Least Square Error Detection for Noncoherent Cooperative Relay Systems and Signal Designs Using Uniquely-Factorable Constellations

LEAST SQUARE ERROR DETECTION FOR NONCOHERENT COOPERATIVE RELAY SYSTEMS AND SIGNAL DESIGNS USING UNIQUELY-FACTORABLE CONSTELLATIONS

BY

LI XIONG, B.Eng.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

Master of Applied Science (2011)	McMaster University
(Electrical & Computer Engineering)	Hamilton, Ontario, Canada

TITLE:	Least Square Error Detection for Noncoherent Cooperative Relay Systems and Signal Designs Using Uniquely-Factorable Constellations
AUTHOR:	Li Xiong B.Eng., (Electronic Information Engineering) University of Science and Technology of China, Hefei, China
SUPERVISOR:	Dr. Jian-Kang Zhang
NUMBER OF PAGES:	xv, 132

Dedications

To my parents

Abstract

In this thesis, noncoherent cooperative amplify-and-forward (AF) half-duplex relay systems and wireless communication systems equipped with a single transmitter antenna and multiple receiver antennas (SIMO) are considered, in which perfect channel information is unavailable at the destination end. For the AF half-duplex relay systems, the use of the least square error (LSE) receiver is proposed for detection. By using perturbation theory on the eigenvalues, an asymptotic formula of pairwise error probability for the LSE detector is derived. The result shows that the full diversity gain function mimics coherent cooperative AF half-duplex relay systems, whereas the coding gain function mimics noncoherent multi-inputs multi-outputs (MIMO) systems. In addition, it is proved that for any given nonzero received signal, the unique blind identification of both the equivalent channel and the transmitted signals in a noise-free case is equivalent to full diversity with the LSE detector in a Gaussian noise environment.

In order to design full diversity noncoherent signals for both systems, a novel concept called a *uniquely factorable constellation* (UFC) is proposed in this thesis. It is proved that such a UFC design guarantees the unique blind identification of channel coefficients and transmitted signals in a noise-free case for the SIMO channel by only processing two received signals, as well as full diversity with the noncoherent

maximum likelihood (ML) receiver in a noisy case. By using the Lagrange's four-square theorem, an algorithm is developed to efficiently and effectively design various sizes of energy-efficient unitary UFCs to optimize the coding gain. In addition, a closed-form optimal energy scale is found to maximize the coding gain for the unitary training scheme based on the commonly-used quadrature amplitude modulation (QAM) constellations.

Based on the signal design criterion and UFCs established in this thesis, the systematic designs of noncoherent full diversity unitary constellations for the noncoherent SIMO systems and the noncoherent AF half-duplex protocol with three nodes are proposed. We also derive the closed-form decision rule for the generalized likelihood ratio test (GLRT) receiver for the relay systems. Comprehensive computer simulations show that error performance of the unitary UFC designed in this thesis is superior to those of the differential schemes, the optimal unitary training schemes presented in this thesis and the signal-to-noise ratio (SNR) efficient training schemes using the QAM constellation for the SIMO systems, which, thus far, performs the best error performance in current literatures. Computer simulations also demonstrate that error performance of the unitary diagonal distributed space-time block codes proposed in this thesis outperforms those of the differential codes and the optimally precoded training schemes for the relay systems.

Acknowledgements

Firstly, I am grateful to my supervisor Dr. Jian-Kang Zhang. This thesis would not have been possible without his consistent guidance and encouragement.

Secondly, I would like to thank Dr. Kon Max Wong and Dr. Jun Chen in the communications research group at McMaster University for their insightful discussions and invaluable suggestions. Besides, I am deeply indebted to Dr. Kon Max Wong, Dr. Dongmei Zhao and Dr. Jian-Kang Zhang for being my defence committee members, as well as sharing their advice and critiques on my thesis.

Thirdly, I offer my regards and blessings to all my group members, Dong Xia, Lisha Wang and Jiaping Liang for their help along the way and the laughters they brought to me. I would also like to thank Ms. Cheryl Gies, the graduate administrative assistant. Her kindness and willingness to help make my study here comfortable and enjoyable.

Lastly, I own my deepest gratitude to my family and my girlfriend for the love and constant support.

Acronyms

AF	Amplify-and-Forward
DDUFC	Diagonal distributed unitary-UFC
GLRT	Generalized Likelihood Ratio Test
LSE	Least Square Error
MIMO	Multi-Input Multi-Output
ML	Maximum Likelihood
PSK	Phase-Shift Keying
QAM	Quadrature Amplitude Modulation
SIMO	Single-Input Multiple-Output
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
UFC	Uniquely-Factorable Constellation

Glossary of Symbols

$\mathbf{a}$	Column vector $\mathbf{a}$
$\mathbf{A}$	Matrix $\mathbf{A}$
$(\cdot)^T$	The transpose of a vector or matrix
$(\cdot)^*$	The complex conjugate of a vector or matrix
$(\cdot)^H$	The Hermitian of a vector or matrix
$\mathrm{Tr}(\cdot)$	The trace operator
$\det(\cdot)$	The determinant operator
$\|\cdot\|_{\mathrm{F}}$	The Frobenius norm of a vector or matrix
$\|\cdot\|_2$	The 2 norm of a vector or matrix
$\mathbf{I}_N$	$N \times N$ identity matrix
Φ	Empty set
ln	Natural logarithm
$\mathrm{E}[\cdot]$	Expectation operator
$\Re\{\cdot\}$	Real part of the variable in the curly bracket
$\otimes$	Kronecker product

j	$\sqrt{-1}$
$f(x) = O(g(x))$	There exists a positive real number M and a real number x_0, such that $\lvert f(x)\rvert \leq M\lvert g(x)\rvert$ for all $x \geq x_0$
$f(x) = o(g(x))$	$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$
$\mathbf{A} \preceq \mathbf{B}$	$\mathbf{A}$, $\mathbf{B}$ are positive semi-definite and $\mathbf{B} - \mathbf{A}$ is also positive semi-definite

Contents

List of Figures

Chapter 1

Introduction

1.1 Noncoherent AF Relay Systems

Over the past several years, various forms of diversity have been employed in practical wireless communication systems to overcome the effects of channel fading. Among them, spatial diversity is most commonly utilized since it can be readily combined with the other forms (such as time, frequency) of diversity. The gain in employing spatial diversity is usually measured by the product (full diversity) of the number of transmitter and receiver antennas in the MIMO system with *linear* flat fading channels. The full spatial diversity can be achieved by the use of space-time block coding. Another form of spatial diversity called cooperative diversity has more recently been proposed for mobile wireless communications [1–4], in which the in-cell mobile users share the use of their antennas to create a virtual array through distributed transmission and signal processing. When channel state information is available at the receiver, a full diversity gain for the coherent cooperative relay system with *product* flat fading channels is characterized by the diversity gain function [5–7] and achieved by utilizing well designed precoders [8, 9] or distributed space-time block codes [5–7, 10, 11]. Unfortunately, full channel state information at the relay nodes and the destination nodes, in practice, is difficult to be attained. If the channel changes slowly, then, the transmitter may have sufficiently long coherence time to send training signals for the accurate estimation of the channel coefficients. However, the fading coefficients in mobile wireless communications may vary rapidly and the coherence time may be so short that it is impossible to allow the reliable estimation of the coefficients. Therefore, the time utilized on transmitting training signals has to be counted since more training signals need to be sent for the precise estimation of the channel [12–14].

Therefore, in recent years, more and more research work has focused on the noncoherent cooperative relay system [15–19], where the channel coefficients are assumed to be unknown at both relay nodes and the destination, but remain unchanged within certain time slots, after which they change to a new independent realization and so on. In spite of the fact that the design criterion of full diversity for noncoherent MIMO systems has been well established with the ML detector and the GLRT detector, this result cannot be directly applied to the noncoherent cooperative relay systems. Since the probability density function of the received signal conditioned the transmitted signal for the noncoherent relay systems is very complicated, there is no explicit decision rule for the ML receiver. In addition, the GLRT detector is also very complicated for the noncoherent relay systems, since the channel now is a product of Gaussian channels, which is not necessarily Gaussian anymore, and the additive noise, which depends on the relay-destination channel coefficients, is not white Gaussian. As a result, unlike the noncoherent MIMO systems, the GLRT receiver for the noncoherent relay systems is not equivalent to the LSE receiver anymore and has no explicit decision rule in general.

1.2 Noncoherent SIMO Systems

In this thesis, we also consider a wireless communication system having a single transmitter antenna and N receiver antennas, where channel state information is not available at either the transmitter or the receiver, and is constant during 2 time slots, after which it changes to new independent values that are fixed for another 2 time slots, and so on. Particularly for $N = 1$, i.e, a single-input-single-output (SISO) system with the fast changing flat Rayleigh fading, in a 1969 technical report, Richters [20] made a rather unexpected conjecture for the channel with a continuous input under an average power constraint that the capacity-achieving input distribution is discrete. In 2001, Abou-Faycal, Trott and Shamai [21] proved rigorously that the conjecture was true. In fact, the optimal input distribution is discrete with a finite number of mass points and one of them located at the origin. This result, afterwards, was extended by Gursor, Poor and Verdu [22, 23] into the fast-changing Rician fading channel. Since then, the extensive studies on the ergodic channel capacities for general MIMO channels have become an important research topic in wireless communications [14, 24–26].

In this thesis, we are interested in the design of constellations from the standing-points of blind signal processing and detection theory for the fast-varying flat Rayleigh fading SIMO channel. It is known that the optimal design of constellations is a classic problem for an additive white Gaussian noise (AWGN) channel [27–32]. However, just as Gallager said [33], the resulting discrete optimization problem is ugly, since it is extremely difficult to be formulated into a tractable optimization problem. On the other hand, the QAM constellation carved from the Gaussian integer ring is very easily designed, very efficiently modulated and demodulated. Hence, it is commonly used in modern digital communication systems. Recently, the hexagonal constellation [34–36] carved from the Eisenstein integer ring has attracted much attention because of the fact that it is more energy-efficient than the QAM constellation [31] and that an efficient demodulation algorithm has been found [36].

Meanwhile, current research on coherent MIMO communications tells us that if perfect channel state information is available at the receiver, then, any signaling scheme for the specific SIMO channel enables coherent full diversity for any constellations and with linear receivers. Unfortunately, this is no longer true for noncoherent communications, even if the commonly-used QAM constellations are transmitted and even if the noncoherent ML receiver is employed, since some signal points of the constellation do not necessarily satisfy the unique factorization in the sense of multiplications (see more details in Theorem 3). Also, it is for the same reason that either the channel coefficients or the transmitted signals are not necessarily able to be uniquely identified, even in a noise-free case. Hence, for the noncoherent SIMO channel, signals must be carefully designed. The unknown of the fading channel at both the transmitter and the receiver requires that the transmitted signals emitting from two distinct time slots must be so correlated that reliable communications with noncoherent full diversity are made possible under a maximum allowable transmission rate.

In addition, attaining perfect channel state information at the receiver, in practice, is a challenging problem [12–14]. Therefore, noncoherent space-time block code designs [37–42] have been recently developed. It has been proved that the unitary constellation is optimal [14, 26, 37, 43] when either SNR is high or coherence time is long. Therefore, most of the noncoherent space-time block code designs have been mainly focused on unitary designs [37–42, 44, 45]. The Cayley [44, 46] transform and the exponential transform [42] are now two well-established transforms which map respective linear dispersion and linear codes into unitary codes. The exponential transform [42] requires that the number of the receiver antennas is not less than that of the transmitter antennas. In general, it cannot guarantee full diversity for the noncoherent ML receiver. The Cayley transform aimed mainly at differential modulation and a differential receiver.

1.3 Contributions of This Thesis

There are two contributions in this thesis. The first contribution is to propose the use of the LSE detector for the noncoherent cooperative relay systems and analyze its asymptotic behavior [47]. The significant advantage of the LSE detector is that it requires statistics of neither the channel nor the noise, which makes it very attractive, particularly for the noncoherent relay systems. Despite the fact that asymptotic formulae of pairwise error probabilities for the noncoherent MIMO systems with the ML and GLRT receivers have been developed by Brehler and Varanasi [43], when these asymptotic formulae are applied to the relay system for given the transmitted signal and the channel gains from the relay to the destination, the resulting channel covariance matrix depends on these relay-destination channels, incurring the fact that the dominant term of the corresponding asymptotic formulae is not integrable when an expectation is taken over the relay-destination channels (see more details in Chapter 3). Therefore, we re-derive a more accurate asymptotic formula using perturbation theory on the eigenvalues. With this, the asymptotic formula of the pairwise error probability for the noncoherent relay systems with the LSE receiver is obtained.

The other contribution is to invent a novel concept, a *uniquely factorable constellation* (UFC) [48], for the systematic designs of noncoherent full diversity unitary constellations for the noncoherent SIMO systems and the noncoherent AF half-duplex protocol with three nodes. By using the Lagrange's four-square theorem, an algorithm is developed to efficiently and effectively design various sizes of energy-efficient unitary UFCs to optimize the coding gain. In addition, a closed-form optimal energy scale is found to maximize the coding gain for the unitary training scheme based on the commonly-used QAM constellations.

Chapter 2

Detection for Noncoherent Space-Time Block Coded MIMO Systems

In this chapter, we first briefly review the ML and GLRT receivers for noncoherent space-time block coded MIMO channels. Then, we introduce the asymptotic formula of pairwise error probability with the GLRT receiver established by Brehler and Varanasi [43]. Finally, using perturbation theory on the eigenvalues, we re-derive a more precise formula for the analysis of the asymptotic behavior of pairwise error probability for the LSE receiver in noncoherent cooperative relay systems.

2.1 Channel Model

Let us first consider a space-time block coded noncoherent MIMO system with M transmitter antennas, N receiver antennas. The channel state information is not known at either the transmitter or at the receiver, and remains constant for T time slots, after which it changes to a new independent realization and so on. The equivalent vector channel model is given by

$$\mathbf{r} = \sqrt{\rho}\mathbf{S}\mathbf{h} + \boldsymbol{\xi} \tag{2.1}$$

where $\mathbf{r}$ denotes an $MN \times 1$ received signal vector, $\mathbf{h}$ denotes an $MN \times 1$ channel vector, $\tilde{\mathbf{S}}$ is a $T \times M$ codeword matrix, $\mathbf{S} = \mathbf{I}_N \otimes \tilde{\mathbf{S}}$, $\boldsymbol{\xi}$ denotes an $MN \times 1$ noise vector and ρ is the signal-to-noise ratio. We assume that the channels linking to the same receiver antenna are correlated among themselves, but are uncorrelated with the channels linking to different receiver antennas, i.e.,

$$\mathrm{E}[\mathbf{h}_\ell \mathbf{h}_n^H] = \begin{cases} \tilde{\boldsymbol{\Sigma}} & \ell = n, \\ \mathbf{0} & \ell \neq n. \end{cases}$$

where $\mathbf{h}_l = (h_{(l-1)M+1}, \cdots, h_{lM})^T$. Then, the covariance matrix of $\mathbf{h}$ is $\boldsymbol{\Sigma} = \mathrm{E}[\mathbf{h}\mathbf{h}^H] = \mathbf{I}_N \otimes \tilde{\boldsymbol{\Sigma}}$. We also assume that the elements ξ_n of $\boldsymbol{\xi}$ are samples of independent circularly-symmetric zero-mean complex Gaussian random variables with each having unit variance.

2.2 ML Detection

Under these assumptions made for the channel model (2.1), the probability density function of the received signal vector $\mathbf{r}$ conditioned on the transmitted signal matrix

$\mathbf{S}$, $f(\mathbf{r}|\mathbf{S})$ is the Gaussian distribution, i.e.,

$$f(\mathbf{r}|\mathbf{S}) = \frac{1}{\pi^{MN}\det(\rho\mathbf{S}\boldsymbol{\Sigma}\mathbf{S}^H + \mathbf{I})} \times \exp\left(-\mathrm{Tr}\big(\mathbf{r}^H(\rho\mathbf{S}\boldsymbol{\Sigma}\mathbf{S}^H + \mathbf{I})^{-1}\mathbf{r}\big)\right)$$

and thus, its likelihood is $-\mathrm{Tr}\big(\mathbf{r}^H(\rho\mathbf{S}\boldsymbol{\Sigma}\mathbf{S}^H + \mathbf{I})^{-1}\mathbf{r}\big) - \ln\det(\rho\mathbf{S}\boldsymbol{\Sigma}\mathbf{S}^H + \mathbf{I}) - MN\ln\pi$. Then, the maximum likelihood receiver for the noncoherent MIMO system is to solve an optimization problem: $\hat{\mathbf{S}} = \arg\min_{\mathbf{S}}\left(G_{\mathbf{s}} + \mathrm{Tr}\left(\mathbf{r}^H\boldsymbol{\Delta}_{\mathbf{s}}\mathbf{r}\right)\right)$, where $\boldsymbol{\Delta}_{\mathbf{s}} = \left(\rho\mathbf{S}\boldsymbol{\Sigma}\mathbf{S}^H + \mathbf{I}\right)^{-1}$ and $G_{\mathbf{s}} = \log\det\left(\rho\mathbf{S}\boldsymbol{\Sigma}\mathbf{S}^H + \mathbf{I}\right)$.

2.3 GLRT Detection

In order to circumvent the difficulties in variance estimation of noise, the generalized likelihood ratio test receiver is proposed. Under the assumptions made in Section 2.1, the probability density function of the received signal matrix $\mathbf{r}$ conditioned on the channel matrix $\mathbf{h}$ and the transmitted signal matrix $\mathbf{S}$ is the Gaussian distribution, i.e., $\frac{1}{\pi^{MN}}\exp\left(-\|\mathbf{r} - \sqrt{\rho}\mathbf{S}\mathbf{h}\|_2^2\right)$, and thus, its likelihood is given by $-\|\mathbf{r} - \sqrt{\rho}\mathbf{S}\mathbf{h}\|_2^2 - MN\ln\pi$. Therefore, the GLRT receiver for the joint estimation of $\mathbf{h}$ and $\mathbf{S}$ is to maximize the likelihood, which is essentially equivalent to solving the following nonlinear least squares optimization problem [49]:

$$\{\hat{\mathbf{h}}, \hat{\mathbf{S}}\} = \arg\min_{\mathbf{h},\mathbf{S}} \|\mathbf{r} - \sqrt{\rho}\mathbf{S}\mathbf{h}\|_2^2$$

We should note that in this case, the GLRT receiver is equivalent to the LSE receiver. The solution of the optimization problem can be obtained by first estimating the transmitted signal matrix $\mathbf{S}$ as

$$\hat{\mathbf{S}} = \arg\max_{\mathbf{S}} \operatorname{tr}\left(\mathbf{r}^H \mathbf{S}\left(\mathbf{S}^H \mathbf{S}\right)^{-1} \mathbf{S}^H \mathbf{r}\right), \tag{2.2}$$

and then, estimating the channel vector $\mathbf{h}$ as $\hat{\mathbf{h}} = (\hat{\mathbf{S}}^H \hat{\mathbf{S}})^{-1} \hat{\mathbf{S}}^H \mathbf{r}/\sqrt{\rho}$. In fact, the GLRT detector for the estimation of the transmitted signal projects the received signal $\mathbf{r}$ on the different subspaces generated by $\mathbf{S}$ and then computes the energies of all the projections and selects the projection maximizing the energy. To study the asymptotic behavior of the pairwise error probability with the GLRT receiver for the noncoherent MIMO systems, we first introduce several necessary propositions and then, establish some lemmas, which facilitate our analysis.

Proposition 1 *Let $\mathbf{A}$ be a Hermitian matrix with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$ and $\mathbf{A}$ be a non-Hermitian perturbation of $\mathbf{A}$, i.e., $\mathbf{E} = \mathbf{A} - \mathbf{A}$ is non-Hermitian. If we let $\tilde{\lambda}_1, \tilde{\lambda}_2, \cdots, \tilde{\lambda}_n$ denote the eigenvalues of $\mathbf{A}$ with $\Re\{\tilde{\lambda}_1\} \leq \Re\{\tilde{\lambda}_2\} \leq \cdots \leq \Re\{\tilde{\lambda}_n\}$, then, we have*

$$\sum_{k=1}^{n} |\tilde{\lambda}_k - \lambda_k|^2 \leq 2||\mathbf{E}||_{\mathrm{F}}^2 \tag{2.3}$$

See [50] for the proof of Proposition 1. We also need the following proposition, whose proof is given in [51].

Proposition 2 *For an $M \times N$ matrix $\mathbf{A}$ with rank N and an $N \times P$ matrix $\mathbf{B}$, the following inequalities hold*

$$\begin{aligned}
||\mathbf{A}||_2 &\leq ||\mathbf{A}||_{\mathrm{F}} \leq \sqrt{N}||\mathbf{A}||_2 \\
||\mathbf{AB}||_2 &\leq ||\mathbf{A}||_2||\mathbf{B}||_2 \\
||\mathbf{AB}||_{\mathrm{F}} &\leq ||\mathbf{A}||_{\mathrm{F}}||\mathbf{B}||_{\mathrm{F}}
\end{aligned}$$

where the F-norm and 2-norm are defined as

$$
\begin{aligned}
||\mathbf{A}||_{\mathrm{F}} &= \sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}|a_{ij}|^2} \\
||\mathbf{A}||_2 &= \sqrt{\lambda_{\max}(\mathbf{A}^H\mathbf{A})}
\end{aligned}
$$

Let us now consider the analysis of the asymptotic behavior of the GLRT receiver. For notation convenience, let $\{\mathbf{S}_i\}$ denote all the distinct codeword matrices, then, the GLRT detector (2.2) can be rewritten as

$$\hat{i} = \arg\min \mathbf{r}^H\mathbf{F}_i\mathbf{r} = \arg\min \nabla_i$$

where $\mathbf{F}_i = -\mathbf{S}_i(\mathbf{S}_i^H\mathbf{S}_i)^{-1}\mathbf{S}_i^H$ and $\nabla_i = \mathbf{r}^H\mathbf{F}_i\mathbf{r}$. The pairwise error probability can be expressed by

$$\Pr(\nabla_j < \nabla_i) = \Pr(\mathbf{x}^H\mathbf{F}_{ij}\mathbf{x} < 0)$$

where $\mathbf{F}_{ij} = \mathbf{F}_i - \mathbf{F}_j$. For discussion simplicity, let $\mathbf{R}_{ij} = \mathbf{S}_i^H\mathbf{S}_j$ and $\mathbf{P}_{\mathbf{w}\hat{\mathbf{w}}} = (\mathbf{W}, \hat{\mathbf{W}})^H(\mathbf{W}, \hat{\mathbf{W}})$. Then, the following lemma is the key to obtaining an asymptotic formula for the pairwise error probability with the GLRT receiver.

Lemma 1 *If we let $\nu_1 \le \cdots \le \nu_{MN}$ denote the eigenvalues of $\mathbf{\Sigma}^{1/2}(\mathbf{R}_{ii}-\mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji})\mathbf{\Sigma}^{1/2}$ and $\lambda_1, \cdots, \lambda_{2MN}$ denote the eigenvalues of $\Sigma_{\mathbf{rr}|i}\mathbf{F}_{ij}$, where $\Re\{\lambda_1\} \le \cdots \le \Re\{\lambda_{2MN}\}$, then, the following asymptotic formulae hold,*

$$
\begin{aligned}
\lambda_k &= -1 + O\left(\frac{\kappa}{\rho\nu_1+\epsilon}\right), \qquad \textit{for } k = 1, \cdots, MN \\
\lambda_k &= \rho\nu_k + 1 + O\left(\frac{\kappa}{\rho\nu_1+\epsilon}\right) \qquad \textit{for } k = MN+1, \cdots, 2MN
\end{aligned}
$$

where $\kappa = \max(1, \sqrt{\lambda_{\max}(\mathbf{\Sigma})})/\min(1, \sqrt{\lambda_{\min}(\mathbf{\Sigma})})$.

The proof of Lemma 1 is provided in Appendix A.

Proposition 3 *Let* $\{\lambda_l\}_{l=1}^{L}$ *be the distinct nonzero eigenvalues of* $\mathbf{\Sigma}_{\mathbf{rr}|i}\mathbf{F}_{ij}$ *with multiplicity* $\{\mu_l\}_{l=1}^{L}$*, and let* $\{\lambda_l\}_{l=1}^{K}$ *be negative and* $\{\lambda_l\}_{l=K+1}^{L}$ *positive eigenvalues, respectively. Then, the pairwise error probability is given by*

$$\Pr\{\nabla_j < \nabla_i\} = -\sum_{k=1}^{K} Res\left(s^{-1}\prod_{l=1}^{L}\lambda_l^{-\mu_l}\left(s+\frac{1}{\lambda_l}\right)^{-\mu_l}, \quad s_k = \frac{-1}{\lambda_k}\right) \tag{2.4}$$

Proposition 3, which is given in [43], establishes a connection between the average error probability and the eigenvalues of matrix $\mathbf{\Sigma}_{\mathbf{rr}|i}\mathbf{F}_{ij}$. Using Proposition 3, Brehler and Varanasi [43] proved the following formula,

$$\mathrm{P}_{\mathrm{GLRT}}(\mathbf{S}_i \to \mathbf{S}_j) = \frac{\begin{pmatrix} 2MN-1 \\ MN \end{pmatrix}\det(\mathbf{R}_{jj})\rho^{-MN}}{\det(\mathbf{\Sigma})\det(\mathbf{P}_{\mathbf{s}_i\mathbf{s}_j})} + o\big(\rho^{-(MN+1)}\big) \tag{2.5}$$

However, we cannot directly apply this formula to the relay systems. The reason is as follows: the conditional received signal is not Gaussian distributed. Even if the received signal is Gaussian distributed for given the transmitted signal and the channel gain from the relay to the destination, the resulting covariance matrix $\mathbf{\Sigma}$ depends on the relay-destination channel, incurring that the dominant term in (2.5) is not integrable when an expectation is taken over the relay-destination channel. See more details in Chapter 3. Therefore, for our purpose, we need to re-derive a more precise asymptotic formula of the GLRT detector for the noncoherent MIMO systems. Now, combining Lemma 1 with Proposition 3 yields

Property 1 *If each of matrices* $\mathbf{S}_{ij} = (\mathbf{S}_i, \mathbf{S}_j)$ *has full column rank for all distinct pairs of* i *and* j*, then, the resulting space-time block code provides full diversity for the GLRT receiver. Moreover, the average pairwise error probability* $\mathrm{P}_{\mathrm{GLRT}}(\mathbf{S}_i \to \mathbf{S}_j)$ *of transmitting* $\mathbf{S}_i$ *and deciding in favor of* $\mathbf{S}_j \neq \mathbf{S}_i$ *has the following asymptotic formula:*

$$\mathrm{P}_{\mathrm{GLRT}}(\mathbf{S}_i \to \mathbf{S}_j) = \frac{\begin{pmatrix} 2MN-1 \\ MN \end{pmatrix} \rho^{-MN}}{\det[\mathbf{\Sigma}(\mathbf{R}_{ii} - \mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}) + \rho^{-1}\mathbf{I}_{MN}]} + O\big(\rho^{-(MN+1)}\big) \qquad (2.6)$$

The proof of Property 1 is provided in Appendix B. We like to make two comments on Property 1.

1. When ρ tends to infinity, (2.6) amounts to,

$$\mathrm{P}_{\mathrm{GLRT}}(\mathbf{S}_i \to \mathbf{S}_j) = \frac{\begin{pmatrix} 2MN-1 \\ MN \end{pmatrix} \det(\mathbf{R}_{jj}) \rho^{-MN}}{\det(\mathbf{\Sigma}) \det(\mathbf{P}_{\mathbf{s}_i\mathbf{s}_j})} + O\big(\rho^{-(MN+1)}\big) \qquad (2.7)$$

2. Compared with the asymptotic formula (2.5), the asymptotic formula (2.6) leads to a more precise characterization of the coding gain as well as of the perturbation term.

3. From the proof of Property 1 we can see that the condition that $\mathbf{P}_{\mathbf{s}_i\mathbf{s}_j}$ has full rank for all the distinct pairs of i and j is actually a necessary and sufficient condition for the GLRT receiver to achieve full diversity.

Chapter 3

LSE Detection for Noncoherent Cooperative Relay Systems

The primary purpose of this chapter is to propose the LSE detector for the noncoherent cooperative relay systems and extend the asymptotic formula for the noncoherent MIMO systems with the GLRT receiver into that for the noncoherent relay systems with the LSE receiver.

3.1 Channel Model

Here, we are interested in the following two kinds of Gaussian product channel models, which usually appear in recently-developing cooperative relay communication systems.

1. Linear and product mixed channels:

$$\mathbf{z}_1 = \sqrt{\rho}\,\mathbf{X}_1\mathbf{h}_1 + \boldsymbol{\eta}_1, \tag{3.1}$$

where $\mathbf{h}_1 = (h, f_1 g_1, f_2 g_2, \cdots, f_M g_M)^T$ and $\mathbf{X}_1$ is a $T_1 \times (M+1)$ transmitted signal matrix with $T_1 \geq M+1$. A typical example for this kind of the channel model is the AF half-duplex protocol proposed in [4], where each node has a single antenna and the coefficients h, f_m and g_m for $m = 1, 2, \cdots, M$ in model (3.1) are respective the channel from the source to the destination (linear channel), the channel from the source to the mth relay and the channel from the mth relay to the destination (product channel). For the block length between the source and each active relay to be 2, the signal matrix $\mathbf{X}_1$ specifically takes the following form:

$$\mathbf{X}_1 = \begin{pmatrix} s_1 & 0 & 0 & \cdots & 0 & 0 \\ s_2 & s_1 & 0 & \cdots & 0 & 0 \\ s_3 & 0 & 0 & \cdots & 0 & 0 \\ s_4 & 0 & s_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 \\ s_{2M-1} & 0 & 0 & \cdots & 0 & 0 \\ s_{2M} & 0 & 0 & \cdots & 0 & s_{2M-1} \end{pmatrix}. \tag{3.2}$$

2. Only relay channels:

$$\mathbf{z}_2 = \sqrt{\rho}\,\mathbf{X}_2\mathbf{h}_2 + \boldsymbol{\eta}_2, \tag{3.3}$$

where $\mathbf{h}_2 = (f_1g_1, f_2g_2, \cdots, f_Mg_M)^T$ and $\mathbf{X}_2$ is a $T_2 \times M$ transmitted signal matrix with $T_2 \geq M$ and a normalized power. For example, the AF protocols proposed in [5–7] belong to this family of the channel model, in which the direct channel between the source and the destination is not allowed and $\mathbf{X}_2$ in (3.3) is the distributed space-time block coded signal matrix.

