Optimal Finite Alphabet NOMA for Uplink Massive MIMO Channels

OPTIMAL FINITE ALPHABET NOMA FOR UPLINK MASSIVE MIMO CHANNELS

BY

YANG YU, B.Sc.

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 (2018) McMaster University
(Electrical & Computer Engineering) Hamilton, Ontario, Canada

TITLE: Optimal Finite Alphabet NOMA for Uplink Massive MIMO Channels

AUTHOR: Yang Yu
B.Sc., (Computer Science and Technology)
Zheng-Zhou University, Zheng Zhou, China

SUPERVISOR: Dr. Jian-Kang Zhang

NUMBER OF PAGES: ix, 58

Dedicate to my family and friends

Abstract

This thesis focuses a noncoherent two-user uplink system with each user having a single antenna and a base station equipped with a large number of antennas. It is assumed that small scale channel fading is Rayleigh fading and varies in every one time slot. For such massive MIMO uplink system, we consider an optimal finite-alphabet non-orthogonal multiple access (NOMA) design with each user utilizing nonnegative binary modulation. A fast noncoherent maximum likelihood (ML) detection algorithm for the sum constellation of the two users and a corresponding closed form symbol error probability (SEP) formula are derived. In addition, the lower and upper bounds on SEP are established to quantitatively characterize how quickly SEP decays when the number of base station antennas goes to infinity. Two important concepts: full receiver diversity and geometrical coding gain, are introduced. Particularly for two users and three users systems, with each user transmitting nonnegative binary constellation, we obtain an optimal closed form sum constellation that maximizes both the receiver diversity gain and geometrical coding gain. Computer simulations validate our theoretical analysis and demonstrate that our proposed optimal constellation attains significant performance gains over the currently available constellation design for the same massive MIMO upink system. Our future work is to develop an algorithm for devising an optimal AUDCG for the considered system in a more general case.

Acknowledgements

First of all, I would like to take this opportunity to express my deepest appreciation to my supervisor, Prof.Jian-Kang Zhang, for the patient guidance, encouragement and advices he has provided throughout my time as his student. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. This work would never have been possible without his help.

Secondly, I would like to thank all the members of our research group. In particular, I would like to express my thanks to Dr.Yan-Yu Zhang, Dr.Zheng Dong, and PhD candidate Pei-Yao Chen, for their kind help during my master program and for their valuable advices.

Furthermore, I would like to thank Dr.Dong-Mei Zhao and Dr.Jun Chen for being members in my defence committee. I appreciate their time for reviewing my thesis and providing valuable feedback.

Last but not least,I must express my very profound gratitude to my parents and my friends for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Notation and abbreviations

e.g.$\mathbf{A}$	Matrices are denoted by boldface characters with uppercase
e.g.$\mathbf{b}$	Column vectors are denoted by boldface characters with lowercase
$\mathbf{A}^T$	The transpose of $\mathbf{A}$
$\mathbf{A}^H$	The Hermitian transpose of $\mathbf{A}$ (i.e., the conjugate and transpose of $\mathbf{A}$)
b_m	The m-th entry of $\mathbf{b}$
$\lVert\mathbf{b}\rVert$	The Euclidean norm of $\mathbf{b}$
$\otimes$	The Kronecker product and $j = \sqrt{-1}$
$\mathbf{a} \leq \mathbf{b}$	$a_k \leq b_k, k = 1, 2, \ldots, n$ with $\mathbf{a}, \mathbf{b} \in \mathbb{R}^{n \times 1}$
$\mathbf{0}$	Column vectors with all zero entries of appropriate dimension
$\mathbf{1}$	Column vectors with all one entries of appropriate dimension

Contents

List of Figures

Chapter 1

Introduction

Today, the recent arrival of big data and artificial intelligence information age has created an explosive demand for knowledge and information exchange between human and human, between human and machine, as well as between machine and machine. This demand has triggered off an tremendous expansion in wireless technologies in which severe technical challenges including the need of transmitting data at huge rates as well as with ultra-reliable low latency, i.e., ultra-reliable low latency communications (URLLC), and the requirement of wireless communications being ubiquitous and indispensable by connecting everyone and almost every device, i.e., massive connection, has been encountered. Recent research on wireless communications has strongly demonstrated that two major potential technologies to meet these three challenges are a non-orthogonal multi-access (NOMA) technology and a massive multiple input and multiple output (MIMO) technology.

1.1 Massive MIMO

Thus far, it has been world-widely recognized that a MIMO system equipped with multiple antennas is an important and very promising technology in wireless communications, where the channel capacity and reliability can be substantially improved by exploiting multi-path scattering from antenna arrays equipped at both transmitter and receiver [56, 68]. Specifically, a massive MIMO (MaMIMO) system, by name, is a special MIMO system where either a transmitter or a receiver or both deploys a massive number of antennas. With recent advances in radio frequency chains and integrated circuit designs, it would be possible to make the MaMIMO systems become realistic in practical wireless communication systems. In recent research on a MaMIMO system, no matter wherever channel state information (CSI) is available, instead of addressing asymptotic system performance *in terms of signal-to-noise ratio (SNR)* in a traditional MIMO system, one always focuses on asymptotic system performance *in terms of the largest number of antennas*. In recent years, information-theoretic research has shown that many significant advantages are brought about in the MaMIMO architectures, such as even extremely higher energy efficiency and spectral efficiency than most of the available systems off-the-shelf [3, 13, 22, 35, 41, 48–50]. With the capability to create a large number of spatial diversity paths, MaMIMO has very recently been treated as one of the most promising techniques for enabling URLLC [7]. Technically speaking, it is due to the massive number of available antennas that the central limit theory plays a so natural and so important role in the MaMIMO systems that the original very complicated random communication situations become stable and simple, particularly in the scenario of multi-user network communications. For example, as the number of base station (BS) antennas tends to infinity, the cross-correlation between the channel vectors of different users will vanish in rich scattering environment (and it is also true for line-of-sight channel with enough angular separation), therefore, resulting in negligible mutual interference if the channel state information (CSI) is perfectly known at the BS. This significantly simplifies the overall system design in the multi-user massive MIMO systems [43, 54], e.g., the maximum-ratio combining (MRC) receiver would be sufficient to approach the sum-capacity. Therefore, MaMIMO techniques have become one of the most promising technologies to meet the ever-increasing bandwidth requirement and connectivity of future wireless communications [3, 6, 19, 35, 38, 43, 54], thus drawing tremendous attention from both researchers and industry [13, 28, 35, 54, 55].

However, enjoying all these aforementioned benefits provided by the MaMIMO heavily relies on the availability of CSI. Unfortunately, perfect CSI at the receiver, in practice, is not easily attainable. If the time between signal fades, i.e., *coherence time*, is sufficiently long, then, the transmitter can send training signals that allow the receiver to estimate the channel coefficients accurately [5, 25, 42]. For mobile wireless communications, however, the fading coefficients may change so rapidly that the coherence time may be too short to allow reliable estimation of the coefficients, especially in a system with a large number of antennas. Therefore, the time spent on sending training signals cannot be ignored because of the necessity of sending more training signals for accurate estimation of the channel [25]. This is exactly the communication situation we will encounter in this thesis, where we consider an uplink MAMIMO system consisting of multiple single antenna users and one base station (B-S) equipped with the large number of antennas, but the channel coefficients change in every time slot. As a consequence, if we use one time slot for training, even if it is not enough, then, we have not time to transmit any information symbols. In fact, such a single input single output channel model was early considered for studying its capacity by Abou-Faycal, Trott and Shamai in [1] and by Gursoy, Poor and Verdú in [21]. Recently, for such fast changing uplink MaMIMO channel, Chowdhury, Manolakos and Goldsmith [12] have proposed an energy-based noncoherent signaling scheme and receiver, in which no instantaneous CSI is needed at either the transmitter or the receiver, and the receiver decodes transmitted information after each received symbol by measuring only the average signal energy across all receiver antennas [11, 12, 30, 65]. Among them, the authors in [12] characterized the scaling law of the noncoherent massive SIMO systems with energy-based detectors over independent and identically distributed channel fading when the number of receiver antenna goes to infinity. The results showed that energy-based noncoherent schemes are able to achieve the same scaling behavior in terms of achievable rates as optimal coherent schemes with an increasing number of antennas. A simple constellation design scheme was also proposed for one-user and two-user cases, which maximized the minimum Euclidean distance of the sum constellation with the energy-based detector [12]. However, for the two-user case, the closed-form optimal solution was only obtained for some specific channels and low-order modulations, and numerical search is needed otherwise. More recently, Zhang et al. have generalized the results attained in [12] to the scenarios with arbitrary numbers of users and channel fading models in [65]. A more comprehensive analysis of the constellation design of noncoherent energy-based massive SIMO systems was conducted in [11], where the constellation design problem was formulated as the maximization of the error exponent of the SEP with respect to the number of receiver antennas, and the formulated problem was resolved under different assumptions on the availability and imperfection of CSI statistics. A similar constellation design problem was also considered in [30], which applied the central limit theorem, instead of the law of large numbers used in the previous papers, to attain a Gaussian approximation of the instantaneous received signal energy, which further leads to a closed-form expression of the detection threshold for the adopted energy-based detector. However, all the aforementioned designs applied certain approximation to the distribution of the received signal by resorting to the asymptotic channel orthogonality and hardening properties of massive MIMO systems. In practice, due to a finite number of receiver antennas, these properties are not strictly satisfied. As such, the approximations will result in error in the performance analysis and the associated constellation design. Furthermore, to the best knowledge of the authors, the optimal diversity and the optimal coding gain for the considered noncoherent systems using the optimal noncoherent maximum likelihood (ML) has not been well investigated. Besides, the condition for achieving the full diversity needs to be better understood. This is one of major motivations for this thesis.

1.2 Non-orthogonal Multiple Access

Multiple access technologies, i.e., methods of strategically distributing scarce spectral resources into multiple users so as to allow them to access a wireless media simultaneously, have been playing an important role in determining the performance of each generation of mobile communication systems. Basically speaking, depending on the manner in which the resources are allocated to users, multiple access technologies can be generally classified into categories: orthogonal multiple access

(OMA) and non-orthogonal multiple access (NOMA). The current generation of cellular networks, i.e., orthogonal frequency-division multiple access (OFDMA) for the fourth generation (4G), and all previous generations including frequency-division multiple access (FDMA) for the first generation (1G), time-division multiple access (T-DMA) for the second generation (2G) and code-division multiple access (CDMA) for the third generation (3G), have primarily adopted the OMA technologies. In all these OMA schemes, the resource is partitioned into orthogonal segments in a time/frequency/code domain, and each resource segment is then assigned to only one single user. Therefore, there is no inter-user interference, thus, leading to low-complexity receivers and scheduling algorithms. In fact, after the resource allocation, a multiple-user problem is divided into several point-to-point problems so that the currently well-established single-user encoder/decoder techniques can be directly applied. However, early information-theoretic studies showed that OMA has lower spectral efficiency than NOMA. In addition, OMA is not scalable as the total number of orthogonal resources and their granularity strictly limits the maximum number of served users.

Unlike OMA, NOMA mainly exploits the feature of each user's signalling to multiplex multiple users together so that they can be served in the same time/frequency/code resources. As such, with proper multi-user detection techniques to deal with the inter-user interference at the receiver side (e.g., successive interference cancellation (SIC) [7]), NOMA is capable of achieving improved spectral efficiency and serving much more users simultaneously. The recent research on NOMA has shown that NOMA has three major advantages, i.e., high spectral efficiency, massive connectivity and low latency and signalling overhead, over the OMA schemes. Despite the fact that the deployment of NOMA, as a new radio access technology for future mobile systems, is relatively new, the performance of NOMA has been studied extensively in the information theory society for various channel topologies such as broadcast channels (BC), multiple access channels and interference channels, with emphasis mainly on the investigation of the channel capacity region under the assumption of unlimited encoding/decoding complexity. Because of this extremely expensive implementation cost, the study of NOMA lies mostly in the theoretical aspects. Thanks to the rapid progress of the radio frequency chain and the processing capability of mobile devices in the past decades, the implementation of NOMA is becoming more and more feasible and thus, has drawn tremendous attention from both academia and industry very recently. In fact, a two-user downlink scenario of NOMA, known as multiuser superposition transmission, has already been incorporated in the 3rd Generation Partnership Project Long Term Evolution-Advanced. Therefore, NOMA has been actively predicted as a promising technology for future wireless communications. Up to now, the vast majority of existing NOMA designs assumed the use of Gaussian input signals. Although the Gaussian input is of great significance both theoretically and practically, its implementation in reality will require huge storage capacity, unaffordable computational complexity and extremely long decoding delay. In addition, the actual transmitted signals in realistic communication systems are drawn from finite-alphabet constellations, such as pulse amplitude modulation (PAM), quadrature amplitude modulation (QAM) and phase shift keying (PSK) modulation. Directly applying the results derived from the Gaussian inputs to the systems with finite-alphabet inputs could lead to a significant performance loss. Usually, the Gaussian input serves as a theoretical benchmark. Therefore, the NOMA design with finite-alphabet inputs is of utmost importance and has attracted considerable efforts. The main principle of these efforts is to ensure that the signal originated from each user can be uniquely decoded from the received sum-signal at the receiver side. Thus far, most of finite alphabet NOMA designs used mutual information as a performance measure and the resulting optimal solutions were specifically numerical [24], from which only limited insights of the relationship between the sum-constellation and each users constellation can be drawn. Essentially, the key to understanding multi-user communications is to understand multi-user interference. Now, it is widely accepted that interference is the central topic of multi-user network information theory [14, 20], where conventional approaches to manage interference are to suppress or to eliminate it. Actually, multiple access is a strategy to manage interference. The classical information-theoretic study on the two-user Gaussian interference channel [23] suggests us that we should treat the interference as noise when it is weak and that the optimal strategy is to decode the interference when it is very strong. In addition, when the level of the power of interference is the same as that of the power of a desired signal, one good strategy is to suppress all the undesired interference into a smaller space that has no overlap with the signal space [8, 9, 59]. Therefore, the original interference channel can be converted into several parallel additive white Gaussian noise (AWGN) channels. However, some recent proposals, which consider interference as a useful resource, attempt to use interference properly to build efficient future wireless communications [37, 47, 67]. For example, interference can be used for boosting up the desired signal [44–46] for energy harvesting [15, 26, 29, 34, 39, 52, 60, 61] or to deteriorate the signal of the eavesdropper for secure communication [66].

