{H}^H \mathbf{H})$ is of *Wishart Distribution* [72, 78]. Let λ_m , $m = 1, \dots, M$ denote the eigenvalues of $(\mathbf{H}^H \mathbf{H})$ ordered as $\lambda_1 > \lambda_2 > \dots > \lambda_M > 0$. (Inequality between adjacent eigenvalues occurs almost surely.) Then [72, 78], the joint PDF of the eigenvalues is given by

$$p(\lambda_1, \dots, \lambda_M) = \alpha(M, N) \exp\left(-\sum_{i=1}^M \lambda_i\right) \prod_{i=1}^M \lambda_i^{N-M} \prod_{i < j}^M (\lambda_i - \lambda_j)^2 \quad (4.5)$$

where $\alpha(M, N) = \pi^{M(M-1)} / [2^{MN/2} \Gamma_{cM}(M) \Gamma_{cM}(N)]$ with $\Gamma_c(\cdot)$ being the Gamma function for complex multivariate defined as $\Gamma_{cM}(\nu) = \pi^{M(M-1)/2} \prod_{i=1}^M \Gamma(\nu - i + 1)$.

Observe that in Eq. (4.4) the random channel appears in the form of matrix $\mathbf{H}^H \mathbf{H}$ only, and the trace term in the denominator can be expressed as

$$\text{tr}\left(\mathbf{I} + \frac{\rho}{M} \mathbf{H}^H \mathbf{H}\right)^{-1} = \sum_{m=1}^M \left(1 + \frac{\rho}{M} \lambda_m\right)^{-1} \quad (4.6)$$

Using the standard bound on the Q -function [87],

$$Q(z) \leq \frac{1}{2} \exp\left(-\frac{z^2}{2}\right), \quad z \geq 0, \quad (4.7)$$

and combining Eqs. (4.6) and (4.7) we have the following upper bound on the minimum BER

$$\mathcal{P}_{\text{min}} \leq \frac{\sqrt{e}}{2} \int \dots \int \exp\left\{-\frac{M}{2 \sum_{m=1}^M \left(1 + \frac{\rho}{M} \lambda_m\right)^{-1}}\right\} p(\lambda_1, \dots, \lambda_M) d\lambda_1 \dots d\lambda_M \quad (4.8)$$

where the joint PDF of the eigenvalues $p(\lambda_1, \dots, \lambda_M)$ is given by Eq. (4.5). For finite number of transmission antennas, by evaluating Eq. (4.8) and focusing on the factor consisting of the lowest order of ρ^{-1} , we obtain the following result:

Theorem 4.2 For a MIMO system ($N \geq M$) transmitting 4-QAM signals coded with the optimal linear STBC described in Theorem 4.1 at full symbol rate, the diversity gain $d(N, M)$ and coding gain $\mathcal{K}(N, M)$ obtained by employing a linear MMSE receiver are respectively given by

$$d(N, M) = N - M + 1 \quad (4.9a)$$

$$\mathcal{K}^{-1}(N, M) = \frac{\alpha(M, N)2^{N-M}e^{-(M-1)/2}(N-M)!}{\alpha(M-1, N+1)}. \quad (4.9b)$$

Proof. See Appendix A. □

Remarks:

1. Although the derivation provided here is for 4-QAM signals only, Theorem 4.2 can be extended to a general square QAM signalling with $2b$ bits per symbol, where b is a positive integer. In this case, the minimum BER achieved by unitary trace-orthogonal code is given by [76, 108, 10, 61]

$$\begin{aligned} \tilde{\mathcal{P}}_{\text{min}} = & \mathbb{E}_{\mathbf{H}} \left\{ \zeta Q \left(\sqrt{\vartheta \left(\frac{M}{\text{tr} \left(\left(\mathbf{I} + \frac{\rho}{M} \mathbf{H}^H \mathbf{H} \right)^{-1} \right)} - 1 \right)} \right) \right. \\ & \left. + \eta Q \left(3 \sqrt{\vartheta \left(\frac{M}{\text{tr} \left(\left(\mathbf{I} + \frac{\rho}{M} \mathbf{H}^H \mathbf{H} \right)^{-1} \right)} - 1 \right)} \right) \right\} \quad (4.10) \end{aligned}$$

at high SNR, where $\zeta = \frac{2^b-1}{b2^{b-1}}$, $\eta = \frac{2^b-2}{b2^{b-1}}$ and $\vartheta = \frac{3 \cdot 2^b}{4^b-1}$. We observe that ζ , η and ϑ are all positive, and Eq. (4.10) is merely a linear transformation of Eq. (4.4) with parameters scaled by positive constants. These constants are independent of ρ , and appear as coefficients of the functions of ρ . Therefore, Eqs. (4.10) and (4.4) have the same power of ρ , i.e., the same diversity order. However, the coding gain will be different depending on the signal constellation.

2. Eq. (4.9a) shows a linear relation between diversity gain and $(N - M)$. This implies that increasing the system's degree of freedom (i.e., M and N) cannot guarantee an improved diversity gain for a linear receiver.

3. It has been shown [104,35,4] that the diversity gain of a MIMO system equipped with ZF receivers and transmitting *uncoded* symbols (i.e., $\mathbf{F} = \mathbf{I}$ in Eq. (4.15)) is also $(N - M + 1)$. Therefore, we conclude that the use of an optimal STBC to minimize BER does not increase the diversity gain when a linear receiver is employed. The lower BER obtained here is due to the increase in coding gain.

The diversity achieved in Eq. (4.9a) is inferior to the full diversity MN obtainable for single-block coded systems employing an ML detector. In the next section, we will present a multi-block transmission scheme in order to increase the diversity of a MIMO systems with a linear MMSE receiver.

4.3 Optimal Linear STBC for Multi-Block Transmission with Linear MMSE Receivers

In the design of STBC for single-block transmission described in Section 4.2.1, a data symbol is distributed among the MT elements of the matrix \mathbf{X} and has no distribution beyond the signal block to which it belongs. For multi-block transmission, the code is designed to have a signal symbol distributed over all the elements of the L block signal matrices \mathbf{X}_ℓ , $\ell = 1, \dots, L$. Intuitively, the more independent channels the same information is transmitted through, the greater is the potential diversity advantage. In this section, we examine the design of multi-block linear STBC for use with a linear MMSE receiver and show how a substantial increase in diversity can be achieved.

Consider L blocks of transmitted signals to be jointly coded. The channel matrix for the ℓ th block is \mathbf{H}_ℓ , which is assumed to remain unchanged during the T time slots and may vary independently after that interval. For full-rate transmission, we require

LMT symbols to be transmitted during these L blocks, i.e., $\{s_i\}$, $i = 1, 2, \dots, LMT$. Each block of coded signal matrix \mathbf{X}_ℓ is constructed such that

$$\mathbf{X}_\ell = \sum_{i=1}^{LMT} s_i \mathbf{C}_{i\ell}, \quad \ell = 1, 2, \dots, L \quad (4.11)$$

where $\mathbf{C}_{i\ell}$ are the $M \times T$ matrices to be designed and are constrained to have the power evenly allocated over all blocks,

$$\sum_{i=1}^{LMT} \text{tr}\{\mathbf{C}_{i\ell}^H \mathbf{C}_{i\ell}\} = MT, \quad \forall \ell \quad (4.12)$$

Using Eq. (4.11), the received signal for each transmitted block is given by

$$\mathbf{Y}_\ell = \sqrt{\frac{\rho}{M}} \mathbf{H}_\ell \left(\sum_{i=1}^{LMT} s_i \mathbf{C}_{i\ell} \right) + \mathbf{W}_\ell, \quad \ell = 1, \dots, L \quad (4.13)$$

Since the received signals from each block contain information from all LMT symbols, we need to jointly detect the signals from all L blocks at the receiver. Therefore, defining the LMT transmitted symbols as a vector $\mathbf{s} = [s_1 \cdots s_{LMT}]^T$ and, vectorizing each block of received data in Eq. (4.13), writing $\mathbf{y}_\ell = \text{vec}(\mathbf{Y}_\ell)$ and $\mathbf{w}_\ell = \text{vec}(\mathbf{W}_\ell)$, each being of dimension $NT \times 1$, and stacking them up as long vectors, we have

$$\underbrace{\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_L \end{pmatrix}}_{\mathbf{y}} = \sqrt{\frac{\rho}{M}} \underbrace{\begin{pmatrix} \mathbf{I}_T \otimes \mathbf{H}_1 & & & \\ & \mathbf{I}_T \otimes \mathbf{H}_2 & & \\ & & \ddots & \\ & & & \mathbf{I}_T \otimes \mathbf{H}_L \end{pmatrix}}_{\mathcal{H}} \underbrace{\begin{pmatrix} \text{vec}(\mathbf{C}_{11}) & \cdots & \text{vec}(\mathbf{C}_{(LMT)1}) \\ \text{vec}(\mathbf{C}_{12}) & \cdots & \text{vec}(\mathbf{C}_{(LMT)2}) \\ \vdots & & \vdots \\ \text{vec}(\mathbf{C}_{1L}) & \cdots & \text{vec}(\mathbf{C}_{(LMT)L}) \end{pmatrix}}_{\mathbf{F}} \mathbf{s} + \underbrace{\begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_L \end{pmatrix}}_{\mathbf{w}} \quad (4.14)$$

Equivalently, we can write

$$\mathbf{y} = \sqrt{\frac{\rho}{M}} \mathcal{H} \mathbf{F} \mathbf{s} + \mathbf{w} \quad (4.15)$$

where \mathbf{y} and \mathbf{w} are $LNT \times 1$ vectors, being respectively the received signal vector and the white Gaussian noise, \mathcal{H} is the equivalent channel matrix of dimension $LNT \times LMT$, and \mathbf{F} is the $LMT \times LMT$ square coding matrix to be designed. We note that the i th column, \mathbf{f}_i , of this coding matrix \mathbf{F} is an LMT -dimensional vector containing the LMT codes for the i th symbol s_i in all L blocks, i.e.,

$$\mathbf{f}_i = [\text{vec}^T(\mathbf{C}_{i1}) \quad \text{vec}^T(\mathbf{C}_{i2}) \quad \cdots \quad \text{vec}^T(\mathbf{C}_{iL})]^T \quad (4.16)$$

Now, the signal vector \mathbf{y} in Eq. (4.15) undergoes MMSE equalization at the receiver. This is carried out by left multiplying both sides of Eq. (4.15) with an equalizer matrix \mathbf{G} [83] as shown in Eq. (2.31), i.e.,

$$\mathbf{G} = \sqrt{\frac{\rho}{M}} (\mathbf{I} + \frac{\rho}{M} \mathbf{F}^H \mathcal{H}^H \mathcal{H} \mathbf{F})^{-1} \mathbf{F}^H \mathcal{H}^H \quad (4.17)$$

The equalized signal $\hat{\mathbf{s}}$ is obtained as

$$\hat{\mathbf{s}} = \mathbf{G} \mathbf{y} = \sqrt{\frac{\rho}{M}} \mathbf{G} \mathcal{H} \mathbf{F} \mathbf{s} + \mathbf{G} \mathbf{w} \quad (4.18)$$

which is then detected by a symbol-by-symbol detector. The *symbol error* covariance matrix of the input to this detector is given by

$$\mathcal{E}_{\text{se}} \triangleq \text{E}\{(\mathbf{s} - \hat{\mathbf{s}})(\mathbf{s} - \hat{\mathbf{s}})^H\} = \left(\mathbf{I} + \frac{\rho}{M} \mathbf{F}^H \mathcal{H}^H \mathcal{H} \mathbf{F} \right)^{-1} \quad (4.19)$$

In order to analyze the *bit error* probability of the detected symbol, it is necessary to specify the mode of modulation of the transmitted signal \mathbf{s} , and here we stipulate the scheme to be 4-QAM. Writing $\boldsymbol{\sigma} = [\mathbf{s}_{\text{re}}, \mathbf{s}_{\text{im}}]^T$, and $\hat{\boldsymbol{\sigma}}$ to be the corresponding equalized signal, then the *bit error* covariance matrix \mathbf{V} is given by

$$\boldsymbol{\mathcal{E}} \triangleq \text{E}\{(\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}})(\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}})^T\} = \left[\mathbf{I} + \frac{\rho}{M} \mathbf{J}^H \hat{\mathbf{F}}^H \hat{\mathcal{H}}^H \hat{\mathcal{H}} \hat{\mathbf{F}} \mathbf{J} \right]^{-1} \quad (4.20)$$

where $\hat{\mathcal{H}}$ is of dimension $2LNT \times 2LMT$, $\hat{\mathbf{F}}$ and \mathbf{J} are both $2LMT \times 2LMT$ and are respectively defined as

$$\hat{\mathcal{H}} \triangleq \begin{pmatrix} \mathcal{H} & \mathbf{0} \\ \mathbf{0} & \mathcal{H}^* \end{pmatrix}; \quad \hat{\mathbf{F}} \triangleq \begin{pmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}^* \end{pmatrix}; \quad \mathbf{J} \triangleq \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{I}_{LMT} & j\mathbf{I}_{LMT} \\ \mathbf{I}_{LMT} & -j\mathbf{I}_{LMT} \end{pmatrix} \quad (4.21)$$

The k th element of the signal vector $\boldsymbol{\sigma}$ has its MSE given by the k th diagonal element of $\boldsymbol{\mathcal{E}}$. Since \mathbf{J} is unitary, an important relationship can be observed directly from Eqs. (4.19), (4.20) and (4.21) that

$$\text{tr } \boldsymbol{\mathcal{E}} = 2 \text{tr } \boldsymbol{\mathcal{E}}_{\text{se}} \quad (4.22)$$

As is in the case of single-block STBC design, the residual ISI from the MMSE equalizer can be approximated by zero-mean Gaussian noise, and hence the error. Average error probability for the k th bit can be written as [77] $\text{E}_{\boldsymbol{\mathcal{H}}} \left\{ Q \left(\sqrt{[\boldsymbol{\mathcal{E}}]_{kk}^{-1}} - 1 \right) \right\}$. The averaged error probability over the $2LMT$ bits is therefore

$$\mathcal{P}_{\text{em}} \approx \text{E}_{\boldsymbol{\mathcal{H}}} \left\{ \frac{1}{2LMT} \sum_{k=1}^{2LMT} Q \left(\sqrt{[\boldsymbol{\mathcal{E}}]_{kk}^{-1}} - 1 \right) \right\} \quad (4.23)$$

The objective here is to obtain the structure of the STBC \mathbf{F} such that it is optimal in BER performance, i.e., $\min_{\mathbf{F}} : \mathcal{P}_{\text{em}}$ subject to the power constraints in Eq. (4.12). The solution to this problem is provided by the following theorem:

Theorem 4.3 *For a MIMO system employing a full-rate linear STBC designed for multi-block transmission and equipped with a linear MMSE receiver, under Gaussian approximation of the interference at the detector, the average BER \mathcal{P}_{em} has a lower bound given by*

$$\mathcal{P}_{\text{em}} \geq E \left\{ Q \left(\sqrt{\frac{LM}{\sum_{\ell=1}^L \text{tr} \left((\mathbf{I} + \frac{\rho}{M} \mathbf{H}_{\ell}^H \mathbf{H}_{\ell})^{-1} \right)} - 1} \right) \right\} \triangleq \mathcal{P}_{\text{mmin}} \quad (4.24)$$

where the expectation is taken over the random channels. Equality in (4.24) holds if and only if the following two conditions are satisfied:

- i) Each of the coding sub-matrices $\mathbf{C}_{i\ell}$ formed from the i th column of \mathbf{F} as shown in Eq. (4.16) is unitary up to a scale, i.e.,

$$\mathbf{C}_{i\ell}\mathbf{C}_{i\ell}^H = \frac{1}{ML}\mathbf{I} \quad (4.25)$$

- ii) The $LMT \times LMT$ matrix \mathbf{F} is unitary, i.e. $\mathbf{F}^H\mathbf{F} = \mathbf{I}$. This condition is equivalent to:

$$\sum_{\ell=1}^L \text{tr}(\mathbf{C}_{i\ell}^H\mathbf{C}_{j\ell}) = \delta_{ij}. \quad (4.26)$$

Proof. See Appendix B. □

Remarks on Theorem 4.3:

1. Theorem 4.1 is a special case of Theorem 4.3, i.e., for $L = 1$, the two theorems are identical. The code satisfying Eqs. (4.25) and (4.26) is called *multi-block unitary trace-orthogonal* code.
2. Theorem 4.3 provides a lower bound on the bit error probability when 4-QAM signals are transmitted through the multi-block coded MIMO systems equipped with a linear MMSE receiver. Following similar arguments as those in [54,55,61], the two conditions can be proved to be necessary and sufficient for minimizing the BER at high SNR for any square QAM signal constellation, and to be optimal in minimizing MSE for a general signalling.
3. As mentioned before, $\mathbf{C}_{i\ell}$ is the code for the i th symbol transmitted through the ℓ th realization of channel. Therefore, the first optimality condition is a condition imposed on the i th symbol such that: i) for each given channel state, the symbol is distributed evenly over all channels using a unitary coding matrix. ii) there is no constraint on the relation between code matrices for different channel realizations, iii) the power allocated to transmit the symbol through each

channel realizations is equal. The second optimality condition keeps orthogonal relations between different symbols while maintaining equal power allocated to these symbols.

4. Eq. (4.24) shows that $\mathcal{P}_{\text{mmin}}$ is a function of L , i.e., $\mathcal{P}_{\text{mmin}}(L)$. In the following, we will prove that $\mathcal{P}_{\text{mmin}}(L) > \mathcal{P}_{\text{mmin}}(L + 1)$, i.e., the detection error probability decreases with the increase of L . For notation simplicity, we write $\mathcal{P}_{\text{mmin}}(L) = \mathbb{E} \left\{ \Upsilon \left(\frac{1}{L} \sum_{\ell=1}^L \xi_{\ell} \right) \right\}$, where $\Upsilon(x) \triangleq Q \left(\sqrt{\frac{1}{x} - 1} \right)$, and $\xi_{\ell} = \frac{1}{M} \text{tr} \left(\mathbf{I} + \frac{\rho}{M} \mathbf{H}_{\ell}^H \mathbf{H}_{\ell} \right)^{-1}$. First, we note that $\Upsilon(x)$ is a convex function w.r.t. x when $0 < x \leq 1$ (e.g., [61]). Now, consider having $(L + 1)$ independent blocks, among which we choose any combination of L different blocks for transmission. In total, there are $(L + 1)$ different combinations. For each combination of L blocks, we jointly transmit signals coded in the way as described in Theorem 4.3 and the expected value of the detection error probability over the random channels is $\mathcal{P}_{\text{mmin}}(L)$. Averaging over all the $L + 1$ combinations, this value remains unchanged, i.e.,

$$\mathcal{P}_{\text{mmin}}(L) = \mathbb{E} \left\{ \Upsilon \left(\frac{1}{L} \sum_{\ell=1}^L \xi_{\ell} \right) \right\} = \frac{1}{L+1} \sum_{i=1}^{L+1} \mathbb{E} \left\{ \Upsilon \left(\frac{1}{L} \sum_{\ell=1, \ell \neq i}^{L+1} \xi_{\ell} \right) \right\} \quad (4.27)$$

Applying Jensen's inequality [9] to the right side of Eq. (4.27) which contains the convex function $\Upsilon(x)$, we arrive at

$$\mathcal{P}_{\text{mmin}}(L) \geq \mathbb{E} \left\{ \Upsilon \left(\frac{1}{L+1} \sum_{i=1}^{L+1} \left(\frac{1}{L} \sum_{\ell=1, \ell \neq i}^{L+1} \xi_{\ell} \right) \right) \right\} \quad (4.28a)$$

$$= \mathbb{E} \left\{ \Upsilon \left(\frac{1}{L+1} \sum_{\ell=1}^{L+1} \xi_{\ell} \right) \right\} = \mathcal{P}_{\text{mmin}}(L+1) \quad (4.28b)$$

Equality in Eq. (4.28a) holds iff all the $(L + 1)$ blocks are identical, which violates the assumption that each of the L blocks changes independently after

T time slots. Therefore, we have strict inequality for independently fading channel blocks.

4.4 Performance Analysis: Multi-Block Transmission Using Optimum Code

In the previous section, we have derived the minimum BER \mathcal{P}_{min} for multi-block transmission in the form of the expected value of Q -function over random channels. We have also proved that \mathcal{P}_{min} decreases as L increases. However, to understand the roles that each parameter plays, we would like to evaluate the expectation in Eq. (4.24). First, by writing $\boldsymbol{\lambda}_\ell = [\lambda_{\ell 1} \cdots \lambda_{\ell M}]$, with $\lambda_{\ell m}$ being the m th eigenvalue for $\mathbf{H}_\ell^H \mathbf{H}_\ell$, \mathcal{P}_{min} can be written as

$$\mathcal{P}_{\text{min}} = \int \cdots \int Q \left(\sqrt{\frac{ML}{\sum_{\ell=1}^L \sum_{m=1}^M (1 + \frac{\rho}{M} \lambda_{\ell m})^{-1}} - 1} \right) p(\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_L) d\boldsymbol{\lambda}_1 \cdots d\boldsymbol{\lambda}_L \quad (4.29)$$

where $p(\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_L)$ denotes the joint PDF of all $\lambda_{\ell m}$, $\ell = 1, \cdots, L$, $m = 1, \cdots, M$. Since the channel matrices \mathbf{H}_ℓ are statistically independent for different ℓ , we have $p(\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_L) = \prod_{\ell=1}^L p(\boldsymbol{\lambda}_\ell)$, where each $p(\boldsymbol{\lambda}_\ell) = p(\lambda_{\ell 1}, \cdots, \lambda_{\ell M})$ satisfies Eq. (4.5). Due to the complexity of the Q -function and $p(\boldsymbol{\lambda}_\ell)$, the integration in Eq. (4.29) is intractable. However, we observe from Eq. (4.29) that \mathcal{P}_{min} depends on the parameters (M, N) , L and ρ , where (M, N) is introduced by multiple antenna system and L is from multi-block design. To examine the properties of \mathcal{P}_{min} in relation to multi-block design, in the following, we evaluate numerically \mathcal{P}_{min} in terms of ρ , L , and (M, N) , keeping $M = N$ so that $d = 1$ for $L = 1$.

A. *Finite L:*

Since the channel eigenvalues are independent from one block to another, the evaluation of Eq. (4.29) with respect to L is relatively simple. However, the channel eigenvalues are not independent within each block, hence this evaluation could be quite tedious for large M . The task can be made simpler following the general procedure outlined below:

- i) We let $\xi_{\ell m} = (1 + \frac{\rho}{M} \lambda_{\ell m})^{-1}$, so that $p_{\xi_{\ell}}(\xi_{\ell 1}, \dots, \xi_{\ell M}) = p(\lambda_{\ell 1}, \dots, \lambda_{\ell M}) / \left| \frac{\partial(\xi_{\ell 1}, \dots, \xi_{\ell M})}{\partial(\lambda_{\ell 1}, \dots, \lambda_{\ell M})} \right|$.
- ii) Let $\zeta_{\ell} = \sum_{m=1}^M \xi_{\ell m}$. Then we can obtain $p_{\zeta_{\ell}}(\zeta_{\ell})$ by successive integration such that, for $\ell = 1, 2, \dots, L$,

$$p_{\zeta_{\ell}}(\zeta_{\ell}) = \int \cdots \int p(\zeta_{\ell} - u_{M-1}, u_{M-1} - u_{M-2}, \dots, u_2 - u_1, u_1) du_1 \cdots du_{M-1} \quad (4.30)$$

where $u_1 = \xi_{\ell M}$, $u_2 = \xi_{\ell(M-1)} + \xi_{\ell M}$, \dots , $u_{M-1} = \xi_{\ell 2} + \dots + \xi_{\ell M}$.

- iii) Let $\bar{\zeta} = \sum_{\ell=1}^L \zeta_{\ell}$. Since the random variables ζ_{ℓ} are independent from block to block, we have,

$$p_{\bar{\zeta}}(\bar{\zeta}) = p_{\zeta_1}(\zeta_1) \star p_{\zeta_2}(\zeta_2) \star \cdots \star p_{\zeta_L}(\zeta_L) \quad (4.31)$$

where “ \star ” stands for convolution which, again, can be efficiently carried out using the FFT algorithm..

- iv) Let $z = 1/\bar{\zeta}$, then $p_{\xi}(\xi_{\ell}) = p(\bar{\zeta}) / \left| \frac{dz}{d\bar{\zeta}} \right|$ and Eq. (4.29) becomes

$$\mathcal{P}_{\text{mmin}} = \int Q \left(\sqrt{LMz - 1} \right) p_z(z) dz \quad (4.32)$$

For finite L and M , the above procedure yields, numerically, the value of $\mathcal{P}_{\text{mmin}}$, from which we can appraise the gain obtained by having multi-block design. In the following, we examine the cases for $N = M = 1$ and $N = M = 2$. The study of these two simple cases helps to illustrate the effects of the multi-block design on the system performance:

Figs. 4.1 and 4.2 show the variation of \mathcal{P}_{min} w.r.t. SNR for $N = M = 1$ and $N = M = 2$ respectively, numerically evaluated using the above procedure. In both cases, we vary the values of L . Observations from these figures not only confirm that \mathcal{P}_{min} decreases as L increases, as stated in Remark 4 on Theorem 4.3, it also shows numerically how much \mathcal{P}_{min} decreases with L in these two cases. We observed, as expected, the slope of the curve is independent of the value of $M(= N)$ for $L = 1$, i.e, the two cases have identical diversity gains which also confirms the result in Theorem 4.2. However, for $L > 1$, \mathcal{P}_{min} also varies as the value of $M(= N)$ changes. From these figures, it is observed that the diversity gain for $M = 2$ is higher (greater negative slope at high SNR) than that for $M = 1$. Fig. 4.3 shows the performance of the two cases at different SNR with increasing L . We observe that at lower SNR, the case of $M = N = 1$ outperforms the case of $M = N = 2$, and each settles at its respective performance level after $L > 30$. On the other hand, at higher SNR, the performance of the case of $M = N = 2$ is better than that of the case of $M = N = 1$. For both cases, the performance improves with increasing L , as predicted by the analysis in Section 4.3. The diversity gains of the two cases can be calculated by evaluating the negative slope of the graphs in Figs. 4.1 and 4.2 at high SNR. The slopes for the relatively high SNR of 22dB, are plotted in Fig. 4.4. It is observed here that while the diversity gain for $M = N = 2$ is higher than that for $M = N = 1$, for both cases, they are increasing with L .

B. For asymptotically large L :

When the code is jointly designed for a large number of blocks, we can analyze the effect of L in the following way. Let the fractional part in the argument of the Q -function in Eq. (4.29) be written as $1/[\frac{1}{L}\sum_{\ell=1}^L \frac{1}{M}\sum_{m=1}^M (1 + \frac{\rho}{M}\lambda_{\ell m})^{-1}]$. As $L \rightarrow \infty$, we can apply the law of large numbers, and write the denominator as

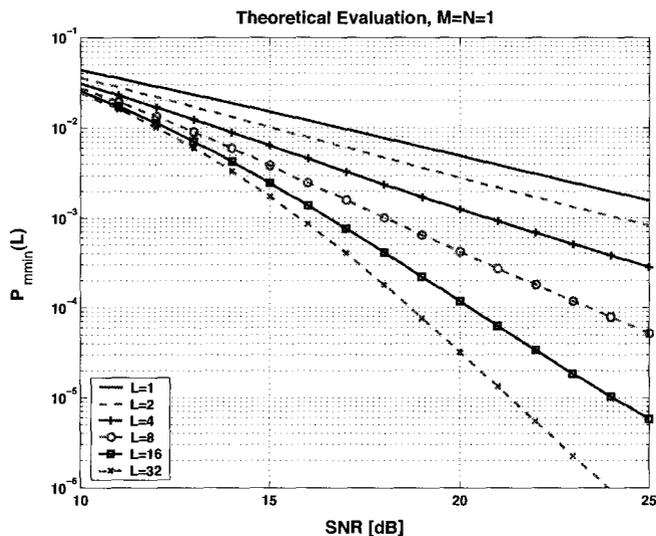


Figure 4.1: P_{\min} vs SNR for different L for the case of $M = N = 1$, and $R = 2$ bits per channel use.

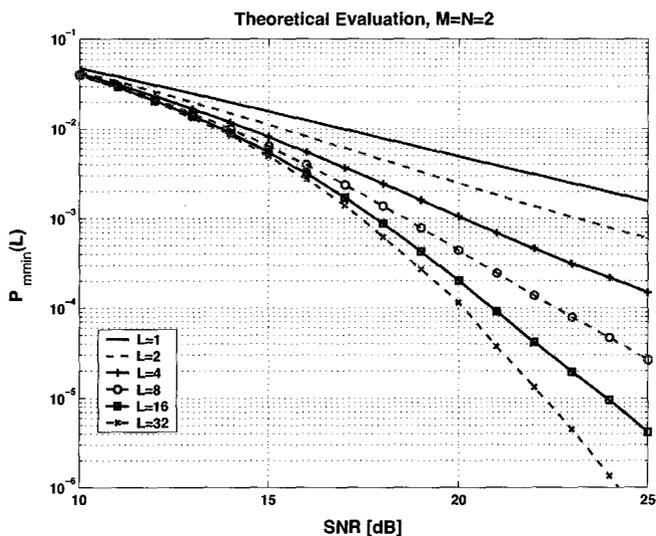


Figure 4.2: P_{\min} vs SNR for different L for the case of $M = N = 2$, and $R = 4$ bits per channel use.