In developing our analysis on an average pairwise error probability, we adopt the following assumptions:

1. *Perfect channel information is not available at either the source node and the relay nodes or the destination*;

2. *The channel coefficients f_m, g_m and h are independently circularly-symmetric complex Gaussian distributed with zero-mean and unit variances*;

3. *For any fixed $\mathbf{g} = (g_1, g_2, \cdots, g_M)^T$, each $\boldsymbol{\eta}_i$ for $i = 1, 2$ is independently circularly-symmetric complex Gaussian noise with the zero mean and covariance matrix*

$$\boldsymbol{\Sigma}_1 = \mathrm{diag}(\mathbf{g}^H\mathbf{B}_{11}\mathbf{g}, \mathbf{g}^H\mathbf{B}_{12}\mathbf{g}, \cdots, \mathbf{g}^H\mathbf{B}_{1T_1}\mathbf{g}) + \mathbf{I}_{T_1}$$

and

$$\boldsymbol{\Sigma}_2 = \mathrm{diag}(\mathbf{g}^H\mathbf{B}_{21}\mathbf{g}, \mathbf{g}^H\mathbf{B}_{22}\mathbf{g}, \cdots, \mathbf{g}^H\mathbf{B}_{2T_2}\mathbf{g}) + \mathbf{I}_{T_2}$$

where $\mathbf{B}_{it_i} = \mathrm{diag}(\delta_{it_i}^{(1)}, \delta_{it_i}^{(2)}, \cdots, \delta_{it_i}^{(M)})$ *with at most one of* $\delta_{it_i}^{(m)}$ *for* $m = 1, 2, \cdots, M$ *being one and the others being zeros.*

4. *The average energy of* $\mathbf{X}_i$*,* E_{x_i} *is normalized into* $E_{x_i} = \sum_{t_i=1}^{T_i} \mathrm{Tr}(\mathbf{B}_{it_i}) + T_i$ *.*

3.2 ML and GLRT Detectors

Despite the fact that it looks that the mathematical channel models between the noncoherent MIMO systems (2.1) and the noncoherent cooperative relay systems (3.1) and (3.3) are almost the same, there are in fact two major differences.

1. The MIMO channel is linear, but the cooperative relay channel is nonlinear, since it involves a variety of product channels from the source to relay and from the relay to the destination.

2. Under the assumptions on the channel models, the probability density function of the received signal vector conditioned on the transmitted signals for the noncoherent MIMO systems is Gaussian distributed, but it is not for the noncoherent cooperative relay systems.

Because of these, the probability density function of the received signal vector $\mathbf{z}_i, (i = 1, 2)$ conditioned on the transmitted signal matrix $\mathbf{X}_i$ has no closed-form formulae and in general, ML detection for the noncoherent relay systems is too complicated to be implemented.

On the other hand, although the conditional probability density function of the received signal vector $\mathbf{z}_i, (i = 1, 2)$ given the channel coefficients and the transmitted signal matrix is still the Gaussian distribution, i.e., $\frac{1}{\pi^{T_i} \det(\mathbf{\Sigma}_i)} \times \exp\left(-(\mathbf{z}_i - \sqrt{\rho}\mathbf{X}_i\mathbf{h}_i)^H \mathbf{\Sigma}_i^{-1}\right.$

$(\mathbf{z}_i - \sqrt{\rho}\mathbf{X}_i\mathbf{h}_i)\big)$, and thus, its likelihood is given by $-(\mathbf{z}_i - \sqrt{\rho}\mathbf{X}_i\mathbf{h}_i)^H \mathbf{\Sigma}_i^{-1}(\mathbf{z}_i - \sqrt{\rho}\mathbf{X}_i\mathbf{h}_i)$ $- \ln\det(\mathbf{\Sigma}_i) - T_i \ln \pi$, the covariance matrix $\mathbf{\Sigma}_i$ depends on the channel coefficients from the relay to the destination. Therefore, unlike the noncoherent MIMO systems, the GLRT detector for the noncoherent cooperative relay systems of maximizing the likelihood over the channel coefficients and subsequently over the transmitted signals is not equivalent to the LSE receiver anymore [17]. As a result, the optimization of the likelihood function is quite involved and no closed-form decision rule can be easily obtained generally. However, for some particular protocols, the GLRT receiver [17] may have a closed-form expression. Hence, suboptimal receivers such as the maximum energy selection receiver [18] and the LSE receiver [19] have been recently proposed for some specific noncoherent cooperative relay systems.

3.3 LSE Detection

In this thesis, instead of employing GLRT detector in noncoherent relay systems, we propose the use of the LSE receiver for more general noncoherent cooperative relay systems. Essentially, the LSE receiver to deal with the problem of jointly and optimally estimating the transmitted signals and channel coefficients for our channel model (3.1) or (3.3) is equivalent to solving the following optimization problem [49]:

$$\{\hat{\mathbf{X}}_i, \hat{\mathbf{h}}_i\} = \arg\min_{\mathbf{X}_i, \mathbf{h}_i} \|\mathbf{z}_i - \sqrt{\rho}\mathbf{X}_i\mathbf{h}_i\|_2^2 = \arg\min_{\mathbf{X}_i} \min_{\mathbf{h}_i} \|\mathbf{z}_i - \sqrt{\rho}\mathbf{X}_i\mathbf{h}_i\|_2^2 \tag{3.4}$$

For the inner minimization, differentiating the objective with respect to $\mathbf{h}_i$ and equating it to zero yields $\hat{\mathbf{h}}_i = ({\mathbf{X}_i}^H\mathbf{X}_i)^{-1}{\mathbf{X}_i}^H\mathbf{z}_i/\sqrt{\rho}$, which, when substituted into (3.4), leads to

$$\hat{\mathbf{X}}_i = \arg\max_{\mathbf{X}_i} \mathbf{z}_i{}^H \mathbf{X}_i \big(\mathbf{X}_i{}^H \mathbf{X}_i\big)^{-1} \mathbf{X}_i{}^H \mathbf{z}_i \tag{3.5}$$

The following propositions and lemmas are presented to assist the derivation of the average pairwise error probabilities.

Proposition 4 *Let $x > 0$. Then, we have*

$$\mathcal{E}(x) = \int_x^{+\infty} \frac{e^{-t}}{t} dt = \gamma - \ln x - \sum_{n=1}^{\infty} \frac{(-1)^n x^n}{n! n}, \tag{3.6}$$

where γ is the Euler constant.

The proof of Proposition 4 is provided in [52].

Property 2 *For a given constant $a > 0$ and integers $m, n \geq 1$, the following asymptotic formulae hold*

$$\begin{aligned}
\int_0^{\infty} \frac{\exp(-t)}{a\rho t + 1} dt &= O(\rho^{-1} \ln \rho) \\
\int_0^{\infty} \frac{t^n \exp(-t)}{a\rho t + 1} dt &= O(\rho^{-1}) \\
\int_0^{\infty} \frac{\exp(-t)}{(t + a\rho^{-1})^n} dt &= O(\ln \rho) \\
\int_{a\rho^{-1}}^{\infty} t^n \exp(-t) dt &= 1 + O(\rho^{-1}) \\
\int_{a\rho^{-1}}^{\infty} \frac{\exp(-t)}{t^n} dt &= O(\ln \rho) \\
\int_0^{\infty} \frac{t^m \exp(-t)}{(t+1)^n (t + a\rho^{-1})} dt &= O(1)
\end{aligned}$$

when signal-to-noise ratio ρ tends to infinity.

The proof of Property 2 is provided in Appendix C.

Proposition 5 *Let*

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \\ \mathbf{K}_{21} & \mathbf{K}_{22} \end{pmatrix} \tag{3.7}$$

be a positive semi-definite matrix. Then, the following two statements are true.

1. *If* $\mathbf{K}_{11}$ *is invertible, then, its Schur complementary matrix* $\mathbf{K}_{22} - \mathbf{K}_{21}\mathbf{K}_{11}^{-1}\mathbf{K}_{12}$ *is also a positive semi-definite matrix and* $\det\big(\mathbf{K}\big) = \det(\mathbf{K}_{11})\det\big(\mathbf{K}_{22}-\mathbf{K}_{21}\mathbf{K}_{11}^{-1}\mathbf{K}_{12}\big)$.

2. *If* $\mathbf{K}_{22}$ *is invertible, then, its Schur complementary matrix* $\mathbf{K}_{11} - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{K}_{21}$ *is also a positive semi-definite matrix and* $\det\big(\mathbf{K}\big) = \det(\mathbf{K}_{22})\det\big(\mathbf{K}_{11}-\mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{K}_{21}\big)$.

See [53] for the proof of Proposition 5.

Lemma 2 *For an* $M \times M$ *semi-definite matrix* $\mathbf{P}$ *with rank* N_1, *if we let* $J(\mathbf{g}, \mathbf{P}) = \det\big(\mathbf{I} + \rho\,\mathrm{diag}(\mathbf{g})^H\mathbf{P}\mathrm{diag}(\mathbf{g})\big)^{-1}$ *and* $F(\rho, \mathbf{P}) = \mathrm{E}_{\mathbf{g}}[J(\mathbf{g}, \mathbf{P})]$, *then,* $F(\rho, \mathbf{P})$ *has the following asymptotic formula,*

$$F\big(\rho, \mathbf{P}\big) = \frac{\ln^{N_1}\rho}{\det\left(\mathbf{I} + \rho\,\mathbf{P}\right)} + O\left(\frac{\ln^{N_1-1}\rho}{\rho^{N_1}}\right) \tag{3.8}$$

when signal-to-noise ratio ρ *tends to infinity.*

The proof of Lemma 2 is given in Appendix D. Now, we formally state the main result of this thesis.

Theorem 1 *The following two statements are true.*

1. *If $\mathbf{P}_{\mathbf{x}_1\hat{\mathbf{x}}_1}$ is invertible for all the distinct pairs of $\mathbf{X}_1$ and $\hat{\mathbf{X}}_1$, then, the resulting code enables full diversity for the LSE detector. Furthermore, the average pairwise error probability of the LSE detector of transmitting $\mathbf{X}_1$ and deciding in favor of $\hat{\mathbf{X}}_1 \neq \mathbf{X}_1$ for the linear and product mixed channels* (3.1), $\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_1 \to \hat{\mathbf{X}}_1)$ *has an asymptotic formula as follows:*

$$\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_1 \to \hat{\mathbf{X}}_1) = \frac{\begin{pmatrix} 2M+1 \\ \\ M+1 \end{pmatrix} \det(\hat{\mathbf{X}}_1^H \mathbf{X}_1) \ln^M \rho}{\det(\mathbf{P}_{\mathbf{x}_1\hat{\mathbf{x}}_1})\rho^{M+1}} + O(\rho^{-M-1} \ln^{M-1} \rho) \quad (3.9)$$

2. *If $\mathbf{P}_{\mathbf{x}_2\hat{\mathbf{x}}_2}$ is invertible for all the distinct pairs of $\mathbf{X}_2$ and $\hat{\mathbf{X}}_2$, then, full diversity is achieved with the LSE detector for the only relay channel* (3.3)*. In addition, an asymptotic formula of the average pairwise error probability is given by*

$$\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2) = \frac{\begin{pmatrix} 2M-1 \\ \\ M \end{pmatrix} \det(\hat{\mathbf{X}}_2^H \mathbf{X}_2) \ln^M \rho}{\det(\mathbf{P}_{\mathbf{x}_2\hat{\mathbf{x}}_2})\rho^{M}} + O(\rho^{-M} \ln^{M-1} \rho) \quad (3.10)$$

The proof of Theorem 1 is postponed until Appendix E. Some perspectives of Theorem 1 are given below:

1. Unlike the noncoherent MIMO systems, the "diversity gain" for the noncoherent cooperative relay systems involves an exponential function as well as the logarithm of SNR, which results from the effect of the Gaussian product channels on the error performance. The order of the logrithm is equal to the number of the relays nodes, whereas the order of SNR in the denominator is equal to the total number of the source-destination channel and the relay nodes. Hence, the diversity gain completely mimics the coherent cooperative relay systems, which should be called the diversity function rather than the diversity order.

2. In addition to full diversity, the coding gain exactly mimics the noncoherent MIMO system and is proportional to the determinant of the autocorrelation of the error matrix formed by distinct pairs of codeword matrices. We can use the rank and the determinant criteria to design the optimal distributed space-time block codes for the noncoherent cooperative relay systems.

3. It can be observed from the proof of Property 1 that in fact, the condition that $\mathbf{P}_{\mathbf{x}_i\hat{\mathbf{x}}_i}$ has full rank for all the distinct pairs of $\mathbf{X}_i$ and $\hat{\mathbf{X}}_i$ is a necessary and sufficient condition for the LSE receiver to extract full diversity.

Recently, Zhang, Huang and Ma [54] have proved that the unique blind identification of the channel and the transmitted signal is equivalent to full diversity for the noncoherent space-time block coded MIMO systems with the GLRT detector. This result can be extended in a straightforward manner into the noncoherent relay systems. For completeness of the exposition, this generalization is given as the following theorem and its proof is also provided.

Theorem 2 *Let each codeword matrix $\mathbf{X}_i$ have full column rank. Then, for an arbitrarily fixed nonzero received signal vector $\mathbf{z}_i = \mathbf{v}_i \neq \mathbf{0}$ without any noise; i.e., $\mathbf{v}_i = \sqrt{\rho}\mathbf{X}_i\mathbf{h}_i$, the transmitted signal codeword matrix $\mathbf{X}_i$ and the equivalent channel vector $\mathbf{h}_i$ is uniquely determined if and only if $\mathbf{P}_{\mathbf{x}_i\hat{\mathbf{x}}_i}$ has full rank for any pair of distinct codeword matrices $\mathbf{X}_i$ and $\hat{\mathbf{X}}_i$.*

PROOF: First, we prove the sufficient condition. Suppose that there exist two pairs of $\mathbf{X}_i, \mathbf{h}_i$ and $\hat{\mathbf{X}}_i, \hat{\mathbf{h}}_i$ for some nonzero received signal vector $\mathbf{v}_i$ such that

$$\mathbf{v}_i = \sqrt{\rho}\mathbf{X}_i\mathbf{h}_i = \sqrt{\rho}\hat{\mathbf{X}}_i\hat{\mathbf{h}}_i \tag{3.11}$$

then, $\mathbf{X}_i$ must be equal to $\hat{\mathbf{X}}_i$. Otherwise, if $\mathbf{X}_i \neq \hat{\mathbf{X}}_i$, then, we would have that $\mathbf{P}_{\mathbf{x}_i\hat{\mathbf{x}}_i}$ has full rank by assumption and that

$$\begin{pmatrix} \mathbf{X} & \hat{\mathbf{X}}_i \end{pmatrix} \begin{pmatrix} \mathbf{h}_i \\ \\ -\hat{\mathbf{h}}_i \end{pmatrix} = \mathbf{0} \tag{3.12}$$

Hence, (3.12) has only zero solution; i.e., $\mathbf{h}_i = \hat{\mathbf{h}}_i = \mathbf{0}$. As a result, $\mathbf{v}_i = \mathbf{0}$, which contradicts with the assumption. Therefore, $\mathbf{X}_i = \hat{\mathbf{X}}_i$ and consequently, $\mathbf{h}_i = \hat{\mathbf{h}}_i = (\mathbf{X}_i^H\mathbf{X}_i)^{-1}\mathbf{X}_i^H\mathbf{v}_i/\sqrt{\rho}$. This completes the proof of the sufficient condition.

Now, we prove the necessary condition. If there existed a pair of distinct codeword matrices $\mathbf{X}_{0,i}$ and $\hat{\mathbf{X}}_{0,i}$ such that $\mathbf{P}_{\mathbf{x}_{0,i}\hat{\mathbf{x}}_{0,i}}$ would not have full column rank, then, the following linear equations with respect to variables $\mathbf{h}_i$ and $-\hat{\mathbf{h}}_i$

$$\begin{pmatrix} \mathbf{X}_{0,i} & \hat{\mathbf{X}}_{0,i} \end{pmatrix} \begin{pmatrix} \mathbf{h}_i \\ \\ -\hat{\mathbf{h}}_i \end{pmatrix} = \mathbf{0} \tag{3.13}$$

would have a nonzero solution $\mathbf{h}_{0,i}$ and $\hat{\mathbf{h}}_{0,i}$. Let $\mathbf{v}_{0,i} = \sqrt{\rho}\mathbf{X}_{0,i}\mathbf{h}_{0,i}$. Then, we would also have $\mathbf{v}_{0,i} = \sqrt{\rho}\hat{\mathbf{X}}_{0,i}\hat{\mathbf{h}}_{0,i}$. In other words, for a given nonzero received signal vector $\mathbf{v}_{0,i}$, equation $\mathbf{v}_{0,i} = \sqrt{\rho}\mathbf{X}_i\mathbf{h}_i$ has two distinct pairs of solutions, which contradicts with the assumption. This completes the proof of Theorem 2. $\square$

Chapter 4

Uniquely-Factorable Constellation Design and Its Applications

4.1 Uniquely-Factorable Constellation for Noncoherent SIMO Channels

In this section, we first discuss the noncoherent SIMO channel model and a transmission scheme. Then, we propose a novel concept, i.e., uniquely-factorable constellation , and prove that it is this kind of the unique factorization that enables the unique blind identification of the channel coefficients in the noise-free case as well as full diversity in the noisy case.

4.1.1 Signaling scheme for SIMO channel

Let us consider wireless communication systems having a single transmitting antenna and multiple receiving antennas with flat fading. For such systems, the discrete-time baseband-equivalent channel model can be represented as

$$\mathbf{r} = \sqrt{\rho}\mathbf{h}s + \boldsymbol{\eta} \tag{4.1}$$

where ρ is the signal-to-noise ratio and the n-th element of $\mathbf{h}$ is the channel coefficient from the transmitter to the n-th receiver for $n = 1, 2, \cdots, N$. We assume that the channel coefficients are constant during two time slots, after which they randomly change to new independent values that are fixed for another two time slots, and so on. Let $\mathbb{U} \subseteq \mathbb{C}^2$ be a given complex two-dimensional constellation to be designed. Then, our noncoherent signaling scheme is now described as follows: randomly, independently and equally likely choose a vector $(x, y)^T$ from the constellation $\mathbb{U}$. During the first time slot, the signal $s = x$ is sent for transmission, i.e.,

$$\mathbf{r}_1 = \sqrt{\rho}\mathbf{h}x + \boldsymbol{\eta}_1$$

During the second time slot, the signal $s = y$ is transmitted through the channel model (4.41) for transmission, i.e.,

$$\mathbf{r}_2 = \sqrt{\rho}\mathbf{h}y + \boldsymbol{\eta}_2$$

Stacking these two received vectors yields

$$\mathbf{r} = \sqrt{\rho}\begin{pmatrix} x\mathbf{I}_N \\ \\ y\mathbf{I}_N \end{pmatrix}\mathbf{h} + \boldsymbol{\eta} = \sqrt{\rho}\mathbf{S}\mathbf{h} + \boldsymbol{\eta} \tag{4.2}$$

where $\mathbf{r} = (\mathbf{r}_1^T, \mathbf{r}_2^T)^T, \mathbf{S} = (x\mathbf{I}_N, y\mathbf{I}_N)^T$ and $\boldsymbol{\xi} = (\boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T)^T$. Now, from the standing-points of blind signal processing and detection theory, a natural question immediately comes up:

Problem 1 *Find conditions on the constellation* $\mathbb{U}$ *such that*

1. *in a noise-free case, for any given nonzero received signal vector* $\mathbf{r} \neq \mathbf{0}$*, the equation reduced from* (4.2)

$$\mathbf{r} = \sqrt{\rho}\mathbf{S}\mathbf{h} \tag{4.3}$$

   with respect to the transmitted symbol variables x *and* y*, and the channel vector* $\mathbf{h}$ *has a unique solution, and*

2. *in a noisy environment, full diversity is enabled for the generalized likelihood ratio test receiver.*

4.1.2 Unique identification and full diversity

First, let us attempt to answer the first question of Problem 1 on the unique identification. In this case when the noise is free, then, during the first time slot, the n-th received signal u_n can be expressed by

$$u_n = \sqrt{\rho}h_n x \tag{4.4a}$$

whereas during the second time slot, the n-th received signal v_n can be written as

$$v_n = \sqrt{\rho}h_n y \tag{4.4b}$$

Eliminating h_n from (4.4) results in

$$\frac{u_n}{v_n} = \frac{x}{y} \tag{4.5}$$

Hence, for any given $\frac{u_n}{v_n}$, (4.5) has a unique solution with respect to x and y if and only if the constellation $\mathbb{U}$ must satisfy such a condition that if $x\hat{y} = \hat{x}y$, then, $x = \hat{x}$ and $y = \hat{y}$, i.e., the *unique factorization of the constellation.*

Now, let us answer the second question of Problem 1 regarding the noncoherent full diversity. According to (2.7), if the GLRT receiver is employed at the receiver in this SIMO system, the pairwise error probability $\mathrm{P}_{\mathrm{GLRT}}(\mathbf{S} \to \hat{\mathbf{S}})$ of transmitting $\mathbf{S}$ and deciding in favor of $\hat{\mathbf{S}} \neq \mathbf{S}$ has the asymptotic formula below:

$$\mathrm{P}_{\mathrm{GLRT}}(\mathbf{S} \to \hat{\mathbf{S}}) = \frac{\begin{pmatrix} 2N-1 \\ \\ N \end{pmatrix} \det^N(\hat{\mathbf{S}}^H\hat{\mathbf{S}})}{\det^N(\mathbf{P}_{\mathbf{s}\hat{\mathbf{s}}})} \times \rho^{-N} + O\big(\rho^{-N}\big)$$

if $\det(\mathbf{P}_{\mathbf{s}\hat{\mathbf{s}}}) \neq 0$. It is not difficult to prove that the statement that $\mathbf{P}_{\mathbf{s}\hat{\mathbf{s}}}$ has full rank for all pairs of distinct codewords $\mathbf{S}$ and $\hat{\mathbf{S}}$ is equivalent to the one that $(\mathbf{S}, \hat{\mathbf{S}})$ has full rank for all pairs of distinct codewords $\mathbf{S}$ and $\hat{\mathbf{S}}$. Hence, the full diversity condition is equivalent to the one that

$$(\mathbf{S}, \hat{\mathbf{S}}) = \begin{pmatrix} x\mathbf{I}_N & \hat{x}\mathbf{I}_N \\ \\ y\mathbf{I}_N & \hat{y}\mathbf{I}_N \end{pmatrix} \tag{4.6}$$

is invertible for any $(x, y)^T \neq (\hat{x}, \hat{y})^T \in \mathbb{U}$. After some algebraic manipulations, we can obtain that the determinant of $(\mathbf{S}, \hat{\mathbf{S}})$ is given by

$$\det\big((\mathbf{S}, \hat{\mathbf{S}})\big) = \big(x\hat{y} - \hat{x}y\big)^N \tag{4.7}$$

Therefore, the full diversity condition is also reduced to the one that for any $(x, y)^T \neq (\hat{x}, \hat{y})^T \in \mathbb{U}$, $x\hat{y} \neq \hat{x}y$, which is equivalent to saying that if $x\hat{y} = \hat{x}y$, then, $x = \hat{x}$ and $y = \hat{y}$, i.e., the *unique factorization of the constellation.*

All the above discussions can be summarized as the following theorem.

Theorem 3 *(Unique Identification and full diversity for SIMO channels) Let $\mathbb{U}$ be a given complex two-dimensional constellation with $|\mathbb{U}| > 1$, and $\mathbf{u}$ and $\mathbf{v}$ be two received signal vectors in the first two time slots transmission from the channel model (4.2) in a noise-free environment; i.e.,*

$$\mathbf{u} = \sqrt{\rho}x\mathbf{h}, \tag{4.8}$$

$$\mathbf{v} = \sqrt{\rho}y\mathbf{h} \tag{4.9}$$

for $(x, y)^T \in \mathbb{U}$. Then, the following three statements are equivalent.

1. *For the arbitrarily given nonzero received signal vector $(\mathbf{u}^T, \mathbf{v}^T)^T$, the channel*

vector $\mathbf{h}$ *and the transmitted symbols* x *and* y *can be uniquely determined.*

2. *In a Gaussian noise environment,* $\mathbb{U}$ *enables full diversity for the GLRT receiver.*

3. $\mathbb{U}$ *satisfies a so-called unique factorization property, i.e.,* $x\tilde{y} \neq \tilde{x}y$ *for any* $(x, y)^T \neq (\tilde{x}, \tilde{y})^T \in \mathbb{U}$.

4.1.3 Uniquely factorable constellation

Theorem 3 motivates us to formally introduce the following new concept:

Definition 1 *Let* $\mathbb{U}$ *be a set composed of some two-dimensional complex column vectors. Then,* $\mathbb{U}$ *is said to form a uniquely-factorable constellation (UFC) if there exist* $(x, y)^T, (\tilde{x}, \tilde{y})^T \in \mathbb{U}$ *satisfying* $x\tilde{y} = \tilde{x}y$*, then, we have* $x = \tilde{x}$ *and* $y = \tilde{y}$.

Example 1 *Here is a UFC example with 2 vectors:*

$$\mathbb{U} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\} \tag{4.10}$$

Example 2 *Here is another UFC example with 4 vectors:*

$$\mathbb{U} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ j \end{pmatrix}, \begin{pmatrix} -j \\ 1 \end{pmatrix}, \begin{pmatrix} j \\ -j \end{pmatrix} \right\} \tag{4.11}$$

Example 3 *(Training UFC). This example shows that the constellation in the traditional training scheme for the scalar AWGN channel is actually a special UFC. Let* $\mathcal{Y}$ *be an arbitrarily given one-diemsional complex constellation. Then, for any fixed nonzero complex number* x_0*, the constellation* $\mathbb{U} = \{(x_0, y)^T : y \in \mathcal{Y}\}$ *derived from* $\mathcal{Y}$ *forms a UFC, since if* $(x_0, y)^T, (x_0, \widetilde{y})^T \in \mathbb{U}$ *and* $x_0\widetilde{y} = x_0 y$*, then,* $y = \widetilde{y}$*.*

In order to systematically design a UFC, we need to develop some necessary conditions which a UFC must satisfy.

Proposition 6 *Let* $\mathbb{U}$ *be a UFC with* $|\mathbb{U}| \geq 2$*. Then, the following statements are true.*

1. $(0, 0)^T \notin \mathbb{U}$.
2. *If* $(x, y)^T \in \mathbb{U}$*, then,* $-(x, y)^T \notin \mathbb{U}$.
3. *If* $(0, y_1)^T \in \mathbb{U}$*, then, for any complex number* $y_2 \neq y_1$*,* $(0, y_2)^T \notin \mathbb{U}$*. Similarly, if* $(x_1, 0)^T \in \mathbb{U}$*, then, for any complex number* $x_2 \neq x_1$*,* $(x_2, 0)^T \notin \mathbb{U}$.
4. *If* $(x, x)^T \in \mathbb{U}$*, then, for any complex number* $y \neq x$*,* $(y, y)^T \notin \mathbb{U}$.

Proof: Statement 1. Since $|\mathbb{U}| \geq 2$, there exists a nonzero vector $(x_0, y_0)^T \in \mathbb{U}$. Now, suppose that $(0, 0)^T \in \mathbb{U}$. Then, we would have $0 \times y_0 = x_0 \times 0 = 0$, but one of x_0 and y_0 is not zero, which contradicts with the assumption that $\mathbb{U}$ is the UFC.

Statement 2. If there existed $(x, y)^T \in \mathbb{U}$ such that $(-x, -y)^T \in \mathbb{U}$, then, we would have $xy = (-x)(-y)$, but $(x, y) \neq (-x, -y)$, since we just know from the Statement 1 that the zero vector does not belong to $\mathbb{U}$.

Statement 3. Assume that $(0, y_1)^T, (0, y_2)^T \in \mathbb{U}$ with $y_1 \neq y_2$. Then, $0 \times y_2 = 0 \times y_1 = 0$ with $y_1 \neq y_2$, which contradicts with the assumption that $\mathbb{U}$ is the UFC. Similarly, we can prove that at least, one of $(x_1, 0)^T$ and $(x_2, 0)^T$ cannot belong to $\mathbb{U}$.

Statement 4. Suppose that $(x, x)^T, (y, y)^T \in \mathbb{U}$ with $x \neq y$. Since $x \times y = y \times x$ with $x \neq y$, $\mathbb{U}$ is not any UFC. This is a contradiction, which completes the proof of Proposition 6. $\square$

4.1.4 Unitary UFC and coding gain

In recent years, extensive research work on noncoherent space-time block codes has been mainly focused on unitary designs [37–42, 44–46, 55], since unitary constellations are optimal [14, 26, 37, 43] when either SNR is high or coherent time is long. It is known that the Cayley transform [44, 46, 55] and the exponential transform [42] are two well-established transforms that convert respective linear dispersion codes and linear space-time block codes into unitary codes for a general MIMO channel. However, particularly for the noncoherent SIMO channel, a unitary constellation can be immediately attained by simply normalizing the nonunitary constellation., i.e., for a given constellation $\mathbb{X} \subseteq \mathbb{C}^2$ with $\mathbf{0} \notin \mathbb{X}$, its unitary constellation, denoted by $\overline{\mathbb{X}}$, is given by

$$\overline{\mathbb{X}} = \left\{ \bar{\mathbf{x}} = \frac{\mathbf{x}}{\|\mathbf{x}\|} : \mathbf{x} \in \mathbb{X} \right\}$$

By Definition 1, we know that $\mathbb{X}$ is a UFC if and only if $\overline{\mathbb{X}}$ is a UFC. Now, applying (2.7) to the constellation $\overline{\mathbb{X}}$ yields

$$\mathrm{P}_{\mathrm{GLRT}}(\bar{\mathbf{x}} \to \hat{\bar{\mathbf{x}}}) = \frac{\begin{pmatrix} 2N-1 \\ N \end{pmatrix}}{\det^N(\mathbf{P}_{\bar{\mathbf{x}}\hat{\bar{\mathbf{x}}}})} \times \rho^{-N} + O\big(\rho^{-N}\big)$$

if $\det(\mathbf{P}_{\bar{\mathbf{x}}\hat{\bar{\mathbf{x}}}}) \neq 0$. Therefore, when SNR is large, the error performance is dominated by the worse case of $\det(\mathbf{P}_{\bar{\mathbf{x}}\hat{\bar{\mathbf{x}}}})$. Following the way similar to coherent MIMO communications [56], we define the coding gain for the unitary constellation $\overline{\mathbb{X}}$ as

$$G(\overline{\mathbb{X}}) = \min_{\bar{\mathbf{x}} \neq \hat{\bar{\mathbf{x}}} \in \overline{\mathbb{X}}} \det(\mathbf{P}_{\bar{\mathbf{x}}\hat{\bar{\mathbf{x}}}}) \tag{4.12}$$

In addition, notice

$$\det(\mathbf{P}_{\bar{\mathbf{x}}\hat{\bar{\mathbf{x}}}}) = |\det((\bar{\mathbf{x}}, \hat{\bar{\mathbf{x}}}))|^2 = \frac{|\det((\mathbf{x}, \hat{\mathbf{x}}))|^2}{\|\mathbf{x}\|^2\|\hat{\mathbf{x}}\|^2}$$

Hence, for discussion convenience, we particularly introduce a definition below.