Inspired by the classic work on interference channels [4, 8], in this thesis, we are interested in understanding the management of interference from the viewpoint of estimation and detection theory. To this end, we would like to revisit some early seminal work on this concept [32, 33] of how to strategically take advantage of the finite alphabet properties of digital communication signals for managing interference for a two-user access binary channel. Essentially, Kasami and Lin's main idea is to carefully design such two finite length codes for the two users that when any sum binary signal of the two user codewords is received in a noiseless environment, each individual user codeword can be uniquely decoded, as well as in a noisy case, the resulting error is correctable. Specifically, such uniquely decodable code (UDC) was explicitly constructed for a two-user binary ensure channel [2, 57]. Then, this important concept was extended to the design of UDC based on trellis modulation for an N-user binary multi-access channel [10], which allows a number of users to access a common receiver simultaneously and outperforms the time sharing method in terms of error probability. Furthermore, the design of trellis-coded UDC was investigated in a complex number domain to extract the desired signal from the superposition of the signal and cochannel interference [36, 51, 62]. In addition, the concept of UDC was also exploited to design variety of multi-resolution modulation schemes for BC and it was shown that they not only outperformed the frequency division scheme by properly designing the resulting constellation [27, 40, 53, 58], but also reduced the transmission delay of the network at the cost of increased transmitting power for fading channels [27]. Recently, Dong et al. [16] developed a punched Farey sequence for the efficient design of an optimal NOMA scheme for a classical two-user uplink channel, with each user having a QAM signalling constraint that maximizes the minimum

Euclidean distance of the received sum-constellation with the ML detector. Very interestingly, the resulting optimal sum constellation is still QAM with an optimally power-loaded scale. The similar idea to this was also utilized for successfully designing an optimal NOMA scheme for the classical Z-channel [18]. In order to better understand the optimal signalling structure of the resulting sum constellation, Zhang et al. [64] used the same performance metric for the design of a finite alphabet downlink multiple-input single-output (MISO) NOMA scheme for a visible light communication system by jointly optimizing each user constellation and a beamformer. It turns out that the resulting optimal sum constellation is the nonnegative uniform PAM constellation scaled by an optimally power-loaded beamformer. However, all these results [16, 18, 65] were attained under the condition that CSI was known at both the transmitter and the receiver. More recently, as we have mentioned in the previous subsection, for our considered MaMIMO system, several simple design schemes for managing finite alphabet multi-user interference were proposed that maximized the minimum Euclidean distance of the sum constellation with the energy-based detector [12]. However, all the aforementioned designs were based on the resulting approximation additive white Gaussian noisy channel with the sum signal as an input by resorting to the asymptotic channel orthogonality and hardening properties of massive MIMO systems, thus, resulting in error in the performance analysis and the associated constellation design. In addition, the energy detector is not optimal. This is another major motivation for this thesis.

1.3 Main Contributions and Future Work

The above factors motivate and enlighten us to perform finite alphabet multi-user interference management from the perspective of digital signal processing. Specifically, in this thesis, we consider an uplink massive MIMO system consisting of one BS equipped with N antennas and K user terminals with each having a single antenna. In order to improve the spectral efficiency of the network, all the users are allowed to transmit their own signals simultaneously to BS with the same frequency. At the receiver, we use the optimal noncoherent ML detector to estimate the transmitted signals for all the users. Our main task in this thesis is to establish a new criterion based on the noncoherent ML detector and to devise such an optimal sum constellation using this criterion that any individual user signal must be uniquely determined from any received noise-free sum signal [24, 25]. In other words, all user signal sets must constitute an additively uniquely decomposable constellation group (AUDCG). Our main contributions of this thesis can be stated as follows:

1. A fast noncoherent maximum likelihood (ML) detection algorithm for the sum constellation of all users and a corresponding closed form symbol error probability (SEP) formula are derived.

2. Two important concepts: full receiver diversity gain and geometrical coding gain for the considered MaMIMO system, are introduced. In addition, the lower and upper bounds on SEP are established to quantitatively characterize how quickly SEP decays when the number of BS antennas goes to infinity, showing that any AUDCG enables the full receiver diversity for the noncoherent ML detector.

3. A novel criterion on the design of an optimal AUDCG for the considered uplink NOMA MaMIMO systems is proposed to maximize the geometrical coding gain. Particularly for two users and three users, with each having binary energy modulation, we attain a closed-form optimal AUDCG.

Our future work is to develop an efficient algorithm for devising an optimal AUDCG that maximizes the geometrical coding gain for the considered NOMA MaMIMO system in a more general case.

Chapter 2

Error Performance Analysis

2.1 System Model and Noncoherent ML Detection

Consider an uplink massive MIMO system consisting of one BS equipped with N antennas and K user terminals with each having a single antenna. In order to improve the spectral efficiency of the network, all the users are allowed to transmit their own signals simultaneously to BS with the same frequency. We consider a fast changing fading channel model in which the channel coefficients are assumed to change every one time slot. Then, a relationship between the input and the output in a baseband is represented by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2.1}$$

where $\mathbf{y}$ is an $N \times 1$ received signal vector, $\mathbf{n}$ is an $N \times 1$ noise matrix arising at BS and $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_K]$ is an $N \times K$ channel matrix, with $\mathbf{h}_k \in \mathbb{C}^{N\times 1}$ being the respective channel links from User k to the BS for $k = 1, 2, \cdots, K$ such that $\mathbf{h}_k = \sqrt{\beta_k}\tilde{\mathbf{h}}_k$.

Here, we assume that β_k are the corresponding large scale fading coefficients and known at both the transmitters and the receiver, and that $\tilde{\mathbf{h}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$. In addition, the $N \times 1$ noise vector $\mathbf{n}$ is assumed to be additive white circularly symmetric complex Gaussian, with each entry having zero mean and variance σ^2. The notation $\mathbf{x} = (x_1, x_2, \cdots, x_K)^T$ in (2.1) denotes a $K \times 1$ transmitted signal vector, where x_k is the transmitted symbol of User k and randomly and equally likely drawn from a finite alphabet constellation to be designed. Under the above assumptions, the probability density function of $\mathbf{y}$ conditioned on $\mathbf{x}$, $f(\mathbf{y}|\mathbf{x})$ is given by

$$f(\mathbf{y}|\mathbf{x}) = \frac{1}{\big(\pi(\sum_{k=1}^{K} \beta_k |x_k|^2 + \sigma^2)\big)^N} \exp\big(-\frac{\|\mathbf{y}\|^2}{\sum_{k=1}^{K} \beta_k |x_k|^2 + \sigma^2}\big),$$

which suggests that in order to assure reliable communications, each user only needs to transmit energy information ($x_k \geq 0$), since the phase information of the transmitted signal is lost. Hence, we can assume that the constellation of User k, $\mathcal{U}_k^{1/2}$ is denoted by $\mathcal{U}_k^{1/2} = \{\sqrt{u_{k,m}}\}_{i=1}^{M_k}$, where $0 \leq u_{k,1} < u_{k,2} < \cdots < u_{k,M_k}$, with each user having own average transmission power constraint: $\frac{1}{M_k} \sum_{m=1}^{M_k} u_{k,m} \leq p_k$. For discussion simplicity, let $\mathcal{E} = \{\sum_{k=1}^{K} \beta_k u_k + \sigma^2 : u_k \in \mathcal{U}_k\}$, which is called a received sum constellation. Now, we can see clearly that in order to achieve reliable communications for such channel, it is necessary to require that the resulting sum constellation must satisfy such additively unique decomposition property that each point in $\mathcal{E}$ is uniquely represented as the linear combination of all user signal points plus noise variance, with each coefficient being the corresponding large scale. This naturally leads us to recalling the following concept introduced in [17, 63, 64]

Definition 1 (AUDCG)**.** *A group of constellations $\{\mathcal{A}_i\}_{i=1}^{L}$ is said to form an additively uniquely decomposable constellation group (AUDCG), denoted by* $\{\sum_{i=1}^{L} \oplus a_i =$

$a_1 \oplus a_2 \oplus \cdots \oplus a_N : a_i \in \mathcal{A}_i\} = \oplus_{i=1}^{L}\mathcal{A}_i = \mathcal{A}_1 \oplus \mathcal{A}_2 \oplus \cdots \oplus \mathcal{A}_L$, *if* $a_i, \tilde{a}_i \in \mathcal{A}_i$ *for* $i = 1, 2, \cdots, L$ *such that* $\sum_{i=1}^{N} a_i = \sum_{i=1}^{L} \tilde{a}_i$*, then, we have* $a_i = \tilde{a}_i$ *for* $i = 1, 2, \cdots, L$. □

For discussion convenience, constellation $\oplus_{i=1}^{N}\mathcal{A}_i$ in Definition 1, denoted by $\mathcal{A}$, is called the *sum-constellation* of all $\mathcal{A}_i$ and each $\mathcal{A}_i$ is called the *i-th sub-constellation* of $\oplus_{i=1}^{N}\mathcal{A}_i$ or *i-th user constellation*. Therefore, in this thesis we need to design the sum constellation $\mathcal{E}$ such that $\mathcal{E} = \oplus_{k=1}^{K}\beta_k\mathcal{U}_k + \sigma^2$. To do that, let us write $\mathcal{E} = \{E_i\}_{i=1}^{M}$, where $M = \prod_{k=1} M_k$ and $E_1 < E_2 < \cdots < E_M$. Our task in this thesis is to design an optimal sum constellation using a noncoherent ML detector. It is known that for the noncoherent MIMO system considered in this thesis, the optimal detector is a noncoherent ML detector, which is to find E such that $f(\mathbf{y}|E)$ is maximized, i.e., $\hat{E} = \underset{E\in\mathcal{E}}{\operatorname{argmax}} f(\mathbf{y}|E)$. Solving this optimization problem is equivalent to solving the following optimization problem:

$$\hat{E} = \underset{E\in\mathcal{E}}{\operatorname{argmin}} \quad \frac{\|\mathbf{y}\|^2}{E} + N \ln E. \tag{2.2}$$

In order to simplify the ML detector (2.2) and to analyze its error performance, we need to establish the following two lemmas.

Lemma 1. *For* $r > 1$*, let two functions be defined by* $\tau(r) = \frac{\ln r}{r-1}$ *and* $\rho(r) = \frac{r \ln r}{r-1}$*, respectively. Then, we have* $0 < \tau(r) < 1$*,* $\rho(r) > 1$*, and* $\tau(r)$ *monotonically decreases and* $\rho(r)$ *monotonically increases.* □

Proof: Let function $f(t)$ be defined by $f(t) = \ln t - t + 1 \quad$ for $0 < t < 1$. Then, its first order derivative $f'(t) = \frac{1}{t} - 1 > 0$ for $0 < t < 1$. Since $f(1) = 0$, we have $f(t) < 0$. Therefore, $\tau'(r) = f\left(\frac{1}{r}\right)\frac{1}{(r-1)^2} < 0$ for $r > 1$, i.e., $u(r)$ monotonically decreases with r increasing. Since $\lim_{r\to 1}\tau(r) = 1$, we have $\tau(r) < 1$. In addition, it is not difficult to have $\tau(r) > 0$ for $r > 1$.

Similarly, in order to prove the statement on $\rho(r)$, let function $g(t)$ be defined by $g(t) = 2t^2 - 3t + 1 - \ln t,\ t > 1$. Then, we have $g'(t) = \frac{(4t+1)(t-1)}{t} > 0$ for $t > 1$, function $g(t)$ is monotonically increasing and thus, $g(t) > g(1) = 0$ for $t > 1$. Hence, we have $\rho'(r) = \frac{g(r)}{r(r-1)} > 0$, i.e., $\rho(r)$ monotonically increases as r increases and as a consequence, $\rho(r) > \rho(1) = 1$. This completes the proof of Lemma 1. □

Lemma 2. *Let* $d_k = \frac{E_k E_{k+1} \ln \frac{E_{k+1}}{E_k}}{E_{k+1} - E_k}$ *for* $k = 1, 2, \cdots, L-1$. *Then, we have* $E_i < d_i < E_{i+1}$. □

Proof: Let function $h(t)$ be defined by $h(t) = t \ln t - t + 1,\ t > 0$. Then, $h'(t) = \frac{1}{t} > 0$ for $t > 0$. Hence, we have $h(t) > h(1) = 0$ if $t > 1$. Since $\frac{E_{i+1}}{E_i} > 1$ and $E_i > 0$, we can attain $d_i - E_i = \frac{E_i^2}{E_{i+1} - E_i} \times h\left(\frac{E_{i+1}}{E_i}\right) > 0$ and $E_{i+1} - d_i = \frac{E_{i+1}^2}{E_{i+1} - E_i} \times h\left(\frac{E_{i+1}}{E_i}\right) > 0$. Thus, we have $E_i < d_i < E_{i+1}$, implying that $d_i < d_{i+1}$ for $i = 1, 2, \cdots, = L-2$. This completes the proof of Lemma 2. □

With the help of the above two lemmas, we are now in a position to present the following fast noncoherent ML detection algorithm for the received sum constellation $\mathcal{E}$.

Algorithm 1: Let the sequence d_i for $i = 1, 2, \cdots, L$ be defined in Lemma 2. Then, the optimal estimates of the transmitted signal using the noncoherent ML detector are determined as $\hat{E} = E_1$ if $\|\mathbf{y}\|^2/N \le d_1$, $\hat{E} = E_i$ if $d_{i-1} < \|\mathbf{y}\|^2/N \le d_i$ for $i = 2, \cdots, M-1$ and $\hat{E} = E_M$ if $\|\mathbf{y}\|^2/N > d_{M-1}$. □

Proof: For notational simplicity, let $L_i = \frac{\|\mathbf{y}\|^2}{E_i} + N \ln E_i$. Then, inequalities $L_i > L_1$ for $i = 2, 3, \cdots, L$ are equivalent to $\frac{\|\mathbf{y}\|^2}{N} < E_1 \rho\left(\frac{E_i}{E_1}\right)$ for $i = 2, 3, \cdots, M$, since $E_M > E_{M-1} > \cdots > E_1$. By Lemma 1, the above inequalities are equivalent to $\frac{\|\mathbf{y}\|^2}{N} < E_1 \rho\left(\frac{E_2}{E_1}\right) = d_1$. Therefore, we have $L_1 = \min\{L_i\}_{i=1}^{L}$, i.e., $\hat{E} = E_1$. In the same token, inequality $L_i > \tau_L$ for $i = 1, 2, \cdots, L-1$ is equivalent to $\frac{\|\mathbf{y}\|^2}{N} > E_M \tau\left(\frac{E_L}{E_i}\right)$ for $i = 1, 2, \cdots, M-1$. By Lemma 1 again, the above inequality is equivalent to $\frac{\|\mathbf{y}\|^2}{N} > E_{M-1}\tau\left(\frac{E_M}{E_{M-1}}\right) = d_{M-1}$. Therefore, we have $L_M = \min\{L_i\}_{i=1}^{M}$, i.e., $\hat{E} = E_L$. Following the same argument, we can arrive at the fact that for any given $1 < k < M$, inequalities $L_k < L_i$ for $i = 1, 2, \cdots, M$ and $i \neq k$ are equivalent to $\frac{\|\mathbf{y}\|^2}{N} > E_k \tau\left(\frac{E_k}{E_i}\right)$ for $i = 1, 2, \cdots, k-1$ and $\frac{\|\mathbf{y}\|^2}{N} < E_k \rho\left(\frac{E_i}{E_k}\right)$for $i = k+1, \cdots, M$. Again, using Lemma 1, we attain $d_{k-1} = E_k \tau\left(\frac{E_k}{E_{k-1}}\right) < \frac{\|\mathbf{y}\|^2}{N} < E_k \rho\left(\frac{E_{k+1}}{E_k}\right) = d_k$. This completes the proof of Algorithm 1. □

Once we have obtained the estimate of the signal point E in the sum constellation $\mathcal{E}$, we can uniquely decompose E so as to attain the estimate of each user signal point.