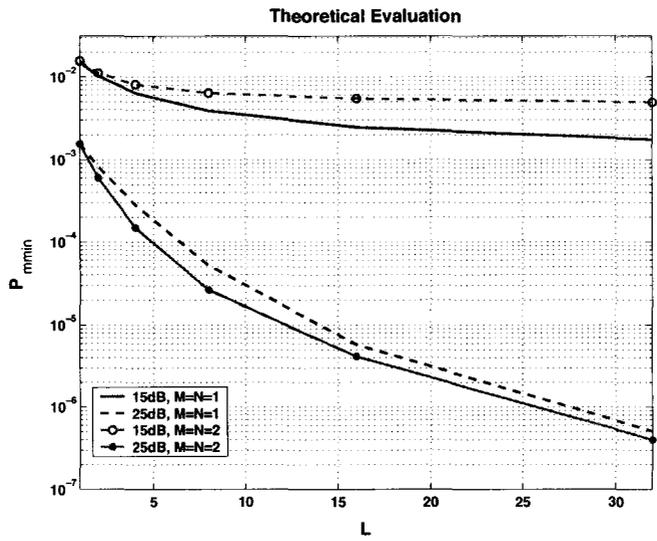


Figure 4.3: P_{\min} vs L .

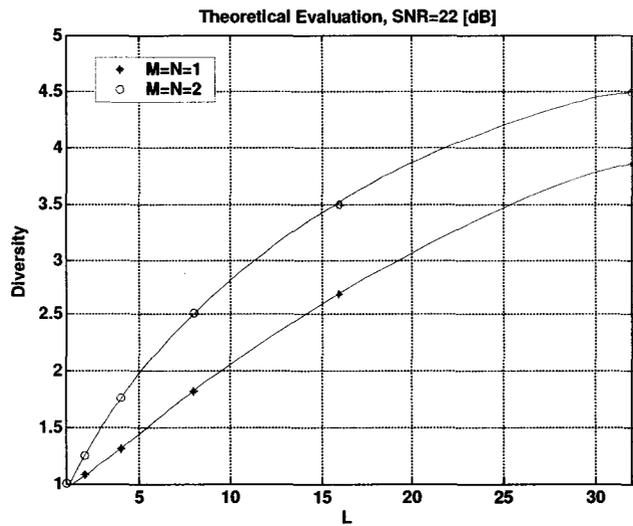


Figure 4.4: Diversity vs L .

$$\bar{\xi} = \mathbb{E} \left\{ \frac{1}{M} \sum_{m=1}^M \frac{1}{1 + \frac{\rho}{M} \lambda_m} \right\} = \frac{1}{M} \sum_{m=1}^M \mathbb{E} \left\{ \frac{1}{1 + \frac{\rho}{M} \lambda_m} \right\} = \mathbb{E} \left\{ \frac{1}{1 + \frac{\rho}{M} \lambda} \right\} \quad (4.33)$$

where the last step is valid for any positive number M since the summand is a constant. The PDF of λ is the marginal density which is obtainable from the joint PDF of the *ordered* eigenvalues $\lambda_1 > \dots > \lambda_M$ in Eq. (4.5) by first deriving the joint PDF of the *unordered* eigenvalues and then integrating, resulting in [92]

$$p(\lambda) = \frac{1}{M} \lambda^{N-M} e^{-\lambda} \sum_{k=0}^{M-1} \frac{k!}{(k+N-M)!} [L_k^{N-M}(\lambda)]^2 \quad (4.34)$$

where $L_k^{N-M}(\lambda) = \sum_{i=0}^k \binom{k+N-M}{k-i} \frac{(-\lambda)^i}{i!}$ is the associated Laguerre polynomial of order $(N-M)$ and degree k , with $\binom{n}{k}$ being the binomial coefficient. Substituting Eq. (4.34) into Eq. (4.33), we obtain

$$\bar{\xi} = \frac{1}{M} \int_0^{\infty} \frac{\lambda^{N-M} \sum_{p=0}^{2(M-1)} a_p \lambda^p}{1 + \frac{\rho}{M} \lambda} e^{-\lambda} d\lambda \quad (4.35)$$

where, for notational simplicity, we have re-written the square of $L_k^{N-M}(\lambda)$ as a polynomial in ascending powers of λ having coefficients a_p . If we let $x = \lambda + \frac{M}{\rho}$, Eq. (4.35) can be written as

$$\begin{aligned} \bar{\xi} &= \rho^{-1} e^{M/\rho} \sum_{p=N-M}^{N+M-2} a_p \int_{\frac{M}{\rho}}^{\infty} \left(x - \frac{M}{\rho} \right)^p \frac{e^{-x}}{x} dx \\ &= \rho^{-1} e^{M/\rho} \sum_{p=N-M}^{N+M-2} \left\{ a_p \sum_{q=0}^p \binom{p}{q} \left(\frac{M}{\rho} \right)^{p-q} \int_{\frac{M}{\rho}}^{\infty} x^{q-1} e^{-x} dx \right\} \end{aligned} \quad (4.36)$$

In the following, we analyze the diversity gain for the two cases:

1) $N = M = 1$: In this case, Eq. (4.36) equals

$$\bar{\xi} = \rho^{-1} e^{(\rho^{-1})} \int_{\rho^{-1}}^{\infty} \frac{e^{-x}}{x} dx = -\rho^{-1} e^{(\rho^{-1})} \text{Ei}(-\rho^{-1}) \quad (4.37)$$

where the exponential integral is defined as [50] $\text{Ei}(\varepsilon) \triangleq \int_{-\infty}^{\varepsilon} \frac{e^t}{t} dt = \gamma + \ln(-\varepsilon) + \sum_{k=1}^{\infty} \frac{\varepsilon^k}{k!k}$, with γ being Euler's constant $\gamma = 0.5772157\dots$. Therefore, Eq. (4.37) can be written as

$$\bar{\xi} = e^{(\rho^{-1})} \rho^{-1} \left(\ln \rho - \gamma - \sum_{k=1}^{\infty} \frac{(-\rho^{-1})^k}{k!k} \right)$$

where the terms inside parenthesis are ordered in descending power of ρ . At high SNR, we have

$$\bar{\xi}|_{\rho \rightarrow \infty} = \rho^{-1} \ln \rho \quad (4.38)$$

Since $z = 1/(LM\bar{\xi})$, applying the upper bound of Q -function in Eq. (4.7) to Eq. (4.32), and using Eq. (4.38), the upper bound on the minimum error probability is given by

$$\mathcal{P}_{\text{min}} \Big|_{\substack{L \rightarrow \infty \\ \rho \rightarrow \infty}} \leq \frac{1}{2e} \exp\left(-\frac{\rho}{2 \ln \rho}\right) \quad (4.39)$$

It can be shown that $\lim_{\rho \rightarrow \infty} \frac{\exp(-\frac{\rho}{2 \ln \rho})}{\rho^{-K}} = 0$ for any finite positive integer K , and we conclude:

Assertion 4.1 *For $N = M = 1$ and for $L \rightarrow \infty$, the diversity gain of the MIMO system with multi-block code design tends to infinity.* \square

The numerical evaluation of Eq. (4.32) with the sample average over L replaced by the expected value substituted into the Q -function is plotted in Fig. 4.5, and indeed confirms the above assertion.

2) $\max(N - M, M - 1) \geq 1$: Let $B_{pq} = a_p M^{p-q} \binom{p}{q}$. Then the term inside the

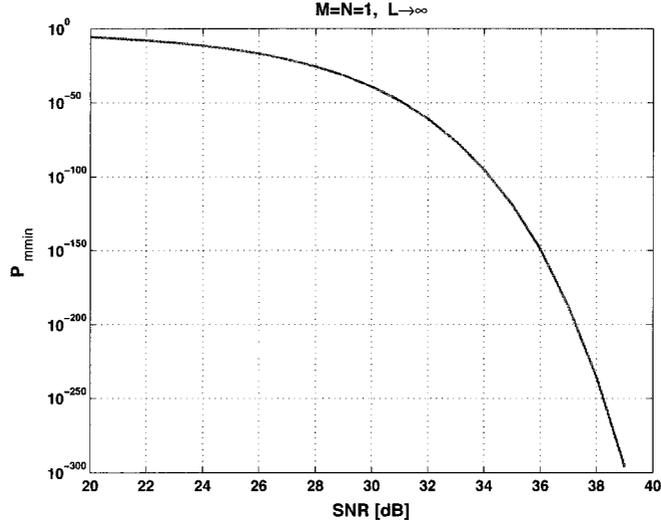


Figure 4.5: Theoretical evaluation of \mathcal{P}_{\min} w.r.t. SNR for $L \rightarrow \infty$

braces in Eq. (4.36) can be integrated. Evaluating it at high SNR, we have

$$\begin{aligned}
 & B_{pq} \rho^{-(1+p-q)} e^{M/\rho} \int_{\frac{M}{\rho}}^{\infty} x^{q-1} e^{-x} dx \\
 &= \begin{cases} B_{pq} \rho^{-(1+p-q)} \sum_{\eta=0}^{q-1} \left(\frac{M}{\rho}\right)^{\eta} \frac{(q-1)!}{\eta!} & q \geq 1 \\ B_{pq} \rho^{-(1+p-q)} e^{M/\rho} \left(\ln \frac{\rho}{M} - \gamma - \sum_{k=1}^{\infty} \frac{(-M\rho^{-1})^k}{k!k} \right) \Big|_{\rho \rightarrow \infty} \\ \approx B_{pq} \rho^{-(1+p)} \ln \rho & q = 0 \end{cases} \quad (4.40)
 \end{aligned}$$

Comparing Eq. (4.40) with Eq. (4.38), we observe that the terms Eq. (4.40) are of higher order of ρ^{-1} than that in Eq. (4.38). Therefore, following from Assertion 4.1 we have,

Assertion 4.2 For $\max(N - M, M - 1) \geq 1$ and for $L \rightarrow \infty$, the diversity gain of the MIMO system with multi-block code design tends to infinity. \square

4.5 Generation and Detection of the Optimum Multi-Block Code

4.5.1 Code Generation

There are various ways of generating space-time code matrices that satisfy the optimality conditions in Theorem 4.3. Here, we present a simple method based on DFT matrices. The basic building block of the code is the $M \times T$ matrix $\mathbf{C}_{i\ell}$ given by Eq. (4.16), $i = 1, 2, \dots, LMT$, $\ell = 1, 2, \dots, L$. To construct these building blocks, the following simple procedure can be taken:

1. We first generate MT matrices $\mathbf{\Omega}_{mt}$, $m = 1, 2, \dots, M$, $t = 1, 2, \dots, T$, $T \geq M$, using the following steps:

- (a) Form a $T \times T$ row permutation matrix such that

$$\mathbf{P} = \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{I}_{T-1} & \mathbf{0} \end{bmatrix}$$

where \mathbf{I}_{T-1} is the $(T-1) \times (T-1)$ identity matrix.

- (b) Form the $M \times M$ normalized DFT matrix $\mathbf{D}_M = [\mathbf{d}_M(1) \ \mathbf{d}_M(2) \ \cdots \ \mathbf{d}_M(M)]$, with $\mathbf{d}_M(m)$ being its m th column.

- (c) Generate the matrices $\mathbf{\Omega}_{mt}$ according to the following equation:

$$\mathbf{\Omega}_{mt} = [\text{Diag}(\mathbf{d}_M(m)) \mid \mathbf{0}] \mathbf{P}^{t-1}, \quad m = 1, 2, \dots, M, \quad t = 1, 2, \dots, T$$

where $\text{Diag}(\mathbf{x})$ is the matrix obtained by putting the M elements of the vector \mathbf{x} into an $M \times M$ diagonal matrix, and $\mathbf{0}$ is an $M \times (T-M)$ zero matrix.

It can be easily verified that all the MT matrices $\mathbf{\Omega}_{mt}$ satisfy the unitarity and trace-orthogonality conditions stated in Theorem 4.1, and hence $\{\mathbf{\Omega}_{mt}\}$ forms an optimal single-block code.

2. Now, generate an $L \times L$ normalized DFT matrix and denote it by \mathbf{D}_L . The set of optimal building block code matrices can now be obtained by

$$[\mathbf{C}_{i1}^T, \mathbf{C}_{i2}^T, \dots, \mathbf{C}_{iL}^T]^T = \mathbf{d}_L(\ell) \otimes \mathbf{\Omega}_{mt},$$

$$m = 1, 2, \dots, M, t = 1, 2, \dots, T, \ell = 1, 2, \dots, L \quad (4.41)$$

where $i = (mt - 1)L + \ell$ and $\mathbf{d}_L(\ell)$ denotes the ℓ th column vector of the matrix \mathbf{D}_L . \square

We now verify that the code matrix generated by the proposed algorithm satisfies the optimal conditions for a multi-block code in Theorem 4.3: Since $\mathbf{\Omega}_{mt}$ is a scaled unitary matrix and since each element of \mathbf{D}_M has equal non-zero magnitude, it follows that $\mathbf{C}_{i\ell}$ is a scaled unitary matrix, and hence the second condition in Theorem 4.3 is satisfied. To simplify our verification of the first condition, we first map the double index mt in an arbitrary one-to-one fashion to the single index, r , where $r = 1, 2, \dots, MT$. Now, for $i = (r_1 - 1)L + \ell_1$ and $j = (r_2 - 1)L + \ell_2$, we have

$$\begin{aligned} \sum_{\ell=1}^L \text{tr}(\mathbf{C}_{i\ell}^H \mathbf{C}_{j\ell}) &= \text{tr}\left(\left(\mathbf{d}_L(\ell_1) \otimes \mathbf{\Omega}_{r_1}\right)^H \left(\mathbf{d}_L(\ell_2) \otimes \mathbf{\Omega}_{r_2}\right)\right) \\ &= \text{tr}\left(\left(\mathbf{d}_L(\ell_1)^H \mathbf{d}_L(\ell_2)\right) \otimes \left(\mathbf{\Omega}_{r_1}^H \mathbf{\Omega}_{r_2}\right)\right) \\ &= \delta_{\ell_1 \ell_2} \cdot \text{tr}(\mathbf{\Omega}_{r_1}^H \mathbf{\Omega}_{r_2}) = \delta_{\ell_1 \ell_2} \delta_{r_1 r_2} = \delta_{ij} \end{aligned}$$

which is the first condition in Theorem 4.3.

4.5.2 Detection of the Optimum Multi-block Code

For an L -block STBC, a total of LMT symbols have to be jointly processed at the receiver. Thus, the complexity of a given receiver is expected to be larger than that of the corresponding receiver for the single-block case. However, as we will show below, for codes with the optimal structures given in Theorem 4.3, the per-symbol complexity

of the linear MMSE receiver is only a little more than that of the single-block linear MMSE receiver.

The computational cost of a linear (multi-block) MMSE receiver is dominated by the cost of obtaining the equalized signal vector

$$\hat{\mathbf{s}} = \mathbf{G}\mathbf{y} = \sqrt{\frac{\rho}{M}} \left(\mathbf{I} + \frac{\rho}{M} \mathbf{F}^H \mathcal{H}^H \mathcal{H} \mathbf{F} \right)^{-1} \mathbf{F}^H \mathcal{H}^H \mathcal{H} \mathbf{y} \quad (4.42)$$

If unitary trace-orthogonal signalling is employed, in accordance with Theorem 4.3, the special structure of the code matrix \mathbf{F} can be used to reduce the cost of evaluating Eq. (4.42). First, the unitarity of \mathbf{F} can be used to rewrite \mathbf{G} as

$$\mathbf{G} = \sqrt{\frac{\rho}{M}} \mathbf{F}^H \text{Blkdiag}(\Psi_1, \dots, \Psi_L) \quad (4.43)$$

where $\text{Blkdiag}(\cdot)$ denotes a block diagonal matrix with diagonal blocks, and

$$\Psi_\ell = \mathbf{I}_T \otimes \left(\mathbf{I}_M + \frac{\rho}{M} \mathbf{H}_\ell^H \mathbf{H}_\ell \right)^{-1} \mathbf{H}_\ell^H, \quad \ell = 1, \dots, L.$$

Eq. (4.43) is arrived at by using the definition of \mathcal{H} in Eq. (4.14) together with the property of Kronecker product that $(\mathbf{C} \otimes \mathbf{D})^{-1} = \mathbf{C}^{-1} \otimes \mathbf{D}^{-1}$. We observe that the expression in Eq. (4.43) contains the inverse of L matrices of size $M \times M$, whereas the expression in Eq. (4.42) contains the inverse of a matrix of size $LMT \times LMT$. If we substitute Eq. (4.43) into Eq. (4.42), the equalized signal vector can be written as

$$\hat{\mathbf{s}} = \sqrt{\frac{\rho}{M}} \mathbf{F}^H \boldsymbol{\gamma}$$

where $\boldsymbol{\gamma}$ is an $LMT \times 1$ vector defined by $\boldsymbol{\gamma} = ((\Psi_1 \mathbf{y}_1)^T, \dots, (\Psi_L \mathbf{y}_L)^T)^T$. We will focus on the number of multiplications required in each operation. Computing $\boldsymbol{\gamma}$ requires $L(8M^3/3 + 2NM^2 + NTM)$ multiplications. Computing $\hat{\mathbf{s}}$ requires a further $(LMT)^2$ multiplications. However if one adopts the particular code structure proposed in Section 4.5.1, the cost of computing $\hat{\mathbf{s}}$ can be substantially reduced. From

Table 4.1: Computation comparison for two detectors in detecting MT symbols for $M = N = T = 2$, and Q-PSK.

Detector	Multiplications
LMMSE, $L = 1$	49
LMMSE, $L = 4$	69
LMMSE, $L = 32$	193
ML, single block	5120

the code construction algorithm in Section 4.5.1, we observe that each column of matrix \mathbf{F} has ML nonzero elements, each of which is generated from two DFT matrices, \mathbf{D}_M and \mathbf{D}_L . As a result, each element of the equalized signal $\hat{\mathbf{s}}$ can be written as

$$\sum_{\ell=0}^{L-1} \sum_{m=0}^{M-1} W_M^{K_1 m} W_L^{K_2 \ell} \xi_{m+M\ell} = \sum_{\ell=0}^{L-1} W_L^{K_2 \ell} \left(\sum_{m=0}^{M-1} W_M^{K_1 m} \xi_{m+M\ell} \right) \quad (4.44)$$

where $K_1 \leq M - 1$ and $K_2 \leq L - 1$ are integers, and W_K is the primitive K th root of unity. Eq. (4.44) can be computed efficiently by standard FFT algorithm. For example, if M and L are self-composite with base 2, then computing $\hat{\mathbf{s}}$ requires $LMT(L \log_2 M + \log_2 L)$ multiplications [21]. By exploiting the FFT algorithm in the computation of $\hat{\mathbf{s}}$, the number of multiplications required to detect each block of MT symbols is reduced to $8M^3/3 + 2NM^2 + MNT + MT(L \log_2 M + \log_2 L)$. For comparison, the ML detector for a single block transmission scheme with a signal of cardinality size μ requires $MT(MT + 1)\mu^{MT}$ multiplications.

To provide a concrete comparison between the computational cost of the multi-block linear MMSE receiver with the signalling scheme provided in Section 4.5.1 and the ML receiver for single block communication, we consider a MIMO system with $M = N = T = 2$ and a signal constellation of size $\mu = 4$. The number of multiplications required to detect each block of MT symbols is provided in Table 4.1. From this

table, it can be seen that with the signalling scheme proposed in Section 4.5.1, the computational cost per-block of multi-block signalling grows slowly with the number of blocks, L , and is much less than that of the ML detector for signal block transmission. A performance comparison of the systems in Table 4.1 will be provided in Example 2 in the ensuing section.

4.6 Numerical Experiments

In this section, we examine the BER performance of the optimum STBC discussed in the previous sections by computer simulations. We present two examples. In Example 1, we examine the performance of single-block transmission and focus on verifying the theoretical analysis of the diversity order provided in Theorem 4.2. In Example 2, we perform experiments on the proposed multi-block transmission scheme to verify the properties of $\mathcal{P}_{\text{mmin}}$ discussed in Section 4.4. In this example, we also compare the BER performance of the multi-block transmission scheme with that of some other available schemes. In both examples, signals are generated by randomly choosing symbols from square QAM constellations in an IID fashion.

Example 1: We evaluate the BER performance of the single-block transmission scheme discussed in Sections 4.2.1 and 4.2.2. Here, our emphasis is on verifying the diversity analysis in Section 4.2.2, for a system with 4-QAM signalling. Performance comparisons with other STBC were provided in [61].

- (1) In the first experiment, we consider a system with $M = T = 2$, and hence the transmission data rate is $R = 4$ bits per channel use (pcu). The averaged BER for $N = 2, 3$ and 4 receiver antennas is plotted against SNR in Fig. 4.6. The negative slopes of BER curves in Fig. 4.6 increase with N and are equal $(N - M + 1)$, verifying the results in Theorem 4.2.

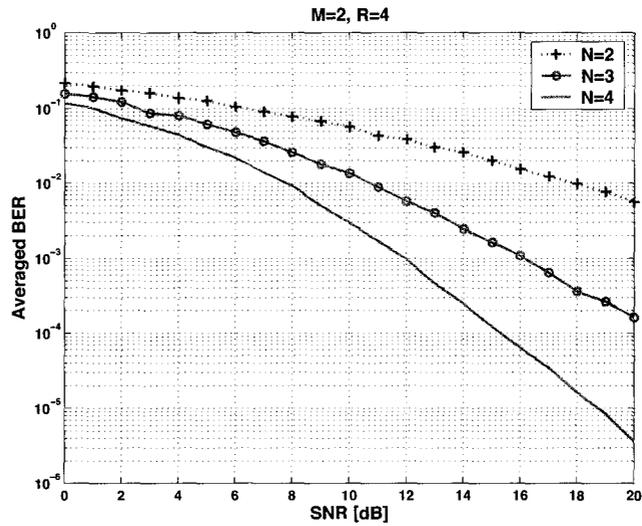


Figure 4.6: Single-block transmission. $M = 2, N = 2, 3, 4$.

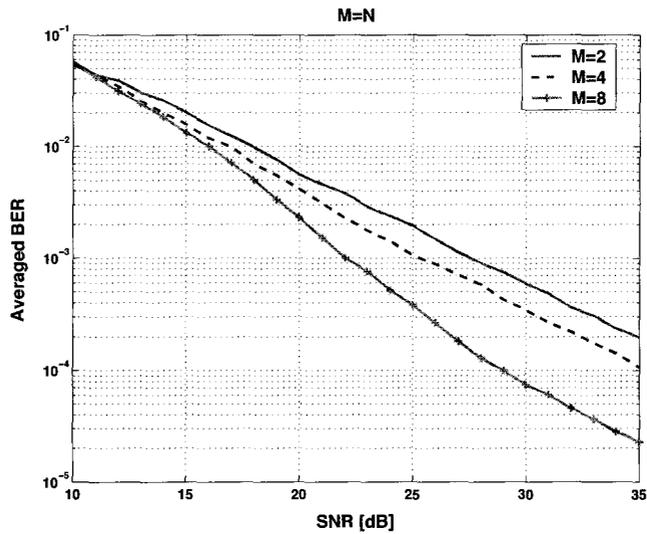


Figure 4.7: Single-block transmission. $M = N = 2, 4, 8$.

(2) In the second experiment, we fix $M = N = T = n$, and we examine the BER performance for $n = 2, 4$ and 8 . The resulting averaged BER curves are plotted in Fig. 4.7. We can observe that at high SNR, all the curves have the same slope, and hence, the same diversity gain. However, the larger the value of M , the smaller the BER that the system achieves, even though the data rate is $2M$ bits pcu. These observations verify the analysis (cf. Theorem 4.2) that showed that diversity gain is only related to $(N - M)$, and that systems with larger M have larger coding gains.

Example 2: In this experiment, we compare the BER performance of the proposed multi-block code design with the numerical evaluation in Section 4.4, and with that of some other available STBCs. We choose the two cases from Section 4.4, namely $M = N = 1$ and $M = N = 2$. In both cases, we consider 4-QAM signalling, resulting in transmission data rates of $R = 2$ and $R = 4$ bits pcu, respectively. The signals are transmitted using the STBC proposed for multi-block transmission in Theorem 4.3. The averaged BER curves for different values of L for these two cases are plotted in Figs. 4.8 and 4.9, respectively. Comparing the analytical results in Figs. 4.1 and 4.2 with the simulation results in Figs. 4.8 and 4.9, we find almost complete agreement.

In order to provide a more comprehensive comparison, we have included in Fig. 4.9 the averaged BER performance of certain single-block MIMO transmission systems with ML detection; namely, systems that transmit i) signals without a STBC (e.g., V-BLAST transmission scheme), and ii) signals coded with the “Golden Code” [73] (which possesses a constant minimum determinant and hence high coding gain). It is well known that the combination of the Golden code transmission and ML detection extracts the full diversity offered by the channel, i.e., $MN = 4$ in this example. However, the ML detector can not achieve full diversity without a STBC in general.

The performance comparisons are shown in Fig. 4.9. We observe that at high SNR

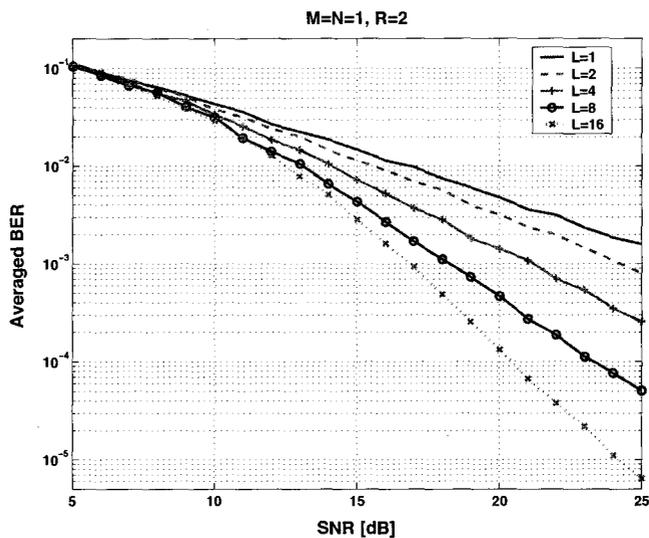


Figure 4.8: Multi-block transmission performance, $M = N = 1$, $R = 2$ bits pcu.

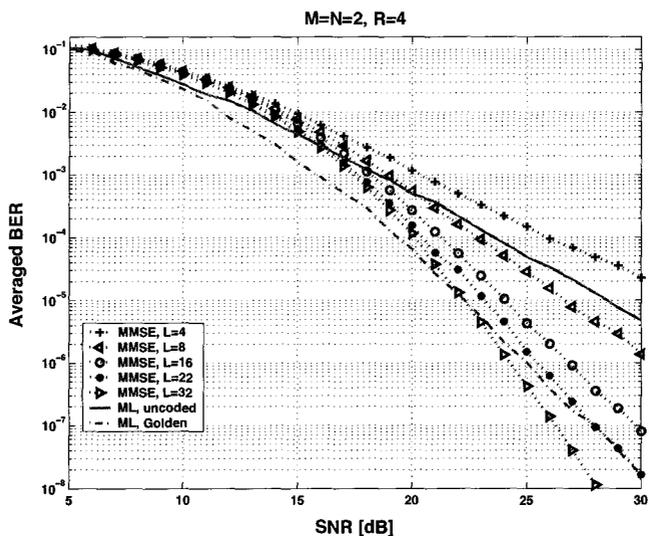


Figure 4.9: Performance of different STBC, $M = N = 2$, $R = 4$ bits pcu.

(above 20 dB), the proposed 8-block codes with linear MMSE receiver perform better than uncoded single-block transmission with ML detection. Now, from Fig. 4.4, we can predict that for $L \geq 22$, the diversity gain for the proposed multi-block scheme with linear reception will be higher than 4, which is the full diversity of the single block system. This is indeed the case in Fig. 4.9. In particular, the performance curve for $L = 22$ shows a slightly higher diversity gain, while a block length of $L = 32$ shows a significantly higher diversity gain, than that of the Golden code with ML detection. It is worth noting that these gains have been achieved with a very low cost as shown in Table 4.1 from Section 4.5.2. On the other hand, when systems are operated at lower SNR regime ($\leq 20\text{dB}$), the combination of ML detection and the Golden code has the best performance.

4.7 Conclusion

In this chapter, we analyze the maximum diversity gain associated with a MIMO system equipped with a linear MMSE receiver for which an STBC of minimum detection error probability has been designed. The calculated diversity gain of $(N - M + 1)$ has been verified by simulations. Indeed, the same diversity gain has also been shown to be achievable in the simpler cases of transmitting uncoded symbols detected by ZF receivers. Here, we showed that employing the optimum unitary trace-orthogonal code together with the MMSE receiver will not increase the diversity gain, rather, the lower BER is achieved by the higher coding gain acquired. To improve the performance of a MIMO system with a linear receiver, we consider multi-block transmission, in which the space-time code spans L realizations of a block-static fading channel. Such a scheme provides additional time diversity, and we designed an optimum STBC that minimizes the probability of error under linear MMSE reception. Our analysis shows

that the diversity order of the proposed system grows with L , while the normalized detection complexity is little more than that for a single-block design with a linear receiver. Thus, if the latency of jointly detecting L blocks of signal can be accommodated, the multi-block code design is an attractive alternative for increasing the diversity gain in a MIMO system.

Chapter 5

Multi-Block Transmission for Linear Receivers

Multi-block transmission can provide more reliable communications compared to single block transmission by taking advantage of time diversity. When an ML detector is employed, the full diversity MNL can be achieved, which increases linearly with L , the number of independent channel realizations a code spans. However, the exhaustive computations in ML detection restrict its practical application. On the other hand, multi-block communication can be employed to improve the system performance for a simple linear receiver.

