PHYSICAL-LAYER NETWORK CODING FOR MIMO SYSTEMS

Ning Xu, B.S., M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2011

APPROVED:

Yan Huang, Major Professor
Shengli Fu, Co-Major Professor
Xinrong Li, Committee Member
Miguel Acevedo, Committee Member
Murali Varanasi, Committee Member
Ian Parberry, Interim Chair of the Department of Computer Science and Engineering
Costas Tsatsoulis, Dean of the College of Engineering
James D. Meernik, Acting Dean of the Toulouse Graduate School

Xu, Ning. Physical-layer network coding for MIMO systems. Doctor of Philosophy (Computer Science and Engineering), May 2011, 112 pp., 7 tables, 45 illustrations, bibliography, 99 titles.

The future wireless communication systems are required to meet the growing demands of reliability, bandwidth capacity, and mobility. However, as corruptions such as fading effects, thermal noise, are present in the channel, the occurrence of errors is unavoidable. Motivated by this, the work in this dissertation attempts to improve the system performance by way of exploiting schemes which statistically reduce the error rate, and in turn boost the system throughput. The network can be studied using a simplified model, the two-way relay channel, where two parties exchange messages via the assistance of a relay in between. In such scenarios, this dissertation performs theoretical analysis of the system, and derives closed-form and upper bound expressions of the error probability. These theoretical measurements are potentially helpful references for the practical system design. Additionally, several novel transmission methods including block relaying, permutation modulations for the physical-layer network coding, are proposed and discussed. Numerical simulation results are presented to support the validity of the conclusions.

ACKNOWLEDGMENTS

I would like to give many thanks to Dr. Shengli Fu and Dr. Yan Huang as my advisors, and Dr. Xinrong Li, Dr. Miguel Acevedo, Dr. Murali Varanasi, as my committee members for their guidance and assistance, without whom none of this would have become mission possible. Also I would like to thank my family, friends, fellow researchers and colleagues, faculty and staff members from Computer Science and Engineering, and Electrical Engineering departments for their great support. And my acknowledgment to the National Science Foundation (NSF), by which the research conducted in this dissertation is partially supported.

The content of certain sections or chapters is reprinted or reproduced under the permission from the publishers of IEEE and John Wiley & Sons, and the corresponding copyright acknowledgments are placed in the footnotes on the first page of individual chapters.

CONTENTS

ACKNOWLEDGMENTS iii

LIST OF TABLES vii

LIST OF FIGURES viii

GLOSSARY xii

CHAPTER 1. INTRODUCTION 1

1.1. Motivations 1

1.2. Research Topics and Contributions 2

1.3. Organization of the Dissertation 3

CHAPTER 2. THEORETICAL BACKGROUND AND SYSTEM MODEL 7

2.1. Two-Way Communication Systems 7

2.1.1. Cooperative Communication Systems 8

2.1.2. Two-Way Relay Channel 10

2.1.3. Coding Techniques 11

2.1.4. Relaying Protocols 13

2.2. Space-Time Coding 14

2.2.1. Alamouti Scheme 15

2.3. Other Related Works 17

2.4. Summary 18

CHAPTER 3. BLOCK RELAYING FOR PHYSICAL-LAYER NETWORK CODING 19

3.1. Background 20

3.2. System Model 21

3.2.1. Amplify-and-Forward 24

3.2.2. Decode-and-Forward 24

3.2.3. Assumptions on the Power Allocation 25

3.3. Performance Analysis 25

3.3.1. Rayleigh Fading Channels 27
3.3.2. AWGN Channels 29
3.3.3. Beneficial SNR Range 30
3.3.4. Optimal Power Allocation 31
3.4. Numerical Results 32
3.4.1. Amplify-and-Forward 33
3.4.2. Decode-and-Forward 34
3.5. Summary 39

CHAPTER 4. POWER MANAGEMENT OF THE DISTRIBUTED DECODE-AND-FORWARD PROTOCOL IN THE PARALLEL RELAY NETWORKS 41
4.1. Motivation 41
4.2. System Model 43
4.3. Performance Analysis and Numerical Results 46
4.3.1. Closed-Form SER Performance of ACDF 46
4.3.2. Upper Bound and Asymptotic Tight Approximation of SER of ACDF 48
4.3.3. Optimized Power Allocation Policy 51
4.4. Summary 55

CHAPTER 5. PERMUTATION MODULATIONS FOR PHYSICAL-LAYER NETWORK CODING 56
5.1. Background 56
5.2. System Model 58
5.2.1. Permutation Modulation 58
5.2.2. Encoding and Decoding in Binary Permutation Modulations 59
5.2.3. Physical-Layer Network Coding 60
5.2.4. Discussion on the Failure Cases of the Proposed Mapping 63
5.3. Performance Analysis 69

5.4. Numerical Results 73
5.5. Summary 81
CHAPTER 6. ON THE PERFORMANCE OF TWO-WAY RELAY CHANNELS USING SPACE-TIME CODES 82
6.1. Introduction 83
6.2. System Model 85
6.2.1. DF Strategy for Single-Antenna TWRCs 85
6.2.2. DF for Space-Time Coded TWRCs 87
6.2.3. AF for Space-Time Coded TWRCs 89
6.2.4. PDF for Space-Time Coded TWRCs 90
6.3. Performance Analysis 90
6.3.1. SER of Single-Antenna TWRCs 91
6.3.2. SER of Space-Time Coded TWRCs 93
6.4. Numerical Results 96
6.5. Summary 100
CHAPTER 7. CONCLUSIONS 101
7.1. Practical Issues 103
BIBLIOGRAPHY 104

LIST OF TABLES

3.1 β_{opt} for $\log_2 M$-MA (16-QAM) in the AWGN channels. 32

4.1 The encoding and transmission sequence for each relay. 44

5.1 The occurrence of codeword addition results in the upper triangle in Fig. 5.4. 65

5.2 The occurrence of codeword addition results in the upper triangle in Fig. 5.5. 67

5.3 The occurrence of codeword addition results in the upper triangle in Fig. 5.6. 68

6.1 Network coding mapping scheme for the BPSK modulation. 86

6.2 Network coding mapping scheme for the QPSK modulation. 87

LIST OF FIGURES

2.1 Improvements of the transmission strategies in the two-way relay channel: from (a) traditional to (b) network coding, and to (c) physical-layer network coding [1]. 9

2.2 A simplified cooperation communication model. 9

2.3 The end-to-end BER performance comparison of detect-and-forward (DF), estimate-and-forward (EF), and MI based forward (MIF) in a parallel realy network with two relays. 14

2.4 The two transmit antenna diversity scheme. 15

3.1 The conventional and $\log_{M_1}(M_2)$-MA block relaying schemes for the two-way relay systems (blocks are indicated by round corner rectangles, (a) (b) can be seen as special cases of $\log_{M_1}(M_2)$-MA for $M_1 = M_2$, namely $\log_2 2$-MA, $\log_{16} 16$-MA respectively). 21

3.2 The end-to-end PER performance of AF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in fading channels when $L = 128$. 33

3.3 The end-to-end throughput performance of AF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in fading channels when $L = 128$. 34

3.4 The end-to-end PER performance of DF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in fading channels when $L = 128$: simulation results versus theoretical values. 35

3.5 The end-to-end throughput performance of DF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in fading channels when $L = 128$: simulation results versus theoretical values. 35

3.6 The end-to-end PER performance of DF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in AWGN channels when $L = 128$: simulation results versus theoretical values. 36

3.7 The end-to-end throughput performance of DF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in AWGN channels when $L = 128$: simulation results versus theoretical values. 36

3.8 The end-to-end throughput comparison between $\log_2 16$-MA (16-PSK) and the conventional methods using BPSK ($B - B$ in the figure legend), 16-PSK ($16 - 16$) throughout in fading channels. The packet sizes L are 64, 128 and 256. 37

3.9 The end-to-end throughput comparison between $\log_2 M$-MA's for $M = 2, 4, 8, 16, 32, 64, 128$ (M-PSK) in fading channels under the packet length $L = 128$. 38

3.10 The end-to-end throughput comparison between schemes using the optimal power allocation policy and the non-optimal ($L = 128$): $\log_2 16$-MA (16-QAM) in AWGN channels. 39

3.11 The end-to-end throughput comparison between schemes using the optimal power allocation policy and the non-optimal ($L = 128$): $\log_2 4$-MA (QPSK) in fading channels. 40

4.1 The fading parallel relay channel model. 43

4.2 Comparison of the closed-form SER, upper bound and asymptotically tight approximation. (QPSK modulation, $P_S = P_{R_1,D} = P_{R_2,D}$, $\sigma^2_{*,*} = 1$, $N_0 = 1$.) 50

4.3 P_S/P and $P_{R_1,D}/P$ ratios plotted against the SNR. ($P_{R_1,D} = P_{R_2,D}$, $\sigma^2_{*,*} = 1$, $N_0 = 1$.) 53

4.4 The end-to-end error performance of ACDF using BPSK. 54

4.5 The end-to-end SER comparison between two intuitive and the optimized power allocation policies using BPSK. 54

4.6 The end-to-end SER comparison between two intuitive and the optimized power allocation policies using QPSK. 55

5.1 The two-way relay channel model. 60

5.2 The mapping scheme of PLNC method II using PM[1,3]. 62

5.3 XOR in decimal formats of the binary indices when $B = 3$. 63

5.4 Failure case 1 of the mapping attempt for BPM[2,3] when $B = 3$. (The occurrence of 30 is 3. It must be grouped with the results with the occurrence of 1. The candidates are 9, 15, 17, 23, all of which share a row or column with 30.) 64

5.5 Failure case 2 of the mapping attempt for BPM[2,3] when $B = 3$. (The same group is filled with the same color, numbers therein of the same style: bold, italic, underlined, or a combination of those. The only valid groups are {6, 29}, {15, 30}, {10, 17, 27}, {3, 18}, {5, 20}, {9, 24}. 12 must be grouped with 23. However they share the same column.) 66

5.6 Failure case 3 of the mapping attempt for BPM[2,3] when $B = 3$. (The same group is filled with the same color, numbers therein of the same style: bold, italic, underlined, or a combination of those. The groups are formed with all distinct results: {12, 27}, {0}, {6, 17, 23}, {15, 30}, {3, 18}, {5, 20}, {10, 29}, {9, 24}. However, this mapping cannot be transformed by row/column swapping to match that in Table 5.3.) 67

5.7 The scatter plot of the ratio $\frac{d^2_{min}(PM[m_1,m_2])}{d^2_{min}(BPSK)}$ (◯: ratio is 1; ×: ratio < 1; □: ratio > 1, a larger size marker indicates a larger value). 74

5.8 The plot of the ratio $\frac{d^2_{min}(PM[m_1,m_2])}{d^2_{min}(BPSK)}$ for $m_1 = m_2$. 75

5.9 The contour plot of the ratio $\frac{d^2_{min}(PM[m_1,m_2])}{d^2_{min}(BPSK)}$. 76

5.10 The SER performance comparison between BPM[3,3], BPM[2,5] and BPSK in one-to-one transmissions for $B = 4$. 76

5.11 The SER performance comparison between BPM[4,4], BPM[3,6] and BPSK in one-to-one transmissions for $B = 6$. 77

5.12 The SER performance comparison between the random codeword selection and the selection based on the maximized distance sum of BPM[2,2] (choosing 4 codewords out of 6). 78

5.13 The SER performance comparison between the fast decoding and the ML based decoding for BPM[5,5]. 79

5.14 The SER performance comparison of PLNC using BPSK and BPM method I (BPM[3,3] in MA, BPM[4,4] in BC) for $B = 3$. 79

5.15 The SER performance comparison of PLNC using BPSK and BPM method II for $B = 2$. 80

5.16 The SER performance comparison of PLNC using BPSK and BPM method II for $B = 3$. 81

6.1 The time scheduling of the two-way relay systems: (a) the conventional four-slot scheme, (b) the three-slot scheme with network coding, (c) the two-slot scheme with the physical-layer network coding. 83

6.2 Communication over a single-antenna TWRC. 85

6.3 Communication over TWRC with space-time codes. 88

6.4 The end-to-end error performance of DF for TWRC with single antenna using BPSK: simulation vs. theoretical upper bound. 97

6.5 The end-to-end error performance of DF for TWRC with single antenna using QPSK: simulation vs. theoretical upper bound. 98

6.6 The end-to-end error performance of DF for TWRC with space-time codes using BPSK: simulation vs. theoretical upper bound. 98

6.7 The end-to-end error performance of DF for TWRC with space-time codes using QPSK: simulation vs. theoretical upper bound. 99

6.8 The SER comparison of AF, PDF and DF using BPSK by simulations. 99

6.9 The SER comparison of AF, PDF and DF using QPSK by simulations. 100

GLOSSARY

AF	amplify-and-forward
BER	bit error rate
BPSK	binary phase-shift keying
DF	decode-and-forward
MIMO	multi-input multi-output (system)
PER	packet error rate
PLNC	physical-layer network coding
PSK	phase-shift keying
QAM	quadrature amplitude modulation
QPSK	quadrature phase-shift keying
SER	symbol error rate
SNR	signal-to-noise ratio
TWRC	two-way relay channel

CHAPTER 1

INTRODUCTION

1.1. Motivations

As the digital age comes marching on, the daily life of the human society is more and more involved in and relies on the exchange of information, and communications to and fro even between remotely located parties. Numerous powerful wireless devices help surpass the space limitation and facilitate instant and comprehensive access to great varieties of resources [2]. This trend continues, calling for more robust and reliable communication, and demanding higher spectral efficiency, network bandwidth and capacity, lower energy consumption, and higher mobility. The future wireless communication systems are required to meet such growing demands [3].

Wireless communication is a very broad topic, covering many subjects and disciplines. The work in this dissertation attempts to address some of the issues in the context of the physical layer of a two-way relayed communication channel model. The communication in most currently used media (air, wires, cables, etc.) is subject to corruptions such as thermal noise of the electronics, the internal structure shifting of the media themselves. Especially for wireless communication, the waveform traveling experiences reflection, refraction, and shadowing, as discovered in certain physics laws. The interaction with the environment potentially causes fading phenomena either in a large scale or a small scale. As the noise and such corruptions exist, the end results of the communication suffer from errors. There are error correcting techniques to detect and repair them or simply request for retransmission. Either way, overhead is introduced. There are other approaches on the system design, such as scheduling, modulation types, to improve the "quality" of the communication in the sense of statistically reducing the error occurrences. This is the interest of this dissertation.

The communication network can be considered as multiple nodes connected by links. Information is being sent as one packet at a time from the source to the destination via one or more intermediate nodes acting as relays [4]. In this process, nodes within each other's vicinity (in terms of the transmission range) can potentially cause interference by simultaneous transmission. This interference was seen only harmful and to be avoided traditionally. Proper scheduling such as time-division protocols allows the nodes to take part in their intended transmission alternately. As a relatively new paradigm in the networking techniques, the physical-layer network coding takes advantage of the additive nature of the electromagnetic waves, and embraces the interferences by performing coding operations to combine the otherwise convoluted messages. In such a way, the communication time and the bandwidth are utilized more efficiently, which in turn boosts the network throughput.

Very often, the entire network can be simplified to one consisting of a source node, a destination node and a relay node, which is sufficient to model and study the system behavior, as it is for the research in this dissertation. Such a model is known as the two-way relay channel. Centering around this model, this dissertation dedicates itself to investigations on many design issues, performance evaluations and discussions.

1.2. Research Topics and Contributions

The work in this dissertation makes contributions to the topics and areas, summarized as below.

- It provides a survey on the latest important research works and results in the area of the physical-layer network coding, which reveals the trend and future direction of this research topic.
- Theoretical performance analysis together with numerical validation is performed on several two-way relay channel models under different channel conditions and system requirements [5, 6, 7, 8, 1]. This serves as potential design references for the real-world applications.

- This work discusses various system design aspects such as allocation policies under a limited total power allowance [7, 8, 1], relaying protocols [5, 6], transmission scheduling [5, 6, 7], modulation/constellation usage [7, 8, 1], etc.
- It addresses design issues and gives theoretically grounded suggestions on parameter choosing.

1.3. Organization of the Dissertation

The rest of the dissertation is organized as follows.

- Chapter 2 (Theoretical Background and System Model): This chapter lays the ground work for the dissertation by giving the detailed problem definition. The typical two-way relay channel model is described. The various aspects of system design and related assumptions include but are not limited to the followings.
  - The link availability: The relay is assumed to be present in the channel to assist the exchange. Some channels are modeled as the terminals cannot communicate with each other directly, whereas in others, links between the terminals and between the transmitting terminal and the relay coexist. The latter case forms a scenario mostly referred to as cooperative relaying in prior works.
  - The number of sources/destinations/relays: This dissertation is primarily concerned with the case where two terminals and a single relay participate in the communications. Chapter 4 as an exception considers two relays instead of one. However, in the existing literatures, cases with multiple terminals at either side, and/or multiple relays have also been studied.
  - The transmission approach: The traditional scheme treats signals from other transmitting parties as interference and makes schedules by alternating transmission between the terminals. The relay forwards messages separately to their intended destinations. With the application of network coding, the relay instead forwards coded messages to the terminals, which later can be recovered. This way, the time consumption for completing the whole process is reduced,

and therefore the system throughput is increased. This network coding technique if applied at the physical layer, allows embracing the interference and turns it into a capacity boost.

  - The relaying strategies: The relaying strategy or protocol is concerned with the processing of the messages received at the relay. There are two types widely used. The amplify-and-forward protocol scales the received signals to its transmission power, serving as a repeater. In contrast, the decode-and-forward protocol attempts to decode the original transmitted signals or their linear combination, which is essentially a process of noise elimination. Both are featured in the discussions in the later chapters.
  - The power allocation: When the terminals are allocated equal power for transmission, the channel is known as a symmetric channel. Otherwise, it is an asymmetric channel. This is the symmetry in terms of transmission power. Both have been investigated in the existing literatures, and some related issues are covered in Chapters 3, 4. There are other parameters which render the channel asymmetric, such as when the channel coefficients statistically follow different distributions. Such topics are mentioned in the literature survey, but do not fall into the scope of this dissertation.

  Additionally, the multi-input multi-output system model is covered. A technique known as the space-time coding is also discussed, which helps combat fading and increase the system throughput without exploiting additional frequencies.

- Chapter 3 (Block Relaying for Physical-Layer Network Coding): I study the scenario of the two-way relay channel where two terminals and a single relay take part in the communication process. There is no direct link available between the two terminals, and thus the delivery is achieved by way of the relay. Traditionally, the same modulation type is used throughout the whole process. The terminals and the relay only interact with symbols of the same regime. I propose a new transmission strategy by means of symbol combination. In other words, symbols

or constellations of a lower order can be accumulated from the first phase to form a higher-order symbol, and the transmission shifts into utilizing this new modulation as it continues to the next phase. In particular, I answer a series of questions, such as: Is this scheme beneficial compared to the traditional approach? Is it always such case? If not, under what conditions does the proposed protocol improve the system performance? The questions are addressed with theoretical analysis, and conclusions are validated by numerical simulations.

- Chapter 4 (Power Management of the Distributed Decode-and-Forward Protocol in the Parallel Relay Networks): I exploit the power distribution among the terminals and the relay in a system known as the parallel relay network. In other words, I look into how the power imbalance affects the system performance. As the whole communication process breaks down to two distinctive stages, theoretically I find that power compensation to the transmitting parties involved in the inferior (in terms of error probability) channel, reduces the overall error rate. This chapter sheds light on how to obtain the exact power allocation policy which yields the optimal performance in the sense of the end-to-end error probability.
- Chapter 5 (Permutation Modulation for Physical-Layer Network Coding): I exploit the novel application of a long ago introduced and widely used scheme, permutation modulations. As the idea does not involve highly complicated operations, it possesses many desirable features. The permutation modulation induced methods have been employed in areas such as source coding, channel coding, and in some cases analog-to-digital conversion techniques. I investigate permutation modulations for the purpose of modulation (in its conventional sense, so called modulations in some cases are actually source coding). I give detailed descriptions of the encoding and decoding procedures. A series of lemmas, a theorem and a corollary are proved on the system design, its performance compared to using conventional modulation types, and its optimality.

- Chapter 6 (On the Performance of Two-Way Relay Channels Using Space-Time Codes): Previous chapters only consider transceivers with a single antenna. In this chapter, the terminals and the relay are each equipped with two antennas, and the space-time coding, in particular, the Alamouti scheme is introduced into such systems. The Alamouti scheme helps combat fading, and reduces the error rate thereby improving the system throughput without exploring additional frequency bandwidth. I propose the versions of decode-and-forward protocols for such two-way relay channels consisting of dual-antenna transceivers. The encoding and decoding procedures are discussed in detail. Additionally, this chapter shows the advantages of the proposed protocols in comparison with the traditional approaches by means of numerical simulations.
- Chapter 7 (Conclusions): I summarize the discussed topics and contributions, and conclude this dissertation.

CHAPTER 2

THEORETICAL BACKGROUND AND SYSTEM MODEL

In this chapter, I discuss the physical-layer network coding technique in the context of the two-way relay channels. Additionally, a very important space-time coding technique, the Alamouti scheme, is briefly introduced here.

The organization of this chapter is as follows. Section 2.1 introduces the general concepts of the two-way communication systems, and covers several variants including cooperative networks and two-way relay networks. The coding schemes and relaying protocols are discussed in Sections 2.1.3 and 2.1.4, respectively. Section 2.2 introduces the space-time coding concepts, and in particular the Alamouti scheme. Finally, other related works are surveyed in Section 2.3.

2.1. Two-Way Communication Systems

Wireless communication has become a crucial part in the functioning of the society. In order for two parties to exchange information or messages, they need to reside within each other's reach, in other words, transmission range. Such type of communication is considered to be line-of-sight [9]. However, the light-of-sight condition is not always available. The transmission energy dissipates as the waves travel to the destination at the rate of approximately the fourth power of the distance. The signal strength when reaching the final destination needs to be above a certain threshold for the receiver to "pick up." This threshold is usually known as the receiver sensitivity. Also, there potentially exists other ongoing transmissions in the channel, and thermal noise from the electronic devices is always present. Thus, in order to maintain decent reception, one way is to generate high transmission power. As an alternative, in an environment where other transmitting parties coexist, collaboration is a possible alternative.

We can consider the "ecology" of communication entities as a network where each party is a node. The available transmission channel between each pair can be seen as a link connecting the two. A message from a source to a destination can be delivered in a multi-hop fashion. The source node chooses an in-between node to forward its message instead of direct transmission. This intermediate node serves as a relay, and it can choose other neighboring nodes as its own relay. In such a hop-by-hop fashion, energy can potentially be conserved as its dissipation is at the rate of an order of the fourth power of the transmission distance [10]. Such a communication chain can be simplified to a three-node model, the source, the destination and the relay. In some cases, such as with the employment of the network coding techniques, the roles of the source and the destination are embodied by the same nodes, and the message flow goes both directions - a two-way communication channel. For this reason, the source and the destination are also known as terminals in those scenarios.

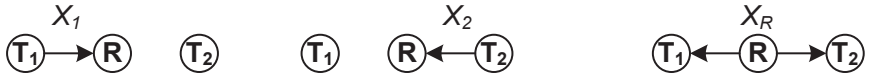
In [11], Shannon first considered a basic two-way communication model consisting of two terminals and investigated effective bi-directional communication in that scenario. Later on, some pioneer work was led and done by van der Meulen [12], Cover and El Gamal [13]. This topic has attracted significant attention recently [5].

Fig. 2.1 shows three different transmission strategies in the two-way relay networks. Without getting into more technical details on the coding techniques, relaying protocols and other issues, first I introduce the system model of a typical network setup and continue to cover certain frequently studied variants in the prior literatures. The system consists of two terminals denoted by T_1 and T_2, respectively. There is a relay R in between to assist the message exchange. Based on the availability of the links, the system has popular variants such as cooperative systems and two-way relay networks (here I only survey and discuss these two common types; there are other types).

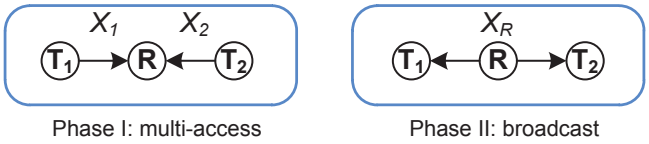
2.1.1. Cooperative Communication Systems

The first type is commonly known as cooperative communication networks [14, 15, 16, 17, 18]. In this setting, the source and the destination are distinguished as they do not transmit at the same time. Let's say T_1 is the source, and T_2 is the sink. Links $T_1 \rightarrow R$ and

(a) Traditional



(b) Network coding



(c) Physical-layer network coding

FIGURE 2.1. Improvements of the transmission strategies in the two-way relay channel: from (a) traditional to (b) network coding, and to (c) physical-layer network coding [1].

$T_1 \to T_2$ are both available. However the coefficients of the two channels are not necessarily subject to the same distribution. In fact, one research topic is to investigate the optimal transmission approach in terms of the power allocation between T_1 and R under a given allowance [19, 20].

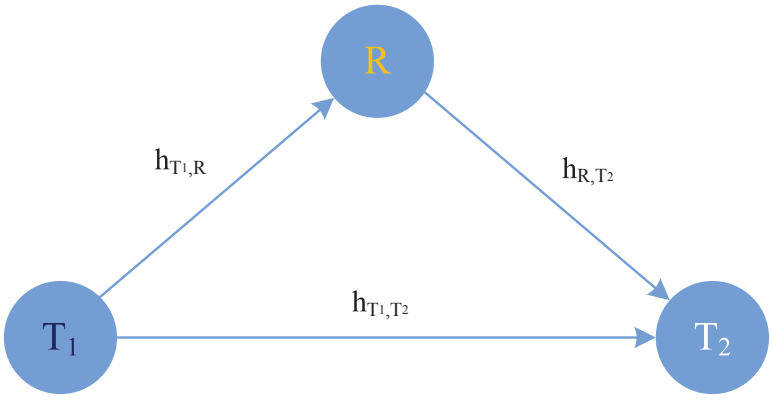


FIGURE 2.2. A simplified cooperation communication model.

Fig. 2.2 shows a simplified cooperative communication model. The basic idea of cooperative communication is that the source and the relay assist each other in sending messages to

the destination collaboratively [19]. In other words, the message originated from the source is repeated in some form by the relay, and reaches the destination besides the original copy. Multiple copies introduces a type of diversity known as the cooperative diversity, which contributes to more reliability of the system [20].

Laneman and Wornell developed cooperative diversity protocols using space-time codes to combat multipath fading [14]. In particular the authors proposed a suite of protocols in which distributed terminals act collectively to average the fading effects, and they demonstrated such protocols achieve full spatial diversity and outperform the repetition-based schemes [15]. The authors described relaying strategies including fixed scheme such as amplify-and-forward, and decode-and-forward (I introduce these concepts shortly in Section 2.1.4), together with selection schemes adapting according to the channel estimations [16]. Also the outage probability of the system was thoroughly studied in [16]. Sendonaris et al. considered such systems in a context of cooperation among the mobile users, and described the new form of diversity in [17], and discussed implementation issues, analyzed its performance and confirmed its advantage in increasing the capacity and improving the system robustness [18]. Su et al. derived the exact closed form symbol-error-rate (SER) performance for the amplify-and-forward protocol, and tight (asymptotically sufficient approximation) upper bound of SER for the decode-and-forward, and further obtained the optimum allocation policy between the source and the relay under a power limit [20]. The authors found in [20] that an equal power allocation is generally not the optimum in cooperative systems though often is also good. In [21], Ahmadzadeh el al. designed a novel constellation using the concept of signal space diversity [22] to help the destination to more efficiently combine the copies sent from the source and the relay.

2.1.2. Two-Way Relay Channel

For the two-way relay channel, the link $T_1 \rightarrow T_2$ is assumed to be unavailable. Thus the transmission takes place solely relying on the relay's assistance. Depending on the scheduling between the two terminals, the relay receives one message from either source at a time in the flat point-to-point system or a system using the conventional network coding techniques; or

else it receives simultaneously transmitted messages, and applies the physical-layer network coding (PLNC). The topic of scheduling is covered shortly in Section 2.1.3. After the relay R receives a message or a combination of two messages, perturbed by the noise, there are several commonly used methods, such as amplify-and-decode (AF), and decode-and-forward (DF), for processing before the next stage. Other similar methods which have been proposed by prior literatures include but not limited to partial-decode-and-forward (PDF) [23], estimate-and-forward (EF) [24, 25, 26], mutual information based forward (MIF) [27]. These protocols are discussed in Section 2.1.4.

2.1.3. Coding Techniques

Without employing any coding techniques, traditionally, the transmission takes place in the form of sending one uncoded symbol/signal at a time, as shown in Fig. 2.1(a). Without loss of generality, let's say T_1 first starts its transmission. T_1 sends a symbol X_1, which is modulated by such as BPSK, QPSK, 16QAM, etc., to the relay R. During the next time slot, T_2 sends its symbol X_2. Then R forwards X_1 and X_2 to their intended destinations respectively using two additional time slots. Thus, a single round of two-way exchange is completed in a total of four time slots. With the application of certain coding techniques, the time consumption can be potentially reduced [28].

2.1.3.1. *Network coding.* Fig. 2.1(b) depicts the approach of the conventional network coding scheme. The relay receives the symbols in the same fashion as in Fig. 2.1(a) where T_1 and T_2 alternate in transmission. For the next phase, instead of forwarding the symbols separately, R can perform network coding operations on these symbols. Let's take the BPSK modulation as an example, say the relay receives and decodes $X_1 = 1, X_2 = 0$. Here for simplicity, we consider zero noise. The network coding is performed by computing bitwise XOR on the two symbols/bits, specifically $X_1 \oplus X_2 = 1 \oplus 0 = 1$. For other combinations, there is $0 \oplus 1 = 1, 0 \oplus 0 = 0, 1 \oplus 1 = 0$. Then a coded new symbol X_R is obtained. The relay utilizes only a single time slot to broadcast this X_R to T_1 and T_2 at the same time. The total time consumption is thus reduced. At the receiver side, again with the same example

$X_1 = 1, X_2 = 0, X_R = 1$, T_1 is able to decode the symbol transmitted by T_2 with its own by computing $X_1 \oplus X_R = 1 \oplus 1 = 0 = X_2$.

2.1.3.2. *Physical-layer network coding.* Using the additive nature of the electromagnetic waves, the network coding can be performed in the physical layer, which is known as the physical-layer network coding [29]. T_1 and T_2 transmit their symbols simultaneously to the relay R, as in Fig. 2.1(c). This stage is usually called multi-access (MA) phase. Zhang et al. researched the synchronization issues in [30, 31]. Here the synchronization in the symbol level and carrier phase level and power control are assumed such that the symbols arrive at the relay with the same phase and amplitude [32]. The processing here is more complex than that in Section 2.1.3.1. I discuss this in the context of AF and DF protocols in Section 2.1.3.3. The basic idea is to obtain the combination (sum) of the two symbols rather than individual symbols explicitly [33]. After obtaining X_R, the relay broadcasts it back to the terminals. The terminals decode the symbols with their own. This second stage is called the broadcast or BC phase.