2.2 System Error Probability Analysis For Sum Constellation

In this section, we analyze the average SEP for the sum constellation with the noncoherent ML detector for the MIMO system. First, we notice that given the transmitted signals, the random variable $Y = \frac{\|\mathbf{y}\|^2}{a_i}$ follows a Chi-square distribution, with its probability density function (PDF) given by $f_Y(y) = \frac{1}{\Gamma(N)} y^{N-1} e^{-y},\ \ y > 0$, and its cumulative distribution function (CDF) is determined by $F(y) = 1 - e^{-y} \sum_{k=0}^{N-1} \frac{y^k}{k!},\ y > 0$. For notational simplicity, let $r_i = \frac{E_{i+1}}{E_i}$, $\tau_i = \tau(r_i)$ and $\rho_i = \rho(r_i)$. From the definition of d_i in Lemma 2, we know $d_i = E_i \rho_i$. Now, using the CDF of the random variable Y and Algorithm 1, we obtain the expression on the probability of correct detection on the i-th symbol E_i, denoted by $\mathrm{P}_{\mathrm{c},i}$ as $\mathrm{P}_{\mathrm{c},1} = F(N\rho_1)$, $\mathrm{P}_{\mathrm{c},i} = F(N\rho_i) - F(N\tau_{i-1})$ for $i = 2, 3, \cdots, M-1$, $\mathrm{P}_{\mathrm{c},M} = 1 - F(N\tau_{M-1})$. Therefore, the average symbol error probability is given by

$$\mathrm{P}_e = \frac{1}{M}\sum_{i=1}^{M-1}\Big(F(N\tau_i) + 1 - F(N\rho_i)\Big), \tag{2.3}$$

2.3 Diversity Analysis for Massive MIMO Systems

Our main task in this section is to quantitatively characterize both the diversity and coding gains for the considered uplink massive MIMO system, which describes how fast SEP decays when the number of the receiver antennas goes to infinity. To do that, we need to establish the following Lemma.

Lemma 3. *Let function $\xi(t)$ be defined by $\xi(t) = \frac{t}{e^{t-1}}$, Then, $\xi(t)$ monotonically increases for $t < 1$, and monotonically decreases for $t > 1$. and Moreover, $\xi\big(\tau(r)\big) = \xi\big(\rho(r)\big)$ holds for $r > 1$, where the functions $\tau(r)$ and $\rho(r)$ are defined in Lemma 1.*

□

Proof: Since $\xi'(t) = \frac{e(1-t)}{e^t}$, we have $\xi'(t) > 0$ for $t < 1$ and $\xi'(t) < 0$ for $t > 1$, i.e., $\xi(t)$ monotonically increases for $t < 1$ and monotonically decreases with increases for $t > 1$. In addition, it can be verified directly that $\xi\big(\rho(r)\big) = \xi\big(\tau(r)\big)$. This completes the proof of Lemma 3. □

Proposition 1. *The optimal SEP derived in Theorem 1 is lower and upper bounded by $\mathrm{P}_{e,\mathrm{L}} < \mathrm{P}_e < \mathrm{P}_{e,\mathrm{U}}$, where*

$$\begin{aligned}\mathrm{P}_{e,\mathrm{U}} &= \frac{4(M-1)\sqrt{N}}{M\sqrt{2\pi}}\, G^{-N}\\ \mathrm{P}_{e,\mathrm{L}} &= \frac{2\sqrt{N-\frac{1}{6}}}{M\sqrt{2\pi}N\rho}\, G^{-N}\end{aligned}$$

with G defined by $G = \frac{e^{\rho(r_{\min})-1}}{\rho(r_{\min})}$. □

Proof: First, we notice that $F(y)$ monotonically increases for $y \geq 0$, since $F'(y) = \frac{y^{N-1}e^{-y}}{(N-1)!} \geq 0$ for $y \geq 0$. Combining this fact with Lemma 1 and (2.3) results in

$$\begin{aligned} \frac{1}{M}\Big(F(N\bar{\tau}) + 1 - F(N\bar{\rho})\Big) &\leq \mathrm{P_e} \\ \leq \frac{(M-1)}{M}\Big(F(N\bar{\tau}) + 1 - F(N\bar{\rho})\Big), \end{aligned} \tag{2.4}$$

where, for notational simplicity, $\bar{\tau} = \tau(r_{\min})$ and $\bar{\rho} = \rho(r_{\min})$. Now, Let us prove the upper bound first. On the one hand, we know that $e^{N\bar{\tau}}$ can be expanded as

$$e^{N\bar{\tau}} = \sum_{k=0}^{3N-1} \frac{(N\bar{\tau})^k}{k!} + \frac{(N\bar{\tau})^{3N} e^{\zeta}}{(3N)!} \tag{2.5}$$

where $0 \leq \zeta \leq N\bar{\tau}$. Therefore, we can rewrite $G\big(N\bar{\tau}\big)$ as

$$\begin{aligned} F\big(N\bar{\tau}\big) &= 1 - \left(\sum_{k=0}^{3N-1} \frac{(N\bar{\tau})^k}{k!} - \sum_{k=N}^{3N-1} \frac{(N\bar{\tau})^k}{k!}\right) e^{-N\bar{\tau}} \\ &= \frac{(N\bar{\tau})^{3N} e^{-(N\bar{\tau}-\xi)}}{(3N)!} + e^{-N\bar{\tau}} \sum_{k=N}^{3N-1} \frac{(N\bar{\tau})^k}{k!} \end{aligned} \tag{2.6}$$

where we have used (2.5). On the other hand, by Lemma 1, since $0 < \bar{\tau} < 1$, sequence $\frac{(N\bar{\tau})^k}{k!}$ for $k = N, N+1, \cdots, 3N-1$ is monotonically decreasing. As a consequence, we have

$$\frac{(N\bar{\tau})^k}{k!} \leq \frac{(N\bar{\tau})^N}{N!} \qquad \text{for } k = N, N+1, \cdots, 3N-1. \tag{2.7}$$

Now, combining (2.7) with (2.7) yields

$$F\big(N\bar{\tau}\big) \leq \frac{(N\bar{\tau})^{3N}}{(3N)!} + e^{-N\bar{\tau}} \times \frac{2N(N\bar{\tau})^N}{N!} \tag{2.8}$$

Then, using the upper bound of the Stirling's inequality in [31]

$$\frac{N^N}{N!} \leq \frac{e^N}{\sqrt{2\pi N}}, \tag{2.9}$$

we can further upper bound (2.8) by

$$\begin{aligned} F\big(N\bar{\tau}\big) &< \frac{1}{\sqrt{6\pi N}} \times \Big(\frac{\bar{\tau}e}{3}\Big)^{3N} + \frac{\sqrt{2N}}{\sqrt{\pi}} \times \Big(\frac{\bar{\tau}}{e^{\bar{\tau}-1}}\Big)^{N} \\ &< \frac{3\sqrt{N}}{\sqrt{2\pi}}\Big(\frac{\bar{\tau}}{e^{\bar{\tau}-1}}\Big)^{N}, \end{aligned} \tag{2.10}$$

where we have used the fact that $\frac{\bar{\tau}e}{3} < \frac{\bar{\tau}}{e^{\bar{\tau}-1}}$ when $0 < \bar{\tau} < 1$. Regarding the second term of P_e, we notice that

$$1 - F\big(N\bar{\rho}\big) = e^{-N\bar{\rho}} \sum_{k=0}^{N-1} \frac{(N\bar{\rho})^k}{k!} \tag{2.11}$$

From Lemma 1 we know that $\bar{\rho} > 1$. In this case, since sequence $\frac{(N\bar{\rho})^k}{k!}$ for $k = 0, 1, \cdots, N-1$ is monotonically increasing, we have that

$$\frac{(N\bar{\rho})^k}{k!} \leq \frac{(N\bar{\rho})^{N-1}}{(N-1)!} \qquad \text{for } k = 0, 1, \cdots, N-1. \tag{2.12}$$

Now, combining (2.11) with (2.12), we can upper bound $1 - G\big(N\bar{\rho}\big)$ by

$$1 - F\big(N\bar{\rho}\big) \leq e^{-N\bar{\rho}} \frac{N(N\bar{\rho})^{N-1}}{(N-1)!} \tag{2.13}$$

Using the upper bound of the Stirling inequality (2.9) again, inequality (2.13) can be further upper bounded by

$$1 - F\big(N\bar{\rho}\big) \;\; \leq \tfrac{\sqrt{N}}{\bar{\rho}\sqrt{2\pi}}\big(\tfrac{\bar{\rho}}{e^{\bar{\rho}-1}}\big)^N < \tfrac{\sqrt{N}}{\sqrt{2\pi}}\big(\tfrac{\bar{\rho}}{e^{\bar{\rho}-1}}\big)^N, \tag{2.14}$$

By Lemma 2, we know

$$G^{-1} = \frac{\bar{\tau}}{e^{\bar{\tau}-1}} = \frac{\bar{\rho}}{e^{\bar{\rho}-1}}. \tag{2.15}$$

Combing (2.10), (2.14), (2.15), and (2.4) completes the proof of Proposition 1 on the upper bound.

Now, we consider to prove the lower bound of Proposition 1. Combining (2.6) with (2.4) and using the lower bound of Stirling's inequality in [31]

$$\frac{N^N}{N!} \geq \frac{e^N\sqrt{N-\frac{1}{6}}}{\sqrt{2\pi}N}, \tag{2.16}$$

and (2.15), we can arrive at the conclusion that

$$\begin{aligned}
\mathrm{P}_e &\geq \frac{1}{M} \times \Big(e^{-N\bar{\tau}}\frac{(N\bar{\tau})^N}{N!} + e^{-N\bar{\rho}}\frac{(N\bar{\rho})^{N-1}}{(N-1)!}\Big) \\
&\geq \frac{1}{M} \times \left(\frac{\sqrt{N-\frac{1}{6}}}{\sqrt{2\pi}N\bar{\rho}}\Big(\frac{\bar{\tau}}{e^{\bar{\tau}-1}}\Big)^N + \frac{\sqrt{N-\frac{1}{6}}}{\sqrt{2\pi}N\bar{\rho}}\Big(\frac{\bar{\rho}}{e^{\bar{\rho}-1}}\Big)^N\right) \\
&= \frac{2}{M} \times \frac{\sqrt{N-\frac{1}{6}}}{\sqrt{2\pi}N\bar{\rho}}\Big(\frac{\bar{\rho}}{e^{\bar{\rho}-1}}\Big)^N.
\end{aligned}$$

Therefore, we complete the proof of the lower bound and thus, Propositon 1. $\square$

This proposition naturally leads us successfully to reveal a significant asymptotic behaviour on the uplink massive MIMO system as N tends to infinity. To see it more clearly, let us introduce the following definition.

Definition 2. *A transmission scheme is said to enable full-receiver diversity for a detector D if, for any given positive numbers* ϵ_L *and* ϵ_U*, there exist positive constants* C_1, C_2 *and* $\mathcal{G}_{\mathrm{D}} > 1$ *independent of* N *such that the probability of detection error,* P_{D} *satisfies the following condition:*

$$C_1 \mathcal{G}_{\mathrm{D}}^{-N(1+\epsilon_L)} \leq P_{\mathrm{D}} \leq C_2 \mathcal{G}_{\mathrm{D}}^{-N(1-\epsilon_U)}.$$

Here, the constant $\mathcal{G}_{\mathrm{D}}$ *is called geometrical coding gain.* □

Chapter 3

Optimal MU Constellation Designs

From Definition 1 and Proposition 1, we know that any nonnegative PAM signalling enables full receiver diversity N for the noncoherent ML detector. In addition, we can obtain $\lim_{N\to\infty} \left(\mathrm{P}_e\right)^{1/N} = G^{-1}$ and thus, the corresponding geometrical coding gain is $\mathcal{G}_{\mathrm{ML}} = G^{-1} = \frac{e^{\tau(r_{\min})-1}}{\tau(r_{\min})} = \frac{e^{\rho(r_{\min})-1}}{\rho(r_{\min})}$, which, essentially, characterizes how rapidly SEP decays when the number of the received antennas goes to infinity. Therefore, we should maximize it. By Lemmas 1 and 3, we notice that $G = \xi(\tau(r_{\min}))$ and thus, maximizing G is equivalent to maximizing $r_{\min}$. Therefore, the design of the optimal sum constellation is formally formulated into the following optimization problem:

Problem 1. *Find a uniquely decomposable sum constellation $\mathcal{E}$ such that the geometrical coding gain is maximized subject to each user power constraint, i.e.,*

$$\max \min\{r_m\}_{m=1}^{M-1}$$

subject to

$$\mathcal{E} = \beta_1\mathcal{U}_1 \oplus \beta_2\mathcal{U}_2 \oplus \cdots \oplus \beta_K\mathcal{U}_K + \sigma^2$$
$$\frac{1}{M_k}\sum_{m=1}^{M_k} u_{k,m} \leq P_k \qquad \textit{for } k = 1, 2, \cdots, K$$

□

In general, it is very hard to attain a closed-form solution to Problem 1. However, particularly for two and three users, with each transmitting a nonnegative binary constellation, we can obtain the optimal solution explicitly. To do that, we need to establish the following lemma.