In Chapter 4, the multi-block transmission scheme was discussed for systems in which the input signals are processed by a *full* rate linear STBC and received by a linear MMSE receiver. The achieved diversity gain d does increase with L and it is the maximum diversity for that transmission scheme, because the proposed code minimizes the detection error probability. However, the function $d(L)$ grows slowly with L and looks like a “log” function. This is due to the use of a linear receiver, which is unable to exploit the total degrees of freedom in a MIMO system. Hence,

from the code design perspective, constraining the code to be a full rate linear STBC may have restrained the system performance for a linear receiver.

With the aim of further improving the performance of the multi-block system as discussed in the previous chapter, in this chapter¹ a scheme with different encoding/decoding functions is proposed. For the same type of signals transmitted through the MIMO channels as was previously considered, the proposed scheme achieves higher diversity gain with the use of a linear receiver. The design is based on the observation that the diversity gain achieved by a linear receiver increases with $(N - M)$, and hence a system that has a taller channel matrix renders high diversity gain. In the proposed transmission scheme, an equivalent channel matrix having dimension $NL \times M$ is constructed by employing multi-block transmission. The achieved diversity gain is $NL - M + 1$, which increases linearly with L .

5.1 Transmission Scheme and System Performance Analysis

Consider the same MIMO channel condition as has been examined in the previous chapter, through which the same input signals are to be transmitted. To further improve the system performance, especially the diversity gain, in this chapter we propose a different multi-block transmission scheme where hybrid precoders and detectors are employed, as shown in Fig. 5.1. As in Chapter 4, in order to utilize time diversity, the signals transmitted through L channel realizations are jointly processed. Since each channel state is assumed to remain unchanged over T channel uses, there are LT channel uses available, and as in Chapter 4, a total of LTM symbols are

¹The work related to this chapter has been presented in *ICASSP 2008* [58] at Las Vegas, April 2008.

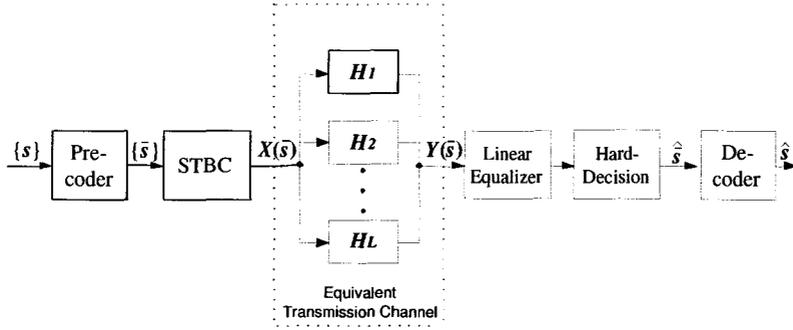


Figure 5.1: Combined precoders/detectors for a multi-block MIMO system.

transmitted by M transmitter antennas in these channel uses. The transmitted symbols $\{s(i), i = 1, 2, \dots, MTL\}$ are independently selected from a constellation \mathcal{S} of cardinality μ . They are first processed by a linear combiner (the precoder in Fig. 5.1) to generate a new set of symbols $\{\bar{s}(k), k = 1, 2, \dots, MT\}$ as

$$\bar{s}(k) = \sum_{\ell=1}^L a_{\ell} s_k(\ell) e^{j\theta_{\ell}}, \quad k = 1, \dots, MT \quad (5.1)$$

where $s_k(\ell)$ are chosen L at a time without repeat, from the input symbols $\{s(i)\}$ to form $\bar{s}(k)$, and the amplitude and phase terms a_{ℓ} and θ_{ℓ} are to be determined. For the averaged signal power to remain unchanged, $\sum_{\ell=1}^L a_{\ell}^2 = L$. Thus, the precoder maps the signal symbol set $\{s(i)\}$ into another symbol set $\{\bar{s}(k)\}$, which is then processed by a linear STBC as $\mathbf{X}(\bar{\mathbf{s}}) = \sum_{k=1}^{MT} \bar{s}(k) \mathbf{C}_k$, where \mathbf{C}_k is an $M \times T$ matrix to be designed. The same coded signals $\mathbf{X}(\bar{\mathbf{s}})$ are then repeatedly transmitted at different blocks of time slots through the channels $\mathbf{H}_{\ell}, \ell = 1, \dots, L$, each having independent channel states. At the receiver, the repeatedly transmitted coded signals are collected and jointly processed by a linear equalizer followed by a hard decision detector to obtain $\{\hat{\bar{s}}(i)\}$, which is then processed by the decoder (that corresponds to the precoder) to obtain the estimate of $\{s(i)\}$.

5.1.1 System Model

We now establish the system model to describe the relation between the input and the output over the L realizations of the MIMO channel. Consider the space-time coded signal $\mathbf{X}(\bar{\mathbf{s}})$ transmitted through the ℓ th state of the channel, \mathbf{H}_ℓ . The corresponding block of received signals, denoted by the $N \times T$ matrix $\mathbf{Y}_\ell(\bar{\mathbf{s}})$, can be written as

$$\mathbf{Y}_\ell(\bar{\mathbf{s}}) = \sqrt{\frac{\rho}{M}} \mathbf{H}_\ell \mathbf{X}(\bar{\mathbf{s}}) + \mathbf{W}_\ell, \quad \ell = 1, \dots, L$$

At the receiver, we wait until the transmission of all the L blocks of signals is complete and we stack them into a tall $NL \times T$ matrix, i.e., $\mathbf{Y}(\bar{\mathbf{s}}) = [\mathbf{Y}_1^T \ \mathbf{Y}_2^T \ \dots \ \mathbf{Y}_L^T]^T$. Correspondingly, we define an $NL \times M$ channel matrix $\mathbf{H} = [\mathbf{H}_1^T \ \mathbf{H}_2^T \ \dots \ \mathbf{H}_L^T]^T$, and an $NL \times T$ noise matrix $\mathbf{W} = [\mathbf{W}_1^T \ \mathbf{W}_2^T \ \dots \ \mathbf{W}_L^T]^T$, and obtain

$$\mathbf{Y}(\bar{\mathbf{s}}) = \sqrt{\frac{\rho}{M}} \mathbf{H} \mathbf{X}(\bar{\mathbf{s}}) + \mathbf{W} \quad (5.2)$$

The stacking of the received signal block matrices \mathbf{Y}_ℓ^T to form $\mathbf{Y}(\bar{\mathbf{s}})$ is equivalent to transmitting $\mathbf{X}(\bar{\mathbf{s}})$ in parallel, as indicated in Fig. 5.1. To facilitate linear equalization, we vectorize the stacked matrix $\mathbf{Y}(\bar{\mathbf{s}})$ in Eq. (5.2) and obtain

$$\mathbf{y} = \text{vec}(\mathbf{Y}) = \sqrt{\frac{\rho}{M}} (\mathbf{I} \otimes \mathbf{H}) \mathbf{F} \bar{\mathbf{s}} + \mathbf{w} \quad (5.3)$$

where $\mathbf{F} = [\text{vec}(\mathbf{C}_1), \text{vec}(\mathbf{C}_2), \dots, \text{vec}(\mathbf{C}_{MT})]$, “ \otimes ” denotes Kronecker product, and \mathbf{w} is the vectorized noise. We can now perform linear equalization followed by a hard decision detector on the received signal vector \mathbf{y} to obtain an estimate of $\bar{\mathbf{s}}$.

5.1.2 Design Criteria and Performance Analysis

To analyze the performance of the scheme, let us first examine the function of the precoder. Suppose the original signals are selected, L at a time without repeat, from a constellation \mathcal{S} of cardinality μ . Then for the signal set $\bar{\mathcal{S}}$ generated by the linear

transformation in Eq. (5.1), there are μ^L elements. It is desired that a_ℓ and θ_ℓ are chosen so that one element \bar{s}_k corresponds uniquely to one group of ordered original signals, $\{s_{k1}, \dots, s_{kL}\}$. In this way, once \bar{s}_k is correctly detected at the receiver, the corresponding L original symbols are also correctly known. Based on that, to examine the system performance, we will consider the error probability in detecting \bar{s}_k .

After linear equalization of the received signal \mathbf{y} , a hard decision is made on each individual equalized symbol to obtain \hat{s}_k . In this procedure, the residual interferences after the linear equalization are treated as noise. For this symbol by symbol detection, the error probability is dominated by the pair-wise error probability (PEP) between the closest pair from the signal set $\{\bar{s}_k\}$, i.e., the worst case PEP. In general, the probability that a transmitted signal symbol \bar{s}_k is detected as \bar{s}_ℓ can be written as

$$P(\bar{s}_k \rightarrow \bar{s}_\ell) = Q\left(\frac{d_{k\ell}}{2}\sqrt{\gamma_k}\right) \quad (5.4)$$

where $Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-x^2/2} dx$, $d_{k\ell}^2 = \|\bar{s}_k - \bar{s}_\ell\|^2$, and γ_k is the SINR at the input of the k th decision device. Eq. (5.4) indicates that the PEP depends on the two parameters $d_{k\ell}$ and γ_k . The distance $d_{k\ell}$ is determined by how the symbols $\{\bar{s}_k\}$ are generated, i.e., it depends only on the precoder. On the other hand, γ_k depends on the choice of the STBC, ρ , the channel states, and the linear equalizer, and is independent of the precoder. Therefore, $d_{k\ell}$ will affect the coding gain, and γ_k will affect both the coding gain and diversity gain of the system. Due to the independence of the precoder and STBC, their design will be considered separately.

We start the design by considering the choice of the STBC. For the signal set $\bar{\mathcal{S}}$, let d_{\min} denotes the minimum distance between any two points. Like $d_{k\ell}$, d_{\min} is determined by the precoder and can be treated as a constant when the design of the STBC is considered. It is obvious that $d_{k\ell}$ is lower bounded by d_{\min} . Now, in improving the system performance, it is desired that the detection error probability averaged over all symbols is kept as low as possible. Hence, we consider minimizing

the PEP averaged over all the signal symbols. That design can be simplified by considering the upper bound generated by the minimum distance, and a design problem is formulated as

$$\min_{\mathbf{F}} : \frac{1}{K} \sum_{k=1}^K \mathbf{E}_{\mathbf{H}} \left\{ Q \left(\frac{d_{\min}}{2} \sqrt{\gamma_k} \right) \right\} \quad (5.5a)$$

$$\text{s.t.} : \text{tr}(\mathbf{F}^H \mathbf{F}) = MT \quad (5.5b)$$

where K is the total number of symbols \bar{s}_k transmitted through the L channel realizations, and the constraint in Eq. (5.5b) is that the code matrix \mathbf{F} maintains the power of the input signals at a constant. Solving Eq. (5.5), and analyzing the minimum error probability, we obtain the following result,

Theorem 5.1 *For the proposed multi-block transmission scheme, the optimal linear STBC is of unitary trace-orthogonal structure, and the achieved diversity gain is $NL - M + 1$.*

Proof. For an MMSE equalizer, the SINR for the k th symbol \bar{s}_k is [61]

$$\gamma_k = \frac{1}{[\mathcal{E}]_{kk}} - 1 \quad (5.6)$$

where the error covariance matrix \mathcal{E} is defined as

$$\mathcal{E} = \left(\mathbf{I} + \frac{\rho}{M} \mathbf{H}^H \mathbf{F}^H \mathbf{F} \mathbf{H} \right)^{-1}$$

Notice here that \mathbf{H} is the equivalent channel with dimension $NL \times M$, each element of which is IID Gaussian distributed. Hence, following similar derivations in [61], the optimization problem in Eq. (5.5) results in the optimal code matrix \mathbf{F} being unitary trace-orthogonal. Applying this optimum code on the system described in Eq. (5.2), the diversity gain is $NL - M + 1$ by Theorem 4.2. \square

Remarks: Theorem 5.1 implies that the system diversity gain does not depend on the choice of the precoder.

We now consider the design of the precoder. From previous discussions, we learn that the error performance is dominated by the worst case PEP, which is further depending on d_{\min} . Therefore, in order to have a small PEP, d_{\min} should be maximized. The design problem can thus be formulated as

$$\max_{\{a_\ell, \theta_\ell, \ell=1, \dots, L\}} : \min\{d_{ij}^2\}, \quad i, j \in \mu^L, \quad i \neq j \quad (5.7a)$$

$$\text{s.t.} : \sum_{\ell=1}^L a_\ell = L \quad (5.7b)$$

The optimal a_ℓ and θ_ℓ depend on the signal constellation and L , and there is no general solution to Eq. (5.7). An example in solving Eq. (5.7) to arrive at an optimum precoder for specific cases is given in Section 5.2.1.

5.2 Design Example and Simulations

In the previous section, a general criterion for the design of the precoder is provided by Eq. (5.7), and the optimal choice of the precoder depends on the structure of the signal constellation and the value of L . Now in this section, we provide an example in the design of the optimal precoder for a 4-PSK constellation and for the conjoining of $L = 2$ blocks of time slots. Since the minimum distance d_{\min} is employed as the objective to be maximized, the optimal precoder designed for a 4-PSK constellation is also optimum for 4-QAM, because the latter is merely a rotated version of the former.

5.2.1 Design Example

For $L = 2$, let $\{s_i\}$ and $\{\bar{s}_i\}$ be the symbol elements in the signal set \mathcal{S} and $\bar{\mathcal{S}}$ respectively. The precoder maps the \mathcal{S} into $\bar{\mathcal{S}}$ by the linear combination: $\bar{s}_k =$

$a_1 s_{k1} e^{j\theta_1} + a_2 s_{k2} e^{j\theta_2}$, where $a_1^2 + a_2^2 = 2$ and s_{k1} and s_{k2} are 4-PSK symbols from \mathcal{S} selected to form \bar{s}_k . Without loss of generality, we can assume $\theta_1 = 0$, $\theta_2 = \theta$ and $a_2 \geq a_1 > 0 \Rightarrow a_2 \geq 1$. Thus, there are three design variables a_1, a_2, θ in the maximization of the minimum distance between any two distinct points in $\bar{\mathcal{S}}$.

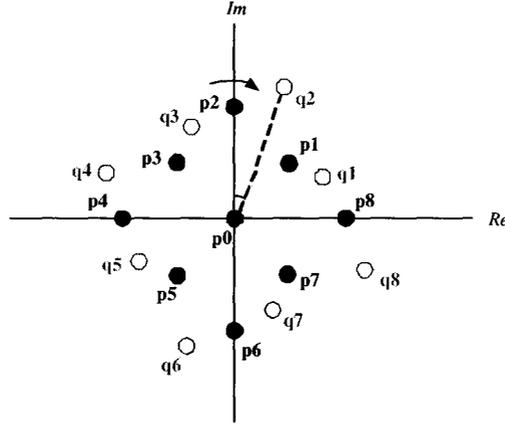
Since the design problem hinges on the distance, we consider the difference between two points $\{\bar{s}_\gamma\}$ and $\{\bar{s}_k\}$ in the set $\bar{\mathcal{S}}$,

$$\bar{s}_\gamma - \bar{s}_k = a_1(s_{\gamma 1} - s_{k1}) + a_2 e^{j\theta}(s_{\gamma 2} - s_{k2}) \quad (5.8)$$

There are two terms in Eq. (5.8), each formed by the difference of two symbols from a 4-PSK constellation. Now, $(s_{\gamma\ell} - s_{k\ell}), \ell = 1, 2$ is the difference between signals chosen from $\{\pm 1, \pm \gamma\}$ forming a set $\mathcal{P} = \{0, \pm 2, (\pm 1 \pm \gamma), \pm 2\gamma\}$ from which the first term in Eq. (5.8) is constructed after scaling by a_1 . This set of symbols is plotted as filled circular dots in Fig. 5.2. Similarly, the second term in Eq. (5.8) is selected from the set \mathcal{Q} (represented by hollow circular dots in Fig. 5.2) which is the set \mathcal{P} scaled by a_2 and rotated by an angle θ . Sets \mathcal{P} and \mathcal{Q} have one common point in the origin. The objective now is to $\max_{a_1, a_2, \theta} : \min\{d_{\gamma k}^2\}$, s.t.: $a_2 > a_1 > 0, a_1^2 + a_2^2 = 2$, where $d_{\gamma k}^2 = \|\bar{s}_\gamma - \bar{s}_k\|^2$ for $\gamma \neq k$. The optimal solution is provided by the following theorem:

Theorem 5.2 *For signals of a 4-PSK or a 4-QAM constellation, when two ($L = 2$) signal blocks are combined for transmission, the optimal parameters for the precoder are $a_1 = \sqrt{1 - \frac{1}{\sqrt{3}}}$, $a_2 = \sqrt{1 + \frac{1}{\sqrt{3}}}$, and $\theta = \frac{\pi}{12}$. The maximized minimum distance is 0.8453.*

Proof. From Fig. 5.2, it is clear that by the symmetry of the two sets \mathcal{P} and \mathcal{Q} , we can limit θ to the range $[0, \pi/4]$. We first obtain the minimum distance and then maximize it. From Eq. (5.8), the distance d_{jk} is the norm of the sum of two elements, one from each of the two sets \mathcal{P} and \mathcal{Q} . Since the elements of both sets are

Figure 5.2: Plotting of $\bar{s}_\gamma - \bar{s}_k$

symmetric about the origin, the minimum distance equals the distance between two neighboring points, one from each of the two sets. From Fig. 5.2, we observe that due to symmetry and $a_1 < a_2$, we only have to consider the distances between two groups of elements: $\{p_1 = \sqrt{2}a_1e^{j\pi/4}, p_2 = j2a_1, p_3 = \sqrt{2}a_1e^{j3\pi/4}\}$; $\{q_0 = 0, q_2 = j2a_2e^{-j\theta}, q_3 = \sqrt{2}a_2e^{j(3\pi/4-\theta)}\}$. There are thus 9 distances that should be considered: $\{d_{10}, d_{12}, d_{13}; d_{20}, d_{22}, d_{23}; d_{30}, d_{32}, d_{33}\}$ where d_{ij} denotes the distance between the two points (p_i, q_j) . From Fig. 5.2, it is obvious that $d_{10} = d_{30} < d_{20}$; $d_{33} < d_{13}$; $d_{22} < d_{32}$. Thus, there are only 5 distances that can be the minimum, i.e., $d_{10}, d_{12}, d_{22}, d_{23}$ and d_{33} . These 5 distances can be easily calculated giving:

$$d_{10}^2 = 2a_1^2 \quad (5.9a)$$

$$d_{12}^2 = 4 + 2a_2^2 - 4a_1a_2(\cos\theta - \sin\theta) \quad (5.9b)$$

$$d_{22}^2 = 8(1 - a_1a_2\cos\theta) \quad (5.9c)$$

$$d_{23}^2 = 4 + 2a_1^2 - 4a_1a_2(\cos\theta + \sin\theta) \quad (5.9d)$$

$$d_{33}^2 = 4(1 - a_1a_2\cos\theta) \quad (5.9e)$$

From Eq. (5.9), clearly $d_{33}^2 \leq d_{22}^2$ and $d_{23}^2 < d_{12}^2$. Therefore, there are three possible minimum distances: d_{10} , d_{23} and d_{33} . Using the constraint $a_1^2 + a_2^2 = 2$, we can compare them in the following,

$$d_{23}^2 - d_{33}^2 = 2a_1(a_1 - 2a_2 \sin \theta) \quad (5.10a)$$

$$d_{10}^2 - d_{33}^2 = 2a_2(2a_1 \cos \theta - a_2) \quad (5.10b)$$

$$d_{23}^2 - d_{10}^2 = 4(1 - \sqrt{2}a_1a_2 \cos(\frac{\pi}{4} - \theta)) \quad (5.10c)$$

Those quantities in Eq. (5.10) can be either positive or negative in the feasible range of the variables, which means that any one d_{10}^2 , d_{23}^2 , or d_{33}^2 can be the minimum.

Closer examination of Eqs. (5.9a), (5.9d) and (5.9e) reveals that increasing one distance will cause the decrease in the others. Therefore, to maximize the minimum, we should make the possible minimums all equal. If such a solution exists, then the quantities in Eq. (5.10) must be equal to zero. Now, we set Eqs. (5.10a) and (5.10b) to be zero (and this implies Eq. (5.10c) equals zero) and obtain $\sin \theta = \frac{a_1}{2a_2}$; $\cos \theta = \frac{a_2}{2a_1}$. Squaring both sides and adding, we have $\frac{a_1^2}{4a_2^2} + \frac{a_2^2}{4a_1^2} = 1$. Combined with the condition that $a_1^2 + a_2^2 = 2$, we arrive at $a_1^2 = 1 - \frac{1}{\sqrt{3}}$ and $a_2^2 = 1 + \frac{1}{\sqrt{3}}$, and the optimal angle is $\theta = \tan^{-1} a_1^2/a_2^2 = \tan^{-1} \frac{\sqrt{3}-1}{\sqrt{3}+1} = \frac{\pi}{12}$. The minimum distance equals $(2 - \frac{2}{\sqrt{3}}) = 0.8453$. \square

The resulting constellation set of $\bar{\mathcal{S}}$ is plotted in Fig. 5.3.

5.2.2 Simulation

In this section, we examine the performance of the multi-block transmission scheme with combined detectors by simulation. For the numerical experiments, we employ the optimal signal design provided in Section 5.2.1 in a MIMO system with $M = N = T = L = 2$. The original signals are randomly selected from a 4-QAM constellation, and the transmission (full) data rate is 4 bits per channel use (pcu). The signals

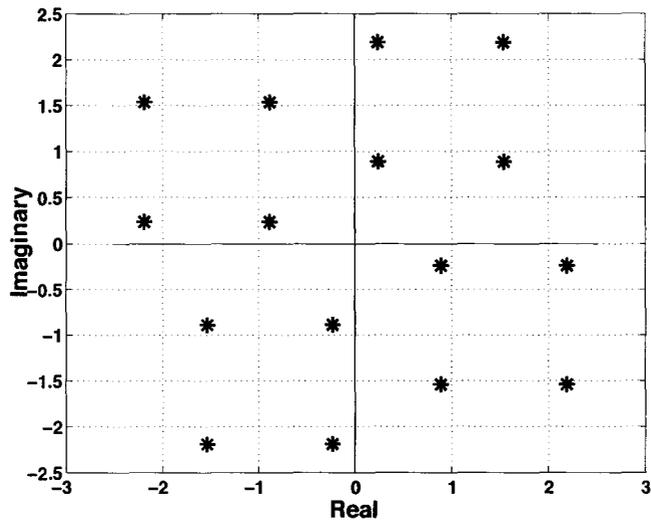


Figure 5.3: The constellation set generated from two 4-PSK.

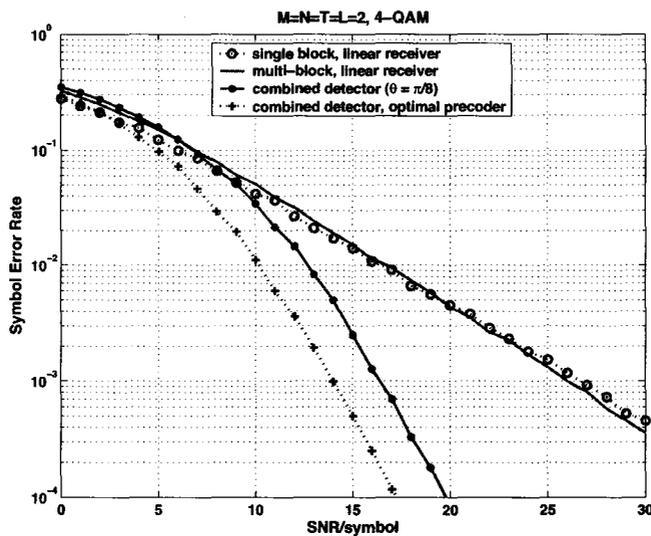


Figure 5.4: Performance comparisons between single block, multi-block, and the proposed schemes.

are first processed by the optimal precoder provided in Section 5.2.1, and then coded by the unitary trace-orthogonal code. The received signals pass through a linear equalizer, a hard decision detector, and a decoder that maps $\hat{\hat{s}}_k$ to obtain the estimate of the original transmitted symbols. The resulting symbol error rate (SER) versus SNR/symbol is plotted as the dotted “+” line in Fig. 5.4. For comparison, we also provide the SER performances three other transmission schemes. The four SER curves in Fig. 5.4 are listed in the following:

1. Single block transmission. The signals are encoded by the unitary trace-orthogonal code [61] and received by a linear receiver. The SER line is blue dotted with circles.
2. Multi-block transmission scheme over $L = 2$ blocks. The code employed here is the optimal multi-block STBC for a linear receiver provided in Chapter 4. The received signals are also processed by a linear receiver. The SER line is green solid in Fig. 5.4.
3. We perform the simulation for the same multi-block transmission with combined detector but using different precoder. Here, in generating $\{\bar{s}\}$, we only shifted the angle, i.e., $\bar{s} = s_i + s_k e^{j\theta}$, with $\theta = \pi/8$. The resulting SER line is red solid with stars.
4. The SER curve for the proposed design provided in Section 5.2.1.

In comparing the four curves in Fig. 5.4, we have the following observations:

- The performance of the proposed scheme having the additional precoder and decoder is significantly better than those two (the first two curves) that do not. The negative slope of the SER curve is much steeper than those for the first two schemes, indicating the higher diversity.

- Now we compare the two SER curves for the same multi-block transmission scheme with different precoders (method 3 and 4). From the analysis in Section 5.1, we know that the design of the precoder will affect the coding gain but not the diversity gain. This is indeed the case in Fig. 5.4, where both curves have the same slope at high SNR, indicating the same diversity gain. The one with the optimal choice of the precoder has superior performance due to the higher coding gain.

5.3 Alternative Choice of the Precoder

In the proposed transmission scheme, the purpose of introducing the precoder and decoder in Fig. 5.1 is to maintain the transmission symbol rate of the multi-block transmission scheme discussed in Chapter 4. In other words, the function of the precoder is to generate a denser constellation $\bar{\mathcal{S}}$ from the original \mathcal{S} . Correspondingly at the receiver end, the decoder maps the detected symbols from $\bar{\mathcal{S}}$ to the original set \mathcal{S} . As a result, the system maintains the same transmission data rate in bits.

Alternatively, the same transmission data rate in bits can be maintained without the precoder/decoder. This can be realized by choosing the original signal symbols directly from known denser constellations. Consider the example examined in Section 5.2, where, we can choose 16-QAM/PSK in order to keep the same rate. These two constellations are plotted in Figs. 5.5 and 5.6 for comparison purposes. From the analysis in Section 5.1, it is clear that the choice of constellations will not affect the diversity, but it may result in different coding gain. Now we compare the three constellations shown in Figs. 5.3, 5.5 and 5.6. To facilitate the comparison, we normalize the three constellations so that the average power for the symbol points in each constellation is unity. Then, the minimum distance $d_{\min}^2 = \min_{i \neq j} : \|s_i - s_j\|^2$

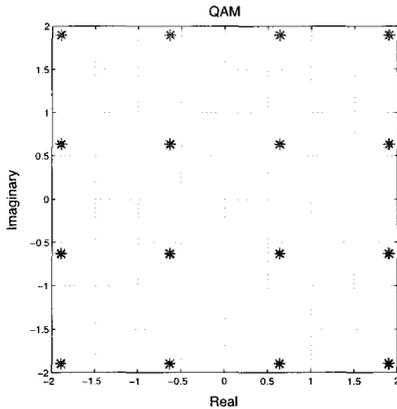


Figure 5.5: 16-QAM

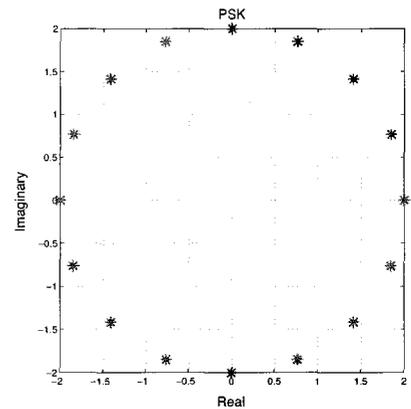


Figure 5.6: 16-PSK

with s_i and s_j being two constellation points, can be found as:

- for the proposed constellation in Fig. 5.3: $d_{\min}^2 = 0.4225$;
- for 16-QAM in Fig. 5.5: $d_{\min}^2 = 0.4000$;
- for 16-PSK in Fig. 5.6: $d_{\min}^2 = 0.1522$.

The proposed constellation has the largest minimum distance and hence it is expected to outperform the other two constellation in detection error performance.

We perform a computer simulation to confirm the observation. Here, we consider the same MIMO channel as described in Section 5.2.2. The signal symbols \bar{s} are directly chosen from a 16 point constellation so that the transmission bit rate is maintained to be 4 bits pcu. The resulting SER curves are plotted in Fig. 5.7, where the signals are picked from the proposed new set in Fig. 5.3, a 16-PSK and a 16-QAM constellation respectively. From this figure, it can be observed that the three curves have the same slope at high SNR, indicating the same diversity gain. Among them, the new proposed constellation has the highest coding gain, as expected.

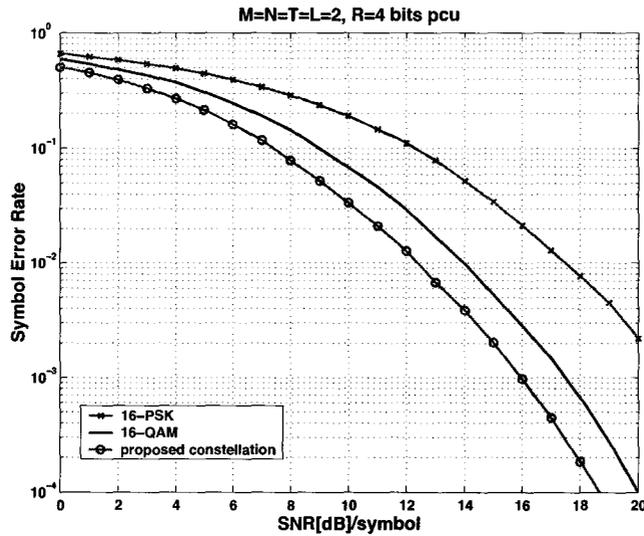


Figure 5.7: SER comparison of the three constellations.