2.1.3.3. *Other coding schemes and extensions to the physical-layer network coding.* There are other coding techniques proposed and discussed in existing papers. Zhang et al. proposed a precise definition of the PLNC and classified it into two categories based on whether the network-code field is finite or infinite [34]. Katti et al. proposed an approach similar to the PLNC, called analog network coding, which does not assume symbol-level and carrier phase synchronization [35]. In [35], the authors demonstrated the practicality of their algorithm by implementation in software defined radios, and showed empirically its advantage in increasing the throughput in comparison to the traditional wireless routing and the digital network coding. Liu et al. presented a new scheme for the asymmetric channels named superimposed XOR based on both XOR in the bit level and superposition coding in the symbol level [36]. The authors derived its achievable rate region, average maximum sum-rate and service delay performance, and showed their proposed scheme outperforms the existing bitwise XOR and symbol-level superposition coding in the two-way relay networks. To and Choi studied

the application of convolutional codes in conjunction with the PLNC, and proposed a low complexity decoding approach using reduced-state trellis of the Viterbi algorithm [37].

2.1.4. Relaying Protocols

2.1.4.1. *Amplify-and-forward.* In the AF protocol, the relay does not attempt to decode the individual symbols or a combination of the symbols that were transmitted from the terminals. As mentioned in the previous discussion, due to the energy dissipation as the electromagnetic waves travel from the source to their destinations, the power of the received signal is smaller than the terminal's transmission power. The relay can be a transceiver of either the same capabilities, or different. In either case, the relay can scale the received signals to its own transmission power, practically amplifying them. Note that, during this process, since the noise is not isolated from the original signals, it is also amplified, and therefore the efficiency of energy use is suboptimal.

2.1.4.2. *Decode-and-forward.* As for DF protocol, the relay performs decoding operations on the received signals. Though it is possible to do exhaustive search among all the constellations to decode the individual symbols from either terminal, a more efficient way is to obtain the combination of the signals. In a Gaussian channel when no fading effects are preset or considered, this combination is simply a sum of two signals [32]. After decoding, the new symbol is forwarded to the terminals in the BC phase. In essence, the decode-and-forward attempts to process and eliminate the noise and fading corruption of the MA phase. However, should error occur in this process, it propagates to the next-hop destination, and continues with the propagation if in a larger network consisting of multiple intermediate nodes.

2.1.4.3. *Other forwarding protocols.* Cui and Kliewer analyzed and optimized the existing AF and DF, and proposed a new suite of absolute value based AF and DF [25]. The authors also proposed an estimate-and-forward inspired by the ideas of both AF and DF, and furthermore derived an optimal relay function by integrating all advantages of AF, DF and EF [38]. Karim et al. proposed a soft forwarding scheme based on the mutual information for a parallel relay network with two terminals and potential multiple relays

in between [27]. Its forwarding function considers both the decision of the symbol and the reliability of this decision measured by the symbol-wise mutual information. The authors showed in [27] their proposed scheme yields signal-to-noise ratio gains compared to existing schemes. Based on their paper, I plotted the performances of the mutual information based forwarding in comparison to the DF and EF protocols in Fig. 2.3.

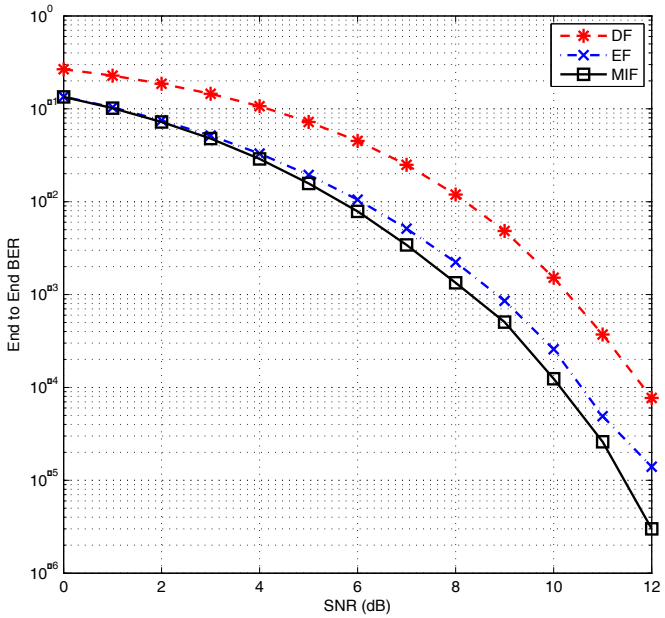


FIGURE 2.3. The end-to-end BER performance comparison of detect-and-forward (DF), estimate-and-forward (EF), and MI based forward (MIF) in a parallel realy network with two relays.

2.2. Space-Time Coding

Performed in both spatial and temporal domains, the space-time coding is a technique that can be used for multiple transmit antennas to introduce correlations between signals transmitted by distinct antennas and at different times [39]. It can achieve the transmit diversity and power gain over their uncoded counterparts without exploiting additional bandwidth. Tarokh et al. introduced the space-time block codes for communication in a MIMO system [40]. In particular, the authors applied the classical mathematical framework of orthogonal designs, discussed the Maximum-Likelihood based decoding algorithm, and demonstrated these codes maximize the achievable transmission rate, and proved the

optimality of many code instances in the sense of trade off between the decoding delay and the number of transmit antennas [40]. This diversity scheme requires the channel state information available to the receivers for the purpose of decoding. Tarokh and Jafarkhani presented in [41, 42] a differential detection method to be used in a system where neither the transmitters nor the receivers have access to the channel information, and discussed the encoding and decoding operations in detail.

2.2.1. Alamouti Scheme

Alamouti described a two-transmitter coding scheme with one or more receivers in [43]. His approach is schematically shown in Fig. 2.4, an example with a single receive antenna for simplicity.

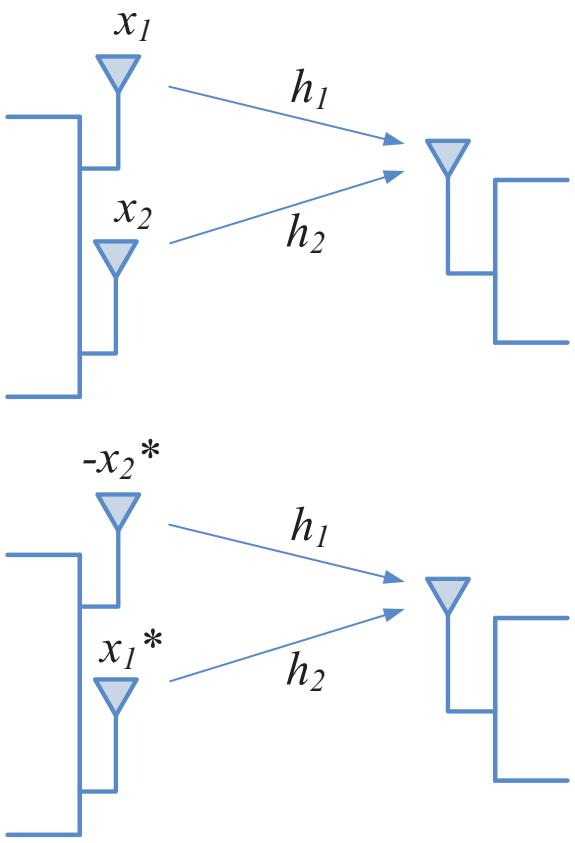


FIGURE 2.4. The two transmit antenna diversity scheme.

The communication is completed in two steps. In the first slot, a signal x_1 is transmitted from the first antenna, at the same time when x_2 is being transmitted by the second antenna.

At the receiver, the received signal can be expressed as

$$y_1 = h_1 x_1 + h_2 x_2 + n_0, \tag{1}$$

where h_1, h_2 are the channel coefficients, and n_0 is the additive white Gaussian noise. Then the next slot, $-x_2^*$ is sent from the first antenna, whereas x_1^* from the second. Assuming the channel states do not change within these two slots (slow fading), the received signal for this slot can be formulated as

$$y_2 = -h_1 x_2^* + h_2 x_1^* + n_1, \tag{2}$$

where again n_1 is the noise at this time instance. The channel information is assumed to be known to the receiver or can be obtained by exchanging pilot symbols. If ideally no noise is present in the system, Eqs. (1)(2) can be solved directly for the two unknowns x_1, x_2. While the thermal noise is always present in the electronic devices, a decoding method based on the Maximum Likelihood criterion can be adopted, which yields the optimal performance given the same a priori probabilities of all the constellations used. The variables for detection can be written as

$$\begin{cases} \tilde{x_1} = h_1^* y_1 + h_2 y_2^*, \\ \tilde{x_2} = h_2^* y_1 - h_1 y_2^*. \end{cases} \tag{3}$$

If we substitute y_1, y_2 with Eqs. (1)(2) respectively, the relations of these detection variables with the original symbols are seen more clearly as below,

$$\begin{cases} \tilde{x_1} = (h_1^2 + h_2^2) x_1 + h_1^* n_0 + h_2 n_1^*, \\ \tilde{x_2} = (h_1^2 + h_2^2) x_2 - h_1^* n_1^* + h_2^* n_0. \end{cases} \tag{4}$$

As mentioned previously, the error probability and furthermore the throughput is improved without sacrificing more frequency bandwidth. Additionally, the decoding in this scheme is efficient and of low complexity. Thus the Alamouti scheme is widely studied, extended and utilized into various system scenarios.

2.3. Other Related Works

The closed-form expressions of the outage probability, maximum sum-rate, and sum-BER of practical PLNC schemes were derived in [44]. Lu et al. investigated the throughput capacity of the PLNC, and provided capacity bounds in comparison with the existing transmission schemes [45]. The authors investigated the capacity of relay networks using multiple antennas in [46, 47]. Nam et al. considered the capacity bound of the use of nested lattice codes in a channel where nodes operate in full-duplex mode [48], and specifically gave the capacity of each node and their sum-rate [49]. Wilson et al. obtained the capacity upper bound of a joint physical-layer network-layer code, which proves advantageous compared to the existing analog network coding schemes at the time of their publication [50].

In [51], the authors addressed the security issue of the PLNC by means of the error probability of a potential eavesdropper, and demonstrated by extensive simulated evaluations that the PLNC provides means against passive eavesdropping.

Ji et al. combined the Alamouti scheme with the PLNC, and derived the approximation and upper bound of the overall outage probability based on the effective end-to-end SNR, and analyzed the gain and diversity order of their proposed scheme [52].

In [53], Wu et al. studied the double space-time diversity system where four transmit antennas are divided into two groups and each group constructs its own orthogonal space time transmit diversity. The exact analytical SER expressions were obtained, and numerical results by simulations confirmed and validated the theoretical results [54]. Zhang et al studied PLNC in the context of a large-scale random wireless network using a generalized physical model, and improved the network capacity bounds by a factor of greater than 1 [55].

Chen et al. combined the transmit antenna selection and maximal-ratio combining in a MIMO system [56]. The authors used the tools of the joint distribution of two or more order statistics [57], and contributed rigorous mathematical modeling and derivation of the closed form error probability performance of such systems [58], and investigated the impact of the antenna selection [59].

There are other related works besides those mentioned previously. They are discussed in the background/introduction section of each chapter as they are in particular closely related to those topics.

2.4. Summary

In this chapter, I covered the preliminaries of the two-way relay channels and the physical-layer network coding. Additionally, some related variants of the network and extensions to the coding scheme were surveyed. The discussion here laid theoretical foundation for the chapters that follow, and provided background for the proposed novel schemes and mathematical modeling therein.

CHAPTER 3

BLOCK RELAYING FOR PHYSICAL-LAYER NETWORK CODING

In this chapter, I study the transmission strategy in a two-way relay channel. In particular, I propose a new class of $\log_{M_1}(M_2)$-MA (multi-access) block relaying protocols where an M_1-ary modulation is used in the MA phase, and an M_2-ary in the broadcast phase. Lower-order (e.g., BPSK) and higher-order (e.g., 16-QAM, 32-PSK) modulations are used respectively for the multi-access phase and the broadcast phase, in contrast to the conventional method of using the same modulation type (of either higher or lower order) throughout. To achieve this, the relay buffers the received symbols from consecutive multi-access phases and performs signal combining in conjunction with the decode-and-forward protocol. I derive the theoretical performance bounds for the Rayleigh fading channels and closed forms for the AWGN channels. The optimal power allocation between the multi-access and broadcast phases is also discussed. Numerical results confirm the advantages of the proposed scheme with an improved throughput in lower signal-to-ratio range.

The organization of the chapter is as follows. Section 3.1 surveys prior works related to this chapter's topic, and gives the motivation behind my work. Section 3.2 gives a detailed description of the system model. In Section 3.3, the theoretical performance in terms of throughput is discussed, for both fading and Gaussian channels. Numerical results are presented in Section 3.4, followed by the conclusion in Section 3.5.

The content in this chapter with the exception of Sections 3.2.1 and Section 3.4.1 is reproduced from N. Xu, and S. Fu, "Block Relaying For Physical-Layer Network Coding," IEEE Journal of Communications (under review), with permission from IEEE.

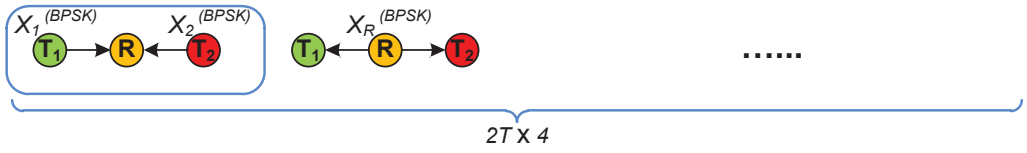
3.1. Background

The topic of the two-way relay channels (TWRC) has attracted significant attention recently. In a TWRC system, terminals as transceiving entities, exchange information through the relay. To achieve the spectral efficiency [60], terminals send symbols to the relay simultaneously first, and then the relay broadcasts the processed symbols back to the terminals. By coordinating the terminals for simultaneous transmission and using the additive nature of the electromagnetic waves, one can achieve equivalent network coding operations at the physical layer[32]. This technique is the physical layer network coding (PLNC). Instead of treating the received symbols from other terminals as interference, PLNC turns it into a capacity boost.

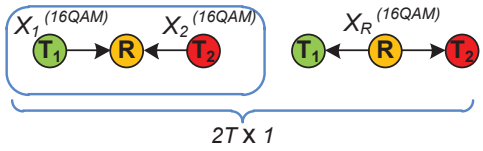
To my best knowledge, in most existing literatures, the same modulation is assumed throughout all stages of the PLNC transmission. There are several recent works on improving the throughput by using more elaborate modulations [36, 61, 62, 63, 26, 64, 65]. Liu et al. discussed the superimposed XOR, a hybrid of the bitwise XOR and the symbol-level superposition coding, for asymmetric channels [36]. It addressed the issue of applying PLNC in a system where the terminals transmit information at different rates. In [61, 62], Koike-Akino et al. investigated a modulation optimizing scheme by way of sphere packing on the denoising map of the relay. [63] proposed non-uniform constellations to lower the complexity of PLNC with high-order PAMs by using binary codes. [26] investigated the constellation design of the differential modulation for PLNC. It proves to be beneficial especially for systems with higher data rate requirements. Yun et al. presented a physical-layer retransmission scheme that applies coding on the packet level to reduce retransmission cost [64]. This approach piggybacks a new packet on a retransmitted packet by using higher modulations, and recovers both by exploiting previously received packets. A novel mapping codebook-based physical network coding scheme is proposed in [65] for asymmetric two-way relay channels where source nodes exchange data in different flow rates. A mapping codebook contains several subcodebooks to be adaptively selected based on the signal phase difference to improve the system performance.

My work tackles the problem from a different angle by exploiting the use of constellations of different orders during different phases. I propose using higher-order constellations (e.g., 16-QAM, 32-PSK) during the broadcast stage while lower-order (e.g., BPSK) during the multi-access stage and investigate its impact on the overall throughput.

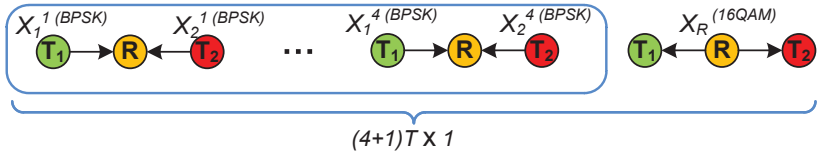
3.2. System Model



(a) Conventional method using BPSK only



(b) Conventional method using 16-QAM only



(c) $\log_2$ 16-MA

FIGURE 3.1. The conventional and $\log_{M_1}(M_2)$-MA block relaying schemes for the two-way relay systems (blocks are indicated by round corner rectangles, (a) (b) can be seen as special cases of $\log_{M_1}(M_2)$-MA for $M_1 = M_2$, namely $\log_2$ 2-MA, $\log_{16}$ 16-MA respectively).

In this section, I describe in detail the proposed $\log_{M_1}(M_2)$-MA (multi-access) block relaying scheme, where an M_1-ary modulation is used in the MA phase, and an M_2-ary in the broadcast phase. To start off with an example, Fig. 3.1 shows three relaying protocols in a two-way relay system, where packets of $L = 4$ (I use 4 here for simplicity while a typical packet size can be 128 or more) bits are exchanged. Fig. 3.1(a) carries out the transmission bit by bit using the BPSK modulation. Utilizing the spectral efficient method [60], a single

round of two-way communication can be accommodated in two time slots. Two terminals ($\mathbf{T_1}, \mathbf{T_2}$) utilize exactly one time slot to transmit their symbols (X_1, X_2) to the relay ($\mathbf{R}$) simultaneously. This is typically known as the MA phase. The next time slot is dedicated for the relay to broadcast a new symbol (X_R) back to $\mathbf{T_1}$ and $\mathbf{T_2}$. This is the broadcast or BC phase. The approach of obtaining X_R is discussed later in this section. Denoting the duration of a single time slot as T, the total time to exchange two packets is $2T \times 4 = 8T$.

Now in order to improve the throughput (I come back to its formal definition in a moment) or in another way, to have more packets correctly received at the receiver side within a certain period of time, an attempt is made by using a higher-order modulation in Fig. 3.1(b). In this approach, the four bits inside a packet can be encapsulated into a 16-QAM symbol to be transmitted at the same time, and the total transmission time is shortened to $2T$. However, this only becomes beneficial with sufficiently high signal-to-noise ratios (SNRs), as it is known from the analyses in previous works, such as [66, 67], that higher-order modulations statistically incur more errors than the lower-order under the same SNR constraint. Especially, this disadvantage results in excessive retransmissions in the low SNR range, which potentially cancels out the advantage of the shorter transmission time per packet.

Fig. 3.1(c) gives an example of my proposed method $\log_{M_1}(M_2)$-MA where $M_1 = 2, M_2 = 16$ in this case. BPSK is used in the MA phase. BPSK symbol pairs are transmitted to the relay in $\log_2 16 = 4$ consecutive slots. The relay combines them into a 16-QAM symbol (X_R) and broadcasts it back to $\mathbf{T_1}, \mathbf{T_2}$ using one slot. The transmission time totals $4T + T = 5T$. This approach shortens the per-packet transmission time compared to Fig. 3.1(a), and achieves a better error rate compared to Fig. 3.1(b). It proves to be the best of the three in terms of throughput in the lower SNR range. I get into the details in Section 3.3.

Having illustrated some examples, I now give the generalized description of the proposed $\log_{M_1}(M_2)$-MA scheme. There are two terminals and a single relay in the system. No direct link is available between the terminals, and packets are exchanged via the assistance of the relay. The communication is broken down into MA and BC phases where M_1-PSK/QAM is used in the MA phase and M_2-PSK/QAM in the BC phase. M_2 should be a power

of M_1. The size of a packet is L bits. To simplify the discussion, I enforce that L is a multiple of $\log_{M_1}(M_2)$, such that, a packet can be exactly divided into symbols of the M_2-ary modulation, and of course M_1-ary too. As for the throughput, which is the primary performance metric of interest, I use a definition similar to that in [68]. The throughput of the system is defined as the size of a packet over the average time to correctly transmit and receive it, assuming an ideal retransmission. [68] also defined a nominal data rate as the packet size over the duration in which a packet can be accommodated. This way, the throughputs are normalized. Given that I consider the modulation types in a lower level than packets, the duration of a packet varies for different schemes. This is where my approach differs. I employ the data rate of a system using BPSK as the benchmark to normalize the throughputs of all investigated systems. Denote the packet duration as T_s, the nominal data rate (R_s) is expressed as $R_s \triangleq \frac{L}{T_s}$. Continuing with the example in Fig. 1(c), the transmission (by either terminal) takes place with R_s rate in the MA phase and with a rate of $4R_s$ for the BC phase using 16-QAM.

There are $\log_{M_1}(M_2)$ time slots of transmission in the MA phase associated with every BC slot. Every symbol in a BC slot is indexed by t ($1 \leq t \leq \frac{L}{\log_2(M_2)}$). To indicate the association, a symbol in an MA slot is designated by two indices t and m ($1 \leq m \leq \log_{M_1}(M_2)$). The MA phase transmission can be formulated as

$$U[m,t] = h_1[m,t]\sqrt{\mathbf{P_1}}X_1[m,t] + h_2[m,t]\sqrt{\mathbf{P_2}}X_2[m,t] + n_R[m,t], \tag{5}$$

where $U[m,t]$ is the received symbol, $X_k[m,t]$ is the transmitted symbol from $\mathbf{T_k}$, $\mathbf{P_k}$ is the transmission power of $\mathbf{T_k}$, $h_k[m,t]$ is the channel coefficient of link $\mathbf{T_k} \rightarrow \mathbf{R}$, and $n_R[m,t]$ is the additive white Gaussian noise (AWGN) with zero mean and σ^2 variance. Note that when $h_1 = h_2 = 1$, the system is an AWGN channel.

During each transmission, the channel state information is assumed known to the receiver. Additionally, I assume that transmitted symbols are backlogged at $\mathbf{T_1}$, $\mathbf{T_2}$ and $\mathbf{R}$ respectively and can be fetched whenever polled. For simplicity, I only consider BPSK in the MA phase

here. The discussion hereafter can also be extended to combining other types of modulation generally.

3.2.1. Amplify-and-Forward

I investigate two relaying strategies, namely amplify-and-forward (AF) and decode-and-forward (DF). Each symbol transmitted in the BC slot is denoted as $X_R[t]$ ($1 \leq t \leq \frac{L}{\log_2 M}$), which is obtained as follows. In the AF protocol, for M-PSK,

$$X_R[t] = \sum_{m=1}^{\log_2 M} U[m-1,t]e^{j\frac{m\pi}{\log_2 M}}, \tag{6}$$

and for the square M-QAM,

$$X_R[t] = \sum_{m=1}^{\log_2 M} [m-1+\mathrm{sgn}(m \mod 2)]U[m,t]e^{\frac{\mathrm{sgn}(m \mod 2)\pi}{2}}, \tag{7}$$

where sgn is the sign function and ($x \mod y$) performs the (x modulo y) operation. To meet the power constraint, before it is sent out, X_R is scaled by α_{AF} as given in Eq.(8).

$$\alpha_{AF} = \begin{cases} \sqrt{\frac{P_S}{\mathbf{P_U} log_2 M}} & , M\text{-PSK} \\ \sqrt{\frac{P_S}{2\mathbf{P_U}\sum_{m=1}^{\frac{\sqrt{M}}{2}}(2m-1)^2}} & , \text{square } M\text{-QAM} \end{cases} \tag{8}$$

where $\mathbf{P_U} = \mathbf{P_1} + \mathbf{P_2} + N_0$ (N_0 is the noise power spectral density).

The received symbol at $\mathbf{T_k}$ at the end of the t-th BC slot is

$$Y_k[t] = h_k[t]\alpha_{AF}\sqrt{\mathbf{P_R}}X_R[t] + n_k[t]. \tag{9}$$

where $\mathbf{P_R}$ is the transmission power of $\mathbf{R}$. The AF protocol requires all CSI available at the terminals in order to cancel the self-interference during the decoding.

3.2.2. Decode-and-Forward

As for the DF protocol, the PLNC technique is employed for processing at the relay $\mathbf{R}$. Instead of decoding $X_1[m,t]$, $X_2[m,t]$ individually, I obtain the combination $X_1[m,t] + X_2[m,t]$, and map it back to an information bit. For more detailed discussions on the PLNC, I recommend [60, 32] for references. After obtaining $\log_{M_1}(M_2)$ consecutive bits from the pairs of $X_1[1,t] + X_2[1,t]$, $\cdots$, $X_1[\log_{M_1}(M_2),t] + X_2[\log_{M_1}(M_2),t]$, $X_R[t]$ is formed by

combining them bitwisely into an M_2-ary symbol. In the BC slot that follows, $\mathbf{R}$ broadcasts $X_R[t]$ to $\mathbf{T_1}$ and $\mathbf{T_2}$. Once again, take the $\log_2 16$-MA in Fig. 3.1(c) as an example. As in [32], $X_1[m,t] + X_2[m,t]$ can be -2, 0, 2 for BPSK. -2 and 2 are mapped to 0 while 0 is mapped to 1. Suppose, after the decoding of $\log_{M_1}(M_2) = \log_2 16 = 4$ consecutive MA slots, I have $X_1[m,t] + X_2[m,t]$ being -2, 2, 0, 0 respectively. The information bits are 0, 0, 1, 1. Thus I have the symbol 0011 with the m-th bit taken from the m-th slot. I modulate it using 16-QAM, and proceed to the BC phase transmission.

The received symbol at $\mathbf{T_k}$ at the end of the t-th BC slot is

$$Y_k[t] = h_k[t]\sqrt{\mathbf{P_R}}X_R[t] + n_k[t]. \tag{10}$$

In DF protocol, eventually, the m-th bit from the other terminal can be obtained using $X_k[t]$ and the m-th bit of the decoded $Y_k[t]$.

3.2.3. Assumptions on the Power Allocation

For a fair comparison, I impose a per-packet power constraint. To be specific, I allow a constant total power for the terminals and the relay altogether to transmit a packet from end to end. This total power per packet can be written as $\frac{3}{2}L\mathbf{P}$. Note that, every round of MA and BC communication, two packets are exchanged. Here I use $\frac{3}{2}$ instead of 1 to avoid introducing too many fractions in the analysis. If equally allocated, the power of $\mathbf{T_1}, \mathbf{T_2}$, and $\mathbf{R}$ is $L\mathbf{P}$. Then the power for per-bit transmission is $\mathbf{P}$. For example, in Fig. 1(c), BPSK is used in the MA phase, and its symbol (only one bit) has a power of $\mathbf{P}$. As for the BC phase, 16-QAM is used, and its symbol (four bits) power is $4\mathbf{P}$. I discuss the power allocation policy in detail in Section 3.3. An equal allocation is assumed hereafter unless explicitly specified otherwise.

3.3. Performance Analysis

I choose the normalized end-to-end throughput as the primary performance metric for the proposed $\log_{M_1}(M_2)$-MA block relaying scheme. Previously, I investigated the performance of the PLNC, in particular when using BPSK modulations [5, 69, 6]. Here, I focus on the case of $M_1 = 2$ when BPSK is used in the MA phase and higher-order modulations

in the BC phase. Whenever not causing confusion, I ignore the subscript of M_2 and use $\log_2 M$ instead. In this section, the theoretical performance bounds are derived for the Rayleigh fading channels and the exact closed forms are obtained for the AWGN channels. Furthermore, I discuss the optimal power allocation between the terminals and the relay under a global power constraint so as to maximize the overall throughput.

Denote P_p, R_p as the end-to-end packet error rate (PER) and the end-to-end throughput, respectively. As mentioned in Section 3.2, the throughput is the size of a packet over the average time of correctly transmitting and receiving it, assuming ideal retransmissions whenever an error occurs [68]. [68] also defined a nominal data rate to normalize the throughput. I use a slightly different approach here, as the data rates in the investigated systems are modulation-dependent. Instead, I use the rate of BPSK as the nominal data rate, which is denoted as $R_s \triangleq L/T_s$ (again T_s is the packet duration using BPSK). The packet duration using M-ary modulations is $\frac{T_s}{\log_2 M}$. With ideal retransmission, the average time to transmit a packet from $\mathbf{T_1}$ to $\mathbf{T_2}$ using $\log_2 M$ is $\frac{T_s+T_s/\log_2 M}{1-P_p} = \frac{\log_2 M+1}{\log_2 M} \cdot \frac{T_s}{1-P_p}$ [68]. The same is true for the link $\mathbf{T_2} \rightarrow \mathbf{T_1}$ due to the system symmetry. Therefore, the throughput can be written as

$$
\begin{aligned}
R_p &= 2L/\left(\frac{\log_2 M+1}{\log_2 M} \cdot \frac{T_s}{1-P_p}\right) \\
&= \frac{L}{T_s}(1-P_p)\frac{2\log_2 M}{\log_2 M+1} \\
&= R_s(1-P_p)\frac{2\log_2 M}{\log_2 M+1}. \qquad (11)
\end{aligned}
$$

Note that, I use $2L$ rather than L in Eq. (11) because transmissions $\mathbf{T_1} \rightarrow \mathbf{T_2}$ and $\mathbf{T_2} \rightarrow \mathbf{T_1}$ take place simultaneously. Now I can express the throughput R_p in terms of the BPSK data rate R_s. As the system is symmetric, and derivation for either side is sufficient. Without loss of generality, I consider $\mathbf{T_2} \rightarrow \mathbf{T_1}$. I only need to derive P_p, and R_p can be computed through Eq. (11).

3.3.1. Rayleigh Fading Channels

I start with the Rayleigh fading channels. As mentioned in Section 3.2, there are $\log_2 M$-MA slots associated with every BC slot. Also, a packet can be exactly divided into M-ary modulated symbols. The PER can be expressed in terms of the symbol error rate (SER). Furthermore, in Gray-coded M-PSK, the bit error rate (BER) follows the independent and identical distribution (i.i.d.) regardless of the bit position [67, 70]. For the square QAM, the BER can be computed given the bit position [66]. Thus, I can further break down PER and express it through the BERs of its every bit in the MA and BC phases.