Lemma 4. *Let $\mathcal{E}^*$ and $\mathcal{U}_k^*$ denote the optimal sum constellation and each user constellation to Problem 1, respectively. Then, we have $u_{k,1}^* = 0$ for $k = 1, 2, \cdots, K$.* □

Proof: First, we notice the following fact: For any given three positive numbers s, s' and s_0 with $s > s' > s_0 \geq 0$, we have

$$\frac{s - s_0}{s' - s_0} \geq \frac{s}{s'}$$

This is because

$$\frac{s - s_0}{s' - s_0} - \frac{s}{s'} = \frac{s_0(s - s')}{(s' - s_0)s'} \geq 0$$

Using this fact, if we let $\bar{\mathcal{U}}_k = \mathcal{U}_k - u_{k,1}$ and $\bar{\mathcal{E}} = \mathcal{E} - \sum_{k=1}^{K} \beta_k u_{k,1}$. then, we have

$\bar{\mathcal{E}} = \beta_1\bar{\mathcal{U}}_1 \oplus \beta_2\bar{\mathcal{U}}_2 \oplus \cdots \oplus \beta_K\bar{\mathcal{U}}_K + \sigma^2$ and

$$\bar{r}_i = \frac{\bar{E}_{i+1}}{\bar{E}_i} = \frac{E_{i+1} - E_0}{E_i - E_0} \geq \frac{E_{i+1}}{E_i} = r_i$$

for $i = 1, 2, \cdots, M-1$, where $E_0 = \sum_{k=1}^{K} \beta_k u_{k,1}$. This completes the proof of Lemma 4. □

Two user case

Here, we consider a two users case, with each user having nonnegative binary modulation, i.e., $K = 2$ and $M_1 = M_2 = 2$. In this case, we can obtain the optimal solution explicitly, which is summarized as the following Theorem 1:

Theorem 1. *The optimal solution to Problem 1 for two users with each transmitting binary constellation is explicitly determined as follows:*

- *If the powers and the large scale fading coefficents of the two users satisfy the following condition:*

$$2p_1^2\beta_1^2 - 2p_2^2\beta_2^2 + 2p_1\beta_1\sigma^2 - p_2\beta_2\sigma^2 + 2p_1\beta_1p_2\beta_2 \leq 0$$

$$2p_2^2\beta_2^2 - 2p_1^2\beta_1^2 + 2p_2\beta_2\sigma^2 - p_1\beta_1\sigma^2 + 2p_1\beta_1p_2\beta_2 \geq 0,$$

then, we have

$$\begin{cases} u_{1,1} = u_{2,1} = 0 \\ u_{1,2} = 2p_1 \\ u_{2,2} = \dfrac{(2p_1\beta_1 - \sigma^2) + \sqrt{(2p_1\beta_1 - \sigma^2)^2 + 16(p_1^2\beta_1^2 + p_1\beta_1\sigma^2)}}{2\beta_2} \end{cases}$$

- *If the powers and the large scale fading coefficents of the two users satisfy the following condition:*

$$2p_1^2\beta_1^2 - 2p_2^2\beta_2^2 + 2p_1\beta_1\sigma^2 - p_2\beta_2\sigma^2 + 2p_1\beta_1 p_2\beta_2 \geq 0$$

$$2p_2^2\beta_2^2 - 2p_1^2\beta_1^2 + 2p_2\beta_2\sigma^2 - p_1\beta_1\sigma^2 + 2p_1\beta_1 p_2\beta_2 \leq 0,$$

then, we obtain

$$\begin{cases} u_{1,1} = u_{2,1} = 0 \\ u_{1,2} = \dfrac{(2p_2\beta_2 - \sigma^2) + \sqrt{(2p_2\beta_2 - \sigma^2)^2 + 16(p_2^2\beta_2^2 + p_2\beta_2\sigma^2)}}{2\beta_1} \\ u_{2,2} = 2p_2 \end{cases}$$

- *If the powers and the large scale fading coefficients of the two users satisfy the following condition:*

$$2p_1^2\beta_1^2 - 2p_2^2\beta_2^2 + 2p_1\beta_1\sigma^2 - p_2\beta_2\sigma^2 + 2p_1\beta_1 p_2\beta_2 \geq 0$$

$$p_1\beta_1 \leq p_2\beta_2,$$

then, we attain

$$\begin{cases} u_{1,2} = u_{2,1} = 0 \\ u_{1,2} = \dfrac{-2p_2\beta_2 + \sqrt{4p_2^2\beta_2^2 + 4(2p_2\beta_2 + \sigma^2)^2}}{2\beta_1} - \dfrac{\sigma^2}{\beta_1} \\ u_{2,2} = 2p_2 \end{cases}$$

- *If the powers and the large scale fading coefficients of the two users satisfy the following condition:*

$$2p_2^2\beta_2^2 - 2p_1^2\beta_1^2 + 2p_2\beta_2\sigma^2 - p_1\beta_1\sigma^2 + 2p_1\beta_1 p_2\beta_2 \geq 0$$

$$p_1\beta_1 \geq p_2\beta_2,$$

then, we have

$$\begin{cases} u_{1,2} = u_{2,1} = 0 \\ u_{1,2} = 2p_1 \\ u_{2,2} = \dfrac{-2p_1\beta_1 + \sqrt{4p_1^2\beta_1^2 + 4(2p_1\beta_1 + \sigma^2)^2}}{2\beta_2} - \dfrac{\sigma^2}{\beta_2} \end{cases}$$

□

Proof: For notational simplicity, let $\bar{p}_1 = 2\beta_1 p_1, \bar{p}_2 = 2\beta_2 p_2$, $\mathcal{A}_1 = \beta_1\mathcal{U}_1 = \{0, a_1\}$ and $\mathcal{A}_2 = \beta_2\mathcal{U}_2 = \{0, a_2\}$ without loss of generality. Consider the following two situations.

(i) Situation 1: $a_1 < a_2$. In this situation, the sum constellation can be determined by

$$\mathcal{E} = \{\sigma^2, a_1 + \sigma^2, a_2 + \sigma^2, a_1 + a_2 + \sigma^2\}$$

Suppose

$$r = \min\{r_i\} = \min\left\{\frac{a_2 + \sigma^2}{a_1 + \sigma^2}, \frac{a_1 + a_2 + \sigma^2}{a_2 + \sigma^2}\right\}$$

Then we have

$$\begin{cases} r \leq \dfrac{a_2 + \sigma^2}{a_1 + \sigma^2} \\ r \leq \dfrac{a_1 + a_2 + \sigma^2}{a_2 + \sigma^2}, \end{cases} \tag{3.1}$$

which is equivalent to

$$\begin{cases} ra_1 - a_2 + (r-1)\sigma^2 \leq 0 & (3.2) \\ -a_1 + (r-1)a_2 + (r-1)\sigma^2 \leq 0 & (3.3) \end{cases}$$

Since $r > 1$, manipulating $(r-1) \times (3.2) + (3.3)$ and $r \times (3.3) + (3.2)$ produces

$$\begin{cases} (r^2 - r - 1)a_1 + (r^2 - r)\sigma^2 \leq 0 & (3.4) \\ (r^2 - r - 1)a_2 + (r^2 - 1)\sigma^2 \leq 0 & (3.5) \end{cases}$$

This implies $r^2 - r - 1 < 0$, i.e., $1 < r < \frac{\sqrt{5}+1}{2}$. Combining (3.4) and (3.5) with the power constraints: $a_1 \leq \bar{p}_1$ and $a_2 \leq \bar{p}_2$, we have

$$\begin{cases} \dfrac{-(r^2 - r)\sigma^2}{r^2 - r - 1} \leq a_1 \leq \bar{p}_1 \\ \dfrac{-(r^2 - 1)\sigma^2}{r^2 - r - 1} \leq a_2 \leq \bar{p}_2, \end{cases} \tag{3.6}$$

which is equivalent to

$$\begin{cases} (\bar{p}_1 + \sigma^2)r^2 - (\bar{p}_1 + \sigma^2)r - \bar{p}_1 \leq 0 \\ (\bar{p}_2 + \sigma^2)r^2 - \bar{p}_2 r - (\bar{p}_2 + \sigma^2) \leq 0 \end{cases}$$

Hence, we obtain

$$\begin{cases} r \leq \dfrac{(\bar{p}_1+\sigma^2)+\sqrt{(\bar{p}_1+\sigma^2)^2+4(\bar{p}_1+\sigma^2)\bar{p}_1}}{2(\bar{p}_1+\sigma^2)} \\ r \leq \dfrac{\bar{p}_2+\sqrt{\bar{p}_2^2+4(\bar{p}_2+\sigma^2)^2}}{2(\bar{p}_2+\sigma^2)} \end{cases} \tag{3.7}$$

and further, we have

$$r \leq \min\left\{\frac{(\bar{p}_1+\sigma^2)+\sqrt{(\bar{p}_1+\sigma^2)^2+4(\bar{p}_1+\sigma^2)\bar{p}_1}}{2(\bar{p}_1+\sigma^2)}, \frac{\bar{p}_2+\sqrt{\bar{p}_2^2+4(\bar{p}_2+\sigma^2)^2}}{2(\bar{p}_2+\sigma^2)}\right\}$$

Therefore, when

$$\frac{(\bar{p}_1+\sigma^2)+\sqrt{(\bar{p}_1+\sigma^2)^2+4(\bar{p}_1+\sigma^2)\bar{p}_1}}{2(\bar{p}_1+\sigma^2)} \leq \frac{\bar{p}_2+\sqrt{\bar{p}_2^2+4(\bar{p}_2+\sigma^2)^2}}{2(\bar{p}_2+\sigma^2)}$$

we attain that the maximum r, denoted by $r_{\max,1}$, is determined by

$$r_{\max,1} = \frac{(\bar{p}_1+\sigma^2)+\sqrt{(\bar{p}_1+\sigma^2)^2+4(\bar{p}_1+\sigma^2)\bar{p}_1}}{2(\bar{p}_1+\sigma^2)} \tag{3.8}$$

and as a result, we have

$$\begin{cases} a_1 = \bar{p}_1 \\ a_2 = \dfrac{(\bar{p}_1-\sigma^2)+\sqrt{(\bar{p}_1-\sigma^2)^2+4(\bar{p}_1^2+2\bar{p}_1\sigma^2)}}{2} \end{cases}$$

However, when

$$\frac{(\bar{p}_1+\sigma^2)+\sqrt{(\bar{p}_1+\sigma^2)^2+4(\bar{p}_1+\sigma^2)\bar{p}_1}}{2(\bar{p}_1+\sigma^2)} > \frac{\bar{p}_2+\sqrt{\bar{p}_2^2+4(\bar{p}_2+\sigma^2)^2}}{2(\bar{p}_2+\sigma^2)}$$

the maximum r, denoted by $r_{\max,2}$ is determined by

$$r_{\max,2} = \frac{\bar{p}_2 + \sqrt{\bar{p}_2^2 + 4(\bar{p}_2 + \sigma^2)^2}}{2(\bar{p}_2 + \sigma^2)} \tag{3.9}$$

and the result a_1 and a_2 are given by

$$\begin{cases} a_1 = \dfrac{-(r^2 - r)\sigma^2}{r^2 - r - 1} = \dfrac{-\bar{p}_2 + \sqrt{\bar{p}_2^2 + 4(\bar{p}_2 + \sigma^2)^2}}{2} - \sigma^2 \\ a_2 = \dfrac{-(r^2 - 1)\sigma^2}{r^2 - r - 1} = \bar{p}_2 \end{cases}$$

(ii) Situation 2: $a_2 < a_1$. In this situation, the sum constellation is represented by

$$\mathcal{E} = \left\{\sigma^2, a_2 + \sigma^2, a_1 + \sigma^2, a_1 + a_2 + \sigma^2\right\}$$

Similarly, if we let

$$r = \min\{r_i\} = \min\{\frac{a_1 + \sigma^2}{a_2 + \sigma^2}, \frac{a_1 + a_2 + \sigma^2}{a_1 + \sigma^2}\}$$

then, we have

$$\begin{cases} r \leq \dfrac{a_1 + \sigma^2}{a_2 + \sigma^2} & (3.10) \\ r \leq \dfrac{a_1 + a_2 + \sigma^2}{a_1 + \sigma^2} & (3.11) \end{cases}$$

then, we have

$$\begin{cases} -a_1 + ra_2 + (r-1)\sigma^2 \leq 0 & (3.12) \\ (r-1)a_1 - a_2 + (r-1)\sigma^2 \leq 0 & (3.13) \end{cases}$$

Since $r > 1$, performing operations $r \times (3.41) + (3.12)$ and $(r-1) \times (3.12) + (3.41)$ results in

$$\begin{cases} (r^2 - r - 1)a_1 + (r^2 - 1)\sigma^2 \leq 0 & (3.14) \\ (r^2 - r - 1)a_2 + (r^2 - r)\sigma^2 \leq 0 & (3.15) \end{cases}$$

Hence, we must have $r^2 - r - 1 < 0$, i.e., $1 < r < \frac{\sqrt{5}+1}{2}$. Combining (3.42) and (3.15) with each user's power constraint yields

$$\begin{cases} \dfrac{-(r^2-1)\sigma^2}{r^2 - r - 1} \leq a_1 \leq \bar{p}_1 \\ \dfrac{-(r^2-r)\sigma^2}{r^2 - r - 1} \leq a_2 \leq \bar{p}_2 \end{cases}$$

which is equivalent to

$$\begin{cases} (\bar{p}_1 + \sigma^2)r^2 - \bar{p}_1 r - (\bar{p}_1 + \sigma^2) \leq 0 \\ (\bar{p}_2 + \sigma^2)r^2 - (\bar{p}_2 + \sigma^2)r - \bar{p}_2 \leq 0 \end{cases}$$

Hence, we have

$$\begin{cases} r \leq \dfrac{\bar{p}_1 + \sqrt{\bar{p}_1^2 + 4(\bar{p}_1 + \sigma^2)^2}}{2(\bar{p}_1 + \sigma^2)} & (3.16) \\ r \leq \dfrac{(\bar{p}_2 + \sigma^2) + \sqrt{(\bar{p}_2 + \sigma^2)^2 + 4(\bar{p}_2 + \sigma^2)\bar{p}_2}}{2(\bar{p}_2 + \sigma^2)} & (3.17) \end{cases}$$

and thus,

$$r \leq \min\left\{ \frac{\bar{p}_1 + \sqrt{\bar{p}_1^2 + 4(\bar{p}_1 + \sigma^2)^2}}{2(\bar{p}_1 + \sigma^2)}, \frac{(\bar{p}_2 + \sigma^2) + \sqrt{(\bar{p}_2 + \sigma^2)^2 + 4(\bar{p}_2 + \sigma^2)\bar{p}_2}}{2(\bar{p}_2 + \sigma^2)} \right\}$$

Therefore, when

$$\frac{\bar{p}_1 + \sqrt{\bar{p}_1^2 + 4(\bar{p}_1 + \sigma^2)^2}}{2(\bar{p}_1 + \sigma^2)} \leq \frac{(\bar{p}_2 + \sigma^2) + \sqrt{(\bar{p}_2 + \sigma^2)^2 + 4(\bar{p}_2 + \sigma^2)\bar{p}_2}}{2(\bar{p}_2 + \sigma^2)}$$

we have

$$r_{\max,3} = \frac{\bar{p}_1 + \sqrt{\bar{p}_1^2 + 4(\bar{p}_1 + \sigma^2)^2}}{2(\bar{p}_1 + \sigma^2)} \tag{3.18}$$

and

$$\begin{cases} a_1 = \bar{p}_1 \\ a_2 = \dfrac{-\bar{p}_1 + \sqrt{\bar{p}_1^2 + 4(\bar{p}_1 + \sigma^2)^2}}{2} - \sigma^2 \end{cases}$$