Definition 2 *A distance between any two vectors* $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}^2$*,* $d(\mathbf{x}_1, \mathbf{x}_2)$*, is defined as*

$$d(\mathbf{x}_1, \mathbf{x}_2) = \frac{|\det(\mathbf{X})|}{||\mathbf{x}_1||\ ||\mathbf{x}_2||} \tag{4.13}$$

where $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2)$ *is a 2-by-2 matrix formed by these two column vectors.*

Definition 3 *The distance of a given constellation $\mathbb{X} \subseteq \mathbb{C}^2$, which is denoted by $D(\mathbb{X})$, is defined as the minimum distance of any two distinct vectors within this constellation, i.e.,*

$$D(\mathbb{X}) = \min_{\mathbf{x}_1, \mathbf{x} \in \mathbb{X}, \mathbf{x}_1 \neq \mathbf{x}_2} d(\mathbf{x}_1, \mathbf{x}_2)$$

Thus, the distance of $\mathbb{X}$ is nothing but the square root of the coding gain for its normalized constellation $\overline{\mathbb{X}}$, i.e., $D(\mathbb{X}) = \sqrt{G(\overline{\mathbb{X}})}$. Hereafter, we mutually use these two concepts in this thesis wherever it is convenient.

4.2 Energy-Efficient Unitary Training Scheme

The main purpose of this section is to simply attain a unitary constellation design by just normalizing the nonunitary training constellation based on the QAM constellation and then, find a closed-form energy scale to maximize the coding gain or distance.

4.2.1 Training scheme

When channel information is not available at either the transmitter or the receiver, a simple and practical way to estimate the channel coefficients is to send training signals. Particularly for the SIMO channel, only one bit is needed for the channel estimation. Hence, the constellation of the training scheme can be represented as

$$\mathbb{T}_{\mathcal{Q}_p} = \left\{ \mathbf{t} \,\middle|\, \mathbf{t} = \begin{pmatrix} 1 \\ \\ t \end{pmatrix} \right\}$$

where t is randomly, independently and equally likely chosen from a certain constellation. In this thesis, we focus on the case when t is randomly, independently and equally likely chosen from the 2^p-ary cross QAM constellation $\mathcal{Q}_p$, a formal definition of which, for the completeness of the exposition, is provided here.

Definition 4 *A cross 2^p-ary QAM constellation $\mathcal{Q}_p$ is defined as follows:*

1. *If p is even, $\mathcal{Q}_p$ is the standard square 2^p-ary QAM constellation, i.e.,*

$$\mathcal{Q}_n = \left\{(2m-1)+(2n-1)j : -2^{\frac{p-2}{2}}+1 \leq m,n \leq 2^{\frac{p-2}{2}}\right\}$$

2. *If $p=3$, then,*

$$\mathcal{Q}_3 = \left\{3+j, 1+j, -1+j, -3+j, -3-j, -1-j, 1-j, 3-j\right\}$$

3. *If p is an odd number exceeding 3, $\mathcal{Q}_p$ is the union of a horizontal rectangular QAM constellation and a vertical rectangular QAM constellation, i.e.,*

$$\begin{aligned}\mathcal{Q}_p = &\left\{(2m-1)+(2n-1)j : -3\times 2^{\frac{p-5}{2}}+1 \leq m \leq 3\times 2^{\frac{p-5}{2}}, -2^{\frac{p-3}{2}}+1 \leq n \leq 2^{\frac{p-3}{2}}\right\}\\ &\bigcup\left\{(2m-1)+(2n-1)j : -2^{\frac{p-3}{2}}+1 \leq m \leq 2^{\frac{p-3}{2}}, -3\times 2^{\frac{p-5}{2}}+1 \leq n \leq 3\times 2^{\frac{p-5}{2}}\right\}\end{aligned}$$

Based on this definition, we can immediately obtain the following lemma.

Lemma 3 *For the 2^p-ary cross QAM $\mathcal{Q}_p$ given in Definition 4, let P denote one of the corner points of $\mathcal{Q}_p$ with the largest energy E_1. If we let $\widetilde{P}$ denote such a nearest neighbor of P that it has the largest energy E_2, then, the following three statements are true:*

1. *If* $p = 3$, *then,* $E_1 = E_2 = 10$.

2. *If* p *is even, then,* $E_1 = 2(2^{\frac{p}{2}} - 1)^2$ *and* $E_2 = (2^{\frac{p}{2}} - 1)^2 + (2^{\frac{p}{2}} - 3)^2$.

3. *If* p *is odd and greater than* 3, *then,* $E_1 = (2^{\frac{p-1}{2}} - 1)^2 + (3 \times 2^{\frac{p-3}{2}} - 1)^2$ *and* $E_2 = (2^{\frac{p-1}{2}} - 3)^2 + (3 \times 2^{\frac{p-3}{2}} - 1)^2$.

Lemma 3 can be verified directly by calculation and thus, its proof is omitted.

4.2.2 Energy-efficient unitary training scheme

Particularly for the noncoherent SIMO channel, the unitary constellation of the training constellation $\mathbb{T}_{\mathcal{Q}_p}$ can be directly obtained by simply normalizing the nonunitary training constellation, i.e.,

$$\overline{\mathbb{T}}_{\mathcal{Q}_p} = \left\{ \frac{1}{\sqrt{1+|t|^2}} \begin{pmatrix} 1 \\ \\ t \end{pmatrix} : t \in \mathcal{Q}_p \right\}$$

Let us now consider an energy scaled version of the training constellation, i.e.,

$$\mathbb{T}_{\mathcal{Q}_p}(\alpha) = \left\{ \mathbf{t}' \,\middle|\, \mathbf{t}' = \begin{pmatrix} 1 \\ \\ \alpha t \end{pmatrix} : t \in \mathcal{Q}_p \right\}$$

for $\alpha > 0$. Then, the corresponding normalized constellation is given by

$$\overline{\mathbb{T}}_{\mathcal{Q}_p}(\alpha) = \left\{ \frac{1}{\sqrt{1+\alpha^2|t|^2}} \begin{pmatrix} 1 \\ \\ \alpha t \end{pmatrix} : t \in \mathcal{Q}_p \right\} \tag{4.15}$$

According to Definition 2, the distance between $\mathbf{t}'_1$ and $\mathbf{t}'_2$ is given by

$$d(\mathbf{t}'_1, \mathbf{t}'_2) = \frac{\alpha|t_1 - t_2|}{\sqrt{1+\alpha^2|t_1|^2}\sqrt{1+\alpha^2|t_2|^2}} \triangleq d(t_1, t_2, \alpha) \tag{4.16}$$

for $t_1, t_2 \in \mathcal{Q}_p$, which is called a *distance function* or *coding gain function.* Here, we are interested in solving the following optimization problem:

Problem 2 *Find an energy scale α such that $D(\mathbb{T}_{\mathcal{Q}_p}(\alpha))$ is maximized, or equivalently, the coding gain of $\overline{\mathbb{T}}_{\mathcal{Q}_p}(\alpha)$ is maximized, i.e.,*

$$\begin{aligned} \hat{\alpha} &= \arg\max_{\alpha} D(\mathbb{T}_{\mathcal{Q}_p}(\alpha)) \\ &= \arg\max_{\alpha} \min_{t_1 \neq t_2 \in \mathcal{Q}_p} d(t_1, t_2, \alpha) \end{aligned} \tag{4.17}$$

The solution to Problem 2 is given in Theorem 4.

Theorem 4 *The optimal solution to Problem 1 is given by*

$$\hat{\alpha} = \frac{1}{\sqrt[4]{E_1 E_2}} \tag{4.18}$$

$$D(\mathbb{T}_{\mathcal{Q}_p}(\hat{\alpha})) = \frac{2}{\sqrt{E_1} + \sqrt{E_2}} \tag{4.19}$$

where E_1 and E_2 are determined as follows:

$$\begin{cases} E_1 = E_2 = 10 \qquad \text{if } p = 3 \\ E_1 = 2(2^{\frac{p}{2}} - 1)^2, E_2 = (2^{\frac{p}{2}} - 1)^2 + (2^{\frac{p}{2}} - 3)^2 \text{ if } p \text{ is even} \\ E_1 = (2^{\frac{p-1}{2}} - 1)^2 + (3 \times 2^{\frac{p-3}{2}} - 1)^2, E_2 = (2^{\frac{p-1}{2}} - 3)^2 + (3 \times 2^{\frac{p-3}{2}} - 1)^2 \text{ if } p > 3 \text{ is odd} \end{cases}$$

PROOF: The proof of Theorem 4 is composed of the following two steps.

Step 1: Solve the inner minimization problem. In order to attain an optimal solution to the inner minimization problem, we split the original feasible domain into the following two disjoint sub-domains:

$$\begin{aligned} \mathcal{D} &= \{(t_1, t_2) : t_1 \neq t_2 \in \mathcal{Q}_p\} \\ &= \{(t_1, t_2) : (|t_1|, |t_2|) \neq (E_1, E_1), t_1 \neq t_2 \in \mathcal{Q}_p\} \\ &\quad \cup \{(t_1, t_2) : (|t_1|, |t_2|) = (E_1, E_1), t_1 \neq t_2 \in \mathcal{Q}_p\} \\ &= \mathcal{D}_1 \cup \mathcal{D}_2 \end{aligned}$$

Therefore, solving the original inner minimization problem is equivalent to first solving the corresponding two sub-minimization problems and then, taking the minimum of these two minimums, i.e.,

$$\min_{(t_1,t_2)\in\mathcal{D}} d(t_1, t_2, \alpha) = \min\{\min_{(t_1,t_2)\in\mathcal{D}_1} d(t_1, t_2, \alpha), \min_{(t_1,t_2)\in\mathcal{D}_2} d(t_1, t_2, \alpha)\} \tag{4.20}$$

1. Solution to minimization problem: $\min_{(t_1,t_2)\in\mathcal{D}_1} d(t_1, t_2, \alpha)$. Let us first reveal

some very interesting optimality properties on the fractional objective funcation.

(a) The numerator of $d(t_1, t_2, \alpha)$ is lower bounded by

$$\alpha|t_1 - t_2| \geq 2\alpha \tag{4.21}$$

where the equality holds, i.e., the minimum of the numerator is achieved when t_1 and t_2 are the nearest neighbors each other.

(b) Since $(t_1, t_2) \neq (E_1, E_1)$, the denominator of $d(t_1, t_2, \alpha)$ is upper bounded by

$$\sqrt{1 + \alpha^2|t_1|^2}\sqrt{1 + \alpha^2|t_2|^2} \leq \sqrt{1 + \alpha^2 E_1}\sqrt{1 + \alpha^2 E_2} \tag{4.22}$$

where the equality holds, i.e., the maximum of the denominator is achieved when one of t_1 and t_2 is located in the corners of the QAM constellation with the largest energy and the other has the second largest energy.

Hence, the above two observations naturally come up with an interesting question: *When do both inequalities* (4.21) *and* (4.22) *hold simultaneously*? It is very amazing to observe that the answer to this question is that one of t_1 and t_2 is the corner point with the largest energy E_1 and the other is its nearest neighbor with the second largest energy E_2. Therefore, the minimum of $d(t_1, t_2, \alpha)$ is given by

$$\min_{(t_1,t_2)\in\mathcal{D}_1} d(t_1, t_2, \alpha) = \frac{2\alpha}{\sqrt{(1 + \alpha^2 E_1)(1 + \alpha^2 E_2)}} \tag{4.23}$$

2. Solution to minimization problem: $\min_{(t_1,t_2)\in\mathcal{D}_2} d(t_1,t_2,\alpha)$. Note that in this case, the denominator of $d(t_1,t_2,\alpha)$ is constant, i.e., $\sqrt{1+\alpha^2|t_1|^2}\sqrt{1+\alpha^2|t_2|^2} = 1+\alpha^2 E_1$. Hence, to solve the problem, we only need to consider the following cases on the numerator:

   (a) Either $p=2$ or $p=3$. In this case, it is very interesting to observe that the lower-bound (4.21) and the upper-bound (4.22) are achieved simultaneously when t_1 is one of the corner point and t_2 is its nearest neighbor, i.e., $|t_1-t_2|=2$. Hence, we have

$$\min_{(t_1,t_2)\in\mathcal{D}_2} d(t_1,t_2,\alpha) = \frac{2\alpha}{1+\alpha^2 E_1} \tag{4.24}$$

   (b) $p=2k$ is even, where $k\geq 2$. In this case, the numerator achieves its minimum, $\sqrt{2E_1}\alpha$. As a result, we obtain

$$\min_{(t_1,t_2)\in\mathcal{D}_2} d(t_1,t_2,\alpha) = \frac{\sqrt{2E_1}\alpha}{1+\alpha^2 E_1} \tag{4.25}$$

   (c) $p=2k+1$ is odd, where $k\geq 2$. Notice that when t_1 and t_2 are located in the same quadrant with the largest energy, the numerator reaches its minimum, $2^{\frac{p-2}{2}}\alpha$. Thus, we have

$$\min_{(t_1,t_2)\in\mathcal{D}_2} d(t_1,t_2,\alpha) = \frac{2^{\frac{p-2}{2}}\alpha}{1+\alpha^2 E_1} \tag{4.26}$$

Now, substituting (4.23), (4.24), (4.25) and (4.26) into (4.20) yields

$$\begin{aligned}
\min_{(t_1,t_2)\in\mathcal{D}} d(t_1,t_2,\alpha) &= \begin{cases} \min\{\frac{2\alpha}{\sqrt{(1+\alpha^2E_1)(1+\alpha^2E_2)}}, \frac{2\alpha}{1+\alpha^2E_1}\} & \text{if } p=2 \text{ or } p=3 \\ \min\{\frac{2\alpha}{\sqrt{(1+\alpha^2E_1)(1+\alpha^2E_2)}}, \frac{\sqrt{2E_1}\alpha}{1+\alpha^2E_1}\} & \text{if } p=2k, k\geq 2 \\ \min\{\frac{2\alpha}{\sqrt{(1+\alpha^2E_1)(1+\alpha^2E_2)}}, \frac{2^{\frac{p-2}{2}}\alpha}{1+\alpha^2E_1}\} & \text{if } p=2k+1, k\geq 2 \end{cases} \\
&= \begin{cases} \frac{2\alpha}{1+\alpha^2E_1} & \text{if } p=2 \text{ or } p=3 \\ \frac{2\alpha}{\sqrt{(1+\alpha^2E_1)(1+\alpha^2E_2)}} & \text{if } p=2k, k\geq 2 \\ \frac{2\alpha}{\sqrt{(1+\alpha^2E_1)(1+\alpha^2E_2)}} & \text{if } p=2k+1, k\geq 2 \end{cases} \\
&= \frac{2\alpha}{\sqrt{(1+\alpha^2E_1)(1+\alpha^2E_2)}}
\end{aligned}$$

Step 2: Solve the outer maximization problem. We know from Step 1 that the distance of the scaled version of the QAM constellation can be rewritten as

$$D(\mathbb{T}_{\mathcal{Q}_p}(\alpha)) = \frac{2}{\sqrt{\alpha^{-2}+E_1+E_2+E_1E_2\alpha^2}} \tag{4.27}$$

Using the geometrical and arithmetical mean inequality: $A^2+B^2 \geq 2AB$, we can derive from (4.27)

$$D(\mathbb{T}_{\mathcal{Q}_p}(\alpha)) \leq \frac{2}{\sqrt{E_1+E_2+2\sqrt{E_1E_2}}} = \frac{2}{\sqrt{E_1}+\sqrt{E_2}}$$

where the equality holds, i.e., the maximum is attained when $\alpha^{-2} = E_1 E_2 \alpha^2$. This is equivalent to

$$\hat{\alpha} = \frac{1}{\sqrt[4]{E_1 E_2}}$$

Hence, the resulting maximum is given by

$$D(\mathbb{T}_{\mathcal{Q}_p}(\hat{\alpha})) = \frac{2}{\sqrt{E_1} + \sqrt{E_2}}$$

This completes the proof of Theorem 4.

Some observations on Theorem 4 are made as follows:

1. It is worth emphasizing a very interesting optimality feature on the fractional distance function (4.16) (or the coding gain function), i.e., the coding gain is attained when the numerator of the objective achieves its minimum, whereas the denominator achieves its second maximum.

2. In spite of the fact that the QAM constellation is generally regarded as a good constellation for the modern digital communication systems, the aforementioned Observation 1) explicitly reveals its drawback for the nocoherent SIMO channel, i.e., the minimum Euclidean distance between the signal points with large energies is the same as the minimum Euclidean distance between the signal points with small energies.

4.3 Energy-Efficient Unitary Uniquely-Factorable Constellations

Just as we have pointed out in the comments of Theorem 4, a good constellation for the noncoherent SIMO channel should be such a constellation that its minimum Euclidean distance is supposed to increase as the energies of the signal points become large. It is this motivation that leads us to developing two algorithms to efficiently and effectively design the UFCs.

4.3.1 Lagrange's four-square theorem

First, we introduce a very famous theorem in additive number theory, the Lagrange's four-square theorem. Lagrange discovered and proved that every positive integer can be represented as the sum of four squares, i.e.,

$$N = a^2 + b^2 + c^2 + d^2 \tag{4.28}$$

where a, b, c, d are integers. A fast algorithm is provided in [57] to find a solution to (4.28). If we let $r_4(N)$ denote the number of solutions to (4.28), then,

$$r_4(N) = 8 \sum_{m|N, 4\nmid m} m \tag{4.29}$$

which is the summation of all divisors of N not divisible by 4. It has been proved that $r_4(N)/8$ is a multiplicative function. Hence, if we let

$$N = 2^{k_0} \times p_1^{k_1} \times p_2^{k_2} \times \cdots \times p_n^{k_n}$$

where $p_1, p_2, \cdots, p_n$ are all odd prime numbers, k_0 is a nonnegative integer, $k_1, \cdots, k_n$ are positive integers, then, the number of different representations $r_4(N)$ is also given by

$$r_4(N) = 24 \prod_{i=1}^{n} \frac{1 - p_i^{k_i+1}}{1 - p_i} \tag{4.30}$$

See more details in [58] on this theorem and some related results.

4.3.2 UFC constructions

Our primary target in this subsection is to find an efficient and effective method for the design of a UFC, $\mathbb{U}_n$, with a given size of 2^n. Theoretically speaking, we should find a UFC from the complex two-dimensional plane $\mathbb{C}^2$ such that its distance is maximized. However, just as we have mentioned in the introduction, this design problem is extremely difficult to be formulated into a tractable optimization problem, even for the AWGN channel [27–33]. Therefore, the basic idea of efficiently and effectively finding a good UFC here is to use the Lagrange's four-square theorem. We start with $N = 2$ and find all the solutions to (4.28). Then, all the combinations of all the solutions a, b, c and d form $(x, y)^T$ as all the possible candidates. Now, by Proposition 6 and the following two rules:

<u>Rule 1</u>: $(0, y)^T \notin \mathbb{U}_n$, since we do not require that infinity belongs to the training-equivalent UFC (see the definition of the training-equivalent UFC in Subsection 4.3.3),

<u>Rule 2</u>: $(x, 0)^T \notin \mathbb{U}_n$, since we do not require that zero belongs to the training-equivalent UFC

we select out such a set of 4 vectors that its distance is maximized. This set is designated as the first UFC $\mathbb{U}_2$ with 4 symbols. Then, we increase N by one and use the the Lagrange's four-square theorem to add another 4 vectors to $\mathbb{U}_2$ so that the resulting set of 8 vectors, which is denoted by $\mathbb{U}_3$, has the maximum distance among all such additions. Continue this process until $\mathbb{U}_n$ with size 2^n is constructed. All these procedures can be summarized as an algorithm below.

Algorithm 1 *This algorithm consists of the following six progressive steps:*

1. *Initialize* $n = 2$, *and* $\mathbb{W} = \Phi$.

2. *Given* $N = 2$, *find all possible combinations of* (a, b, c, d) *to generate vectors according to Rules 1 and 2.*

3. *Check every vector. If the distance between this vector and any other vector already in* $\mathbb{W}$ *is not equal to 0, i.e., the unique factorization condition is satisfied, then, add it into* $\mathbb{W}$.

4. *Find all the subsets* $\mathbb{X}_n$ *of* $\mathbb{W}$ *with size* 2^n *such that* $\mathbb{U}_{n-1} \subseteq \mathbb{X}_n$ *and* $D(\mathbb{X}_n) > D(\mathbb{T}_{\mathcal{Q}_n}(\hat{\alpha}))$, *go to 5). Otherwise, go back to 2) and increase* N *by 1.*

5. *Among all the candidates* $\mathbb{X}_{n_1}, \mathbb{X}_{n_2}, \cdots, \mathbb{X}_{n_M}$, *select the index* n_m *such that* $\mathbb{X}_{n_m}$ *has the largest distance and then, let* $\mathbb{U}_n = \mathbb{X}_{n_m}$.

6. *Go to 2) and increase* n *by 1.*

In Step 5), if there is more than one candidate having the largest distance, then, choose the one that makes the corresponding training-equivalent UFC as symmetric as possible. Some UFCs with 4, 8, 16, 32 and 64 symbols designed using Algorithm 1 are given in the Appendix F. Other larger sizes of UFCs can be found in [59]. All these constellations enjoy some nice geometrical properties described as follows.

Proposition 7 *Let the UFC $\mathbb{U}_n$ be designed by Algorithm 1. If we let*

$$\mathcal{Z}_n = \left\{ \frac{y}{x} : (x, y)^T \in \mathbb{U}_n \right\} \tag{4.31}$$

then, $\mathcal{Z}_n$ satisfies a certain of rotation-invariant properties, i.e., if $z \in \mathcal{Z}_n$, then, $e^{j\theta} z \in \mathcal{Z}_n$, where $\theta = \pi, \pm\pi/2$.

4.3.3 Unitary and Training-Equivalent UFCs

After we have designed a UFC $\mathbb{U}_n$, a unitary UFC is immediately obtained by normalizing the nonunitary UFC $\mathbb{U}_n$, i.e.,

$$\overline{\mathbb{U}}_n = \left\{ \frac{1}{\sqrt{|x|^2 + |y|^2}} \begin{pmatrix} x \\ \\ \\ y \end{pmatrix} : \ (x, y)^T \in \mathbb{U}_n \right\} \tag{4.32}$$

In addition, a new constellation resulting from $\mathbb{U}_n$ is defined as

$$\mathbb{T}_{\mathcal{Z}_n} = \left\{ \mathbf{z} = \begin{pmatrix} 1 \\ \\ \\ z \end{pmatrix} : z = \frac{y}{x}, (x, y)^T \in \mathbb{U}_n \right\}$$

$\mathbb{T}_{\mathcal{Z}_n}$ is nothing but the collection of every vector in $\mathbb{U}_n$ divided by its first element. As a result, the first element in the new constellation is always 1, which is exactly the same as training scheme using the one-dimensional constellation $\mathcal{Z}_n$ and thus, is called a training-equivalent constellation, whereas $\mathcal{Z}_n$ itself is called a training-equivalent

UFC. In addition, since the one-dimensional constellation $\mathcal{Z}_n$ can be plotted on a complex plane, it provides us with intuitive understanding of the constellation. So, by symmetry mentioned in Subsection 4.3.2, we mean that the constellation $\mathcal{Z}_n$ is geometrically symmetric. In addition, notice that the distance between any two vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ in the training-equivalent constellation $\mathbb{T}_{\mathcal{Z}_n}$ is

$$\begin{aligned} d(\mathbf{z}_1, \mathbf{z}_2) &= \frac{|y_1/x_1 - y_2/x_2|}{\sqrt{1+|y_1/x_1|^2}\sqrt{1+|y_2/x_2|^2}} \\ &= \frac{|y_1x_2 - y_2x_1|}{\sqrt{|x_1|^2+|y_1|^2}\sqrt{|x_2|^2+|y_2|^2}} \\ &= d(\mathbf{u}_1, \mathbf{u}_2) \end{aligned}$$

for $\mathbf{u}_1 = (x_1, y_1)^T, \mathbf{u}_2 = (x_2, y_2)^T \in \mathbb{U}$. Hence, the distance is maintained, i.e.,

$$D(\mathbb{T}_{\mathcal{Z}_n}) = D(\mathbb{U}_n)$$

It is for this reason that from now on, we will work only on the training-equivalent constellation instead of the original one.

4.3.4 Optimal unitary UFC

Similar to the definitions of $\mathbb{T}_{\mathcal{Q}_p}(\alpha)$ and $\overline{\mathbb{T}}_{\mathcal{Q}_p}(\alpha)$, the energy-scaled version of the training-equivalent UFC, $\mathbb{T}_{\mathcal{Z}_n}(\beta)$, and its unitary UFC, $\overline{\mathbb{T}}_{\mathcal{Z}_n}(\beta)$, are defined, respectively, as

$$\begin{aligned}
\mathbb{T}_{\mathcal{Z}_n}(\beta) &= \left\{ \begin{pmatrix} 1 \\ \\ \beta z \end{pmatrix} : z \in \mathcal{Z}_n \right\} \\
\overline{\mathbb{T}}_{\mathcal{Z}_n}(\beta) &= \left\{ \frac{1}{\sqrt{1+\beta^2|z|^2}} \begin{pmatrix} 1 \\ \\ \beta z \end{pmatrix} : z \in \mathcal{Z}_n \right\}
\end{aligned}$$

Then, the distance between any two vectors $\mathbf{z}'_1$ and $\mathbf{z}'_2$ in $\mathbb{T}_{\mathcal{Z}_n}(\beta)$ is equal to

$$d(\mathbf{z}'_1, \mathbf{z}'_2) = \frac{\beta|z_1 - z_2|}{\sqrt{1+\beta^2|z_1|^2}\sqrt{1+\beta^2|z_2|^2}} \triangleq d(z_1, z_2, \beta) \tag{4.33}$$

Hence, we desire to solve the following optimization problem:

Problem 3 *Find an energy scale β such that the distance of the constellation $\mathbb{T}_{\mathcal{Z}_n}(\beta)$ is maximized, i.e.,*

$$\begin{aligned}
\hat{\beta} &= \arg\max_{\beta} D(\mathbb{T}_{\mathcal{Z}_n}(\beta)) \\
&= \arg\max_{\beta} \min_{z_1 \neq z_2 \in \mathcal{Z}_n} \frac{\beta|z_1 - z_2|}{\sqrt{1+\beta^2|z_1|^2}\sqrt{1+\beta^2|z_2|^2}} \\
&= \arg\max_{\beta} \min_{z_1 \neq z_2 \in \mathcal{Z}_n} d(z_1, z_2, \beta)
\end{aligned} \tag{4.34}$$

In order to efficiently solve Problem 3, we first notice that the objective function $d(z_1, z_2, \beta)$ is symmetric with respect to the variables z_1 and z_2, i.e., $d(z_1, z_2, \beta) = d(z_2, z_1, \beta)$. Hence, we can always assume $|z_1| \geq |z_2|$ without loss of generality. Moreover, since the constellation $\mathcal{Z}_n$ satisfies the rotation-invariant Proposition 7, $d(z_1, z_2, \beta) = d(e^{j\theta} z_1, e^{j\theta} z_2, \beta)$ for any $z_1, z_2 \in \mathcal{Z}_n$, where $\theta = \pi, \pm\pi/2$.

Lemma 4 *Let a constellation $\mathcal{V}$ contain such 12 points v_k for $k = 1, 2, \cdots, 12$ that $v_1 = 2a, v_2 = -2a, v_3 = 2aj, v_4 = -2aj, v_5 = a + aj, v_6 = a - aj, v_7 = -a + aj, v_8 = -a - aj, v_9 = a, v_{10} = -a, v_{11} = aj, v_{12} = -aj$ and the other points satisfy $|v_k| < a$ for $k \geq 13$, where a is positive. If we define the distance of v_1 to the other points in $\mathcal{V}$ as*

$$g(v_1, \beta) = \min_{v_1 \neq v \in \mathcal{V}} d(v_1, v, \beta) \tag{4.35}$$

then, we have $g(v_1, \beta) = d(v_1, v_9, \beta)$.