More generally, a good constellation is the one for which the minimum distance is maximized with a constrained power [4]. How to design such a good constellation has been an interesting topic in communications for some time (e.g., [4]), but the rigorous development of contributions to that topic is out of the scope of this thesis.

5.4 Complexity Analysis

In the system model described in Eq. (5.2), the equivalent channel matrix is of dimension $NL \times M$, which is rank M with probability one. Hence, when a linear receiver is employed, the detection complexity of the linear equalizer is of order $O(M^3)$ [61], which is the same as that for a single block system with a linear receiver.

5.5 Conclusion and Discussion

In this chapter, we have proposed a special multi-block MIMO communication scheme for a linear receiver. The system takes advantage of time diversity provided by a multi-block transmission and the simplicity of a linear receiver. The system achieves a diversity gain of $NL - M + 1$, which increases linearly with L .

It is possible that for a large L , the proposed system may suffer from a low coding gain that results in a poor error performance at low SNR. This is due to the fact that a dense constellation is needed to compensate for the transmission data rate. However in the case of a small L , such as the $L = 2$ case as discussed in Section 5.2, a significant improvement in the error performance can be obtained with essentially the same computational cost as that for the single block transmission. This makes the proposed scheme practically attractive. A multi-block transmission with a small value of L is more likely to be employed in practice, since a large L requires longer delay in detection. Such delays are difficult to tolerate in latency sensitive applications, such as interactive multimedia.

The idea of combining a linear receiver and the multi-block transmission may be realized in various ways. The successful application relies on the efficient utilization of the greater degrees of freedom embedded in a multi-block system compared to that of a traditional single block transmission. In spite of the significant improvement obtained here, the transmission scheme presented in this thesis is only one of several possible ways to combine linear receivers with a multi-block transmission strategy.

Chapter 6

Cyclotomic STBC – a linear dispersion FRFD unitary trace-orthogonal code

In previous chapters, the optimality of unitary trace-orthogonal code are exploited in the communication systems that employs a linear receiver. In fact, the application of unitary trace-orthogonality is not constrained to these systems. In this chapter¹, we design full rate full diversity unitary trace-orthogonal codes for a MIMO system having ML detection. Specifically, we consider linear dispersion (LD) unitary trace-orthogonal codes as defined in Eq. (2.40).

Starting from both the information theoretic and detection error viewpoints, we first establish that unitary trace-orthogonality is a desirable property for general LD codes. By imposing the this structure on a LD code and applying cyclotomic number theory, we establish, for an *arbitrary number of transmitter and receiver antennas*, a systematic and simple method to *jointly* design a unitary cyclotomic matrix, the

¹The work related to this chapter has been published in *IEEE Trans. Signal Processing* [113]

Diophantine number and the corresponding constellation for a LD code. As a result, this enables us to construct full diversity *rectangular* cyclotomic LD codes with any symbol transmission rate less than or equal to the number of transmitter antennas. In addition, for the case when the number of transmitter antennas is greater than the number of receiver antennas, by taking advantage of the delay, we also arrive at the design of a special trace-orthonormal full diversity cyclotomic space-time block code. For the number of transmitter antenna being equal to 2^m , the code can be proved to minimize the worst case pair-wise error probability of an ML detector for a μ -ary QAM signal constellation and therefore, achieves optimal coding gain. Computer simulations show that these codes have bit-error performance advantages over currently available codes.

6.1 Introduction

To achieve good performance and good transmission rate in MIMO communications, Hassibi and Hochwald [37] proposed LD codes in which the transmitted codeword is a linear combination of certain weighted matrices. The key to LD code design is that the basis matrices are optimized such that the resulting codes maximize the *ergodic* capacity of the LD coded MIMO system. Unfortunately, for the LD codes proposed in [37], good error probability performance is not strictly guaranteed. To bridge the gap between multiplexing and diversity gains, Heath and Paulraj [39] proposed a LD code design using frame theory which typically performs well in terms of both ergodic capacity as well as error performance. However, their design still cannot guarantee full diversity.

More recent research [65, 22, 64, 85] based on number theory has shown that it is possible to design full rate full diversity (FRFD) linear *space-time block codes* which

are also *information lossless* (i.e., the ergodic capacity of the space-time coded MIMO channel is equal to that of the original MIMO channel). For this type of design, there are two important issues:

1. Design in different layers – The purpose in space-time coding is to transform the original information symbols to be transmitted each time by multiplying the M symbols by a *rotation matrix* to generate a code matrix. This code matrix may be scaled by a *Diophantine number* ψ_n [22] and then further multiplied by a circulant matrix to form various *layers* of transmission. Therefore, the joint design of the rotation matrix, the signal constellation, and the Diophantine number for an arbitrary number of transmitter antennas and receiver antennas becomes an important issue. The rotation matrix in [22, 64] was originated from [32], where the number of transmitter antennas is restricted to $2^m 3^n$ (m and n being integers), and the corresponding constellations are limited only to PAM or QAM. Recently, Wang [100] *et al* have obtained a result for *cyclotomic* linear diagonal space-time block code design in which a systematic method to jointly design a generating matrix and corresponding constellation set was proposed. However, the resulting optimal generating matrix is not unitary, and the number of transmitter antennas is limited only to a specific rational integer.
2. Restriction on the number of transmitters and channel uses – the signal matrix in [22, 64] is square, i.e., the number of channel uses T must be equal to the number of transmitter antennas. This is particularly restrictive in the case when the channel is not known to either the transmitter or the receiver, since, under this condition, a non-coherent receiver which needs a non-square space-time code design (e.g. [42]) has to be employed. Several examples of non-square coding matrices for non-coherent MIMO systems are given in [46], but how

these codes were obtained was not shown.

In this Chapter, both of the above issues will be addressed. We develop a systematic and efficient method to jointly design the rotation matrix, the constellation, and the Diophantine number for an arbitrary number of transmitter antennas and receiver antennas. We also generalize the signal matrix to a rectangular matrix so that the severe restriction on the number of channel uses is relaxed. Since there is a large number of parameters to be considered for the design (particularly if the numbers of transmitter antennas and channel uses are large, and/or the constellation is sizable), it will be very much more efficient to confine our consideration to only the possible necessary conditions satisfied by a good space-time code so that the parameter space can be dramatically reduced. To this end, we first establish the architecture of a good LD code from the viewpoints of information theory and detection theory, and then focus our design around such structures with the aim of developing a systematic and efficient way of finding a good space-time code.

6.2 Channel model with linear dispersion codes

Consider a LD code as introduced in Chapter 2

$$\mathbf{X}(\mathbf{s}) = \sum_{k=1}^K \mathbf{A}_k s_k + \sum_{k=1}^K \mathbf{B}_k s_k^* \quad (6.1)$$

where \mathbf{s} denotes a $K \times 1$ transmission symbol vector $\mathbf{s} = [s_1, s_2, \dots, s_K]^T$ containing the K information symbols selected from an alphabet \mathcal{S} to be transmitted within the T time slots, and \mathbf{A}_k and \mathbf{B}_k each denotes an $M \times T$ matrix.

At the receiver, the received signal can be written as

$$\mathbf{Y} = \sqrt{\frac{\rho}{M}} \mathbf{H} \mathbf{X}(\mathbf{s}) + \mathbf{W} \quad (6.2)$$

\mathbf{Y} and \mathbf{W} denote the $N \times T$ received signal and noise matrices, respectively. An equivalent model can be established by vectoring both sides of Eq. (6.2) resulting

$$\text{vec}(\mathbf{Y}) = \sqrt{\frac{\rho}{M}} \mathcal{H}(\mathbf{A}\mathbf{s} + \mathbf{B}\mathbf{s}^*) + \text{vec}(\mathbf{W}) \quad (6.3)$$

where $\mathcal{H} = \mathbf{I}_T \otimes \mathbf{H}$, and

$$\mathbf{A} = [\text{vec}(\mathbf{A}_1), \text{vec}(\mathbf{A}_2), \dots, \text{vec}(\mathbf{A}_K)]$$

$$\mathbf{B} = [\text{vec}(\mathbf{B}_1), \text{vec}(\mathbf{B}_2), \dots, \text{vec}(\mathbf{B}_K)]$$

Combining Eq. (6.3) and its conjugate, and applying a transformation matrix \mathbf{T} to the combined equation results in a model which involves only real multi-inputs and real multi-outputs such that

$$\begin{pmatrix} \text{vec}(\mathbf{Y}_{\text{re}}) \\ \text{vec}(\mathbf{Y}_{\text{im}}) \end{pmatrix} = \sqrt{\frac{\rho}{M}} \mathbf{T}^H \tilde{\mathcal{H}} \mathcal{F} \mathbf{T} \begin{pmatrix} \mathbf{s}_{\text{re}} \\ \mathbf{s}_{\text{im}} \end{pmatrix} + \begin{pmatrix} \text{vec}(\mathbf{W}_{\text{re}}) \\ \text{vec}(\mathbf{W}_{\text{im}}) \end{pmatrix} \quad (6.4)$$

where the *unitary* transformation matrix is defined as $\mathbf{T} = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{I} & j\mathbf{I} \\ \mathbf{I} & -j\mathbf{I} \end{pmatrix}$, the

precoder matrix is given by $\mathcal{F} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix}$, the matrix $\tilde{\mathcal{H}}$ is defined as $\tilde{\mathcal{H}} = \text{diag}(\mathcal{H}, \mathcal{H}^*)$, and $[\cdot]_{\text{re}}$ and $[\cdot]_{\text{im}}$ denote the real and imaginary parts of a quantity, respectively. Eq. (6.4) is useful in the analysis of the detection performance of an LD code at the receiver as well as the ergodic capacity of the LD coded channel.

6.3 Good structures for LD codes: Unitary trace-orthogonality

In this section, we examine the structure of an LD code used in MIMO communications. From an information-theoretic viewpoint, we found that a good LD code

should have an *inter*-unitary structure, while from a detection error viewpoint, we establish that a good LD code should have an *intra*-unitary structure. We called a code having the combined inter-unitary and intra-unitary structure a *unitary trace orthogonal* LD code [109, 110].

6.3.1 An Information-theoretic Viewpoint

In the design of space-time LD codes, it is important that the ergodic capacity of the equivalent channel of Eq. (6.4) employing the LD code be made as close as possible to that of the original uncoded MIMO channel. Now, the ergodic capacity of the channel without space-time coding [92] is given by $C = \text{E} [\log \det(\mathbf{I}_N + \frac{\rho}{M} \mathbf{H}\mathbf{H}^H)]$, whereas, for the channel of Eq. (6.4), the corresponding ergodic channel capacity has been evaluated [37, 39] and is given by $C_{\text{LDC}} = \max_{\mathcal{F}} \frac{1}{2T} \text{E} [\log \det(\mathbf{I}_{2K} + \frac{\rho}{M} \tilde{\mathcal{H}}\mathcal{F}\mathcal{F}^H\tilde{\mathcal{H}}^H)]$. A code is said to be *capacity optimum* if it renders the same ergodic capacity of the coded channel as that of the original uncoded channel. Now, in obtaining such an optimal \mathcal{F} , we have the following assertion:

Theorem 6.1 *Let $K = MT$ (i.e., consider full symbol rate transmission). Then, subject to the power constraint: $\sum_{k=1}^K \text{tr}(\mathbf{A}_k\mathbf{A}_k^H + \mathbf{B}_k\mathbf{B}_k^H) = MT$, \mathcal{F} is capacity-optimal if and only if \mathcal{F} is unitary, or equivalently, the following conditions hold*

$$\text{tr}(\mathbf{A}_k\mathbf{A}_{k'}^H + \mathbf{B}_{k'}\mathbf{B}_k^H) = \delta(k - k') \quad (6.5a)$$

$$\text{tr}(\mathbf{B}_k\mathbf{A}_{k'}^H + \mathbf{B}_{k'}\mathbf{A}_k^H) = 0 \quad (6.5b)$$

for $k, k' = 1, 2, \dots, K$, where $\delta(k - k')$ is the Kronecker delta. ■

It must be pointed out here that several authors [81, 39, 37, 55] have noted the importance of the above conditions in Eq. (6.5). When a linear STBC was considered, that \mathcal{F} being unitary is both *necessary and sufficient* for optimal capacity in

Theorem 6.1 was shown in [55].

6.3.2 A Detection Error Viewpoint

We now establish a structure of the code necessary to minimize the detection error. For the development of the analysis, we assume that ML detection is employed at the receiver.

Given a channel realization \mathbf{H} , the probability of transmitting \mathbf{s} and deciding in favor of $\mathbf{s}' \neq \mathbf{s}$ by the ML detector is given by

$$P(\mathbf{s} \rightarrow \mathbf{s}' | \mathbf{H}) = Q\left(\frac{\nu}{2}\right), \quad \mathbf{s} \neq \mathbf{s}' \quad (6.6)$$

where $Q(z) = (1/\sqrt{2\pi}) \int_z^\infty e^{-\zeta^2/2} d\zeta$ and ν is the Euclidean distance between the received code words $\mathcal{H}\mathcal{F}\mathbf{s}$ and $\mathcal{H}\mathcal{F}\mathbf{s}'$ which is a function of $\mathbf{e} = \mathbf{s} - \mathbf{s}'$, i.e.,

$$\begin{aligned} \nu^2 &= \frac{\rho}{M} \text{tr} [\mathbf{X}^H(\mathbf{e}) \mathbf{H}^H \mathbf{H} \mathbf{X}(\mathbf{e})] \\ &= \frac{\rho}{M} [\text{vec}(\mathbf{H})]^H [\mathbf{X}^*(\mathbf{e}) \mathbf{X}^T(\mathbf{e}) \otimes \mathbf{I}_N] \text{vec}(\mathbf{H}) \end{aligned} \quad (6.7)$$

with the coding matrix $\mathbf{X}(\cdot)$ defined in Eq. (6.1).

To evaluate the probability in Eq. (6.6), we invoke an alternative expression for the Q -function [12], [87], $Q(z) = \frac{1}{\pi} \int_0^{\pi/2} \exp\left(-\frac{z^2}{2\sin^2\theta}\right) d\theta$, ($z \geq 0$).

Applying this in Eq. (6.6) and taking the average over the Gaussian-distributed random vector $\text{vec}(\mathbf{H})$, we obtain the average pair-wise error probability at the ML detector as

$$P(\mathbf{s} \rightarrow \mathbf{s}') = \frac{1}{\pi} \int_0^{\pi/2} \frac{d\theta}{\det\left(\mathbf{I}_M + \frac{\rho}{8M\sin^2\theta} \mathbf{X}(\mathbf{e}) \mathbf{X}^H(\mathbf{e})\right)^N} \quad (6.8)$$

Eq. (6.8) yields an expression of the exact pairwise error probability, which can be employed as a design criterion to obtain good space-time codes [91]. In the following, we explore the properties of a good code by establishing the lower bound of the worst

case pair-wise error probability and then by examining the necessary code structure for the lower bound to be reached.

Let us first define the *minimum distance of the constellation* \mathcal{S} as $\nu_{\min}(\mathcal{S}) = \min_{s, s' \in \mathcal{S}, s \neq s'} |s - s'|$ where s be an element of \mathbf{s} . The following theorem [110] provides us with an expression on the universal lower bound of the worst case pair-wise error probability for any linear dispersion code:

Theorem 6.2 *Let \mathcal{S}_{re} and \mathcal{S}_{im} denote two constellations consisting of, respectively, the real and the imaginary parts of the elements in constellation \mathcal{S} . If \mathcal{S} satisfies the following geometrical property,*

$$\nu_{\min}(\mathcal{S}) = \nu_{\min}(\mathcal{S}_{\text{re}}) = \nu_{\min}(\mathcal{S}_{\text{im}}) \quad (6.9)$$

then, for any LD code \mathcal{F} with a power budget $\text{tr}(\mathcal{F}\mathcal{F}^H) \leq MT$, the lower bound of the worst case pairwise error probability of the ML detector is given by

$$\max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K, \mathbf{s} \neq \mathbf{s}'} P_{\mathcal{F}}(\mathbf{s} \rightarrow \mathbf{s}') \geq J \left(\frac{\rho T \nu_{\min}^2(\mathcal{S})}{8MK} \right) \quad (6.10)$$

where $J(a)$ is given by $J(a) = \frac{1}{\pi} \int_0^{\pi/2} \left(1 + \frac{a}{\sin^2 \theta}\right)^{-MN} d\theta$ for $a > 0$. Furthermore, a necessary condition for the lower bound to be achieved is that the pair of matrices, \mathbf{A}_k and \mathbf{B}_k , associated with each individual symbol must satisfy the following conditions

$$\mathbf{A}_k \mathbf{A}_k^H + \mathbf{B}_k \mathbf{B}_k^H = \frac{T}{K} \mathbf{I}_M, \quad \mathbf{A}_k \mathbf{B}_k^H + \mathbf{B}_k \mathbf{A}_k^H = \mathbf{0} \quad (6.11)$$

■

Proof. Define $M \times M$ coding matrices as $\mathbf{F}_k = \mathbf{A}_k + \mathbf{B}_k$ and $\mathbf{F}_{K+k} = j(\mathbf{A}_k - \mathbf{B}_k)$ for $k = 1, \dots, K$. Then, $[\mathbf{A}_k^T \ \mathbf{B}_k^T]^T = \frac{1}{\sqrt{2}} \mathbf{T}^* [\mathbf{F}_k^T \ \mathbf{F}_{K+k}^T]^T$ where \mathbf{T} is the transformation matrix mentioned in the Section 6.2. With this relationship, it is easy to verify that the conditions in Eq. (6.11) are equivalent to:

$$\mathbf{F}_k \mathbf{F}_k^H = \frac{T}{K} \mathbf{I}_M, \quad \mathbf{F}_{K+k} \mathbf{F}_{K+k}^H = \frac{T}{K} \mathbf{I}_M \quad (6.12)$$

Consider the case when the error is $|e_{k'}| = \nu_{\min}(\mathcal{S})$ and $e_k = 0$, $k = 1, 2, \dots, K$, $k \neq k'$, the denominator in the integral of Eq. (6.8) can be written as

$$\det \left(\mathbf{I} + \frac{\rho \mathbf{X}_{\mathcal{F}}(\mathbf{e}) \mathbf{X}_{\mathcal{F}}^H(\mathbf{e})}{8M \sin^2 \theta} \right) = \det \left(\mathbf{I} + \frac{\rho \nu_{\min}^2(\mathcal{S}) \mathbf{F}_{k'} \mathbf{F}_{k'}^H}{8M \sin^2 \theta} \right) \quad (6.13)$$

Using Hardamard's inequality [43] and then employing the relationship between the arithmetic mean and the geometrical mean, we have

$$\begin{aligned} \det \left(\mathbf{I} + \frac{\rho \nu_{\min}^2(\mathcal{S}) \mathbf{F}_{k'} \mathbf{F}_{k'}^H}{8M \sin^2 \theta} \right) &\leq \prod_{\ell=1}^M \left(1 + \frac{\rho \nu_{\min}^2(\mathcal{S}) [\mathbf{F}_{k'} \mathbf{F}_{k'}^H]_{\ell\ell}}{8M \sin^2 \theta} \right) \\ &\leq \left(1 + \frac{\rho \nu_{\min}^2(\mathcal{S}) \text{tr}(\mathbf{F}_{k'} \mathbf{F}_{k'}^H)}{8M^2 \sin^2 \theta} \right)^M \end{aligned} \quad (6.14)$$

where, respectively, equalities in the first and second parts of Eq. (6.14) hold if and only if:

- (i) $\mathbf{F}_{k'} \mathbf{F}_{k'}^H$ is a diagonal matrix, and
- (ii) $[\mathbf{F}_{k'} \mathbf{F}_{k'}^H]_{\ell\ell} = [\mathbf{F}_{k'} \mathbf{F}_{k'}^H]_{mm}$, for $\ell, m = 1, 2, \dots, M$.

The upper bound on the right side of Eq. (6.14) is still a function of the design parameters $\mathbf{F}_{k'}$. Now, if in particular, this coding matrix is the one which has been allocated the minimum power among all the coding matrices $\{\mathbf{F}_k\}$, i.e., let \mathbf{F}_{m0} be the coding matrix such that $\text{tr}(\mathbf{F}_{m0} \mathbf{F}_{m0}^H) = \min [\text{tr}(\mathbf{F}_k \mathbf{F}_k^H)]$, $1 \leq k \leq 2K$, we then have

$$\text{tr}(\mathbf{F}_{m0} \mathbf{F}_{m0}^H) \leq \frac{1}{2K} \sum_{k=1}^{2K} \text{tr}(\mathbf{F}_k \mathbf{F}_k^H) = \frac{MT}{K} \quad (6.15)$$

where the last part of Eq. (6.15) is due to the power constraint. Equality in (6.15) holds if and only if the following condition is satisfied

$$\text{tr}(\mathbf{F}_k \mathbf{F}_k^H) = \frac{MT}{K} \quad \text{for } k = 1, 2, \dots, 2K \quad (6.16)$$

Combining Eq. (6.15) with the upper bound on the right side of (6.14) yields

$$\left(1 + \frac{\rho \nu_{\min}^2(\mathcal{S}) \text{tr}(\mathbf{F}_{m0} \mathbf{F}_{m0}^H)}{8M^2 \sin^2 \theta} \right)^M \leq \left(1 + \frac{\rho T \nu_{\min}^2(\mathcal{S})}{8MK \sin^2 \theta} \right)^M \quad (6.17)$$

for which equality holds if and only if the condition in Eq. (6.16) holds.

The upper bound in Eq. (6.17) is now independent of the design parameter and is therefore the maximum value attainable by $\det(\mathbf{I} + \frac{\rho \nu_{\min}^2(\mathcal{S}) \mathbf{F}_{m0} \mathbf{F}_{m0}^H}{8M^2 \sin^2 \theta})$ for the coding matrix allocated the lowest power. On the other hand, we note that Conditions (i) and (ii) above as well as Eq. (6.16) are simultaneously satisfied if

$$\mathbf{F}_{k'} \mathbf{F}_{k'}^H = \frac{T}{K} \mathbf{I}_M \quad \text{for } k' = 1, 2, \dots, 2K \quad (6.18)$$

which, due to the fact that Conditions (i) and (ii) are met, will render the upper bound on the right side of Eq. (6.14) attainable for all the coding matrices $\mathbf{F}_k, k = 1, \dots, 2K$. As stated before (see Eq. (6.12)), these conditions of Eq. (6.18) are equivalent to the conditions that $\mathbf{A}_m \mathbf{A}_m^H + \mathbf{B}_m \mathbf{B}_m^H = \frac{T}{K} \mathbf{I}_M$ and $\mathbf{A}_m \mathbf{B}_m^H + \mathbf{B}_m \mathbf{A}_m^H = \mathbf{0}$. Hence, by substituting this upper bound value in Eq. (6.17) into Eq. (6.8), the probability of the particular case of error when $|e_m| = \nu_{\min}(\mathcal{S})$ and $e_k = 0, k = 1, 2, \dots, K, k \neq m$ is lower bounded by $J\left(\frac{\rho T \nu_{\min}^2(\mathcal{S})}{8MK}\right)$. Therefore, we can conclude that the *worst case* pairwise error probability is bounded by²

$$\max_{s, s' \in \mathcal{S}^K, s \neq s'} P_{\mathcal{F}}(s \rightarrow s') \geq J\left(\frac{\rho T \nu_{\min}^2(\mathcal{S})}{8MK}\right)$$

□

Theorem 6.2 establishes from a detection viewpoint that *each individual code matrix of a good LD code should have an intra-unitary structure*. The following comments on the result should be noted:

1. For the lower bound of Eq. (6.10) to be valid, the constellation \mathcal{S} has to satisfy Eq. (6.9). If the real and imaginary parts have different powers, then, when two

²We note that a similar lower bound for a *linear space-time block code* based on the Chernoff bound was given in [15].

real symbol errors occur, it will not be possible for us to claim that the worst case pair-wise error probability is lower-bounded by the error caused by the two corresponding unitary matrices.

2. The square q -ary QAM satisfies the condition in Eq. (6.9), whereas non-square μ -ary QAM and q -ary PSK constellation $\{\exp(\frac{2k\pi j}{q})\}_{k=0}^{q-1}$ do not. Although the constellation for LD codes is limited to satisfy the condition in Eq. (6.9), from the proof of Theorem 6.2, we see that in the case of linear STBC, Theorem 6.2 is true for any constellation.
3. It can be verified that under the condition of an allowable symbol rate, orthogonal STBC can achieve the lower bound in Eq. (6.10). In Section 6.6, we will design a family of high symbol rate linear STBC achieving this lower bound.

6.3.3 Unitary Trace-orthogonal LD codes

Theorem 6.1 and Theorem 6.2 above together suggest that from the information-theoretic and the detection error viewpoints, we should strive for both inter-unitary and intra-unitary structures in the design of an LD code. We call the such structures of an LD code *unitary trace orthogonality* [54, 109, 110] which can be formally defined as:

Definition 6.1 *Let $T \geq M$. A sequence of $M \times T$ matrices \mathbf{A}_k and \mathbf{B}_k , $k = 1, 2, \dots, K$ and $K \leq MT$, is said to constitute a trace-orthonormal LD code if the*

following conditions are satisfied,

$$\mathbf{A}_k \mathbf{A}_k^H + \mathbf{B}_k \mathbf{B}_k^H = \frac{T}{K} \mathbf{I}_M \quad (6.19a)$$

$$\mathbf{A}_k \mathbf{B}_k^H + \mathbf{B}_k \mathbf{A}_k^H = \mathbf{0} \quad (6.19b)$$

$$\text{tr}(\mathbf{A}_k \mathbf{A}_{k'}^H + \mathbf{B}_{k'} \mathbf{B}_k^H) = \frac{MT}{K} \delta(k - k') \quad (6.19c)$$

$$\text{tr}(\mathbf{B}_k \mathbf{A}_{k'}^H + \mathbf{B}_{k'} \mathbf{A}_k^H) = 0 \quad (6.19d)$$

for $k, k' = 1, 2, \dots, K$. In particular, when $K = MT$, it is said to constitute a trace-orthonormal LD code of full transmission symbol rate. ■

Remarks:

1. The conditions for designing *complex orthogonal space-time block codes* are [1, 94, 30, 90, 29]:

$$\begin{aligned} \mathbf{A}_k \mathbf{A}_{k'}^H + \mathbf{B}_{k'} \mathbf{B}_k^H &= \frac{T}{K} \delta(k - k') \mathbf{I}_M \\ \mathbf{B}_k \mathbf{A}_{k'}^H + \mathbf{B}_{k'} \mathbf{A}_k^H &= \mathbf{0} \end{aligned}$$

These conditions do not imply, but are implied by, Eqs. (6.19a), (6.19c) and (6.19d). Thus, trace-orthonormal LD codes can be viewed as a generalization of complex orthogonal STBC.

2. The space-time codes in [22, 64] are trace-orthonormal linear STBC.
3. Eqs. (6.19c) and (6.19d) are equivalent to \mathcal{F} being column-wise orthonormal and thus, by Theorem 6.1, a trace-orthonormal full dimension LD code is information lossless [39, 37, 17].
4. Eqs. (6.19a) and (6.19b) ensure that the LD code satisfies the necessary conditions of being intra-unitary as required in Theorem 6.2.

5. The conditions of Eqs. (6.19a) – (6.19d) ensure that the signal transmission power for each symbol is evenly allocated, and simultaneously ensure that the output interference and noise power of each subchannel after the linear equalizer (zero-forcing or minimum mean square error) are evenly distributed if $N \geq M$ and $K = MT$ [55, 54, 61].

6.4 Construction of trace orthonormal LD codes

Since unitary trace orthogonality is desirable for an LD code, in transforming the information symbols, we aim to arrive at a code with such a structure. In this section, we develop a scheme for the construction of trace-orthonormal LD codes. First, we introduce the following definition:

Definition 6.2 Let $T = LM$, where L is a positive integer. A $2T \times 2T$ unitary matrix \mathbf{V} satisfying

$$\mathbf{V} = \begin{pmatrix} \mathbf{\Theta} & \mathbf{\Phi} \\ \mathbf{\Phi}^* & \mathbf{\Theta}^* \end{pmatrix} \quad (6.20)$$

is said to be of V -structure if:

- a) The entries of the $T \times T$ matrices $\mathbf{\Theta}$ and $\mathbf{\Phi}$ satisfy

$$\sum_{n=0}^{L-1} |\theta_{(Mn+m),t}|^2 + \sum_{n=0}^{L-1} |\phi_{(Mn+m),t}|^2 = \frac{1}{M} \quad (6.21)$$

for $t = 1, \dots, T$; $m = 1, \dots, M$.