However, when

$$\frac{\bar{p}_1 + \sqrt{\bar{p}_1^2 + 4(\bar{p}_1 + \sigma^2)^2}}{2(\bar{p}_1 + \sigma^2)} > \frac{(\bar{p}_2 + \sigma^2) + \sqrt{(\bar{p}_2 + \sigma^2)^2 + 4(\bar{p}_2 + \sigma^2)\bar{p}_2}}{2(\bar{p}_2 + \sigma^2)}$$

we attain

$$r_{\max,4} = \frac{(\bar{p}_2 + \sigma^2) + \sqrt{(\bar{p}_2 + \sigma^2)^2 + 4(\bar{p}_2 + \sigma^2)\bar{p}_2}}{2(\bar{p}_2 + \sigma^2)} \tag{3.19}$$

and

$$\begin{cases} a_1 = \dfrac{(\bar{p}_2 - \sigma^2) + \sqrt{(\bar{p}_2 - \sigma^2)^2 + 4(\bar{p}_2^2 + 2\bar{p}_2\sigma^2)}}{2} \\ a_2 = \bar{p}_2 \end{cases}$$

Finally, summarizing the above discussions completes the proof of Theorem 1. $\square$

Three user case

In this case, each user has nonnegative binary modulation, i.e., $K = 3$ and $M_1 = M_2 = M_3 = 2$. By Lemma 4, we can assume that $\mathcal{B}_k = \beta_k \mathcal{U}_k = \{0, b_k\}$ for $k = 1, 2, 3$. This case is more complicated than the two users case, since we have to consider all 6 user permutations individually. Therefore, in order to obtain the optimal solution to Problem 1 and for presentation clarity and concision, let us define π_1, π_2 and π_3 as such permutation π of three numbers $1, 2$ and 3 that $b_{\pi_1} < b_{\pi_2} < b_{\pi_3}$, and correspondingly, let the resulting permutated constellation be denoted by $\mathcal{A}_i = \mathcal{B}_{\pi_i}$, with $a_i = b_{\pi_i}$ for $i = 1, 2, 3$. Also, let $\bar{p}_i = 2\beta_{\pi_i} p_{\pi_i}$. Under this users permutation π, i.e., $a_1 < a_2 < a_3$, let us now consider the following two situations.

(i) Situation 1: $a_1 + a_2 < a_3$. In this situation, the sum constellation can be represented by

$$\mathcal{E} = \{0, a_1, a_2, a_1 + a_2, a_3, a_1 + a_3, a_2 + a_3, a_1 + a_2 + a_3\} + \sigma^2$$

Then, we have

$$r = \min\{r_m\}_{m=1}^{7} = \min\left\{\frac{a_1 + \sigma^2}{\sigma^2}, \frac{a_2 + \sigma^2}{a_1 + \sigma^2}, \frac{a_1 + a_2 + \sigma^2}{a_2 + \sigma^2}, \frac{a_3 + \sigma^2}{a_1 + a_2 + \sigma^2}, \right.$$
$$\left. \frac{a_1 + a_3 + \sigma^2}{a_3 + \sigma^2}, \frac{a_2 + a_3 + \sigma^2}{a_1 + a_3 + \sigma^2}, \frac{a_1 + a_2 + a_3 + \sigma^2}{a_2 + a_3 + \sigma^2}\right\}$$

Since

$$\frac{a_1 + \sigma^2}{\sigma^2} > \frac{a_1 + a_3 + \sigma^2}{a_3 + \sigma^2}, \frac{a_2 + \sigma^2}{a_1 + \sigma^2} > \frac{a_2 + a_3 + \sigma^2}{a_1 + a_3 + \sigma^2},$$
$$\frac{a_1 + a_2 + \sigma^2}{a_2 + \sigma^2} > \frac{a_1 + a_2 + a_3 + \sigma^2}{a_2 + a_3 + \sigma^2}, \frac{a_1 + a_3 + \sigma^2}{a_3 + \sigma^2} > \frac{a_1 + a_2 + a_3 + \sigma^2}{a_2 + a_3 + \sigma^2}$$

we have

$$r = \min\{r_m\}_{m=1}^{7} = \min\left\{\frac{a_3+\sigma^2}{a_1+a_2+\sigma^2}, \frac{a_2+a_3+\sigma^2}{a_1+a_3+\sigma^2}, \frac{a_1+a_2+a_3+\sigma^2}{a_2+a_3+\sigma^2}\right\}$$

Hence, we obtain

$$\begin{cases} r \le \dfrac{a_3+\sigma^2}{a_1+a_2+\sigma^2} \\ r \le \dfrac{a_2+a_3+\sigma^2}{a_1+a_3+\sigma^2} \\ r \le \dfrac{a_1+a_2+a_3+\sigma^2}{a_2+a_3+\sigma^2} \end{cases}$$

which is equivalent to

$$\begin{cases} ra_1 + ra_2 - a_3 - (1-r)\sigma^2 \le 0 & (3.20) \\ ra_1 - a_2 + (r-1)a_3 - (1-r)\sigma^2 \le 0 & (3.21) \\ -a_1 + (r-1)a_2 + (r-1)a_3 - (1-r)\sigma^2 \le 0 & (3.22) \end{cases}$$

Since $r > 1$, manipulating $(r-1) \times$ (3.20) + (3.21) and $(r-1) \times$ (3.20) + (3.22) produces

$$\begin{cases} r^2a_1 + (r^2-r-1)a_2 \le (r-r^2)\sigma^2 & (3.23) \\ (r^2-r-1)a_1 + (r^2-1)a_2 \le (r-r^2)\sigma^2 & (3.24) \end{cases}$$

Hence, we must have $r^2 - r - 1 < 0$, i.e., $1 < r < \frac{\sqrt{5}+1}{2}$. Taking operatons on $(r^2-1) \times$ (3.23) + $(1+r-r^2) \times$ (3.24) yiels

$$(2r^3 - 2r - 1)a_1 \le (r^2 - r^3)\sigma^2 \tag{3.25}$$

Thus, we must have $2r^3 - 2r - 1 < 0$ for $r > 1$, which is equivalent to $1 < r < \frac{\sqrt[3]{54+6\sqrt{33}}+\sqrt[3]{54-6\sqrt{33}}}{6} = \bar{r}_1$, since $\bar{r}_1$ is the root of function $2r^3 - 2r - 1$ and it monotonically increases for $r > 1$. Then, performing operations on $(1 + r - r^2) \times (3.23) + r^2 \times (3.24)$ results in (3.26)

$$(2r^3 - 2r - 1)a_2 \leq (r - r^3)\sigma^2 \tag{3.26}$$

Similarly, by performing $(3.20) + r \times (3.21)$ and $(3.21) + (3.22)$, we have

$$\begin{cases} (r + r^2)a_1 + (r^2 - r - 1)a_3 \leq (1 - r^2)\sigma^2 & (3.27) \\ (r^2 - r - 1)a_1 + (r^2 - r)a_3 \leq (r - r^2)\sigma^2 & (3.28) \end{cases}$$

Then, manipulating $(1 + r - r^2) \times (3.27) + (r + r^2) \times (3.28)$ produces

$$(2r^3 - 2r - 1)a_3 \leq (-r^3 - r^2 + r + 1)\sigma^2 \tag{3.29}$$

Combining (3.25), (3.26) and (3.29) with the power constraints $a_1 \leq \bar{p}_1, a_2 \leq \bar{p}_2, a_3 \leq \bar{p}_3$ leads to

$$\begin{cases} \bar{p}_1 \geq a_1 \geq \dfrac{r^3\sigma^2 - r^2\sigma^2}{1 + 2r - 2r^3} \\ \bar{p}_2 \geq a_2 \geq \dfrac{-r\sigma^2 + r^3\sigma^2}{1 + 2r - 2r^3} \\ \bar{p}_3 \geq a_3 \geq \dfrac{(r^3 + r^2 - r - 1)\sigma^2}{1 + 2r - 2r^3} \end{cases}$$

which is equivalent to

$$\begin{cases} (\sigma^2 + 2\bar{p}_1)r^3 - \sigma^2 r^2 - 2\bar{p}_1 r - \bar{p}_1 \le 0 \\ (\sigma^2 + 2\bar{p}_2)r^3 - (\sigma^2 + 2\bar{p}_2)r - \bar{p}_2 \le 0 \\ (\sigma^2 + 2\bar{p}_3)r^3 + \sigma^2 r^2 - (\sigma^2 + 2\bar{p}_3)r - \sigma^2 - \bar{p}_3 \le 0 \end{cases} \tag{3.30}$$

In order to obtain the solution to (3.30), we let $f_1(r) = (\sigma^2+2\bar{p}_1)r^3 - \sigma^2 r^2 - 2\bar{p}_1 r - \bar{p}_1$, $f_2(r) = (\sigma^2+2\bar{p}_2)r^3 - (\sigma^2+2\bar{p}_2)r - \bar{p}_2$ and $f_3(r) = (\sigma^2+2\bar{p}_3)r^3 + \sigma^2 r^2 - (\sigma^2+2\bar{p}_3)r - \sigma^2 - \bar{p}_3$. Since the first order derivative of $f_1(r)$, $f_1'(r) = 3(\sigma^2+2\bar{p}_1)r^2 - 2\sigma^2 r - 2\bar{p}_1 = r\sigma^2(3r-2) + 2\bar{p}_1(3r^2-1) > 0$ for $1 < r < \bar{r}$, function $f_1(r)$ monotonically increases with r increasing with $1 < r < \bar{r}$. Besides, since $f_1(1) < 0$, $f_1(\bar{r}_1) = \bar{r}_1^2(\bar{r}_1 - 1)\sigma^2 > 0$, function $f_1(r)$ has one and only one root R_1 in inside $1 < r < \bar{r}_1$. Similarly, we can prove that each of functions $f_2(r)$ and $f_3(r)$ has one and only one root in inside $1 < r_1 < \bar{r}_1$, whose roots are denoted by R_2 and R_3, respectively. Therefore, the solution to (3.30) is determined by $r \le \min\{R_1, R_2, R_3\}$. As a consequence, the maximum r in this situation, denoted by $R_{\pi,1}$, is given by

$$R_{\pi,1} = \min\{R_1, R_2, R_3\}$$

(ii) Situation 2: $a_1 + a_2 > a_3$ In this situation, by Lemma 4, the sum constellation can be determined by

$$\mathcal{E} = \{0, a_1, a_2, a_3, a_1+a_2, a_1+a_3, a_2+a_3, a_1+a_2+a_3\} \tag{3.31}$$

Correspondingly, we have

$$r = \min\left\{\frac{a_1+\sigma^2}{\sigma^2}, \frac{a_2+\sigma^2}{a_1+\sigma^2}, \frac{a_3+\sigma^2}{a_2+\sigma^2}, \frac{a_1+a_2+\sigma^2}{a_3+\sigma^2},\right.$$
$$\left.\frac{a_1+a_3+\sigma^2}{a_1+a_2+\sigma^2}, \frac{a_2+a_3+\sigma^2}{a_1+a_3+\sigma^2}, \frac{a_1+a_2+a_3+\sigma^2}{a_2+a_3+\sigma^2}\right\}$$

Since

$$\frac{a_1+\sigma^2}{\sigma^2} > \frac{a_1+a_2+a_3+\sigma^2}{a_2+a_3+\sigma^2}, \quad \frac{a_2+\sigma^2}{a_1+\sigma^2} > \frac{a_2+a_3+\sigma^2}{a_1+a_3+\sigma^2}, \quad \frac{a_3+\sigma^2}{a_2+\sigma^2} > \frac{a_1+a_3+\sigma^2}{a_1+a_2+\sigma^2}$$

we attain

$$r = \min\left\{\frac{a_1+a_2+\sigma^2}{a_3+\sigma^2}, \frac{a_1+a_3+\sigma^2}{a_1+a_2+\sigma^2}, \frac{a_2+a_3+\sigma^2}{a_1+a_3+\sigma^2}, \frac{a_1+a_2+a_3+\sigma^2}{a_2+a_3+\sigma^2}\right\}$$

Therefore, we have

$$\begin{cases} r \le \dfrac{a_1+a_2+\sigma^2}{a_3+\sigma^2} \\ r \le \dfrac{a_1+a_3+\sigma^2}{a_1+a_2+\sigma^2} \\ r \le \dfrac{a_2+a_3+\sigma^2}{a_1+a_3+\sigma^2} \\ r \le \dfrac{a_1+a_2+a_3+\sigma^2}{a_2+a_3+\sigma^2} \end{cases} \tag{3.32}$$

which is equivalent to

$$\begin{cases} a_1 + a_2 - ra_3 + (1-r)\sigma^2 \geq 0 & (3.33) \\ (1-r)a_1 - ra_2 + a_3 + (1-r)\sigma^2 \geq 0 & (3.34) \\ -ra_1 + a_2 + (1-r)a_3 + (1-r)\sigma^2 \geq 0 & (3.35) \\ a_1 + (1-r)a_2 + (1-r)a_3 + (1-r)\sigma^2 \geq 0 & (3.36) \end{cases}$$

Notice that if both inequalities (3.33) and (3.34) holds, then, (3.36) also holds. Therefore, we only consider $(3.33), (3.34), (3.35)$. Since $r > 1$, performing $(r-1) \times (3.33) + (3.34)$ and $r \times (3.33) + (3.35)$ results in

$$\begin{cases} -a_2 + (-r^2 + r + 1)a_3 + r(1-r)\sigma^2 \geq 0 & (3.37) \\ (r+1)a_2 + (-r^2 - r + 1)a_3 + (1-r)(r+1)\sigma^2 \geq 0 & (3.38) \end{cases}$$

Inequality (3.37) implies $-r^2 + r + 1 \geq 0$. Then, performing $(r+1) \times (3.37) + (3.38)$ produces

$$(-r^3 - r^2 + r + 2)a_3 + (1-r)(r+1)^2\sigma^2 \geq 0 \qquad (3.39)$$

This implies $r^3 + r^2 - r - 2 < 0$, which is equivalent to $1 < r < \frac{\sqrt[3]{172 - 4\sqrt[3]{177} - \sqrt[3]{172 + 12\sqrt{177}}} - 2}{6} = \bar{r}_2$, since $\bar{r}_2$ is the root of function $r^3 + r^2 - r - 2$ and it monotonically increases for $r > 1$. Then, manipulating $(r^2 + r - 1) \times (3.37) + (-r^2 + r + 1) \times (3.38)$ leads to

$$(-r^3 - r^2 + r + 2)a_2 + \sigma^2(1 - r^3) \geq 0 \qquad (3.40)$$

Operate (3.33) $\times r +$ (3.34) and (3.35) $\times r +$ (3.34) results in

$$\begin{cases} a_1 + (1-r^2)a_3 + (1-r^2)\sigma^2 \geq 0 & (3.41) \\ (-r^2 - r + 1)a_1 + (-r^2 + r + 1)a_3 + (1-r^2)\sigma^2 \geq 0 & (3.42) \end{cases}$$