Proof: For clarity, the diagram of the constellation $\mathcal{V}$ is plotted in Fig. 4.1. Let us first consider the distances of v_1 to v_2, v_3 and v_4. Since

$$d_1 \triangleq d(v_1, v_3, \beta) = d(v_1, v_4, \beta) = \frac{2\sqrt{2}a\beta}{\sqrt{1 + 4a^2\beta^2}\sqrt{1 + 4a^2\beta^2}}$$

$|v_1 - v_2| > |v_1 - v_3| = |v_1 - v_4|$ and $|v_2| = |v_3| = |v_4|$, we can arrive at the fact that $d(v_1, v_2, \beta) > d_1$. Hence, we have

$$\min\{d(v_1, v_2, \beta), d(v_1, v_3, \beta), d(v_1, v_4, \beta)\} = d_1 \tag{4.36}$$

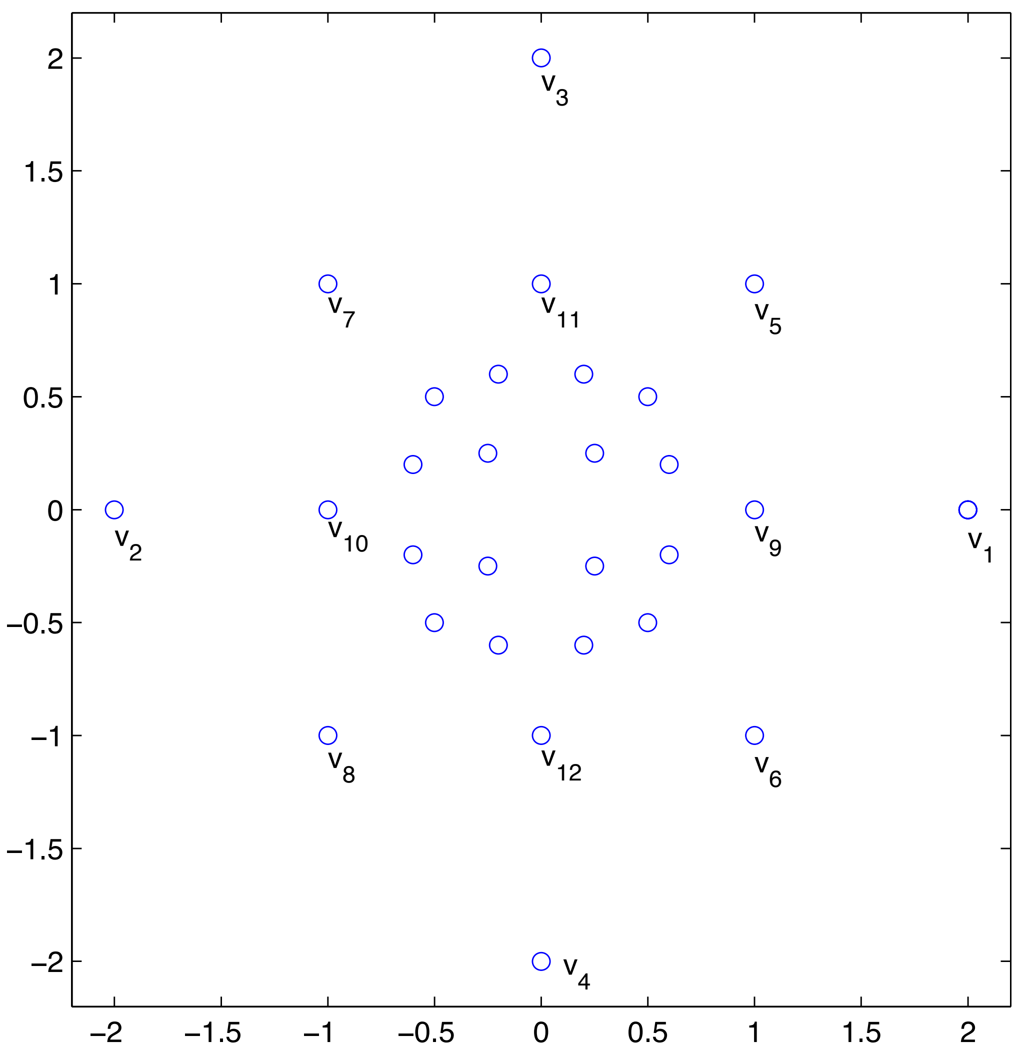


Figure 4.1: Set $\mathcal{V}$ on the complex plane

Now, we consider the distances of v_1 to v_5, v_6, v_7 and v_8. Notice that

$$d_2 \triangleq d(v_1, v_5, \beta) = d(v_1, v_6, \beta) = \frac{\sqrt{2}a\beta}{\sqrt{1+4a^2\beta^2}\sqrt{1+2a^2\beta^2}} \tag{4.37}$$

Since $|v_1 - v_7| = |v_1 - v_8| > |v_1 - v_5| = |v_1 - v_6|$ and $|v_5| = |v_6| = |v_7| = |v_8|$, we have $d(v_1, v_7, \beta), d(v_1, v_8, \beta) > d_2$. Therefore, we derive that

$$\min\{d(v_1, v_5, \beta), d(v_1, v_6, \beta), d(v_1, v_7, \beta), d(v_1, v_8, \beta)\} = d_2 \tag{4.38}$$

Finally, we consider all the distances of v_9, v_{10}, v_{11} and v_{12} to v_1. Since

$$d_3 \triangleq d(v_1, v_9, \beta) = \frac{a\beta}{\sqrt{1+4a^2\beta^2}\sqrt{1+a^2\beta^2}} \tag{4.39}$$

$|v_1 - v_i| > |v_1 - v_9|$ and $|v_i| = |v_9|$ for $i = 10, 11, 12$, we attain $d(v_1, v_i, \beta) > d_3$. For the other points $v_k \in \mathcal{V}$ with $k \geq 13$, since $|v_1 - v_k| > |v_1 - v_9|$ and $|v_k| < a = |v_9|$, we have that $d(v_1, v_k, \beta) > d_3$. Therefore, the minimum distance of v_1 to the other points in $\mathcal{V}$ is determined by

$$g(v_1, \beta) = \min\{d_1, d_2, d_3\} \tag{4.40}$$

Since

$$\begin{aligned} \frac{d_1}{d_2} &= \sqrt{\frac{4+8a^2\beta^2}{1+4a^2\beta^2}} > 1 \\ \frac{d_2}{d_3} &= \sqrt{\frac{2+2a^2\beta^2}{1+2a^2\beta^2}} > 1 \end{aligned}$$

we obtain $d_1 > d_2 > d_3$ and thus, $g(v_1, \beta) = d(v_1, v_9, \beta)$. This completes the proof of Lemma 4. □

From the proof of Lemma 4, we can immediately obtain the following corollary.

Corollary 1 *We have*

$$d(v_1, v_5, \beta) = d(v_1, v_6, \beta) < d(v_1, v_3, \beta) = d(v_1, v_4, \beta) < d(v_1, v_2, \beta)$$

The proof of Corollary 1 is ommited. Now, by taking advantage of the geometrical properties of $\mathcal{Z}_n$ and the distance function $d(z_1, z_2, \beta)$, Lemma 4 and its corollary, we can find all the solutions to Problem 3 with the size 2^n ranging from $n = 2$ to $n = 6$.

Theorem 5 *The solutions to Problem 3 for* $n = 2, 3, 4, 5$ *and* 6 *are given as follows:*

$$\hat{\beta} = \begin{cases} 1 & \text{if } n = 2 \text{ or } 6 \\ \frac{1}{\sqrt[4]{2}} & \text{if } n = 3 \text{ or } 4 \\ \sqrt{\frac{2+\sqrt{61}}{6}} & \text{if } n = 5 \end{cases}$$

$$D(\mathbb{T}_{\mathcal{Z}_n}(\hat{\beta})) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } n = 2 \\ \frac{1}{1+\sqrt{2}} & \text{if } n = 3 \\ \frac{1}{\sqrt{5+3\sqrt{2}}} & \text{if } n = 4 \\ \sqrt{\frac{57}{431+79\sqrt{61}}} & \text{if } n = 5 \\ \frac{1}{3\sqrt{6}} & \text{if } n = 6 \end{cases}$$

The proof of Theorem 5 is given in Appendix G. In spite of the fact that Theorem 5 explicitly gives the optimal solutions to Problem 3 only for small size of UFCs, if we carefully read through its proof, we find that it actually provides us with an elegant machinery to efficiently attain the optimal solution to Problem 3 for sizable UFCs. The key here is to gradually reduce the size of the constellation by properly making use of the geometrical properties of $\mathcal{Z}_n$ and the objective function, Lemma 4 and Corollary 1. Some major steps are highlighted as the following algorithm:

Algorithm 2 *The optimal solution to Problem 3 is found based on the following 5 steps.*

1. *Sort all the elements in the constellation $\mathcal{Z}_n$ by descending magnitude order such that $|z_1| \geq |z_2| \geq \cdots \geq |z_{2^n}|$.*

2. *Starting z_k with $k = 1$, go through $z_{k+1}, z_{k+2}, \cdots, z_{2^n}$ to find*

$$g(z_k, \beta) = \min_{k+1 \leq i \leq 2^n} d(z_k, z_i, \beta)$$

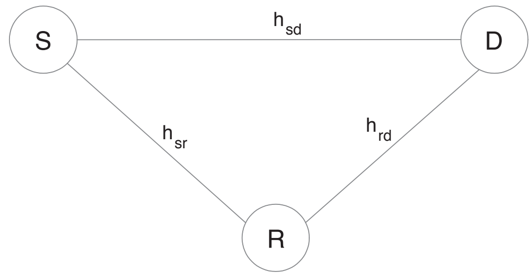


Figure 4.2: A single relay system

3. *Go back to 2) and increase k by 4 until $k = 2^n - 3$.*

4. *Compare all functions $g(z_k, \beta)$ for $k = 1, 5, \cdots, 2^n - 3$ to obtain $D(\mathbb{T}_{\mathcal{Z}_n}(\beta))$, i.e.,*

$$D(\mathbb{T}_{\mathcal{Z}_n}(\beta)) = \min_k g(z_k, \beta)$$

5. *Maximize $D(\mathbb{T}_{\mathcal{Z}_n}(\beta))$ over the energy scale variable β.*

4.4 Diagonal Distributed UFC Space-time Block Codes

In this section, we particularly consider a noncoherent AF half-duplex cooperative relay system with three nodes. Using Theorem 1 just established in Chapter 3 and the UFCs constructed in Subsection 4.3.2, we propose the design of full diversity unitary diagonal distributed space-time block codes with the LSE receiver. We also derive the closed-form decision rule for the GLRT receiver for such a system with the proposed code.

4.4.1 Design of unitary UFC codes

An AF half-duplex cooperative relay system with three nodes: a source S, a destination node D and a relay node R, is shown in Fig. 4.2. Each node has only a single antenna that cannot transmit and receive simultaneously. The channel gain from the source to the destination is denoted by h_{sd}, whereas those from the source to the relay and from the relay to the destination are denoted by h_{sr} and h_{rd}, respectively. It is assumed that the channel gains are completely unknown at the destination, but remain unchanged within four transmission time slots, after which they change to new independent values that are fixed for next four time slots, and so on. We are interested in the AF half-duplex protocols introduced in [4, 10, 11]. Particularly, we adopt the orthogonal cooperative transmission scheme proposed in [11], Then, the channel model can be written as

$$\mathbf{r} = \sqrt{\rho}\mathbf{S}\mathbf{h} + \boldsymbol{\eta} \tag{4.41}$$

where $\mathbf{h} = (h_{sd}, h_{sr}h_{rd})^T, \boldsymbol{\eta} = (\eta_1, \eta_2 + h_{rd}\eta_3, \eta_4, \eta_5 + h_{rd}\eta_6)^T$ and $\mathbf{S} = (s_1\mathbf{I}_2, s_2\mathbf{I}_2)^T \times \sqrt{3/E_s}$, with a pair of (s_1, s_2) randomly, independently and equally-likely drawn from a certain constellation $\mathbb{S}$ to be designed and E_s being the average energy of $\mathbb{S}$. We assume that the noise $\eta_i, i = 1, \cdots, 6$ are independent circularly-symmetric complex Gaussian random variables with each having zero mean and unit variance. Then, the covariance matrix $\mathbf{D}$ of vector $\boldsymbol{\eta}$ is $\mathbf{D} = \text{diag}\{1, 1 + |h_{rd}|^2, 1, 1 + |h_{rd}|^2\}$. Here, we aim at the design of a full diversity unitary diagonal distributed space-time block code for this system. From Theorem 1, we know that in order to achieve full diversity, we need to design such a constellation $\mathbb{S}$ that the matrices $\mathbf{P}_{\mathbf{s}\hat{\mathbf{s}}} = (\mathbf{S}, \hat{\mathbf{S}})^H(\mathbf{S}, \hat{\mathbf{S}})$ are invertible for all the distinct pairs of $\mathbf{S}$ and $\hat{\mathbf{S}}$. Notice that in this case,

$$\mathbf{P}_{\mathbf{s}\hat{\mathbf{s}}} = \frac{3}{E_s} \times \begin{pmatrix} s_1^*\mathbf{I}_2 & s_2^*\mathbf{I}_2 \\ \hat{s}_1^*\mathbf{I}_2 & \hat{s}_2^*\mathbf{I}_2 \end{pmatrix} \begin{pmatrix} s_1\mathbf{I}_2 & \hat{s}_1\mathbf{I}_2 \\ s_2\mathbf{I}_2 & \hat{s}_2\mathbf{I}_2 \end{pmatrix}$$

Since

$$\det(\mathbf{P}_{\mathbf{s}\hat{\mathbf{s}}}) = (s_1\hat{s}_2 - \hat{s}_1 s_2)^2 (3/E_s)^4 \tag{4.42}$$

the fact that the matrices $\mathbf{P}_{\mathbf{s}\hat{\mathbf{s}}} = (\mathbf{S}, \hat{\mathbf{S}})^H(\mathbf{S}, \hat{\mathbf{S}})$ are invertible for all the distinct pairs of $\mathbf{S}$ and $\hat{\mathbf{S}}$ is equivalent to the fact that $s_1\hat{s}_2 \neq \hat{s}_1 s_2$ for all $(s_1, s_2)^T \neq (\hat{s}_1, \hat{s}_2)^T \in \mathbb{S}$, which is nothing but the UFC. Therefore, designing the constellation $\mathbb{S}$ such that full diversity is achieved with the LSE receiver is equivalent to designing the constellation $\mathbb{S}$ to be a UFC. Fortunately, various sizes of UFCs have been designed in Subsection 4.3.2. Once the UFC has been constructed, a unitary UFC can be immediately attained by simply normalizing the nonunitary UFC (4.32). Hence, the resulting full diversity unitary diagonal distributed space-time block code is generated from the unitary UFC $\mathbb{S} = \bar{\mathbb{U}}$. Such a code is called diagonal distributed unitary-UFC (DDUFC) code.

Since the implementation of the ML receiver is intractable, we derive the decision rules for the GLRT receiver as well as for the LSE receiver in the following two subsections. In order to arrive at a general decoding algorithm, we consider an arbitrary constellation $\mathbb{S}$ instead of a unitary constellation.

4.4.2 GLRT detection

We first examine the case when the GLRT detector is employed at the receiver end. Conditioned on the channel coefficients $\mathbf{h}$ and transmitted signals $\mathbf{S}$, the probability density function of the received signal $\mathbf{z}$ is the Gaussian distribution, i.e., $\frac{1}{\pi^2 \det(\mathbf{D})} \times \exp\left(-(\mathbf{z}-\sqrt{\rho}\mathbf{Sh})^H\mathbf{D}^{-1}(\mathbf{z}-\sqrt{\rho}\mathbf{Sh})\right)$, and thus, its likelihood is given by $-(\mathbf{z}-\sqrt{\rho}\mathbf{Sh})^H\mathbf{D}^{-1}(\mathbf{z}-\sqrt{\rho}\mathbf{Sh})-2\ln\pi-2\ln(1+|h_{rd}|^2)$. By maximizing the likelihood function, we can obtain the estimated channel coefficient $\mathbf{h}$ and transmitter signal $\mathbf{S}$ as

$$\begin{aligned}\{\hat{\mathbf{S}},\hat{\mathbf{h}}\} &= \arg\min_{\mathbf{S},\mathbf{h}}\ (\mathbf{z}-\sqrt{\rho}\mathbf{Sh})^H\mathbf{D}^{-1}(\mathbf{z}-\sqrt{\rho}\mathbf{Sh})+2\ln(1+|h_{rd}|^2) \\ &= \arg\min_{\mathbf{S}}\min_{\mathbf{h}_{rd}}\min_{\mathbf{f}}\ (\mathbf{z}-\sqrt{\rho}\mathbf{SAf})^H\mathbf{D}^{-1}(\mathbf{z}-\sqrt{\rho}\mathbf{SAf}) \\ &\quad +2\ln(1+|h_{rd}|^2) \end{aligned} \tag{4.43}$$

where $\mathbf{A}=\mathrm{diag}(1,h_{rd})$ and $\mathbf{f}=(h_{sd},h_{sr})^T$. For the innermost minimization problem, differentiating the objective with respect to $\mathbf{f}$ and equating it to zero yields $\hat{\mathbf{f}}=(\mathbf{A}^H\mathbf{S}^H\mathbf{D}^{-1}\mathbf{SA})^{-1}\mathbf{A}^H\mathbf{S}^H\mathbf{D}^{-1}\mathbf{z}/\sqrt{\rho}$. Then, substituting $\hat{\mathbf{f}}$ and $\mathbf{A}$ back into (4.43) results in

$$\begin{aligned}\{\hat{\mathbf{S}},\hat{h}_{rd}\} &= \arg\min_{\mathbf{S}}\min_{h_{rd}}\ \big\{\mathbf{z}^H\mathbf{D}^{-1}\mathbf{z}-\mathbf{z}^H\mathbf{D}^{-1}\mathbf{SA}(\mathbf{A}^H\mathbf{S}^H\mathbf{D}^{-1}\mathbf{SA})^{-1}\mathbf{A}^H\mathbf{S}^H\mathbf{D}^{-1}\mathbf{z} \\ &\quad +2\ln(1+|h_{rd}|^2)\big\} \\ &= \arg\min_{\mathbf{S}}\min_{h_{rd}}\ \Big\{\frac{|s_1z_4-s_2z_2|^2}{(1+|h_{rd}|^2)(|s_1|^2+|s_2|^2)}-\frac{|s_1^*z_1+s_2^*z_3|^2}{|s_1|^2+|s_2|^2} \\ &\quad +2\sigma^2\ln(1+|h_{rd}|^2)\Big\} \end{aligned} \tag{4.44}$$

Letting $x = (1 + |h_{rd}|^2)^{-1}$ and equating the derivative of (4.44) with respect to x to zero leads to

$$x = \frac{2\sigma^2}{\bar{e}}$$

where $\bar{e} = \frac{|s_1 z_4 - s_2 z_2|^2}{|s_1|^2 + |s_2|^2}$. Therefore, the estimated magnitude of the channel coefficient h_{rd} is determined by

$$|\hat{h}_{rd}|^2 = \begin{cases} \frac{\bar{e}}{2\sigma^2} - 1 & \text{if } \bar{e} \geq 2\sigma^2, \\ \\ 0 & \text{if } 0 \leq \bar{e} < 2\sigma^2 \end{cases} \tag{4.45}$$

Substituting the result back into (4.44), the estimated transmitted signal $\hat{\mathbf{S}}$ is given by

$$\hat{\mathbf{S}} = \min_{(s_1, s_2)^T \in \mathbb{S}} \begin{cases} 2\sigma^2 \ln(\bar{e}/2) - \frac{|s_1^* z_1 + s_2^* z_3|^2}{|s_1|^2 + |s_2|^2} + 2\sigma^2 - 4\sigma^2 \ln \sigma, & \text{if } \quad \bar{e} \geq 2\sigma^2; \\ \\ \bar{e} - \frac{|s_1^* z_1 + s_2^* z_3|^2}{|s_1|^2 + |s_2|^2}, & \text{if } \quad 0 \leq \bar{e} < 2\sigma^2 \end{cases} \tag{4.46}$$

4.4.3 LSE detection

If the LSE receiver is used at the destination node, the following optimization problem needs to be solved,

$$\{\hat{\mathbf{S}}, \hat{\mathbf{h}}\} = \arg \min_{\mathbf{S}, \mathbf{h}} \; \|\mathbf{z} - \sqrt{\rho} \mathbf{S} \mathbf{h}\|_2^2 \tag{4.47}$$

Differentiating the quadratic function with respect to **h** and equating the result to zero, we can obtain the estimated channel coefficients

$$\hat{\mathbf{h}} = (\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H\mathbf{z}/\sqrt{\rho} \tag{4.48}$$

then, the estimated signal matrix is found out to be

$$\begin{aligned} \hat{\mathbf{S}} &= \arg\max_{\mathbf{S}} \ \mathbf{z}^H\mathbf{S}(\mathbf{S}^H\mathbf{S})^{-1}\mathbf{S}^H\mathbf{z} \\ &= \arg\max_{(s_1,s_2)^T\in\mathbb{S}} \frac{|s_1^*z_1 + s_2^*z_3|^2 + |s_1^*z_2 + s_2^*z_4|^2}{|s_1|^2 + |s_2|^2} \end{aligned} \tag{4.49}$$

It is worthwhile to point out that in general, the variance of the noise must be available for the GLRT detector, whereas the LSE detector doesn't require that information.

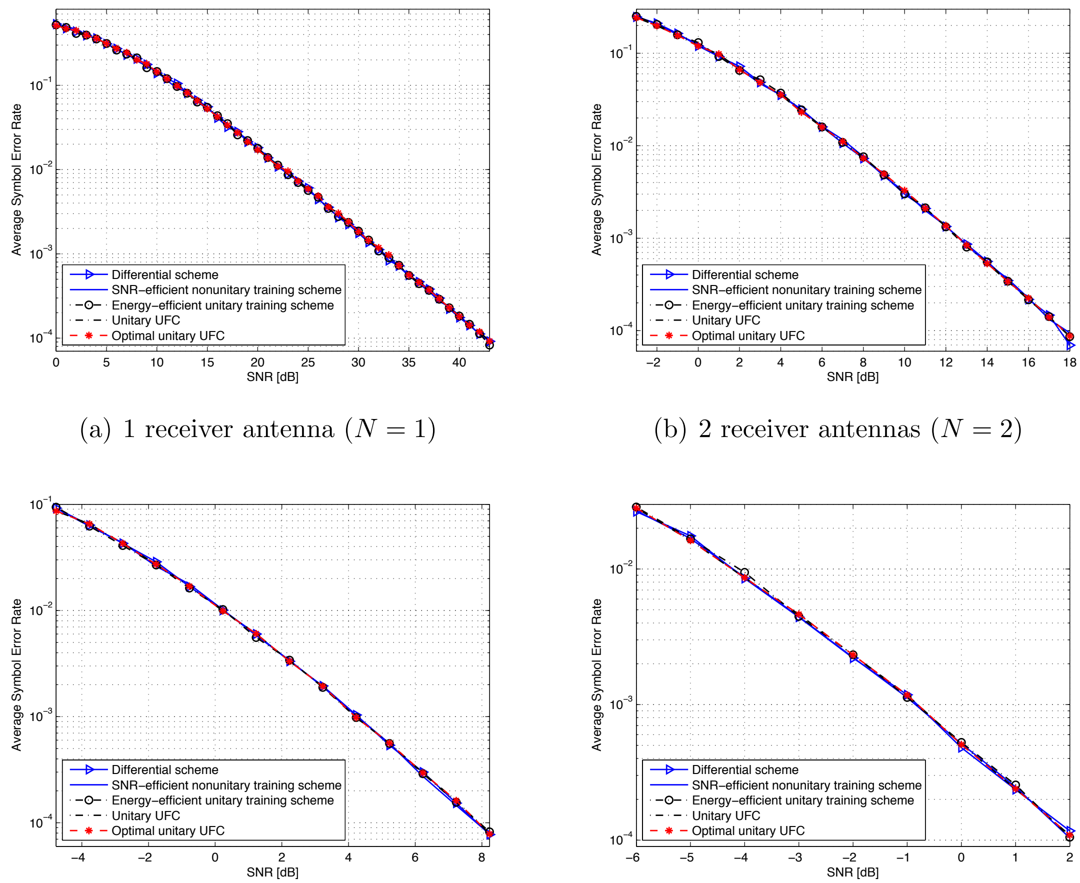
Chapter 5

Computer Simulations and Discussions

5.1 Computer Simulations for SIMO Systems

In this section, we perform computer simulations and examine error performance of the unitary UFC design proposed in this thesis by comparing it with other schemes in the literatures which can be used in the SIMO system, where channel state information is completely unknown at both the transmitter and the receiver. The coherence time is $T = 2$ and the number of the receiver antennas ranges from $N = 1$ to $N = 4$. All the schemes that we like to compare here are discussed as follows:

(a) *Differential scheme based on Phase Shift Keying (PSK) constellations.* For the necessity of performance comparison and decoding with the GLRT receiver, the



(a) 1 receiver antenna ($N = 1$)

(b) 2 receiver antennas ($N = 2$)

(c) 3 receiver antennas ($N = 3$)

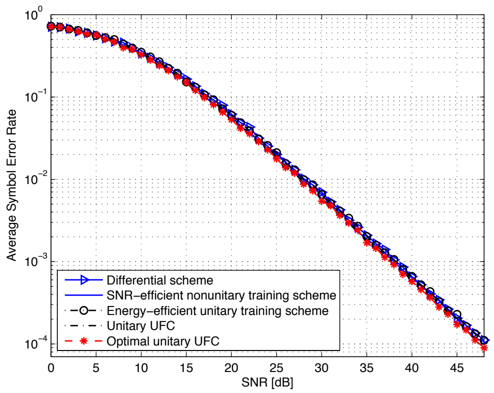
(d) 4 receiver antennas ($N = 4$)

Figure 5.1: Performance comparison for transmission bit rate $R_b = 1$ bits per channel use

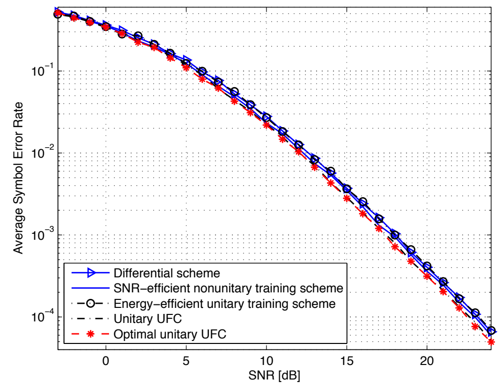
unitary codeword matrix is expressed by

$$\mathbf{S}_a = \frac{1}{\sqrt{2}} \times \begin{pmatrix} \mathbf{I}_N \\ \\ s_a \mathbf{I}_N \end{pmatrix} \tag{5.1}$$

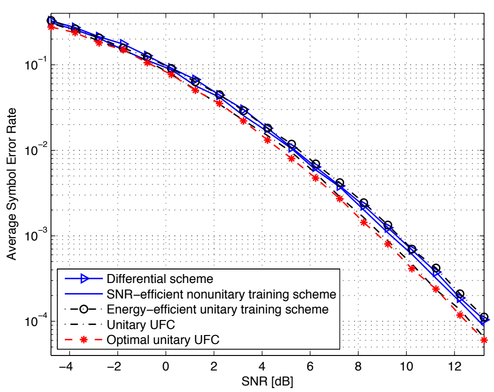
where s_a is randomly, independently and equally likely chosen from the 2^K-ary PSK



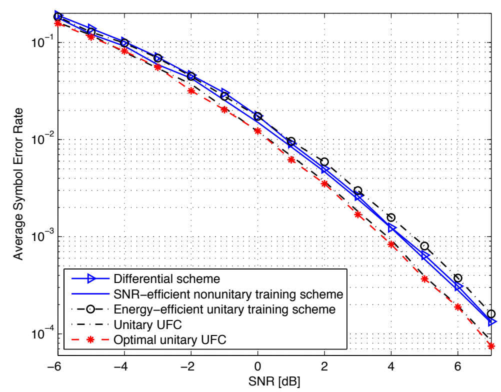
(a) 1 receiver antenna ($N = 1$)



(b) 2 receiver antennas ($N = 2$)



(c) 3 receiver antennas ($N = 3$)

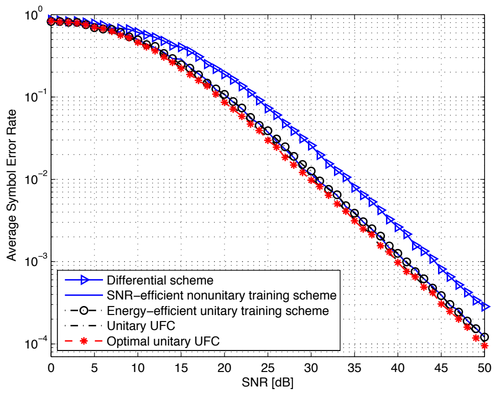


(d) 4 receiver antennas ($N = 4$)

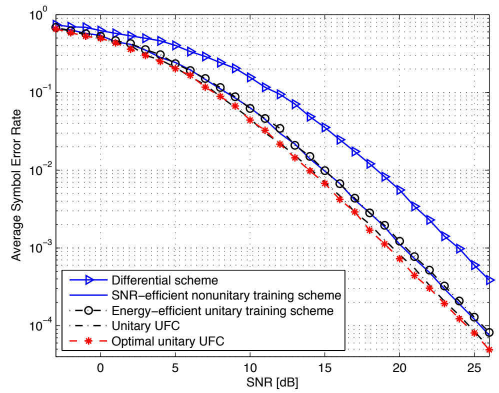
Figure 5.2: Performance comparison for transmission bit rate $R_b = 1.5$ bits per channel use

constellation and the normalization constant assures $\mathrm{E}\big[\mathrm{tr}\big(\mathbf{S}_a^H \mathbf{S}_a\big)\big] = N$.