- b) The cross terms between $\mathbf{\Theta}$ and $\mathbf{\Phi}$ satisfy

$$\sum_{n=0}^{L-1} (\theta_{(Mn+m),t} \phi_{(Mn+m),t}^* + \theta_{(Mn+m),t}^* \phi_{(Mn+m),t}) = 0 \quad (6.22)$$

for $t = 1, \dots, T$; $m = 1, \dots, M$, where $\theta_{(Mn+m),t}$ and $\phi_{(Mn+m),t}$ are the $[(Mn+m), t]$ th elements of Θ and Φ respectively. ■

The following is an example for a V-structured matrix:

Example 6.1 For $M = 2, L = 1$ and $T = 2$, a 4×4 V-structured matrix can be constructed as

$$\mathbf{V} = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos \theta & \cos \theta & -j \sin \theta & j \sin \theta \\ \sin \theta & \sin \theta & j \cos \theta & -j \cos \theta \\ j \sin \theta & -j \sin \theta & \cos \theta & \cos \theta \\ -j \cos \theta & j \cos \theta & \sin \theta & \sin \theta \end{pmatrix}$$

where $\Theta = \begin{pmatrix} \cos \theta & \cos \theta \\ \sin \theta & \sin \theta \end{pmatrix}$ and $\Phi = \begin{pmatrix} -j \sin \theta & j \sin \theta \\ j \cos \theta & -j \cos \theta \end{pmatrix}$. It can be verified that both Eqs. (6.21) and (6.22) are satisfied. ■

That V-structured matrices are closed under the Kronecker product operation and that a new V-structured matrix with a larger dimension can be generated from a given V-structured matrix with a smaller dimension are stated in the following proposition:

Proposition 6.1 Let $T_1 = L_1 M$ and $T_2 = L_2 T_1$. If \mathbf{V}_1 is a $2T_1 \times 2T_1$ V-structured matrix formed with Θ_1 and Φ_1 according to Eq. (6.20) and satisfies Eqs. (6.21) and (6.22), then, for any $L_2 \times L_2$ unitary matrix \mathbf{U} , the following matrix \mathbf{V}_2 is a $2T_2 \times 2T_2$ V-structured matrix,

$$\mathbf{V}_2 = \begin{pmatrix} \mathbf{U} \otimes \Theta_1 & \mathbf{U} \otimes \Phi_1 \\ \mathbf{U}^* \otimes \Phi_1^* & \mathbf{U}^* \otimes \Theta_1^* \end{pmatrix} \quad (6.23)$$

■

We now present two methods for the systematic construction of such a V-structured matrix, which can be easily verified by Definition 6.2 of a V-structured matrix.

Construction 6.1 Let \mathbf{U} denote an $M \times M$ unitary matrix with all its elements having equal magnitude. Let $\mathbf{U}(m : n, :)$ denote a matrix formed by taking row m to row n of \mathbf{U} , while retaining all its columns. We now form the $T \times T$ matrices Θ and Φ such that

$$\Theta = \begin{pmatrix} \mathbf{U}(1 : M_1, :) \\ \mathbf{0}_{M_2 \times M} \end{pmatrix}, \quad \Phi = \begin{pmatrix} \mathbf{0}_{M_1 \times M} \\ \mathbf{U}^*(M_1 + 1 : M, :) \end{pmatrix} \quad (6.24)$$

where the nonnegative integers M_1 and M_2 satisfy $M_1 + M_2 = M$. Then, the matrix \mathbf{V} defined by Eq. (6.20) with such Θ and Φ is of V-structure. ■

Construction 6.2 Let Θ and Φ denote a pair of $M \times M$ real matrices satisfying the conditions i) $\theta_{k,m}^2 + \phi_{k,m}^2 = 1/M$, for $k, m = 1, 2, \dots, M$, ii) $\Theta\Theta^T + \Phi\Phi^T = \mathbf{I}_M$, and iii) $\Theta\Phi^T + \Phi^T\Theta = \mathbf{0}$. Then, the matrix \mathbf{V} given by

$$\mathbf{V} = \begin{pmatrix} \Theta & j\Phi \\ -j\Phi & \Theta \end{pmatrix} \quad (6.25)$$

is a $2M \times 2M$ V-structured matrix. ■

As mentioned in Sections 6.1 and 6.2, different “layers” can be generated by multiplying a code matrix by a circulant matrix \mathbf{P} so that the appropriate symbols of the coded signal vectors can be selected to be transmitted at a given time instant. An important property of a circulant matrix is stated in the following lemma which will be useful in the ensuing sections:

Lemma 6.1 For any $P \times P$ circulant generator matrix \mathbf{P} , the diagonal entries of its m th power \mathbf{P}^m are all zero, i.e.,

$$[\mathbf{P}^m]_{pp} = \delta(m)$$

for $m = 0, 1, \dots, P-1$ and $p = 1, 2, \dots, P$. ■

PROOF: First we notice that \mathbf{P} can be decomposed as $\mathbf{P} = \mathbf{D}_P^H \mathbf{\Lambda} \mathbf{D}_P$, where $\mathbf{\Lambda} = \text{diag}(1, e^{-\frac{j2\pi}{P}}, \dots, e^{-\frac{j2(P-1)\pi}{P}})$ and \mathbf{D}_P denotes the $P \times P$ discrete Fourier transform matrix. Then, the p th diagonal element of its m th power is given by

$$[\mathbf{P}^m]_{pp} = [\mathbf{D}_P^H \mathbf{\Lambda}^m \mathbf{D}_P]_{pp} = \frac{1}{P} \sum_{\ell=0}^{P-1} e^{-\frac{j2\pi\ell m}{P}} = \delta(m)$$

□

With a V-structured matrix constructed as above, together with different circulant matrices to create the different layers of transmission, we are now in a position to present the systematic generation of unitary trace-orthogonal LD codes. This method applies when the number of time slots T is a multiple of the number of transmitter antennas M and the total number of information symbols K is a multiple of the number of transmission time slots T .

Theorem 6.3 *Let $T = LM$, and $K = RT$ with $R \leq M$. Suppose we have R V-structured matrices \mathbf{V}_r , $r = 1, \dots, R$, each of dimension $2T \times 2T$ and each of form given by Eq. (6.20).*

By taking the t -th column vectors of the component matrices $\mathbf{\Theta}_r$ and $\mathbf{\Phi}_r$ in \mathbf{V}_r , we now form the matrices $\mathbf{\Pi}_t$ and $\mathbf{\Delta}_t$, each being an alignment of L diagonal sub-matrices, such that

$$\mathbf{\Pi}_t = \begin{pmatrix} \theta_{t,1} & & & \theta_{t,(M+1)} & & & & & \theta_{t,(L-1)M+1} \\ & \ddots & & & \ddots & & \dots & & \ddots \\ & & \theta_{t,M} & & & \theta_{t,2M} & & & \theta_{t,LM} \end{pmatrix}$$

$$\mathbf{\Delta}_t = \begin{pmatrix} \phi_{t,1} & & & \phi_{t,(M+1)} & & & & & \phi_{t,(L-1)M+1} \\ & \ddots & & & \ddots & & \dots & & \ddots \\ & & \phi_{t,M} & & & \phi_{t,2M} & & & \phi_{t,LM} \end{pmatrix}$$

for $r = 1, \dots, R$ and $t = 1, \dots, T$. Let two sequences of matrices $\mathbf{A}_{r,t}$ and $\mathbf{B}_{r,t}$ be given by

$$\mathbf{A}_{r,t} = \sqrt{\frac{MT}{K}} \mathbf{P}_M^{r-1} \mathbf{\Pi}_t, \quad \mathbf{B}_{r,t} = \sqrt{\frac{MT}{K}} \mathbf{P}_M^{r-1} \mathbf{\Delta}_t \quad (6.26)$$

then, the matrix family $\{\mathbf{A}_{r,t}, \mathbf{B}_{r,t}\}$ constitutes a unitary trace-orthogonal LD code with a symbol rate R per channel use. ■

The proof of this theorem is shown in Appendix C. It should be noted that Theorem 6.3 is applicable to any type of modulated symbol. In the following, an example is provided in employing the above theorem to generate a unitary trace-orthogonal LD code:

Example 6.2 Using the V -structured matrix in Example 6.1 and employing Theorem 6.3, we have

$$\begin{aligned} \mathbf{A}_{11} = \mathbf{A}_{22} &= \frac{1}{\sqrt{2}} \begin{pmatrix} \cos \theta & 0 \\ 0 & \sin \theta \end{pmatrix} \\ \mathbf{A}_{12} = \mathbf{A}_{21} &= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & \sin \theta \\ \cos \theta & 0 \end{pmatrix} \\ \mathbf{B}_{11} = -\mathbf{B}_{22} &= \frac{j}{\sqrt{2}} \begin{pmatrix} -\sin \theta & 0 \\ 0 & \cos \theta \end{pmatrix} \\ \mathbf{B}_{12} = -\mathbf{B}_{21} &= \frac{j}{\sqrt{2}} \begin{pmatrix} 0 & \cos \theta \\ -\sin \theta & 0 \end{pmatrix} \end{aligned}$$

It can be seen that the family of matrices $\{\mathbf{A}_{r,t}, \mathbf{B}_{r,t}\}$ constitutes a trace-orthonormal LD code with a symbol rate $R = 2$ per channel use.

6.5 Design of full diversity LD codes

In this section, we consider the design of *full diversity* LD codes using the construction of unitary trace-orthogonal LD codes developed in the previous section. We will show how the V-structured matrices can be applied to generate full diversity unitary trace-orthogonal code. The following fundamentals of algebraic number theory will be employed in this section.

6.5.1 Some basic definitions and results in algebraic number theory

For completeness of the exposition, we begin by briefly introducing some definitions and results extracted from [8, 32, 6, 66, 31, 71, 7], which are the key to the design of full diversity LD codes.

Definition 6.3 (*The Euler Function*). For a rational positive integer n , i.e., $n \in \{1, 2, \dots\}$, the Euler function $\varphi(n)$ is defined as the number of all positive rational integers that do not exceed n and are prime to n . ■

Lemma 6.2 *The Euler function has the following properties:*

1. $\varphi(1) = 1$;
2. $\varphi(p^\alpha) = p^{\alpha-1}(p-1)$, where p is prime and α is a positive rational integer;
3. (*Multiplicative Property*) If $\gcd(m, n) = 1$ where “gcd” stands for the greatest common divisor, i.e., if m and n are relatively prime rational integers, then $\varphi(mn) = \varphi(m)\varphi(n)$;
4. Let $m = \prod_{k=1}^r p_k^{\alpha_k}$ where p_k are prime and α_k are positive rational integers. Then, $\varphi(m) = m \prod_{k=1}^r (1 - \frac{1}{p_k})$.

We denote the rational integer ring $\{0, \pm 1, \pm 2, \dots\}$ by \mathbb{Z} , the Gaussian integer ring, $\{a + jb; a, b \in \mathbb{Z}\}$ by $\mathbb{Z}[j]$, and the Eisenstein ring, $\{a + \zeta_3 b; a, b \in \mathbb{Z}\}$ by $\mathbb{Z}[\zeta_3]$ with $\zeta_n = \exp(j\frac{2\pi}{n})$. We also denote the corresponding quotient fields by \mathbb{Q} , by $\mathbb{Q}(j)$, i.e., $\mathbb{Q}(j) = x + jy; x, y \in \mathbb{Q}$, and by $\mathbb{Q}(\zeta_3)$, i.e., $\mathbb{Q}(\zeta_3) = x + \zeta_3 y; x, y \in \mathbb{Q}$, respectively. (Here, we use square brackets for rings and parentheses for fields).

Definition 6.4 An n th “order” cyclotomic polynomial is defined as

$$\Phi_n(x) = \prod_{k=1, \gcd(n,k)=1}^{n-1} (x - \zeta_n^k)$$

The degree of $\Phi_n(x)$ is the Euler function [66], $\varphi(n)$. Furthermore, $\mathbb{Q}(\zeta_n)$ is called a cyclotomic field and its integer ring $\mathbb{Z}[\zeta_n]$ is called a cyclotomic ring. ■

In the following, we use \mathbb{I}_n to collectively denote \mathbb{Z} , or $\mathbb{Z}[j]$, or $\mathbb{Z}[\zeta_3]$ and use \mathbb{K} to collectively denote their respective quotient fields, \mathbb{Q} , $\mathbb{Q}(j)$, or $\mathbb{Q}(\zeta_3)$.

Definition 6.5 (Algebraic Number and Algebraic Integer) A field \mathbb{L} is said to be an extension of field \mathbb{K} provided that \mathbb{K} is a subfield of \mathbb{L} . An element $\theta \in \mathbb{C}$ is said to be algebraic over \mathbb{K} provided that θ is a root of some nonzero polynomial $f(x) = \sum_{\ell=0}^d a_\ell x^\ell \in \mathbb{K}[x]$, the ring of polynomials with coefficients in \mathbb{K} , and θ is not a root of such a polynomial of degree less than d . Such a polynomial is called a minimal polynomial. If $f(x)$ can be chosen monic (having the coefficient of the highest power term being unity), with coefficients in \mathbb{I}_n , the number θ is said to be integral over \mathbb{I}_n . ■

Definition 6.6 (Trace and Norm). Let θ be algebraic over \mathbb{K} with a degree d . Let σ_ℓ for $\ell = 1, 2, \dots, d$ be d automorphisms of \mathbb{K} in \mathbb{C} that fix \mathbb{K} pointwise (i.e., $\sigma_\alpha = \alpha$). Then, the trace $\text{Tr}(\theta)$ and the norm $\text{Nr}(\theta)$ of θ from \mathbb{K} are defined, respectively, as $\text{Tr}(\theta) = \sum_{\ell=1}^d \sigma_\ell(\theta)$ and $\text{Nr}(\theta) = \prod_{\ell=1}^d \sigma_\ell(\theta)$. ■

We state the results of the following two lemmas which are presented in [16] and [100], respectively and are important to the development of the ensuing materials of this chapter.

Lemma 6.3 *Let \mathcal{O} be the set of elements of \mathbb{K} which are integral over \mathbb{I}_n . Then, \mathcal{O} is the number ring (subset of integer ring) of \mathbb{K} over \mathbb{I}_n and both the trace and norm of $\theta \in \mathcal{O} \setminus \{0\}$ belong to $\mathbb{I}_n \setminus \{0\}$. As a result, $|\text{Tr}(\theta)|, |\text{Nr}(\theta)| \geq 1$. (Here, $\mathcal{O} \setminus \{0\}$ denotes the set \mathcal{O} minus the element 0). ■*

Lemma 6.4 *Let $P = LJ$ and $L_t = \frac{\varphi(P)}{\varphi(L)}$. Then, all the L_t automorphisms σ_i , $1 \leq i \leq L_t$, of field $\mathbb{Q}(\zeta_P)$ which fix subfield $\mathbb{Q}(\zeta_L)$ can be represented by $\sigma_i(\zeta_P) = \zeta_P^{1+P_iL}$ for $1 \leq i \leq L_t$, where P_i , $1 \leq i \leq L_t$, are the integers that satisfy $0 = P_1 < P_2 < \dots < P_{L_t} \leq J - 1$ and $1 + P_iL$ and P are co-prime for $1 \leq i \leq L_t$. Furthermore, if we let \mathbf{G}_{LJ} denote a generating matrix such that $[\mathbf{G}_{LJ}]_{mn} = \{\sigma_m(\zeta_P^n)\}_{1 \leq m, n \leq L_t}$ and $\mathbf{x} = [x_1, x_2, \dots, x_{L_t}]^T = \mathbf{G}_{LJ} \mathbf{s}$, where $\mathbf{s} \in \mathbb{Z}^{L_t}[\zeta_L]$, then, the generating matrix \mathbf{G}_{LJ} is of full diversity over $\mathbb{Z}^{L_t}[\zeta_L]$, i.e., $\prod_{k=1}^{L_t} x_k \neq 0$ for any nonzero symbol vector \mathbf{s} belonging to $\mathbb{Z}^{L_t}[\zeta_L]$. ■*

For a properly fixed positive rational integer L_t , Wang *et al* [100] optimized L and J such that the resulting cyclotomic lattice is as dense as possible. Unfortunately, the resulting optimal generating matrix is not unitary and the number of transmitter antennas, L_t , is limited to be a specific rational integer. Nonetheless, in this chapter, Lemma 6.4 forms the basis of the joint design of a cyclotomic rotation matrix, the corresponding constellation and the Diophantine number such that our resulting LD code guarantees full diversity. Here, we employ Lemma 6.4 together with the following lemma, whose proof is provided in Appendix D:

Lemma 6.5 *Let $M = \prod_{k=1}^r p_k^{\alpha_k}$, where each p_k is prime and $\alpha_k \geq 1$. Then, for a*

positive rational integer $P = LM$, its Euler number $\varphi(P) = M\varphi(L)$ if and only if $L = L_1 \prod_{k=1}^r p_k^{\beta_k}$, where each $\beta_k \geq 1$ and L_1 is prime to M . ■

Jointly applying Lemma 6.5 and Lemma 6.4 immediately results in the following full diversity cyclotomic rotation matrix and a corresponding cyclotomic ring³:

Corollary 6.1 Let $M = \prod_{k=1}^r p_k^{\alpha_k}$, where each p_k is prime and $\alpha_k \geq 1$, and let $P = LM$, where $L = L_1 \prod_{k=1}^r p_k^{\beta_k}$, $\beta_k \geq 1$ with L_1 being prime to M . The matrix $\mathcal{R}_0 = \mathbf{D}_M^H \text{diag}(1, \zeta_P, \dots, \zeta_P^{M-1})$ is unitary and is of full diversity over the cyclotomic ring $\mathbb{Z}[\zeta_L]$. ■

Proof. Using Lemmas 6.4 and 6.5, where $J = M, P = LM$ (and thus $L_t = M$ due to the property of the Euler function in [66, 106]), we obtain $P_i = i-1$ for $i = 1, 2, \dots, M$, since in this case, $1 + (i-1)L$ is prime to P . As a result, $\sigma_m(\zeta_P) = \zeta_P^{1+(m-1)L}$ and hence, $\mathcal{R}_0(P, M) = \mathbf{D}_M^H \text{diag}(1, \zeta_P, \dots, \zeta_P^{M-1})$. □

Using Corollary 6.1, for a fixed M , we can choose L as small as possible such that

1. $L = \prod_{k=1}^r p_k$ if $M \neq 2^n$ for any rational positive integer n ;
2. $L = 4$ or $L = 6$, if $M = 2^n$ for some rational positive integer n .

It has been shown [8, 32, 6, 31, 7, 14] that a unitary rotation matrix can be obtained for the particular case of $M = 2^m 3^n$. Here, we see that we can apply Corollary 6.1 to the Gaussian integer ring $\mathbb{Z}[\zeta_4]$ or the Eisentein integer ring $\mathbb{Z}[\zeta_6]$ and arrive at the same result of a unitary rotation matrix. Hence, we can conclude that Corollary 6.1 is a generalization of the available result for the particular case of $M = 2^m 3^n$.

³Corollary 6.1 appears similar to the proposition given in [65]. However, Corollary 6.1 not only shows the existence of, but also provides us with an explicit and flexible choice of the cyclotomic ring from which the transmitted symbols should be selected. In addition, along with Lemma 6.4, Corollary 1 provides an explicit and flexible choice of the Diophantine number. The latter remark can be further seen in Theorem 6.4.

6.5.2 Design of full diversity LD codes

We are now in a position to design a full diversity unitary trace-orthogonal LD code. First, the following result of matrix theory [43] is introduced to proceed with the design:

Lemma 6.6 *For $M \times M$ nonnegative definite matrices Ω_1 and Ω_2*

$$\det(\Omega_1 + \Omega_2) \geq \det(\Omega_1) + \det(\Omega_2) \quad (6.27)$$

For Ω_1 being positive definite, equality in (6.27) holds if and only if $\Omega_2 = \mathbf{0}$. ■

The main result for the design of a full diversity code is presented as follows:

Theorem 6.4 *Let $K = RT$ with $T = M\tilde{L}$ and $0 < R \leq M$. Also, let $M = \prod_{m=1}^d q_m^{\lambda_m} \prod_{k=1}^{d_M} p_k^{\alpha_k}$, $R = \prod_{m=1}^d q_m^{\mu_m} \prod_{\ell=1}^{d_R} \tilde{p}_\ell^{\beta_\ell}$ and $L = L_0 \prod_{i=1}^d q_m^{\tau_m} \prod_{k=1}^{d_M} p_k^{\gamma_k} \prod_{\ell=1}^{d_R} \tilde{p}_\ell^{\rho_\ell}$, where q_m, p_k and \tilde{p}_ℓ are primes, $\alpha_k, \beta_\ell, \gamma_i, \rho_\ell \geq 1$ and L_0 is prime to both M and R . If we choose the V -structured matrices in Theorem 6.3 as follows:*

$$\Theta_r = \zeta_{LRM^2}^{r-1} \mathbf{U} \otimes \begin{pmatrix} \mathcal{R}_0[1 : M_1, :] \\ \mathbf{0}_{M_2 \times M} \end{pmatrix} \quad (6.28a)$$

$$\Phi_r = \zeta_{LRM^2}^{r-1} \mathbf{U} \otimes \begin{pmatrix} \mathbf{0}_{M_1 \times M} \\ \mathcal{R}_0^*[M_1 + 1 : M, :] \end{pmatrix} \quad (6.28b)$$

for $r = 1, \dots, R$, where $M_1 + M_2 = M, M_1, M_2 \geq 0$, \mathbf{U} is an arbitrarily given $\tilde{L} \times \tilde{L}$ unitary matrix and \mathcal{R}_0 is the unitary rotation matrix defined in Corollary 6.1, then, the resulting signal matrix $\mathbf{X}(\mathbf{s})$ provides full diversity over any constellation carved from $\mathbb{Z}^K[\zeta_L]$ with a symbol transmission rate R symbols per channel use. ■

The proof of Theorem 6.4 is provided in Appendix E.

Remarks on Theorem 6.4:

1. When $M_2 = 0$, the resulting code reduces to a full diversity linear space-time block code [22, 64].
2. $\zeta_{LRM^2}^{r-1}$ in Eq. (6.28) scales the rotation matrix and is thus a Diophantine number. Therefore, Theorem 6.4 not only provides a unified framework to systematically design a full diversity cyclotomic LD code at any symbol transmission rate less than or equal to the number of transmitter antennas, but also provides a method to properly select the rotation matrices and the corresponding Diophantine number.
3. Different V-structured matrices result in different coding matrices $\mathbf{X}(\mathbf{s})$ and hence, different coding gains. How to choose a V-structured matrix such that the resulting coding signal matrix has a good coding gain is still an open problem.
4. If $R = M$, our resulting LD code is information lossless and provides full rate and full diversity. The case of $N < M$ will be considered in the following section.

6.6 Design of full diversity linear space-time block codes for $N < M$

The design of LD codes provided in the previous section requires the number of receiver antennas to be larger than or equal to the number of transmitter antennas to maintain full rate and full diversity. In this section, we examine the design of linear space-time block codes for $N < M$. We propose the *unitary trace-orthogonal linear cyclotomic space-time block codes* by taking advantage of the delay. The symbol rate here equals N symbols per channel use. In particular, when the number of the transmitter antennas is equal to 2^m , we show that our proposed code minimizes the

worst case pair-wise error probability of the ML detector for an μ -ary QAM signal; i.e., the optimal coding gain is achieved. The conditions and results of the design are stated in the following theorem.

Theorem 6.5 *Let $M = \prod_{k=1}^r p_k^{\alpha_k}$, $L = L_1 \prod_{k=1}^r p_k^{\beta_k}$, where each p_k is prime, $\alpha_k, \beta_k \geq 1$ and L_1 is prime to M . Also, let $\gcd(M, N) = D$, $K = NT$, $T = \frac{MT_1}{D}$ and $T_2 = \frac{NT_1}{D}$ with $T_1 = \lceil \frac{(M-1)D}{M-N} \rceil$. If we let $\mathbf{R}_m = [\text{diag}(\mathcal{R}_0[:, m]), \mathbf{0}_{M \times (T-M)}]$ for $m = 1, 2, \dots, M$, where $\mathcal{R}_0[:, m]$ denotes the m -th column vector of matrix $\mathcal{R}_0 = \mathbf{D}_M \text{diag}(1, \zeta_P, \dots, \zeta_P^{M-1})$ defined in Corollary 6.1 of the previous section, and $\mathbf{A}_{t,m} = \sqrt{\frac{MT}{K}} \mathbf{R}_m \mathbf{P}_T^{t-1}$ for $t = 1, 2, \dots, T_2$, where \mathbf{P}_T is a $T \times T$ circulant matrix, then we have the following three statements:*

1. *The matrix family $\{\mathbf{A}_{t,m}\}$ constitutes a unitary trace-orthogonal linear triangular space-time block code with a symbol rate N per channel use.*
2. *The resulting coding signal matrix provides full diversity over any constellation carved from a cyclotomic ring $\mathbb{Z}[\zeta_L]$.*
3. *In particular, when the number of transmitter antennas is equal to 2^m , the so designed code minimizes the worst case pair-wise error probability for μ -ary QAM; i.e.,*

$$\mathcal{A} = \arg \min_{\text{tr}(\mathcal{F}\mathcal{F}^H) \leq K} \max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K, \mathbf{s} \neq \mathbf{s}'} P_{\mathcal{F}}(\mathbf{s} \rightarrow \mathbf{s}')$$

$$\min_{\text{tr}(\mathcal{F}\mathcal{F}^H) \leq K} \max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K, \mathbf{s} \neq \mathbf{s}'} P_{\mathcal{F}}(\mathbf{s} \rightarrow \mathbf{s}') = J \left(\frac{\rho \nu_{\min}^2}{8MN} \right)$$

■

Proof.

- 1) The proof of Statement 1 follows that of Theorem 6.3.
- 2) For Statement 2, it suffices to prove that the code matrix $\mathbf{X}(\mathbf{e})\mathbf{X}^H(\mathbf{e})$ is invertible

for any nonzero vector $\mathbf{e} \in \mathbb{Z}^K[\zeta_L]$, where $\mathbf{e} = \mathbf{s} - \mathbf{s}'$. Since $\mathbf{e} \neq \mathbf{0}$, there exists some layer i for $1 \leq i \leq T_2$ such that $\mathbf{e}_i = [e_{M(i-1)+1} \ e_{M(i-1)+2} \ \cdots \ e_{M(i-1)+M}]^T \neq \mathbf{0}$. By code construction, we can always assume that $\mathbf{X}_i = \mathbf{V}\mathbf{e}_i$ is located in the diagonal line of the triangular submatrix, say, $\mathbf{X}_1(\mathbf{e})$, consisting of the first M columns of the signal coding matrix $\mathbf{X}(\mathbf{e})$. If this is not the case, then we permute the column or row of $\mathbf{X}(\mathbf{e})$ to obtain this placement. This permutation does not change the determinant of $\mathbf{X}(\mathbf{e})\mathbf{X}^H(\mathbf{e})$. In other words, without loss of generality, we can always assume that $\mathbf{X}(\mathbf{e})$ can be written as $\mathbf{X}(\mathbf{e}) = [\mathbf{X}_1(\mathbf{e}) \ ; \ \mathbf{X}_2(\mathbf{e})]$ where $\mathbf{X}_1(\mathbf{e})$ is a triangular matrix with non-zero diagonal entries $[\mathbf{X}_1(\mathbf{e})]_k = \sigma_i(k) \neq 0$ for $k = 1, 2, \dots, M$. Together with this structure as well as Lemma 6.6 in the previous section, we obtain

$$\begin{aligned}
\det(\mathbf{X}(\mathbf{e})\mathbf{X}^H(\mathbf{e})) &= \det(\mathbf{X}_1(\mathbf{e})\mathbf{X}_1^H(\mathbf{e}) + \mathbf{X}_2(\mathbf{e})\mathbf{X}_2^H(\mathbf{e})) \\
&\geq \det(\mathbf{X}_1(\mathbf{e})\mathbf{X}_1^H(\mathbf{e})) + \det(\mathbf{X}_2(\mathbf{e})\mathbf{X}_2^H(\mathbf{e})) \\
&\geq \det(\mathbf{X}_1(\mathbf{e})\mathbf{X}_1^H(\mathbf{e})) \\
&= \prod_{k=1}^M |\sigma_i(k)|^2 > 0
\end{aligned} \tag{6.29}$$

where equality in both inequalities holds iff $\mathbf{X}_2(\mathbf{e}) = \mathbf{0}$. This completes the proof of Statement 2.

3) We now prove Statement 3. Let us first establish an upper bound of the worst case pair-wise error probability for this code. To that end we apply Minkowski's inequality [43, p.482] which states that if $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ are both $M \times M$ positive definite matrices, then,

$$\det(\mathbf{\Omega}_1 + \mathbf{\Omega}_2)^{1/M} \geq \det(\mathbf{\Omega}_1)^{1/M} + \det(\mathbf{\Omega}_2)^{1/M} \tag{6.30}$$

where equality holds if and only if $\mathbf{\Omega}_2 = c\mathbf{\Omega}_1$ for some constant c . Now, putting $\mathbf{\Omega}_1 = \mathbf{I}_M$ and $\mathbf{\Omega}_2 = \mathbf{X}(\mathbf{e})\mathbf{X}^H(\mathbf{e})$ in Eq. (6.30) and applying Eq. (6.29) and Lemma 6.3, we

obtain, for any nonzero vector \mathbf{e} and nonzero θ in the interval $[0, \pi/2]$,

$$\begin{aligned}
& \det \left(\mathbf{I}_M + \frac{\rho}{8M \sin^2 \theta} \mathbf{X}(\mathbf{e}) \mathbf{X}^H(\mathbf{e}) \right)^{1/M} \\
& \geq 1 + \frac{\rho}{8M \sin^2 \theta} \det \left(\mathbf{X}(\mathbf{e}) \mathbf{X}^H(\mathbf{e}) \right)^{1/M} \\
& \geq 1 + \frac{\rho |\text{Nr}(\mathbf{e}_i)|^{2/M}}{8MN \sin^2 \theta} \\
& \geq 1 + \frac{\rho \nu_{\min}^2}{8MN \sin^2 \theta}
\end{aligned} \tag{6.31}$$

Here, the first equality holds iff $\mathbf{X}(\mathbf{e}) \mathbf{X}^H(\mathbf{e})$ is a scaled identity matrix and the second equality holds iff $\mathbf{X}_2(\mathbf{e}) = \mathbf{0}$. Therefore, both equalities hold iff \mathbf{s} and \mathbf{s}' are neighbor points, i.e., $\|\mathbf{s} - \mathbf{s}'\| = \nu_{\min}$. We conclude from this that $\det \left(\mathbf{I}_M + \frac{\rho}{8M \sin^2 \theta} \mathbf{X}(\mathbf{e}) \mathbf{X}^H(\mathbf{e}) \right) \geq \left(1 + \frac{\rho \nu_{\min}^2}{8MN \sin^2 \theta} \right)^M$ which results in

$$\max_{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K, \mathbf{s} \neq \mathbf{s}'} P_A(\mathbf{s} \rightarrow \mathbf{s}') \leq J \left(\frac{\rho \nu_{\min}^2}{8MN} \right) \tag{6.32}$$

where the equality holds iff $\|\mathbf{s} - \mathbf{s}'\| = \nu_{\min}$. Combining (6.32) with Theorem 6.2 in Section 6.3B yields

$$\min_{\text{tr}(\mathcal{F}^H \mathcal{F}) \leq K} \max_{\substack{\mathbf{s}, \mathbf{s}' \in \mathcal{S}^K \\ \mathbf{s} \neq \mathbf{s}'}} P_{\mathcal{F}}(\mathbf{s} \rightarrow \mathbf{s}') = J \left(\frac{\rho \nu_{\min}^2}{8MN} \right)$$

This completes the proof of Statements 3 and hence, Theorem 6.5. \square

6.7 Design examples and simulations

In this section, we present some examples of unitary trace-orthogonal space-time code designs and examine their performance in comparison to other codes available in literature.