Then, manipulating (3.41) $\times (-r^2 + r + 1) +$ (3.42) $\times (r^2 - 1)$ produces

$$(-r^3 - r^2 + r + 2)a_1 + (1-r^2)\sigma^2 \geq 0 \tag{3.43}$$

Combing (3.39), (3.40), (3.43) with the power constraints, $a_1 \leq \bar{p}_1, a_2 \leq \bar{p}_2, a_3 \leq \bar{p}_3$, we have:

$$\begin{cases} \bar{p}_1 \geq a_1 \geq \dfrac{-(1-r^2)\sigma^2}{-r^3 - r^2 + r + 2} \\ \bar{p}_2 \geq a_2 \geq \dfrac{-(1-r^3)\sigma^2}{-r^3 - r^2 + r + 2} \\ \bar{p}_3 \geq a_3 \geq \dfrac{(r^3 + r^2 - r - 1)\sigma^2}{-r^3 - r^2 + r + 2} \end{cases} \tag{3.44}$$

which is equivalent to

$$\begin{cases} \bar{p}_1 r^3 + (\bar{p}_1 + \sigma^2)r^2 - \bar{p}_1 r - (2\bar{p}_1 + \sigma^2) \leq 0 \\ (\bar{p}_2 + \sigma^2)r^3 + \bar{p}_2 r^2 - \bar{p}_2 r - (2\bar{p}_2 + \sigma^2) \leq 0 \\ (\bar{p}_3 + \sigma^2)r^3 + (\bar{p}_3 + \sigma^2)r^2 - (\bar{p}_3 + \sigma^2)r - (2\bar{p}_3 + \sigma^2) \leq 0 \end{cases} \tag{3.45}$$

In this case, the range of r can be denoted by

$$\begin{cases} -r^2 + r + 1 \geq 0 \\ -r^3 - r^2 + r + 2 > 0 \end{cases}$$

Let $f_4(r) = \bar{p}_1 r^3 + (\bar{p}_1 + \sigma^2)r^2 - \bar{p}_1 r - (2\bar{p}_1 + \sigma^2)$, $f_5(r) = (\bar{p}_2 + \sigma^2)r^3 + \bar{p}_2 r^2 - \bar{p}_2 r - (2\bar{p}_2 - \sigma^2)$, $f_6(r) = (\bar{p}_3 + \sigma^2)r^3 + (\bar{p}_3 + \sigma^2)r^2 - (\bar{p}_3 + \sigma^2)r - (2\bar{p}_3 + \sigma^2)$. Then, we have $f_4'(r) = 3\bar{p}_1 r^2 + (2r-1)\bar{p}_1 + \sigma^2 r > 0$, $f_5'(r) = (3\bar{p}_2 + \sigma^2)(r^2 - 1) + 2\sigma^2 r^2 + 2\bar{p}_2 r > 0$ and $f_6'(r) = (\bar{p}_3 + \sigma^2)(3r^2 + 2r - 1) > 0$ for $1 < r < \bar{r}$, all these three functions $f_4(r), f_5(r)$ and $f_6(r)$ monotonically increase with r increasing. In addition, since $f_i(1) < 0$ and $f_i(\bar{r}_2) > 0$ for $i = 4, 5, 6$, each function $f_i(r)$ has one and only one root, denoted by R_i, inside $1 < r_1 < \bar{r}_2$. Therefore, the solution to (3.45) is given by $r \leq \min\{R_4, R_5, R_6\}$. This leads to the fact that the maximum r is determined by

$$R_{\pi,2} = \min\{R_4, R_5, R_6\}$$

Now, putting the results on these two situations together gives us the overall maximum r for the given user permutation π, $R_\pi = \max\{R_{\pi,1}, R_{\pi,2}\}$.

Chapter 4

Simulation and Discussion

In this chapter, we carry out simulations to verify our theoretical results by comparing the SEP of our proposed optimal system with that of the existing minimum-distance constellation design scheme in [12], i.e., a unipolar L-level PAM constellation (L-PAM) with equal power distance, where the average conditional received signal energy given the transmitted signal is approximated as a Gaussian random variable with constant variance for the large number of antennas by the central limit theorem. To make al the comparisons as fair and comprehensive as possible, we detect the transmitted signal with the energy detector proposed in [12] and our proposed fast ML detector in Algorithm 1 in all the simulations. More details are revealed to show the SEP comparison for the MIMO system over Rayleigh fading channels, with the number of receiver antennas increasing.

4.1 Two-User Case

Firstly, let us consider the two-user case. Figs 4.1-4.4, and Figs 4.5 and 4.6 show the error performance of our optimal constellation scheme and the minimum-distance constellation design scheme in the noncoherent MIMO system over Rayleigh fading channels with the number of receiver antennas increasing at SNR=0dB, SNR=10dB, SNR=20dB, respectively. We can observe that both schemes approximately exhibits exponential decay as the number of receiver antenna increases. From these figures, we can see that the proposed optimal constellation yields much better error performance than the minimum-distance PAM constellation. In addition, the error performance also depends on SNR and the better error performance is shown as SNR increases, which can be observed from Fig.4.1 and Figs 4.5 and 4.6. Moreover, Figs 4.1-4.4 show that the error performance for both constellation varies as different values of $\beta_1, \beta_2, p_1, p_2$ when SNR is fixed. Specifically, as shown in Figs 4.3 and 4.4, better error performance can be achieved in the case when $\beta_1 < \beta_2, p_1 > p_2$, when comparing with the case when $\beta_1 < \beta_2, p_1 < p_2$. Similarly, when comparing with the case when $\beta_1 > \beta_2, p_1 > p_2$, better error performance can be obtained in the case when $\beta_1 > \beta_2, p_1 < p_2$.

4.2 Three-User Case

Then, we consider the three-user case. Figs 4.7-4.11 show the SEP in the considered MIMO system over Rayleigh fading channels with the number of receiver antennas increasing at SNR=0dB, SNR=10dB, SNR=20dB, respectively. Similarly, better error performance is observed as the number of receiver antennas or SNR increases. Moreover, we also observe that error performance is better in the case when $\beta_1 p_1 \neq \beta_2 p_2 \neq \beta_3 p_3$ (shown in Fig 4.7, Fig 4.9, Fig 4.11) comparing with the case when $\beta_1 p_1 = \beta_2 p_2 = \beta_3 p_3$ (shown in Fig 4.8, Fig 4.10, Fig 4.12). In this case, our proposed optimal constellation scheme also obtain much better error performance than the existing minimum-distance PAM constellation scheme.

Computer simulations validate our theoretical analysis and demonstrate that our proposed optimal constellation attains significant performance gains over the currently available constellation design for the same massive MIMO upink system. In the future, our task is to develop an efficient algorithm for devising an optimal AUD-CG that maximizes the geometrical coding gain for the considered NOMA MaMIMO system in a more general case.