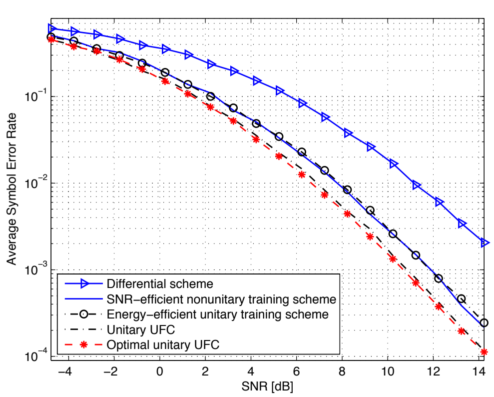
(b) *SNR-efficient nonunitary training scheme based on QAM constellations.* The



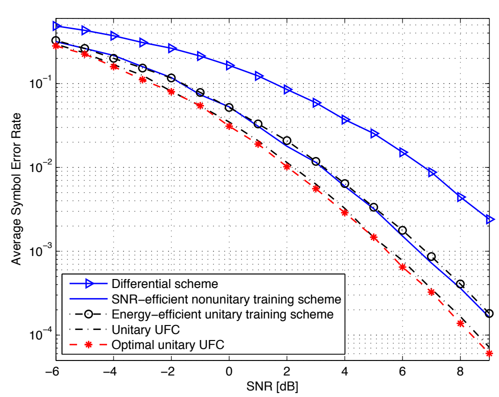
(a) 1 receiver antenna ($N = 1$)



(b) 2 receiver antennas ($N = 2$)



(c) 3 receiver antennas ($N = 3$)



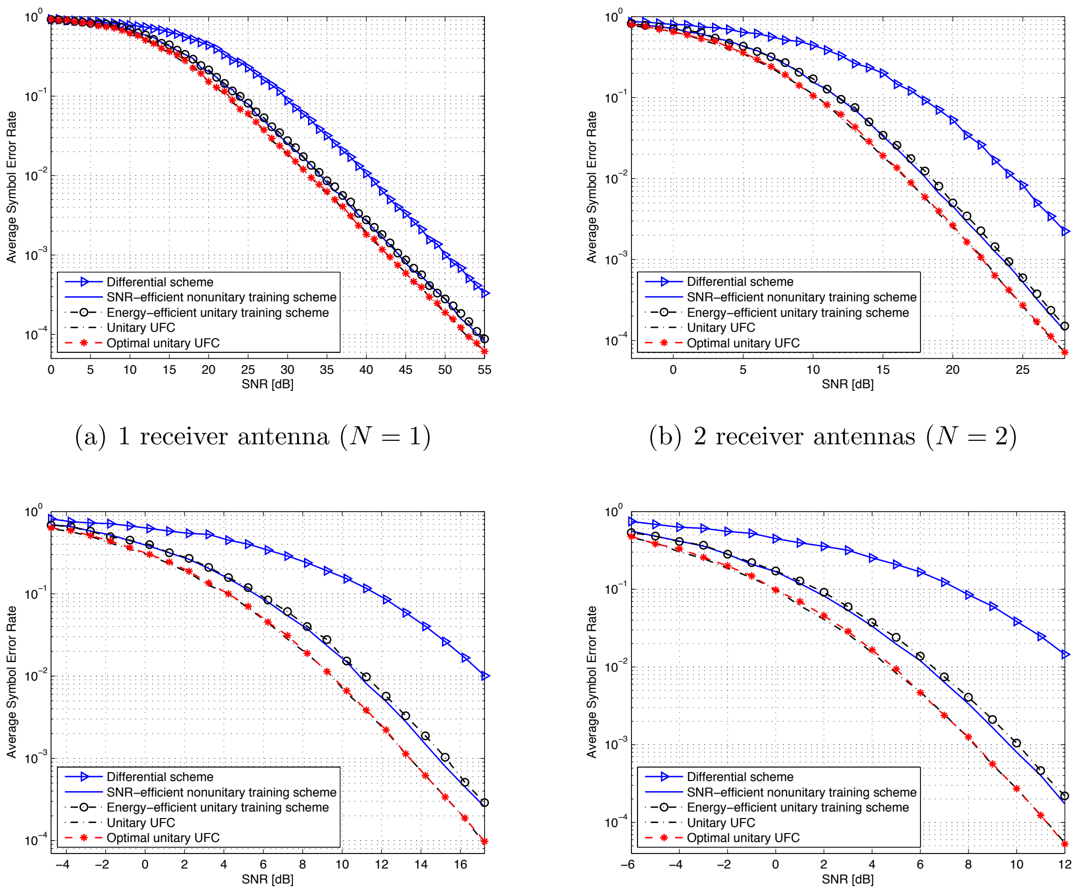
(d) 4 receiver antennas ($N = 4$)

Figure 5.3: Performance comparison for transmission bit rate $R_b = 2$ bits per channel use

codeword matrices for this SNR-efficient training scheme are characterized by

$$\mathbf{S}_b = \frac{1}{\sqrt{2E_b}} \times \begin{pmatrix} \sqrt{E_b}\mathbf{I}_N \\ \\ s_b\mathbf{I}_N \end{pmatrix} \tag{5.2}$$

where s_b is randomly, independently and equally likely chosen from the 2^K-ary cross



(a) 1 receiver antenna ($N = 1$)

(b) 2 receiver antennas ($N = 2$)

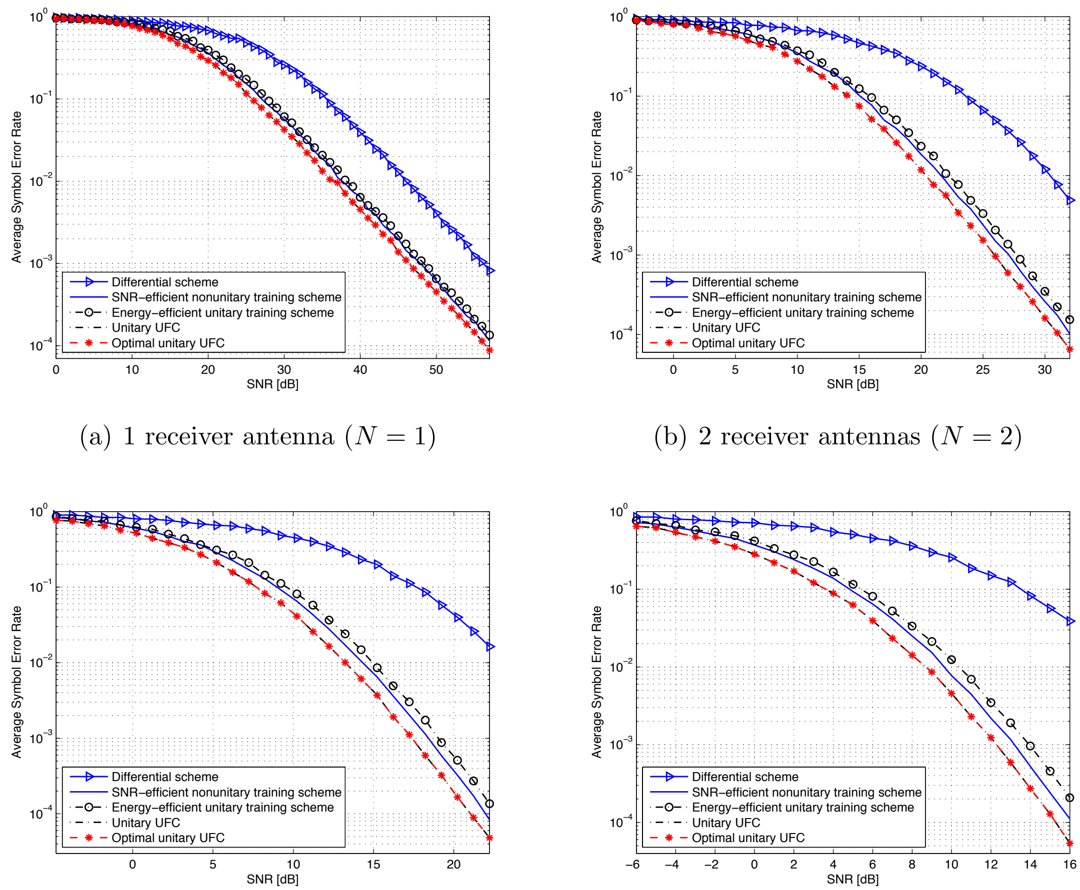
(c) 3 receiver antennas ($N = 3$)

(d) 4 receiver antennas ($N = 4$)

Figure 5.4: Performance comparison for transmission bit rate $R_b = 2.5$ bits per channel use

QAM constellation. The energy constant E_b is normalized in such a way that $\mathrm{E}\left[\mathrm{tr}\left(\mathbf{S}_b^H \mathbf{S}_b\right)\right] = N$. Here, the optimal average energy distribution over the training phase and communication phase is attained by maximizing the training efficiency [14, 60, 61].

(c) *Energy-efficient unitary training scheme based on QAM constellations.* This



(a) 1 receiver antenna ($N = 1$)

(b) 2 receiver antennas ($N = 2$)

(c) 3 receiver antennas ($N = 3$)

(d) 4 receiver antennas ($N = 4$)

Figure 5.5: Performance comparison for transmission bit rate $R_b = 3$ bits per channel use

design is proposed in this thesis and the codeword matrices are represented by

$$\mathbf{S}_c = \frac{1}{\sqrt{1+\hat{\alpha}^2|s_c|^2}} \times \begin{pmatrix} \mathbf{I}_N \\ \\ \hat{\alpha} s_c \mathbf{I}_N \end{pmatrix}, \qquad s_c \in \mathcal{Q}_K \tag{5.3}$$

where the energy scale $\hat{\alpha}$ is determined by Theorem 4.

(d) *Unitary UFC*. The constellation is designed in this thesis and the codeword matrices are of the form:

$$\mathbf{S}_d = \frac{1}{\sqrt{|x|^2 + |y|^2}} \times \begin{pmatrix} x\mathbf{I}_N \\ \\ y\mathbf{I}_N \end{pmatrix}, \qquad (x, y)^T \in \mathbb{U}_K \tag{5.4}$$

(e) *Optimal unitary UFC*. The optimal constellation design is proposed in this thesis and the codeword matrices are characterized by

$$\mathbf{S}_e = \frac{1}{\sqrt{1 + \hat{\beta}^2 |s_e|^2}} \times \begin{pmatrix} \mathbf{I}_N \\ \\ \hat{\beta} s_e \mathbf{I}_N \end{pmatrix}, \qquad s_e \in \mathcal{Z}_K \tag{5.5}$$

where the optimal energy scale $\hat{\beta}$ is given by Theorem 5.

It can be seen that the above five transmission schemes have the same spectrum efficiency, i.e., each transmission rate is $R_b = K/2$ bits per channel use. To make all the comparisons fair, we decode all the codes using the GLRT detector, i.e.,

$$\hat{\mathbf{S}} = \arg\max_{\mathbf{S}} \mathbf{r}^H \mathbf{S} \left(\mathbf{S}^H \mathbf{S}\right)^{-1} \mathbf{S}^H \mathbf{r}.$$

For $K = 2, 3, 4, 5$ and 6, the coding gains for all the constellations are listed in Table 5.1 and the average codeword error rates versus SNR are shown Fig. 5.1 to Fig. 5.5. It is observed that the optimal unitary UFC designed in this thesis performs the best error performance among all the five transmission schemes.

Table 5.1: Coding Gains for Different Constellations

K	$G(\mathbb{S}_a)$	$G(\mathbb{S}_b)$	$G(\mathbb{S}_c)$	$G(\mathbb{S}_d)$	$G(\mathbb{S}_e)$
2	0.5	0.5	0.5	0.5	0.5
3	0.1461	0.125	0.1	0.1667	0.1716
4	0.0381	0.0714	0.0730	0.1	0.1082
5	0.0096	0.0370	0.0335	0.0476	0.0543
6	0.0024	0.0143	0.0117	0.0185	0.0185

5.2 Computer Simulations for Three-Nodes Relay Systems

In this section, we carry out computer simulations and examine error performance of the DDUFC space-time block code designed in this thesis with the GLRT and LSE detectors. We compare this new code with the differential scheme and optimally precoded training method for the noncoherent AF half-duplex relay system.

(f) *Differential schemes.* First, we compare the DDUFC code with the differential coding scheme proposed in [15]. Despite the fact that the binary phase shift-keying (BPSK) constellation was only used in [15], it can be generalized in a straightforward manner into the 2^K-PSK constellation. Basically, the transmission scheme in [15] is that the relay scales the received signals from the source before transmitting them to the destination. Therefore, the transmitted signal $\mathbf{S}_f$ in (4.41) can be represented as

$$\mathbf{S}_f = \sqrt{E_f} \times \begin{pmatrix} \mathbf{I}_2 \\ \\ s_f\mathbf{I}_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \\ 0 & 1/\sqrt{E_f+1} \end{pmatrix}$$

where $s_f \in 2^K$-PSK and $E_f = \rho - 1 + \sqrt{\rho^2+\rho+1}$. The noise vector is $\boldsymbol{\eta} = (\eta_1, \eta_2 + h_{rd}\eta_3/\sqrt{E_f+1}, \eta_4, \eta_5 + h_{rd}\eta_6/\sqrt{E_f+1})^T$ with covariance matrix $\mathbf{D} = \mathrm{diag}\{1, 1 + |h_rd|^2/(E_f+1), 1, 1 + |h_{rd}|^2/(E_f+1)\}$. If the GLRT detector is employed at the destination node, then, after proper optimization over $\mathbf{f}$, which is the exactly same procedure as shown in Subsection 4.4.2, the objective function evaluates to

$$\begin{aligned}\{\hat{s}_f, \hat{h}_{rd}\} \quad = \quad & \arg\min_{s_f}\min_{h_{rd}} \Big\{ \frac{1+E_f}{1+E_f+|h_{rd}|^2}\frac{|z_2 s_f - z_4|^2}{2} \\ & -\frac{|z_1 + s_f^* z_3|^2}{2} + 2\ln\left(\frac{1+E_f+|h_{rd}|^2}{1+E_f}\right)\Big\} \end{aligned} \tag{5.6}$$

Following a similar procedure, the estimated magnitude of the channel coefficient h_{rd} is given by

$$|\hat{h}_{rd}|^2 = \begin{cases} \sqrt{\left(\frac{|z_2 s_f - z_4|^2}{2} - 1\right)(1+E_f)} & \text{if } |z_2 s_f - z_4|^2 \geq 4, \\ \\ 0 & \text{if } \quad 0 \leq |z_2 s_f - z_4|^2 < 4 \end{cases} \tag{5.7}$$

and the estimated signal is determined by

$$\hat{s}_f = \min_{s_f} \begin{cases} 4\ln|z_2 s_f - z_4| - \frac{|z_1 + s_f^* z_3|^2}{2} & \text{if } |z_2 s_f - z_4|^2 \geq 4, \\ |z_2 s_f - z_4|^2 - |z_1 + s_f^* z_3|^2 & \text{if } 0 \leq |z_2 s_f - z_4|^2 < 4 \end{cases} \tag{5.8}$$

(g) *Precoded training schemes.* The second transmitting scheme we would like to compare here is the precoded training scheme. For this scheme, the information symbols s_{g1} and s_{g2} transmitted from the source node are first precoded by a rotation matrix $\mathbf{F}$, i.e.,

$$\begin{pmatrix} s_{1\text{pre}} \\ s_{2\text{pre}} \end{pmatrix} = \mathbf{F} \begin{pmatrix} s_{g1} \\ s_{g2} \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

with $\alpha \in [0, 2\pi]$. Correspondingly, the transmitted signal matrix is represented by

$$\mathbf{S}_g = \begin{pmatrix} \frac{\sqrt{6}}{2} \times \mathbf{I}_2 \\ \frac{1}{\sqrt{E_g}} \times \mathbf{G} \end{pmatrix}, \quad \mathbf{G} = \begin{pmatrix} s_{1\text{pre}} & 0 \\ s_{2\text{pre}} & s_{1\text{pre}} \end{pmatrix}$$

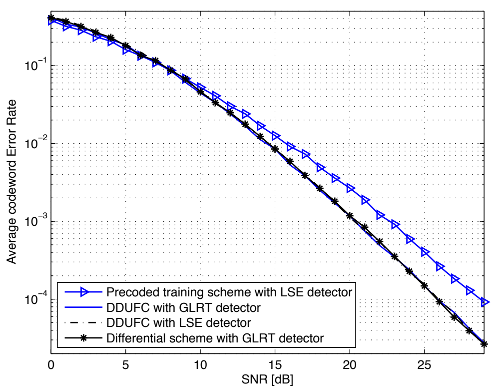
where E_g denote the average energy per information symbol. The optimal rotation angle [9] for this system with the QAM constellation of size K is given by $\alpha = \tan^{-1}(1/\sqrt{K})$. At the receiver end, the transmitted signal is estimated using LSE detector, i.e.,

$$\{\hat{s}_{1\text{pre}}, \hat{s}_{2\text{pre}}\} = \arg\max_{\mathbf{S}_g}\{\mathbf{r}^H\mathbf{S}_g(\mathbf{S}_g^H\mathbf{S}_g)^{-1}\mathbf{S}_g^H\mathbf{r}\}$$

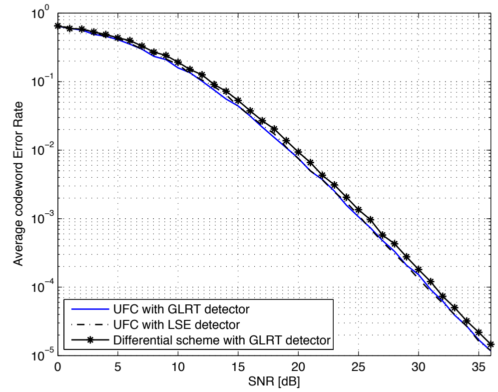
(h) *Diagonal distributed unitary-UFC codes* As described in Section 4.4, the unitary signal matrix is characterized by (4.32). However, for notation consistence in this section, we change notation $\mathbf{S}$ for the unitary UFC codeword matrix into $\mathbf{S}_h$, i.e., $\mathbf{S}_h = \mathbf{S}$. Hence, we have

$$\mathbf{S}_h = \sqrt{\frac{3}{|x|^2 + |y|^2}} \times \begin{pmatrix} x\mathbf{I}_2 \\ \\ y\mathbf{I}_2 \end{pmatrix} \tag{5.9}$$

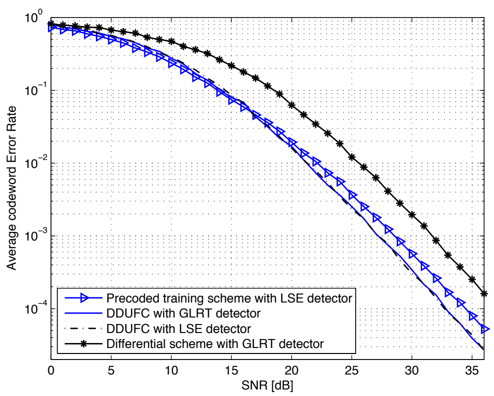
The codeword error rates versus SNR with a variety of transmission bit rates are plotted in Fig. 5.6. It can be observed that among all the schemes which we compare, the DDUFC code performs the best performance. In addition, there is almost no difference between the GLRT receiver and the LSE receiver for the DDUFC coded systems. As the size of constellation increases, the gap between the error rate curves also increases. Specifically when three 64-points constellations are used, the DDUFC code outperforms the differential scheme by about 13dB at the error rate of 10^{-3}. In addition, another interesting observation is that in spite of the fact that the symbol rate of the DDUFC code is 1/4 symbols per channel use and that of the precoded training scheme is 1/2 symbol per channel use, the error performance of the DDUFC code is still better than that of the precoded training scheme with both transmission bit rates being 0.5bits and 1bit per channel use (see Figs. 5.6(a) and 5.6(c)) and with the same LSE detector. When the transmission bit rate is increased to 2bits per channel use, within the region of SNR from 0dB to 40dB, i.e., the error rate being greater than 10^{-3} shown in Fig. 5.6(f), the error performance of the DDUFC code is worse than that of the precoded training scheme, whereas it is better when SNR is larger than 40dB.



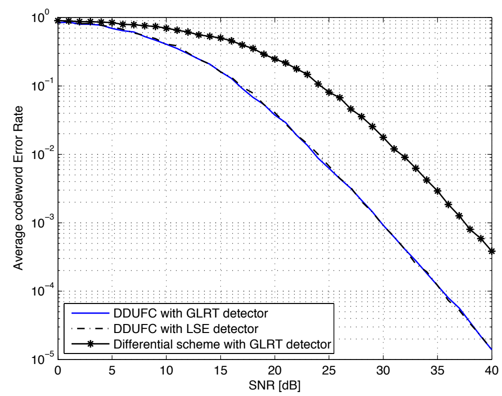
(a) Transmission bit rate $R_b = 0.5$ bits per channel use



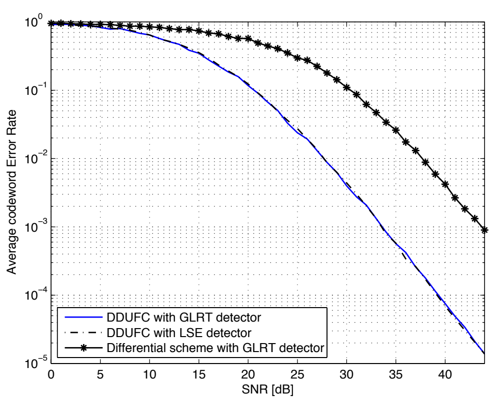
(b) Transmission bit rate $R_b = 0.75$ bits per channel use



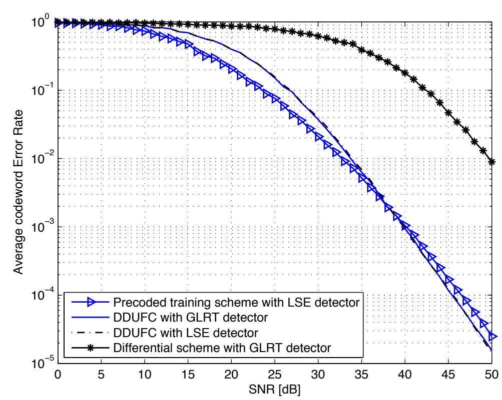
(c) Transmission bit rate $R_b = 1$ bits per channel use



(d) Transmission bit rate $R_b = 1.25$ bits per channel use



(e) Transmission bit rate $R_b = 1.5$ bits per channel use



(f) Transmission bit rate $R_b = 2$ bits per channel use

Figure 5.6: Average codeword error rate comparison with different transmission bit rates

Chapter 6

Conclusion and Future Work

In this thesis, we have first considered the noncoherent cooperative AF half-duplex relay systems and then, discussed the noncoherent SIMO systems, where full channel state information is completely unknown at both the source and the destination sides, but remains constant for a period of coherence time, after which it changes to a new independent realization that are fixed for the next period of coherence time, and so on. We have analyzed the asymptotic behavior of the pairwise error probability for the relay systems with the LSE receiver. A novel signal design method using the UFC has been proposed for the systematic constructions of the full diversity energy-efficent unitary constellations for the noncoherent SIMO channel and the unitary diagonal distributed space-time block codes for the relay channel with three nodes.

6.1 Asymptotic Performance Analysis of AF Relay Systems

For the AF half-duplex relay systems, we have proposed the use of the LSE receiver for detection. Since the currently available asymptotic analysis of the pairwise error probability for the noncoherent MIMO systems with the GLRT detector cannot be directly applied to the noncoherent cooperative relay systems with the LSE detector, we have re-derived the more accurate asymptotic formula using perturbation theory on the eigenvalues. With this, we have established the asymptotic formula of the pairwise error probability for the noncoherent relay systems with the LSE receiver. The result demonstrates that the full diversity gain function imitates coherent cooperative AF half-duplex relay systems, whereas the coding gain function imitates noncoherent MIMO systems. In addition, we have rigorously proved that for any given nonzero received signal, the unique blind identification of both the equivalent channel and the transmitted signals in the noise-free case for the AF relay systems is equivalent to full diversity with the LSE detector in the Gaussian noise environment.

6.2 Energy-Efficient Unitary UFC Designs

For the SIMO systems, we have first considered the design of the optimal unitary training constellation based on the commonly-used QAM constellations to maximize the coding gain. A deep investigation of the fractional coding gain function has revealed that the coding gain is achieved when the numerator achieves the minimum and meanwhile, the denominator achieves the second maximum. Therefore, a technical approach developed in this thesis to maximizing the coding gain is to appropriately design an energy scale to geometrically compress the first two largest energy points in the corner of the QAM constellations in the denominator of the objective. Closed-form optimal energy scale and coding gain have been attained. The result explicitly exposes a significant drawback of the QAM constellations for the nocoherent SIMO channel, i.e., the minimum Euclidean distance between the signal points with large energies is the same as the minimum Euclidean distance between the signal points with small energies. It is this drawback that has greatly motivated us to invent the novel concept, uniquely factorable constellation. We have proved that the UFC design assures the unique blind identification of channel coefficients and transmitted signals in the noise-free case for the SIMO systems by only processing two received signals, as well as full diversity for the GLRT receiver in the noisy case. By using the Lagrange's four-square theorem, an algorithm has been developed to efficiently and effectively construct various sizes of energy-efficient unitary UFCs to optimize the coding gain.

Particularly for the noncoherent AF half-duplex protocol with three nodes, we have used the full diversity criterion and UFCs established in this thesis for the systematic design of the full diversity unitary diagonal distributed space-time block codes. Furthermore, we have derived the closed-form decision rule for the GLRT receiver for this specific protocol. The comprehensive computer simulations have shown that error performance of the unitary UFC designed in this thesis is superior to those of the differential schemes, the optimal unitary training schemes presented in this thesis and SNR-efficient training schemes using the QAM constellation for the SIMO systems, which, thus far, performs the best error performance in current literatures.

The computer simulations have also demonstrated that error performance of the unitary diagonal distributed space-time block codes designed in this thesis outperforms those of the differential codes and the optimally precoded training schemes for the relay systems.

6.3 Future Work

As we have observed, the concept of the UFC is the key to the systematic design of energy-efficient full diversity unitary constellations for the noncoherent SIMO channel. However, some constructions and properties on the UFC which have been presented in this thesis are just initiative. More research and deeper investigations need to be done. Our future work will first focus on the following two aspects:

1. Instead of a pair of coprime PSK constellations, whether could the UFC constructed in this thesis be utilized to systematically design full diversity noncoherent space-time block codes for a general MIMO system by following the strategy similar to [54]?

2. The construction of the UFCs has been derived from the Gaussian integer ring using the Lagrange's four-square theorem. How about the Lagrange's four-square theorem in the Eisenstein integer ring? since the hexagonal constellations carved from the Eisenstein integer ring are more energy-efficient than the QAM constellations carved from the Gaussian integer ring [31].

In addition, in spite of the fact that full diversity space-time block code designs have recently been developed for the noncoherent MIMO systems by using a pair of coprime PSK constellations [54], this result cannot be directly applied to the noncoherent cooperative relay systems, since different protocols have different signal matrix structures. On the other hand, it is known that the PSK constellation is not as energy-efficient as the QAM constellation. Therefore, our future work will then concentrate on the systematic design of full diversity distributed space-time block codes for a general noncoherent AF half-duplex relay system based on the UFCs.

Appendix A

Proof of Lemma 1

Let $\mathbf{S}_{ij} = (\mathbf{S}_i, \mathbf{S}_j)$. Then, matrix $\mathbf{\Sigma}_{\mathbf{rr}|i}\mathbf{F}_{ij}$ can be rewritten as

$$\begin{aligned}
\mathbf{\Sigma}_{\mathbf{rr}|i}\mathbf{F}_{ij} &= (\rho\mathbf{S}_i\mathbf{\Sigma}\mathbf{S}_i^H + \mathbf{I}_{MN})[\mathbf{S}_i(\mathbf{S}_i^H\mathbf{S}_i)^{-1}\mathbf{S}_i^H - \mathbf{S}_j(\mathbf{S}_j^H\mathbf{S}_j)^{-1}\mathbf{S}_j^H] \\
&= \mathbf{S}_i(\rho\mathbf{\Sigma} + \mathbf{R}_{ii}^{-1})\mathbf{S}_i^H - \rho\mathbf{S}_i\mathbf{\Sigma}\mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{S}_j^H - \mathbf{S}_j\mathbf{R}_{jj}^{-1}\mathbf{S}_j^H \\
&= \mathbf{S}_{ij}\begin{bmatrix} \rho\mathbf{\Sigma} + \mathbf{R}_{ii}^{-1} & -\rho\mathbf{\Sigma}\mathbf{R}_{ij}\mathbf{R}_{jj}^{-1} \\ \mathbf{0}_{MN} & -\mathbf{R}_{jj}^{-1} \end{bmatrix}\mathbf{S}_{ij}^H
\end{aligned}$$

Since the nonzero eigenvalues of matrices $\mathbf{AB}$ and $\mathbf{BA}$ are equal, we can equivalently move $\mathbf{S}_{ij}$ to the right to form a matrix $\mathbf{M}_{ij}$ given by

$$\mathbf{M}_{ij} = \begin{bmatrix} \rho\mathbf{\Sigma}\mathbf{R}_{ii} + \mathbf{I}_{MN} - \rho\mathbf{\Sigma}\mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji} & \mathbf{R}_{ii}^{-1}\mathbf{R}_{ij} \\ -\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji} & -\mathbf{I}_{MN} \end{bmatrix}$$

Then, the eigenvalues of $\mathbf{M}_{ij}$ are equal to those of $\mathbf{\Sigma}_{\mathbf{rr}|i}\mathbf{F}_{ij}$. For discussion simplicity, let

$$\begin{aligned}
\mathbf{A} &= \mathbf{\Sigma}^{1/2}(\mathbf{R}_{ii} - \mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji})\mathbf{\Sigma}^{1/2} \\
\tilde{\mathbf{A}} &= (\rho\mathbf{A} + \mathbf{I}_{MN})^{-1}\mathbf{\Sigma}^{-1/2}\mathbf{R}_{ii}^{-1}\mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}\mathbf{\Sigma}^{1/2} \\
\mathbf{T}_1 &= \begin{bmatrix} \mathbf{\Sigma}^{1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_M \end{bmatrix} \\
\mathbf{T}_2 &= \begin{bmatrix} \mathbf{I}_{MN} & -(\rho\mathbf{A} + \mathbf{I}_{MN})^{-1}\mathbf{\Sigma}^{-1/2}\mathbf{R}_{ii}^{-1}\mathbf{R}_{ij} \\ \mathbf{0} & \mathbf{I}_{MN} \end{bmatrix} \\
\mathbf{T}_3 &= \begin{bmatrix} \mathbf{I}_{MN} & \mathbf{0} \\ -\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}\mathbf{\Sigma}^{1/2}(\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1} & \mathbf{I}_{MN} \end{bmatrix}
\end{aligned}$$

and

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \\ \mathbf{C}_3 & \mathbf{C}_4 \end{bmatrix} = \mathbf{T}_3^{-1}\mathbf{T}_2^{-1}\mathbf{T}_1^{-1}\mathbf{M}_{ij}\mathbf{T}_1\mathbf{T}_2\mathbf{T}_3$$

where each submatrix of matrix $\mathbf{C}$ is given by

$$\begin{aligned}
\mathbf{C}_1 &= \rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}} + (\mathbf{I}_{MN} - \tilde{\mathbf{A}})\tilde{\mathbf{A}}(\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1} \\
\mathbf{C}_2 &= (\rho\mathbf{A} + \mathbf{I}_{MN})^{-1}\mathbf{\Sigma}^{-1/2}\mathbf{R}_{ii}^{-1}\mathbf{R}_{ij} - \tilde{\mathbf{A}}(\rho\mathbf{A} + \mathbf{I}_{MN})^{-1}\mathbf{\Sigma}^{-1/2}\mathbf{R}_{ii}^{-1}\mathbf{R}_{ij} \\
\mathbf{C}_3 &= \mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}\mathbf{\Sigma}^{1/2}[\mathbf{I}_{MN} - \tilde{\mathbf{A}} - (\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}(\mathbf{I}_{MN} - \tilde{\mathbf{A}})\tilde{\mathbf{A}}](\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1} \\
\mathbf{C}_4 &= -\mathbf{I}_{MN} + \mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}\mathbf{\Sigma}^{1/2}[\mathbf{I}_{MN} - (\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}(\mathbf{I}_{MN} - \tilde{\mathbf{A}})] \\
&\quad \times(\rho\mathbf{A} + \mathbf{I}_{MN})^{-1}\mathbf{\Sigma}^{-1/2}\mathbf{R}_{ii}^{-1}\mathbf{R}_{ij}
\end{aligned}$$