Following Eq. (5), I use t ($1 \leq t \leq \frac{L}{\log_2 M}$) as the index of a BC slot, and for a MA slot associated with it, t and m ($1 \leq m \leq \log_2 M$). Given that a bit of a BC symbol is obtained from decoding the combination of two corresponding MA symbols, I also use m to index the bits in a BC symbol. Let $P_{ma}[m,t]$ be the BER of the m,t-designated MA slot, and $P_{bc}[m,t]$ be the BER of the m-th bit of the t-th received symbol in the BC phase. Then I have

$$P_p = 1 - \prod_{t=1}^{\frac{L}{\log_2 M}} \prod_{m=1}^{\log_2 M} \left[P_{ma}[m,t]P_{bc}[m,t] + (1 - P_{ma}[m,t])\,(1 - P_{bc}[m,t])\right]. \tag{12}$$

For all t's, $P_{ma}[m,t]$ is subject to i.i.d., and so is $P_{bc}[m,t]$. I can omit the parameter t and simplify Eq. (12) to

$$P_p = 1 - \prod_{m=1}^{\log_2 M} \left[P_{ma}[m]P_{bc}[m] + (1 - P_{ma}[m])\,(1 - P_{bc}[m])\right]^{\frac{L}{\log_2 M}}. \tag{13}$$

All $P_{ma}[m]$ follows the i.i.d.. As my previous works [5, 6] indicate, $P_{ma}[m]$ can be tightly upper bounded by $\hat{P}_{ma}$ given in

$$\hat{P}_{ma} = \frac{1}{2(\gamma_{ma} + 1)}, \tag{14}$$

where γ_{ma} is the SNR per symbol. For BPSK, $\gamma_{ma} = \mathbf{P}/N_0$ (N_0 is the noise spectral density). As mentioned in Section 3.2, $\mathbf{T_1}, \mathbf{T_2}$, and $\mathbf{R}$ each has a power of an $L\mathbf{P}$ for L-bit packet. In Eq. (13),

$$P_{ma}[m]P_{bc}[m] + (1 - P_{ma}[m])\,(1 - P_{bc}[m])$$

$$
\begin{aligned}
&= 1 - P_{bc}[m] - (1 - 2P_{bc}[m])P_{ma}[m] \\
&\leq 1 - P_{bc}[m] - (1 - 2P_{bc}[m])\hat{P}_{ma} \\
&= \hat{P}_{ma}P_{bc}[m] + \left(1 - \hat{P}_{ma}\right)(1 - P_{bc}[m]), \qquad (15)
\end{aligned}
$$

as $1 - 2P_{bc}[m]$ is always non-negative. Then the upper bound of P_p, denoted as $\hat{P}_p$, can be expressed as

$$
\hat{P}_p = 1 - \prod_{m=1}^{\log_2 M} \left[\hat{P}_{ma}P_{bc}[m] + \left(1 - \hat{P}_{ma}\right)(1 - P_{bc}[m])\right]^{\frac{L}{\log_2 M}}. \qquad (16)
$$

For the square M-QAM modulations, [66] derived the error probability for AWGN channels. For fading channels, the conditional error probability of the BC phase given the fading amplitude r can be computed as

$$
\begin{aligned}
P_{bc}[m|r] = \frac{1}{\sqrt{M}} \sum_{i=0}^{(1-2^{-m})\sqrt{M}-1} &\left\{(-1)^{\left\lfloor \frac{i \cdot 2^{m-1}}{\sqrt{M}} \right\rfloor}(2^{m-1} - \left\lfloor \frac{i \cdot 2^{m-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor) \right. \\
&\left. \cdot \operatorname{erfc}\left((2i+1)\sqrt{\frac{3\gamma_{bc}r^2}{2(M-1)}}\right)\right\}, \qquad (17)
\end{aligned}
$$

where $\gamma_{bc} = \mathbf{P}\log_2 M/N_0$ per symbol and $\operatorname{erfc}(\cdot)$ is the complementary error function. Given the probability distribution function of $h_1[m,t]$ being $P_{h_1}(r) = 2re^{-r^2}$ $(r > 0)$, $P_{bc}[m]$ can be obtained by integrating $P_{bc}[m|r]$ with regard to r,

$$
P_{bc}[m] = \int_0^\infty P_{bc}[m|r]P_{h_1}(r)dr. \qquad (18)
$$

With the help of [71, Eq. (6.283)] and the variable substitution, I rewrite Eq. (18) as

$$
\begin{aligned}
P_{bc}[m] = \frac{1}{\sqrt{M}} \sum_{i=0}^{(1-2^{-m})\sqrt{M}-1} &\left\{(-1)^{\left\lfloor \frac{i \cdot 2^{m-1}}{\sqrt{M}} \right\rfloor}\left(2^{m-1} - \left\lfloor \frac{i \cdot 2^{m-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor\right) \right. \\
&\left. \cdot \left(1 - \frac{1}{\sqrt{1 + \frac{2(M-1)}{3(2i+1)^2\gamma_{bc}}}}\right)\right\}. \qquad (19)
\end{aligned}
$$

Then $\hat{P}_p$ can be computed by Eq. (16). Furthermore, R_p can be lower bounded as

$$
R_p \geq \check{R}_p = R_s(1 - \hat{P}_p)\frac{2\log_2 M}{\log_2 M + 1}. \qquad (20)
$$

As for Gray-coded M-PSK, the average $P_{bc}[m]$ is identical regardless of the bit position m and can be expressed as [67]

$$P_{bc}[m] = P_{bc} = \frac{1}{\log_2 M} \sum_{k=1}^{M-1} \left(2 \left| \frac{k}{M} - \left\lfloor \frac{k}{M} \right\rceil \right| + 2 \sum_{i=2}^{\log_2 M} \left| \frac{k}{2^i} - \left\lfloor \frac{k}{2^i} \right\rceil \right| \right) p_b(k), \tag{21}$$

where

$$p_b(k) = \begin{cases} 1 - \mathrm{I}(\frac{\pi}{M}), & k = 0, \\ \frac{1}{2} \left(\mathrm{I}(\frac{(2k-1)\pi}{M}) - \mathrm{I}(\frac{(2k+1)\pi}{M}) \right), & k = 1, \ldots, \frac{M}{2} - 1, \\ \mathrm{I}\left(\frac{(M-1)\pi}{M} \right), & k = \frac{M}{2} \end{cases} \tag{22}$$

and $\lfloor * \rceil$ is the operation of rounding to the closest integer, and $p_b(k) = p_b(M-k)$ for $k = \frac{M}{2} + 1, \ldots, M-1$. $\mathrm{I}(\cdot)$ is defined as

$$\mathrm{I}(\theta) = 1 - \frac{\theta}{\pi} - \frac{\frac{1}{2} + \frac{1}{\pi} \arctan \frac{\cot \theta}{\sqrt{1 + \frac{1}{\gamma_{bc} \sin \theta}}}}{\sqrt{1 + \frac{1}{\gamma_{bc} \sin \theta}}}. \tag{23}$$

Thus Eq. (16) becomes

$$\hat{P}_p = 1 - \left[\hat{P}_{ma} P_{bc} + (1 - \hat{P}_{ma})(1 - P_{bc}) \right]^L. \tag{24}$$

$\check{R}_p$ can also be derived for Gray-coded M-PSK accordingly.

3.3.2. AWGN Channels

The derivation for the AWGN channels follows that for the fading channels, by setting all the channel coefficients to 1. The BER of any MA slot is [69]

$$P_{ma} = \frac{3}{2\pi} \int_0^{\frac{\pi}{2}} \exp(-\frac{\gamma_{ma}}{\sin^2 \theta}) d\theta - \frac{1}{2\pi} \int_0^{\frac{\pi}{2}} \exp(-\frac{9\gamma_{ma}}{\sin^2 \theta}) d\theta. \tag{25}$$

For the BC phase, I also consider the square M-QAM and M-PSK modulations. In the square M-QAM, $P_{bc}[m]$ of an AWGN channel equals Eq. (17) by setting $r = 1$ [70]. Given the MA slots being identical and the symmetry along the two dimensions of the square M-QAM constellation, from Eq. (12) I have

$$P_p = 1 - \prod_{m=1}^{\log_2 M} \left[P_{ma} P_{bc}[m] + (1 - P_{ma}) (1 - P_{bc}[m]) \right]^{\frac{L}{\log_2 M}}. \tag{26}$$

As for M-PSK using Gray coding, $P_{bc}[m]$ or P_{bc} takes the same form as Eq. (21) with a different definition of $p_b(k)$ [72]. Now

$$p_b(k) = \frac{1}{2\pi}\left(\int_0^{\pi(1-\frac{2k-1}{M})} \exp(-\frac{\gamma_{bc}\sin^2\frac{(2k-1)\pi}{M}}{\sin^2\theta})d\theta - \int_0^{\pi(1-\frac{2k+1}{M})} \exp(-\frac{\gamma_{bc}\sin^2\frac{(2k+1)\pi}{M}}{\sin^2\theta})d\theta\right), \tag{27}$$

regardless of the bit position m. Thus, Eq. (12) is simplified to

$$P_p = 1 - \left[P_{ma}P_{bc} + (1-P_{ma})(1-P_{bc})\right]^L. \tag{28}$$

Computation of the overall throughput R_p is performed using Eq. (11).

3.3.3. Beneficial SNR Range

Two factors affect the end-to-end throughput, as shown from Eq. (11), namely the PER and the modulation type M. Continuing with the previous section, I still only consider BPSK in the MA phase, and general modulation types in the BC phase. $\frac{\log_2 M}{\log_2 M+1}$ increases as M grows larger. However, for higher-order modulations, their error performance is comparatively poorer, resulting in a worse end-to-end PER. For smaller SNRs, the error probabilities are rather large for high-order modulations, which is not compensated by the gain from $\frac{\log_2 M}{\log_2 M+1}$. Thus, lower-order modulations prove to be beneficial than the higher-order in the low SNR range. As the SNR becomes sufficiently large, few errors occur in the transmission, and the PER approaches zero. Then the gain from $\frac{\log_2 M}{\log_2 M+1}$ dominates in Eq. (11). Asymptotically, the higher-order modulations are advantageous.

Substitute the P_p in Eq. (11) with Eq. (26), and I have for the AWGN channels

$$R_p = 2R_s \cdot \frac{\log_2 M}{\log_2 M + 1} \prod_{m=1}^{\log_2 M} \left[P_{ma}P_{bc}[m] + (1-P_{ma})(1-P_{bc}[m])\right]^{\frac{L}{\log_2 M}}. \tag{29}$$

For the fading channels, it is $\hat{P}_{ma}$ instead of P_{ma}. Denote the factor other than $2R_s$ as $\mathcal{F}(M)$. For two systems using M'-ary and M''-ary modulations respectively for their BC phases, the "inflection" SNR value (dividing the regions of one scheme's beneficial and not compared to the other) can be obtained by solving $\mathcal{F}(M') = \mathcal{F}(M'')$. As this equation does not generally

have a tractable analytical solution, I do not go into the details of solving a particular pair of schemes numerically.

3.3.4. Optimal Power Allocation

As mentioned in Section 3.2, the total power consumed by $\mathbf{T_1}, \mathbf{T_2}$ and $\mathbf{R}$ altogether to transmit a packet from end to end is constantly $\frac{3}{2}L\mathbf{P}$. Transmissions via links $\mathbf{T_1} \to \mathbf{T_2}$ and $\mathbf{T_2} \to \mathbf{T_1}$ take place simultaneous. The total power is then $3L\mathbf{P}$ per packet, and equivalently $3\mathbf{P}$ per bit. Up till this point, the system setting has been an equal allocation policy. The power of the BPSK symbols in the MA phase are $\mathbf{P_1} = \mathbf{P_2} = \mathbf{P}$ whereas for the M-ary modulations in the BC phase, $\mathbf{P_R} = \mathbf{P}\log_2 M$. I still consider a symmetric system where $\mathbf{P_1} = \mathbf{P_2}$. More generally, let β be a ratio ($0 < \beta < 3/2$), I allocate $\beta\mathbf{P}$ per bit to $\mathbf{T_1}, \mathbf{T_2}$, and $(3-2\beta)\mathbf{P}$ per bit to $\mathbf{R}$. Then $\mathbf{P_1} = \mathbf{P_2} = \beta\mathbf{P}$, $\mathbf{P_R} = (3-2\beta)\mathbf{P}\log_2 M$.

My goal is to find such a ratio β in order to maximize the end-to-end throughput, or to minimize the end-to-end PER. Let $\mathcal{G}^M_{qam}(\beta) = \prod_{m=1}^{\log_2 M}\left[P_{ma}P_{bc}[m] + (1-P_{ma})\,(1-P_{bc}[m])\right]$ and $\mathcal{G}^M_{psk}(\beta) = P_{ma}P_{bc} + (1-P_{ma})\,(1-P_{bc})$. Also $\hat{\mathcal{G}}^M_{qam}(\beta), \hat{\mathcal{G}}^M_{psk}(\beta)$ for their counterparts with P_{ma} replaced by $\hat{P}_{ma}$. It can be seen from Eq. (16) that, for a given M, $\hat{\mathcal{G}}^M_{qam}(\beta)$ reaches its maximum when $\hat{P}_p$ reaches its minimum varying β. Same observations can be made for $\hat{\mathcal{G}}^M_{psk}(\beta)$ and $\hat{P}_p$ from Eq. (24), $\mathcal{G}^M_{psk}(\beta)$ and P_p from Eq. (28), and also $\mathcal{G}^M_{qam}(\beta)$ and P_p from Eq. (26). Then substituting $\gamma_{ma} = \beta\mathbf{P}/N_0, \gamma_{bc} = (3-2\beta)\mathbf{P}\log_2 M/N_0$ into Eqs. (16) (24) (26) and (28), respectively, the problem can be formulated as solving $\operatorname{argmax}_\beta \hat{G}^M_{psk}(\beta)$, $\operatorname{argmax}_\beta \hat{G}^M_{qam}(\beta)$, $\operatorname{argmax}_\beta G^M_{psk}(\beta)$ or $\operatorname{argmax}_\beta G^M_{qam}(\beta)$ depending on the modulation type and channel condition (fading or AWGN). To obtain $\operatorname{argmax}_\beta \hat{G}^M_*(\beta)$ ($*$ represents either psk or qam), I derive the partial derivative of $\hat{G}^M_*(\beta)$ with regard to β, and solve the equation $\partial\hat{G}^M_*(\beta)/\partial\beta = 0$. It is the same way with $\operatorname{argmax}_\beta G^M_*(\beta)$.

For example, I look at $\log_2 4$-MA (QPSK) in the Rayleigh fading channel,

$$\hat{\mathcal{G}}^2_{psk} = \hat{P}_{ma}P_{bc} + (1-\hat{P}_{ma})\,(1-P_{bc})\,, \tag{30}$$

where

$$\hat{P}_{ma} = \frac{1}{2(\beta\mathbf{P}/N_0+1)}, \quad \text{(rf. Eq. (14))},$$

$$P_{bc} = \frac{1}{2}\left(1-\sqrt{\frac{2(3-2\beta)\mathbf{P}/N_0}{2(3-2\beta)\mathbf{P}/N_0+1}}\right), \quad \text{(rf. Eq. (21))}. \tag{31}$$

Solving $\partial\hat{\mathcal{G}}^2_{psk}/\partial\beta = 0$ and choosing the valid value ($0 < \beta < 3/2$), I have the optimal $\beta_{opt} = 1$. Based on the results from [5, 6, 67], $\hat{P}_{ma}$ is close to P_{bc} under the same SNR. Therefore, when $\beta = 1$, an equal allocation ensures the best overall performance.

However, for higher-order modulations, it is not generally mathematically tractable. For them, I do not seek to explicitly give analytical solutions. The optimal β values can always be calculated using numerical methods. As an example, the numerically computed β_{opt} values for $\log_2 16$-MA using 16-QAM are summarized in Table 3.1. For $\log_2 16$-MA using 16-QAM, P_{bc} is worse than P_{ma} for the same SNR, due to the use of higher-order constellations. As the BC phase is the bottleneck of the system, more power is allocated to $\mathbf{R}$. For example, at SNR = 8 dB, $3-2\beta_{opt} = 1.5224$, i.e., $\mathbf{R}$ takes almost half of the $3\mathbf{P}$ per bit. Also, the error performance of the MA phase improves faster than that of the BC phase as the SNR increases. Thus, for higher SNRs, the portion $(3-2\beta_{opt})$ is also larger.

TABLE 3.1. β_{opt} for $\log_2 M$-MA (16-QAM) in the AWGN channels.

SNR (dB)	0	2	4	6	8	10	12
β_{opt}	0.9384	0.8799	0.8211	0.7739	0.7388	0.7140	0.6973
$3-2\beta_{opt}$	1.1232	1.2402	1.3578	1.4522	1.5224	1.5720	1.6054

3.4. Numerical Results

In this section, I present the numerical results. The end-to-end throughput is the primary metric of the system performances. It is measured in terms of the nominal data rate R_s of the BPSK modulation. The throughput values are plotted against SNR which is defined as $\mathbf{P}/N_0$.

For the notation $\mathbf{P}$ (per-bit transmission power), here I briefly revisit some assumptions of the transmission power in Section 3.2. I assume $3L\mathbf{P}$ for the exchange of two packets (two packets travel via links $\mathbf{T_1} \to \mathbf{T_2}$, $\mathbf{T_2} \to \mathbf{T_1}$ simultaneously). L is the packet length. Thus, the power allocation per bit for the terminals $\mathbf{T_1}, \mathbf{T_2}$ and the relay $\mathbf{R}$ together is $3\mathbf{P}$. Take Fig. 3.1 ($L = 4$) as an example. Using BPSK throughout both phases (Fig. 3.1(a)), the packet is transmitted bit by bit, and each BPSK symbol (bit) has a power of $\mathbf{P}$. It takes eight time slots (four for MA, BC respectively) for the exchange of two packets. The total power consumed is $(2\mathbf{P} + \mathbf{P}) \times 4 = 12\mathbf{P} = 3L\mathbf{P}$. As for using 16-QAM only, as in Fig. 3.1(b), the power of each symbol (four bits) is $4\mathbf{P}$, and the transmission is completed in two time slots. The power consumption is $(4\mathbf{P} + 4\mathbf{P}) + 4\mathbf{P} = 12\mathbf{P} = 3L\mathbf{P}$. Similarly, for the $\log_2 16$-MA scheme (Fig. 3.1(c)), BPSK is used in the MA phase ($\mathbf{P}$ per symbol, four time slots), and 16-QAM for the BC phase ($4\mathbf{P}$ per symbol, one slot). Therefore, the total power consumption is $(\mathbf{P} + \mathbf{P}) \times 4 + 4\mathbf{P} \times 1 = 12\mathbf{P} = 3L\mathbf{P}$. From the above analysis, for all these three schemes, the transmission power of the terminal/relay is $\mathbf{P}$ per bit.

3.4.1. Amplify-and-Forward

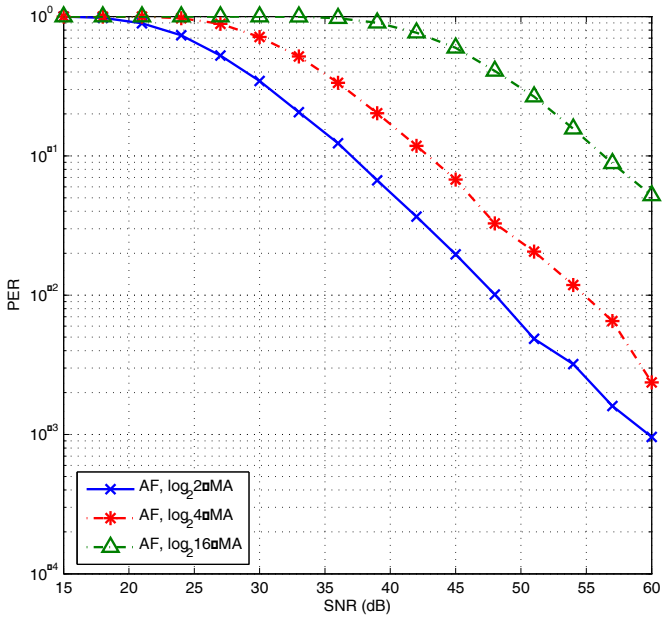


FIGURE 3.2. The end-to-end PER performance of AF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in fading channels when $L = 128$.

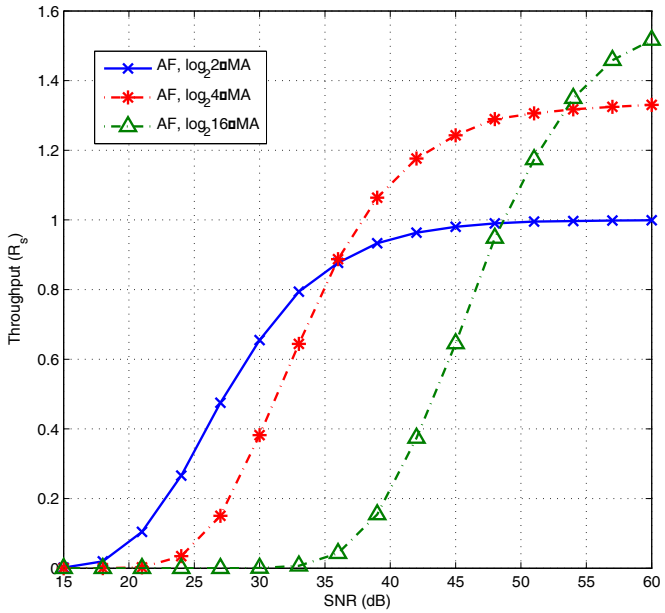


FIGURE 3.3. The end-to-end throughput performance of AF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in fading channels when $L = 128$.

Fig. 3.2 shows the end-to-end PER comparison among three schemes, whereas Fig. 3.3 displays that for the end-to-end throughput. It is seen that AF benefits from $log_2 M$-MA ($M > 2$) in the high SNR range in terms of the overall throughput. Raising the modulation to a higher order results in an increase of the overall error rate. However, due to the signal combining and scaling at the relay, as the order of the constellations grows higher, this advantage diminishes.

3.4.2. Decode-and-Forward

Figs. 3.4, 3.5 compare the simulation results with the upper bounds obtained for each $\log_2 M$-MA scheme in fading channels, respectively. QPSK (4-QAM) is used for $\log_2 4$-MA and 16-QAM for $\log_2 16$-MA. As seen from the figure, the theoretical bounds merge tightly with the numerical values. Thus, the bound performance suffices for asymptotic analysis in place of the exact closed forms. The same pattern can be observed here that modulations of higher-order tend to have a better asymptotic throughput while proving less efficient in low SNRs.

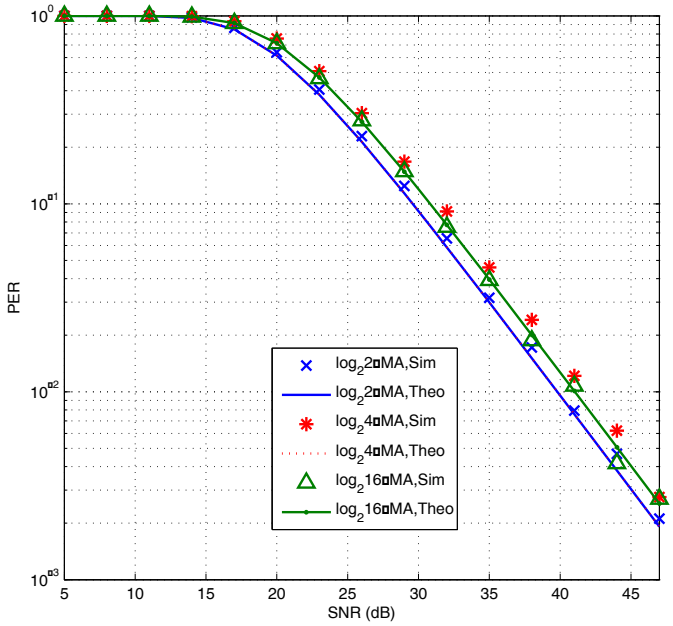


FIGURE 3.4. The end-to-end PER performance of DF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in fading channels when $L = 128$: simulation results versus theoretical values.

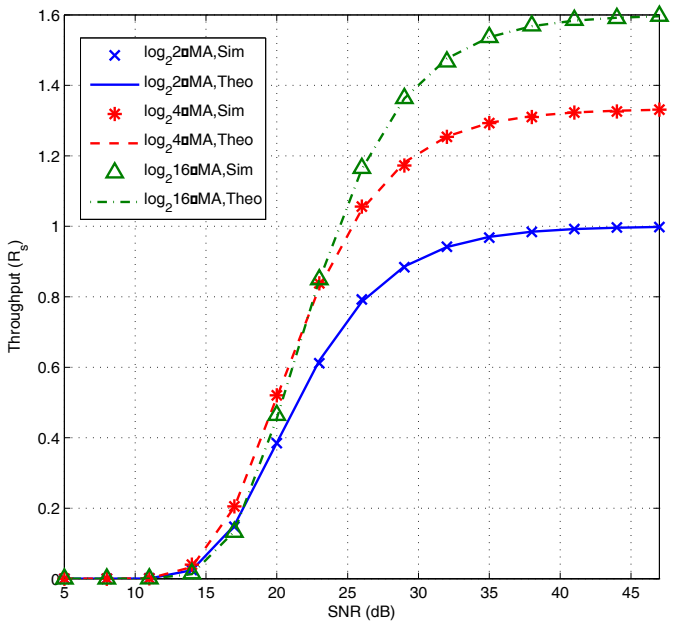


FIGURE 3.5. The end-to-end throughput performance of DF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in fading channels when $L = 128$: simulation results versus theoretical values.

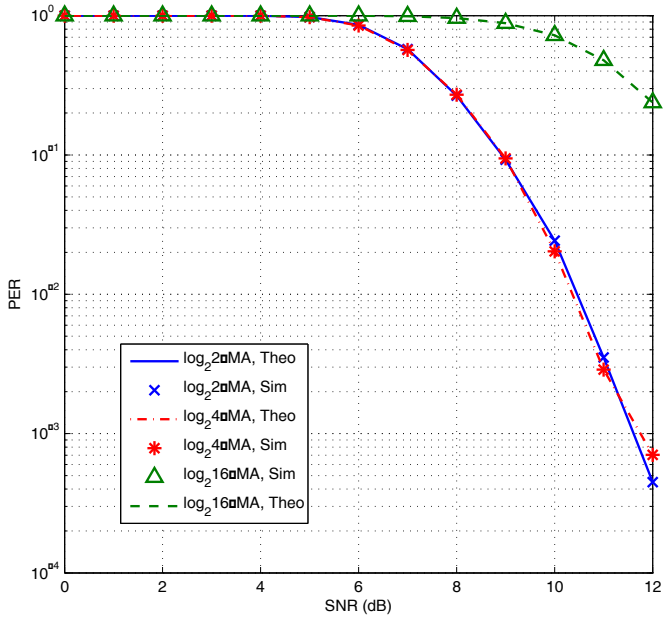


FIGURE 3.6. The end-to-end PER performance of DF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in AWGN channels when $L = 128$: simulation results versus theoretical values.

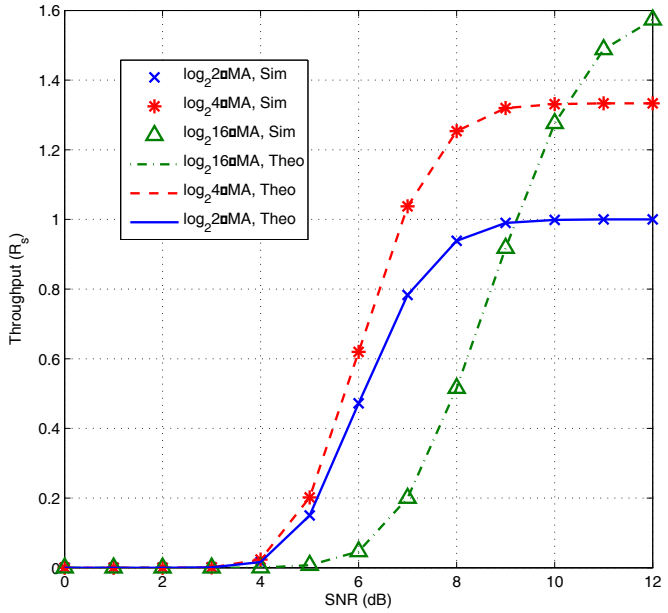


FIGURE 3.7. The end-to-end throughput performance of DF, $\log_2 M$-MA (QPSK/4-QAM for $M = 4$, 16-QAM for $M = 16$) in AWGN channels when $L = 128$: simulation results versus theoretical values.

Figs. 3.6, 3.7 perform the same comparison on the same three schemes in the AWGN channels. I see that the exact theoretical values are validated by the numerical results.

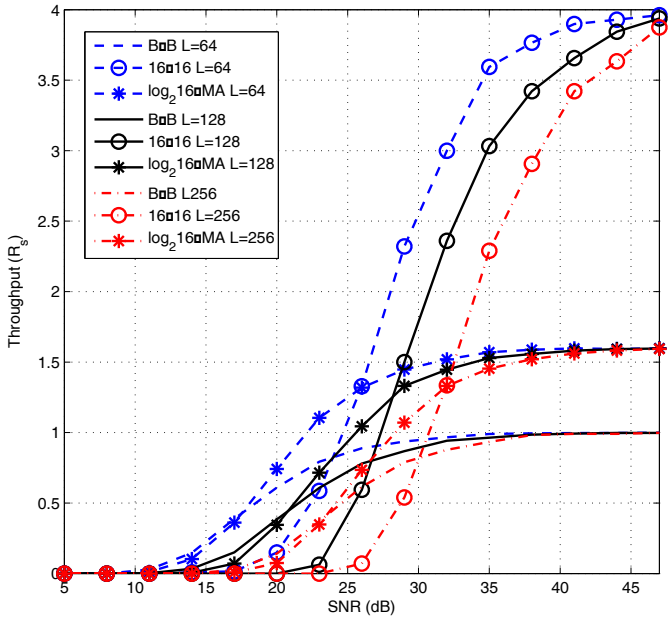


FIGURE 3.8. The end-to-end throughput comparison between $\log_2 16$-MA (16-PSK) and the conventional methods using BPSK ($B-B$ in the figure legend), 16-PSK ($16-16$) throughout in fading channels. The packet sizes L are 64, 128 and 256.

In Fig. 3.8, the throughput of the proposed $\log_2 16$-MA scheme using 16-PSK is plotted for packet lengths 64, 128 and 256. For comparison, I choose BPSK and 16-PSK as two extreme cases among the conventional methods (whose order not larger than 16). BPSK sees a lower error probability and longer transmission time per packet, whereas 16-PSK has a higher error rate and shorter time to accommodate a packet. My proposed $\log_2 16$-MA approach makes comprise between PER and the per-packet time consumption. This proves to be beneficial in the low SNR range. For example, when the packet size $L = 64$, from around 26 dB SNR and below, $\log_2 16$-MA surpasses $16-16$. As the SNR becomes smaller, the throughputs of all three schemes approach zero, and $B-B$ gains a very slim margin over $\log_2 16$-MA. Overall, $\log_2 16$-MA can be seen as the best scheme in the below 26 dB

SNR region. For even larger packet sizes, I see the right end of this beneficial region expands towards higher SNR; it is 28 dB when $L = 128$ and 32 dB when $L = 256$.

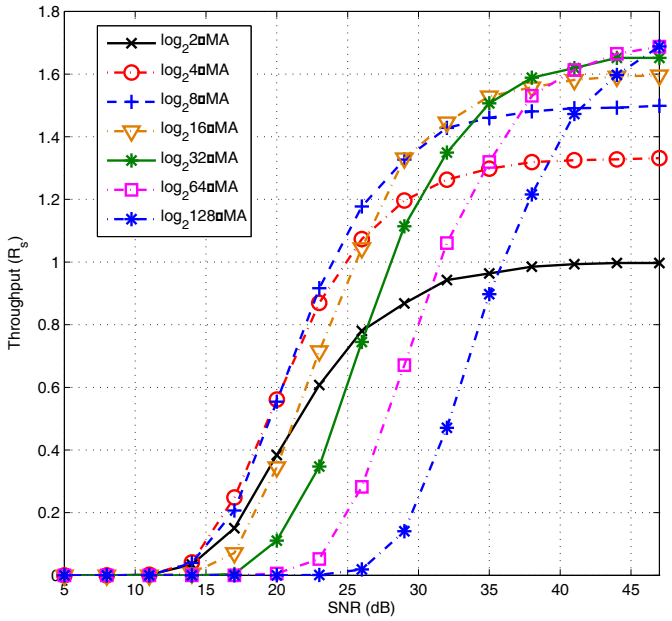


FIGURE 3.9. The end-to-end throughput comparison between $\log_2 M$-MA's for $M = 2, 4, 8, 16, 32, 64, 128$ (M-PSK) in fading channels under the packet length $L = 128$.