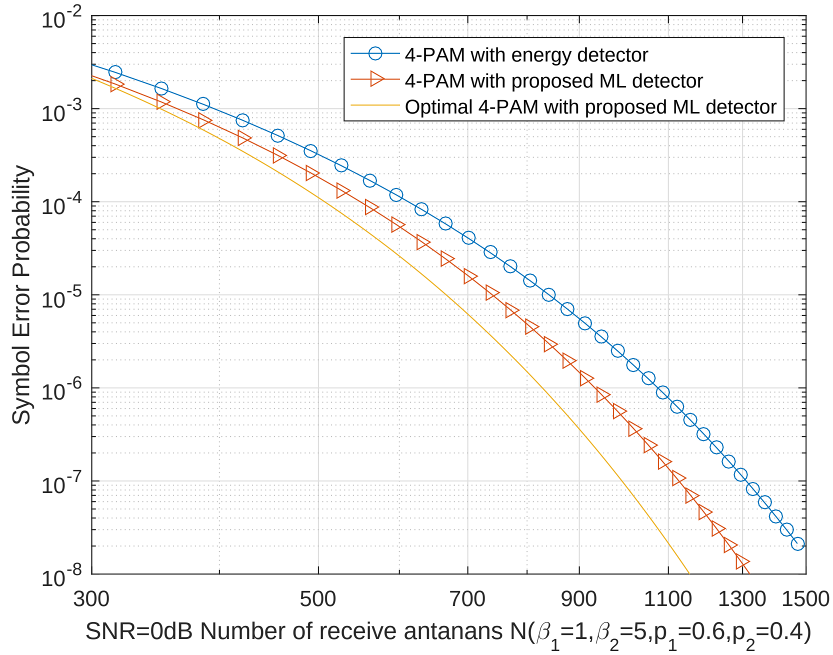

Figure 4.1: Error performance comparison for 2-bit(L=4) at SNR=0dB

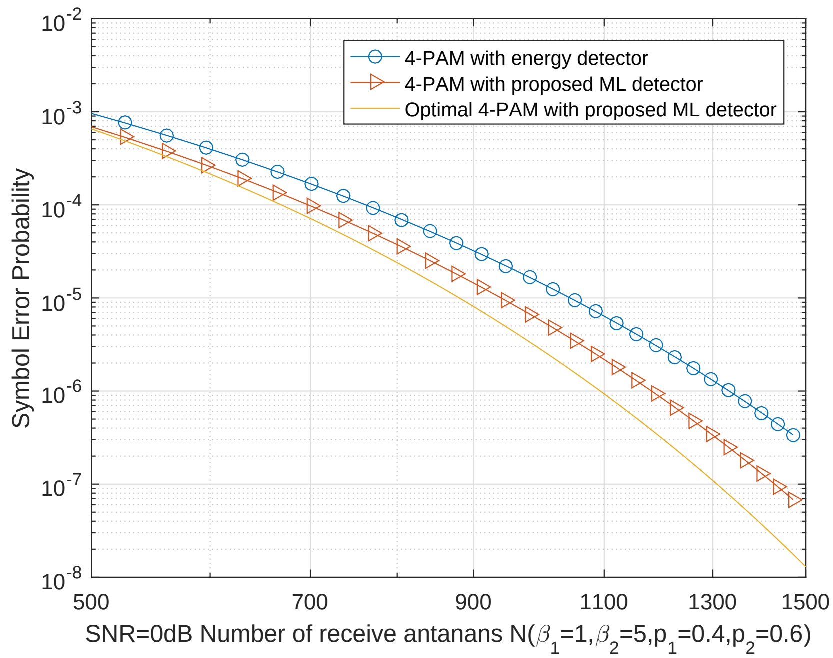

Figure 4.2: Error performance comparison for 2-bit(L=4) at SNR=0dB

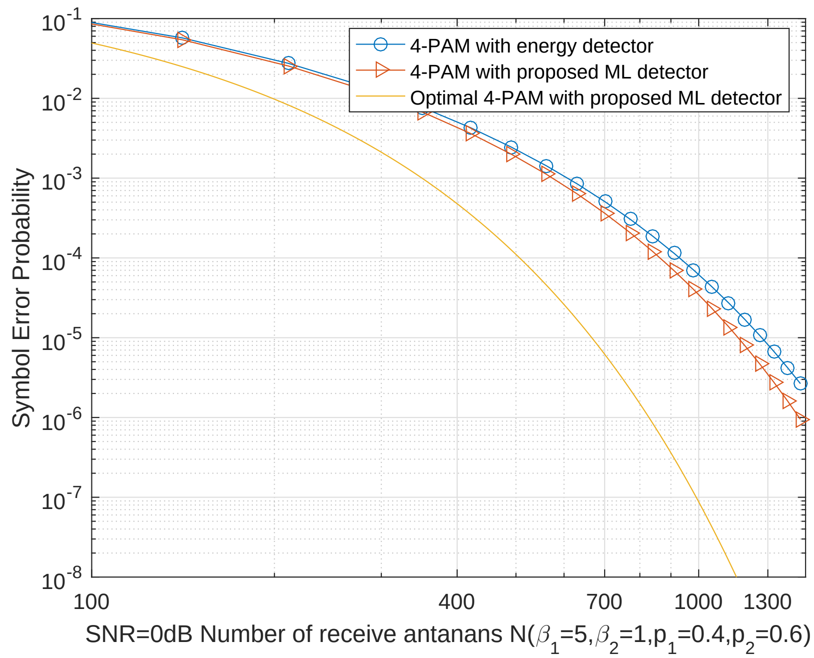

Figure 4.3: Error performance comparison for 2-bit(L=4) at SNR=0dB

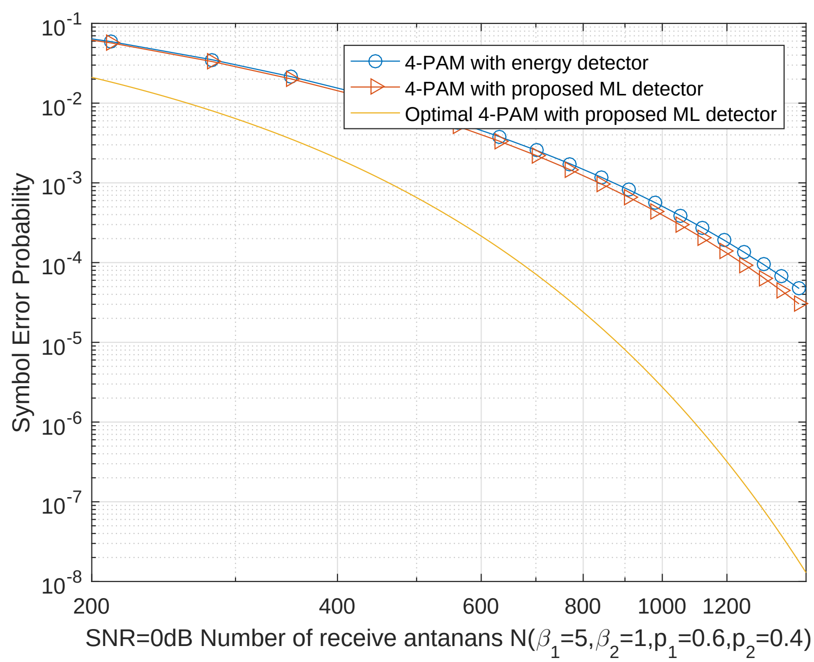

Figure 4.4: Error performance comparison for 2-bit(L=4) at SNR=0dB

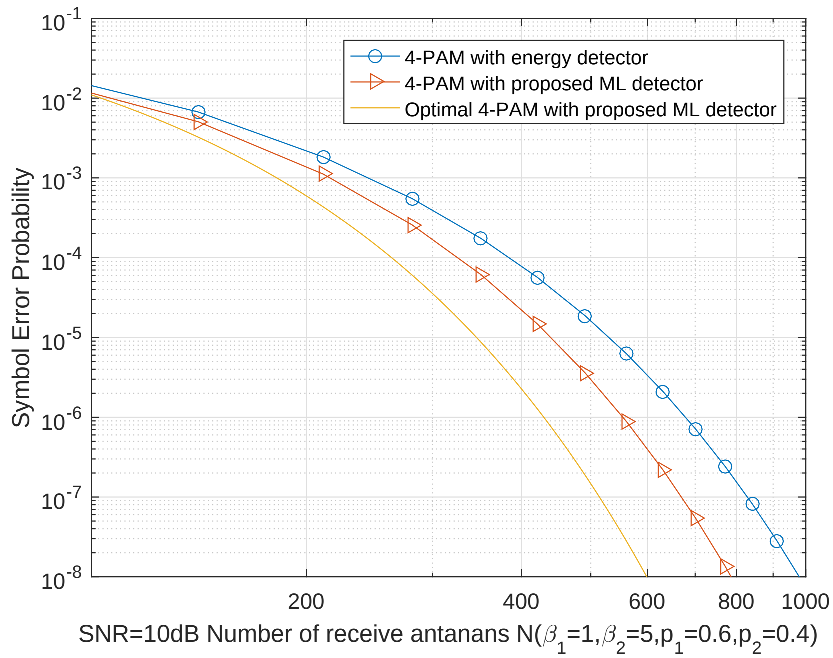

Figure 4.5: Error performance comparison for 2-bit(L=4) at SNR=10dB

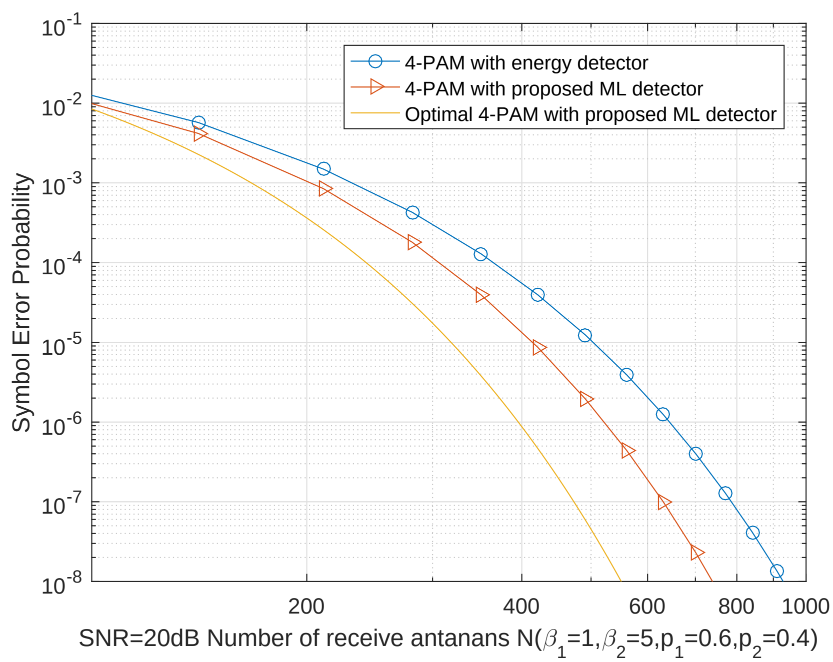

Figure 4.6: Error performance comparison for 2-bit(L=4) at SNR=20dB

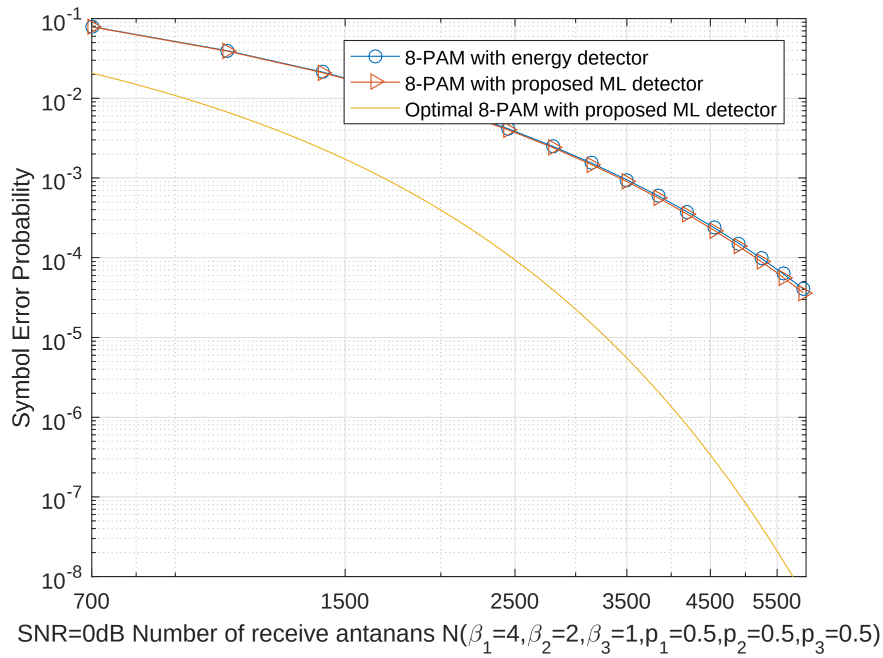

Figure 4.7: Error performance comparison for 3-bit(L=8) at SNR=0dB

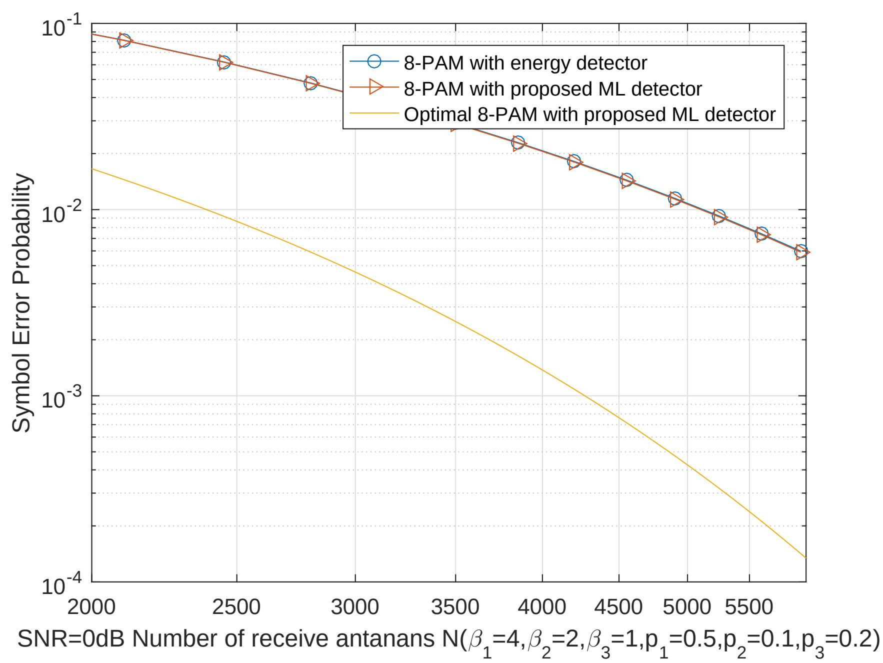

Figure 4.8: Error performance comparison for 3-bit(L=8) at SNR=0dB

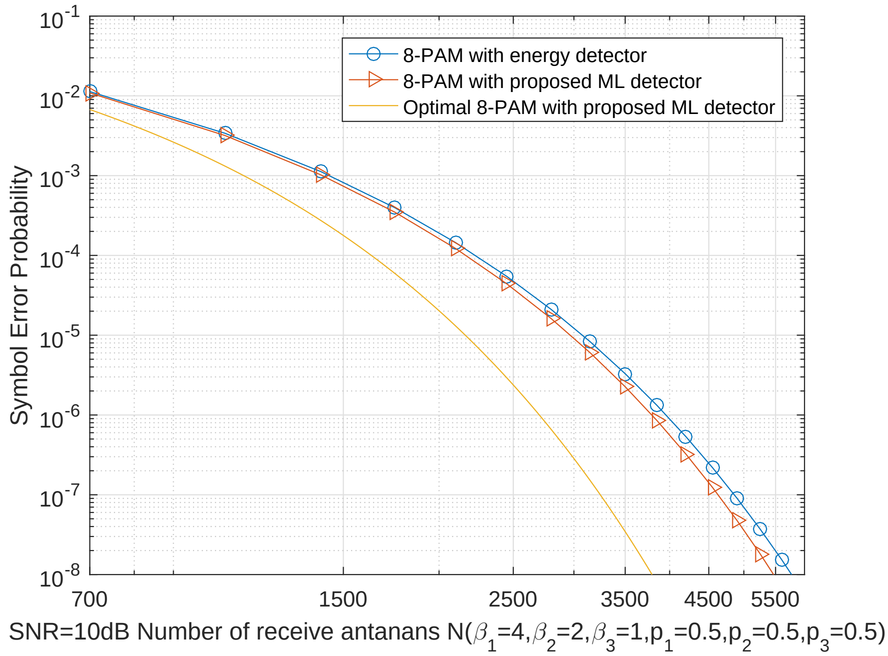

Figure 4.9: Error performance comparison for 3-bit(L=8) at SNR=10dB

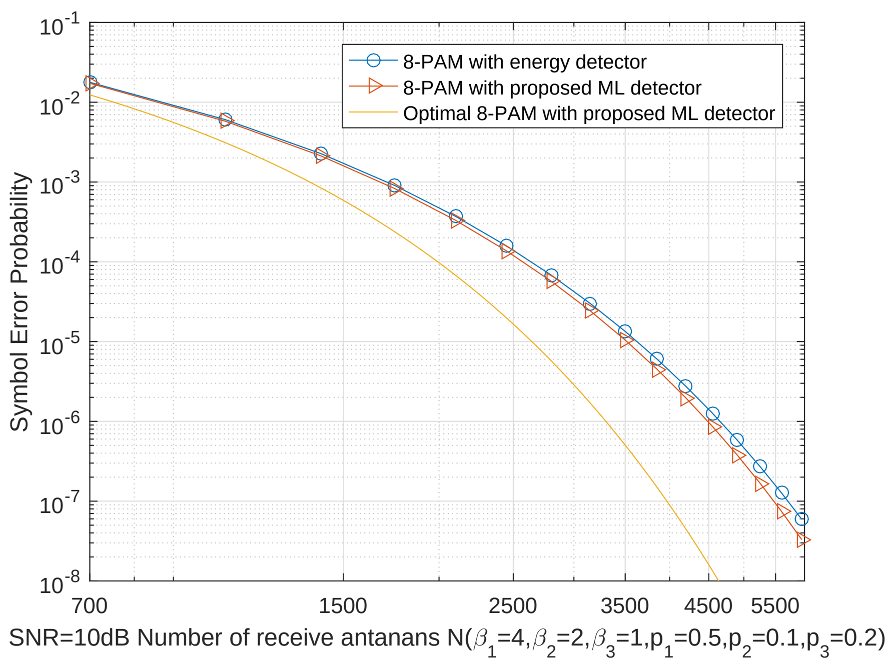

Figure 4.10: Error performance comparison for 3-bit(L=8) at SNR=10dB

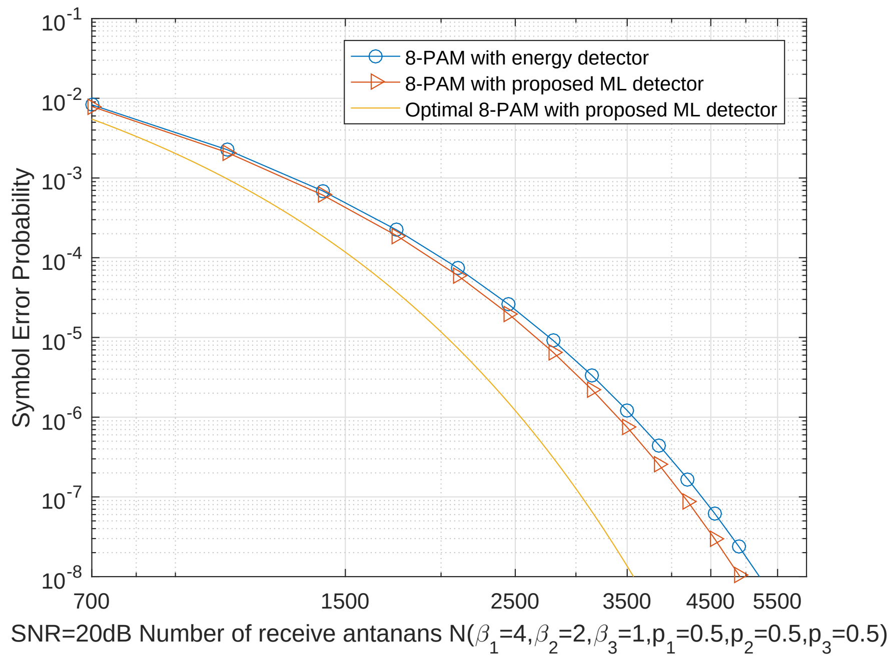

Figure 4.11: Error performance comparison for 3-bit(L=8) at SNR=20dB

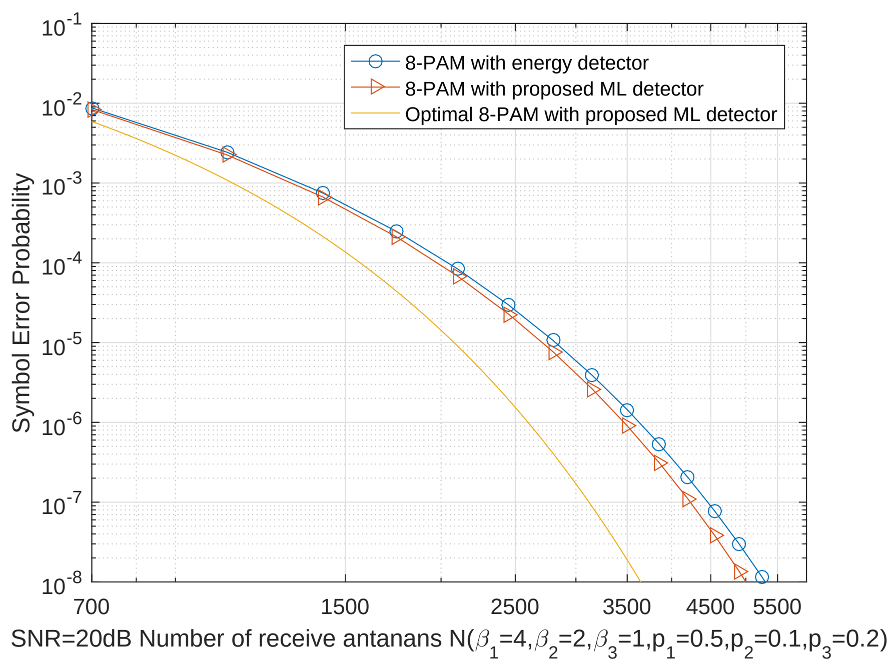

Figure 4.12: Error performance comparison for 3-bit(L=8) at SNR=20dB

Bibliography

[1] Abou-Faycal, I. C., Trott, M. D., and Shamai, S. (2001). The capacity of discrete time memoryless Rayleigh-fading channels. *IEEE Trans. Inf. Theory*, **44**, 1290–1301.