Since eigenvalues do not change under any similarity transformation, the eigenvalues of matrix $\mathbf{C}$ are identical to those of $\mathbf{M}_{ij}$. In addition, let

$$\begin{aligned}
\tilde{\mathbf{C}} &= \begin{bmatrix} \mathbf{C}_1 - \rho\mathbf{A} - \mathbf{I}_{MN} & \mathbf{C}_2 \\ \mathbf{C}_3 & \mathbf{C}_4 + \mathbf{I}_{MN} \end{bmatrix} \\
\mathbf{D} &= \begin{bmatrix} \rho\mathbf{A} + \mathbf{I}_{MN} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_{MN} \end{bmatrix} \\
\mathbf{E} &= \mathbf{C} - \mathbf{D} = \tilde{\mathbf{C}}
\end{aligned} \tag{A.1}$$

Now, according to Proposition 2, we can bound the 2-norm of matrix $\tilde{\mathbf{A}}$ by

$$||\tilde{\mathbf{A}}||_2 \leq ||(\rho\mathbf{A} + \mathbf{I}_{MN})^{-1}||_2||\mathbf{\Sigma}^{-1/2}||_2||\mathbf{R}_{ii}^{-1}\mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}||_2||\mathbf{\Sigma}^{1/2}||_2 = \frac{c\kappa}{\rho\nu_1 + 1}$$

where c denotes the largest singular value of matrix $\mathbf{R}_{ii}^{-1}\mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}$. As a result, the asymptotic behavior of the F-norm of matrix $\tilde{\mathbf{A}}$ is

$$||\tilde{\mathbf{A}}||_{\mathrm{F}} = O\left(\frac{\kappa}{\rho\nu_1 + 1}\right) \tag{A.2}$$

By the triangular inequality of norm, we have

$$||\mathbf{I}_{MN} - \tilde{\mathbf{A}}||_{\mathrm{F}} \leq ||\mathbf{I}_{MN}||_{\mathrm{F}} + ||\tilde{\mathbf{A}}||_{\mathrm{F}} = \sqrt{MN} + O\left(\frac{\kappa}{\rho\nu_1 + 1}\right) \tag{A.3}$$

and

$$||(\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}||_{\mathrm{F}} = O\left(\frac{\kappa}{\rho\nu_1 + \epsilon}\right) \tag{A.4}$$

where $\epsilon = 1 - \frac{c\kappa}{\rho\nu_1+1}$. Utilizing (A.2), (A.4), (A.3) and Proposition 2, we can obtain

$$\begin{aligned} ||(\mathbf{I}_{MN} - \tilde{\mathbf{A}})\tilde{\mathbf{A}}(\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}||_{\mathrm{F}} &\leq \left[\sqrt{MN} + O\left(\frac{\kappa}{\rho\nu_1 + 1}\right)\right] \\ &\quad \times O\left(\frac{\kappa}{\rho\nu_1 + 1}\right) O\left(\frac{\kappa}{\rho\nu_1 + \epsilon}\right) \\ &= O\left(\frac{\kappa}{\rho\nu_1 + \epsilon}\right) \end{aligned}$$

Similarly, we can derive

$$\begin{aligned} &||\mathbf{I}_{MN} - \tilde{\mathbf{A}} - (\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}(\mathbf{I}_{MN} - \tilde{\mathbf{A}})\tilde{\mathbf{A}}||_{\mathrm{F}}^2 \\ &\leq 2||\mathbf{I}_{MN} - \tilde{\mathbf{A}}||_{\mathrm{F}}^2 + 2||(\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}(\mathbf{I}_{MN} - \tilde{\mathbf{A}})\tilde{\mathbf{A}}||_{\mathrm{F}}^2 \\ &\leq 2\left[\sqrt{MN} + O\left(\frac{\kappa}{\rho\nu_1 + 1}\right)\right]^2 \left[1 + O\left(\frac{\kappa^2}{(\rho\nu_1 + \epsilon)^2}\right) O\left(\frac{\kappa^2}{(\rho\nu_1 + 1)^2}\right)\right] \\ &= 2\left[MN + O\left(\frac{\kappa^2}{(\rho\nu_1 + \epsilon)^2}\right)\right] \end{aligned}$$

and

$$
\begin{aligned}
&||\mathbf{I}_{MN} - (\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}(\mathbf{I}_{MN} - \tilde{\mathbf{A}})||_{\mathrm{F}}^2 \\
&\leq 2MN + 2||(\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}(\mathbf{I}_{M} - \tilde{\mathbf{A}})||_{\mathrm{F}}^2 \\
&\leq 2MN + O\left(\frac{2\kappa^2}{(\rho\nu_1+\epsilon)^2}\right)\left[\sqrt{MN} + O\left(\frac{\kappa}{\rho\nu_1+1}\right)\right]^2 \\
&= 2MN + O\left(\frac{\kappa^2}{(\rho\nu_1+\epsilon)^2}\right)
\end{aligned}
$$

As a result, the norm of each submatrix of $\tilde{\mathbf{C}}$ is bounded by

$$
\begin{aligned}
||\mathbf{C}_1 - \rho\mathbf{A} - \mathbf{I}_{MN}||_{\mathrm{F}}^2 &= ||-\tilde{\mathbf{A}} + (\mathbf{I}_{MN} - \tilde{\mathbf{A}})\tilde{\mathbf{A}}(\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}||_{\mathrm{F}}^2 \\
&\leq 2||\tilde{\mathbf{A}}||_{\mathrm{F}}^2 + 2||(\mathbf{I}_{MN} - \tilde{\mathbf{A}})\tilde{\mathbf{A}}(\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}||_{\mathrm{F}}^2 \\
&= O\left(\frac{\kappa^2}{(\rho\nu_1+1)^2}\right) + O\left(\frac{\kappa^2}{(\rho\nu_1+\epsilon)^2}\right) \\
&= O\left(\frac{\kappa^2}{(\rho\nu_1+\epsilon)^2}\right) \\
||\mathbf{C}_2||_{\mathrm{F}}^2 &= ||-(\rho\mathbf{A} + \mathbf{I}_{MN})^{-1}\mathbf{\Sigma}^{-1/2}\mathbf{R}_{ii}^{-1}\mathbf{R}_{ij} \\
&\quad + \tilde{\mathbf{A}}(\rho\mathbf{A} + \mathbf{I}_{MN})^{-1}\mathbf{\Sigma}^{-1/2}\mathbf{R}_{ii}^{-1}\mathbf{R}_{ij}||_{\mathrm{F}}^2 \\
&\leq \left[\sqrt{MN} + O\left(\frac{\kappa}{\rho\nu_1+1}\right)\right]^2 O\left(\frac{\kappa^2}{(\rho\nu_1+1)^2}\right) \\
&= O\left(\frac{\kappa^2}{(\rho\nu_1+1)^2}\right) \\
||\mathbf{C}_3||_{\mathrm{F}}^2 &= ||\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}\mathbf{\Sigma}^{1/2}[\mathbf{I}_{MN} - \tilde{\mathbf{A}} - (\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}(\mathbf{I}_{MN} - \tilde{\mathbf{A}})\tilde{\mathbf{A}}] \\
&\quad \times(\rho\mathbf{A} + \mathbf{I}_{MN} - \mathbf{A})^{-1}||_{\mathrm{F}}^2 \\
&\leq \left[MN + O\left(\frac{\kappa^2}{(\rho\nu_1+\epsilon)^2}\right)\right] O\left(\frac{\kappa^2}{(\rho\nu_1+\epsilon)^2}\right) \\
&= O\left(\frac{\kappa^2}{(\rho\nu_1+\epsilon)^2}\right)
\end{aligned}
$$

and

$$\begin{aligned}
||\mathbf{C}_4 + \mathbf{I}_{MN}||_{\mathrm{F}}^2 &= ||\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}\mathbf{\Sigma}^{1/2}[\mathbf{I}_{MN} - (\rho\mathbf{A} + \mathbf{I}_{MN} - \tilde{\mathbf{A}})^{-1}(\mathbf{I}_{MN} - \tilde{\mathbf{A}})] \\
&\quad \times(\rho\mathbf{A} + \mathbf{I}_{MN})^{-1}\mathbf{\Sigma}^{-1/2}\mathbf{R}_{ii}^{-1}\mathbf{R}_{ij}||_{\mathrm{F}}^2 \\
&\leq \left[2MN + O\left(\frac{\kappa^2}{(\rho\nu_1 + \epsilon)^2}\right)\right] O\left(\frac{\kappa^2}{(\rho\nu_1 + 1)^2}\right) \\
&= O\left(\frac{\kappa^2}{(\rho\nu_1 + \epsilon)^2}\right)
\end{aligned}$$

Therefore, the norm of matrix $\tilde{\mathbf{C}}$ is bounded by

$$||\tilde{\mathbf{C}}||_{\mathrm{F}} = O\left(\frac{\kappa}{\rho\nu_1 + \epsilon}\right) \tag{A.5}$$

In addition, since the eigenvalues of $\mathbf{T}_3^{-1}\mathbf{C}\mathbf{T}_3$ satisfy condition: $\Re\{\tilde{\lambda}_1\} \leq \cdots \leq \Re\{\tilde{\lambda}_{2MN}\}$ and the eigenvalues of $\mathbf{D}$ are $\lambda_1 \leq \cdots \leq \lambda_{2MN}$ with $\lambda_1 = \cdots = \lambda_{MN} = -1$ and $\lambda_{MN+1} = \rho\nu_1 + 1, \cdots, \lambda_{2MN} = \rho\nu_{MN} + 1$, by Proposition 1, we have

$$\sum_{k=1}^{2MN} |\tilde{\lambda}_k - \lambda_k|^2 \leq 2||\mathbf{E}||_{\mathrm{F}}^2 \tag{A.6}$$

By (A.5), the F-norm of the perturbation matrix has the following asymptotic behavior,

$$||\mathbf{E}||_{\mathrm{F}} = ||\tilde{\mathbf{C}}||_{\mathrm{F}} = O\left(\frac{\kappa}{\rho\nu_1 + \epsilon}\right)$$

As a consequence, each individual term in the sum (B.9) is given by

$$\tilde{\lambda}_k = \lambda_k + O\left(\frac{\kappa}{\rho\nu_1 + \epsilon}\right) \tag{A.7}$$

for $k = 1, 2, \cdots, 2MN$. Substituting λ_k into (A.7) above completes the proof. $\square$

Appendix B

Proof of Property 1

By Lemma 1, we can let $\bar{\nu}_\ell$ for $\ell = 1, 2, \cdots, L-1$ denote all distinct eigenvalues of ν_l for $l = 1, 2, \cdots, MN$ with multiplicity $\mu_{\ell+1}$, and $\bar{\lambda}_i$ for $i = 1, 2, \cdots, L$ denote all distinct eigenvalues of λ_l for $l = 1, 2, \cdots, 2MN$ with multiplicity μ_i. In addition, $\bar{\lambda}_1 = -1 + O\left(\frac{\kappa}{\rho\nu_1+\epsilon}\right), \mu_1 = MN, \bar{\lambda}_\ell = \rho\bar{\nu}_{\ell-1} + 1 + O\left(\frac{\kappa}{\rho\nu_1+\epsilon}\right), \ell = 2, \cdots, L$ and $\sum_{\ell=2}^{L} \mu_\ell = MN$. By Proposition 3, the pairwise error probability can be written as

$$\Pr\{\nabla_j < \nabla_i\} = -\frac{1}{(MN-1)! \prod_{\ell=1}^{L} \bar{\lambda}_\ell^{\mu_\ell}} \lim_{s \to -\bar{\lambda}_1^{-1}} \frac{d^{MN-1}}{ds^{MN-1}} \left(s^{-1} \prod_{\ell=1}^{L-1} \left(s + \bar{\lambda}_\ell^{-1}\right)^{-\mu_\ell} \right) \quad \text{(B.8)}$$

Since κ and ν_1 are constant, we have $O\left(\frac{\kappa}{\rho\nu_1+\epsilon}\right) = O\left(\frac{1}{\rho}\right)$. Then the product of eigenvalues becomes

$$\begin{aligned}\prod_{\ell=1}^{L}\bar{\lambda}_\ell^{-\mu_\ell} &= \prod_{k=1}^{MN}[\rho\nu_k+1+O(1/\rho)]^{-1}[-1+O(1/\rho)]^{-MN} \\ &= (-1)^{MN}\prod_{k=1}^{MN}(\rho\nu_k+1)^{-1}+O(\rho^{-MN-1}) \\ &= 1/\det(\mathbf{D})+O(\rho^{-MN-1})\end{aligned}$$

where $\mathbf{D}$ is given in (A.1), and the derivative term in (B.8) can be expressed as

$$\frac{d^{MN-1}}{ds^{MN-1}}\left(s^{-1}\prod_{\ell=2}^{L}\left(s+\bar{\lambda}_\ell^{-1}\right)^{-\mu_\ell}\right) = \sum f(n_2,\cdots,n_{L-1})\frac{(-1)^{MN-1}}{s^{n_1+1}\prod\limits_{\ell=2}^{L}\left(s+\bar{\lambda}_\ell^{-1}\right)^{\mu_\ell+n_\ell}} \quad \text{(B.9)}$$

where

$$f(n_2,\cdots,n_L) = \frac{(MN-1)!}{n_2!\cdots n_L!}\prod_{\ell=2}^{L}\frac{(\mu_\ell+n_\ell-1)!}{(\mu_\ell-1)!}$$

and the summation in (B.9) is over all possible combinations of nonnegative integers $n_1,\cdots,n_L$ such that $\sum\limits_{\ell=1}^{L} n_l = MN-1$. Since $-\bar{\lambda}_1^{-1} = -1/(-1+O(1/\rho)) = 1+O(1/\rho)$, we have $\lim_{s\to-\bar{\lambda}_1^{-1}} s^{-n_1-1} = 1+O(1/\rho)$ and $\lim_{s\to-\bar{\lambda}_1^{-1}}(s+\bar{\lambda}_\ell^{-1})^{-\mu_\ell-n_\ell} = [1+O(1/\rho)+(\rho\nu_k+1+O(1/\rho))^{-1}]^{-\mu_\ell-n_\ell} = [1+O(1/\rho)]^{-\mu_\ell-n_\ell} = 1+O(1/\rho)$. Substituting these results into (B.8) yields

$$\begin{aligned}\Pr\{\nabla_j<\nabla_i\} &= -\frac{(-1)^{MN-1}}{(MN-1)!\prod\limits_{\ell=1}^{L}\bar{\lambda}_\ell^{\mu_\ell}}\sum f(n_2,\cdots,n_L)(1+O(1/\rho))\prod_{\ell=2}^{L}(1+O(1/\rho)) \\ &= -\frac{(-1)^{MN-1}}{(MN-1)!}[1/\det(\mathbf{D})+O(\rho^{-MN-1})](1+O(1/\rho))\sum f(n_2,\cdots,n_L)\end{aligned}$$

Finally, after some calculation and simplification, we can arrive at the conclusion,

$$
\begin{aligned}
\Pr\{\nabla_j < \nabla_i\} &= \frac{\begin{pmatrix} 2MN-1 \\ \\ MN \end{pmatrix}}{\det[\rho\boldsymbol{\Sigma}^{1/2}(\mathbf{R}_{ii} - \mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji})\boldsymbol{\Sigma}^{1/2} + \mathbf{I}_{MN}]} + O(\rho^{-MN-1}) \\
&= \frac{\begin{pmatrix} 2MN-1 \\ \\ MN \end{pmatrix}\rho^{-MN}}{\det[\boldsymbol{\Sigma}(\mathbf{R}_{ii} - \mathbf{R}_{ij}\mathbf{R}_{jj}^{-1}\mathbf{R}_{ji}) + \rho^{-1}\mathbf{I}_{MN}]} + O(\rho^{-MN-1})
\end{aligned}
$$

This completes the proof of Property 1. $\square$

Appendix C

Proof for Proproposition 2

Here, we provide a proof only for the last equation in Property 2, since the proofs for the other equations are very similar. Using partial fractions, we have

$$\frac{t^m}{(t+1)^n(t+a\rho^{-1})} = \sum_{p=1}^{n} \frac{c_p}{(t+1)^p} + \frac{c_0}{t+a\rho^{-1}}$$

where the residues are given by

$$\begin{aligned} c_0 &= O(\rho^{-m}) \\ c_p &= \frac{1}{(n-p)!}\frac{d^{n-p}}{dt^{n-p}}\left.\frac{t^m}{t+a\rho^{-1}}\right|_{t=-1} \end{aligned} \tag{C.10}$$

The derivative term in (C.10) can be rewritten as

$$\begin{aligned}
\frac{d^{n-p}}{dt^{n-p}}\frac{t^m}{t+a\rho^{-1}} &= \frac{d^{n-p}}{dt^{n-p}}\frac{(t+a\rho^{-1}-a\rho^{-1})^m}{t+a\rho^{-1}} \\
&= \frac{1}{(n-p)!}\frac{d^{n-p}}{dt^{n-p}}\left[\frac{(-a\rho^{-1})^m}{t+a\rho^{-1}}+\sum_{j=1}^{m} b_j(t+a\rho^{-1})^{j-1}(-a\rho^{-1})^{m-j}\right] \\
&= \frac{(a\rho^{-1})^m}{(t+a\rho^{-1})^{n-p+1}} \\
&\quad +\frac{1}{(n-p)!}\sum_{j=1}^{m} b_j(-a\rho^{-1})^{m-j}\frac{d^{n-p}}{dt^{n-p}}(t+a\rho^{-1})^{j-1} \qquad \text{(C.11)}
\end{aligned}$$

Now, substituting (C.11) into (C.10) and letting $t=-1$ lead to

$$c_p = O(1)$$

Hence, we obtain

$$\begin{aligned}
\int_0^\infty \frac{t^m\exp(-t)}{(t+1)^n(t+a\rho^{-1})}dt &= \sum_{p=1}^{n} c_p\int_0^\infty \frac{\exp(-t)}{(t+1)^p}dt + c_0\int_0^\infty \frac{\exp(-t)}{t+a\rho^{-1}}dt \\
&= O(1)+O(\ln\rho/\rho^m) \\
&= O(1)
\end{aligned}$$

This completes the proof of the last equation in Property 2 and thus, of Property 2 itself. □

Appendix D

Proof for Lemma 2

We prove this lemma by the induction on M. When $M = 1$, in this case, $N_1 = 1$, $\mathbf{P} = p_{11} \neq 0$, $\text{diag}(\mathbf{g}) = g_1$ and hence, $F(\rho, \mathbf{P})$ becomes

$$\begin{aligned} F\big(\rho, \mathbf{P}\big) &= \mathrm{E}\left[\frac{1}{1 + \rho\, p_{11}|g_1|^2}\right] \\ &= \int_0^\infty \frac{e^{-z}dz}{1 + \rho\, p_{11} z} \\ &= \frac{e^{\left(p_{11}\rho\right)^{-1}}}{p_{11}\rho} \mathcal{E}\big((p_{11}\rho)^{-1}\big) \end{aligned} \tag{D.12}$$

Using Proposition 4, we can have

$$\mathcal{E}((p_{11}\rho)^{-1}) = \ln\rho + \gamma + \ln p_{11} + O(\rho^{-1}) \tag{D.13}$$

In addition, the Taylor expansion of $e^{\left(p_{11}a(\theta,\rho)\right)^{-1}}$ gives us

$$e^{\left(p_{11}\rho\right)^{-1}} = 1 + O(\rho^{-1}) \tag{D.14}$$

Substituting (D.14) and (D.13) into (D.12) yields

$$F\big(\rho, \mathbf{P}\big) = \frac{\ln\rho + \gamma + \ln p_{11}}{p_{11}\rho} + O\Big(\frac{\ln\rho}{\rho^2}\Big) \tag{D.15}$$

Hence, Lemma 2 is true for $M = 1$. Now we assume that Lemma 2 is true for $M = L$. In the following we will prove that it is also true for $M = L + 1$. For notation simplicity, let

$$\mathbf{P} = \begin{pmatrix} p_{11} & \mathbf{P}_{12} \\ & \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{pmatrix} \tag{D.16}$$

where $\mathbf{P}_{21} = \mathbf{P}_{12}^H$. By Proposition 5, we have that,

$$\begin{aligned} &\det\big(\mathbf{I} + \rho \mathrm{diag}(\mathbf{g})^H \mathbf{P} \mathrm{diag}(\mathbf{g})\big) \\ &= \big(1 + \rho|g_1|^2 \overline{p}_{22}\big) \det\big(\mathbf{I} + \rho(\mathrm{diag}(\overline{\mathbf{g}}_1))^H \mathbf{P}_{22} \mathrm{diag}(\overline{\mathbf{g}}_1)\big) \end{aligned} \tag{D.17}$$

where $\overline{\mathbf{g}}_1 = (g_2, g_3, \cdots, g_M)^T$ and

$$\begin{aligned} \overline{p}_{22} \;\; = \;\; & p_{11} - \rho \mathbf{P}_{12} \mathrm{diag}(\overline{\mathbf{g}}_1)(\mathbf{I} + (1+\rho)(\mathrm{diag}(\overline{\mathbf{g}}_1))^H \mathbf{P}_{22} \mathrm{diag}(\overline{\mathbf{g}}_1))^{-1} \\ & \times \big(\mathrm{diag}(\overline{\mathbf{g}}_1)\big)^H \mathbf{P}_{21} \end{aligned} \tag{D.18}$$

On the other hand, we notice that if we let $\bar{\mathbf{I}}_1 = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$, then, applying Proposition 5 to matrix $\left(\bar{\mathbf{I}}_1 + \rho\big(\mathrm{diag}(1,\mathbf{g})\big)^H \mathbf{P}\mathrm{diag}(1,\mathbf{g}\right)$ twice yields

$$\begin{aligned}
& \det\left(\bar{\mathbf{I}}_1 + \rho\big(\mathrm{diag}(1,\mathbf{g})\big)^H \mathbf{P}\mathrm{diag}(1,\mathbf{g})\right) \\
= \;& \rho\bar{p}_{22}\det\left(\mathbf{I} + \rho(\mathrm{diag}(\overline{\mathbf{g}}_1))^H \mathbf{P}_{22}\mathrm{diag}(\overline{\mathbf{g}}_1)\right) \\
= \;& \rho\bar{p}_{11}\det\left(\mathbf{I} + \rho(\mathrm{diag}(\overline{\mathbf{g}}_1))^H \big(\mathbf{P}_{22} - p_{11}^{-1}\mathbf{P}_{21}\mathbf{P}_{12}\big)\mathrm{diag}(\overline{\mathbf{g}}_1)\right)
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\bar{p}_{22} &= \frac{p_{11}\det\left(\mathbf{I} + (\mathrm{diag}(\overline{\mathbf{g}}_1))^H \big(\mathbf{P}_{22} - p_{11}^{-1}\mathbf{P}_{21}\mathbf{P}_{12}\big)\mathrm{diag}(\overline{\mathbf{g}}_1)\right)}{\det\left(\mathbf{I} + (\mathrm{diag}(\overline{\mathbf{g}}_1))^H \mathbf{P}_{22}\mathrm{diag}(\overline{\mathbf{g}}_1)\right)} \\
&= p_{11}\det\left(\mathbf{I} - \rho p_{11}^{-1}\mathbf{P}_{21}\mathbf{P}_{12}\mathrm{diag}(\overline{\mathbf{g}}_1)(\mathrm{diag}(\overline{\mathbf{g}}_1))^H\big(\mathbf{I} + \mathbf{P}_{22}(\mathrm{diag}(\overline{\mathbf{g}}_1))^H\mathrm{diag}(\overline{\mathbf{g}}_1)\big)^{-1}\right) \\
&= p_{11}\det\Big(\mathbf{I} - p_{11}^{-1}\mathbf{P}_{21}\mathbf{P}_{12}\mathbf{P}_{22}^{-1} + \\
&\quad p_{11}^{-1}\mathbf{P}_{21}\mathbf{P}_{12}\big(\mathbf{I} + \rho\mathbf{P}_{22}(\mathrm{diag}(\overline{\mathbf{g}}_1))^H\mathrm{diag}(\overline{\mathbf{g}}_1)\big)^{-1}\Big) \qquad \text{(D.19)}
\end{aligned}$$

Similar to the discussion of (D.12) , we can have

$$\mathrm{E}\left[\frac{1}{1+\bar{p}_{22}|g_1|^2}\right] = \frac{\ln\rho}{1+\bar{p}_{22}\rho} + O(\rho^{-1}) \qquad \text{(D.20)}$$

Since g_1 and $\overline{\mathbf{g}}_1$ are independent, using (D.17) and (D.20) we obtain

$$\begin{aligned}
F\big(\rho,\mathbf{P}\big) &= \mathrm{E}_{\overline{\mathbf{g}}_1}\left[\frac{1}{\det\left(\mathbf{I}+\rho(\mathrm{diag}(\overline{\mathbf{g}}_1))^H\mathbf{P}_{22}\mathrm{diag}(\overline{\mathbf{g}}_1)\right)}\mathrm{E}_{g_1}\left[\frac{1}{1+\rho\overline{p}_{22}|g_1|^2}\right]\right] \\
&= \mathrm{E}_{\overline{\mathbf{g}}_1}\left[\frac{\ln\rho}{\big(1+\rho\overline{p}_{22}\big)\det\left(\mathbf{I}+\rho(\mathrm{diag}(\overline{\mathbf{g}}_1))^H\mathbf{P}_{22}\mathrm{diag}(\overline{\mathbf{g}}_1)\right)}\right] \\
&\quad +O\Big(\rho^{-1}\mathrm{E}_{\overline{\mathbf{g}}_1}\big[J(\overline{\mathbf{g}}_1,\mathbf{P}_{22})\big]\Big) \\
&= \mathrm{E}_{\overline{\mathbf{g}}_1}\left[\frac{\ln\rho}{\det\left(\mathbf{I}+\rho(\mathrm{diag}(1,\overline{\mathbf{g}}_1))^H\mathbf{P}\mathrm{diag}(1,\overline{\mathbf{g}}_1)\right)}\right] \\
&\quad +O\left(\frac{\ln^{r_1-1}\rho}{\rho^{r_1}}\right)
\end{aligned}$$

where we have used the induction hypothesis and $\big(1+\overline{p}_{22}\big)\det\big(\mathbf{I}+(\mathrm{diag}(\overline{\mathbf{g}}_1))^H\mathbf{P}_{22}\mathrm{diag}(\overline{\mathbf{g}}_1)\big)$ $=\det\big(\mathbf{I}+\rho(\mathrm{diag}(1,\overline{\mathbf{g}}_1))^H\mathbf{P}\mathrm{diag}(1,\overline{\mathbf{g}}_1)\big)$. Again, by Proposition 5, we have

$$\begin{aligned}
&\det\big(\mathbf{I}+\rho(\mathrm{diag}(1,\overline{\mathbf{g}}_1))^H\mathbf{P}\mathrm{diag}(1,\overline{\mathbf{g}}_1)\big) \\
&= \big(1+\rho p_{11}\big)\det\big(\mathbf{I}+\rho(\mathrm{diag}(\overline{\mathbf{g}}_1))^H\overline{\mathbf{P}}_{22}\mathrm{diag}(\overline{\mathbf{g}}_1)\big)
\end{aligned}\tag{D.21}$$

where $\overline{\mathbf{P}}_{22}$ is defined by

$$\overline{\mathbf{P}}_{22}=\mathbf{P}_{22}-\rho\big(1+\rho p_{11}\big)^{-1}\mathbf{P}_{21}\mathbf{P}_{12}\tag{D.22}$$

Therefore, we obtain

$$\begin{aligned}
&\mathrm{E}_{\overline{\mathbf{g}}_1}\left[\frac{\ln\rho}{\det\left(\mathbf{I}+\rho(\mathrm{diag}(1,\overline{\mathbf{g}}_1))^H\mathbf{P}\mathrm{diag}(1,\overline{\mathbf{g}}_1)\right)}\right] \\
=\;& \frac{\ln\rho}{\big(1+\rho p_{11}\big)}\mathrm{E}_{\overline{\mathbf{g}}_1}\left[\frac{1}{\det\big(\mathbf{I}+\rho(\mathrm{diag}(\overline{\mathbf{g}}_1))^H\overline{\mathbf{P}}_{22}\mathrm{diag}(\overline{\mathbf{g}}_1)\big)}\right] \\
=\;& \frac{\ln^{N_1}\rho}{\det\left(\mathbf{I}+\rho\mathbf{P}\right)}+O\left(\frac{\ln^{N_1-1}\rho}{\rho^{N_1}}\right)
\end{aligned}\tag{D.23}$$

This completes the proof of Lemma 2. $\square$

Appendix E

Proof of Theorem 1

We just prove the second statement in Theorem 1, since the first statement can be proved similarly. From the definition of pairwise error probability, we have

$$\begin{aligned} \mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2) &= \Pr\Big\{ {\mathbf{z}_2}^H \hat{\mathbf{X}}_2 \big(\hat{\mathbf{X}}_2^H \hat{\mathbf{X}}_2\big)^{-1} \hat{\mathbf{X}}_2^H \mathbf{z}_2 > \\ & {\mathbf{z}_2}^H \mathbf{X}_2 \big({\mathbf{X}_2}^H \mathbf{X}_2\big)^{-1} {\mathbf{X}_2}^H \mathbf{z}_2 \Big\} \end{aligned} \tag{E.24}$$

The received signal vector $\mathbf{z}_2$ conditioned on the channel coefficient $\mathbf{g}$ and the transmitted signal matrix $\mathbf{X}_2$ is Gaussian distributed with zero mean and covariance matrix $\mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2}$ being given by $\mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2} = \rho \mathbf{X}_2 \mathbf{G}\mathbf{G}^H {\mathbf{X}_2}^H + \mathbf{\Sigma}_2$, and thus, the conditional probability density function $f_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2)$ is determined by $f_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) = \frac{1}{\pi^{T_2}\det(\mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2})} \times \exp\left(-{\mathbf{z}_2}^H \mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2}^{-1} \mathbf{z}_2\right)$. Since $\mathbf{g}$ is Gaussian distributed with zero mean and unit variance, the probability density function $f_{\mathbf{z}_2|\mathbf{X}_2}(\mathbf{z}_2)$ of $\mathbf{z}_2$ conditioned on the transmitted signal matrix $\mathbf{X}_2$ is given by

$$f_{\mathbf{z}_2|\mathbf{X}_2}(\mathbf{z}_2) = \frac{1}{\pi^M} \int_{\mathbb{C}} f_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) \exp\left(-\mathbf{g}^H \mathbf{g}\right) d\mathbf{g}$$