The throughputs of $\log_2 M$-MA schemes for $M = 2, 4, 8, 16, 32, 64, 128$ (M-PSK) with the packet size $L = 128$ are compared in Fig. 3.9. As the order of the modulation type used in the BC phase increases, the "inflection" SNR is higher where its throughput turns from disadvantageous to advantageous compared to a certain lower-order scheme. As discussed in Section 3.3.3, for low SNRs, the gain from $\frac{\log_2 M}{\log_2 M+1}$(see Eq. (11)) is not yet able to compensate the poor PER performance that high-order $\log_2 M$-MA's suffer from. For a larger constellation size M, it takes a higher SNR level for $(1 - P_p)$ to become the dominant factor. For example, $\log_2 16$-MA surpasses $\log_2 2$-MA in throughput after passing 20 dB, whereas $\log_2 64$-MA does only after reaching past 31 dB. The throughput difference between certain schemes are not visible in the figure, as it is rather small. Based on the results in my experiments, at SNR = 11 dB, $\log_2 2$-MA's throughput is 0.0020, higher than $\log_2 4$-MA's

0.0013, though by an almost negligible margin. The conclusion holds for those two schemes although their curves appear to overlap in the figure.

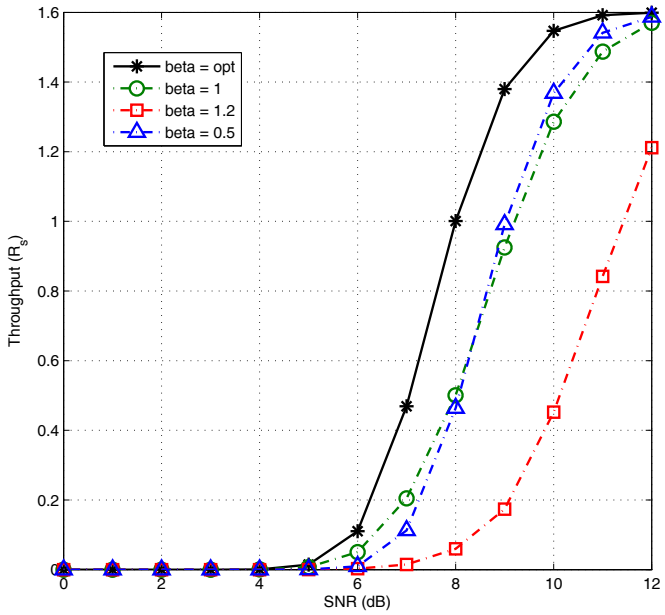


FIGURE 3.10. The end-to-end throughput comparison between schemes using the optimal power allocation policy and the non-optimal ($L = 128$): $\log_2 16$-MA (16-QAM) in AWGN channels.

For $\log_2 16$-MA (16-QAM) in AWGN channels, the throughput values obtained by using the optimal power allocation are shown in Fig. 3.10. The β_{opt} values are summarized in Table 3.1. Several non-optimal policies, namely $\beta = 0.5, 0.8, 1.2$ (constantly), are plotted for comparison. The optimality of the policy described in Table 3.1 is validated. As mentioned in Section 3.3.4, due to the error rate bottleneck of the BC phase using higher-order modulations, more power is allocated in compensation. Additionally, towards high SNRs, as the MA phase improves faster than the BC phase, the portion for the relay is even larger. Fig. 3.11 confirms that the equal allocation policy ($\beta = 1$) for $\log_2 4$-MA (QPSK) in fading channels is optimal, although by a small margin over the non-optimal.

3.5. Summary

In this chapter, I proposed a new class of $\log_{M_1} M_2$-MA block relay for TWRC systems. Instead of forwarding the received symbols for each MA phase, the relay buffers the symbols

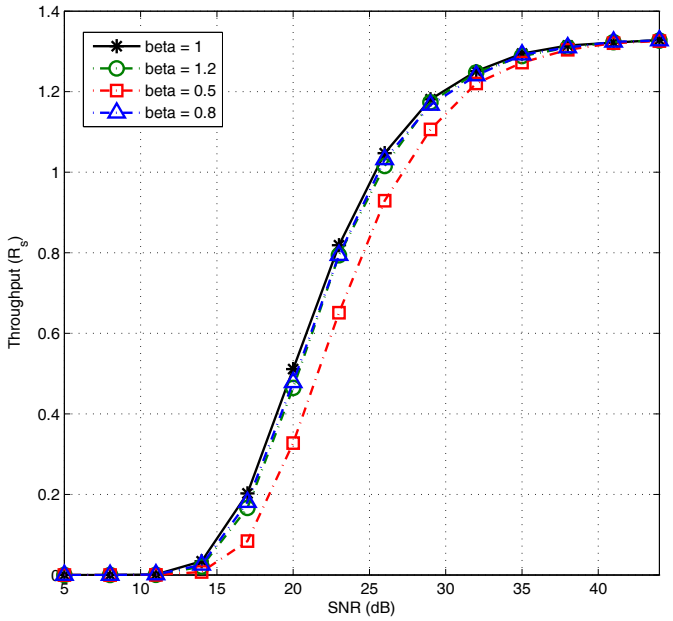


FIGURE 3.11. The end-to-end throughput comparison between schemes using the optimal power allocation policy and the non-optimal ($L = 128$): $\log_2 4$-MA (QPSK) in fading channels.

from multiple MA phases and forward the combination through a higher modulation. In particular, I focused on the case of $M_1 = 2$ using BPSK in the MA phase. I obtained the performance bounds for $\log_2 M$-MA block relaying in the fading channels and the exact closed forms for the AWGN channels. The beneficial SNR range for the proposed scheme, and the optimal power allocation strategy between the terminals and the relay were also discussed. Numerical results confirmed the advantages of the proposed scheme in terms of higher throughputs in low SNRs.

CHAPTER 4

POWER MANAGEMENT OF THE DISTRIBUTED DECODE-AND-FORWARD PROTOCOL IN THE PARALLEL RELAY NETWORKS

In this chapter, I study the power allocation polices to achieve the minimum error probability in a wireless fading parallel relay network with limited energy. In particular, the performance of the distributed Alamouti-coded decode-and-forward (ACDF) protocol is investigated. First I derive the exact closed-form symbol error rate (SER) of the protocol. Then in order to reveal the relationship between the error rate performance and the power allocation in a more mathematically simplistic fashion, I further obtain an upper bound and an approximation asymptotically tight at high signal-to-noise ratios (SNRs). With the assistance of these tools, I derive the goal function and its solution for the optimized power management under a global energy constraint. Two basic methods of fixed allocation are compared with the scheme with the optimized allocation. Numerical results showed the proposed power allocation helps achieve the best error performance.

The organization of the chapter is as follows. Section 4.1 gives the motivation behind the work in this chapter. Section 4.2 gives a detailed description of the overall communication system. In Section 4.3, I derive the closed-form end-to-end error probability of the ACDF scheme together with its upper bound and asymptotically tight approximation. Then follows the discussion on the optimized allocation policy minimizing the SER. The conclusions are drawn in Section 4.4.

4.1. Motivation

The power allocation serves as an effective means to improve the performance of an energy constrained relay network [73]. As wireless sensor networks draw significant attention from the research communities, there have been lots of work addressing power-related issues, primarily allocation strategies under certain global constraints. In [20][19], Su et al.

studied the optimum power allocation for the wireless cooperative network using amplify-and-forward (AF) and decode-and-forward (DF) protocols. Their results showed that equal power strategy is not optimum for cooperative communications in general. They discovered that optimum allocation only depends on the links related to the relay. In [73], the authors investigated the power control rule of the Alamouti-coded amplify-and-forward (ACAF) transmission protocol in a system where a single source-destination pair communicates over a layer of two parallel relay nodes. By analytical evaluation and simulation, they revealed the on-off nature of the transmission of each relay. Power control prevents a relay from amplifying highly noisy signals to the destination. Hasna and Alouini tackled a similar problem by splitting the communication link from the source to the destination into several hops [74][75]. One of the main advantages is that it distributes the use of power throughout the hops which prolongs the battery life and lowers the interference introduced to the rest of the network. Other related work such as [76][77] optimized the power control policy based on sum rate or outage probability. In [76], the authors proposed an optimal power allocation in the two-way relay channel employing the physical-layer network coding (PLNC) protocol by maximizing the achievable sum-rate under a sum-power constraint in a Rayleigh fading channel environment. [77] presented two power allocation schemes for bidirectional AF relaying over the Rayleigh fading channels through the exploitation of the channel mean strength. Previous work such as [78, 79, 80] studied some different aspects of the distributed Alamouti code. [78] did not investigate the channel conditions and modulation types, and simply relied on the observation on the pairwise error probability. [79] studied the distributed AF scheme using PAM or QAM modulations based on the assumption of the locations of the relays in reference to the destination. [80] only discussed the error performance, and the obtained closed forms were not suitable for deriving power control strategies. My work refines this research in the power management by exploring various parameters and obtained more tractable mathematical solutions.

In this chapter, I study power allocation polices to achieve the minimum error probability in a fading parallel relay channel with energy constraints. In particular, the power allocation

strategy to boost the performance of the distributed Alamouti-coded [43] decode-and-forward protocol (ACDF) is investigated. To my best knowledge, this topic has not been discussed in the literature.

4.2. System Model

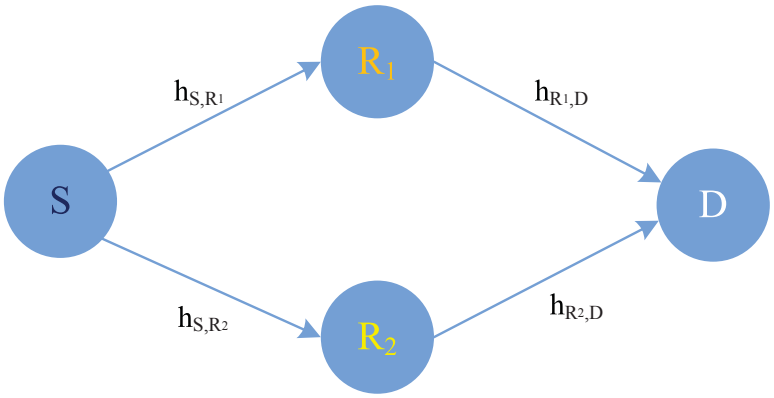


FIGURE 4.1. The fading parallel relay channel model.

In this section, I give a detailed description of the system. Fig. 4.1 shows a simplified schematic of a fading parallel relay channel. The source S is to send information to the destination D via two relays R_1 and R_2, as no direct communication link is available. Consider two consecutive transmissions of symbols as one block (BL), then the whole process of symbol transferring from S to D can be broken down to two blocks, namely $S \to R_1(R_2)$ and $R_1(R_2) \to D$, where each block consists of two time slots.

The transmission power of S, R_1, R_2 is denoted as P_S, P_{R_1} and P_{R_2}, respectively, which is subject to $P_S + P_{R_1} + P_{R_2} \leq P$ where P is the total available power.

During the first block BL-I, S broadcasts two symbols $X_S(1), X_S(2)$ to the relays. The slow Rayleigh fading is assumed such that within each block the channel condition does not vary. Yet this assumption is only crucial for BL-II. The received symbols at R_i can be written as

$$U_{R_i}(t) = h_{S,R_i}\sqrt{P_S}X_S(t) + n_{R_i}(t), \tag{32}$$

where h_{S,R_i} represents the fading channel coefficient of link $S \to R_i$, $n_{R_i}(t)$ is the additive white Gaussian noise (AWGN) with zero mean and N_0 variance, $t = 1, 2$, and $i = 1, 2$. For simplicity, I assume that all noises share the same variance. The channel coefficients are modeled as $h_{*,*} \sim \mathcal{CN}(0, \sigma^2_{*,*})$ where $\mathcal{CN}(0, \sigma^2_{*,*})$ is the complex Gaussian distribution with zero mean and $\sigma^2_{*,*}$ variance for links $S \to R_1, S \to R_2, R_1 \to D$, and $R_2 \to D$, respectively.

In each transmission, the channel state information (CSI), together with the transmission power allocation, is assumed to be available at the receiver. In practice, CSI can be estimated by exchanging and examining pilot symbols [81], which is beyond the scope of this chapter.

The decoding is performed individually at each relay based on the maximum likelihood (ML) criteria. If both relays decode the symbols correctly, BL-II transmission takes place using a distributed Alamouti space-time code [43]. When a single relay alone decodes the symbols successfully, this relay transmits the decoded symbols during BL-II transmission, whereas the erroneous relay does not. As for the worst scenario when both fail to successfully decode the symbols, neither proceed to the next phase. For the analytical tractability, I assume an ideal DF protocol in which it is known whether the symbols are correctly decoded. Also, I do not consider retransmission, such that transmission skips BL-II to the next round once decoding error occurs.

Denote the symbols to transmit by R_i as $X_{R_i}(t)$. In BL-II, the encoding and transmission sequence for each relay is shown in Table 4.1, which is an Alamouti scheme applied in a distributed fashion. As previously mentioned, the relay only transmits during BL-II when

TABLE 4.1. The encoding and transmission sequence for each relay.

	R_1	R_2
$t = 1$	$X_{R_1}(1)$	$X_{R_2}(2)$
$t = 2$	$-X^*_{R_1}(2)$	$X^*_{R_2}(1)$

it obtains the correctly decoded symbols. The values for each entry in Table 4.1 can be

summarized as

$$X_{R_i}(t) = \begin{cases} X_S(t) & \text{, correct decoding,} \\ 0 & \text{, o.w.} \end{cases} \tag{33}$$

where $i, t = 1, 2$. It still follows the scheme described in Table 4.1 when only a single relay participates in BL-II transmission. In this way, a synchronization mechanism is not needed, and the distributed collaboration between relays is achieved.

Using the distributed Alamouti code, the received symbols at the destination are expressed as

$$\begin{bmatrix} Y_D(1) \\ Y_D(2) \end{bmatrix} = \begin{bmatrix} h_{R_1,D}\sqrt{P_{R_1}}X_{R_1}(1) + h_{R_2,D}\sqrt{P_{R_2}}X_{R_2}(2) + n_D(1) \\ -h_{R_1,D}\sqrt{P_{R_1}}X^*_{R_1}(2) + h_{R_2,D}\sqrt{P_{R_2}}X^*_{R_2}(1) + n_D(2) \end{bmatrix}. \tag{34}$$

When a single relay alone takes part in BL-II, e.g., $Y_D(2)$ receives $-h_{R_1,D}\sqrt{P_{R_1}}X^*_{R_1}(2)+n_D(2)$ only from R_1 or $h_{R_2,D}\sqrt{P_{R_2}}X^*_{R_2}(1) + n_D(2)$ only from R_2. Eq. (34) holds given Eq. (33). When both relays commit errors, they skip the BL-II transmission, and only noise is present at the channel. Therefore, Eq. (34) can be used as a universal expression of the received symbols for all scenarios.

To detect the Alamouti coded symbols, the combiner builds the detection variable as:

$$\begin{aligned} \tilde{X}(1) &= \alpha_1^* Y_D(1) + \alpha_2 Y_D(2)^*, \\ \tilde{X}(2) &= \alpha_2^* Y_D(1) - \alpha_1 Y_D(2)^*, \end{aligned} \tag{35}$$

where

$$\alpha_i = \frac{h_{R_i,D}\sqrt{P_{R_i}}}{N_0}. \tag{36}$$

As for the reception from a single relay, the detection symbols are specified as

$$\begin{aligned} \tilde{X}(1) &= \alpha_1^* Y_D(1), \\ \tilde{X}(2) &= -\alpha_1 Y_D(2)^*, \end{aligned} \tag{37}$$

for R_1 only, and

$$\tilde{X}(1) = \alpha_2 Y_D(2)^*,$$

$$\tilde{X}(2) = \alpha_2^* Y_D(1), \tag{38}$$

for R_2 only. No detection is performed when the BL-II transmission is skipped by both relays. At this point, the original transmitted symbols can be decoded by the ML detector at the destination.

4.3. Performance Analysis and Numerical Results

In this section, I first derive the closed-form end-to-end error probability of the ACDF scheme. Then its upper bound and asymptotically tight approximation are obtained as mathematically simplified tools to reveal the relationship between the power allocation and the error rate performance.

4.3.1. Closed-Form SER Performance of ACDF

I start the derivation from Eq. (32). The effective signal-to-noise ratio (SNR) of the detection variables after BL-I can be formulated as [82]

$$\gamma_{I,i} = \frac{P_S |h_{S,R_i}|^2}{N_0}, \text{ for } i = 1, 2. \tag{39}$$

Here due to the limited space, I only consider M-PSK modulation. With the effective SNR, the conditional SER given the channel coefficients can be written as [83]

$$P_M^{h_{S,R_i}} = \frac{1}{\pi} \int_0^{(M-1)\pi/M} \exp(-\frac{g_M \gamma_{I,i}}{\sin^2 \theta}) d\theta, \tag{40}$$

where $i = 1, 2$, and $g_M = \sin^2(\pi/M)$. The error probability of a single time slot of BL-I can be computed as

$$P_{I,i} = \int_{-\infty}^{\infty} P_M^{h_{S,R_i}} dh_{S,R_i} = \mathcal{F}(1 + \frac{g_M P_S \sigma_{S,R_i}^2}{N_0 \sin^2 \theta}), \tag{41}$$

where

$$\mathcal{F}(f(\theta)) = \frac{1}{\pi} \int_0^{(M-1)\pi/M} \frac{1}{f(\theta)} d\theta. \tag{42}$$

For the Alamouti coded BL-II, the effective SNR of the detection variables can be formulated as

$$\gamma_{II} = \frac{P_{R_1}|h_{R_1,D}|^2 + P_{R_2}|h_{R_2,D}|^2}{N_0}. \tag{43}$$

With the effective SNR, the conditional SER given the channel coefficients can be written as

$$P_M^{h_{R_1,D},h_{R_2,D}} = \frac{1}{\pi}\int_0^{(M-1)\pi/M} \exp(-\frac{g_M\gamma_{II}}{\sin^2\theta})d\theta. \tag{44}$$

Similarly, the error probability of Alamouti coded BL-II can be expressed

$$\begin{aligned} P_{II} =& \iint\limits_{h_{R_1,D},h_{R_2,D}} P_M^{h_{R_1,D},h_{R_2,D}} dh_{R_1,D}dh_{R_2,D} \\ =& \mathcal{F}\left((1+\frac{g_M P_{R_1}\sigma_{R_1}^2}{N_0\sin^2\theta})(1+\frac{g_M P_{R_2}\sigma_{R_2,D}^2}{N_0\sin^2\theta})\right). \end{aligned} \tag{45}$$

As for the uncoded BL-II, it can be seen as a repetition of BL-I with potentially different transmission power P_{R_i} and channel coefficient $h_{R_i,D}$ depending on the power allocation policy and the instantaneous channel condition. Similar to the previous derivations, Eq. (39) $\sim$ (42), we have

$$P_{II,i} = \int_{-\infty}^{\infty} P_M^{h_{R_i,D}} dh_{Ri,D} = \mathcal{F}(1+\frac{g_M P_{R_i}\sigma_{R_i,D}^2}{N_0\sin^2\theta}). \tag{46}$$

Then the end-to-end SER for ACDF described in Section 4.2 can be computed as Eq. (47).

$$\begin{aligned} P_M =& 1 - \left(\prod_{i=1}^{2}(1-P_{I,i})^2(1-P_{II}) + \sum_{i=1}^{2} P_{I,3-i}(2-P_{I,3-i})(1-P_{I,i})^2(1-P_{II,i})\right) \\ =& 1 - \left(1-\mathcal{F}(1+\frac{\varphi_{S,R_1}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}(1+\frac{\varphi_{S,R_2}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}\left((1+\frac{\varphi_{R_1,D}}{\sin^2\theta})(1+\frac{\varphi_{R_2,D}}{\sin^2\theta})\right)\right) \\ & -\mathcal{F}(1+\frac{\varphi_{S,R_2}}{\sin^2\theta})\left(2-\mathcal{F}(1+\frac{\varphi_{S,R_2}}{\sin^2\theta})\right)\left(1-\mathcal{F}(1+\frac{\varphi_{S,R_1}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}(1+\frac{\varphi_{R_1,D}}{\sin^2\theta})\right) \\ & -\mathcal{F}(1+\frac{\varphi_{S,R_1}}{\sin^2\theta})\left(2-\mathcal{F}(1+\frac{\varphi_{S,R_1}}{\sin^2\theta})\right)\left(1-\mathcal{F}(1+\frac{\varphi_{S,R_2}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}(1+\frac{\varphi_{R_2,D}}{\sin^2\theta})\right), \end{aligned} \tag{47}$$

where

$$\begin{aligned} \varphi_{S,R_i} &= \frac{g_M P_S \sigma^2_{S,R_i}}{N_0}, \quad i\text{=}1,\ 2, \\ \varphi_{R_i,D} &= \frac{g_M P_{R_i} \sigma^2_{R_i,D}}{N_0}, \quad i\text{=}1,\ 2. \end{aligned} \tag{48}$$

This closed-form formula can be numerically evaluated rather easily.

4.3.2. Upper Bound and Asymptotic Tight Approximation of SER of ACDF

The closed form I just derived involves integration and is prohibitive for unveiling the relationship between SER and power allocation. Now I obtain the upper bound and asymptotically tight approximation at high SNRs.

Note that the integrands in $P_{I,i}, P_{II}, P_{II,i}$ all reach their maximum at $\sin\theta = 1$. Subsequently, all factors in the final expression of Eq. (47) reach their minimum except for $\mathcal{F}(1+\frac{\varphi_{S,R_i}}{\sin^2\theta})$ $(i=1,2)$. I can use $(1-\mathcal{F}(*))$ to bound $\mathcal{F}(*)$ as in reality the error rate is always smaller than its corresponding success rate. Additionally, substitute $\sin\theta = 1$, I have the upper bound in Eq. (49).

$$\begin{aligned} P_M <& 1-\left(1-\mathcal{F}(1+\frac{\varphi_{S,R_1}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}(1+\frac{\varphi_{S,R_2}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}\left((1+\frac{\varphi_{R_1,D}}{\sin^2\theta})(1+\frac{\varphi_{R_2,D}}{\sin^2\theta})\right)\right) \\ &-\left(1-\mathcal{F}(1+\frac{\varphi_{S,R_2}}{\sin^2\theta})\right)\left(2-\mathcal{F}(1+\frac{\varphi_{S,R_2}}{\sin^2\theta})\right)\left(1-\mathcal{F}(1+\frac{\varphi_{S,R_1}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}(1+\frac{\varphi_{R_1,D}}{\sin^2\theta})\right) \\ &-\left(1-\mathcal{F}(1+\frac{\varphi_{S,R_1}}{\sin^2\theta})\right)\left(2-\mathcal{F}(1+\frac{\varphi_{S,R_1}}{\sin^2\theta})\right)\left(1-\mathcal{F}(1+\frac{\varphi_{S,R_2}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}(1+\frac{\varphi_{R_2,D}}{\sin^2\theta})\right) \\ \leq& 1-\left(1-\mathcal{F}(1+\varphi_{S,R_1})\right)^2\left(1-\mathcal{F}(1+\varphi_{S,R_2})\right)^2\left(1-\mathcal{F}\left((1+\varphi_{R_1,D})(1+\varphi_{R_2,D})\right)\right) \\ &-\left(1-\mathcal{F}(1+\varphi_{S,R_2})\right)\left(2-\mathcal{F}(1+\varphi_{S,R_2})\right)\left(1-\mathcal{F}(1+\varphi_{S,R_1})\right)^2\left(1-\mathcal{F}(1+\varphi_{R_1,D})\right) \\ &-\left(1-\mathcal{F}(1+\varphi_{S,R_1})\right)\left(2-\mathcal{F}(1+\varphi_{S,R_1})\right)\left(1-\mathcal{F}(1+\varphi_{S,R_2})\right)^2\left(1-\mathcal{F}(1+\varphi_{R_2,D})\right) \\ =& 1-\prod_{i=1}^{2}\left(1-\frac{M-1}{M(1+\varphi_{S,R_i})}\right)^2\left(1-\frac{M-1}{M\prod_{j=1}^{2}(1+\varphi_{R_j,D})}\right) \\ &-\sum_{i=1}^{2}\left(1-\frac{M-1}{M(1+\varphi_{S,R_{i-3}})}\right)\left(2-\frac{M-1}{M(1+\varphi_{S,R_{i-3}})}\right)\times \end{aligned}$$

$$\left(1-\frac{M-1}{M(1+\varphi_{S,R_i})}\right)^2\left(1-\frac{M-1}{M(1+\varphi_{R_i,D})}\right).$$

(49)

As SNR gets higher, the other element other than 1 in the variable of $\mathcal{F}(*)$ becomes larger. Thus, I can ignore the 1, and obtain the approximation which is asymptotically tight at high SNRs in Eq. (50).

$$\begin{aligned}
P_M \approx & 1-\left(1-\mathcal{F}(\frac{\varphi_{S,R_1}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}(\frac{\varphi_{S,R_2}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}\left((\frac{\varphi_{R_1,D}}{\sin^2\theta})(\frac{\varphi_{R_2,D}}{\sin^2\theta})\right)\right) \\
& -\mathcal{F}(\frac{\varphi_{S,R_2}}{\sin^2\theta})\left(2-\mathcal{F}(\frac{\varphi_{S,R_2}}{\sin^2\theta})\right)\left(1-\mathcal{F}(\frac{\varphi_{S,R_1}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}(\frac{\varphi_{R_1,D}}{\sin^2\theta})\right) \\
& -\mathcal{F}(\frac{\varphi_{S,R_1}}{\sin^2\theta})\left(2-\mathcal{F}(\frac{\varphi_{S,R_1}}{\sin^2\theta})\right)\left(1-\mathcal{F}(\frac{\varphi_{S,R_2}}{\sin^2\theta})\right)^2\left(1-\mathcal{F}(\frac{\varphi_{R_2,D}}{\sin^2\theta})\right) \\
= & 1-\prod_{i=1}^{2}(1-\frac{\mathcal{A}}{\varphi_{S,R_i}})^2(1-\mathcal{B}\prod_{j=1}^{2}\frac{1}{\varphi_{R_j,D}}) \\
& -\sum_{i=1}^{2}\frac{\mathcal{A}}{\varphi_{S,R_{3-i}}}\left(2-\frac{\mathcal{A}}{\varphi_{S,R_{3-i}}}\right)\left(1-\frac{\mathcal{A}}{\varphi_{S,R_i}}\right)^2(1-\frac{\mathcal{A}}{\varphi_{R_i,D}})
\end{aligned} \tag{50}$$

where,

$$\begin{aligned}
\mathcal{A} &= \frac{1}{\pi}\int_0^{(M-1)\pi/M}\sin^2\theta d\theta = \frac{M-1}{2M}+\frac{\sin\frac{2\pi}{M}}{4\pi} \\
\mathcal{B} &= \frac{1}{\pi}\int_0^{(M-1)\pi/M}\sin^4\theta d\theta = \frac{3(M-1)}{8M}+\frac{\sin\frac{2\pi}{M}}{4\pi}-\frac{\sin\frac{4\pi}{M}}{32\pi}
\end{aligned} \tag{51}$$

Now I use a numeric example to validate the effectiveness of the obtained upper bound and tight approximation. Consider using QPSK modulation ($M=4$). Let $P_S=P_{R_1}=P_{R_2}$, $\gamma=\frac{P}{3N_0}$, all $\sigma_{*,*}^2=1$, $N_0=1$. Then $\varphi_{S,R_i}=\varphi_{R_i,D}=\frac{1}{2}\gamma$ $(i=1,2)$. After simplification, I have $P_{QPSK}^1, P_{QPSK}^2, P_{QPSK}^3$ as the closed-form SER, upper bound, and asymptotic approximation, respectively in Eq. (52). As shown in Fig. 4.2, the upper bound is parallel to the exact SER, which indicates two curves have the same diversity order. The upper bound well represents the trend of the SER when the SNR changes. The approximation curves merges with the

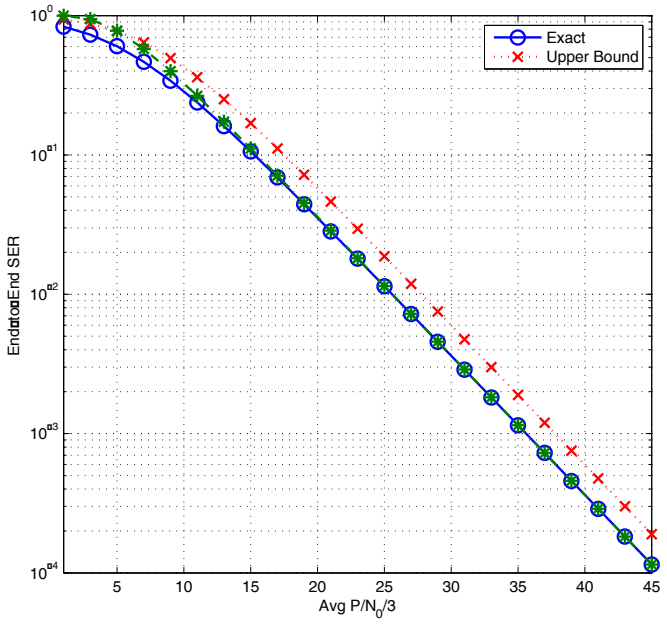


FIGURE 4.2. Comparison of the closed-form SER, upper bound and asymptotically tight approximation. (QPSK modulation, $P_S = P_{R_1,D} = P_{R_2,D}$, $\sigma^2_{*,*} = 1$, $N_0 = 1$.)

closed form as the SNR get higher, which validates its asymptotic tightness.