[2] Ahlswede, R. and Balakirsky, V. (1999). Construction of uniquely decodable codes for the two-user binary adder channel. *IEEE Trans. Inf. Theory*, **45**(1), 326–330.

[3] Andrews, J. G., Buzzi, S., Choi, W., Hanly, S. V., Lozano, A., Soong, A. C. K., and Zhang, J. C. (2014). What will 5G be? *IEEE J. Sel. Areas Commun.*, **32**(6), 1065–1082.

[4] Avestimehr, A. S., Diggavi, S. N., and Tse, D. N. C. (2011). Wireless network information flow: A deterministic approach. *IEEE Trans. Inf. Theory*, **57**(4), 1872–1905.

[5] Biguesh, M. and Gershman, A. B. (2006). Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals. *IEEE Trans. Signal Process.*, **54**(3), 884–893.

[6] Boccardi, F., Heath, R. W., Lozano, A., Marzetta, T. L., and Popovski, P. (2014). Five disruptive technology directions for 5G. *IEEE Commun. Mag.*, **52**(2), 74–80.

[7] Brahmi, N., Yilmaz, O. N. C., Helmersson, K. W., Ashraf, S. A., and Torsner, J. (2015). Deployment strategies for ultra-reliable and low-latency communication in factory automation. In *IEEE Globecom Workshops*.

[8] Bresler, G., Parekh, A., and Tse, D. N. C. (2010). The approximate capacity of the many-to-one and one-to-many Gaussian interference channels. *IEEE Trans. Inf. Theory*, **56**(9), 4566–4592.

[9] Cadambe, V. R. and Jafar, S. A. (2008). Interference alignment and degrees of freedom of the K-user interference channel. *IEEE Trans. Inf. Theory*, **54**(8), 3425–3441.

[10] Chevillat, P. (1981). N-user trellis coding for a class of multiple-access channels. *IEEE Trans. Inf. Theory*, **27**(1), 114–120.

[11] Chowdhury, M., Manolakos, A., and Goldsmith, A. (2016a). Energy-based modulation for noncoherent massive SIMO systems. *IEEE Trans. Wireless Commun.*, **15**, 7831–7846.

[12] Chowdhury, M., Manolakos, A., and Goldsmith, A. (2016b). Scaling laws for noncoherent energy-based communications in the SIMO MAC. *IEEE Trans. Inf. Theory*, **62**, 1980–1992.

[13] Chuah, C.-N., Tse, D., Kahn, J., and Valenzuela, R. (2002). Capacity scaling in MIMO wireless systems under correlated fading. *IEEE Trans. Inf. Theory*, **48**(3), 637–650.

[14] Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*. Wiley-Interscience, 2nd edition.

[15] Ding, Z., Krikidis, I., Sharif, B., and Poor, H. V. (2014). Wireless information and power transfer in cooperative networks with spatially random relays. *IEEE Trans. Wireless Commun.*, **13**(8), 4440–4453.

[16] Dong, Z., Chen, H., Zhang, J.-K., Huang, L., and Vucetic, B. (????). Uplink non-orthogonal multiple access with finite-alphabet inputs. *accepted by IEEE Trans. Wireless Commun.*

[17] Dong, Z., Zhang, Y.-Y., Zhang, J.-K., and Gao, X.-C. (2016). Quadrature amplitude modulation division for multiuser MISO broadcast channels. *IEEE J. Sel. Topics Signal Processing*, **10**, 1551–1566.

[18] Dong, Z., Chen, H., Zhang, J.-K., and Huang, L. (2017). On non-orthogonal multiple access with finite-alphabet inputs in z-channels. *IEEE J. Sel. Areas Commun.*, **35**, 2829–2845.

[19] Elijah, O., Leow, C., Rahman, T., Nunoo, S., and Iliya, S. (2015). A comprehensive survey of pilot contamination in massive MIMO-5G system. *IEEE Commun. Surveys Tuts.*, **PP**(99), 1–1.

[20] Gamal, A. E. and Kim, Y.-H. (2012). *Network Information Theory*. Cambridge Univ. Press.

[21] Gursoy, M. C., Poor, H. V., and Verdu, S. (2005). The noncoherent Rician fading channel-part i: Structure of the capacity-achieving input. *IEEE Trans. Wireless Commun.*, **4**, 2193–2206.

[22] Hachem, W., Khorunzhiy, O., Loubaton, P., Najim, J., and Pastur, L. (2008). A

new approach for mutual information analysis of large dimensional multi-antenna channels. *IEEE Trans. Inf. Theory*, **54**(9), 3987–4004.

[23] Han, T. S. and Kobayashi, K. (1981). A new achievable rate region for the interference channel. *IEEE Trans. Inf. Theory*, **27**(1), 49–60.

[24] Harshan, J. and Rajan, B. (2011). On two-user gaussian multiple access channels with finite input constellations. *IEEE Trans. Inf. Theory*, **57**(3), 1299–1327.

[25] Hassibi, B. and Hochwald, B. M. (2003). How much training is needed in multiple-antenna wireless links? *IEEE Trans. Inf. Theory*, **49**(4), 951–963.

[26] Hossain, E., Rasti, M., Tabassum, H., and Abdelnasser, A. (2014). Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective. *IEEE Wireless Commun.*, **21**(3), 118–127.

[27] Hossain, J., Alouini, M.-S., and Bhargava, V. (2007). Multi-user opportunistic scheduling using power controlled hierarchical constellations. *IEEE Trans. Wireless Commun.*, **6**(5), 1581–1586.

[28] Hoydis, J., ten Brink, S., and Debbah, M. (2013). Massive MIMO in the UL/DL of cellular networks: How many antennas do we need? *IEEE J. Sel. Areas Commun.*, **31**(2), 160–171.

[29] Huang, K. and Lau, V. K. N. (2014). Enabling wireless power transfer in cellular networks: Architecture, modeling and deployment. *IEEE Trans. Wireless Commun.*, **13**(2), 902–912.

[30] Jing, L., Carvalho, E. D., Popovski, P., and Martinez, A. O. (2016). Design

and performance analysis of noncoherent detection systems with massive receiver arrays. *IEEE Trans. Wireless Commun.*, **64**, 5000–5010.

[31] Jozsef, S. and Debnath, L. (2000). On certain inequalities involving the constant e and their applications, jvol. 249, no. 2, pp. 569. *Math. Anal. Appl.,*, **249**, 569–582.

[32] Kasami, T. and Lin, S. (1976). Coding for a multiple-access channel. *IEEE Trans. Inf. Theory*, **22**(2), 129–137.

[33] Kasami, T. and Lin, S. (1978). Bounds on the achievable rates of block coding for a memoryless multiple-access channel. *IEEE Trans. Inf. Theory*, **24**(2), 187–197.

[34] Krikidis, I., Timotheou, S., Nikolaou, S., Zheng, G., Ng, D. W. K., and Schober, R. (2014). Simultaneous wireless information and power transfer in modern communication systems. *IEEE Commun. Mag.*, **52**(11), 104–110.

[35] Larsson, E., Edfors, O., Tufvesson, F., and Marzetta, T. (2014). Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.*, **52**(2), 186–195.

[36] Li, Y., Murata, H., and Yoshida, S. (1998). Coding for multi-user detection in interference channel. In *Proc. IEEE GLOBECOM 98*, volume 6, pages 3596–3601 vol.6.

[37] Liu, Y., Zhang, Y., Yu, R., and Xie, S. (2015). Integrated energy and spectrum harvesting for 5G wireless communications. *IEEE Network*, **29**(3), 75–81.

[38] Lu, L., Li, G., Swindlehurst, A., Ashikhmin, A., and Zhang, R. (2014). An overview of massive MIMO: Benefits and challenges. *IEEE J. Sel. Topics Signal Process.*, **8**(5), 742–758.

[39] Lu, X., Wang, P., Niyato, D., Kim, D. I., and Han, Z. (2015). Wireless networks with RF energy harvesting: A contemporary survey. *IEEE Commun. Surveys Tuts.*, **17**(2), 757–789.

[40] Malladi, D. (2012). Hierarchical modulation for communication channels in single-carrier frequency division multiple access. US Patent 8,259,848.

[41] Marcenko, V. A. and Pastur, L. A. (1967). Distributions of eigenvalues for some sets of random matrices. *Math USSR-Sbornik*, **1**, 457–483.

[42] Marzetta, T. L. (2006). How much training is required for multiuser MIMO? In *Proc. Asilomar Conf. Signals Systems Comput.*, pages 359–363.

[43] Marzetta, T. L. (2010). Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Trans. Wireless Commun.*, **9**(11), 3590–3600.

[44] Masouros, C. and Alsusa, E. (2009). Dynamic linear precoding for the exploitation of known interference in MIMO broadcast systems. *IEEE Trans. Wireless Commun.*, **8**(3), 1396–1404.

[45] Masouros, C. and Ratnarajah, T. (2012). Interference as a source of green signal power in cognitive relay assisted co-existing MIMO wireless transmissions. *IEEE Trans. Commun.*, **60**(2), 525–536.

[46] Masouros, C. and Zheng, G. (2015). Exploiting known interference as green signal power for downlink beamforming optimization. *IEEE Trans. Signal Process.*, **63**(14), 3628–3640.

[47] Masouros, C., Ratnarajah, T., Sellathurai, M., Papadias, C. B., and Shukla,

A. K. (2013). Known interference in the cellular downlink: a performance limiting factor or a source of green signal power? *IEEE Commun. Mag.*, **51**(10), 162–171.

[48] McKay, M., Collings, I., and Tulino, A. (2010). Achievable sum rate of MIMO MMSE receivers: A general analytic framework. *IEEE Trans. Inf. Theory*, **56**(1), 396–410.

[49] Mestre, X., Fonollosa, J., and Pages-Zamora, A. (2003). Capacity of MIMO channels: asymptotic evaluation under correlated fading. *IEEE J. Sel. Areas Commun.*, **21**(5), 829–838.

[50] Moustakas, A., Simon, S., and Sengupta, A. (2003). MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large N analysis. *IEEE Trans. Inf. Theory*, **49**(10), 2545–2561.

[51] Murata, H. and Yoshida, S. (1997). Trellis-coded cochannel interference canceller for microcellular radio. *IEEE Trans. Commun.*, **45**(9), 1088–1094.

[52] Ng, D. W. K., Lo, E. S., and Schober, R. (2013). Wireless information and power transfer: Energy efficiency optimization in OFDMA systems. *IEEE Trans. Wireless Commun.*, **12**(12), 6352–6370.

[53] Ramchandran, K., Ortega, A., Uz, K., and Vetterli, M. (1993). Multiresolution broadcast for digital HDTV using joint source/channel coding. *IEEE J. Sel. Areas Commun.*, **11**(1), 6–23.

[54] Rusek, F., Persson, D., Lau, B. K., Larsson, E., Marzetta, T., Edfors, O., and Tufvesson, F. (2013). Scaling up MIMO: Opportunities and challenges with very large arrays. *IEEE Signal Process. Mag.*, **30**(1), 40–60.

[55] Shepard, C., Yu, H., Anand, N., Li, L. E., Marzetta, T. L., Yang, R., and Zhong, L. (2012). Argos: Practical many-antenna base stations. In *Proc. ACM Int. Conf. Mobile Computing, Networking, MobiCom'12*, Istanbul, Turkey.

[56] Telatar, I. (1999). Capacity of multiple antenna Gaussian channels. *Europ. Trans. Telecommu.*, **10**, 585–595.

[57] van den Braak, P. and van Tilborg, H. (1985). A family of good uniquely decodable code pairs for the two-access binary adder channel. *IEEE Trans. Inf. Theory*, **31**(1), 3–9.

[58] Wei, L.-F. (1993). Coded modulation with unequal error protection. *IEEE Trans. Commun.*, **41**(10), 1439–1449.

[59] Wu, Y., Shitz, S. S., and Verdú, S. (2015). Information dimension and the degrees of freedom of the interference channel. *IEEE Trans. Inf. Theory*, **61**(1), 256–279.

[60] Zeng, Y. and Zhang, R. (2015). Optimized training design for wireless energy transfer. *IEEE Trans. Commun.*, **63**(2), 536–550.

[61] Zhang, R. and Ho, C. K. (2013). MIMO broadcasting for simultaneous wireless information and power transfer. *IEEE Trans. Wireless Commun.*, **12**(5), 1989–2001.

[62] Zhang, W., D'Amours, C., and Yongacoglu, A. (2004). Trellis coded modulation design for multi-user systems on AWGN channels. In *Proc. IEEE 59th Vehi. Tech. Conf. (VTC Spring'04)*, volume 3, pages 1722–1726 Vol.3.

[63] Zhang, Y.-Y., Yu, H.-Y., Zhang, J.-K., and Zhu, Y.-J. (2016). Signal-cooperative

multilayer modulated vlc systems for automotive applications. *IEEE Photonics Journal*, **8**, 1–9.

[64] Zhang, Y.-Y., Yu, H.-Y., and Zhang, J.-K. (2018a). Block precoding for peak-limited miso broadcast vlc: constellation-optimal structure and addition-unique designs. *IEEE J. Sel. Areas Commun.*, **36**, 78–90.

[65] Zhang, Y.-Y., Zhang, J.-K., and Yu, H.-Y. (2018b). Physically securing energy based massive MIMO MAC via joint alignment of multi-user constellations and artificial noise,. *IEEE J. Sel. Areas Commun.*, **36**, 829–844.

[66] Zheng, G., Krikidis, I., Li, J., Petropulu, A. P., and Ottersten, B. (2013). Improving physical layer secrecy using full-duplex jamming receivers. *IEEE Trans. Signal Process.*, **61**(20), 4962–4974.

[67] Zheng, G., Krikidis, I., Masouros, C., Timotheou, S., Toumpakaris, D. A., and Ding, Z. (2014). Rethinking the role of interference in wireless networks. *IEEE Commun. Mag.*, **52**(11), 152–158.

[68] Zheng, L. and Tse, D. (2003). Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *IEEE Trans. Inf. Theory*, **49**(5), 1073–1096.