For discussion convenience, let $\mathbb{E}$ denote the erroneous decision region, i.e.,

$$\mathbb{E} = \{\mathbf{z}_2 : {\mathbf{z}_2}^H \hat{\mathbf{X}}_2 \big(\hat{\mathbf{X}}_2^H \hat{\mathbf{X}}_2\big)^{-1} \hat{\mathbf{X}}_2^H \mathbf{z}_2 > {\mathbf{z}_2}^H \mathbf{X}_2 \big({\mathbf{X}_2}^H \mathbf{X}_2\big)^{-1} {\mathbf{X}_2}^H \mathbf{z}_2\}$$

Then, the pair wise error probability (E.24) can be represented as

$$\begin{aligned}
\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2) &= \int_{\mathbb{E}} f_{\mathbf{z}_2|\mathbf{X}_2}(\mathbf{z}_2) d\mathbf{z}_2 \\
&= \frac{1}{\pi^M} \int_{\mathbb{C}} \int_{\mathbb{E}} f_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) \exp\left(-\mathbf{g}^H \mathbf{g}\right) d\mathbf{z}_2 d\mathbf{g} \qquad \text{(E.25)}
\end{aligned}$$

In order to analyze the asymptotic behavior of $\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2)$, we first notice that

$$\mathbf{\Sigma}_{\mathrm{L}} = \rho \mathbf{X}_2 \mathbf{G} \mathbf{G}^H {\mathbf{X}_2}^H + \mathbf{I}_{T_2} \preceq \mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2} \preceq \rho \mathbf{X}_2 \mathbf{G} \mathbf{G}^H {\mathbf{X}_2}^H + (|g_M|^2 + 1)\mathbf{I}_{T_2} = \mathbf{\Sigma}_{\mathrm{U}}$$

where we assume that $|g_M|$ has the largest magnitude in $\mathbf{g}$. Therefore, we have

$$\frac{1}{\pi^{T_2} \det(\mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2})} \exp\left(-{\mathbf{z}_2}^H \mathbf{\Sigma}_{\mathrm{L}}^{-1} \mathbf{z}_2\right) \le f_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) \le \frac{1}{\pi^{T_2} \det(\mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2})} \exp\left(-{\mathbf{z}_2}^H \mathbf{\Sigma}_{\mathrm{U}}^{-1} \mathbf{z}_2\right)$$

which can be rewritten as

$$\frac{\det(\mathbf{\Sigma}_{\mathrm{L}})}{\det(\mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2})} F_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) \le f_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) \le \frac{\det(\mathbf{\Sigma}_{\mathrm{U}})}{\det(\mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2})} G_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2)$$

with

$$\begin{aligned}
F_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) &= \frac{1}{\pi^{T_2} \det(\mathbf{\Sigma}_{\mathrm{L}})} \times \exp\left(-{\mathbf{z}_2}^H \mathbf{\Sigma}_{\mathrm{L}}^{-1} \mathbf{z}_2\right) \\
G_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) &= \frac{1}{\pi^{T_2} \det(\mathbf{\Sigma}_{\mathrm{U}})} \times \exp\left(-{\mathbf{z}_2}^H \mathbf{\Sigma}_{\mathrm{U}}^{-1} \mathbf{z}_2\right)
\end{aligned}$$

By doing so, the pairwise error probability $\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2)$ can be lower and upper bounded, respectively, by

$$\begin{aligned}\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2) &\geq \frac{1}{\pi^M}\int_{\mathbb{C}} \frac{\det(\mathbf{\Sigma}_{\mathrm{L}})}{\det(\mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2})} \times \exp\left(-\mathbf{g}^H\mathbf{g}\right) \\ &\quad \times \int_{\mathbb{E}} F_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) d\mathbf{z}_2 d\mathbf{g} \qquad \text{(E.26a)} \\ \mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2) &\leq \frac{1}{\pi^M}\int_{\mathbb{C}} \frac{\det(\mathbf{\Sigma}_{\mathrm{U}})}{\det(\mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2})} \times \exp\left(-\mathbf{g}^H\mathbf{g}\right) \\ &\quad \times \int_{\mathbb{E}} G_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) d\mathbf{z}_2 d\mathbf{g} \qquad \text{(E.26b)}\end{aligned}$$

At this moment, it is important to realize that the integral $\int_{\mathbb{E}} F_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) d\mathbf{z}_2$ is actually the pairwise error probability of the GLRT detector for the space-time block coded MIMO system with M transmitter antennas, a single receiver antenna, the codeword matrix $\mathbf{X}_2$, the circularly-symmetric zero-mean complex Gaussian channel having covariance matrix $\mathbf{\Sigma}_{hh} = \mathbf{G}\mathbf{G}^H$, signal-to-noise ratio ρ, and independent circularly-symmetric zero-mean complex Gaussian noise having covariance matrix $\mathbf{I}_{T_2}$, whereas the integral $\int_{\mathbb{E}} G_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2) d\mathbf{z}_2$ is the pairwise error probability of the GLRT detector for the same MIMO system, but the noise covariance matrix is $(|g_M|^2+1)\mathbf{I}_{T_2}$ instead of $\mathbf{I}_{T_2}$, and signal-to-noise ratio $\rho/(|g_M|^2+1)$ instead of ρ. Applying Property 1 to the both integrals yields

$$\int_{\mathbb{E}} F_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2)d\mathbf{z}_2 = \frac{\begin{pmatrix} 2M-1 \\ M \end{pmatrix}}{\det[\rho\mathbf{A}\mathbf{G}\mathbf{G}^H + \mathbf{I}_M]} + O\left(\frac{\sqrt{\kappa}}{(\rho|g_1|^2c_1+\epsilon)^{M+1}}\right) \quad \text{(E.27a)}$$

$$\begin{aligned}\int_{\mathbb{E}} G_{\mathbf{z}_2|\mathbf{X}_2,\mathbf{g}}(\mathbf{z}_2)d\mathbf{z}_2 &= \frac{\begin{pmatrix} 2M-1 \\ M \end{pmatrix}}{\det[\rho\mathbf{A}\mathbf{G}\mathbf{G}^H + (|g_M|^2+1)\mathbf{I}_M]} \\ &\quad + O\left(\frac{(|g_M|^2+1)\sqrt{\kappa}}{(\rho\frac{|g_1|^2c_1}{|g_M|^2+1}+\epsilon)^{M+1}}\right) \quad \text{(E.27b)}\end{aligned}$$

where matrix $\mathbf{R}_{\mathbf{x}_2\hat{\mathbf{x}}_2} = \mathbf{X}_2^H\mathbf{X}_2, \mathbf{A} = \mathbf{R}_{\mathbf{x}_2\mathbf{x}_2} - \mathbf{R}_{\mathbf{x}_2\hat{\mathbf{x}}_2}\mathbf{R}^{-1}_{\hat{\mathbf{x}}_2\hat{\mathbf{x}}_2}\mathbf{R}_{\hat{\mathbf{x}}_2\mathbf{x}_2}, \sqrt{\kappa} = \frac{\max\{|g_M|,1\}}{\min\{|g_1|,1\}}$, and c_1 is the smallest eigenvalue of matrix $\mathbf{A}$. By letting $t_i = |g_i|^2, i = 1, \cdots, M$, the ratio of the two determinants in (E.26b) can be bounded by

$$\frac{\det(\boldsymbol{\Sigma}_{\mathrm{U}})}{\det(\boldsymbol{\Sigma}_{\mathbf{z}_2\mathbf{z}_2})} \leq \frac{\det(\boldsymbol{\Sigma}_{\mathrm{U}})}{\det(\boldsymbol{\Sigma}_{\mathrm{L}})} = \frac{\det[(t_M+1)\rho\mathbf{X}_2\mathbf{G}\mathbf{G}^H{\mathbf{X}_2}^H + (t_M+1)\mathbf{I}_{T_2}]}{\det[\rho\mathbf{X}_2\mathbf{G}\mathbf{G}^H{\mathbf{X}_2}^H + \mathbf{I}_{T_2}]}$$

It can be seen that

$$\frac{\det(\boldsymbol{\Sigma}_{\mathrm{U}})}{\det(\boldsymbol{\Sigma}_{\mathbf{z}_2\mathbf{z}_2})} \leq (t_M+1)^{T_2} = 1 + \sum_{k=1}^{T_2} c_k t_M^k$$

where c_k for $k = 1, 2, \cdots, T_2$ are constants. In addition, the denominator in the dominant term of (E.27b) can be rewritten as

$$\frac{1}{\det(\rho \mathbf{A}\mathbf{G}\mathbf{G}^H + (t_M + 1)\mathbf{I}_M)} \leq \frac{1}{\det(\rho \mathbf{A}\mathbf{G}\mathbf{G}^H + \mathbf{I}_M)} \leq \prod_{i=1}^{M} \frac{1}{\rho \nu_1 t_i + 1}$$

where ν_1 is the smallest eigenvalue of matrix $\mathbf{A}$. Now, the error probability can be upper bounded by

$$\begin{aligned}
\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2) \leq & \begin{pmatrix} 2M-1 \\ M \end{pmatrix} \left\{ \mathrm{E}_{\mathbf{g}} \left[\frac{1}{\det(\rho \mathbf{A}\mathbf{G}\mathbf{G}^H + \mathbf{I}_M)} \right] \right. \\
& \left. + \frac{1}{\pi^M} \int_{\mathbb{C}} \sum_{k=1}^{T_2} c_k t_M^k \prod_{i=1}^{M} \frac{\exp(-t_i)}{\rho \nu_1 t_i + 1} dg_1 \cdots dg_M \right\} \\
= & \begin{pmatrix} 2M-1 \\ M \end{pmatrix} \mathrm{E}_{\mathbf{g}} \left[\frac{1}{\det(\rho \mathbf{A}\mathbf{G}\mathbf{G}^H + \mathbf{I}_M)} \right] \\
& + \begin{pmatrix} 2M-1 \\ M \end{pmatrix} \left[\sum_{k=1}^{T_2} c_k \int_0^{\infty} \frac{t_M^k \exp(-t_M)}{\rho \nu_1 t_M + 1} dt_M \right] \\
& \times \left[\prod_{i=1}^{M-1} \int_0^{\infty} \frac{\exp(-t_i)}{\rho \nu_1 t_i + 1} dt_i \right] \qquad \text{(E.28)}
\end{aligned}$$

Using the results in Proposition 2 and Lemma 2, we simplify the upper bound into

$$P_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2) \leq \frac{\begin{pmatrix} 2M-1 \\ M \end{pmatrix} \ln^M \rho}{\det(\mathbf{I}_M + \rho \mathbf{A})} + O\left(\frac{\ln^{M-1}\rho}{\rho^M}\right) \tag{E.29}$$

On the other hand, the ratio of the two determinants in (E.26a) can be lower bounded by

$$\frac{\det(\mathbf{\Sigma}_{\mathrm{L}})}{\det(\mathbf{\Sigma}_{\mathbf{z}_2\mathbf{z}_2})} \geq \frac{1}{(t_M+1)^{T_2}} = 1 + \sum_{k=1}^{T_2} c'_k \frac{t_M^k}{(1+t_M)^{T_2}}$$

where c'_k are also constants. Hence, the pairwise error probability can be lower bounded by

$$
\begin{aligned}
\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \rightarrow \hat{\mathbf{X}}_2) \;\geq\; & \begin{pmatrix} 2M-1 \\ M \end{pmatrix} \left\{ \mathrm{E}_{\mathbf{g}}\left[\frac{1}{\det(\rho\mathbf{A}\mathbf{G}\mathbf{G}^H + \mathbf{I}_M)}\right] \right. \\
& \left. + \frac{1}{\pi^M}\int_{\mathbb{C}} \sum_{k=1}^{T_2} c'_k \frac{t_M^k}{(1+t_M)^{T_2}} \prod_{i=1}^{M} \frac{\exp(-t_i)}{\rho\nu_1 t_i + 1} dg_1 \cdots dg_M \right\} \\
= & \begin{pmatrix} 2M-1 \\ M \end{pmatrix} \mathrm{E}_{\mathbf{g}}\left[\frac{1}{\det(\rho\mathbf{A}\mathbf{G}\mathbf{G}^H + \mathbf{I}_M)}\right] \\
& + \begin{pmatrix} 2M-1 \\ M \end{pmatrix} \left[\sum_{k=1}^{T_2} c'_k \int_0^{\infty} \frac{t_M^k \exp(-t_M)}{(1+t_M)^{T_2}(\rho\nu_1 t_M + 1)} dt_M\right] \\
& \times \left[\prod_{i=1}^{M-1} \int_0^{\infty} \frac{\exp(-t_i)}{\rho\nu_1 t_i + 1} dt_i\right] \\
= & \begin{pmatrix} 2M-1 \\ M \end{pmatrix} \mathrm{E}_{\mathbf{g}}\left[\frac{1}{\det(\rho\mathbf{A}\mathbf{G}\mathbf{G}^H + \mathbf{I}_M)}\right] \\
& + \left[\sum_{k=1}^{T_2} c'_k \int_0^{\infty} \frac{t_M^k \exp(-t_M)}{(1+t_M)^{T_2}(t_M + \rho^{-1}\nu_1^{-1})} dt_M\right] O\left(\frac{\ln^{M-1}\rho}{\rho^M}\right)
\end{aligned}
$$

Using Proposition 2 and Lemma 2, we have,

$$\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2) \geq \frac{\begin{pmatrix} 2M-1 \\ M \end{pmatrix} \ln^M \rho}{\det(\mathbf{I}_M + \rho \mathbf{A})} + O\left(\frac{\ln^{M-1} \rho}{\rho^M}\right) \tag{E.30}$$

Now, combining (E.29) with (E.30) results in

$$\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2) = \frac{\begin{pmatrix} 2M-1 \\ M \end{pmatrix} \ln^M \rho}{\det(\mathbf{I}_M + \rho \mathbf{A})} + O\left(\frac{\ln^{M-1} \rho}{\rho^M}\right) \tag{E.31}$$

Therefore, when ρ tends to infinity, the asymptotic behaviour of the average pairwise error probability of the LSE detector is given by

$$\mathrm{P}_{\mathrm{LSE}}(\mathbf{X}_2 \to \hat{\mathbf{X}}_2) = \frac{\begin{pmatrix} 2M-1 \\ M \end{pmatrix} \ln^M \rho}{\rho^M \det(\mathbf{R}_{\mathbf{x}_2\mathbf{x}_2} - \mathbf{R}_{\mathbf{x}_2\hat{\mathbf{x}}_2}\mathbf{R}_{\hat{\mathbf{x}}_2\hat{\mathbf{x}}_2}^{-1}\mathbf{R}_{\hat{\mathbf{x}}_2\mathbf{x}_2})} + O\left(\frac{\ln^{M-1} \rho}{\rho^M}\right) \tag{E.32}$$

Noting that $\det(\mathbf{R}_{\mathbf{x}_2\mathbf{x}_2} - \mathbf{R}_{\mathbf{x}_2\hat{\mathbf{x}}_2}\mathbf{R}_{\hat{\mathbf{x}}_2\hat{\mathbf{x}}_2}^{-1}\mathbf{R}_{\hat{\mathbf{x}}_2\mathbf{x}_2}) = \det(\mathbf{P}_{\mathbf{x}_2\hat{\mathbf{x}}_2})/\mathbf{R}_{\hat{\mathbf{x}}_2\hat{\mathbf{x}}_2}$, we completes the proof of Theorem 1. □

Appendix F

UFCs designed by Algorithm 1

F.1 4-UFC

$$\mathbb{U}_2 = \begin{pmatrix} 1 \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -j \end{pmatrix}$$

F.2 8-UFC

$$\mathbb{U}_3 = \begin{pmatrix} 1 \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ \\ 1+j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1-j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 1-j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1+j \end{pmatrix}$$

F.3 16-UFC

$$\mathbb{U}_4 = \begin{pmatrix} 1 \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ \\ 1+j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1-j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 1-j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1+j \end{pmatrix}$$

$$\begin{pmatrix} 1+j \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ j \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ -j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 2j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -2j \end{pmatrix}$$

F.4 32-UFC

$$\mathbb{U}_5 = \begin{pmatrix} 1 \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ \\ 1+j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1-j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 1-j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1+j \end{pmatrix}$$

$$\begin{pmatrix} 1+j \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ j \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ -j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 2j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -2j \end{pmatrix}$$

$$\begin{pmatrix} 1+2j \\ \\ 1+j \end{pmatrix}, \begin{pmatrix} 1+2j \\ \\ -1-j \end{pmatrix}, \begin{pmatrix} 1+2j \\ \\ 1-j \end{pmatrix}, \begin{pmatrix} 1+2j \\ \\ -1+j \end{pmatrix}$$

$$\begin{pmatrix} 2+j \\ \\ 1+j \end{pmatrix}, \begin{pmatrix} 2+j \\ \\ -1-j \end{pmatrix}, \begin{pmatrix} 2+j \\ \\ 1-j \end{pmatrix}, \begin{pmatrix} 2+j \\ \\ -1+j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ \\ 2+2j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -2-2j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 2-2j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -2+2j \end{pmatrix}$$

$$\begin{pmatrix} 2+2j \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 2+2j \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 2+2j \\ \\ j \end{pmatrix}, \begin{pmatrix} 2+2j \\ \\ -j \end{pmatrix}$$

F.5 64-UFC

$$\mathbb{U}_6 = \begin{pmatrix} 1 \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ \\ 1+j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1-j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 1-j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1+j \end{pmatrix}$$

$$\begin{pmatrix} 1+j \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ j \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ -j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 2j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -2j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ \\ 1+2j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1-2j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 1-2j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -1+2j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ \\ 2+j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -2-j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ 2-j \end{pmatrix}, \begin{pmatrix} 1 \\ \\ -2+j \end{pmatrix}$$

$$\begin{pmatrix} 1+2j \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 1+2j \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 1+2j \\ \\ j \end{pmatrix}, \begin{pmatrix} 1+2j \\ \\ -j \end{pmatrix}$$

$$\begin{pmatrix} 2+j \\ \\ 1 \end{pmatrix}, \begin{pmatrix} 2+j \\ \\ -1 \end{pmatrix}, \begin{pmatrix} 2+j \\ \\ j \end{pmatrix}, \begin{pmatrix} 2+j \\ \\ -j \end{pmatrix}$$

$$\begin{pmatrix} 1+j \\ \\ 1+2j \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ -1-2j \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ 1-2j \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ -1+2j \end{pmatrix}$$

$$\begin{pmatrix} 1+j \\ \\ 2+j \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ -2-j \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ 2-j \end{pmatrix}, \begin{pmatrix} 1+j \\ \\ -2+j \end{pmatrix}$$

$$\begin{pmatrix} 1+2j \\ \\ 1+j \end{pmatrix}, \begin{pmatrix} 1+2j \\ \\ -1-j \end{pmatrix}, \begin{pmatrix} 1+2j \\ \\ 1-j \end{pmatrix}, \begin{pmatrix} 1+2j \\ \\ -1+j \end{pmatrix}$$

$$\begin{pmatrix} 2+j \\ \\ 1+j \end{pmatrix}, \begin{pmatrix} 2+j \\ \\ -1-j \end{pmatrix}, \begin{pmatrix} 2+j \\ \\ 1-j \end{pmatrix}, \begin{pmatrix} 2+j \\ \\ -1+j \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 2+2j \end{pmatrix}, \begin{pmatrix} 1 \\ -2-2j \end{pmatrix}, \begin{pmatrix} 1 \\ 2-2j \end{pmatrix}, \begin{pmatrix} 1 \\ -2+2j \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1+2j \end{pmatrix}, \begin{pmatrix} 2 \\ -1-2j \end{pmatrix}, \begin{pmatrix} 2 \\ 1-2j \end{pmatrix}, \begin{pmatrix} 2 \\ -1+2j \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 2+j \end{pmatrix}, \begin{pmatrix} 2 \\ -2-j \end{pmatrix}, \begin{pmatrix} 2 \\ 2-j \end{pmatrix}, \begin{pmatrix} 2 \\ -2+j \end{pmatrix}$$

$$\begin{pmatrix} 2+2j \\ 1 \end{pmatrix}, \begin{pmatrix} 2+2j \\ -1 \end{pmatrix}, \begin{pmatrix} 2+2j \\ j \end{pmatrix}, \begin{pmatrix} 2+2j \\ -j \end{pmatrix}$$

Appendix G

Proof of Theorem 5

G.1 $n = 2$

The diagram of constellation $\mathcal{Z}_2$ is plotted on the complex plane in Fig. G.1. Since there are only four points, we have from Corollary 1 that $D(\mathbb{T}_{\mathcal{Z}_2}(\beta)) = g(1, \beta)$, where

$$g(1, \beta) = d(1, j, \beta) = \frac{\sqrt{2}\beta}{1 + \beta^2}$$

Notice that

$$D(\mathbb{T}_{\mathcal{Z}_2}(\beta)) = \frac{\sqrt{2}}{\beta^{-1} + \beta} \leq \frac{\sqrt{2}}{2\sqrt{\beta^{-1}\beta}} = \frac{\sqrt{2}}{2}$$

where the equality in the inequality is achieved when $\beta^{-1} = \beta$. Thus,

$$\begin{aligned} \hat{\beta} &= 1 \\ D(\mathbb{T}_{\mathcal{Z}_2}(\hat{\beta})) &= \frac{\sqrt{2}}{2} \end{aligned}$$

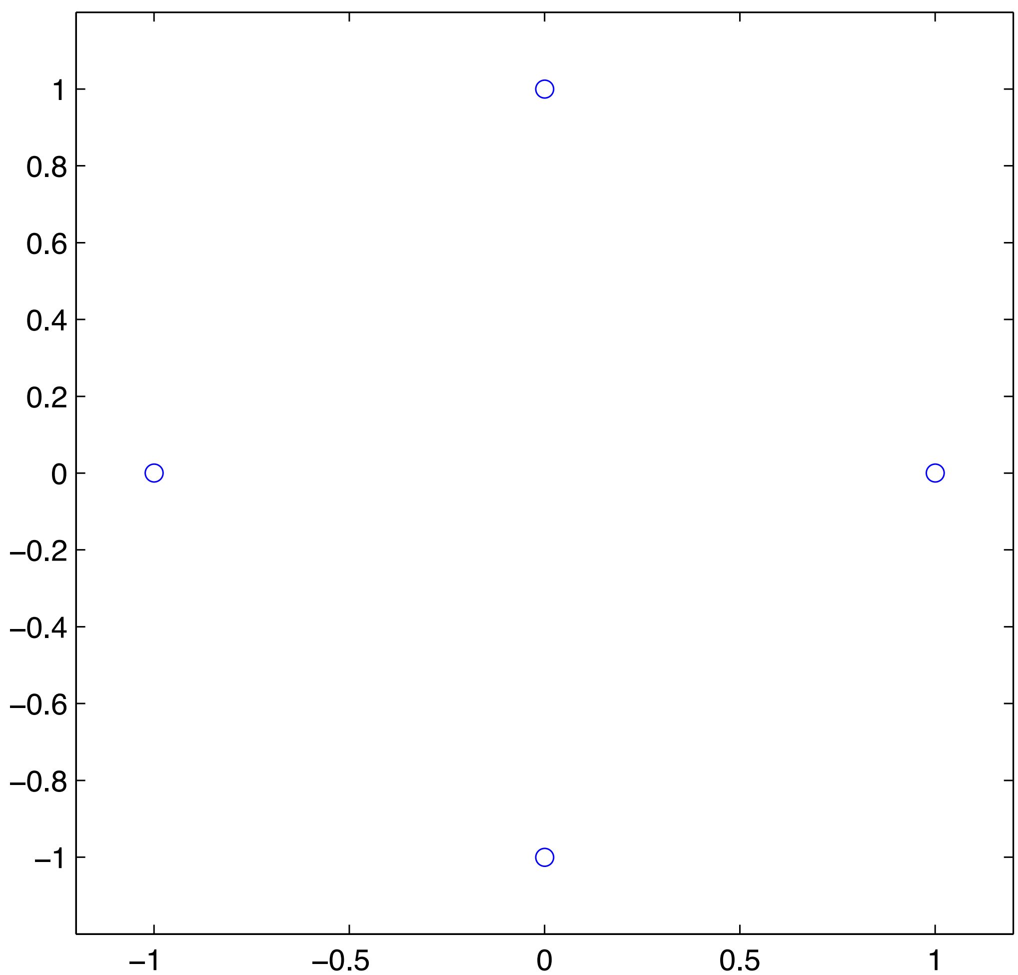


Figure G.1: 4 symbols training-equivalent UFC $\mathcal{Z}_2$

G.2 $n = 3$

The diagram of $\mathcal{Z}_3$ is shown in Fig. G.2. By Definition 3, it can be verified that $D(\mathbb{X}) = D(e^{j\theta}\mathbb{X})$. In other words, rotation does not change the distance. Since $e^{-j\pi/4}\mathcal{Z}_3$ has the same geometrical structure as the subset of $\mathcal{V}$ in Fig. 4.1, consisting of the outer 8 points v_k for $k = 1, 2, \cdots, 8$, using Corollary 1 gives us

$$
\begin{aligned}
g(1+j,\beta) &= d(1+j,1,\beta) = \frac{\beta}{\sqrt{1+2\beta^2}\sqrt{1+\beta^2}} \\
g(1,\beta) &= d(1,j,\beta) = \frac{\sqrt{2}\beta}{\sqrt{1+\beta^2}\sqrt{1+\beta^2}}
\end{aligned}
$$

and thus,

$$D(\mathbb{T}_{\mathcal{Z}_3}(\beta)) = \min\{g(1+j,\beta), g(1,\beta)\}$$

Since $g(z_1,\beta) < g(z_5,\beta)$ for any positive β, we have

$$D(\mathbb{T}_{\mathcal{Z}_3}(\beta)) = g(1+j,\beta) = \frac{1}{\sqrt{\beta^{-2}+2\beta^2+3}} \le \frac{1}{1+\sqrt{2}}$$

where we have used the geometrical and arithmetical mean inequality with the equality holding when $\beta^{-2} = 2\beta^2$. Therefore, we obtain

$$
\begin{aligned}
\hat{\beta} &= \frac{1}{\sqrt[4]{2}} \\
D(\mathbb{T}_{\mathcal{Z}_3}(\hat{\beta})) &= \frac{1}{1+\sqrt{2}}
\end{aligned}
$$

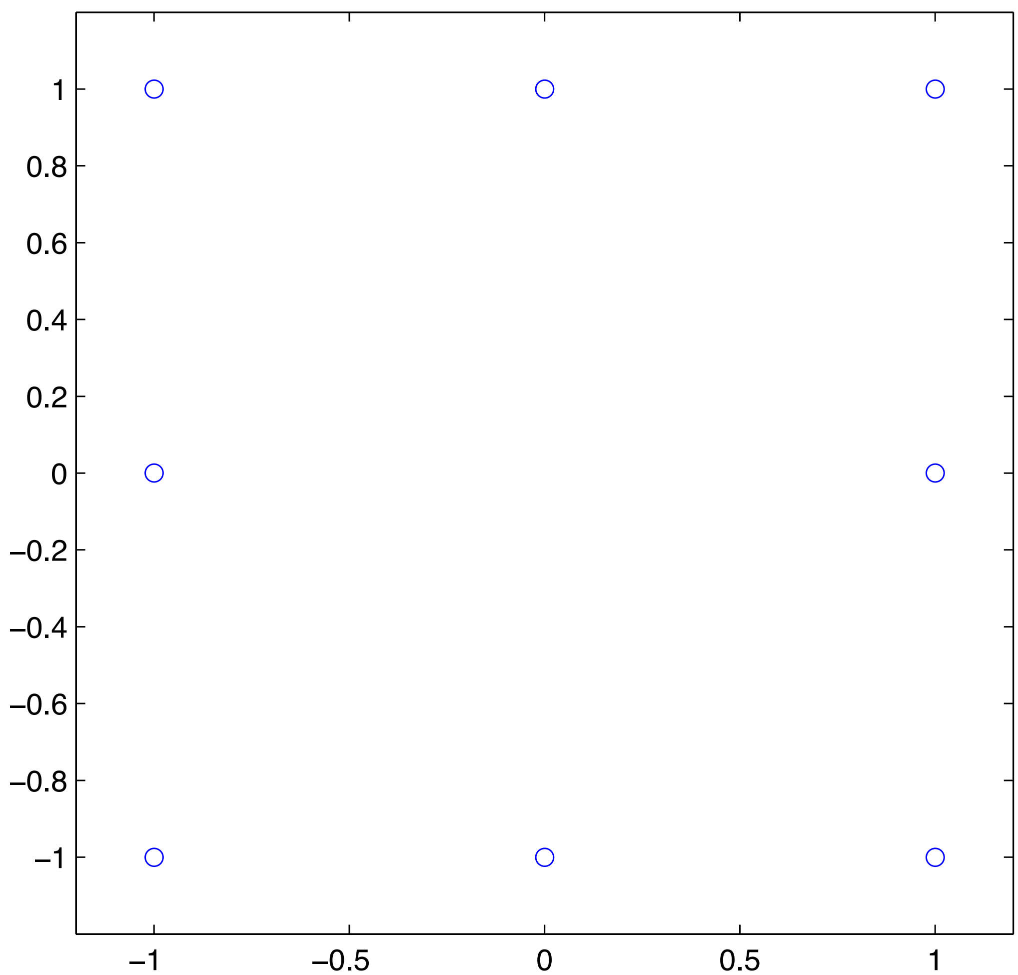


Figure G.2: 8 symbols training-equivalent UFC $\mathcal{Z}_3$

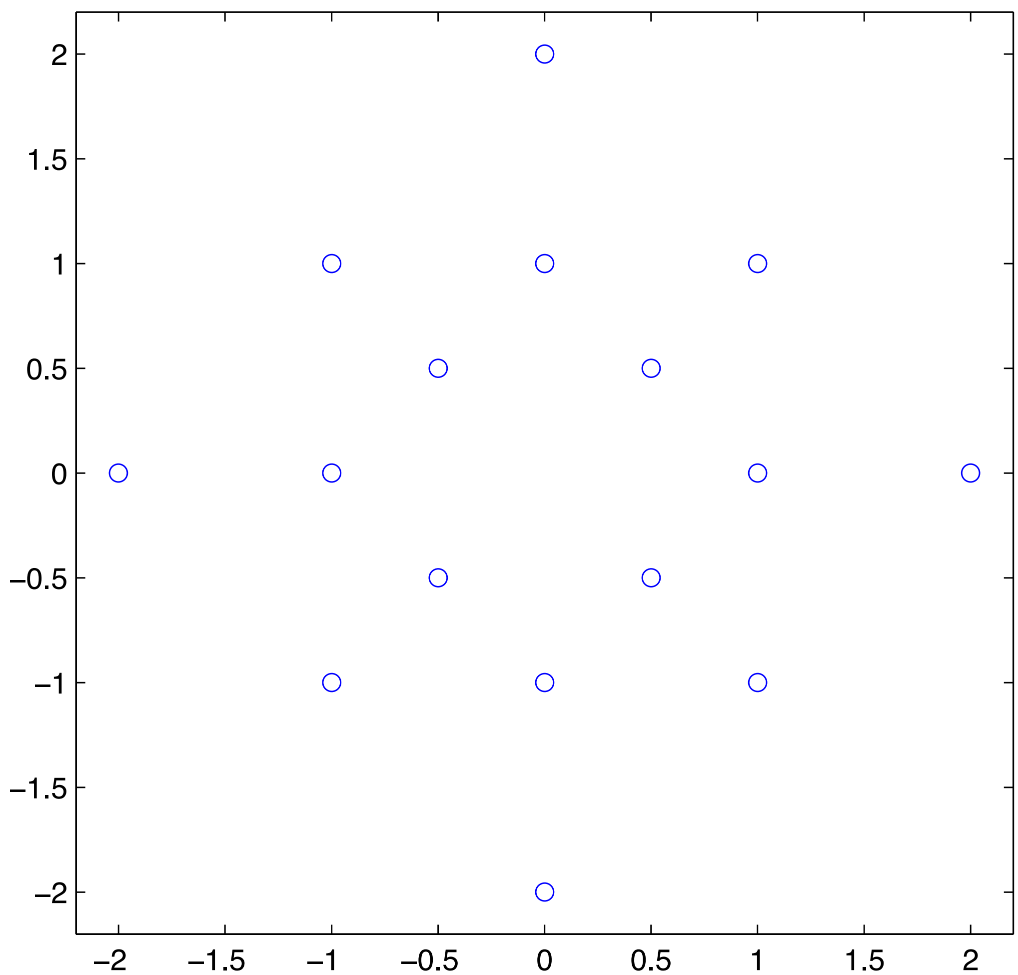


Figure G.3: 16 symbols training-equivalent UFC $\mathcal{Z}_4$

G.3 $n = 4$

For discussion clarity, the diagram of constellation $\mathcal{Z}_4$ is shown in Fig. G.3. There are four layers and each layer contains 4 symbols with equal energy. First, utilizing Lemma 4 and Corollary 1 with $v_1 = 2$ yields

$$g(2, \beta) = \frac{\beta}{\sqrt{1+\beta^2}\sqrt{1+4\beta^2}}$$

Then, delete the most outer 4 points, $2, -2j, -2, 2j$ and consider the remaining set, $\mathcal{V}_1 = \mathcal{Z}_4 - \{2, -2j, -2, 2j\}$. Now, applying Lemma 4 and Corollary 1 into $e^{-j\pi/4}\mathcal{V}_1$ with starting point $\sqrt{2}$ produces

$$g(1+j, \beta) = \frac{\frac{\sqrt{2}}{2}\beta}{\sqrt{1+0.5\beta^2}\sqrt{1+2\beta^2}}$$