$$P^1_{QSPK}=1-\left(1-\frac{1}{\pi}\int_0^{3\pi/4}\frac{2\sin^2\theta}{2\sin^2\theta+\gamma}d\theta\right)^2\left(1-\frac{1}{\pi}\int_0^{3\pi/4}\frac{2\sin^2\theta}{2\sin^2\theta+\gamma}d\theta\right)^2\times$$
$$\left(1-\frac{1}{\pi}\int_0^{3\pi/4}\frac{4\sin^4\theta}{(2\sin^2\theta+\gamma)^2}d\theta\right)$$
$$-\frac{2}{\pi}\int_0^{3\pi/4}\frac{2\sin^2\theta}{2\sin^2\theta+\gamma}d\theta\left(2-\frac{1}{\pi}\int_0^{3\pi/4}\frac{2\sin^2\theta}{2\sin^2\theta+\gamma}d\theta\right)\left(\frac{1}{\pi}\int_0^{3\pi/4}\frac{2\sin^2\theta}{2\sin^2\theta+\gamma}d\theta\right)^3,$$

$$P^2_{QSPK}=1-\left(1-\frac{3}{4+2\gamma}\right)^4\left(1-\frac{3}{(\gamma+2)^2}\right)-2\left(1-\frac{3}{4+2\gamma}\right)^4\left(2-\frac{3}{4+2\gamma}\right)$$
$$=1-\left(1-\frac{3}{4+2\gamma}\right)^4\left(5-\frac{3}{(\gamma+2)^2}-\frac{3}{4+2\gamma}\right),$$

$$P^3_{QSPK}=1-\left(1-\frac{3\pi+2}{4\pi\gamma}\right)^4\left(1-\frac{9\pi+8}{8\pi\gamma^2}\right)-\frac{3\pi+2}{2\pi\gamma}\left(2-\frac{3\pi+2}{4\pi\gamma}\right)\left(1-\frac{3\pi+2}{4\pi\gamma}\right)^3.$$

(52)

4.3.3. Optimized Power Allocation Policy

I now proceed to discuss the optimized asymptotic power allocation strategy. I use the tight approximation obtained in Eq. (50), as it is mathematically less obscure than the closed form in Eq. (47) and its curve merges quickly with the one plotted from Eq. (47), as shown in previous discussions. The goal is to allocate power among the relays and the source such that the error probability P_M is minimized. Within all the parameters relevant to Eq. (50), the constant factors here include the global power constraint, specified by $P_S + P_{R_1} + P_{R_2} \leq P$ where P is the total available power. Furthermore, the average SNR $\gamma = \frac{P}{3N_0}$ is fixed. Given a type of channel, I consider the noise power N_0 and the variances of the fading coefficients $\sigma_{*,*}$ invariant. In practice, the fading and the noise distributions could change due to the environment. In this case, an adaptive power control can be developed based on the tight approximation. However, this is beyond the scope of this chapter. I only consider an invariant distribution of noise and fading, to be specific, the AWGN and the Rayleigh fading. Also a particular modulation type, in other words, M is known before the allocation is carried out. Thus, the task here becomes an optimization problem wherein the goal function is the asymptotic P_M with variables P_S, P_{R_1}, P_{R_2} under the constraint of $P_S + P_{R_1} + P_{R_2} \leq P$. I denote the asymptotic P_M as $\mathcal{P}(P_S, P_{R_1}, P_{R_2})$ to emphasize its relationships with the individual power allowances. The goal is to find the set of parameters P_S, P_{R_1}, P_{R_2} in order to achieve the minimized end-to-end error probability, which now can be formulated as in Eq. (53).

$$
\begin{aligned}
& \operatorname{argmin}_{P_S, P_{R_1}, P_{R_2}} \mathcal{P}(P_S, P_{R_1}, P_{R_2}) \\
= \; & \operatorname{argmin}_{P_S, P_{R_1}, P_{R_2}} \left(1 - \prod_{i=1}^{2} \left(1 - \frac{\mathcal{C}_i}{P_S} \right)^2 \left(1 - \frac{\mathcal{C}_3}{P_{R_1} P_{R_2}} \right) \right. \\
& \left. - \sum_{i=1}^{2} \frac{\mathcal{C}_{3-i}}{P_S} \left(2 - \frac{\mathcal{C}_{3-i}}{P_S} \right) \left(1 - \frac{\mathcal{C}_i}{P_S} \right)^2 \left(1 - \frac{\mathcal{C}_{3+i}}{P_{R_i}} \right) \right),
\end{aligned}
\tag{53}
$$

where,

$$\mathcal{C}_1 = \frac{\mathcal{A}N_0}{g_M\sigma^2_{S,R_1}}, \tag{54}$$

$$\mathcal{C}_2 = \frac{\mathcal{A}N_0}{g_M\sigma^2_{S,R_2}}, \tag{55}$$

$$\mathcal{C}_3 = \frac{\mathcal{B}N_0^2}{g_M^2\sigma^2_{R_1,D}\sigma^2_{R_2,D}}, \tag{56}$$

$$\mathcal{C}_4 = \frac{\mathcal{A}N_0}{g_M\sigma^2_{R_1,D}}, \tag{57}$$

$$\mathcal{C}_5 = \frac{\mathcal{A}N_0}{g_M\sigma^2_{R_2,D}}. \tag{58}$$

I first show another numeric example, and then discuss a more general scenario. Let $P_{R_1} = P_{R_2}$, then I obtain the partial derivative of the goal function in Eq. (53) in regard to P_S, which is $\frac{\partial\mathcal{P}}{\partial P_S}$. Let $\frac{\partial\mathcal{P}}{\partial P_S} = 0$, then I can derive the condition for the goal function to reach its minimum. Due to the limited space, I only provide the equations for the given channel conditions and modulation types as they are much simpler than the general results. Consider the BPSK modulation ($M = 2$), and again $\sigma^2_{*,*} = 1$. Then $N_0 = 1$. $\mathcal{C}_1 = \mathcal{C}_2 = \mathcal{C}_4 = \mathcal{C}_5 = \frac{1}{4}, \mathcal{C}_3 = \frac{3}{16}$. Now after a series of simplification, the zero requirement on the partial derivative reduces to solving the cubic equation below,

$$\begin{aligned} 0 = {} & 32P_S^6 - 16(6P+11)P_S^5 + 6(16P^2+24P+23)P_S^4 \\ & -8P(4P^2+34P-11)P_S^3 + (112P^3+38P+5)P_S^2 \\ & -6P^2(7P-2)P_S + 2P^2(2P-1). \end{aligned} \tag{59}$$

Substituting with P values, I can obtain different P_S/P ratio numerically. For example, $P_S/P \approx 0.3051, 0.2828, 0.2536, 0.2506$ for $P = 3, 10, 20, 30$ dB respectively. A complete set of results is shown in Fig. 4.3. For each given P, a pair of ratios P_S/P and $P_{R_1,D}/P$ can be calculated. As seen from the figure, P_S/P asymptotically approaches 1/4.

More generally, without constraining $P_{R_1,D} = P_{R_2,D}$, by varying the transmission power of the source and each relay, I obtain a mesh plot in Fig. 4.4 for the BPSK modulation when $\sigma^2_{*,*} = 1$, $N_0 = 1$. It is shown from the contours projected at the bottom pane that the

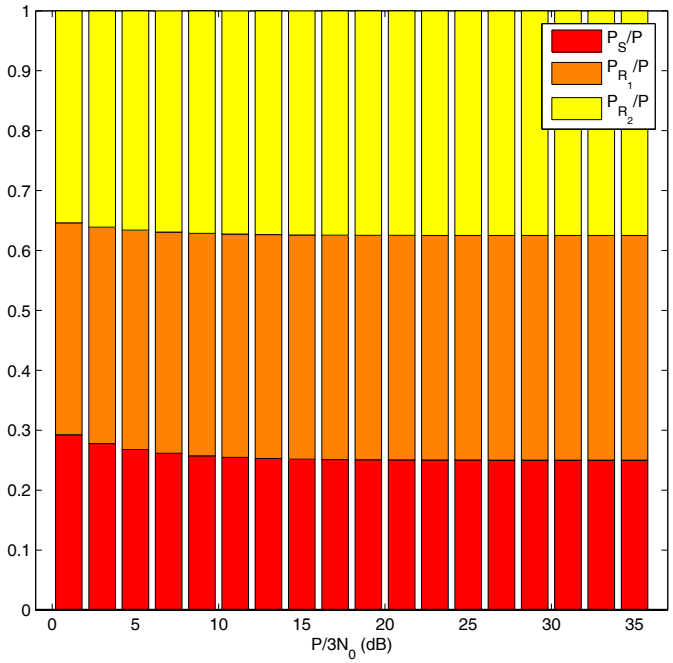


FIGURE 4.3. P_S/P and $P_{R_1,D}/P$ ratios plotted against the SNR. ($P_{R_1,D} = P_{R_2,D}$, $\sigma^2_{*,*} = 1$, $N_0 = 1$.)

optimized allocation for relays is symmetrical. This, to some extent, justifies the numerical example given in the previous sections by imposing an equal power allocation on the relays. The optimized area enclosed by the inner most contour lies in the low ratio as it is known that for the same given SNR, BL-II (Alamouti code) achieves a better performance than the conventional uncoded transmission.

Finally, I show comparisons between the performances of two intuitive and the optimized power allocation policies. In a system using BPSK, it is seen from Fig. 4.5 that the equal allocation among the source and two relays yields the worst SER. The system performs better when the total power of relays equals that of the source. The optimized allocation policy achieves the best error rate, and its advantage is more significant at higher SNR values. As for the QPSK modulation ($M = 4, \sigma^2_{*,*} = 1, N_0 = 1$), now $\mathcal{C}_1 = \mathcal{C}_2 = \mathcal{C}_4 = \mathcal{C}_5 = \frac{3}{4} + \frac{1}{\pi}, \mathcal{C}_3 = \frac{9}{8} + \frac{1}{\pi}$. With similar derivation, the zero requirement on the partial derivative for QPSK can be obtained. For example, $P_S/P \approx 0.7026$, 0.6988, 0.6926 0.6920, 0.6919 for $P =$ 3, 5, 15, 25, 35 dB, respectively. The ratio P_S/P does not vary much and approaches 0.69

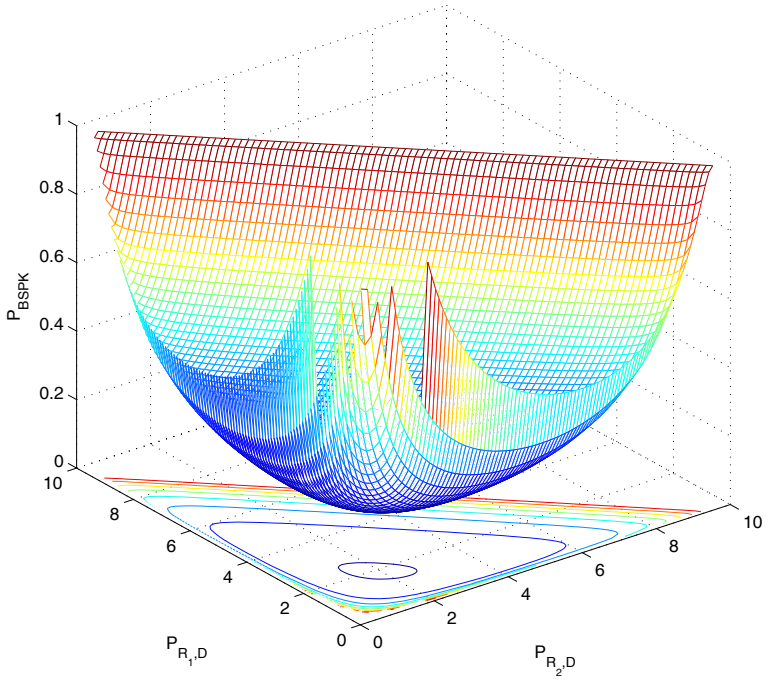


FIGURE 4.4. The end-to-end error performance of ACDF using BPSK.

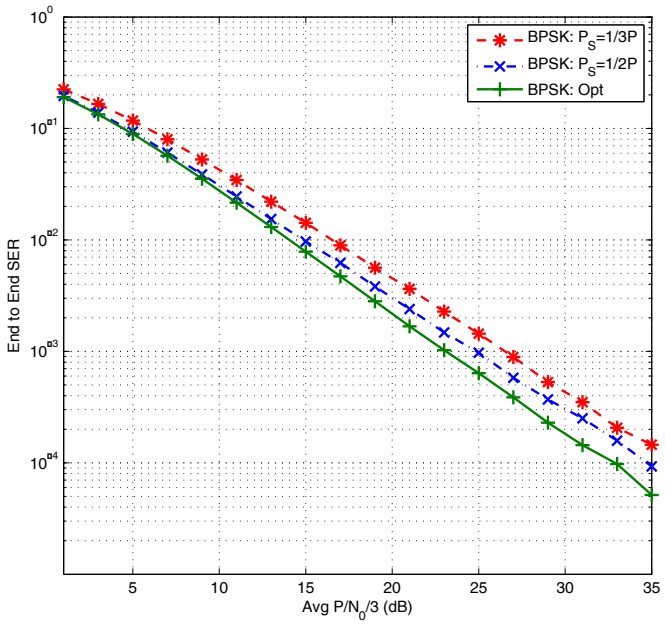


FIGURE 4.5. The end-to-end SER comparison between two intuitive and the optimized power allocation policies using BPSK.

asymptotically. The same observation can be made on Fig. 4.6 for the QPSK modulation. Additionally, the exact closed form SER in Eq. (52) and the plot in Fig. 4.2 are also verified here.

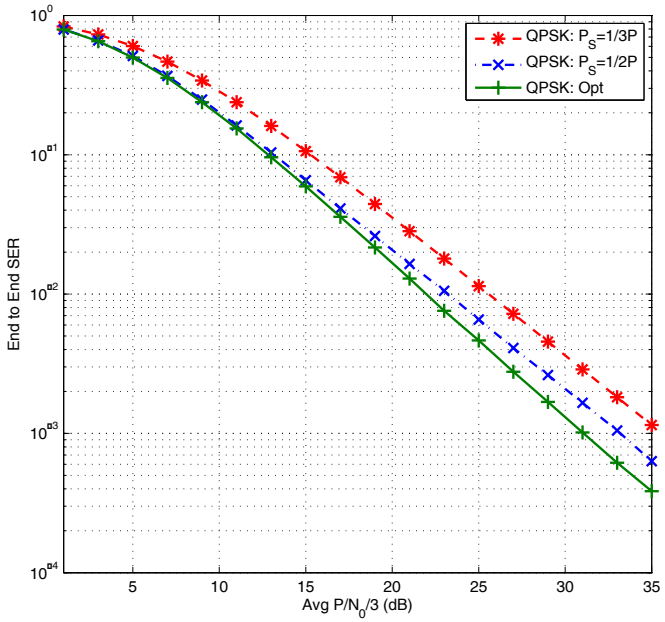


FIGURE 4.6. The end-to-end SER comparison between two intuitive and the optimized power allocation policies using QPSK.

4.4. Summary

In this chapter, I investigated the optimized transmission power allocation strategies aiming at minimizing the end-to-end SER performance of the ACDF protocol in the energy constrained fading parallel relay network. I derived the exact closed-form SER of the system, and further obtained an upper bound to represent the trend in a more mathematically simplistic fashion, and also an approximation asymptotically tight at high SNRs. These tools help reveal the relationship between the SER performance and the power allocation. Finally, I derived the goal function and the optimized solution. Numerical results confirm my work by showing the performance improvement in terms of minimizing the end-to-end SER.

CHAPTER 5

PERMUTATION MODULATIONS FOR PHYSICAL-LAYER NETWORK CODING

In this chapter, I introduce a novel method of the physical-layer network coding by utilizing the binary permutation modulations instead of the traditional modulation types. I compare the end-to-end performance of the proposed scheme with the traditional approach using BPSK, as BPSK yields the smallest error probability among the conventional modulation types. Both theoretical analysis in the explicit analytic forms and numerical results are given. In particular, I prove a series of lemmas and a theorem on the performance and the optimality of the proposed scheme. It is shown that the proposed class of permutation codes outperforms the conventional approach using BPSK in the physical-layer network coding.

The rest of this chapter is organized as follows. The system model is described in detail in Section 5.2. In Section 5.3, the error rate performance of my proposed scheme is theoretically analyzed with comparison to the traditional approach using BPSK. Numerical results are presented in Section 5.4, followed by the conclusion in Section 5.5.

5.1. Background

David Slepian first introduced a class of codes and its corresponding decoders, permutation modulations, for digital information transmission in 1965 [84]. Described in the context of Shannon's Gaussian communication channel, the idea of permutation modulations is as follows. Consider a sequence of signals as a vector in a multi-dimensional Euclidean space, the coding is performed by swapping and rearranging (sign changing is allowed in a variant scheme) the scalar components within in all possible ways. Each vector is called a codeword.

The content in this chapter with the exception of Section 5.2.4 is reproduced from N. Xu, S. Fu, and Y. Huang, "Permutation Modulation For Physical-Layer Network Coding," IEEE Journal of Wireless Communications (under review), with permission from IEEE.

In this section, the two terms "vector" and "codeword" are used interchangeably whenever not causing confusion. All the vectors resulting from the coding operations are referred to as permutation codes, while the very codeword from which the permutation codes are obtained is known as the initial vector.

Such a system has many appealing features [84]. Some common codes, such as pulse code modulations, simplex codes, and biorthogonal codes are all special cases of permutation modulations [85]. Given that all codewords are the permutations of the initial vector, the same energy is required for transmitting them. At the receiving end, the decoder based on Maximum Likelihood criterion is relatively easy to construct, and the possible sent sequences or codewords need not to be generated locally. Additionally, the system is resilient to small changes in the transmitter parameters.

The topic of permutation modulations has been investigated by numerous works. One of the most important topics in the early stage research is the problem of choosing the initial vector for the optimal performances. This problem was stated in [84] without giving explicit solutions. For a certain given energy allowance, there exist many codewords. Assuming that all possible sent messages occur with the same a priori probability of transmission, then the Maximum Likelihood decoding minimizes the average error probability [86]. The Euclidean distance between codewords is the dominant factor affecting the average error rate of the system. The distribution of energy within the components of the vector becomes essentially an equivalent problem. Under the same energy constraint, a system with the initial vector yielding the largest minimum distance achieves its optimality. Karlof formulated this problem as integer programming, and derived a numerical algorithm for searching for the optimal solution [86]. Later, Ingemarsson derived explicit formulas for the optimal initial vectors [87].

The error performance of PM coding was studied in [88, 89]. Gaarder considered the binary permutation modulations in the fading channel, and found the upper bounds of the error probability [88]. The author showed that with the proper order of diversity, any form of binary permutation modulations asymptotically approaches zero while increasing the memory of the transmitter at the speed of exponential decrement. Also, the transmitter's rate is

constrained to less than the signal-to-noise ratio per unit bandwidth from the receiver's perspective. Nordio and Viterbo presented efficient ML decoder construction methods for both Variant I and II in different fading scenarios, such as AWGN (no fading, only Gaussian noise is present), fast fading and slow fading channels [89]. The authors gave exact expressions of the error probability of the independent Rayleigh fading channels. In recent works such as [90], Silva and Finamore proposed a scheme that extends the Slepian's scalar PM codes into the case where the codeword components are also multi-dimensional vectors, which proved more advantageous.

The same permutation codes can be used for the purpose of source quantization as well as modulation, the latter of which is the focus of this chapter's work.

5.2. System Model

5.2.1. Permutation Modulation

Signals in a channel can be viewed as vectors in a Euclidean vector space. In permutation modulations, the coding is performed by rearranging (also sign changing in certain schemes) the scalar components within a particular vector in all possible ways. Each vector is called a codeword. I use "vector" and "codeword" interchangeably hereafter unless causing confusion. The very vector from which all codewords are obtained is known as the initial vector, denoted as x_0.

DEFINITION 5.2.1. *The initial vector x_0 is expressed as*

$$x_0 = (\underbrace{\mu_1, \ldots, \mu_1}_{m_1}, \underbrace{\mu_2, \ldots, \mu_2}_{m_2}, \cdots, \underbrace{\mu_k, \ldots, \mu_k}_{m_k}), \tag{60}$$

where k is an integer, μ_i is a real number for $1 \leq i \leq k$, and without loss of generality $\mu_1 < \mu_2 < \cdots < \mu_k$.

The number of components in all the codewords is $n = \sum_{i=1}^{k} m_k$.

DEFINITION 5.2.2. *The energy of a codeword is defined as*

$$E \triangleq \sum_{i=1}^{k} m_k \mu_k^2. \tag{61}$$

The energy E of all codewords is the same. Slepian proposed two variants of the permutation modulation scheme [84]. In variant I, only permutation of the component are performed without changing their signs whereas variant II allows sign changing at the same time. In this chapter, I only consider variant I of the binary permutation modulations (BPM) where $k = 2$ due to its simplicity.

5.2.2. Encoding and Decoding in Binary Permutation Modulations

All the codewords associated with the same x_0 (including x_0 itself) collectively form a set of size M. M can be computed as

$$M = \frac{(m_1 + m_2)!}{m_1!\, m_2!}. \tag{62}$$

With a given pair of m_1, m_2, the binary permutation code is noted as BPM$[m_1, m_2]$ hereafter. For a BPM$[m_1, m_2]$ scheme, the number of possible codewords is M as computed in Eq. (62). Therefore, a single vector can accommodate $\lfloor \log_2 M \rfloor$ bits of information. I designate another symbol for this number.

Definition 5.2.3. *Denote B as the number of information bits accommodated in a codeword of BPM$[m_1, m_2]$, then*

$$B \triangleq \left\lfloor \log_2 \frac{(m_1 + m_2)!}{m_1!\, m_2!} \right\rfloor. \tag{63}$$

The encoding procedure is straightforward: choose B codewords out of the M (I discuss the effect of the choices in a moment), and label them with the binary representation of $0, \cdots, B-1$, respectively; then I map messages with the codewords by their binary labels. The unchosen codewords are practically invalid and discarded.

All the codewords are assumed to have the same a priori probability. Then the decoder based on the maximum likelihood (ML) criterion yields the optimal performance in terms of the error rate [86]. Owing to the unique feature of permutation modulations, a low-complexity decoding other than the exhaustive search can be used [89]. In this method, decoding is performed by replacing the smallest k_1 components of the received vector with μ_1, and the remaining k_2 components with μ_2 (Note I already assumed $\mu_1 < \mu_2$ previously).

This is an ML decoder when $B = \log_2 M$. Additionally, the receivers do not need to locally generate the valid codewords. However, when $B < \log_2 M$, as some codewords have been purged from the actual transmission, the decoding using this method results in invalid codewords. Its performance is suboptimal. Going back to the encoding operation, the choice of the B codewords does not affect the optimal error probability of the system even when $B < \log_2 M$. Yet in order to achieve the optimality , I turn to the exhaustive search based on the ML decoding in such cases, and all possible transmitted codewords must be generated at the receiving ends. I show in Section 5.3 that such optimal BPMs yield better system performances than BPSK, which has the lowest error probability among traditional M-ary modulations (e.g., QPSK, 16QAM).

5.2.3. Physical-Layer Network Coding

Now I briefly visit the concept of the physical-layer network coding (PLNC) before discussing its application with permutation modulations. It is a recent research topic which has received increasing attention for the past few years. Here I only cover the important basics. Please refer to works such as [32, 68] for more details.

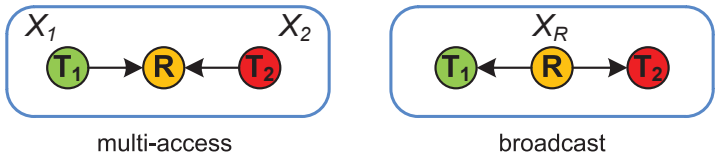


FIGURE 5.1. The two-way relay channel model.

A two-way relay channel is shown in Fig. 5.1, where two terminals T_1 and T_2 exchange messages with the assistance of a relay R. No direct link is available between T_1, T_2. Traditionally, the interference was seen only harmful to the transmission, and the terminals send symbols to the relay alternately to avoid such interference. With the help of the network coding technique, the relay is able to broadcast an encoded symbol X_R obtained from the received X_1, X_2. The PLNC takes advantage of the additive nature of the electromagnetic waves, and turns the interference into more capacity boost by allowing simultaneous transmissions to the relay [32]. Typically, the first stage of the whole transmission is known as

the multi-access (MA) phase, and the second as the broadcast (BC) phase. In a Gaussian channel, I model the MA phase as

$$U[t] = X_1[t] + X_2[t] + n_0[t] \tag{64}$$

and the BC phase as

$$Y_k[t] = X_R[t] + n_k[t] \tag{65}$$

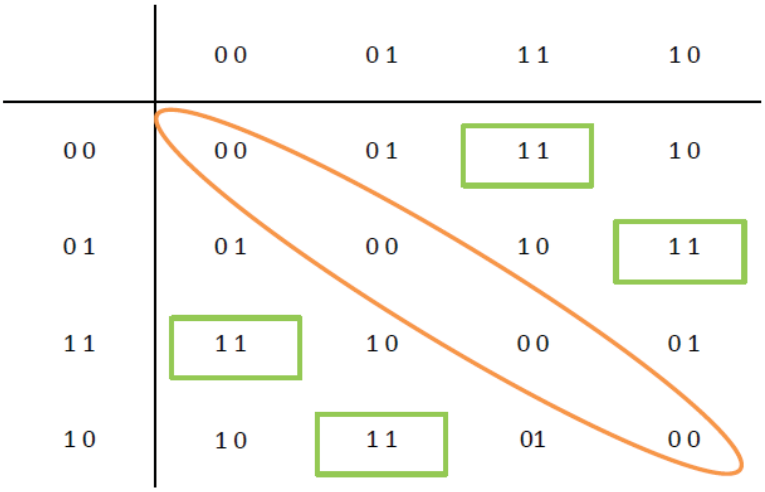
where t is the index of the symbols (or components in PMs, $1 \leq t \leq n$), $k = 1, 2$, U is the received symbol at the relay, and n_k, n_0 are the additive white Gaussian noise.

I use BPMs (not necessarily the same) for both MA, BC phases. Compared to the conventional modulations (e.g., BPSK), my proposed method using BPMs differs in how received symbols are processed at the relay. I propose two methods for this purpose. In method I, I adopt a BPM for the BC phase, BPM (BC), which is of higher order than that in the MA phase. T_1 and T_2 send out one component of a codeword at a time. The transmitted symbol is from the set $\{\mu_1, \mu_2\}$. The decision rules for the MA phase can be described as follows,

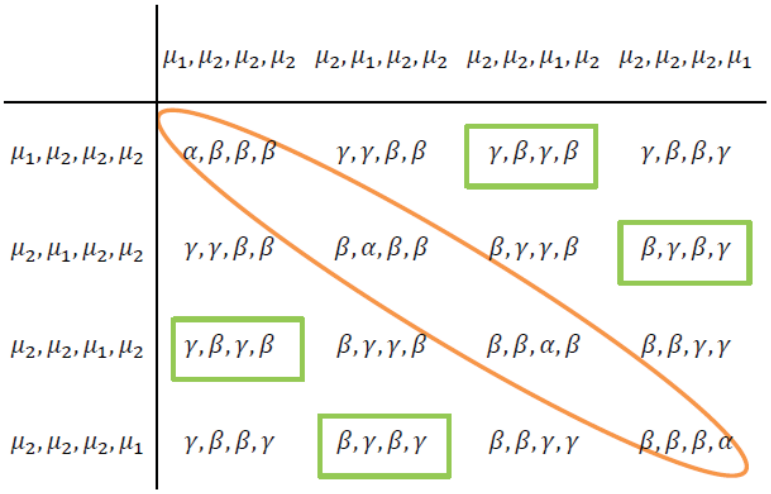
$$\phi(U[t]) = \begin{cases} U[t] < \frac{3\mu_1+\mu_2}{2} \ \cup \ U[t] > \frac{\mu_1+3\mu_2}{2}, & \text{decode as } 0, \\ \frac{3\mu_1+\mu_2}{2} \leq U[t] \leq \frac{\mu_1+3\mu_2}{2}, & \text{decode as } 1. \end{cases} \tag{66}$$

Note that in the decision rules described in Eq. (66), 0 indicates the transmitted components $(X_1[\cdot], X_2[\cdot])$ are the same value whereas 1 indicates they are different. After accumulating symbols of indices from 1 to n, a coded vector of n components of $\{0, 1\}$ is obtained.

Now for the BC phase, I utilize a BPM of higher order to accommodate n rather than only B bits in the MA phase. Let's look at a simple example: use BPM[1,1] in the MA phase. Now $B = 1$ as BPM[1,1] consists of exactly two codewords $\{\mu_1, \mu_2\}$ and $\{\mu_2, \mu_1\}$. Suppose T_1, T_2 both send the same codeword $\{\mu_1, \mu_2\}$, and zero noise is present for simplicity. As in Eq. (64), $U[1] = X_1[1] + X_2[1] = 2\mu_1, U[2] = X_1[2] + X_2[2] = 2\mu_2$. According to the decision rules in Eq. (66), $U[1], U[2]$ are mapped to 00, i.e., received components are the same value.



(a) XOR of binary indices.



(b) Addition of codewords ($\alpha = 2\mu_1, \beta = 2\mu_2, \gamma = \mu_1 + \mu_2$).

FIGURE 5.2. The mapping scheme of PLNC method II using PM[1,3].

Then I send this 00 using a BPM which accommodates two bits, e.g. BPM[1,3], BPM[2,2]. I discuss the choice of such BPMs (BC) in Section 5.3.

I only apply Method II to the cases where $m_1 = 1$ or $m_2 = 1$. For such BPMs, there exists a mapping that associate the 2^n possible received vectors to 2^B codewords, and therefore the same BPM can used in both MA and BC phases. An example, the mapping of BPM[1,3], is shown in Fig. 5.2. Fig. 5.2(a) gives the XOR results of the binary indices of the codewords. For those pairs of codewords, which has the same XOR value, I map them into the same region

in Fig. 5.2(b). For instance, as marked out in the figure, additions $\{\mu_1+\mu_2, 2\mu_2, \mu_1+\mu_2, 2\mu_2\}$ ($\{\gamma, \beta, \gamma, \beta\}$) and $\{2\mu_2, \mu_1+\mu_2, 2\mu_2, \mu_1+\mu_2\}$ ($\{\beta, \gamma, \beta, \gamma\}$) are both mapped to 11. Mapping to the regions of the XOR results ensures that the same BPM can accommodate the number of bits. Also, it is not difficult to see that for any BPM$[m_1, m_2]$ ($m_1 = 1$ or $m_2 = 1$), the additions of codewords in the upper triangle do not duplicate each other. Therefore there is no ambiguity for the decoding at the end of the BC phase when the terminal tries to obtain the vector sent by the other side using its own. Duplicate additions cause the mapping to become "many-to-many," causing ambiguity and impossible decoding at the receivers.

5.2.4. Discussion on the Failure Cases of the Proposed Mapping

0	1	3	2	6	7	5	4
1	0	2	3	7	6	4	5
3	2	0	1	5	4	6	7
2	3	1	0	4	5	7	6
6	7	5	4	0	1	3	2
7	6	4	5	1	0	2	3
5	4	6	7	3	2	0	1
4	5	7	6	2	3	1	0

FIGURE 5.3. XOR in decimal formats of the binary indices when $B = 3$.

However, based on the numerical testing on some BPM examples with $m_1, m_2 \geq 2$, I have found that such mappings do not always exist for any given BPM. For instance, I provide some failure examples of BPM[2,3] as below. BPM[2,3] consists of ten codewords, out of which eight are chosen for the coding. Then there are $\binom{10}{8} = 45$ ways to make the selection.

0	12	10	9	20	18	17	30
12	0	6	5	24	30	29	18
10	6	0	3	30	24	27	20
9	5	3	0	29	27	24	23
20	24	30	29	0	6	5	10
18	30	24	27	6	0	3	12
17	29	27	24	5	3	0	15
30	18	20	23	10	12	15	0

FIGURE 5.4. Failure case 1 of the mapping attempt for BPM[2,3] when $B = 3$. (The occurrence of 30 is 3. It must be grouped with the results with the occurrence of 1. The candidates are 9, 15, 17, 23, all of which share a row or column with 30.)