Example 1: In this example, we consider a MIMO system having two transmitter

antennas and two receiver antennas. We design an information lossless full rate full diversity linear dispersion code over Gaussian integer ring $\mathbb{Z}[i]$. A coded signal matrix

is given by $\mathbf{X}(\mathbf{s}) = \frac{1}{\sqrt{2}} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$ where

$$x_{11} = (s_1 + s_2) \cos \theta + (s_2^* - s_1^*) \sin \theta, \quad x_{12} = e^{\frac{i\pi}{4}} ((s_3 + s_4) \sin \theta + (s_4^* - s_3^*) \cos \theta),$$

$$x_{21} = e^{\frac{j\pi}{4}} ((s_3 + s_4) \cos \theta + (s_3^* - s_4^*) \sin \theta), \quad \text{and } x_{22} = (s_1 + s_2) \sin \theta + (s_1^* - s_2^*) \cos \theta.$$

Thus, we have

$$\begin{aligned} 2 \det(\mathbf{X}(\mathbf{e})) &= \frac{(s_1 + s_2)^2 - (s_1^* - s_2^*)^2}{2} \sin 2\theta \\ &\quad + (s_1 + s_2)(s_1^* - s_2^*) \cos 2\theta \\ &\quad - \frac{(s_3 + s_4)^2 - (s_3^* - s_4^*)^2}{2} j \sin 2\theta \\ &\quad - (s_3 + s_4)(s_3^* - s_4^*) j \cos 2\theta \end{aligned}$$

Now choosing $\sin 2\theta = \frac{1}{\sqrt{5}}$ and $\cos 2\theta = \frac{2}{\sqrt{5}}$, we have

$$4\sqrt{5} \det(\mathbf{X}(\mathbf{e})) = (s_1 + s_2 + 2(s_1^* - s_2^*))^2 - 5(s_1^* - s_2^*)^2 - j((s_3 + s_4 + 2(s_3^* - s_4^*))^2 - 5(s_3^* - s_4^*)^2).$$

Let $S_1 = (s_1 + s_2 + 2(s_1^* - s_2^*))^2$, $S_2 = (s_1^* - s_2^*)^2$, $S_3 = (s_3 + s_4 + 2(s_3^* - s_4^*))^2$ and $S_4 = (s_3^* - s_4^*)^2$, then, $4\sqrt{5} \det(\mathbf{X}(\mathbf{e})) = S_1^2 - 5S_2^2 - j(S_3^2 - 5S_4^2) = (S_1^2 - jS_3^2) - 5(S_2^2 - jS_4^2)$. It can be easily verified by checking the remainders of 5 dividing S_k^2 that $\det(\mathbf{X}(\mathbf{e})) = 0$ iff $\mathbf{e} = \mathbf{0}$. On the other hand, since $((S_1^2 - jS_3^2) - 5(S_2^2 - jS_4^2)) - ((S_1^2 - S_2^2) - j(S_3^2 - S_4^2))$ is divisible by 4, while both $S_1^2 - S_2^2$ and $S_3^2 - S_4^2$ are divisible by 4, we conclude that $4\sqrt{5} \det(\mathbf{X}(\mathbf{e}))$ is divisible by 4. As a result, $|4\sqrt{5} \det(\mathbf{X}(\mathbf{e}))| \geq 4$, which⁴ implies $|\det(\mathbf{X}(\mathbf{e}))| \geq \frac{1}{\sqrt{5}}$. Fig. 6.1 shows that the error performance comparison of our code with those proposed in [1, 17, 107, 5].

Example 2: For three transmitter antennas and three receiver antennas, we can directly apply Theorem 6.4 to design an information lossless full rate full diversity linear space-time block code. In this case, the determinant of a signal matrix is

⁴More recently, Dayal and Varanasi [19], Belfiore, Rekaya and Viterbo [5] constructed a linear space-time block code that has the same coding gain with a non-vanishing determinant.

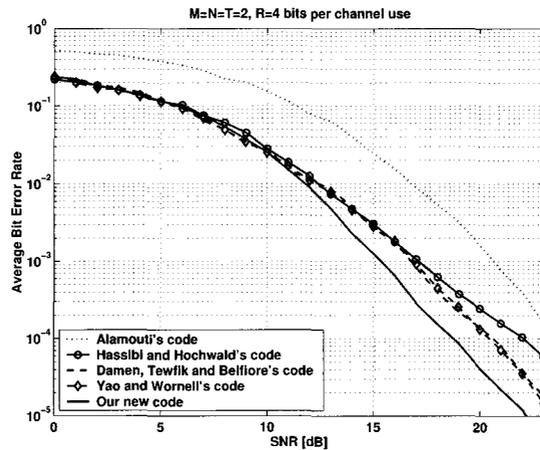


Figure 6.1: The error performance comparison of our new code with the current available codes in [1, 37, 17, 107].

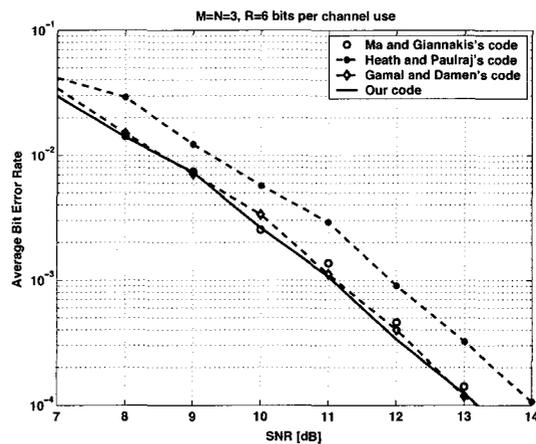


Figure 6.2: The error performance comparison of our new code with the current available codes in [39, 22, 64].

simple to evaluate. Thus, we can first take care of the coefficients of the expansion of its determinant and then apply Theorem 6.3, where the V-structured matrices are chosen as $\mathbf{V}_1 = \mathbf{D}_3^H \text{diag}(1, \zeta_9, \zeta_9^2)$, $\mathbf{V}_2 = \zeta_{27} \mathbf{P}_3 \mathbf{D}_3^H \text{diag}(1, \zeta_9, \zeta_9^2)$, $\mathbf{V}_3 = \zeta_{27}^2 \mathbf{P}_3^2 \mathbf{D}_3^H \text{diag}(1, \zeta_9, \zeta_9^2)$. The resulting coding signal matrix is

$$\mathbf{X}(\mathbf{s}) = \begin{pmatrix} S_{11} & \zeta_{27}^2 S_{33} & \zeta_{27} S_{22} \\ \zeta_{27} S_{23} & S_{12} & \zeta_{27}^2 S_{31} \\ \zeta_{27}^2 S_{32} & \zeta_{27} S_{21} & S_{13} \end{pmatrix}$$

where $\mathbf{S}_k = [S_{k1}, S_{k2}, S_{k3}]^T = \mathbf{D}_3^H \text{diag}(1, \zeta_9, \zeta_9^2) \mathbf{s}_k$ with $\mathbf{s}_k = [s_{3k-2}, s_{3k-1}, s_{3k}]^T$ for $k = 1, 2, 3$. Its determinant is given by

$$\begin{aligned} \det(\boldsymbol{\Sigma}(\mathbf{s})) &= S_{11}S_{12}S_{13} + \zeta_9 S_{21}S_{22}S_{23} + \zeta_9^2 S_{31}S_{32}S_{33} \\ &\quad - S_{11}S_{21}S_{31} - S_{12}S_{22}S_{32} - S_{13}S_{23}S_{33} \\ &= \text{Nr}(\theta_1) + \zeta_9 \text{Nr}(\theta_2) + \zeta_9^2 \text{Nr}(\theta_3) - \text{Tr}(\theta_1\theta_2\theta_3) \end{aligned}$$

where $\theta_k = s_{3k-2} + s_{3k-1}\zeta_9 + s_{3k}\zeta_9^2$ for $k = 1, 2, 3$. Since $\mathbf{s}_k \in \mathbb{Z}^3[\zeta_3]$, by Lemma 6.3, both the norm $\text{Nr}(\mathbf{s}_k)$ and the trace $\text{Tr}(\mathbf{s}_k)$ belong to the Eisenstein ring $\mathbb{Z}[\zeta_3]$ for $k = 1, 2, 3$. By Lemma 6.4, $\{1, \zeta_9, \zeta_9^2\}$ is independent over the Eisenstein cyclotomic ring $\mathbb{Z}[\zeta_3]$. Therefore, $\det(\mathbf{X}(\mathbf{s}))$ is nonzero over any constellation carved from $\mathbb{Z}^9[\zeta_3] \setminus 0$, and hence, $\mathbf{X}(\mathbf{s})$ provides full rate full diversity without information loss over any constellation carved from $\mathbb{Z}[\zeta_3]$. Such designed code has the Diophantine number with a smaller degree than that constructed by directly using Theorem 6.4. An error performance comparison of our code with those in [39, 22, 64] is shown in Fig 6.2. It can be seen that the performance of our code is slightly better than those in [22, 64]. This is because the choice of the full diversity rotation matrix and the Diophantine number affects performances. However, it is difficult to determine the choice of the rotation matrices and the Diophantine numbers to yield better performance.

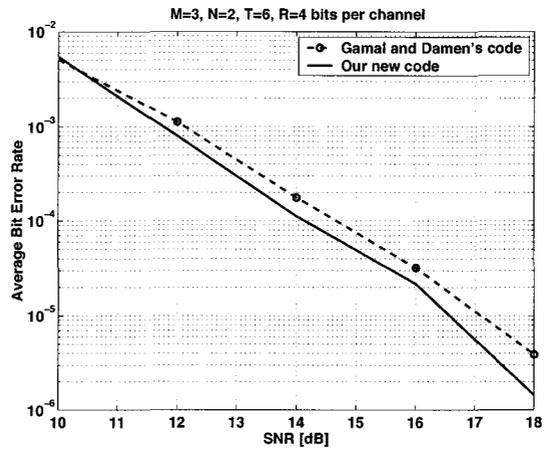


Figure 6.3: Comparison of our code with the code in [22].

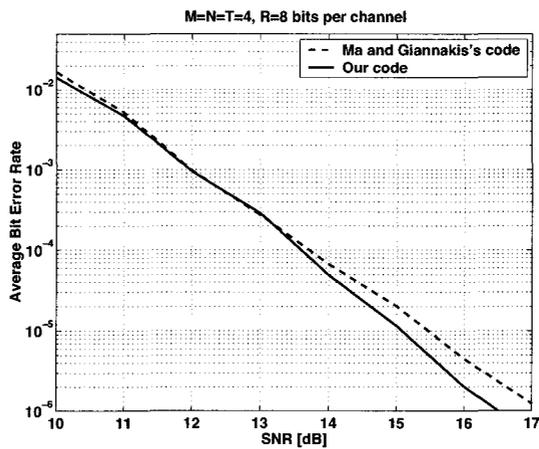


Figure 6.4: Comparison of our code with the code in [64].

Example 3: For a MIMO system having three transmitter antennas and two receiver antennas, we design a full diversity linear space-time block code with a symbol rate being equal to two symbols per channel use. The code is shown as follows

$$\mathbf{X}(\mathbf{s}) = \frac{\sqrt{2}}{2} \begin{pmatrix} X_{11} & 0 & 0 & X_{41} & X_{31} & X_{21} \\ X_{22} & X_{12} & 0 & 0 & X_{42} & X_{32} \\ X_{33} & X_{23} & X_{13} & 0 & 0 & X_{43} \end{pmatrix}$$

where

$$\begin{pmatrix} X_{i1} \\ X_{i2} \\ X_{i3} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta_3 & \zeta_3^2 \\ 1 & \zeta_3^2 & \zeta_3^4 \end{pmatrix} \begin{pmatrix} s_{i1} \\ s_{i2}\zeta_9 \\ s_{i3}\zeta_9^2 \end{pmatrix}$$

for $i = 1, 2, 3, 4$. It can be shown that $\det(\mathbf{X}(\mathbf{e})\mathbf{X}^H(\mathbf{e})) \neq 0$ for any constellation carved from $\mathbb{Z}[\zeta_3] \setminus 0$. An error performance comparison of our code with that in [22] is provided in Fig 6.3, where it can be observed that the performance of our code is better. This is obtained with a longer decoding delay.

Example 4: In this example, we compare our code with the code presented in [64] for a MIMO system having four transmitter antennas and four receiver antennas. Both codes employ the same rotation matrix in each layer, however, the proposed code adopts the Diophantine number according to Theorem 6.4, which is different from that in [64]. Simulation results are shown in Fig. 6.4.

6.8 Conclusion

In this chapter, we examine the design of LD space-time codes applied to a MIMO communication system from both the information-theoretic and detection error viewpoints. We arrive at the conclusion that a good LD code should have a unitary

and trace-orthogonal structure. This desirable structure has prompted us to come up with a method to construct such a code family by the way of the V-structured matrix. Furthermore, by applying cyclotomic field theory, we have developed a systematic method to jointly design the rotation matrix, the Diophantine number and constellation. This has led us to arrive at an efficient design of full diversity rectangular LD code at a symbol transmission rate less than or equal to the number of the transmitter antennas. In particular, when the number of the transmitter antennas is $M = 2^m$, and the number of the receiver antennas is less than M , we proposed a linear space-time block code design that minimizes the worst case average pair-wise error probability of the ML detector for a QAM signal. As a result, this code achieves the optimal coding gain.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

In this thesis, the design of STBCs for MIMO communication systems has been addressed. More specifically, a family of unitary trace-orthogonal codes has been considered and its performance in a variety of communication applications has been analyzed. In particular, codes from this family were shown to be optimal in a number of senses.

The combination of unitary and trace-orthogonal structures was first proposed as the optimal structure for a full rate STBC to achieve the minimum detection error probability for a linear receiver when the MIMO system has $N \geq M$ as reviewed in Chapter 2. In [61, 55], the unitary trace-orthogonal condition was proved to be both necessary and sufficient for optimality. Furthermore, when the communication system is studied from information theory point of view, trace-orthogonality was proved [55] to be the condition for a STBC to be information lossless.

For a MISO communication system, a general criterion for designing *full diversity* STBCs for a *linear receiver* was presented in Chapter 3. Also in this chapter, a

member from the unitary trace-orthogonal family, the Toeplitz code was proposed as the first non-orthogonal full diversity code for linear receivers. The proposed code is superior to orthogonal STBC in that its symbol transmission data rate is higher than orthogonal code (for $M \geq 3$) and can asymptotically approach unity. In addition, the proposed Toeplitz STBC was proved to asymptotically approach the optimal diversity v.s. multiplexing tradeoff curve with the use of square QAM signals. When applied to a system having correlated channel fading coefficients, the structure of the Toeplitz code enables an efficient computation of the optimal beamforming matrix that minimizes the detection error. This is realized by transforming the direct formulation of the design problem into a convex optimization problem.

For a general MIMO system with $N \geq M$, a linear receiver cannot extract full diversity. Since a linear receiver achieves its minimum detection error with the use of the unitary trace-orthogonal code, the associated diversity gain is therefore the maximum achievable diversity gain for a linear receiver. By analyzing the expression of the error probability, this maximum diversity gain was shown to be equal to $N - M + 1$ in Chapter 4. This result provides a theoretical measure of the performance gap between the optimal ML receiver and the simple linear receiver.

To improve the performance of a linear receiver, a multiple block transmission scheme was studied in Chapter 4. In that scheme, the signal symbols are coded so that they span several (block) channel realizations, and as such, it enjoys more degrees of freedom than a conventional signal block scheme. However, the application of an ML detector to this system incurs an extremely high computational cost. On the other hand, a linear receiver can be applied here for its simplicity. By minimizing the detection error probability of a linear receiver applied with a multiple block scheme, the optimal full rate code was derived in Chapter 4. The optimal code is a multi-block version of the unitary trace-orthogonal code. Additionally, it was proved in

Chapter 4 that the achieved minimum error probability is an decreasing function of L , the number of blocks that a code spans. Further studies showed that the diversity gain grows with L . The analysis in Chapter 4 also showed that the special structure of the multi-block unitary trace-orthogonal code reduces the detection complexity to the same order as that for single block transmission. Hence, a low cost linear receiver and a multi-block transmission scheme may outperform an ML detector with single block transmission. The compromise here is the increased decoding delay.

By relaxing the code from the constraint of *full* symbol rate transmission, a multi-block transmission scheme specially designed for a linear receiver was proposed in Chapter 5. Noticing that the diversity gain for a linear receiver is increasing with $N - M$, the proposed scheme therefore constructs a “tall” ($N > M$) equivalent channel matrix. It was proved in Chapter 5 that the proposed system has significantly improved diversity gain than the single block or multi-block systems with full symbol rate. Since the equivalent channel matrix maintains the same number of columns as the single block channel, the detection complexity remains almost the same as that of the single block case.

All the above-mentioned work was focused on the application of the unitary trace-orthogonal structure to linear STBCs and linear receivers. In fact, the structure can also be extended to more general codes (e.g., linear dispersion codes), and to more complicated detectors, e.g., ML detectors. When LD codes were considered in Chapter 6, the unitary trace-orthogonal structure was also shown to be the optimum from points of view of both information theory (information lossless) and detection error probability. For a MIMO system having an ML detector, a properly designed STBC is able to extract full diversity. However, designing such codes is generally complicated due to the existence of a huge number of design parameters. In Chapter 6, by constraining the code to be unitary trace-orthogonal code, the number of unknown

parameters was reduced. As a result, a systematic method in generating full diversity ML code was presented.

7.2 Future Work

The best STBC for ML detection: In designing STBC with the aim of minimizing the detection error probability for an ML detector, there are two gains to be optimized: diversity gain and coding gain. The maximum diversity gain equals MN and it has been achieved by numerous STBCs. The coding gain (for ML detectors) is associated with the determinant of the STBCs. Fruitful results have been presented in designing a STBC with non-vanishing determinant. However, the best achievable coding gain is still unknown. For example, when the signals are chosen from a constellation having integer points, e.g., QAM constellations, the minimum distance is equal to one. For such a constellation, does there exist a full diversity code that enables the minimum determinant to be equal to 1? If it exists, it is the best code.

The best communication strategy for linear receivers: In this thesis, the optimal full rate STBC was designed for linear receivers with the optimality proved to be both necessary and sufficient in minimizing the detection error probability. Further, when the system has a full rate transmission, the maximum achievable diversity gain by a linear receiver was analyzed. On the other hand in Chapter 5, the full rate constraint was relaxed, and a much more reliable transmission scheme for linear receivers was proposed. This leads to the following question: What is the best transmission scheme for a linear receiver? The answer to this question is strongly influenced by the fact that due to its simplicity, a linear receiver cannot take advantage of the full degrees of freedom offered by a MIMO system.. This is the key factor in understanding the improved system performance obtained by relaxing the full rate constraint.

More importantly, does there exist a fundamental tradeoff for linear receivers that is analogous to the diversity v.s. multiplexing gain tradeoff for optimal receivers?

Code design for Ricean fading channels: The channel conditions studied in this thesis are assumed to be Rayleigh fading. However in general scenarios, the channel mean may not be perfectly zero, and this renders a Ricean fading model. Recently, emerging research [44,67,68,99,69,97,45,2,70] has been carried out for such a model. Due to the complicated mathematical formulations, the code design problems for this model are hard to dealt with. In [59], two problems for Ricean fading channels have been considered: i) the optimal input covariance that achieves the ergodic channel capacity, and ii) the optimal beamforming that minimizes the MSE at the output of a linear MMSE equalizer. For both problems, the conditions in achieving the optimality have been proposed and proved. However, these conditions are expressed in the form of the expectation over complicated random functions, which can not be easily applied in obtaining the optimal parameters. Furthermore, obtaining the expected values of these functions is mathematically intractable. In [59], the close approximations for both functions have been presented by employing Taylor series expansion, which enables a fast computation of the design parameters numerically. However, the closed form solutions to the original problems are still unknown. These solutions could be obtained by analyzing and simplifying the optimal conditions proposed in [59] to reduce the complexity in evaluating the expectations.

Appendix A

Proof of Theorem 4.2

Substituting the PDF of the eigenvalues in Eq. (4.5) into the RHS of Eq. (4.8) and denoting it by $\eta(\rho)$, we have

$$\eta(\rho) = \frac{\alpha(M, N)\sqrt{e}}{2} \int_0^\infty d\lambda_1 \int_0^{\lambda_1} d\lambda_2 \cdots \int_0^{\lambda_{M-2}} \exp\left(-\sum_{i=1}^{M-1} \lambda_i\right) \prod_{i=1}^{M-1} \lambda_i^{(N-M)} \prod_{i<j}^{M-1} (\lambda_i - \lambda_j)^2 d\lambda_{M-1} \\ \cdot \int_0^{\lambda_{M-1}} \exp\left[-\left(\frac{M/2}{\sum_{i=1}^M (1 + \frac{\rho}{M}\lambda_i)^{-1}} + \lambda_M\right)\right] \lambda_M^{(N-M)} \prod_{i=1}^{M-1} (\lambda_i - \lambda_M)^2 d\lambda_M \quad (\text{A.1})$$

The proof of the theorem requires successively integrating Eq. (A.1) by parts. In the following, we examine the result of each stage of integration by parts and select the non-zero term which has the lowest order of ρ^{-1} . To proceed, we first introduce the following function sequences to simplify the notation. For $1 \leq m \leq M$, we define

$$\mu_m = \sum_{i=1}^m \left(1 + \frac{\rho}{M}\lambda_i\right)^{-1}, \quad \mu_0 = 0 \quad (\text{A.2a})$$

$$\phi_m = \left(\frac{M/2}{\mu_{m-1} + (M - m + 1) \left(1 + \frac{\rho}{M}\lambda_m\right)^{-1}} + (M - m + 1)\lambda_m\right) \quad (\text{A.2b})$$

$$\theta_m = \lambda_m^{(N-M)(M-m+1)} \prod_{i=1}^{m-1} (\lambda_i - \lambda_m)^{2(M-m+1)} \quad \text{with } \theta_1 = \lambda_1^{(N-M)M} \quad (\text{A.2c})$$

From Eq. (A.2b), we have $\phi'_m(\lambda_m) = \left(\frac{\rho(M-m+1)/2}{(\mu_{m-1}(1+\frac{\rho}{M}\lambda_m)+(M-m+1))^2} + (M-m+1) \right)$, where $(\cdot)'$ denotes the first derivative w.r.t. λ_m . Since ϕ'_m is also a function of other eigenvalues λ_i , $i \leq m$, we define

$$\psi_{mi}(\lambda_i) = \frac{\rho(M-m+1)/2}{(\mu_{i-1}(1+\frac{\rho}{M}\lambda_i)+(M-i+1))^2} + (M-m+1), \quad \text{for } i \leq m \quad (\text{A.3})$$

Since the lower limit of integration is $\lambda_m = 0$, we examine the values of ϕ_m and ϕ'_m at this limit. From Eq. (A.2b), we have $\phi_m|_0 = \left(\frac{M/2}{\mu_{m-1}+(M-m+1)} \right)$. Now, the eigenvalues are ordered by strict inequality such that $\lambda_1 > \lambda_2 > \dots > \lambda_M$. Therefore, for $i < m$, λ_i is finite and at high SNR, $(1 + \frac{\rho}{M}\lambda_i)^{-1}$ is negligible with respect to 1. Hence, $\mu_{m-1}|_{\lambda_m=0} \approx 0$ and therefore we have $\exp(-\phi_m)|_0 = \exp\left(-\frac{M/2}{M-m+1}\right) + o(1)$. Similarly, $(\phi'_m)^{-1}|_{\lambda_m=0} = 2(M-m+1)\rho^{-1} + o(\rho^{-1})$ for a finite number of transmitter antennas.

To evaluate Eq. (A.1), we use “integration by parts” such that for general functions Φ and Θ of λ ,

$$\int_{\Lambda_1}^{\Lambda_2} e^{-\Phi} \Theta \, d\lambda \triangleq A + I \quad (\text{A.4})$$

where $A \triangleq -\frac{e^{-\Phi}}{\Phi'} \Theta \Big|_{\Lambda_1}^{\Lambda_2}$ and $I \triangleq \int_{\Lambda_1}^{\Lambda_2} e^{-\Phi} \frac{\partial}{\partial \lambda} \left(\frac{1}{\Phi'} \Theta \right) d\lambda$. The proof of the theorem now follows the procedure as outlined below:

a) $N - M = 0$: Let I_{M1} denote the inner-most integral in Eq. (A.1). Integrating I_{M1} by parts we have,

$$\begin{aligned} I_{M1} &= \int_0^{\lambda_{M-1}} \exp(-\phi_M(\lambda_M)) \theta_M(\lambda_M) d\lambda_M \\ &= -\frac{\exp(-\phi_M)}{\phi'_M(\lambda_M)} \theta_M \Big|_0^{\lambda_{M-1}} + \int_0^{\lambda_{M-1}} \exp(-\phi_M(\lambda_M)) \frac{\partial}{\partial \lambda_M} \left(\frac{1}{\phi'_M(\lambda_M)} \theta_M(\lambda_M) \right) d\lambda_M \\ &\triangleq A_{M1} + I_{M2} \end{aligned} \quad (\text{A.5})$$

where $A_{M1} = -\frac{\exp(-\phi_M)}{\phi'_M(\lambda_M)} \theta_M \Big|_0^{\lambda_{M-1}}$ and

$I_{M2} = \int_0^{\lambda_{M-1}} \exp(-\phi_M(\lambda_M)) \frac{\partial}{\partial \lambda_M} \left(\frac{1}{\phi'_M(\lambda_M)} \theta_M(\lambda_M) \right) d\lambda_M$. Now, the value of A_{M1} is

the difference of the two terms obtained by substituting the upper and lower limits. For the upper limit $\lambda_M = \lambda_{M-1}$, θ_M always contains the factor $(\lambda_{M-1} - \lambda_{M-1})$ and is therefore zero. For the lower limit $\lambda_M = 0$, θ_M contains the factor $\lambda_M^{(N-M)} \Big|_{\lambda_M=0}$. If $N - M = 0$, since $0^0 = 1$, then we have $\theta_M \Big|_{\lambda_M=0} = \prod_{i=1}^{M-1} \lambda_i^2$. Substituting $e^{-\phi_M} \Big|_{\lambda_M=0}$ and $\phi'_M(\lambda_M) \Big|_{\lambda_M=0}$, we have $A_{M1} = \frac{2}{\rho} \exp\left(-\frac{M}{2}\right) \prod_{i=1}^{M-1} \lambda_i^2 + o(\rho^{-1})$. Further integration w.r.t. $\lambda_{M-1}, \dots, \lambda_1$ will not increase the order of ρ^{-1} in this term and thus this contains the lowest order of ρ^{-1} . If $(N - M) \neq 0$, then $A_{M1} = 0$ and we have to examine I_{M2} .

b) $(N - M) > 0$: In this case, for Eq. (A.5), $A_{M1} = 0$ and we examine I_{M2} . Integrating I_{M2} by parts,

$$I_{M2} = \int_0^{\lambda_{M-1}} \exp(-\phi_M) \frac{\partial}{\partial \lambda_M} \left(\frac{1}{\phi'_M(\lambda_M)} \theta_M \right) d\lambda_M = A_{M2} + I_{M3} \quad (\text{A.6})$$

where

$$\begin{aligned} A_{M2} &= \frac{-e^{-\phi_M}}{\phi'_M} \frac{\partial}{\partial \lambda_M} \left(\frac{1}{\phi'_M} \theta_M \right) \Big|_0^{\lambda_{M-1}} = \frac{-e^{-\phi_M}}{\phi'_M} \left[-\frac{\phi''_M}{(\phi'_M)^2} \theta_M + \frac{1}{\phi'_M} \theta'_M \right] \Big|_0^{\lambda_{M-1}} \\ &= \frac{-e^{-\phi_M}}{(\phi'_M)^2} \theta'_M \Big|_0^{\lambda_{M-1}} \end{aligned} \quad (\text{A.7a})$$

$$I_{M3} = \int_0^{\lambda_{M-1}} e^{-\phi_M} \frac{\partial}{\partial \lambda_M} \left[\frac{1}{\phi'_M} \frac{\partial}{\partial \lambda_M} \left(\frac{1}{\phi'_M} \theta_M \right) \right] d\lambda_M \quad (\text{A.7b})$$

The last step of Eq. (A.7a) is obtained by observing that the first term is zero for both limits, since $\theta_M \Big|_0 = \theta_M \Big|_{\lambda_{M-1}} = 0$. Now,

$$\theta'_M \Big|_0^{\lambda_{M-1}} = -(N - M) \lambda_M^{(N-M-1)} \Big|_0 \prod_{i=1}^{M-1} \lambda_i^2 \quad (\text{A.8})$$

where this result is due to the fact that the other terms of the differential all vanish after the limits are substituted, i.e., $\theta'_M \Big|_0^{\lambda_{M-1}}$ is only non-zero in the limit $\lambda_M \rightarrow 0$.