Then, delete 4 points, $1+j, 1-j, -1-j, -1+j$, from $\mathcal{V}_1$ and consider the remaining set $\mathcal{V}_2 = \mathcal{V}_1 - \{1+j, 1-j, -1-j, -1+j\}$. Again, applying Lemma 4 and Corollary 1 into $\mathcal{V}_2$ with starting point 1 yields

$$g(1, \beta) = \frac{\frac{\sqrt{2}}{2}\beta}{\sqrt{1+0.5\beta^2}\sqrt{1+\beta^2}}$$

Following the same strategy, after we have deleted another four points, $1, -j, -1, j$, from $\mathcal{V}_2$, we find that there are only four points left with the same geometrical structure as the case of $n = 2$ and hence,

$$g(0.5+0.5j, \beta) = \frac{\beta}{\sqrt{1+0.5\beta^2}\sqrt{1+0.5\beta^2}}$$

Overall, the distance of the constellation $\mathbb{T}_{\mathcal{Z}_4}(\beta)$ is determined by

$$D(\mathbb{T}_{\mathcal{Z}_4}(\beta)) = \min\{g(2,\beta), g(1+j,\beta), g(1,\beta), g(0.5+0.5j,\beta)\}$$

Since $g(2,\beta) < g(0.5+0.5j,\beta)$ and $g(1+j,\beta) < g(1,\beta)$ for any $\beta > 0$, we attain

$$D(\mathbb{T}_{\mathcal{Z}_4}(\beta)) = \min\{g(2,\beta), g(1+j,\beta)\}$$

On the other hand, when $\beta \leq \frac{1}{\sqrt[4]{2}}$, $g(2,\beta) \leq g(1+j,\beta)$, whereas $g(2,\beta) \geq g(1+j,\beta)$ when $\beta \geq \frac{1}{\sqrt[4]{2}}$, we have

$$D(\mathbb{T}_{\mathcal{Z}_4}(\beta)) = \begin{cases} g(1+j,\beta) = \frac{\frac{\sqrt{2}}{2}\beta}{\sqrt{1+0.5\beta^2}\sqrt{1+2\beta^2}}, & \beta \leq \frac{1}{\sqrt[4]{2}} \\ g(2,\beta) = \frac{\beta}{\sqrt{1+\beta^2}\sqrt{1+4\beta^2}}, & \beta > \frac{1}{\sqrt[4]{2}} \end{cases}$$

Since $g(1+j,\beta)$ is monotonically increasing when $\beta \leq \frac{1}{\sqrt[4]{2}}$ and $g(2,\beta)$ is monotonically decreasing when $\beta > \frac{1}{\sqrt[4]{2}}$, the maximum of $D(\mathbb{T}_{\mathcal{Z}_4}(\beta))$ is obtained at this turning point $\beta = \frac{1}{\sqrt[4]{2}}$. Therefore, we arrive at the fact that

$$\begin{aligned} \hat{\beta} &= \frac{1}{\sqrt[4]{2}} \\ D(\mathbb{T}_{\mathcal{Z}_4}(\hat{\beta})) &= \frac{1}{\sqrt{5+3\sqrt{2}}} \end{aligned}$$

G.4 $n = 5$

The diagram of constellation $\mathcal{Z}_5$ is plotted in Fig. G.4. Following the argument similar

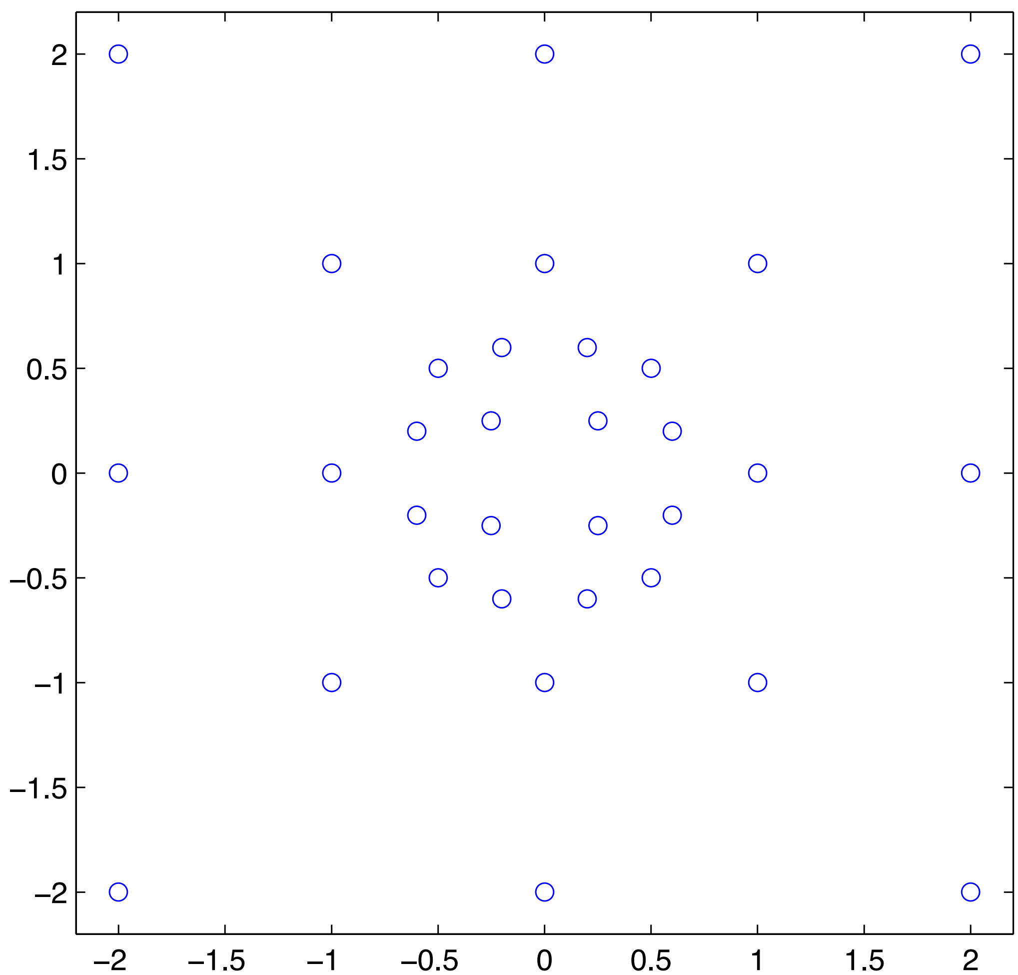


Figure G.4: 32 symbols training-equivalent UFC $\mathcal{Z}_5$

to the case of $n = 4$ and taking advantage of Lemma 4 and Corollary 1 three times lead to

$$\begin{aligned} g(2+2j,\beta) &= \frac{\sqrt{2}\beta}{\sqrt{1+2\beta^2}\sqrt{1+8\beta^2}} \\ g(2,\beta) &= \frac{\beta}{\sqrt{1+\beta^2}\sqrt{1+4\beta^2}} \\ g(1+j,\beta) &= \frac{\frac{\sqrt{2}}{2}\beta}{\sqrt{1+0.5\beta^2}\sqrt{1+2\beta^2}} \end{aligned}$$

In addition, it can be verified that $j, -j, -1, -0.5+0.5j, -0.5-0.5j$ are not closer to $z = 1$ than points $z = 0.5+0.5j$ or $z = 0.5-0.5j$ are, whose distances to $z = 1$ are equal, given by

$$d(1, 0.5-0.5j, \beta) = d(1, 0.5+0.5j, \beta) = \frac{\frac{\sqrt{2}}{2}\beta}{\sqrt{1+0.5\beta^2}\sqrt{1+\beta^2}}$$

whereas the distances of two points $0.6+0.2j$ and $0.6-0.2j$ to 1 are also equal, i.e.,

$$d(1, 0.6+0.2j, \beta) = d(1, 0.6-0.2j, \beta) = \frac{\frac{\sqrt{5}}{5}\beta}{\sqrt{1+0.4\beta^2}\sqrt{1+\beta^2}} < d(1, 0.5+0.5j, \beta)$$

For the other points z inside the disk $|z| < |0.6+0.2j|$, since $|1-z| > |1-(0.6+0.2j)|$, we have $d(1, z, \beta) > d(1, 0.6+0.2j, \beta)$. Therefore, we obtain

$$g(1,\beta) = \frac{\frac{\sqrt{5}}{5}\beta}{\sqrt{1+0.4\beta^2}\sqrt{1+\beta^2}}$$

Similarly, we can attain

$$\begin{aligned}
g(0.5+0.5j,\beta) &= \frac{\frac{\sqrt{10}}{10}\beta}{\sqrt{1+0.4\beta^2}\sqrt{1+0.5\beta^2}} \\
g(0.6+0.2j,\beta) &= \min\left(\frac{\frac{\sqrt{8}}{8}\beta}{\sqrt{1+0.125\beta^2}\sqrt{1+0.4\beta^2}}, \frac{0.4\beta}{\sqrt{1+0.4\beta^2}\sqrt{1+0.4\beta^2}}\right) \\
g(0.2+0.6j,\beta) &= g(0.6+0.2j,\beta) \\
g(0.25+0.25j,\beta) &= \frac{0.5\beta}{\sqrt{1+0.125\beta^2}\sqrt{1+0.125\beta^2}}
\end{aligned}$$

Over all, the distance of the scaled training-equivalent constellation, $D(\mathbb{T}_{\mathcal{Z}_5}(\beta))$ is the minimum among the above eight functions. However, since $g(0.25+0.25j,\beta) > g(2,\beta)$ and $g(1+j,\beta), g(1,\beta), g(0.6+0.2j,\beta), g(0.2+0.6j,\beta) > g(0.5+0.5j,\beta)$ for any $\beta > 0$, then, $D(\mathbb{T}_{\mathcal{Z}_5}(\beta))$ can be simplified to

$$D(\mathbb{T}_{\mathcal{Z}_5}(\beta)) = \min\left\{g(2+2j,\beta), g(2,\beta), g(0.5+0.5j,\beta)\right\} \tag{G.33}$$

In addition, notice that

$$\begin{aligned}
g(0.5+0.5j,\beta) \le g(2+2j,\beta) \quad &\Longleftrightarrow \quad 0 < \beta \le \sqrt{\frac{2+\sqrt{61}}{6}} \\
g(0.5+0.5j,\beta) \le g(2,\beta) \quad &\Longleftrightarrow \quad 0 < \beta \le \sqrt{1+\sqrt{5.5}} \\
g(2+2j,\beta) \le g(2j,\beta) \quad &\Longleftrightarrow \quad 0 < \beta \le \frac{\sqrt[4]{2}}{2}
\end{aligned}$$

Equation (G.33) can be further simplified into

$$D(\mathbb{T}_{\mathcal{Z}_5}(\beta)) = \begin{cases} g(0.5+0.5j,\beta) = \frac{\frac{\sqrt{10}}{10}\beta}{\sqrt{1+0.4\beta^2}\sqrt{1+0.5\beta^2}}, & \beta \leq \sqrt{\frac{2+\sqrt{61}}{6}} \\ g(2+2j,\beta) = \frac{\sqrt{2}\beta}{\sqrt{1+2\beta^2}\sqrt{1+8\beta^2}}, & \beta > \sqrt{\frac{2+\sqrt{61}}{6}} \end{cases}$$

Because of the fact that $g(0.5+0.5j,\beta)$ is monotonically increasing when $\beta \leq \sqrt{\frac{2+\sqrt{61}}{6}}$ and $g(2+2j,\beta)$ is monotonically decreasing when $\beta > \sqrt{\frac{2+\sqrt{61}}{6}}$, the maximum of $D(\mathbb{T}_{\mathcal{Z}_5}(\beta))$ is achieved when $\beta = \sqrt{\frac{2+\sqrt{61}}{6}}$. Thus, we have

$$\begin{aligned} \hat{\beta} &= \sqrt{\frac{2+\sqrt{61}}{6}} \\ D(\mathbb{T}_{\mathcal{Z}_5}(\hat{\beta})) &= \sqrt{\frac{57}{79\sqrt{61}+431}} \end{aligned}$$

G.5 $n = 6$

The diagram of constellation is shown in Fig. G.5. Following the strategy much

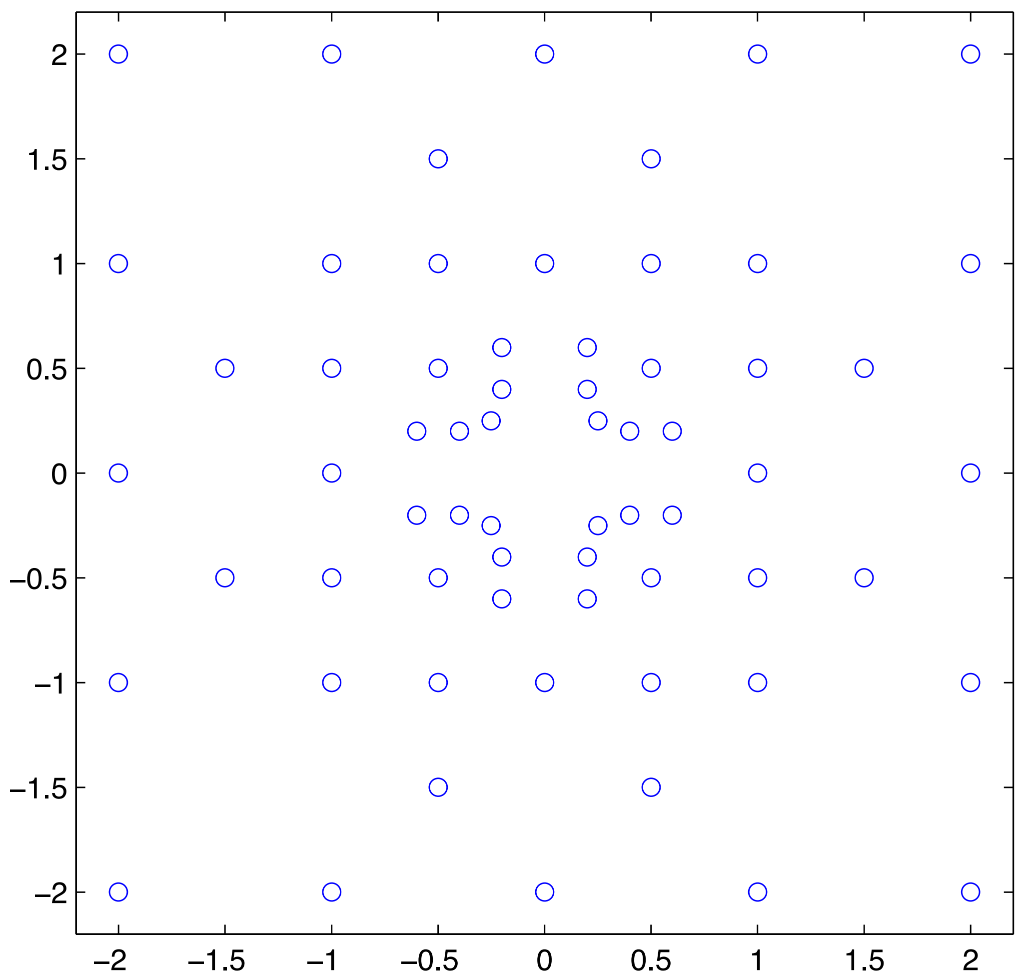


Figure G.5: 64 symbols training-equivalent UFC $\mathcal{Z}_6$

similar to the cases of $n = 4$ and $n = 5$, we can obtain

$$
\begin{aligned}
g(2+2j,\beta) &= \frac{\beta}{\sqrt{1+8\beta^2}\sqrt{1+5\beta^2}} \\
g(1+2j,\beta) &= \frac{\frac{\sqrt{2}}{2}\beta}{\sqrt{1+5\beta^2}\sqrt{1+2.5\beta^2}} \\
g(2+j,\beta) &= g(1+2j,\beta) \\
g(2,\beta) &= \frac{\frac{\sqrt{2}}{2}\beta}{\sqrt{1+4\beta^2}\sqrt{1+2.5\beta^2}} \\
g(1.5+0.5j,\beta) &= \frac{0.5\beta}{\sqrt{1+2.5\beta^2}\sqrt{1+1.25\beta^2}} \\
g(0.5+1.5j,\beta) &= g(1.5+0.5j,\beta) \\
g(1+j,\beta) &= \frac{0.5\beta}{\sqrt{1+2\beta^2}\sqrt{1+1.25\beta^2}} \\
g(1+0.5j,\beta) &= \frac{0.5\beta}{\sqrt{1+1.25\beta^2}\sqrt{1+\beta^2}} \\
g(0.5+j,\beta) &= g(1+0.5j,\beta) \\
g(1,\beta) &= \frac{\frac{\sqrt{5}}{5}\beta}{\sqrt{1+\beta^2}\sqrt{1+0.4\beta^2}} \\
g(0.5+0.5j,\beta) &= \frac{\frac{\sqrt{10}}{10}\beta}{\sqrt{1+0.5\beta^2}\sqrt{1+0.4\beta^2}} \\
g(0.6+0.2j,\beta) &= \frac{0.2\beta}{\sqrt{1+0.4\beta^2}\sqrt{1+0.2\beta^2}} \\
g(0.2+0.6j,\beta) &= g(0.6+0.2j,\beta) \\
g(0.4+0.2j,\beta) &= \frac{\frac{\sqrt{10}}{20}\beta}{\sqrt{1+0.2\beta^2}\sqrt{1+0.125\beta^2}} \\
g(0.2+0.4j,\beta) &= g(0.4+0.2j,\beta) \\
g(0.25+0.25j,\beta) &= \frac{0.5\beta}{\sqrt{1+0.125\beta^2}\sqrt{1+0.125\beta^2}}
\end{aligned}
$$

The distance of $\mathcal{Z}_6$ is the minimum among all the above 11 functions. Comparing any two of them results in

$$D(\mathbb{T}_{\mathcal{Z}_6}(\beta)) = \begin{cases} g(0.4+0.2j,\beta) = \frac{\frac{\sqrt{10}}{20}\beta}{\sqrt{1+0.2\beta^2}\sqrt{1+0.125\beta^2}}, & \beta \leq 1; \\ g(2+2j,\beta) = \frac{\beta}{\sqrt{1+8\beta^2}\sqrt{1+5\beta^2}}, & \beta > 1; \end{cases}$$

Using the same argument, we conclude that the maximum of $D(\mathbb{T}_{\mathcal{Z}_6}(\beta))$ is reached at the turning point, i.e.,

$$\begin{aligned} \hat{\beta} &= 1 \\ D(\mathbb{T}_{\mathcal{Z}_6}(\hat{\beta})) &= \frac{1}{3\sqrt{6}} \end{aligned}$$

This completes the proof of Theorem 5. $\square$

Bibliography

[1] T. M. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, Sep. 1979.

[2] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity — Part I: System description," *IEEE Trans. Comm.*, vol. 51, pp. 1927–1938, Nov. 2003.

[3] ——, "User cooperation diversity — Part II: Implementation aspects and performance analysis," *IEEE Trans. Comm.*, vol. 51, pp. 1939–1948, Nov. 2003.

[4] K. Azarian, H. E. Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4152–4157, Dec. 2005.

[5] Y. Chang and Y. Hua, "Diversity analysis of orthogonal space-time modulation for distributed wireless relays," in *Int. Conf. Acoust., Speech, Signal Process.*, vol. 4, Montreal, Canada, April 2004, pp. 561–564.

[6] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 3524–3536, Dec. 2006.

[7] G. Wang, J. K. Zhang, M. Amin, and K. M. Wong, "Nested distributed space-time encoding protocol for wireless networks with high energy efficiency," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 521–531, Feb. 2008.

[8] Y. Ding, J. K. Zhang, and K. M. Wong, "The amplify-and-forward half-duplex cooperative systems: pairwise error probability and precoder design," *IEEE Trans. Signal Processing*, vol. 55, pp. 605–617, Feb. 2007.

[9] ——, "Optimal precoder for amplify-and-forward half-duplex relay system," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2890–2895, Aug. 2008.

[10] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2415–2425, Oct. 2003.

[11] ——, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2062–3080, Dec. 2004.

[12] T. L. Marzetta, "BLAST training: estimation channel characteristics for high-capacity space-time wireless," in *Proc. 37th Annu. Allerton Conf. Communications, Control, and Computing*, Sept. 22-24 1999.

[13] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theory*, vol. 49, pp. 951–963, Apr. 2003.

[14] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 359–383, Feb. 2002.

[15] Q. Zhao and H. Li, "Performance of differential modulation with wireless relays in Rayleigh fading channels," *IEEE Trans. Commun. Letters*, vol. 9, pp. 343–345, Apr. 2005.

[16] R. Annavajjala, P. C. Cosman, and L. B. Miltstein, "On the performance of optimum noncoherent amplify-and-forward reception for cooperative diversity," in *Proc. MILCOM*, Oct. 2005.

[17] M. G. Souryal, "Non-coherent amplify-and-forward relay generalized likelihood ratio test receiver," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2020–2327, Jul. 2010.

[18] G. Farhadi and N. C. Beaulieu, "A low complexity receiver for noncoherent space-time codes in amplify-and-forward cooperative systems," *IEEE Trans. Commun.*, vol. 58, pp. 2499–2504, Sept. 2010.

[19] S. Ma and J.-K. Zhang, "Noncoherent diagonal distributed space-time block codes for amplify-and-forward half-duplex cooperative relay channels," *IEEE Trans. Vehicular Technology*, vol. 60, pp. 2400–2405, June 2011.

[20] J. S. Richters, *"Communication over fading dispersive channels"*. MIT Res. Lab. Electronics, Cambridge, MA, Tech. Rep. 464, Nov. 30, 1967.

[21] I. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1290–1301, May 2001.

[22] M. C. Gursoy, H. V. Poor, and S. Verdu, "The noncoherent Rician fading channelPart I: structure of the capacity-achieving input," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2193–2206, Sept. 2005.

[23] ——, "Noncoherent Rician fading channelPart II: spectral efficiency in the low-power regime," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2207–2221, Sept. 2005.

[24] I. Teletar, "Capacity of multi-antenna Gaussian channels," *AT & T Bell Labs, Tech. Rep.*, June. 1995.

[25] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antenna," *Wireless Personal Communications*, vol. 6, pp. 311–335, March 1998.

[26] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat-fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.

[27] G. D. Forney, "Coset codes–Part I: Introduction and geometrical classification," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1123–1151, Sept. 1988.

[28] ——, "Coset codes–Part II: Binary lattices and related codes," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1152–1187, Sept. 1988.

[29] G. D. Forney and L.-F. Wei, "Multidimensional constellations–Part I: Introduction, figures of merit, and generalized cross constellations," *IEEE Journal on Selected Areas in Communications*, vol. 7, pp. 877–892, Aug. 1989.

[30] G. D. Forney, “Multidimensional constellations–Part II: Voronoi constellations,” *IEEE Journal on Selected Areas in Communications*, vol. 7, pp. 941–958, Aug. 1989.

[31] J. H. Conway and N. J. A. Sloane, *Sphere Packing, Lattices and Groups.* New York: Springer-Verlag, 1998.

[32] 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.

[33] R. G. Gallager, *Principles of Digital Communications.* Cambridge: Cambridge University Press, 2008.

[34] M. K. Simon and J. G. Smith, “Hexagonal multiple phase-and-amplitude-shift keyed signal sets,” *IEEE Trans. Commun.*, vol. 21, pp. 1108–1115, Oct. 1973.

[35] S. H. Han, J. M. Cioffi, and J. H. Lee, “On the use of hexagonal constellation for peak-to-average power ratio reduction of an OFDM signal,” *IEEE Trans. Wireless Commun.*, vol. 7, pp. 781–786, Mar. 2008.

[36] W. H. Mow, “Fast decoding of the hexagonal lattice with applications to power efficient multi-level modulation systems,” in *Proc. Singapore, ICCS/ISITA*, Singapore, Nov. 1992, pp. 370–373.

[37] 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.

[38] V. Tarokh and I.-M. Kim, "Existence and construction of noncoherent unitary space-time codes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 3112–3117, Dec. 2002.

[39] M. M. McCloud, M. Brehler, and M. K. Varanasi, "Signal design and convolutional coding for noncoherent space-time communication on the block-Rayleigh-fading channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1186–1194, May 2002.

[40] D. Warrier and U. Madhow, "Spectrally efficient noncoherent communication," *IEEE Trans. Inform. Theory*, vol. 48, pp. 651–668, March 2002.

[41] W. Zhao, G. Leus, and G. Giannakis, "Orthogonal design of unitary constellations for uncoded and trellis coded non-coherent space-time systems," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1319–1327, June 2004.

[42] I. Kammoun and J. C. Belfiore, "A new family of Grassmann space-time codes for non-coherent MIMO systems," *IEEE Commun. Letters*, vol. 7, pp. 528–530, Nov. 2003.

[43] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receiver in Rayleigh-fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2383–2999, Sept. 2001.

[44] Y. Jing and B. Hassibi, "Unitary space-time modulation via Cayley transformation," *IEEE Trans. Signal Processing*, vol. 51, pp. 2891–2904, Nov. 2003.

[45] R. Gohary and T. Davidson, "Noncoherent mimo communication: Grassmannian constellations and efficient detection," *IEEE Trans. Inform. Theory*, vol. 55, pp. 1176 –1205, March 2009.

[46] B. Hassibi and B. M. Hochwald, "Cayley differential unitary space-time codes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1485–1503, June 2002.

[47] L. Xiong and J.-K. Zhang, "Submitted to least square error detection for noncoherent cooperative relay systems," *IEEE Trans. Vehicular Technology*, Aug. 2011.

[48] ——, "Submitted to energy-efficient uniquely-factorable constellation designs for noncoherent SIMO channels," *IEEE Trans. Vehicular Technology*, Aug. 2011.

[49] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM J. Num. Anal.*, vol. 10, pp. 413–432, 1973.

[50] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory.* London, UK: Academic Press Limited, 1990.

[51] G. Golub and C. F. V. Loan, *Matrix Computations (The third edition).* 2715 North Charles street, Baltimore Maryland: The John Hopkins University Press, 1996.

[52] N. N. Lebedev, *Special functions and their applications.* Englewood Cliffs, NJ: Prentice-Hall.

[53] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge, MA: Cambridge University Press, 1985.

[54] J.-K. Zhang, F. Huang, and S. Ma, "Full diversity blind space-time block codes," *accepted by IEEE Trans. Inform. Theory.*

[55] F. Oggier and B. Hassibi, "Algebraic differential Cayley space-time codes," *IEEE Trans. Inform. Theory*, vol. 53, no. 5, pp. 1911–1919, May 2007.

[56] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time codes for high date rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[57] M. O. Rabin and J. O. Shallit, "Randomized algorithms in number theory," *Communications on Pure and Applied Mathematics*, vol. 39, pp. 239–256, 1986.

[58] L.-K. Hua, *Introduction to Number Theory.* Berlin ; New York: Springer-Verlag, 1982.

[59] L. Xiong and J.-K. Zhang, *UFCs.* Department of Electrical and Computer Engineering, McMaster University, 2011, http://www.ece.mcmaster.ca/ jkzhang/.

[60] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1804–1824, Jul. 2002.

[61] P. Dayal, M. Brehler, and M. K. Varanasi, "Leveraging coherent space-time codes for noncoherent communication via training," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2058–2080, Sept 2004.