Before stepping into the details of the necessary conditions to support this mapping scheme, I explain how these figures are generated. First of all, the information capacity for BPM[2,3] is $B = 3$, which means that there are $2^3 = 8$ codewords selected for the coding. It is very tedious to write out all the exact results of the XOR operation on the binary indices and addition of the vectors. Instead, here I use the decimal representation of the binary XOR values. For example, in Fig. 5.3, the binary representations are Gray-coded. The first row is the XOR result of 000 with $000, 001, 011, 010, 110, 111, 101, 100$, respectively. I get the same result by computing the XOR of a number with zeros. In this case, the decimal representations are $0, 1, 3, 2, 6, 7, 5, 4$. It is done in the same fashion for all the entries in that mapping table in Fig. 5.3.

As for the addition results, I handle it in a slightly different way. To my best knowledge, there is no simple Gray coding solution for these investigated permutation codes. To mitigate

the effect of labeling, one can always equivalently perform column/row swapping similar to that in the Gaussian elimination on either table of the mapping. Also, based on the decision rules summarized in Eq. (66), I do not need to distinguish $2\mu_1$ from $2\mu_2$ as they fall into the same decision region. Effectively, this performs XOR operation on the components of a codeword. If the addition of two components from two codewords is $\mu_1 + \mu_2$, it is written as 1 indicating that, the two components from which it is obtained, are different. For $2\mu_1$ or $2\mu_2$, it is interpreted as 0. This way, I have another set of binary representations of more bits ($(m_1 + m_2)$ bits in this case). Again, using their correspondent decimal representations, I have obtained Fig. 5.4. As the original two codeword values are not important to the testing, they are omitted as the table headers. For example, in Fig. 5.4, 30 is the binary 11110, which is the XOR of 01100 and 10010, the second codeword 12 and the sixth 18, respectively.

TABLE 5.1. The occurrence of codeword addition results in the upper triangle in Fig. 5.4.

value	**0**	**3**	**5**	**6**	**9**	**10**	**12**	**15**	**17**	**18**	**20**	**23**	**24**	**27**	**29**	**30**
occurrences	8	2	2	2	1	2	2	1	1	2	2	1	3	2	2	3

Having explained the meaning and presentation of the figures, now let's look at several cases where the table does not meet the requirement of the proposed mapping method. It is not difficult to see the tables in Figs. 5.4, 5.5, and 5.6 are diagonally symmetric. For simplicity, I only look at the upper triangle (The entries along the diagonal is all zeros). Table 5.1 lists the occurrences of each addition value in the upper triangle. For example, 30 appears three times and is filled by the color blue (shaded if printed as black and white) in Fig. 5.4. In order for the proposed mapping to exist, it is not difficult to see that a number of such addition values must together form a group of the same size as a potential mapping region in Fig. 5.3. In the case of BPM[2,3], all XORs above the diagonal is repeated exactly four times. Then the sums of codewords must form groups of size 4. Let's start with 30 as it can only be grouped with a value of only one occurrence while values such as 3 can

be grouped with other twice occurrent or two once occurrent's. Four once occurrent values $9, 15, 17, 23$ according to Table 5.1 are also marked out with the color red (shaded in black and white version) in Fig. 5.4. One necessary condition to form the proposed mapping is that values grouped together do not occupy the same column or row in the upper triangle. The reason is straightforward: if such is allowed that two distinct values take the same row, for instance, then at the receiver's side, there are two possibilities for the decoding of this vector. By this rule, all four candidates are eliminated. Therefore, no method II type mappings are supported by this instance of BPM[2,3].

0	***12***	*10*	9	*20*	**18**	30	29
12	0	6	*5*	24	30	**18**	*17*
10	6	0	**3**	30	24	*20*	23
9	5	3	0	29	*27*	23	*20*
20	24	30	29	0	6	*10*	9
18	30	24	27	6	0	***12***	15
30	18	20	23	10	12	0	**3**
29	17	23	20	9	15	3	0

FIGURE 5.5. Failure case 2 of the mapping attempt for BPM[2,3] when $B = 3$. (The same group is filled with the same color, numbers therein of the same style: bold, italic, underlined, or a combination of those. The only valid groups are {6, 29}, {15, 30}, {10, 17, 27}, {3, 18}, {5, 20}, {9, 24}. 12 must be grouped with 23. However they share the same column.)

Fig. 5.5 shows another failed instance of BPM[2,3]. This time, the attempt has gone as far as grouping most values, {6, 29}, {15, 30}, {10, 17, 27}, {3, 18}, {5, 20}, {9, 24}. These are the only valid groups I was able to form with the table entries according to the

TABLE 5.2. The occurrence of codeword addition results in the upper triangle in Fig. 5.5.

value	**0**	**3**	**5**	**6**	**9**	**10**	**12**	**15**	**17**	**18**	**20**	**23**	**24**	**27**	**29**	**30**
occurrences	8	2	1	2	2	2	2	1	1	2	3	2	2	1	2	3

occurrence statistics in Table 5.2. Now, look at 12, whose block is colored in red and the number is underlined normal font. The only remaining value is 23. However, 23 takes the same column as 12. Thus, they cannot be grouped together. This attempt has exhausted all the possible choices of forming valid groups only to end in failure.

0	**10**	*9*	20	***18***	*17*	**30**	27
10	0	***3***	**30**	*24*	27	20	*17*
9	3	0	**29**	27	*24*	*23*	***18***
20	30	29	0	*6*	5	**10**	**15**
18	24	27	6	0	***3***	12	*9*
17	27	24	5	3	0	**15**	**10**
30	20	23	10	12	15	0	5
27	17	18	15	9	10	5	0

FIGURE 5.6. Failure case 3 of the mapping attempt for BPM[2,3] when $B = 3$. (The same group is filled with the same color, numbers therein of the same style: bold, italic, underlined, or a combination of those. The groups are formed with all distinct results: {12, 27}, {0}, {6, 17, 23}, {15, 30}, {3, 18}, {5, 20}, {10, 29}, {9, 24}. However, this mapping cannot be transformed by row/column swapping to match that in Table 5.3.)

TABLE 5.3. The occurrence of codeword addition results in the upper triangle in Fig. 5.6.

value	**0**	**3**	**5**	**6**	**9**	**10**	**12**	**15**	**17**	**18**	**20**	**23**	**24**	**27**	**29**	**30**
occurrences	8	2	2	1	2	3	1	2	2	2	2	1	2	3	1	2

For a third case, let's look at Fig. 5.6. This time, I succeed in grouping the values based on Table 5.3, such that within each group, no two values occupy the same row or column in the upper triangle. The groups are {12, 27}, {0}, {6, 17, 23}, {15, 30}, {3, 18}, {5, 20}, {10, 29}, {9, 24}, and again filled with different colors and with different styled fonts (bold, italic, underlined, or a combination of those) for the numbers. The next step, these regions must match those in Fig. 5.3. However, by performing all possible row/column swapping (same as that in Gaussian Elimination) on all the table entries, the search terminates without success. This is the only valid grouping result. Failed to match the regions of the XORs of the indices, this instance of BPM[2,3] is shown not to support the proposed mapping.

Now I summarize two necessary conditions from numerical testing performed (not limited to BPM[2,3]) as below.

- The same sum values must be mapped to the same XOR of the corresponding indices.
- Different sum values can be mapped to the same XOR of the corresponding indices.

Certain rules mentioned in the previous discussions, such as no values from the same group occupying the same row or column, are implied in these necessary conditions. Sharing the same row or column renders the mapping to be a many-to-many relationship, which prevents the receivers from correct decoding.

Additionally, the same testing failed on BPM[2,2], BPM[3,3], etc. As shown in Section 5.3, the optimal BPM satisfies $m_1 = 1$ or $m_2 = 1$. Therefore I do not pursue this and get into more details.

5.3. Performance Analysis

In this section, I theoretically analyze the performance of the proposed scheme, and show the advantages of using BPMs in PLNC. BPSK is chosen for comparison as it has the lowest bit error probability among the traditional M-ary modulations. The symbol error rate (SER) is used as the primary performance metric. When using the Maximum Likelihood decoding, the minimum distance d_{min} between the codewords is the main factor affecting the error performance: a larger d_{min} yields a smaller SER. Note that, although for a codeword of BPM$[m_1, m_2]$, the length or the number of the components of the vector is $n \triangleq m_1 + m2$, the information it conveys is actually $B = \lfloor \log_2 \frac{(m_1+m_2)!}{m_1!\, m_2!} \rfloor$ bits. Therefore, for a fair comparison, when I consider a symbol using BPSK for transmission, the symbol size is B rather than n.

Now I introduce several lemmas before reaching the main theorem on the performance of BPMs.

LEMMA 5.3.1. *For any binary permutation modulation BPM$[m_1, m_2]$, the optimal μ_1, μ_2 under a given energy allowance satisfies the following,*

$$m_1\mu_1 + m_2\mu_2 = 0. \tag{67}$$

PROOF. As defined in Definition 5.2.2, the energy of a codeword is $E = m_1\mu_1^2 + m_2\mu_2^2$. Solving μ_1 in terms of μ_2, I have $\mu_1 = \pm\frac{1}{m_1}\sqrt{E - m_2\mu_2^2}$. The minimum Euclidean distance between any two codewords is $d_{min} = 2|\mu_1 - \mu_2|$. There is only the energy constraint on the values of μ_1, μ_2. I can also see that, if μ_1 (μ_2) is a valid solution then $-\mu_1$ ($-\mu_2$) is too. In order to maximize d_{min}, I choose μ_1, μ_2 of different signs. Without loss of generality, I assume here $\mu_1 \leq 0, \mu_2 \geq 0$ (they cannot be zeros at the same time unless $E = 0$). Thus, $d_{min} = 2|\mu_1 - \mu_2| = 2(\mu_2 - \mu_1) = 2\left(\mu_2 + \frac{1}{m_1}\sqrt{E - m_2\mu_2^2}\right)$. Obtain the derivative of d_{min} as follows

$$\begin{aligned}
\frac{\mathrm{d}d_{min}}{\mathrm{d}\mu_2} &= 2\left(1 + \frac{1}{m_1} \times \frac{1}{2} \times (-2m_2\mu_2)\frac{1}{\sqrt{E - m_2\mu_2^2}}\right) \\
&= 2\left(1 - \frac{m_2\mu_2}{m_1\sqrt{E - m_2\mu_2^2}}\right)
\end{aligned}$$

$$= \quad 2\left(1+\frac{m_2\mu_2}{m_1\mu_1}\right). \tag{68}$$

Therefore d_{min} reaches its maximum when $\frac{\mathrm{d}d_{min}}{\mathrm{d}\mu_2}=0$, i.e., $m_1\mu_1+m_2\mu_2=0$. □

Immediately following Lemma 5.3.1 with $E=m_1\mu_1^2+m_2\mu_2^2$, I have the lemma below.

LEMMA 5.3.2. *Denote E as the energy of a codeword in BPM$[m_1,m_2]$, then the optimal allocation for μ_1,μ_2 is*

$$\begin{cases} \mu_1=-\sqrt{E}\sqrt{\frac{m_2}{m_1(m_1+m_2)}} \\ \mu_2=\sqrt{E}\sqrt{\frac{m_1}{m_2(m_1+m_2)}} \end{cases}. \tag{69}$$

Now I compare BPM$[m_1,m_2]$ (using the optimal allocation just derived) with BPSK for a given pair of m_1,m_2. For a scheme using BPSK, each component of a vector is represented by a BPSK symbol. Again the energy of a vector is E. It has B bits of information and therefore B components. Then the magnitude A of each BPSK component is $A=\sqrt{\frac{E}{B}}$. The minimum distance between two vectors, denoted as $d_{min}(BPSK)$, is $2A$, given that a BPSK component is either A or $-A$. As in the proof of Lemma 5.3.1, the minimum distance between two BPM$[m_1,m_2]$ codewords, denoted as $d_{min}(BPM[m_1,m_2])$ here, is $2\left|\mu_2-\mu_1\right|$. Based on Eq. (69), I can calculate this distance for the optimal BPM scheme given m_1,m_2. Then I can show which scheme is better by comparing the two d_{min}'s.

LEMMA 5.3.3. *BPM$[m_1,m_2]$ yields better SER performance than BPSK when m_1,m_2 satisfy the following condition,*

$$\frac{d_{min}^2(BPM[m_1,m_2])}{d_{min}^2(BPSK)}=\frac{(m_1+m_2)}{2m_1m_2}\left\lfloor\log_2\frac{(m_1+m_2)!}{m_1!\,m_2!}\right\rfloor>1. \tag{70}$$

From the system design perspective, more often, the number of information bits B is given as a parameter or requirement rather than m_1,m_2. For a given B, it is not difficult to see that there are multiple pairs of m_1,m_2 satisfying Eq. (63). Now I prove the lemma on its solution region as below.

LEMMA 5.3.4. *The m_1, m_2 solutions of Eq. (63) for any given B are contained in regions symmetric with respect to the line $m_1 = m_2$ on the m_1, m_2-plane. Among such regions, those not above $m_1 = m_2$ ($m_1 \geq m_2$) are described by the following inequalities,*

$$\begin{cases} m_2 = 1, & 2^B - 1 \leq m_1 < 2^{B+1} - 1, \\ m_2 > 1, & m_1 < 2^{B+2-m_2} - 1. \end{cases} \tag{71}$$

PROOF. It is easily seen by observing Eq. (63), that m_1 and m_2 are interchangeable in the equation, and therefore, the solution regions are symmetric with respect to the line $m_1 = m_2$ on the m_1, m_2-plane. Then without loss of generality, I consider the case when $m_1 \geq m_2$. Eq. (63) can be rewritten as

$$2^B \leq \frac{(m_1 + m_2)!}{m_1!\, m_2!} < 2^{B+1}. \tag{72}$$

I prove this lemma in the following two cases.

Case I : $m_2 = 1$. Now Eq. (72) becomes $2^B \leq m_1 + 1 < 2^{B+1}$. Subtracting 1 from all the operands, I have $2^B - 1 \leq m_1 \leq 2^{B+1} - 2$, which is the first pair of inequality in Eq. (71).

Case II : $m_2 > 1$. Let $\mathcal{R} = \prod_{i=2}^{m_2} (1 + \frac{m_1}{i})$, then

$$\frac{(m_1 + m_2)!}{m_1!\, m_2!} = \frac{(m_1 + m_2) \cdots (m_1 + 2) \times (m_1 + 1)}{m_2 \cdots 2 \times 1} = \mathcal{R} \times (m_1 + 1). \tag{73}$$

Again $m_1 \geq m_2 \geq 2$, I have $\mathcal{R}(m_1, m_2) \geq \prod_{i=2}^{m_2} (1 + 1) = 2^{m_2 - 1}$. Additionally, $\frac{(m_1+m_2)!}{m_1!\, m_2!} < 2^{B+1}$, which is required by Eq. (72). Thus following Eq. (73), I have

$$(m_1 + 1) = \frac{1}{\mathcal{R}} \times \frac{(m_1 + m_2)!}{m_1!\, m_2!} < \frac{1}{2^{m_2 - 1}} \times 2^{B+1} = 2^{B+2-m_2}. \tag{74}$$

Thus, I have obtained $m_2 > 1, m_1 < 2^{B+2-m_2} - 1$, which is the second pair of inequality in Eq. (71).

□

THEOREM. *For any BPM$[m_1, m_2]$ scheme carrying exactly B bits of information and no more, it yields the optimal SER performance for $m_1 = 2^B - 1, m_2 = 1$ or $m_1 = 1, m_2 = 2^B - 1$.*

PROOF. As the regions containing all m_1, m_2 pairs are symmetric with regard to the line $m_2 = m_1$ in the m_1, m_2-plane (Lemma 5.3.4), without loss of generality, I investigate the case when $m_1 \geq m_2$.

First of all, I fix $m_2 = 1$. Consider two different values m_1 and $m_1^{'}$ where $2^B - 1 \leq m_1 < m_1^{'} < 2^{B+1} - 1$. Based on Eq. (69), I know the optimum values of μ_1, μ_2 respectively for BPM$[m_1, m_2]$ and BPM$[m_1^{'}, m_2]$. Then

$$\begin{aligned} \frac{d_{min}^2(BPM[m_1, m_2])}{d_{min}^2(BPM[m_1^{'}, m_2])} &= \frac{(m_1+1)^2}{m_1 + m_1^2} \times \frac{m_1^{'} + (m_1^{'})^2}{(m_1^{'}+1)^2} \\ &= \frac{m_1^{'}+1}{m_1^{'}} \times \frac{m_1}{m_1+1} \\ &> \frac{m_1+1}{m_1} \times \frac{m_1}{m_1+1} \\ &= 1. \end{aligned} \tag{75}$$

Thus, for $m_2 = 1$ and a given B, the smallest valid m_1 yields the largest d_{min}, and furthermore the SER. In this case, $\min(m_1) = 2^B - 1$.

Inside the rest of the solution region ($1 < m_2 <= m_1 < 2^B - 1$), suppose a certain BPM$[m_1, m_2]$ has the largest d_{min}. If it yields a larger d_{min} than BPM$[2^B - 1, 1]$, then this particular BPM$[m_1, m_2]$ is the optimal scheme; otherwise, all m_1, m_2 pairs have a smaller d_{min} than BPM$[2^B - 1, 1]$, and BPM$[2^B - 1, 1]$ is the optimal scheme to be found. Now I proceed to perform the comparison. Again $1 < m_2 <= m_1 < 2^B - 1$, then

$$\begin{aligned} \frac{d_{min}^2(BPM[2^B-1, 1])}{d_{min}^2(BPM[m_1, m_2])} &= \frac{2^B}{2^B-1} \times \frac{m_1 m_2}{m_1 + m_2} \\ &= \frac{2^B}{2^B-1} \times \frac{1}{\frac{1}{m_1} + \frac{1}{m_2}} \quad \left(\because \frac{1}{m_1} \leq \frac{1}{m_2} \leq \frac{1}{2}, \frac{1}{m_1} + \frac{1}{m_2} \leq 1 \right) \\ &\geq \frac{2^B}{2^B-1} \\ &> 1. \end{aligned} \tag{76}$$

Thus the optimal scheme is BPM$[2^B - 1, 1]$ when $m_1 \geq m_2$. Due to the symmetry of solution regions, I also have BPM$[1, 2^B - 1]$, and therefore completing the proof. □

COROLLARY. *There exists at least one pair of binary permutation codes which can be used in the physical-layer network coding, such that its end-to-end error probability is smaller than the corresponding BPSK scheme under any given SNR.*

PROOF. Denote P_{ma}, P_{bc} as the error probabilities of the MA and BC phases respectively, then the end-to-end error rate can be expressed as $P_{ete} = P_{ma}(1 - P_{bc}) + (1 - P_{ma})P_{bc}$. I can analyze the overall SER based on the individual performances of the MA, BC phases. Lemmas 5.3.1 to 5.3.4 and the Theorem apply to the BC phase. According to [32, 6], the theoretical evaluation of the one-to-one transmission serves as a tight upper bound for the MA phase. Therefore, based on the lemmas and the theorem, for a given B required by the system, I can always find two BPMs that both perform better than BPSK for the PLNC using the proposed method I (cf. Section 5.2). Additionally, the optimal BPM$[2^B - 1, 1]$ or BPM$[1, 2^B - 1]$ can be used in PLNC method II. □

5.4. Numerical Results

In this section, I present several numerical results. The signal-to-noise ratio (SNR) is defined as the energy of a vector/codeword over the noise energy.

Fig. 5.7 shows a scatter plot of ratios of d^2_{min} by plugging in different m_1, m_2 values into Eq. (70). Those marked by ◯ have a ratio of 1, which potentially indicates equal error probabilities based on the Maximum Likelihood criterion. Markers × are those whose ratio is smaller than 1. In this case, the BPM scheme has worse error performance than its BPSK counterpart. As can be seen in the figure, most of the BPMs appear more advantageous than the BPSK modulation. □ markers show these beneficial BPM schemes. A larger d_{min} results in greater separation between the constellations under the same symbol power constraint, and furthermore reduced error probability. For this reason, I use four different colors and sizes to label the BPMs falling into different ratio ranges. A larger size marker represents a larger ratio. As marked out, BPM$[1, 2]$/BPM$[2, 1]$ ($B = 1$) has worse performance than BPSK. BPM$[1, 1]$ ($B = 1$) is in fact BPSK for $B = 1$. For $B = 2$, BPM$[2, 2]$ ($B = 2$) is

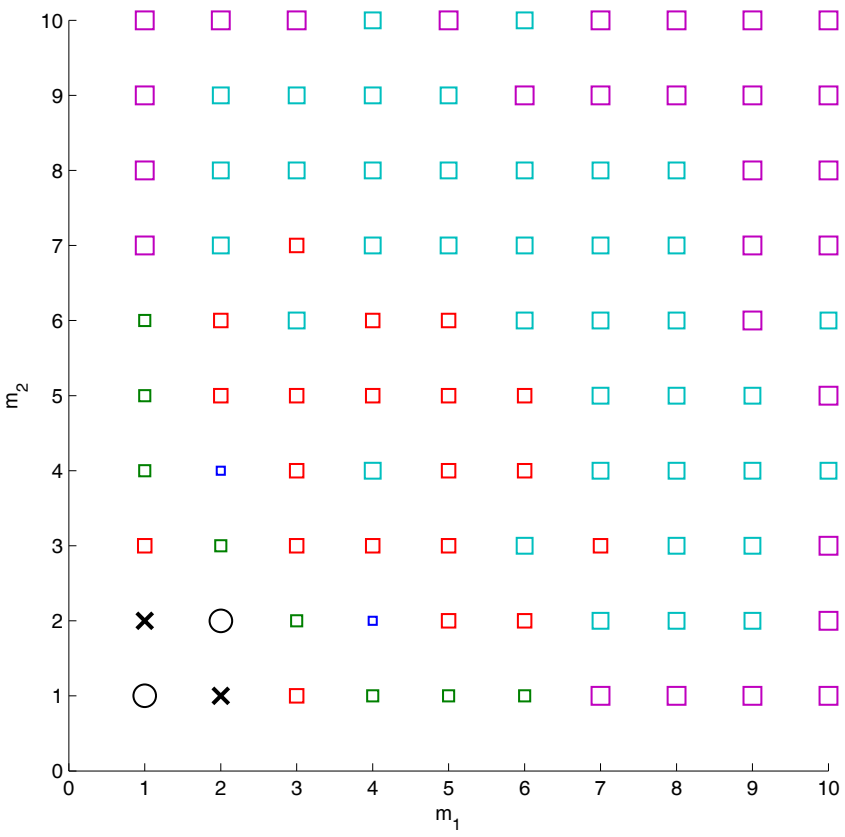


FIGURE 5.7. The scatter plot of the ratio $\frac{d^2_{min}(PM[m_1,m_2])}{d^2_{min}(BPSK)}$ (◯: ratio is 1; ×: ratio < 1; □: ratio > 1, a larger size marker indicates a larger value).

a tie with BPSK while the optimal BPM[1, 3] is still better than BPSK. This supports the corollary in Section 5.3.

The introduction of various marker shapes and sizes in fact brings in a third dimension to Fig. 5.7. Now taking the values along the line of $m_1 = m_2$ on the m_1, m_2-plane, I have Fig. 5.8. It shows that the ratio value does not always increase with m_1 or m_2 (in the case of this figure, $m_1 = m_2$). It is a result of multiple factors. For example, at $m_1 = m_2 = 5$, there is a dip on the increasing curve. By calculation, I find that, there are 70 codewords in BPM[4,4] and 64 of them, which takes up 91% of the set, are used to support a symbol of $B = 6$ bits of information. For BPM[5,5] where $B = 7$, only 128, that is 50% of the 252 codewords is used for coding and modulation. As for BPM[6,6], this percentage is higher, 55%, from 512 out of 924 codewords. The performance dent of BPM[5,5] is partially due to

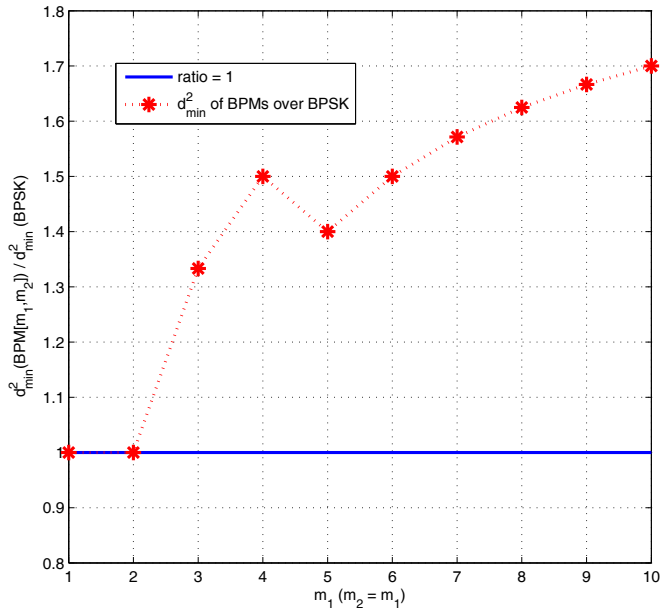


FIGURE 5.8. The plot of the ratio $\frac{d^2_{min}(PM[m_1,m_2])}{d^2_{min}(BPSK)}$ for $m_1 = m_2$.

the inefficient use of its codewords. I can look at this issue from another perspective. For BPM[5,5], I compute $\log_2 M = 7.98$, which is almost 8. However, in reality, only 7 bits can be accommodated, and effectively, the power which has been allocated for that additional 0.98 unit is unusable. This aspect only partially contributes to the performance difference. The error probability is the result of the interplay of various factors. However, it is not the focus of this chapter. I do not get into more details on this.

Fig. 5.9 is obtained by interpolation on the values in Fig. 5.7. The color/shade on the grids indicate the magnitude of the ratio value. The contour plot shows the trend of the ratio change varying m_1, m_2. It forms a radiant-like pattern centering on $m_1 = m_2 = 1$ in the first quadrant of the plane. It is shown that the BPMs continue to gain more advantage over the traditional BPSK asymptotically. However, with the increase of the order of BPMs, more memory is demanded of the transceivers to accommodate a symbol, as it requires all the components of the symbol to be received in order to perform the decoding operations.

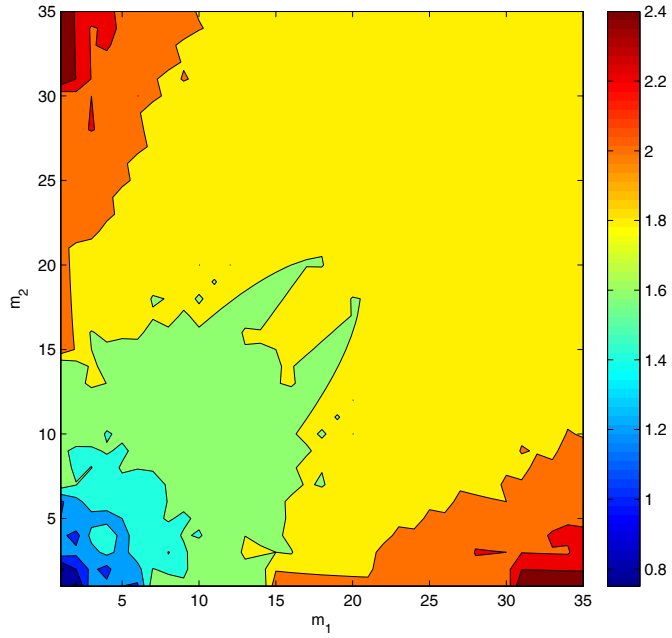


FIGURE 5.9. The contour plot of the ratio $\frac{d^2_{min}(PM[m_1,m_2])}{d^2_{min}(BPSK)}$.

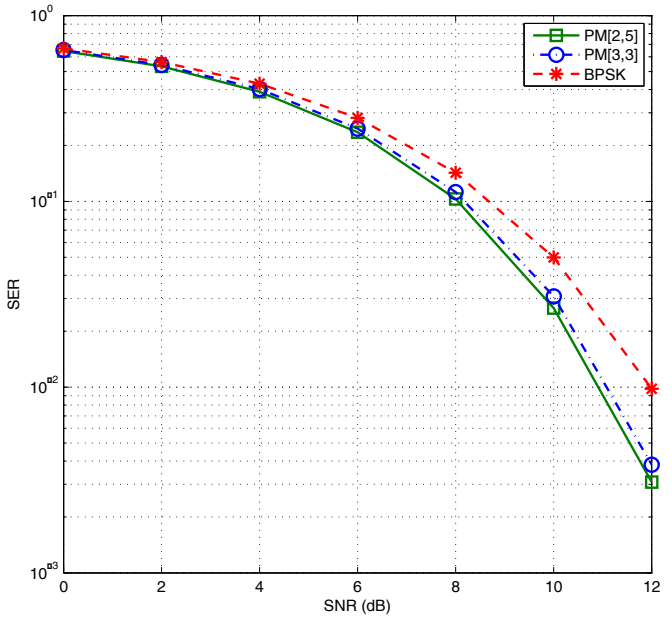


FIGURE 5.10. The SER performance comparison between BPM[3,3], BPM[2,5] and BPSK in one-to-one transmissions for $B = 4$.

In Fig. 5.10, the SER performance is compared between BPSK and two BPMs yielding the information capacity $B = 4$. The figure shows that the BPMs outperform BPSK in the one-to-one transmission. This is consistent with the conclusion in the proved theorem. Also, the

ratio $d^2_{min}(BPM[3,3])/d^2_{min}(BPSK) = 1.33$, $d^2_{min}(BPM[2,5])/d^2_{min}(BPSK) = 1.40$. Due to the coarse granularity of the marker size, this difference is not well reflected in Fig. 5.7. The larger d_{min} of BPM[2,5] lends itself to a smaller error probability, which confirms the previous discussions on the effect of the minimum distance between codewords on the error probability based on the ML decoding criterion.

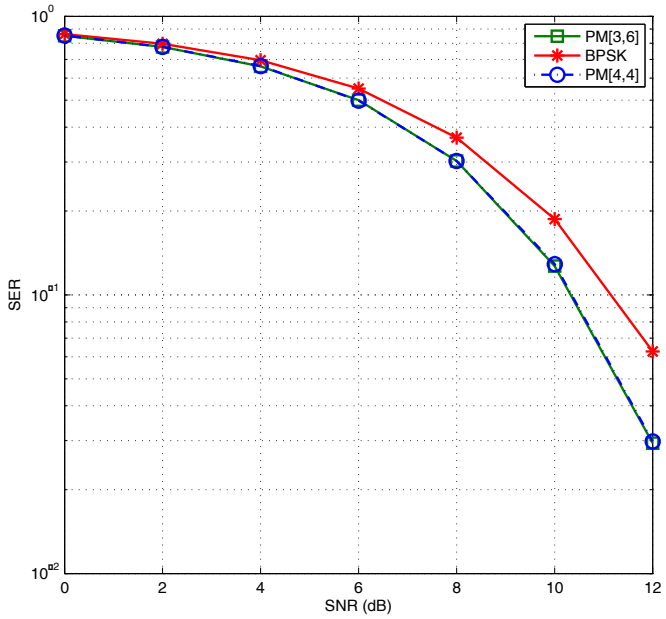


FIGURE 5.11. The SER performance comparison between BPM[4,4], BPM[3,6] and BPSK in one-to-one transmissions for $B = 6$.