Thus, using Eq. (A.8) in Eq. (A.7a), we have

$$A_{M2} = \left(\frac{2}{\rho}\right)^2 \exp\left(-\frac{M}{2}\right) (N-M) \left[\lambda_M^{(N-M-1)} \Big|_0 \prod_{i=1}^{M-1} \lambda_i^2 \right] + o(\rho^{-2}) \quad (\text{A.9})$$

which contains $(\rho^{-1})^2$. Similar to the case of A_{M1} , if $N-M-1=0$, then A_{M2} contains the lowest order of ρ^{-1} . Further integration w.r.t. $\lambda_1, \dots, \lambda_{M-1}$ will not change the order of ρ^{-1} . However, if $N-M-1 > 0$, then $A_{M2} = 0$ and we have to integrate I_{M3} in Eq. (A.7b) by parts, i.e.,

$$I_{M3} = A_{M3} + I_{M4}$$

where

$$A_{M3} = -\frac{e^{-\phi_M}}{\phi'_M} \frac{\partial}{\partial \lambda_M} \left[\frac{1}{\phi'_M} \frac{\partial}{\partial \lambda_M} \left(\frac{1}{\phi'_M} \theta_M \right) \right] \Big|_0^{\lambda_{M-1}} = \frac{e^{-\phi_M}}{(\phi'_M)^3} \theta''_M \Big|_{\lambda_{M-1}}^0 \quad (\text{A.10a})$$

$$I_{M4} = \int_0^{\lambda_{M-1}} e^{-\phi_M} \frac{\partial}{\partial \lambda_M} \left[\frac{1}{\phi'_M} \frac{\partial}{\partial \lambda_M} \left(\frac{1}{\phi'_M} \frac{\partial}{\partial \lambda_M} \left(\frac{1}{\phi'_M} \theta_M \right) \right) \right] d\lambda_M \quad (\text{A.10b})$$

where Eq. (A.10a) is obtained by observing that both θ_M and θ'_M are equal to zero at the upper and lower limits under the condition $N-M-1 > 0$. Now,

$$\theta''_M \Big|_{\lambda_{M-1}}^0 = \left[(N-M)(N-M-1) \lambda_M^{(N-M-2)} \Big|_0 \prod_{i=1}^{M-1} \lambda_i^2 \right] + \left[2 \lambda_{M-1}^{(N-M)} \prod_{i=1}^{M-2} (\lambda_i - \lambda_{M-1})^2 \right] \quad (\text{A.11})$$

Eq. (A.11) is the result after the limits have been substituted into the second derivative. Here, $\theta''_M \Big|_{\lambda_{M-1}}^0$ consists of two terms. If $N-M-2=0$, then the first term is independent of ρ . From Eq. (A.10a), the factor $(e^{-\phi_M(0)}/(\phi'_M(0))^3)$, which contains the factor $(\rho^{-1})^3$, is independent of λ_{M-1} or any other eigenvalues and the order of ρ^{-1} will not increase upon further integration w.r.t. $\lambda_{M-1}, \dots, \lambda_1$. Hence, this is the lowest order of ρ^{-1} . On the other hand, if $N-M-2 > 0$, the first term in Eq. (A.11) is zero. Now, we can obtain the lowest order term in two different ways:

- i) Consider I_{M4} of Eq. (A.10b) and perform integration by parts. Here, we focus on reducing the power index of λ_M^{N-M-2} to zero so that the integration has non-zero result at $\lambda_M = 0$. The order of ρ^{-1} will not increase upon further integration w.r.t. the other eigenvalues.
- ii) When the upper limit $\lambda_M = \lambda_{M-1}$ is substituted, the second term in Eq. (A.11) is non-zero. Putting this term back in Eq. (A.10a), together with the factor $[e^{-\phi_M}/(\phi'_M)^3] \Big|_{\lambda_M=\lambda_{M-1}}$ which is a function of ρ and λ_{M-1} , this whole quantity of A_{M3} has to be further integrated w.r.t. $\lambda_{M-1}, \dots, \lambda_1$, and therefore increasing the order of ρ^{-1} .

Thus, the problem of seeking the lowest order of ρ^{-1} has been reduced to the following questions. Is it the terms in Step i) for which, in integrating I_{M4} by parts, the index of $\lambda_M \Big|_0$ is reduced to zero after differentiation w.r.t. λ_M ? Or is it the term in Step ii) which yields the term involving the lowest order of ρ^{-1} after all the integrations w.r.t. $\lambda_{M-1}, \dots, \lambda_1$? We will examine the questions in parts c) and d) below.

- c) The order of ρ^{-1} by Step i): The first integration of I_{M1} in Eq. (A.5) yields $(\phi'_M)^{-1}$ together with θ_M which involves no reduction in the index. The second integration I_{M2} yields $(\phi'_M)^{-2}$ together with θ'_M and reduces the index by 1. Thus, to reduce the index of $\lambda_M \Big|_0$ from $(N-M)$ to zero, we need to differentiate the function θ_M $(N-M)$ times. This yields $(\phi'_M)^{-(N-M+1)}$, which contains the SNR term $\rho^{-(N-M+1)}$. The non-zero component of the integration is $[(N-M)!(2/\rho)^{N-M+1} e^{-\frac{M}{2}} \prod_{i=1}^{M-1} \lambda_i^2]$, which is independent of λ_M . Substituting this non-zero term into Eq. (A.1), we obtain the term in $\eta(\rho)$ containing the lowest

order of ρ^{-1} , namely

$$A_0 = (N - M)! \frac{\alpha(M, N) \exp\left(-\frac{M-1}{2}\right) \left(\frac{2}{\rho}\right)^{N-M+1}}{2} \int_0^\infty d\lambda_1 \int_0^{\lambda_1} d\lambda_2 \cdots \int_0^{\lambda_{M-2}} \exp\left(-\sum_{i=1}^{M-1} \lambda_i\right) \prod_{i=1}^{M-1} \lambda_i^{N-M+2} \prod_{i<j}^{M-1} (\lambda_i - \lambda_j)^2 d\lambda_{M-1} \quad (\text{A.12})$$

where $\alpha(M, N)$ was defined in Eq. (4.5). Observe that the integration in Eq. (A.12) will not increase the order of ρ^{-1} . Now, multiplying both the numerator and denominator of Eq. (A.12) by $\alpha(M - 1, N + 1)$, we can write

$$A_0 = (N - M)! \frac{\alpha(M, N) \exp\left(-\frac{M-1}{2}\right) \left(\frac{2}{\rho}\right)^{(N-M+1)}}{2\alpha(M - 1, N + 1)} \left\{ \alpha(M - 1, N + 1) \int_0^\infty d\lambda_1 \int_0^{\lambda_1} d\lambda_2 \cdots \int_0^{\lambda_{M-2}} \exp\left(-\sum_{i=1}^{M-1} \lambda_i\right) \prod_{i=1}^{M-1} \lambda_i^{N-M+2} \prod_{i<j}^{M-1} (\lambda_i - \lambda_j)^2 d\lambda_{M-1} \right\} \quad (\text{A.13a})$$

$$= (N - M)! \frac{\alpha(M, N) \exp\left(-\frac{M-1}{2}\right) 2^{N-M}}{\alpha(M - 1, N + 1)} \rho^{-(N-M+1)} \quad (\text{A.13b})$$

The last step in Eq. (A.13b) is obtained by observing that the integrand inside the braces of Eq. (A.13a) is the same PDF as Eq. (4.5), and thus its integration results in unity. From Eq. (A.13b), we observe that for high SNR the order of ρ^{-1} is given by

$$d_1 = N - M + 1 \quad (\text{A.14})$$

This is thus the lowest order of ρ^{-1} contained in the term obtained when the lower limit $\lambda_M = 0$ is substituted.

- d) The order of ρ^{-1} obtained by Step ii): Putting the second term of Eq. (A.11) into Eq. (A.10a) and then the result into Eq. (A.1), and taking the terms involving $\lambda_1, \dots, \lambda_{M-2}$ outside, the integral w.r.t. λ_{M-1} is

$$I_{(M-1)1} = 2 \int_0^{\lambda_{M-2}} \frac{e^{-\phi_{M-1}}}{(\psi_{M(M-1)}(\lambda_{M-1}))^3} \theta_{M-1} d\lambda_{M-1} \quad (\text{A.15})$$

where ϕ_{M-1} and θ_{M-1} have been defined in a general form in Eqs. (A.2b) and (A.2c), respectively. We note that the indices of the factors λ_{M-1} and $(\lambda_i - \lambda_{M-1})$ in θ_{M-1} have increased by one extra fold and by 2, respectively, in comparison to the corresponding indices in θ_M . Eq. (A.15) has a similar structure to Eq. (A.5), and hence integrating Eq. (A.15) by parts follows a similar pattern, i.e.,

$$I_{(M-1)1} = 2(A_{(M-1)1} + I_{(M-1)2}) \quad (\text{A.16})$$

where $A_{(M-1)1} = \frac{-\exp(-\phi_{M-1})}{(\psi_{M(M-1)}(\lambda_{M-1}))^3 \phi'_{M-1}} \theta_{M-1} \Big|_0^{\lambda_{M-2}}$ and $I_{(M-1)2} = \int_0^{\lambda_{M-2}} e^{-\phi_{M-1}} \left(\frac{\theta_{M-1}}{(\psi_{M(M-1)}(\lambda_{M-1}))^3 \phi'_{M-1}} \right)' d\lambda_{M-1}$. Due to the property of $\theta_{M-1} \Big|_0^{\lambda_{M-2}} = 0$, we only have to examine the effect of $I_{(M-1)2}$ on the index of ρ^{-1} . Again, applying integration by parts of Eq. (A.4) to $I_{(M-1)2}$, we have

$$I_{(M-1)2} = A_{(M-1)2} + I_{(M-1)3} \quad (\text{A.17})$$

where $A_{(M-1)2} = -\frac{e^{-\phi_{M-1}}}{\phi'_{M-1}} \left(\frac{1}{(\psi_{M(M-1)}(\lambda_{M-1}))^3 \phi'_{M-1}} \theta_{M-1} \right)' \Big|_0^{\lambda_{M-2}}$ and $I_{(M-1)3} = \int_0^{\lambda_{M-2}} e^{-\phi_{M-1}} \left(\frac{1}{\phi'_{M-1}} \left(\frac{1}{(\psi_{M(M-1)}(\lambda_{M-1}))^3 \phi'_{M-1}} \theta_{M-1} \right)' \right)' d\lambda_{M-1}$. The term $A_{(M-1)2}$ is zero since θ'_{M-1} contains the term $4\lambda_{M-1}^{2(N-M)} (\lambda_{M-2} - \lambda_{M-1})^3$ which yields zero at either limit of 0 or λ_{M-2} . A non-zero result is obtained from $I_{(M-1)3}$ at the limit of $\lambda_{M-1} = \lambda_{M-2}$ by having three more steps of integration by parts on $I_{(M-1)3}$, each involving the differentiation of $(\lambda_{M-2} - \lambda_{M-1})^3$ and resulting in the power index being reduced to zero. This non-zero term is given by

$$\frac{2 \cdot 4! e^{-\phi_{M-2}} \theta_{M-2}}{(\psi_{M(M-2)}(\lambda_{M-2}))^3 (\psi_{(M-1)(M-2)}(\lambda_{M-2}))^5} \quad (\text{A.18})$$

Further integration of Eq. (A.18) w.r.t. $\lambda_{M-2}, \dots, \lambda_1$ repeats the same procedure as in Eq. (A.15) and reveals similar structures, i.e., each time the factors in the denominator increases by $(\phi'_m)^{2(M-m+1)+1}$ and the subscript of θ reduces by 1. Continuing the procedure until the integration w.r.t. λ_1 , we have,

$$I_1 = \int_0^\infty \frac{\left(\prod_{i=1}^{M-1} (2i)! \right) e^{-\phi_1(\lambda_1)}}{(\psi_{M1}(\lambda_1))^3 (\psi_{(M-1)1}(\lambda_1))^5 \dots \psi_{21}(\lambda_1)^{2(M-1)+1}} \theta_1 d\lambda_1 \quad (\text{A.19})$$

where, from Eqs. (A.2b) and (A.2c), $\phi_1(\lambda_1) = \left(\frac{1}{2} + \left(\frac{\rho}{2M} + M\right)\lambda_1\right)$ and $\theta_1 = \lambda_1^{(N-M)M}$. We note that $e^{-\phi_1(\lambda_1)}\big|_{\lambda_1=\infty} = 0$, and the non-zero result of integrating I_1 comes only from the lower limit $\lambda_1 = 0$. Again, this term is arrived at by repeated integration by parts until the index of the term $\lambda_1^{(N-M)M}$ in θ_1 reaches 0. The resulting non-zero term with minimum steps of integration by parts is

$$A_1 = \frac{((N-M)M)! \prod_{i=1}^{M-1} (2i)! e^{-1/2}}{(\psi_{M1}(\lambda_1))^3 (\psi_{(M-1)1}(\lambda_1))^5 \dots \psi_{21}(\lambda_1)^{(2(M-1)+1)} \phi_1'(\lambda_1)^{((N-M)M+1)}} \bigg|_{\lambda_1=0} \quad (\text{A.20})$$

Now, we can evaluate the power index of ρ^{-1} in Eq. (A.20). We observe that each ψ or ϕ' contains ρ (unity power index). Thus, ignoring the constant coefficients, the power index of ρ^{-1} in A_1 is the sum of the exponents of the functions ψ and ϕ' in the denominator of Eq. (A.20), i.e.,

$$\sum_{i=1}^{M-1} (2i+1) + (N-M)M + 1 = MN \triangleq d_2 \quad (\text{A.21})$$

Eq. (A.21) provides the order of ρ^{-1} obtained by continuing the integration of $I_{(M-1)1}$ w.r.t. $\lambda_{M-1}, \dots, \lambda_1$.

- e) We now compare the two lowest orders d_1 and d_2 established respectively in Eqs. (A.14) and (A.21). We note that $d_2 - d_1 = MN - (N - M + 1) = (N + 1)(M - 1) \geq 0$ since $N, M > 0$. Therefore, d_1 is the lowest order of ρ^{-1} in the result of evaluating $\eta(\rho)$ of Eq. (A.1), and is therefore the diversity order for the single-block MIMO system. \square

Appendix B

Proof of Theorem 4.3

We can prove Theorem 4.3 by first deriving several consecutively achievable lower bounds on the error probability given in Eq. (4.23) until we reach a minimum independent of the design parameters. This is followed by examining the conditions required on the optimal \mathbf{F} for reaching each of the lower bounds.

B.0.0.1 First lower bound

Using the convexity property of the Q -function [61] in Eq. (4.23), we obtain the first lower bound on \mathcal{P}_{em} by employing Jensen's inequality [9], i.e.,

$$\begin{aligned}\mathcal{P}_{\text{em}} &= \mathbb{E}_{\mathcal{H}} \left\{ \frac{1}{2LMT} \sum_{k=1}^{2LMT} Q \left(\sqrt{\frac{1}{[\mathbf{V}]_{kk}} - 1} \right) \right\} \\ &\geq \mathbb{E} \left\{ Q \left(\sqrt{\frac{2LMT}{\sum_{k=1}^{2LMT} [\mathbf{V}]_{kk}} - 1} \right) \right\} \end{aligned} \quad (\text{B.1a})$$

$$= \mathbb{E} \left\{ Q \left(\sqrt{\frac{LMT}{\text{tr}(\mathbf{V}_{\text{se}})} - 1} \right) \right\} \quad (\text{B.1b})$$

where \mathbf{V}_{se} is the symbol error covariance matrix given by Eq. (4.19). Eq. (B.1b) holds due to $\text{tr}(\mathbf{V}) = \text{tr}(2\mathbf{V}_{\text{se}})$, as shown in Eq (4.22). Equality in (B.1a) holds iff

the diagonal elements of \mathbf{V} are all equal, i.e., all the MSE of the bits are equal. Since we are transmitting 4-QAM symbols having symmetry in both real and imaginary parts, this is equivalent to

$$[\mathbf{V}_{\text{se}}]_{ii} = [\mathbf{V}_{\text{se}}]_{jj}, \quad \forall i, j = 1, \dots, LMT \quad (\text{B.2})$$

B.0.0.2 Second lower bound

The following two lemmas [43] are provided here to facilitate the development of the second lower bound:

Lemma B.1 *For any square matrix $\mathbf{\Gamma}$ we have $\text{tr}\left((\mathbf{I} + \mathbf{\Gamma}\mathbf{\Gamma}^H)^{-1}\right) = \text{tr}\left((\mathbf{I} + \mathbf{\Gamma}^H\mathbf{\Gamma})^{-1}\right)$* ■

Lemma B.2 *For any nonsingular Hermitian symmetric positive semi-definite (PSD) matrix $\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_{11} & \mathbf{Z}_{12} \\ \mathbf{Z}_{12}^H & \mathbf{Z}_{22} \end{bmatrix}$ we have*

$$\text{tr}(\mathbf{Z}^{-1}) \geq \text{tr}(\mathbf{Z}_{11}^{-1}) + \text{tr}(\mathbf{Z}_{22}^{-1})$$

where equality holds iff $\mathbf{Z}_{12} = \mathbf{0}$, i.e., iff \mathbf{Z} is block diagonal. ■

Lemma B.2 can be repeatedly applied to extend to a nonsingular Hermitian symmetric PSD matrix partitioned into multiple submatrices.

Applying Lemma B.1 to the error covariance matrix \mathbf{V}_{se} in Eq. (4.19) yields

$$\text{tr}(\mathbf{V}_{\text{se}}) = \text{tr}\left(\left(\mathbf{I} + \frac{\rho}{M}\mathbf{\Xi}\mathbf{F}\mathbf{F}^H\mathbf{\Xi}^H\right)^{-1}\right) \quad (\text{B.3})$$

where we define $\mathbf{\Xi} \triangleq (\mathcal{H}^H\mathcal{H})^{\frac{1}{2}}$. From the definition of \mathcal{H} in Eq. (4.14), we can see that $\mathbf{\Xi}$ is an $LMT \times LMT$ block diagonal matrix. Each diagonal block $\tilde{\mathbf{\Xi}}_\ell \triangleq \left[\mathbf{I}_T \otimes (\mathbf{H}_\ell^H\mathbf{H}_\ell)^{\frac{1}{2}}\right]$, $\ell = 1, \dots, L$, is also block diagonal of dimension $MT \times MT$, containing T identical $M \times M$ sub-blocks of $\mathbf{\Xi}_\ell \triangleq (\mathbf{H}_\ell^H\mathbf{H}_\ell)^{\frac{1}{2}}$.

Now, writing $\mathbf{B} = \mathbf{F}\mathbf{F}^H$ in Eq. (B.3), we note that this $LMT \times LMT$ nonsingular Hermitian symmetric PSD matrix can be partitioned into L^2 blocks of $MT \times MT$ submatrices \mathbf{B}_{ij} , $i, j = 1, \dots, MT$. Denoting by $\mathbf{B}_{\ell\ell}$ the ℓ th $MT \times MT$ submatrix on the diagonal of \mathbf{B} and applying Lemma B.2, a lower bound for $\text{tr}(\mathbf{V}_{se})$ in Eq. (B.3) can be obtained, namely

$$\text{tr}(\mathbf{V}_{se}) \geq \sum_{\ell=1}^L \text{tr} \left(\left(\mathbf{I}_{MT} + \frac{\rho}{M} \tilde{\Xi}_{\ell} \mathbf{B}_{\ell\ell} \tilde{\Xi}_{\ell}^H \right)^{-1} \right) \quad (\text{B.4})$$

Equality in Eq. (B.4) holds iff $\mathbf{B}_{ij} = \mathbf{0}$, for $i, j = 1, \dots, L$, $i \neq j$. But each $\mathbf{B}_{\ell\ell}$ can be further partitioned into T^2 submatrices $\mathbf{B}_{\ell\ell}^{ij}$, $i, j = 1, \dots, T$, of dimension $M \times M$. Denoting by $\mathbf{B}_{\ell\ell}^{tt}$ the t th $M \times M$ submatrix on the diagonal of $\mathbf{B}_{\ell\ell}$ and applying Lemma B.2 to each of the bracketed term in Eq. (B.4) results in

$$\text{tr}(\mathbf{V}_{se}) \geq \sum_{\ell=1}^L \sum_{t=1}^T \text{tr} \left(\left(\mathbf{I}_M + \frac{\rho}{M} \Xi_{\ell} \mathbf{B}_{\ell\ell}^{tt} \Xi_{\ell}^H \right)^{-1} \right) \quad (\text{B.5})$$

where equality in Eq. (B.5) holds iff

$$\mathbf{B}_{\ell\ell}^{ij} = \mathbf{0} \quad \text{for } i, j = 1, \dots, T, \quad i \neq j, \quad (\text{B.6})$$

i.e., $\mathbf{B} = \mathbf{F}\mathbf{F}^H$ is block diagonal having non-zero $M \times M$ diagonal submatrices.

Furthermore, for any positive definite matrix \mathbf{Z} , $\text{tr}(\mathbf{Z}^{-1})$ is convex with respect to \mathbf{Z} [9]. Thus, applying Jensen's inequality to this convex function in the inner sum of Eq. (B.5) and employing Lemma B.1 leads to

$$\frac{1}{T} \sum_{t=1}^T \text{tr} \left(\left(\mathbf{I}_M + \frac{\rho}{M} \Xi_{\ell} \mathbf{B}_{\ell\ell}^{tt} \Xi_{\ell}^H \right)^{-1} \right) \geq \text{tr} \left(\left(\mathbf{I}_M + \frac{\rho}{M} \Xi_{\ell} \bar{\mathbf{B}}_{\ell\ell} \Xi_{\ell}^H \right)^{-1} \right) \quad (\text{B.7})$$

where $\bar{\mathbf{B}}_{\ell\ell} \triangleq \frac{1}{T} \sum_{t=1}^T \mathbf{B}_{\ell\ell}^{tt}$. Equality in (B.7) holds iff all $\mathbf{B}_{\ell\ell}^{tt}$ are equal, i.e.,

$$\mathbf{B}_{\ell\ell}^{tt} = \bar{\mathbf{B}}_{\ell\ell} \quad \text{for } t = 1, 2, \dots, T \quad (\text{B.8})$$

Combining Eq. (B.5) with Eq. (B.7) results in

$$\text{tr}(\mathbf{V}_{\text{se}}) \geq T \sum_{\ell=1}^L \text{tr} \left(\left(\mathbf{I}_M + \frac{\rho}{M} \mathbf{\Xi}_\ell \bar{\mathbf{B}}_{\ell\ell} \mathbf{\Xi}_\ell^H \right)^{-1} \right) \quad (\text{B.9})$$

with equality holding iff Eqs. (B.6) and (B.8) are satisfied. Since the function $f(x) = Q(\sqrt{x^{-1}-1})$ is monotonically increasing with x , substituting Eq. (B.9) in Eq. (B.1b) establishes the second lower bound for the average asymptotic BER \mathcal{P}_{em} such that,

$$\mathcal{P}_{\text{em}} \geq \mathbb{E} \left\{ Q \left(\sqrt{\frac{ML}{\sum_{\ell=1}^L \text{tr} \left(\left(\mathbf{I}_M + \frac{\rho}{M} \mathbf{\Xi}_\ell \bar{\mathbf{B}}_{\ell\ell} \mathbf{\Xi}_\ell^H \right)^{-1} \right)} - 1} \right) \right\} \quad (\text{B.10})$$

Again, equality in (B.10) holds iff the conditions in Eqs. (B.6) and (B.8) are met simultaneously.

B.0.0.3 Final lower bound

Eq. (B.10) still depends on the design variable \mathbf{F} and therefore needs to be further minimized to obtain a constant achievable lower bound. We observe that the elements of the channel matrix \mathbf{H}_ℓ are of IID distribution, hence the stochastic properties of \mathbf{H}_ℓ remain unchanged by the multiplication of a unitary matrix. By pre- and post-multiplying permutation matrices to the eigenvalue matrix of $\bar{\mathbf{B}}_{\ell\ell}$ in Eq. (B.10) and following similar procedures of averaging over all $M!$ permutations, as in [92, 61], we arrive at

$$\mathcal{P}_{\text{em}} \geq \mathbb{E} \left\{ Q \left(\sqrt{\frac{ML}{\sum_{\ell=1}^L \text{tr} \left(\left(\mathbf{I}_M + \frac{\rho}{M} \mathbf{H}_\ell^H \mathbf{H}_\ell \right)^{-1} \right)} - 1} \right) \right\} \quad (\text{B.11})$$

with equality holding iff

$$\bar{\mathbf{B}}_{\ell\ell} = \mathbf{I}, \quad \ell = 1, \dots, L \quad (\text{B.12})$$

The lower bound in Eq. (B.11) is independent of the design matrix \mathbf{F} and is, therefore, a genuine lower bound on the value that can be achieved by a designer, in the sense that it can be computed prior to the design.

B.0.0.4 Optimal Linear STBC

The minimum BER given in Eq. (B.11) can be achieved iff the four conditions in Eqs. (B.2), (B.6), (B.8) and (B.12) are satisfied simultaneously. Eqs. (B.6), (B.8) and (B.12) jointly imply that $\mathbf{B} = \mathbf{F}\mathbf{F}^H = \mathbf{I}$, i.e., \mathbf{F} must be unitary. Eq. (B.2) requires all the diagonal elements of \mathbf{V}_{se} to be equal, i.e.,

$$[\mathbf{V}_{\text{se}}]_{ii} = \frac{1}{LMT} \sum_{i=1}^{LMT} [\mathbf{V}_{\text{se}}]_{ii} = \frac{1}{LMT} \text{tr}(\mathbf{V}_{\text{se}}) \quad (\text{B.13})$$

and this must hold for any channel realization.

Now, Eq. (4.19) can be re-written as $\mathbf{V}_{\text{se}} = \left(\mathbf{F}^H \mathbf{F} + \frac{\rho}{M} \mathbf{F}^H \mathcal{H}^H \mathcal{H} \mathbf{F} \right)^{-1} = \mathbf{F}^H \left(\mathbf{I} + \frac{\rho}{M} \mathcal{H}^H \mathcal{H} \right)^{-1} \mathbf{F}$, where we have used the condition that \mathbf{F} must be unitary. Thus, the i th diagonal element of \mathbf{V}_{se} is

$$\begin{aligned} [\mathbf{V}_{\text{se}}]_{ii} &= \mathbf{f}_i^H \left(\mathbf{I}_{LMT} + \frac{\rho}{M} \mathcal{H}^H \mathcal{H} \right)^{-1} \mathbf{f}_i = \text{tr} \left(\sum_{\ell=1}^L \mathbf{C}_{i\ell}^H \left(\mathbf{I}_M + \frac{\rho}{M} \mathbf{H}_\ell^H \mathbf{H}_\ell \right)^{-1} \mathbf{C}_{i\ell} \right) \\ &= \text{tr} \left(\sum_{\ell=1}^L \left(\mathbf{I}_M + \frac{\rho}{M} \mathbf{H}_\ell^H \mathbf{H}_\ell \right)^{-1} \mathbf{C}_{i\ell} \mathbf{C}_{i\ell}^H \right) \end{aligned} \quad (\text{B.14})$$

where \mathbf{f}_i is the i th column of the coding matrix \mathbf{F} and $\mathbf{C}_{i\ell}$ is the coding submatrix in the ℓ th block for the i th symbol. The second step in Eq. (B.14) is possible because of the relationship between \mathbf{f}_i and $\mathbf{C}_{i\ell}$ in Eq. (4.16). On the other hand, the right hand side of Eq. (B.13) can also be written as

$$\text{tr}(\mathbf{V}_{\text{se}}) = \text{tr} \left(\mathbf{I}_{LMT} + \frac{\rho}{M} \mathcal{H}^H \mathcal{H} \right)^{-1} = T \text{tr} \left(\sum_{\ell=1}^L \mathbf{I}_M + \frac{\rho}{M} \mathbf{H}_\ell^H \mathbf{H}_\ell \right)^{-1} \quad (\text{B.15})$$

Substituting Eqs. (B.14) and (B.15) respectively into the left and right sides of Eq. (B.13), the condition becomes

$$\text{tr} \left(\sum_{\ell=1}^L \left(\mathbf{I}_M + \frac{\rho}{M} \mathbf{H}_\ell^H \mathbf{H}_\ell \right)^{-1} \mathbf{C}_{i\ell} \mathbf{C}_{i\ell}^H \right) = \frac{1}{ML} \text{tr} \left(\sum_{\ell=1}^L \left(\mathbf{I}_M + \frac{\rho}{M} \mathbf{H}_\ell^H \mathbf{H}_\ell \right)^{-1} \right) \quad (\text{B.16})$$

Thus, for any channel realization \mathbf{H}_ℓ , the condition in Eq. (B.2) holds iff $\mathbf{C}_{i\ell} \mathbf{C}_{i\ell}^H = \frac{1}{ML} \mathbf{I}_M$. \square

Appendix C

Proof of Theorem 6.3

First we note that

$$\begin{aligned}
 \text{tr}(\mathbf{A}_{r,t}\mathbf{A}_{r',t'}^H) &= \frac{MT}{K} \text{tr}(\mathbf{P}_M^{r-1}\mathbf{\Pi}_t\mathbf{\Pi}_{t'}^H\mathbf{P}_M^{-r'+1}) \\
 &= \frac{MT}{K} \text{tr}(\mathbf{P}_M^{r-r'}\mathbf{\Pi}_t\mathbf{\Pi}_{t'}^H) \\
 &= \frac{MT}{K} \sum_{k=1}^M [\mathbf{P}_M^{r-r'}\mathbf{\Pi}_t\mathbf{\Pi}_{t'}^H]_{kk} \\
 &= \frac{MT}{K} \sum_{k=1}^M [\mathbf{P}_M^{r-r'}]_{kk} [\mathbf{\Pi}_t\mathbf{\Pi}_{t'}^H]_{kk}
 \end{aligned} \tag{C.1}$$

where in the last step of Eq. (C.1) we have used the fact that $\mathbf{\Pi}_t\mathbf{\Pi}_{t'}^H$ is a *diagonal* matrix. Similarly, we have

$$\text{tr}(\mathbf{B}_{r',t'}\mathbf{B}_{r,t}^H) = \frac{MT}{K} \sum_{k=1}^M [\mathbf{P}_M^{r-r'}]_{kk} [\mathbf{\Delta}_t\mathbf{\Delta}_{t'}^H]_{kk} \tag{C.2}$$

Combining Lemma 6.1 and Eq. (C.2) with Eq. (C.1) yields

$$\text{tr}(\mathbf{A}_{r,t}\mathbf{A}_{r',t'}^H + \mathbf{B}_{r',t'}\mathbf{B}_{r,t}^H) = \frac{MT}{K} \delta(r-r')\delta(t-t')$$

where $\delta(n-m)$ denotes the unit impulse function, and we have used the fact that each matrix \mathbf{V}_r is V-structured. Similarly, we can obtain

$$\text{tr}(\mathbf{A}_{r,t} \mathbf{B}_{r',t'}^H + \mathbf{B}_{r',t'} \mathbf{A}_{r,t}^H) = 0$$

Appendix D

Proof of Lemma 6.5

First we prove the sufficient condition. Suppose that $L = L_1 \prod_{k=1}^r p_k^{\beta_k}$, where $\gcd(p_k, L_1) = 1$ for $k = 1, 2, \dots, r$. Then, in this case, employing the multiplicative property of the Euler function yields $\varphi(P) = \varphi(L_1) \prod_{k=1}^r \varphi(p_k^{\alpha_k + \beta_k}) = \left(\varphi(L_1) \prod_{k=1}^t p_k^{\beta_k - 1} (p_k - 1) \right) \left(\prod_{k=1}^r p_k^{\alpha_k} \right) \varphi(L)M$, which gives the proof of the sufficient condition.