Similarly to Fig. 5.10, the SER performance is compared between BPSK and two BPMs yielding the information capacity $B = 6$ in Fig. 5.11. This is another case of the numerical support of the prove theorem. Note that, the ratio $d^2_{min}(BPM[4,4])/d^2_{min}(BPSK) = 1.50$, $d^2_{min}(BPM[3,6])/d^2_{min}(BPSK) = 1.50$. The equal d_{min} is reflected in the figure as perfectly overlapping curves of BPM[4,4] and BPM[3,6].

Fig. 5.12 shows the effect of choosing the codewords on the error probability with the example of a one-to-one transmission using BPM[2,2]. For BPM[2,2], $M = 6, B = 2$, only 4 out of the 6 vectors generated in this code are used for modulation. For the optimal selection, I propose to choose those which collectively yield the maximized sum of d^2_{min}. It is a straightforward strategy. The random selection is marked by ◯ and the optimal selection

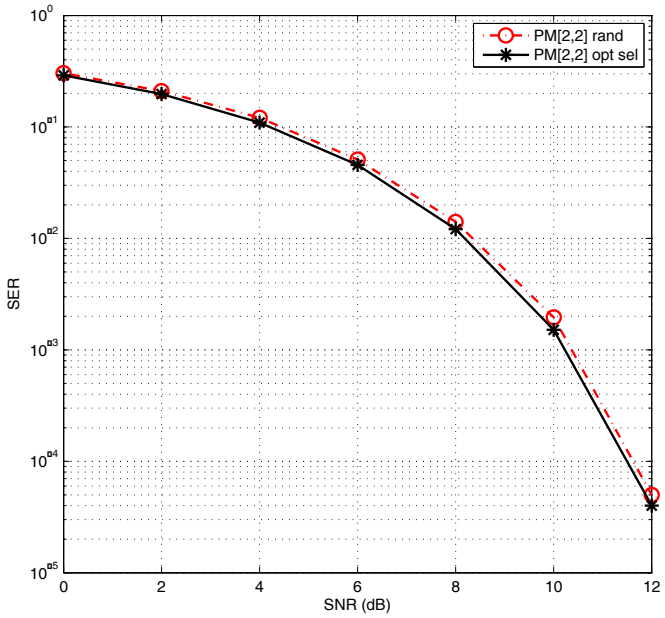


FIGURE 5.12. The SER performance comparison between the random codeword selection and the selection based on the maximized distance sum of BPM[2,2] (choosing 4 codewords out of 6).

by $*$. The figure shows that such selection does not significantly improve the error rate. As the order of BPM becomes large, the complexity of such selection method explodes as all combinations take part in a pairwise comparison. Due to the moderate benefit, I did not seek to incorporate the optimal selection into PLNC incurring potentially large computation overhead.

Fig. 5.13 compares the performance of BPM[5,5] using the fast decoding and the ML decoding. The fast decoding is discussed in Section 5.2.2. The coding procedure is performed by replacing the largest k_2 components of the received vector with μ_2, and the remaining k_1 components with μ_1. However, when $B < \log_2 M$, this method results in invalid codewords as some codewords have been purged from the actual transmission. For BPM[5,5], $B = 7 < \log_2 M = \log_2 252 = 7.98$. The fast decoding is shown to yield inferior performance compared to the ML decoding.

The SER of PLNC method I, using BPM[3,3] for the MA phase and BPM[4,4] for the BC phase, is compared with using BPSK in Fig. 5.14. In BPM[3,3], $B = 4, n = 6$. After

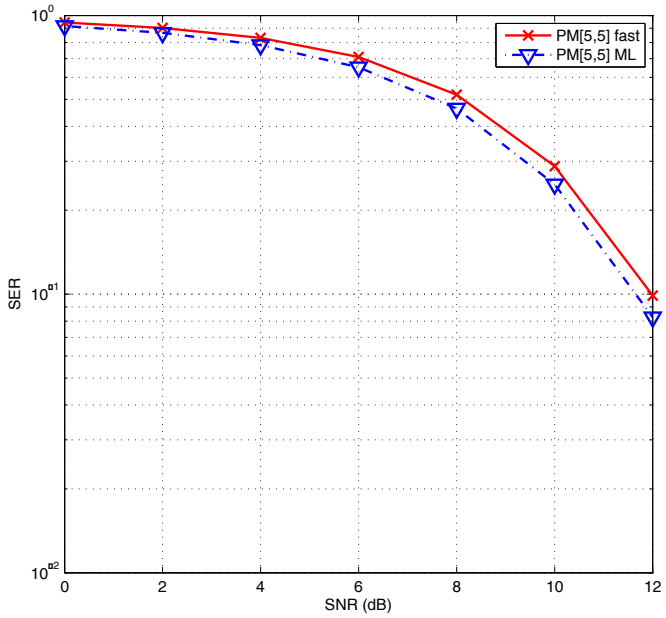


FIGURE 5.13. The SER performance comparison between the fast decoding and the ML based decoding for BPM[5,5].

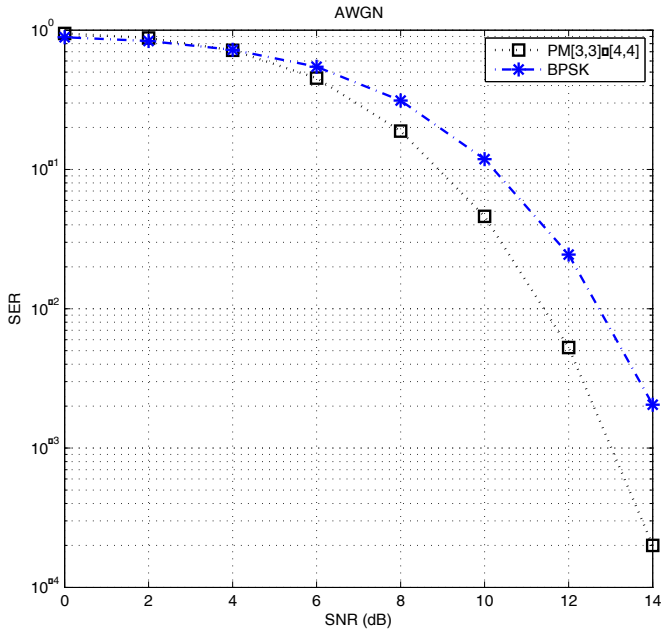


FIGURE 5.14. The SER performance comparison of PLNC using BPSK and BPM method I (BPM[3,3] in MA, BPM[4,4] in BC) for $B = 3$.

receiving all the $n = 6$ components of a vector in the MA phase, the relay decodes the components based on the decision rules described in Eq. (66). For the BC phase, the relay

chooses a modulation scheme of a higher order, BPM[4,4] to accommodate $n = 6$ bits instead of $B = 4$ in the MA phase. Note that, the total power allocated for a single codeword for BPM[4,4] is still the same as BPM[3,3]. PLNC method I (BPM[3,3] for the MA phase and BPM[4,4] for the BC phase) is shown to achieve better performance compared to BPSK.

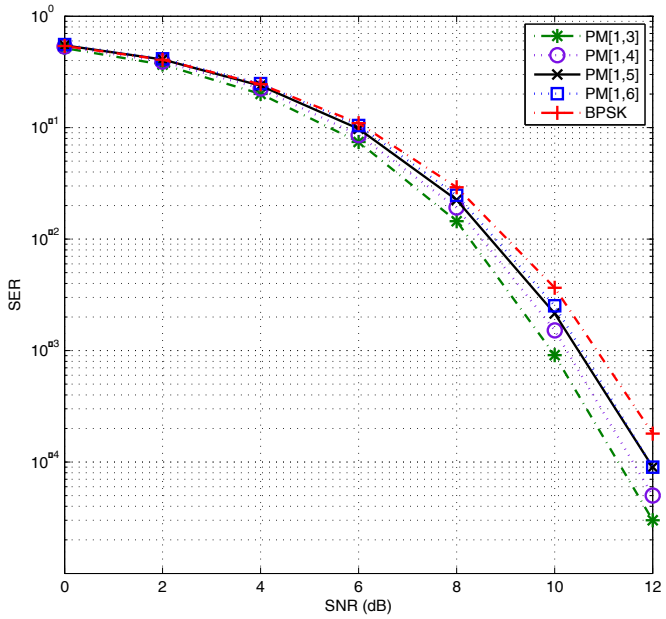


FIGURE 5.15. The SER performance comparison of PLNC using BPSK and BPM method II for $B = 2$.

As for PLNC method II, Fig. 5.15 shows several BPMs compared with BPSK. The plots confirm the optimality (based on the Theorem) of BPM$[1, 3]$ when $B = 2$. I see that as d_{min} increases (d_{min}(BPSK) $<$ d_{min}(BPM[1,6]) $< \cdots <$ d_{min}(BPM[1,3]), refer to Fig. 5.7), the SER becomes smaller under the same SNR. The same comparison is performed in Fig. 5.15 for $B = 3$. Again, the optimal (based on the Theorem) BPM$[1, 7]$ is confirmed by yielding the best error probability. I can also see that given the same SNR, as d_{min} increases (d_{min}(BPSK) $<$ d_{min}(BPM[1,14]) $<$ d_{min}(BPM[1,11]) $<$ d_{min}(BPM[1,7]), refer to Fig. 5.7), the SER becomes smaller.

A comparison to similar Fig. 5.15 is made for the case of $B = 3$ in Fig. 5.16. Again, I can see the proposed PLNC method II outperforms the traditional approach using BPSK for transmitting every bit individually.

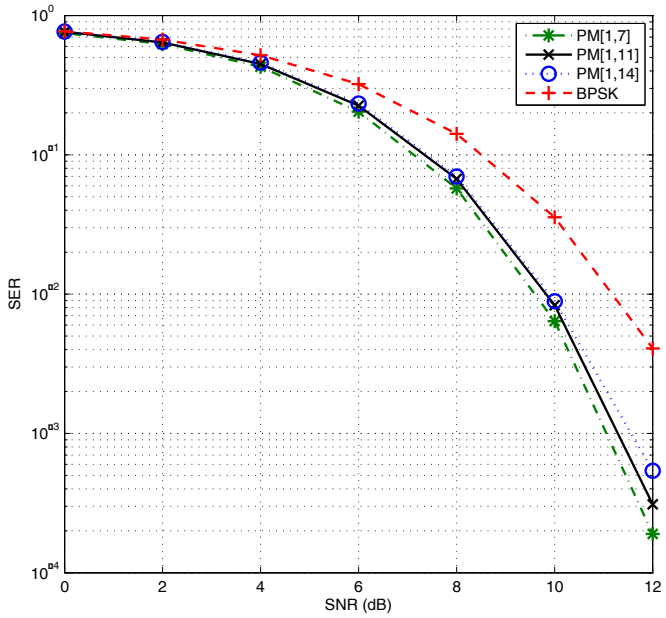


FIGURE 5.16. The SER performance comparison of PLNC using BPSK and BPM method II for $B = 3$.

5.5. Summary

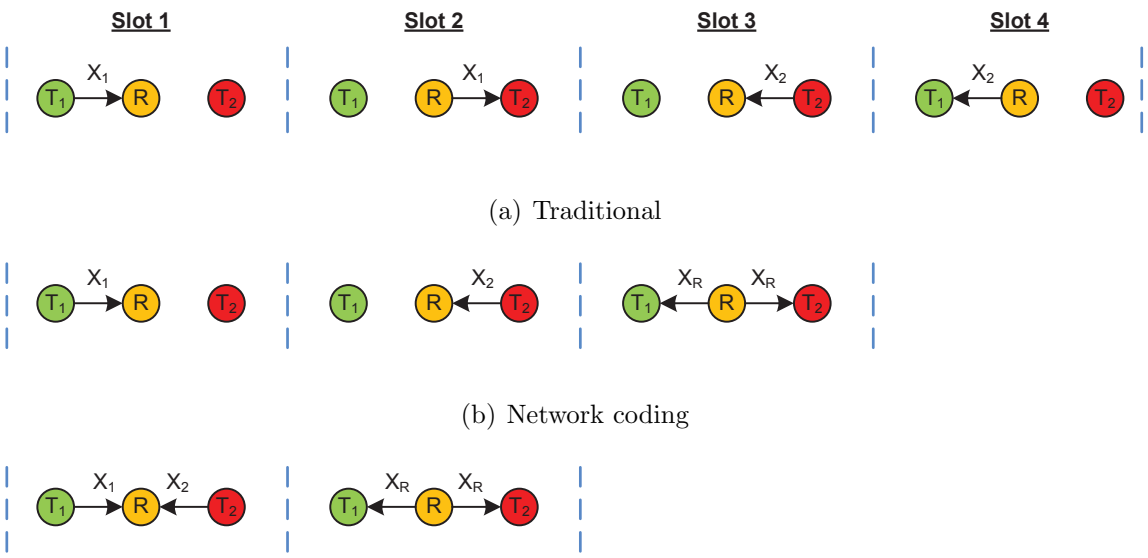
In this chapter, I introduced a novel method of the physical-layer network coding by utilizing the binary permutation modulations instead of the traditional modulation types. I compared the end-to-end performance of the proposed scheme with the traditional approach using BPSK, as BPSK yields the smallest error probability among the conventional modulation types. Both the theoretical analysis in the explicit analytic forms and the numerical results were given. In particular, I proved a series of lemmas and a theorem on the performance and optimality of the proposed scheme. It is shown that the proposed class of permutation codes outperforms the conventional approach using BPSK in the physical-layer network coding.

CHAPTER 6

ON THE PERFORMANCE OF TWO-WAY RELAY CHANNELS USING SPACE-TIME CODES

In this chapter, I investigate the application of the space-time codes in conjunction with the physical-layer network coding in the context of the two-way relay channels. In particular, I extend the decode-and-forward (DF) protocol into the aforementioned system settings, and discuss the encoding and decoding operations, and derive the theoretical bounds for its end-to-end error probability performance. Additionally, two other protocols namely amplify-and-decode (AF), and partial-decode-and-forward (PDF), are extended for comparison. I present numerical results confirming the advantage of the proposed DF protocol over AF and PDF.

The organization of the chapter is as follows. First the background and motivation is provided in Section 6.1. Section 6.2 gives detailed description of the model of the single-antenna and the space-time coded systems, and the relaying protocols including AF, PDF and DF. Theoretical performance analysis is performed in Section 6.3 where upper bounds of the end-to-end error probabilities are derived for both DF schemes using BPSK and QPSK. The numerical results are presented in Section 6.4, followed by the summary in Section 6.5.



(a) Traditional

(b) Network coding

(c) Physical-layer network coding

FIGURE 6.1. The time scheduling of the two-way relay systems: (a) the conventional four-slot scheme, (b) the three-slot scheme with network coding, (c) the two-slot scheme with the physical-layer network coding.

The two-way relay channel has been a widely investigated topic due to its representation of various types of systems and compact abstraction of large-scale networks. Here let's briefly revisit the two-way relay system. The two terminals namely T_1, T_2 wish to exchange information. However, no communication link is available for direction transmission. Therefore, a single relay is needed in between to assist this process, and messages are delivered in a two-hop fashion. The evolution of time and bandwidth efficient transmission protocols for such systems are shown in Fig. 6.1. The traditional way takes four time slots to complete a single round of two-way exchange in Fig. 6.1(a). In this approach, the terminals alternate in occupying the relay as the receiver. Simultaneous transmission is prohibited to avoid collision and interference. With the help of network coding [32, 91, 92, 93], such interference can be potentially embraced and turned into throughput increase. By performing bitwise algebraic operations on the received symbols, the relay obtains a new symbol with information

from two parties encapsulated, and forwards it in a single take. Using such an approach, the total time consumption for a single round is reduced to three time slots in Fig. 6.1(b). If the network coding techniques are applied at the physical layer, the terminals are enabled to participate in communication with the relay at the same time in Fig. 6.1(c), which further boosts the system throughput [32].

Efforts have been made by the researchers in other design aspects to improve the network performances. An amplify-and-forward (AF) protocol, one of the methods concerned with the processing of the received symbols at the relay, scales the entire received signal to its own transmission power, or in other word, amplifies it [94]. The noise is also amplified in this process, and passed on during the next stage transmission [95]. A decode-and-forward (DF) relaying attempts to decode the symbols or a combination of the symbols before forwarding them [94, 95]. The noise is possibly eliminated while processing complexity is incurred at this point and incorrect detection is also unavoidable.

Recent research has attempted to incorporate multiple antennas into such systems, or create collaboration between multiple relays such that space-time diversities can be taken advantage of. As discussed in Chapter 2, such diversity scheme improves the system throughput by combating fading effects and increases reliability [96]. In [23, 26, 78], the authors discussed applying space-time codes in a distributed fashion with single-antenna transceivers. However, the error performance was not addressed, and the channel does not support half-duplex communications. Min et al. studied the outage probability of a relay network where only the source terminal is equipped with multiple antennas [97].

Motivated by these works, I propose using the space-time coding in conjunction with the physical-layer network coding in a two-way relay channel where terminals and the relay are each equipped with two antennas. Specifically, I describe a new class of DF methods, and obtain the performance bounds, and present numerical results to support its advantage over the existing schemes.

6.2. System Model

In this section, I describe the system model and the proposed DF schemes, together with the extension of two existing schemes AF and PDF [23] for comparison.

6.2.1. DF Strategy for Single-Antenna TWRCs

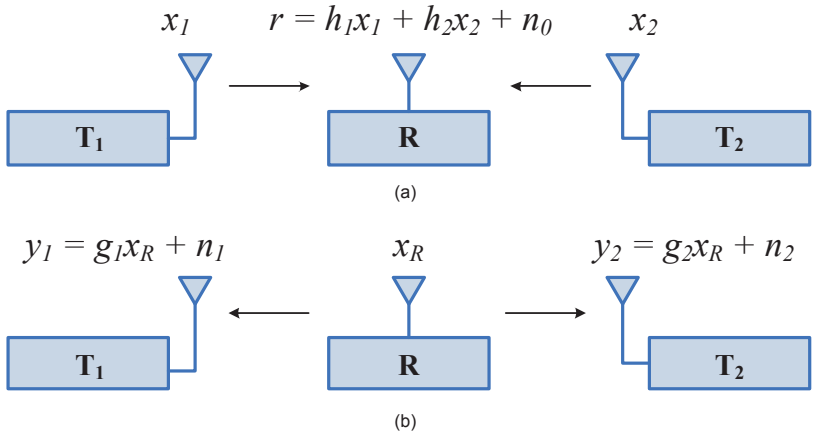


FIGURE 6.2. Communication over a single-antenna TWRC.

An illustrating example of communication over the single-antenna TWRC system employing DF relaying is displayed in Fig. 6.2. With the assistance of a single relay R, information is exchanged between two terminals, T_1 and T_2 with no direct link available, which is completed within two transmission phases [25]. During phase I as shown in Fig. 6.2(a), two symbols x_1 and x_2 with energy $\sqrt{E_s}$ per symbol from T_1 and T_2 respectively are transmitted simultaneously to the relay. The received symbol is

$$r = h_1x_1 + h_2x_2 + n_0, \tag{77}$$

where both h_i and n_0 complex random variables: h_i is the fading channel coefficient from T_i to the relay, for $i = 1, 2$, and n_0 is the AWGN at the relay with zero mean and variance of σ_0^2 per dimension. For each transmission, the receiver is assumed to have perfect channel state information (CSI). Using the maximum likelihood (ML) decoding, the relay chooses a symbol pair $\langle \hat{x}_1, \hat{x}_2 \rangle$, and then encodes them by network coding into x_R with energy $\sqrt{E_R}$

TABLE 6.1. Network coding mapping scheme for the BPSK modulation.

Received Symbols		**Encoded Symbol**
$\hat{\boldsymbol{x}}_{\boldsymbol{1}}$	$\hat{\boldsymbol{x}}_{\boldsymbol{2}}$	$\boldsymbol{x_R}$
-1	-1	-1
-1	1	1
1	-1	1
1	1	-1

per symbol. In phase II, R broadcasts x_R to T_1 and T_2, as shown in Fig. 6.2(b). As a simple example of network coding, let $E_s = E_R = 1$, and the mapping for BPSK modulation is given in Table 6.1, in which the encoder does not distinguish $\langle 1, 1\rangle$ from $\langle -1, -1\rangle$, nor $\langle -1, 1\rangle$ from $\langle 1, -1\rangle$. Based on the same idea, a possible mapping for QPSK (Gray coding) is given in Table 6.2 (For notational simplicity, $E_s = E_R = 2$).

The received symbols at the terminal sides are

$$\begin{cases} y_1 = g_1 x_R + n_1 \\ y_2 = g_2 x_R + n_2, \end{cases} \tag{78}$$

where complex random variable g_i is the fading channel coefficient from the relay to T_i, and n_i is the AWGN at T_i with zero mean and variance of σ_i^2 per dimension, for $i = 1, 2$. For notational simplicity, I assume the fading channel is reciprocal for both directions, i.e., $h_1 = g_1$, $h_2 = g_2$, and all noises share the same variance σ^2, i.e., $\sigma^2 = \sigma_0^2 = \sigma_1^2 = \sigma_2^2$. It does not affect the results of error probability derivation without such assumptions, merely to make the intermediate expression more complex. As can be seen from Section 6.3 that derivations for phases I and II are separate procedures which do not rely on channel coefficients being the same for both directions, and the error probabilities obtained are a function of the signal-to-noise ratio (SNR) rather than the signal or noise power alone. Each terminal decodes the received symbol according to the ML criterion and extracts the symbol forwarded by the relay with its own. The end-to-end error probability refers to the rate at which symbols transmitted by the terminal on one side are not correctly extracted by the other side.

TABLE 6.2. Network coding mapping scheme for the QPSK modulation.

Received Symbols		**Encoded Symbol**
$\hat{\boldsymbol{x}}_{\mathbf{1}}$	$\hat{\boldsymbol{x}}_{\mathbf{2}}$	$\boldsymbol{x}_{\boldsymbol{R}}$
$-1-i$	$-1-i$	$-1-i$
$-1-i$	$-1+i$	$-1+i$
$-1-i$	$1-i$	$1-i$
$-1-i$	$1+i$	$1+i$
$-1+i$	$-1-i$	$-1+i$
$-1+i$	$-1+i$	$-1-i$
$-1+i$	$1-i$	$1+i$
$-1+i$	$1+i$	$1-i$
$1-i$	$-1-i$	$1-i$
$1-i$	$-1+i$	$1+i$
$1-i$	$1-i$	$-1-i$
$1-i$	$1+i$	$-1+i$
$1+i$	$-1-i$	$1+i$
$1+i$	$-1+i$	$1-i$
$1+i$	$1-i$	$-1+i$
$1+i$	$1+i$	$-1-i$

6.2.2. DF for Space-Time Coded TWRCs

Now I consider the half-duplex space-time coded system, displayed in Fig. 6.3. Terminals T_i $(i = 1, 2)$ and the relay are each equipped with two antennas which cannot transmit and receive at the same time. Space-time coding is applied at both terminals and the relay to improve transmission quality. The whole process of communication can be further broken down to two phases each accounting for two time slots. Phase I consists of two streams of Alamouti scheme, as illustrated in Fig. 6.3(a). Let x_{ij} denote the symbol transmitted by the jth antenna of T_i, $i, j = 1, 2$, in the first time slot of phase I, then in the second time slot T_i transmits $-x_{i2}^*$ from its first antenna, and x_{i1}^* from its second antenna simultaneously to the

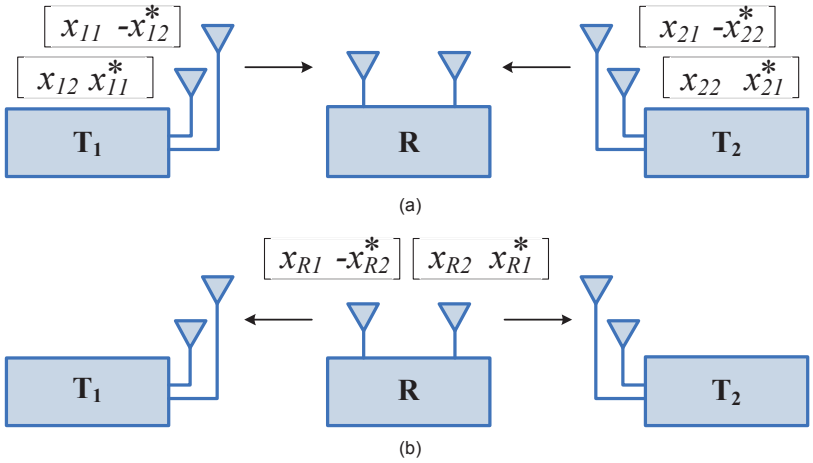


FIGURE 6.3. Communication over TWRC with space-time codes.

relay, where $*$ is the complex conjugate operation. The symbols received at the relay can be written as

$$\begin{bmatrix} r_{1,1} \\ r_{1,2}^* \\ r_{2,1} \\ r_{2,2}^* \end{bmatrix} = \sum_{i=1}^{2} \begin{bmatrix} h_{i1:1} & h_{i2:1} \\ h_{i2:1}^* & -h_{i1:1}^* \\ h_{i1:2} & h_{i2:2} \\ h_{i2:2}^* & -h_{i1:2}^* \end{bmatrix} \begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix} + \begin{bmatrix} z_{1,1} \\ z_{1,2}^* \\ z_{2,1} \\ z_{2,2}^* \end{bmatrix}, \tag{79}$$

where $r_{i,t}$ is the received symbol at the relay's ith antenna during the tth time slot, $h_{ij:k}$ is the fading channel coefficient from T_i's jth antenna to the relay's kth antenna, x_{ij} is the symbol transmitted from T_i's jth antenna during the first time slot, and $z_{i,t}$ is the AWGN at the relay's ith antenna during the tth time slot.

The relay decodes the received symbol using ML criterion to obtain $\langle \hat{x}_{11}, \hat{x}_{12}, \hat{x}_{21}, \hat{x}_{22} \rangle$. Then, network coding is performed on $\langle \hat{x}_{11}, \hat{x}_{21} \rangle$ and $\langle \hat{x}_{12}, \hat{x}_{22} \rangle$ respectively to get x_{R1} and x_{R2}, which the relay broadcasts to both terminals using Alamouti scheme during phase II, as shown in Fig. 6.3(b). Assuming the channel varies slowly enough, fading remains constant

within the two transmission phases. Thus, at the end of phase II I have at T_i

$$\begin{bmatrix} y_{i1,1} \\ y^*_{i1,2} \\ y_{i2,1} \\ y^*_{i2,2} \end{bmatrix} = \begin{bmatrix} h_{i1:1}\ h_{i2:1} \\ h^*_{i2:1} - h^*_{i1:1} \\ h_{i1:2}\ \ h_{i2:2} \\ h^*_{i2:2} - h^*_{i1:2} \end{bmatrix} \begin{bmatrix} x_{R1} \\ x_{R2} \end{bmatrix} + \begin{bmatrix} z_{i1,1} \\ z^*_{i1,2} \\ z_{i2,1} \\ z^*_{i2,2} \end{bmatrix}. \tag{80}$$

Then T_i performs ML decoding and extracts transmitted symbols from the other terminal with its own symbols.

6.2.3. AF for Space-Time Coded TWRCs

Now I extend the AF protocol introduced in [23] to the space-time coded two-way relay network. The first phase of communication is carried out the same way as DF and PDF, as is discussed in Section 6.2.4. Note that these relaying strategies are primarily concerned with the processing of the received symbols at the relay, and can potentially make an impact on the next phase transmission but not the phase before. In this case, Eq. (79) is also applicable to AF.

As mentioned in Section 6.1, the relay amplifies the received signal without decoding the originally transmitted symbols. The scale of the amplification is given by

$$\alpha_{AF} \equiv \sqrt{\frac{E_R}{2\left(2E_s + N_0\right)}}\,. \tag{81}$$

The authors in [23] proposed a set of unit power precoding matrices to be applied on the received signals before the next stage. These precoding matrices in Eq. (82), result in the minimum pairwise error probability, and are adopted here too.

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ A_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{82}$$

The symbols transmitted by the relay's ith antenna are formulated as,

$$\begin{bmatrix} x_{Ri,1} \\ x_{Ri,2} \end{bmatrix} = \alpha_{AF} A_i \begin{bmatrix} r_{i,1} \\ r_{i,2} \end{bmatrix}. \tag{83}$$

Thus at the receiving T_i, the received symbols are as below,

$$\begin{bmatrix} y_{i1,1} \\ y_{i1,2} \\ y_{i2,1} \\ y_{i2,2} \end{bmatrix} = \begin{bmatrix} h_{i1:1} & 0 & h_{i1:2} & 0 \\ 0 & h_{i1:1} & 0 & h_{i1:2} \\ h_{i2:1} & 0 & h_{i2:2} & 0 \\ 0 & h_{i2:1} & 0 & h_{i2:2} \end{bmatrix} \begin{bmatrix} x_{R1,1} \\ x_{R1,2} \\ x_{R2,1} \\ x_{R2,2} \end{bmatrix} + \begin{bmatrix} z_{i1,1} \\ z_{i1,2} \\ z_{i2,1} \\ z_{i2,2} \end{bmatrix}, \tag{84}$$

where $z_{ij,t}$ represents the Gaussian noise at the jth antenna of T_i at the end of the tth time slot. Then the ML decoding is performed when the terminal extracts the information from the other side with its own transmitted symbols.

6.2.4. PDF for Space-Time Coded TWRCs

As mentioned at beginning of Section 6.2.3, the first phase transmission is the same regardless of which relaying protocol is used. Also, Eq. (79) holds for PDF. The relay decodes and obtains $\langle \hat{x}_{11}, \hat{x}_{12}, \hat{x}_{21}, \hat{x}_{22} \rangle$. Then a noiseless version of the received signal is reconstructed as

$$\begin{bmatrix} \hat{r}_{i,1} \\ \hat{r}^*_{i,2} \end{bmatrix} = \begin{bmatrix} h_{11:i} & h_{12:i} & h_{21:i} & h_{22:i} \\ h^*_{12:i} & -h^*_{11:i} & h^*_{22:i} & -h^*_{21:i} \end{bmatrix} \begin{bmatrix} \hat{x}_{11} \\ \hat{x}_{12} \\ \hat{x}_{21} \\ \hat{x}_{22} \end{bmatrix}. \tag{85}$$

Now that the total energy as expressed in Eq. (85) are twice as that of the relay's transmission power, $\frac{2(E_s+E_s)}{E_R} = 2$, a fraction $\frac{1}{2}$ is applied to offset it. The precoding matrices in Eq. (82) are also used here. For the next stage, the relay broadcasts the following,

$$\begin{bmatrix} x_{Ri,1} \\ x_{Ri,2} \end{bmatrix} = \frac{1}{2} A_i \begin{bmatrix} \hat{r}_{i,1} \\ \hat{r}_{i,2} \end{bmatrix}. \tag{86}$$

Then in the system using the PDF protocol, the decoding and the symbol extraction are performed in the same way as AF.