Now we prove the necessary condition. Suppose that $\varphi(P) = \varphi(L)M$. Then, we claim that L and M are not coprime. Otherwise, if L and M were coprime, then, using Result 1 would result in $\varphi(P) = \varphi(L)\varphi(M)$. By the definition of the Euler function, we know that $\varphi(M) < M$ if $M > 1$. This would lead to $\varphi(P) = \varphi(L)\varphi(M) < \varphi(L)M$ for $M > 1$, which would contradict with the assumption. Therefore, L and M are not coprime. Let p_1, p_2, \dots, p_t ($t \geq 1$) be the common prime divisors of L and M ; i.e., L and M can be decomposed as $L = \prod_{k=1}^t p_k^{\gamma_k} L_1$ and $M = \prod_{k=1}^t p_k^{\alpha_k} M_1$, where $t \geq 1$, $\alpha_k, \gamma_k \geq 1$, each p_k is coprime to both L_1 and M_1 , and $\gcd(L_1, M_1) = 1$. Hence,

$\varphi(P) = \varphi(L_1)\varphi(M_1) \prod_{k=1}^t \varphi(p_k^{\alpha_k + \gamma_k})$. Since $\varphi(p_k^{\alpha_k + \gamma_k}) = p_k^{\alpha_k + \gamma_k - 1}(p_k - 1)$,

$$\begin{aligned}
 \varphi(P) &= \left(\varphi(L_1) \prod_{k=1}^t p_k^{\gamma_k - 1}(p_k - 1) \right) \left(\prod_{k=1}^t p_k^{\alpha_k} \varphi(M_1) \right) \\
 &= \varphi(L) \left(\prod_{k=1}^t p_k^{\alpha_k} \varphi(M_1) \right) \\
 &\leq \varphi(L) \left(\prod_{k=1}^t p_k^{\alpha_k} M_1 \right) = \varphi(L)M
 \end{aligned} \tag{D.1}$$

where the equality in Eq. (D.1) holds if and only if $M_1 = 1$. Therefore, in this case,

$L = L_1 \prod_{k=1}^t p_k^{\beta_k}$. This completes the proof of Lemma 6.5.

Appendix E

Proof of Theorem 6.4

By Theorem 6.3, the LD code constructed in Theorem 6.4 is trace-orthonormal. In the following, we prove that such resulting signal coding matrix $\mathbf{X}(\mathbf{s})$ provides full diversity. To this end, first generate the DAST code. Let

$$\boldsymbol{\sigma}_{r,k} = \begin{pmatrix} \mathcal{R}_0[1 : M_1, :] \\ \mathbf{0}_{M_2 \times M} \end{pmatrix} \mathbf{s}_{r,k} + \begin{pmatrix} \mathbf{0}_{M_1 \times M} \\ \mathcal{R}_0^*[M_1 + 1 : M, :] \end{pmatrix} \mathbf{s}_{r,k}^*$$

where $\mathbf{s}_{r,k} = [s_{(r-1)T+(k-1)M+1}, \dots, s_{(r-1)T+(k-1)M+M}]^T$. Then, the DAST code is generated by the matrix $\mathcal{D}(\mathbf{s}_{r,k}) \triangleq \text{diag}(\boldsymbol{\sigma}_{r,k})$. By circularly placing stream data of each DAST code, we form the following matrix

$$\mathcal{X}(\mathbf{s}_k) = \sum_{r=1}^R \mathbf{P}^{r-1} \mathcal{D}(\mathbf{s}_{r,k}) \quad \text{for } k = 1, 2, \dots, \tilde{L} \quad (\text{E.1})$$

where $\mathbf{s}_k = [\mathbf{s}_{1,k}^T, \mathbf{s}_{2,k}^T, \dots, \mathbf{s}_{M,k}^T]^T$. A signal matrix $\mathbf{X}(\mathbf{s})$ can be built by Theorem 6.3

$$\mathbf{X}(\mathbf{s}) = [\mathbf{X}_1(\mathbf{s}), \mathbf{X}_2(\mathbf{s}), \dots, \mathbf{X}_{\tilde{L}}(\mathbf{s})] \quad (\text{E.2})$$

where $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \dots, \mathbf{s}_{\tilde{L}}^T]^T = [s_1, s_2, \dots, s_{TM}]^T$ and the k th sub-signal matrix $\mathbf{X}_k(\mathbf{s})$ is given by

$$\mathbf{X}_\ell(\mathbf{s}) = \sum_{k=1}^{\tilde{L}} u_{k,\ell} \mathcal{X}(\mathbf{s}_k) \quad \text{for } \ell = 1, 2, \dots, \tilde{L} \quad (\text{E.3})$$

We have from Eq. (E.2) and Eq. (E.3) that

$$\mathbf{X}(\mathbf{e})\mathbf{X}^H(\mathbf{e}) = \sum_{\ell=1}^{\tilde{L}} \mathbf{X}_\ell(\mathbf{e})\mathbf{X}_\ell^H(\mathbf{e}) = \sum_{k=1}^{\tilde{L}} \mathcal{X}(\mathbf{e}_k)\mathcal{X}^H(\mathbf{e}_k) \quad (\text{E.4})$$

where $\mathbf{e}_k = \mathbf{s}_k - \mathbf{s}'_k$ and $\mathbf{e} = [\mathbf{e}_1^T, \mathbf{e}_2^T, \dots, \mathbf{e}_{\tilde{L}}^T]^T$. This shows that, if we construct the V-structured matrix using Construction 1 and then use it to obtain the signal matrix $\mathbf{X}(\mathbf{s})$ through Theorem 6.3, then the autocorrelation matrix $\mathbf{X}(\mathbf{e})\mathbf{X}^H(\mathbf{e})$ does not depend on the choice of the unitary matrix \mathbf{U} . Now it suffices to prove that for any point \mathbf{e} there exists at least one k such that $\mathcal{X}(\mathbf{e}_k)$ is of full rank. Since $\mathbf{e} \neq \mathbf{0}$, there exists at least one k such that $\mathbf{e}_k \neq \mathbf{0}$ and as a result, there exists the maximum \bar{r} such that $\mathbf{e}_{\bar{r},k} \neq \mathbf{0}$. In other words, $\mathbf{e}_{r,k} = \mathbf{0}$ for $\bar{r} \leq r \leq R$, but $\mathbf{e}_{\bar{r},k} \neq \mathbf{0}$, where $\mathbf{e}_{r,k} = \mathbf{s}_{r,k} - \mathbf{s}'_{r,k}$. Therefore, $\det(\mathcal{D}(\mathbf{e}_{\bar{r},k})) \neq 0$. Now following the proof in [22, 64], where the Diophantine numbers are chosen as $\psi_1 = 1, \psi_2 = \zeta_{LRM^2} \cdots, \psi_{R-1} = \zeta_{LRM^2}^{R-1}$, and expanding the determinant of $\mathcal{X}(\mathbf{e}_k)$ according to the power of ζ_{LRM} , we have that $\det(\mathcal{X}(\mathbf{e}_k)) = \sum_{\ell=1}^{\bar{r}} c_\ell \zeta_{LRM}^{\ell-1}$, where $c_\ell \in \mathbb{Z}[\zeta_{LM}]$ for $\ell = 1, 2, \dots, \bar{r}$. By Corollary 6.1, $c_{\bar{r}} = \det(\mathcal{D}(\mathbf{s}_{\bar{r},k})) \neq 0$. In addition, by Lemma 6.4 and the properties of Euler functions, we know that $1, \zeta_{LRM}, \zeta_{LRM}^2, \dots, \zeta_{LRM}^{R-1}$ is linearly independent over $\mathbb{Z}[\zeta_{LM}]$ and as a result, we claim that the determinant of $\mathcal{X}(\mathbf{e}_k)$ is not zero. Therefore, using Lemma 6.6, we conclude from Eq. (E.4) that $\det(\mathbf{X}(\mathbf{e})\mathbf{X}^H(\mathbf{e})) = \det\left(\sum_{k=1}^{\tilde{L}} \mathcal{X}(\mathbf{e}_k)\mathcal{X}^H(\mathbf{e}_k)\right) \geq \sum_{k=1}^{\tilde{L}} \det(\mathcal{X}(\mathbf{e}_k)\mathcal{X}^H(\mathbf{e}_k)) > 0$ for any $\mathbf{e} \in \mathbb{Z}^K[\zeta_L] \setminus \{0\}$. This completes the proof of Theorem 6.4.

Bibliography

- [1] S. M. Alamouti, “A simple transmitter diversity scheme for wireless communications,” *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.
- [2] E. K. S. Au, S. Jin, M. R. McKay, X. G. W. H. Mow, and I. B. Collings, “Analytical performance of MIMO-SVD systems in Ricean fading channels with channel estimation error and feedback delay,” *IEEE Trans. Wireless Communi.*, vol. 7, pp. 1315–1325, Apr. 2008.
- [3] S. Barbarossa and A. Fasano, “Trace-orthogonal space-time coding for multiuser systems,” in *Proc. IEEE Conf. Acoustics, Speech, and Signal Processing Conference*, Philadelphia, PA, USA, March 2005.
- [4] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communications*, 3rd ed. Kluwer Academic, 2004.
- [5] J. C. Belfiore, G. R. Rekaya, and E. Viterbo, “The Golden code: A 2×2 full rate space-time code with non-vanishing determinants,” in *Proceedings IEEE International Symposium on Information Theory*, Chicago, June 2004.
- [6] K. Boulle and J.-C. Belfiore, “Modulation schemes designed for the Rayleigh fading channel,” in *Proc. CISS*, Princeton, NJ, USA, Mar. 1992.

- [7] J. Boutros and E. Viterbo, "Signal space diversity: A power-and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1453–1467, July 1998.
- [8] J. Boutros, E. Viterbo, C. Rastello, and J. C. Belfiore, "Good lattice constellations for both Rayleigh fading and Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 502–518, Mar. 1996.
- [9] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [10] K. Cho and D. Yoon, "On the general BER expression of one and two dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, pp. 1074–1080, July 2002.
- [11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley and Sons, 1991.
- [12] J. W. Craig, "A new, simple, and exact result for calculating the probability of error for two-dimensional signal constellations," in *Proc. IEEE Milit. Commun. Conf.*, Mclean, VA, Oct. 1991, pp. 571–575.
- [13] M. G. D.-S. Shiu, G.J. Foschini and J. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Trans. Comm.*, vol. 48, pp. 502–513, March 2000.
- [14] M. O. Damen, H. E. Gamal, and N. C. Beaulieu, "Systematic construction of full diversity algebraic constellations," *IEEE Trans. Inform. Theory*, vol. 49, pp. 3344–3340, Dec 2003.

- [15] —, “Linear threaded algebraic space-time constellation,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 2372–2388, Oct. 2003.
- [16] M. O. Damen, K. A. Meraim, and J. C. Belfiore, “Diagonal algebraic space-time block codes,” *IEEE Trans. Inform. Theory*, vol. 48, pp. 628–636, Mar. 2002.
- [17] M. O. Damen, A. Tewfik, and J. C. Belfiore, “A construction of a space-time code based on number theory,” *IEEE Trans. Inf. Theory*, vol. 48, pp. 753–760, Mar. 2002.
- [18] D. N. Dao and C. Tellambura, “Optimal rotations for quasi-orthogonal STBC with two-dimensional constellations,” in *Proceedings IEEE Global Communications Conference*, St. Louis, U.S.A., Nov.-Dec. 2005, pp. 1290–1294.
- [19] P. Dayal and M. K. Varanasi, “An optimal two transmit antenna space-time code and its stacked extensions,” in *Asilomar Conf. on Signals, Systems and Computers*, Monterey, CA, Nov. 2003.
- [20] S. N. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. R. Calderbank, “Great expectations: The value of spatial diversity in wireless networks,” *IEEE Proceedings*, vol. 92, pp. 219–270, Feb. 2004.
- [21] P. Duhamel and M. Vetterli, “Fast Fourier transforms: A tutorial review and a state of the art,” *Proc. Signal Processing*, vol. 19, pp. 259–299, 1990.
- [22] H. El Gamal and M. O. Damen, “Universal space-time coding,” *IEEE Trans. Inf. Theory*, vol. 49, pp. 1097–1119, May 2003.
- [23] H. El Gamal and A. R. Hammons, “On the design of algebraic space-time codes for MIMO block fading channels,” *IEEE Trans. Inf. Theory*, vol. 49, pp. 151–163, Jan. 2003.

- [24] P. Elia, K. Kumar, S. Pawar, P. V. Kumar, and H. Lu, "Explicit space-time codes achieving the diversity-multiplexing gain tradeoff," *IEEE Trans. Inform. Theory*, vol. 52, pp. 3869–3884, Sept. 2006.
- [25] C. F. Mecklenbräuker and M. Rupp, "Generalized Alamouti codes for trading quality of service against data rate in MIMO UMTS," *EURASIP Journal Applied Signal Processing*, pp. 662–675, 2004.
- [26] G. D. Forney and G. U. Ungerboeck, "Modulation and coding for linear Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2384–2415, May 1998.
- [27] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Comm.*, vol. 6, pp. 311–335, Mar. 1998.
- [28] H. E. Gamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity–multiplexing tradeoff of MIMO channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 968–985, June 2004.
- [29] G. Ganesan and P. Stoica, "Space-time block codes: a maximum SNR approach," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1650–1656, May. 2001.
- [30] A. V. Geramita and J. Seberry, "Orthogonal design, quadratic forms and hadamard matrices," in *Lecture Notes in Pure and Applied Mathematics*, New York: Marcel Dekker INC, 1979.
- [31] X. Giraud and J. C. Belfiore, "Constellations matched to the Rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 42, pp. 106–115, Jan. 1996.

- [32] X. Giraud, E. Boutillon, and J. C. Belfiore, "Algebraic tools to build modulation schemes for fading channels," *IEEE Trans. Inform. Theory*, vol. 43, pp. 938–952, May 1997.
- [33] R. H. Gohary and T. N. Davidson, "Design of linear dispersion codes: Some asymptotic guidelines and their implementation," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2892–2906, Nov. 2005.
- [34] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. The John Hopkins University Press, 1996.
- [35] D. Gore, R. W. Heath Jr., and A. Paulraj, "On performance of the zero forcing receivers in presence of transmit correlation," in *Proc. Int. Symp. Inf. Theory*, Lausanne, Switzerland, June-July 2002.
- [36] D. Gore, S. Sandhu, and A. Paulraj, "Delay diversity code for frequency selective channels," *Electronics Letters*, vol. 37, pp. 1230–1231, Sept. 2001.
- [37] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1804–1824, July 2002.
- [38] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs: Prentice Hall, 1991.
- [39] R. W. Heath and A. J. Paulraj, "Linear dispersion codes for MIMO systems based on frame theory," *IEEE Trans Signal Processing*, vol. 50, pp. 2429–2441, Oct. 2002.
- [40] T. Hehn, R. Schober, and W. H. Gerstacker, "Optimal delay diversity for frequency-selective fading channels," *IEEE Trans. commun.*, vol. 4, pp. 2289–2298, Sept. 2005.

- [41] J. Hiltunen, C. Hollanti, and J. Lahtonen, "Dense full-diversity matrix lattices for four transmit antenna MISO channel," in *Proceedings IEEE International Symposium on Information Theory*, Adelaide, Australia, Sept. 2005, pp. 1290–1294.
- [42] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communication in Rayleigh flat-fading," *IEEE Trans. Inform. Theory*, vol. 46, pp. 543–564, Mar. 2000.
- [43] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.
- [44] S. K. Jayaweera and H. V. Poor, "On the capacity of multiple-antenna systems in Rician fading," *IEEE Trans. wireless communi.*, vol. 4, pp. 1102–1111, May 2005.
- [45] S. Jin, X. Gao, and X. You, "On the ergodic capacity of rank-1 Ricean-fading MIMO channels," *IEEE Trans. Inf. Theory*, vol. 4, pp. 502–517, Feb. 2007.
- [46] I. Kammoun and J. C. Belfiore, "A new family of Grassmann space-time codes for non-coherent MIMO systems," *IEEE Commun. Letters*, to appear.
- [47] S. Karmakar and B. Rajan, "Multi-group decodable STBCs from clifford algebras," in *Proceedings 2006 IEEE Information Theory Workshop*, Chengdu, China, Oct. 2006, pp. 448–452.
- [48] T. Kiran and B. S. Rajan, "STBC-schemes with nonvanishing determinant for certain number of transmit antennas," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2984–2992, Aug. 2005.

- [49] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*. Orlando: Academic Press, 1985.
- [50] N. N. Lebedev, *Special Functions and Their Applications*. Prentice-Hall, 1965.
- [51] X.-B. Liang, "Orthogonal designs with maximal rates," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2468–2503, Oct. 2003.
- [52] X.-B. Liang and X.-G. Xia, "Nonexistence of rate one space-time blocks from generalized complex linear processing orthogonal designs for more than two transmit antennas," in *International Symposium on Inform. Theory*, Washington DC, June. 2001.
- [53] J. Liu, J.-K. Zhang, and K. M. Wong, "Optimal linear code design in improving BER for MIMO systems with MMSE receivers," in *Proc. ICSP'04*, Beijing, China, Sept. 2004.
- [54] —, "Design of optimal orthogonal linear codes in MIMO systems for MMSE receiver," in *Int. Conf. Acoust., Speech, Signal Process.*, Montreal, Canada, May 2004.
- [55] J. Liu, *Linear STBC design for MIMO systems with MMSE receivers*. A thesis for M.A.Sc level, McMaster University, 2004, <http://grads.ece.mcmaster.ca/~jliu>.
- [56] J. Liu, T. N. Davidson, and K. M. Wong, "Diversity analysis and design of space-time multi-block codes for MIMO systems equipped with linear MMSE receivers," *submitted to IEEE Trans. Inform. Theory*, under second round of review.

- [57] ———, “Design of space-time multi-block codes for MIMO systems with linear MMSE receivers,” in *Proc. ISIT2006*, Seattle, USA, July 2006.
- [58] J. Liu and K. M. Wong, “Combined receiver for multi-block transmission,” in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008)*, Las Vegas, USA, Mar.-April 2008.
- [59] ———, “Optimum MIMO precoder designs for Ricean fading channels,” in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008)*, Las Vegas, USA, Mar.-April 2008.
- [60] J. Liu, J.-K. Zhang, and K. M. Wong, “Properties and applications of the Toeplitz space-time block codes in MISO transmission systems,” *to appear, IEEE Trans. Inform. Theory*.
- [61] ———, “On the design of minimum BER linear space-time block codes for MIMO systems with MMSE receivers,” *IEEE Trans. Signal Processing*, vol. 54, no. 8, pp. 3147–3158, Aug. 2006.
- [62] H.-F. Lu, “Constructions of multiblock space-time coding schemes that achieve the diversity-multiplexing tradeoff,” *IEEE Trans. Inf. Theory*, vol. 54, pp. 3790–3796, Aug. 2008.
- [63] K. Lu, S. Fu, and X.-G. Xia, “Closed form designs of complex orthogonal space-time block codes of rates $k + 1/2k$ for $2k - 1$ or $2k$ transmit antennas,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 4340–4347, Dec. 2005.
- [64] X. Ma and G. B. Giannakis, “Full-diversity full rate complex-field space-time coding,” *IEEE Trans. Sig. Proc.*, vol. 51, pp. 2917–2930, Nov. 2003.

- [65] X.-L. Ma and G. Giannakis, "Complex field coded MIMO systems: performance, rate, and trade-offs," *Wireless communications and mobile computing*, vol. 2, pp. 693–717, 2002.
- [66] D. A. Marcus, *Number fields*. New York: Springer-Verlag, 1977.
- [67] M. R. McKay and I. B. Collings, "General capacity bounds for spatially correlated Rician MIMO channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3121–3145, Sept. 2005.
- [68] ———, "Improved general lower bound for spatially-correlated Rician MIMO capacity," *IEEE Communi. Letter*, vol. 10, pp. 162–164, Mar. 2006.
- [69] ———, "On the capacity of frequency-flat and frequency-selective Rician MIMO channels with single-ended correlation," *IEEE Trans. Wireless Communi.*, vol. 5, pp. 2038–2043, Aug. 2006.
- [70] M. R. McKay, A. Zanella, I. B. Collings, and M. Chiani, "Error probability and sinr analysis of optimum combining in rician fading," *to appear, IEEE Trans. wireless communi.*, http://ihome.ust.hk/~eemckay/TCom_OC_Rician_Fading_Accepted.pdf.
- [71] P. Morandi, *Field and Galois Theory*. New York: Springer, 1996.
- [72] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, 1982.
- [73] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space time block codes," *IEEE Trans. Inform. Theory*, vol. 52, pp. 3885–3902, Sept. 2006.
- [74] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. McGraw-Hill, 2002.

- [75] R. F. Pawula, S. O. Rice, and J. H. Roberts, "Distribution of the phase angle between two vectors perturbed by Gaussian noise," *IEEE Trans. Commun.*, vol. 30, pp. 1828–1841, Aug 1982.
- [76] V. Poor, *An Introduction to Signal Detection and Estimation (2nd ed.)*. New York: Springer-Verlag, 1989.
- [77] J. G. Proakis, *Digital Communications*, 4th ed. Mc-Graw Hill, 2001.
- [78] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Eigenvalues and condition numbers of complex random matrices," *SIAM J. Matrix Analysis and Applications*, vol. 26, pp. 441–456, 2004.
- [79] G. Rekaya, J. C. Belfiore, and E. Viterbo, "Algebraic 3×3 , 4×4 and 6×6 space-time codes with non-vanishing determinants," in *ISITA*, Parma, Italy, Oct. 2004.
- [80] S. Sandhu and A. J. Paulraj, "Space-time block coding: A capacity perspective," *IEEE Commun. Letters*, vol. 4, pp. 384–386, Dec. 2000.
- [81] —, "Unified design of linear space-time block codes," in *Proc. Global Telecommunication Conference*, vol. 2, Nov. 2001, pp. 1073–1077.
- [82] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers Part II: Blind channel estimation, synchronization, and direct equalization," *IEEE Trans. Signal Process.*, vol. 47, pp. 2007–2022, July 1999.
- [83] —, "Redundant filterbank precoders and equalizers Part I: Unification and optimal designs," *IEEE Trans. Signal Processing*, vol. 47, pp. 1988–2006, July 1999.

- [84] N. Seshadri and J. Winters, "Two signaling schemes for improving the error performance of FDD transmission systems using transmitter antenna diversity," in *Proc. 1993 IEEE VTC*, May 1993, pp. 508–511.
- [85] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, "Full-diversity, high rate space-time block codes from division algebras," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2596–2616, Oct. 2003.
- [86] N. Sharma and C. B. Papadias, "Full rate full-diversity linear quasi-orthogonal space-time codes for any number of transmit antennas," *EURASIP J. Applied Signal Processing*, pp. 1246–1256, 2004.
- [87] M. K. Simon and M.-S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. of the IEEE*, vol. 86, pp. 1860–1877, Sept. 1998.
- [88] W. Su and X.-G. Xia, "Two generalized complex orthogonal space-time block codes of rates 7/11 and 3/5 for 5 and 6 transmit antennas," in *International Symposium on Inform. Theory*, Washington DC, June. 2001.
- [89] —, "Signal constellations for quasi-orthogonal space-time block codes with full diversity," *IEEE Trans. Inform. Theory*, vol. 50, pp. 2331–2347, Oct. 2004.
- [90] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July. 1999.
- [91] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

- [92] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Euro. Trans. Telecomm.*, vol. 10, pp. 585–595, Nov./Dec. 1999.
- [93] C. Tepedelenlioglu04, "Maximum multipath diversity with linear equalization in precoded OFDM systems," *IEEE Trans. Inf. Theory*, vol. 50, pp. 232–235, Jan. 2004.
- [94] O. Tirkkonen and A. Hottinen, "Square-matrix embeddable space-time codes for complex signal constellations," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1122–1126, Feb. 2002.
- [95] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," *Proceedings of the IEEE*, vol. 86, pp. 1951–1968, Oct. 1998.
- [96] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [97] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. wireless communi.*, vol. 5, pp. 662–671, Mar. 2006.
- [98] A. D. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory*, vol. 13, pp. 260–269, April 1967.
- [99] M. Vu and A. Paulraj, "Optimal linear precoders for MIMO wireless correlated channels with non-zero mean in space-time coded systems," *IEEE Trans. Signal Processing*, vol. 54, pp. 2318–2332, June 2006.

- [100] G.-Y. Wang, H.-Y. Liao, H.-Q. Wang, and X.-G. Xia, "Systematic and optimal cyclotomic space-time code designs based on high dimension lattices," in *Proc. Global Telecommunicatoion Conference*, 2003.
- [101] H. Wang and X.-G. Xia, "Upper bounds of rates of complex orthogonal space-time block codes," in *International Symposium on Inform. Theory*, Lausanne, Switzerland, June. 2002.
- [102] F. S. Weinstein, "Simplified relationships for the probability distribution of the phase of a sine wave in narrow-band normal noise," *IEEE Trans. Inform. Theory*, vol. 20, pp. 658–661, Sept 1974.
- [103] J. H. Winters, "The diversity gain of transmit diversity in wireless systems with Rayleigh fading," in *Proceedings ICC/SUPERCOMM*, New Orleans, LA, May 1994, pp. 1121–1125.
- [104] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commu.*, vol. 42, pp. 1740–1751, Feb. 1994.
- [105] H. S. Witsenhausen, "A determinant maximization problem occurring in the theory of data," *SIAM J. App. Math.*, vol. 29, pp. 515–522, Nov 1975.
- [106] Y. Xin, Z. Wang, and G. B. Giannakis, "Space-time diversity systems based on linear constellation precoding," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 294–309, Mar. 2003.
- [107] H. Yao and G. W. Wornell, "Achieving the full MIMO diversity-vs-multiplexing frontier with rotation-based space-time codes," in *41th Annual Allerton Conf. on Comm. Control, and Comput.*, Monticello, IL, Oct. 2003.

- [108] J. Zhang, E. K. P. Chong, and D. N. C. Tse, "Output MAI distributions of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Trans. Inf. Theory*, vol. 47, pp. 1128–1144, Mar. 2001.
- [109] J.-K. Zhang, J. Liu, and K. M. Wong, "Trace-orthonormal full diversity cyclotomic linear dispersion codes," in *Proceedings IEEE International Symposium on Information Theory*, Chicago, June 2004.
- [110] —, "Trace-orthonormal full diversity triangular cyclotomic space-time minimizing worst case pair-wise error probability," in *Int. Conf. Sensor Array and Multichannel Signal Processing*, Sitges, Barcelona, Spain, July 2004.
- [111] —, "Linear Toeplitz space-time block codes," in *Proceedings IEEE International Symposium on Information Theory*, Adelaide, Australia, Sept. 2005.
- [112] —, *Trace-Orthogonal Full Diversity Cyclotomic Space-Time Codes*. John Wiley & Sons: Space-Time Processing for MIMO Communications, A. B. Gershman and N. D. Sidiropoulos, 2005.
- [113] —, "Trace-orthonormal full-diversity cyclotomic spacetime codes," *IEEE Trans. Signal Processing*, vol. 55, no. 2, pp. 618–630, Feb. 2007.
- [114] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antennas channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073–1096, May 2003.
- [115] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1673–1690, July 2003.