6.3. Performance Analysis

In this section, I study the SER performance of DF protocols for the single-antenna and the space-time encoded systems.

6.3.1. SER of Single-Antenna TWRCs

In the single-antenna system, the relay employing the conventional decoding chooses a symbol pair $\langle \hat{x}_1, \hat{x}_2 \rangle$ from the set $\langle \pm\sqrt{E_s}, \pm\sqrt{E_s} \rangle$ so that $h_1 x_1 + h_2 x_2$ has the minimum Euclidean distance from the received symbol r. Assuming that perfect CSI can be obtained at the receiver, the detection variables for x_1 and x_2 can be formulated as

$$\begin{cases} \tilde{s}_1 = h_1^*(r - h_2\tilde{x}_2) \\ \tilde{s}_2 = h_2^*(r - h_1\tilde{x}_1) \ , \end{cases} \tag{87}$$

where $\tilde{x}_i$ is the decision variable of x_i at the relay. The SER of detecting x_i is the same for $i = 1, 2$ due to the symmetry of the transmission system; thus only that of x_1 is derived. Define

$$\begin{aligned} X_1 &= r - h_2\tilde{x}_2 \\ Y_1 &= h_1 \ , \end{aligned} \tag{88}$$

where the complex random variables X_1 and Y_1 are correlated and statistically independent. $\tilde{s}_1$ can be recast as

$$\tilde{s}_1 = X_1 Y_1^* \ . \tag{89}$$

Let $Z_R = Re(\tilde{s}_1)$ and $Z_I = Im(\tilde{s}_1)$, then the envelope R and phase Θ can be expressed as

$$\begin{aligned} R &= \sqrt{Z_R^2 + Z_I^2} \\ \Theta &= \tan^{-1}\left(\frac{Z_I}{Z_R}\right) \ . \end{aligned} \tag{90}$$

According to [98], the probability density function of the joint distribution of R and Θ can be obtained by performing double Fourier transform on the characteristic function of the joint distribution of Z_R and Z_I. The probability density function of Θ can then be obtained by integrating the joint probability density function over R. Given h_2 and correctly decoded $\tilde{x}_2$, the conditional probability density function of Θ can be written as Eq. (91) which is an

even function [98].

$$p(\theta\,|\tilde{x}_2, h_2\,) = \frac{1-|\mu|^2}{2\pi}\left[\frac{1}{1-|\mu|^2\cos^2\theta} + \frac{|\mu|\cos\theta\cos^{-1}(-|\mu|\cos\theta)}{(1-|\mu|^2\cos^2\theta)\sqrt{1-|\mu|^2\cos^2\theta}}\right], \tag{91}$$

where

$$\begin{aligned}\gamma_s &= \frac{E_s}{N_0}\\ \mu &= \sqrt{\frac{\gamma_s}{\gamma_s+1}}\,.\end{aligned} \tag{92}$$

For the BPSK modulation, the error probability of x_1 can be calculated as in Eq. (93),

$$\begin{aligned}P_e^{I,R}\,(\,x_1|\,x_2) &= P(\frac{1}{2}\pi < \theta < \frac{3}{2}\pi) = 2P(\frac{1}{2}\pi < \theta < \pi)\\ &= 2\sum_{\tilde{x}_2=\pm\sqrt{E_s}} P(\tilde{x}_2)\int_{\frac{\pi}{2}}^{\pi}\int_{-\infty}^{\infty} p(\theta|\tilde{x}_2, h_2)p(h_2)\,dh_2\,d\theta\\ &= \frac{1}{\pi}\left[\sqrt{\frac{\sin^2\theta}{\gamma_s^{-1}+\sin^2\theta}}\cot^{-1}\left(\frac{-\cos\theta}{\sqrt{\gamma_s^{-1}+\sin^2\theta}}\right)+\theta-\pi\right]\Bigg|_{\frac{\pi}{2}}^{\pi}\\ &= \frac{1}{2}\left(1-\sqrt{\frac{\gamma_s}{\gamma_s+1}}\right),\end{aligned} \tag{93}$$

and the SER at the relay is

$$\begin{aligned}P_e^{I,R} &= P_e^{I,R}\,(\,x_1|\,x_2)\,P\,(x_2) + P_e^{I,R}\,(\,x_2|\,x_1)\,P\,(x_1)\\ &= 2P_e^{I,R}\,(\,x_1|\,x_2)\,P\,(x_2)\\ &< 2P_e^{I,R}\,(\,x_1|\,x_2)\,\big(1-P_e^{BPSK}\,(x_2)\big)\\ &= \frac{1}{2\,(\gamma_s+1)}\,,\end{aligned} \tag{94}$$

where [98]

$$P_e^{BPSK} = \frac{1}{2}\left(1-\sqrt{\frac{\gamma_s}{\gamma_s+1}}\right)\,. \tag{95}$$

Furthermore, the end-to-end SER for the BPSK modulation can be upper bounded as

$$P_e^{I,T} < \frac{1}{2\,(\gamma_s+1)}\left[1-\frac{1}{2}\left(1-\sqrt{\frac{\gamma_s}{\gamma_s+1}}\right)\right] + \frac{1}{2}\left(1-\sqrt{\frac{\gamma_s}{\gamma_s+1}}\right)\left[1-\frac{1}{2\,(\gamma_s+1)}\right]$$

$$= \frac{1}{2}\left(1 - \frac{\gamma_s}{\gamma_s+1}\sqrt{\frac{\gamma_s}{\gamma_s+1}}\right) . \tag{96}$$

Similarly for QPSK, the conditional error probability of x_1 given x_2 is derived in Eq. (97).

$$\begin{aligned} P_e^{I,R}(x_1|x_2) &= P(\frac{1}{4}\pi < \theta < \frac{3}{4}\pi) = 2P(\frac{1}{4}\pi < \theta < \pi) \\ &= 2\sum_{\tilde{x}_2=\sqrt{\frac{E_s}{2}}(\pm 1 \pm i)} P(\tilde{x}_2) \int_{\frac{\pi}{4}}^{\pi}\int_{-\infty}^{\infty} p(\theta|\tilde{x}_2, h_2)p(h_2)\, dh_2\, d\theta \\ &= \frac{1}{\pi}\left[\sqrt{\frac{\sin^2\theta}{\gamma_s^{-1}+\sin^2\theta}}\cot^{-1}\left(\frac{-\cos\theta}{\sqrt{\gamma_s^{-1}+\sin^2\theta}}\right) + \theta - \pi\right]\Bigg|_{\frac{\pi}{4}}^{\pi} \\ &= \frac{3}{8} - \frac{1}{2}\sqrt{\frac{\gamma_s}{\gamma_s+2}} + \frac{1}{2\pi}\sqrt{\frac{\gamma_s+1}{\gamma_s+2}}\cot^{-1}\left(\sqrt{\frac{\gamma_s}{\gamma_s+2}}\right). \end{aligned} \tag{97}$$

Thus, the SER at the relay is upper bounded as,

$$\begin{aligned} P_e^{I,R} &= 2P_e^{I,R}(x_1|x_2)\,P(x_2) \\ &< 2P_e^{I,R}(x_1|x_2)\left(1 - P_e^{QPSK}(x_2)\right) \\ &= \frac{15}{32} - \frac{1}{4}\sqrt{\frac{\gamma_s}{\gamma_s+2}} - \frac{\gamma_s}{2(\gamma_s+2)} + \frac{1}{4\pi}\sqrt{\frac{\gamma_s+1}{\gamma_s+2}}\cot^{-1}\left(\sqrt{\frac{\gamma_s}{\gamma_s+2}}\right) \\ &\quad + \frac{\gamma_s(\gamma_s+1)}{\pi(\gamma_s+2)}\cot^{-1}\left(\sqrt{\frac{\gamma_s}{\gamma_s+2}}\right) - \frac{\gamma+1}{2\pi^2(\gamma+2)}\left(\cot^{-1}\left(\sqrt{\frac{\gamma_s}{\gamma_s+2}}\right)\right)^2. \end{aligned} \tag{98}$$

The upper bound of the end-to-end SER for QPSK can be obtained as

$$P_e^{I,T} < 2P_e^{I,R}(1 - P_e^{QPSK}). \tag{99}$$

The upper bound of the end-to-end SER for higher modulations can be obtained similarly.

6.3.2. SER of Space-Time Coded TWRCs

The SER at the relay and the terminals for TWRC with space-time codes can be derived in a similar way. Again bearing the system symmetry in mind, I only derive the SER of x_{11} detection. At the relay the detection variable of x_{11} is expressed as

$$\tilde{s}_{11} = \sum_{k=1}^{L} X_k Y_k^* , \tag{100}$$

where

$$
\begin{aligned}
Y_1 &= h_{11:1}, Y_2 = h^*_{12:1}, Y_3 = h_{11:2}, Y_4 = h^*_{12:2} \\
X_1 &= r_{1,1} - h_{21:1}\hat{x}_{21} - h_{22:1}\hat{x}_{22} \\
X_2 &= \left(r_{1,2} - h^*_{22:1}\hat{x}_{21} + h^*_{21:1}\hat{x}_{22}\right)^* \\
X_3 &= r_{2,1} - h_{21:2}\hat{x}_{21} - h_{22:2}\hat{x}_{22} \\
X_4 &= \left(r_{2,2} - h^*_{22:2}\hat{x}_{21} + h^*_{21:2}\hat{x}_{22}\right)^*
\end{aligned} \tag{101}
$$

and L is the number of channels, in this case, $L = 4$. The conditional probability density function of Θ [98] given $h_{21:1}$, $h_{21:2}$, $h_{22:1}$, $h_{22:2}$, and correctly decoded $\tilde{x}_{21}$ and $\tilde{x}_{22}$ is

$$
\begin{aligned}
\rho_{x_{11}} &= p(\theta \,|\tilde{x}_{21}, \tilde{x}_{22}, h_{21:1}, h_{21:2}, h_{22:1}, h_{22:2}\,) \\
&= \frac{(-1)^{L-1}\left(1-|\mu|^2\right)^L}{2\pi\,(L-1)!} \left\{ \frac{\partial^{L-1}}{\partial b^{L-1}} \left[\frac{1}{b - |\mu|^2 \cos^2\theta} + \frac{|\mu|\cos\theta \cos^{-1}\left(-\frac{|\mu|\cos\theta}{\sqrt{b}}\right)}{\left[b - |\mu|^2\cos^2\theta\right]^{3/2}} \right] \right\} \Bigg|_{b=1}
\end{aligned} \tag{102}
$$

where for all k,

$$
\mu = \sqrt{\frac{\gamma_s}{\gamma_s + 2}}\,. \tag{103}
$$

For BPSK, the conditional error probability of x_{11} at the relay $P_e^{II,R}(\,x_{11}|\,x_{21}, x_{22})$ given that x_{21} and x_{22} are correctly decoded, is derived in Eq. (104).

$$
\begin{aligned}
& P_e^{II,R}(\,x_{11}|\,x_{21}, x_{22}) \\
&= 2 \sum_{\tilde{x}_{21}, \tilde{x}_{22} = \pm\sqrt{\frac{E_s}{2}}} P(\tilde{x}_{21}) P(\tilde{x}_{22}) \\
&\qquad \int_{\frac{\pi}{2}}^{\pi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho_{x_{11}} p(h_{21:1}) p(h_{21:2}) p(h_{22:1}) p(h_{22:2})\, dh_{21:1} dh_{21:2}\, dh_{22:1}\, dh_{22:2}\, d\theta
\end{aligned}
$$

$$= \frac{-8/3\pi}{(\gamma_s+2)^4}\frac{\partial^3}{\partial b^3}\left\{\frac{2\gamma_s^{-1}+1}{b(2\gamma_s^{-1}+1)-1}\left[\frac{\sin\theta\cot^{-1}\left(\frac{-\cos\theta}{\sqrt{b\left(2\gamma_s^{-1}+1\right)-\cos^2\theta}}\right)}{\sqrt{b\left(2\gamma_s^{-1}+1\right)-\cos^2\theta}}-\theta+\pi\right]\right\}\Bigg|_{\frac{\pi}{4}}^{\pi}\Bigg|_{b=1}$$

$$= \frac{1}{2}-\frac{1}{2}\sqrt{\frac{\gamma_s}{\gamma_s+2}}\left[1+\frac{1}{\gamma_s+2}+\frac{3}{2\left(\gamma_s+2\right)^2}+\frac{5}{2\left(\gamma_s+2\right)^3}\right], \tag{104}$$

and therefore I have

$$\begin{aligned} & P_e^{II,R} \\ &= 2P_e^{II,R}\left(x_{11}|\, x_{21},x_{22}\right)P\left(x_{21},x_{22}\right) \\ &< 2P_e^{II,R}\left(x_{11}|\, x_{21},x_{22}\right)\left(1-P_e^{AL}(x_{21},x_{22})\right) \\ &= \frac{1}{2}-\frac{\gamma_s}{2\left(\gamma_s+2\right)}\left[1+\frac{2}{\gamma_s+2}+\frac{4}{\left(\gamma_s+2\right)^2}+\frac{8}{\left(\gamma_s+2\right)^3}+\frac{29}{4\left(\gamma_s+2\right)^4}\right. \\ &+ \left.\frac{15}{2\left(\gamma_s+2\right)^5}+\frac{25}{4\left(\gamma_s+2\right)^6}\right] \end{aligned} \tag{105}$$

where $P_e^{AL}(x_{21},x_{22})$ represents the SER of the two-receiver Alamouti scheme using BPSK formed by either terminal as given by [98, 99]. I can also derive the upper bound for the end-to-end SER as follows:

$$\begin{aligned} & P_e^{II,T} \\ &< 2P_e^{AL}(x_{R1})\left(1-P_e^{II,R}\right) \\ &= \frac{1}{2}-\frac{\gamma_s}{2\left(\gamma_s+2\right)}\sqrt{\frac{\gamma_s}{\gamma_s+2}}\left[1+\frac{3}{\gamma_s+2}+\frac{15}{2\left(\gamma_s+2\right)^2}+\frac{35}{2\left(\gamma_s+2\right)^3}+\frac{105}{4\left(\gamma_s+2\right)^4}\right. \\ &+ \left.\frac{147}{4\left(\gamma_s+2\right)^5}+\frac{357}{8\left(\gamma_s+2\right)^6}+\frac{285}{8\left(\gamma_s+2\right)^7}+\frac{225}{8\left(\gamma_s+2\right)^8}+\frac{125}{8\left(\gamma_s+2\right)^9}\right], \end{aligned} \tag{106}$$

where $P_e^{AL}(x_{R1})$ represents the SER of symbol x_{R1} transmitted by the relay using the single-receiver Alamouti scheme and the BPSK modulation.

Again, the conditional error probability of x_{11} for QPSK is given by Eq. (107).

$$
\begin{aligned}
& P_e^{II,R}(x_{11}|\, x_{21}, x_{22}) \\
&= 2 \sum_{\tilde{x}_{21},\tilde{x}_{22}=\frac{\sqrt{E_s}}{2}(\pm 1 \pm i)} P(\tilde{x}_{21}) P(\tilde{x}_{22}) \\
&\qquad \int_{\frac{\pi}{2}}^{\pi} \int_{-\infty}^{\infty} \rho_{x_{11}} p(h_{21:1}) p(h_{21:2}) p(h_{22:1}) p(h_{22:2})\, dh_{21:1}\, dh_{21:2}\, dh_{22:1}\, dh_{22:2}\, d\theta \\
&= \frac{-8/3\pi}{(\gamma_s+2)^4} \frac{\partial^3}{\partial b^3} \left\{ \frac{(2\gamma_s^{-1}+1)/b}{2\gamma_s^{-1}+1-\frac{1}{b}} \left[\frac{\sin\theta \cot^{-1}\left(\frac{-\cos\theta}{\sqrt{b\left(2\gamma_s^{-1}+1\right)-\cos^2\theta}} \right)}{\sqrt{b\left(2\gamma_s^{-1}+1\right)-\cos^2\theta}} - \theta + \pi \right] \right\}_{\frac{\pi}{4}}^{\pi} \Bigg|_{b=1} .
\end{aligned}
$$

(107)

Then at the relay I have the relation derived as

$$
P_e^{II,R} = 2P_e^{II,R}\left(x_{11}|\, x_{21}, x_{22}\right) P\left(x_{21}, x_{22}\right) < 2P_e^{II,R}\left(x_{11}|\, x_{21}, x_{22}\right) \left(1 - P_e^{AL}(x_{21}, x_{22})\right).
$$

(108)

The upper bound for the end-to-end SER of QPSK can be obtained in the same way as BPSK. The explicit closed form is rather long. Instead the numerical results computed directly from the theoretical bound are given in Figs. 6.6 and 6.7. Similarly, the bounds for higher modulations can also be derived.

6.4. Numerical Results

In this section, I present some numerical results obtained from the simulations. These results show that the proposed DF protocols outperform the AF and PDF when used in a dual-antenna relay network. Certain assumptions are made in regard to the system model. It is assumed that ideally perfect channel information is known to the receivers, and the transmission powers for terminals T_1, T_2 and the relay R are the same. Essentially, this is a symmetric network. The end-to-end SER is chosen as the primary performance metric, and its values are plotted against the SNRs.

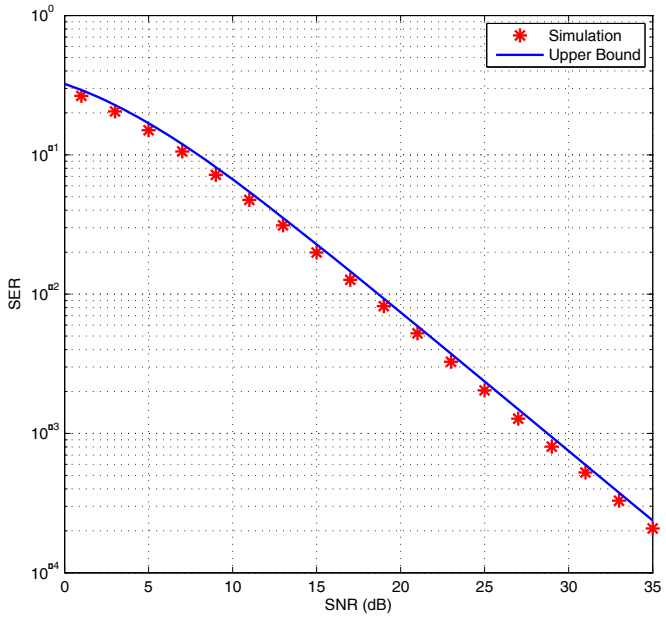


FIGURE 6.4. The end-to-end error performance of DF for TWRC with single antenna using BPSK: simulation vs. theoretical upper bound.

Fig. 6.4 compares the simulated end-to-end error probability performance of DF with its theoretical upper bound, in the single-antenna TWRC using BPSK modulations. The figure validates the mathematical derivations in Section 6.3.1, and shows the derived upper bound serves as a tight approximation for the values obtained from simulations.

Fig. 6.5 performs the same comparison as Fig. 6.4 on the DF protocols using the QPSK modulation in the same system scenario. Again, the theoretical upper bound is shown to be a good approximation for the simulated values and suffices to provide insight for the system design considerations.

Figs. 6.6, 6.7 compare the simulated error rate with its upper bounds for DF using BPSK and QPSK respectively in the space-time coded two-way relay network. The validity of the theoretically obtained bounds is confirmed by the simulation results as the curves pair very closely together in the figures.

The last two figures Figs. 6.8 and 6.9 perform comparison of DF protocols with AF and PDF. As discussed in Section 6.1, the noise is amplified and propagated together with the original signals or the combinations of the original signals. The efficiency of the energy use

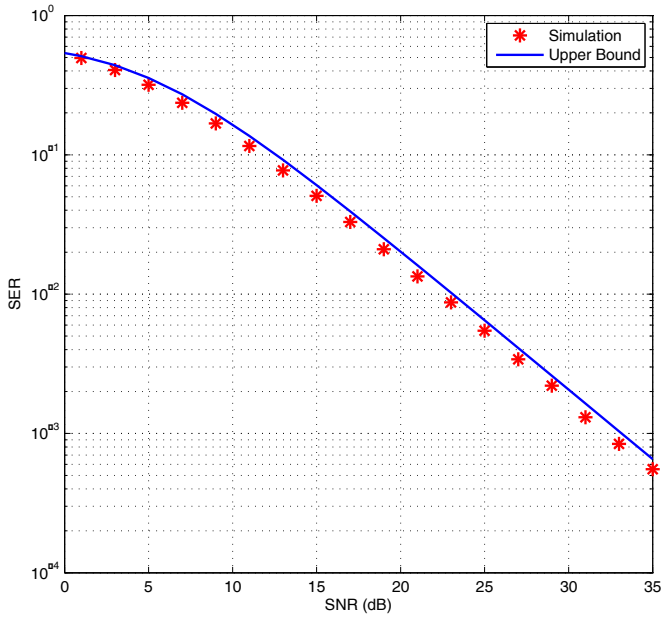


FIGURE 6.5. The end-to-end error performance of DF for TWRC with single antenna using QPSK: simulation vs. theoretical upper bound.

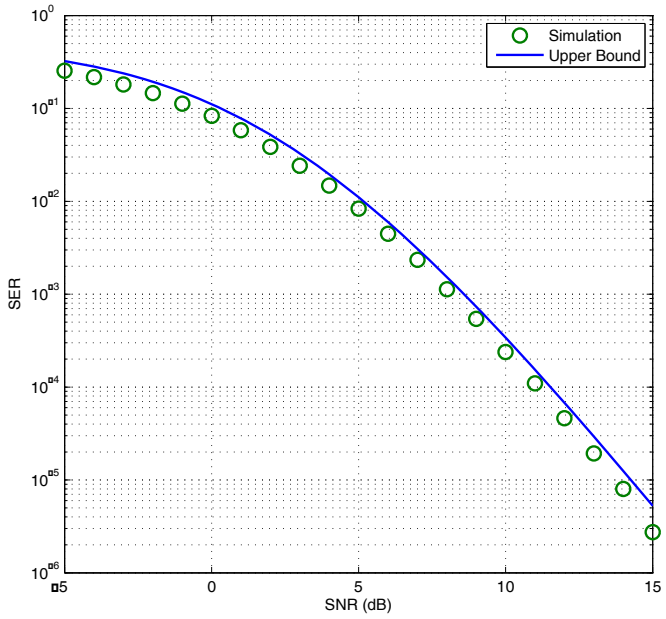


FIGURE 6.6. The end-to-end error performance of DF for TWRC with space-time codes using BPSK: simulation vs. theoretical upper bound.

is less compared to PDF and DF, and AF yields the worst error rate among three schemes in either single-antenna TWRC or the space-time coded counterpart. As for the PDF protocol, it gains an advantage over AF due to its noise processing at the relay, but does not benefit

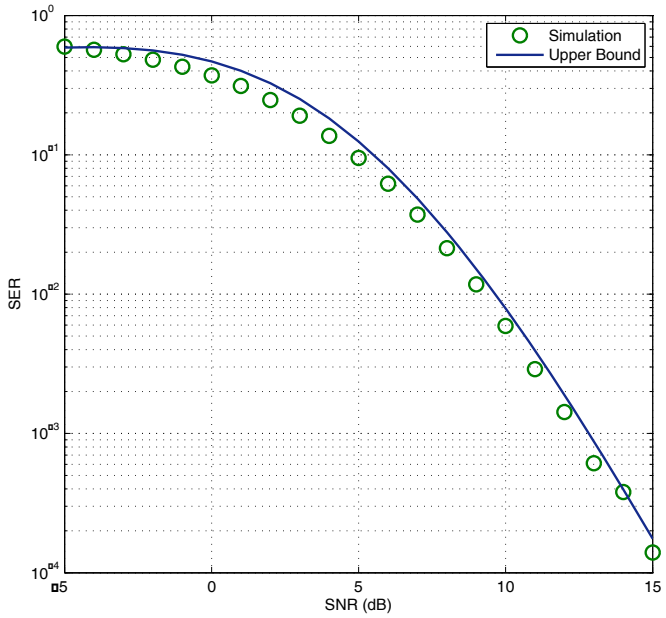


FIGURE 6.7. The end-to-end error performance of DF for TWRC with space-time codes using QPSK: simulation vs. theoretical upper bound.

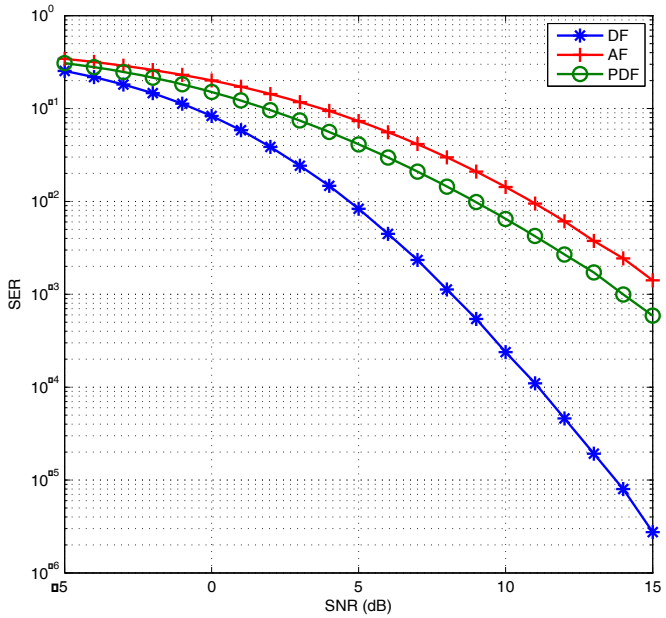


FIGURE 6.8. The SER comparison of AF, PDF and DF using BPSK by simulations.

from the coding gain of DF. Thus, DF proves to be the most beneficial as shown in the simulation results.

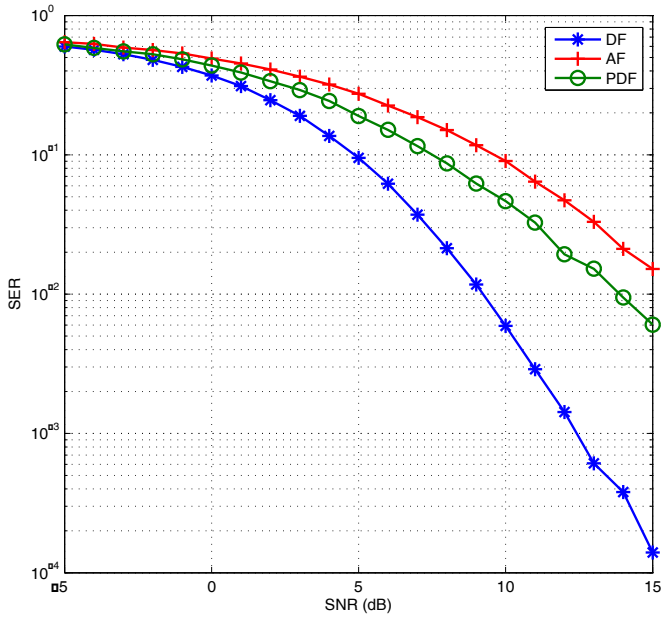


FIGURE 6.9. The SER comparison of AF, PDF and DF using QPSK by simulations.

6.5. Summary

In this chapter, I introduced a new class of DF protocols as the space-time coded physical-layer network coding. It was shown by numerical simulations that the proposed schemes outperform the existing AF and PDF protocols when extended to the dual-antenna relay network. The theoretical upper bounds were also derived in particular for the BPSK and QPSK modulations.

CHAPTER 7

CONCLUSIONS

In this dissertation, I studied the physical-layer network coding in the two-way relay channels. The physical-layer network coding takes advantage of the additive nature of the electromagnetic waves, and embraces the interferences by performing coding operations to combine the otherwise conflicted messages. In particular, I proposed a new transmission scheme namely block relaying and a novel application of permutation modulations in conjunction with the physical-layer network coding technique. Also, I performed extensive theoretical analysis on the performances of the two-way relay channel and the parallel relay channel, and obtained the power allocation strategies in order to minimize the end-to-end error probability or equivalently to maximize the end-to-end throughput. Eventually, I extended the Alamouti scheme of the space-time coding to the two-way relay networks, and discussed the implementations of the amplify-and-forward, partial-decode-and-forward, and decode-and-forward protocols and their performances.

The content and contributions of this dissertation are summarized as follows.

- Chapter 1: I introduced the motivation behind this work. As fading and noise are present in the channels, errors are inevitably introduced to the received symbols. The dissertation aims at reducing the error rate and improving the system performance.
- Chapter 2: I provided a detailed overview of the two-way relay channels. I described various aspects of the system design and related assumptions including the link availabilities between the terminals and the relay, the number of sources / destination / relays, transmission approach, relaying strategy, power allocation policy. The latest important research works and results in the area of the physical-layer network coding were surveyed and discussed.

- Chapter 3: In this chapter, I proposed a new class of $\log_{M_1} M_2$-MA block relay for TWRC systems. Instead of forwarding the received symbols for each MA phase, the relay buffers the symbols from multiple MA phases and forwards the combination through a higher modulation. In particular, I focused on the case of $M_1 = 2$ using BPSK in the MA phase. I obtained the performance bounds for $\log_2 M$-MA block relaying in the fading channels and the exact closed forms for the AWGN channels. The beneficial SNR range for the proposed scheme, and the optimal power allocation strategy between the terminals and the relay were also discussed. Numerical results confirmed the advantages of the proposed scheme in terms of higher throughput in low SNRs.
- Chapter 4: I investigated the optimized transmission power allocation strategies aiming at minimizing the end-to-end SER performance of the ACDF protocol in the energy constrained fading parallel relay network. I derived the exact closed-form SER of the system, and further obtained an upper bound to represent the trend in a more mathematically simplistic fashion, and also an approximation asymptotically tight at high SNRs. These tools help reveal the relationship between the SER performance and the power allocation. Finally, I derived the goal function and the optimized solution. Numerical results confirm this work by showing performance improvement in terms of minimizing the end-to-end SER.
- Chapter 5: I introduced a novel method of the physical-layer network coding by utilizing the binary permutation modulations instead of the traditional modulation types. I compared the end-to-end performance of the proposed scheme with the traditional approach using BPSK, as BPSK yields the smallest error probability among the conventional modulation types. Both theoretical analysis in the explicit analytic forms and numerical results were given. In particular, I proved a series of lemmas and a theorem on the performance and optimality of the proposed scheme. It is shown that the proposed class of permutation codes outperforms the conventional approach using BPSK in the physical-layer network coding.

- Chapter 6: In this chapter, I introduced a new class of DF protocols as the space-time coded physical-layer network coding. It was shown by numerical simulations that the proposed schemes outperform the existing AF and PDF protocols when extended to the dual-antenna relay network. The theoretical upper bounds were also derived in particular for the BPSK and QPSK modulations.

7.1. Practical Issues

While there are many advantages and desired benefits from using the network coding or the physical-layer network coding techniques, some practical issues exist in the implementation of such protocols. For example, in the AWGN channel, the use of the physical-layer network coding requires strict synchronization between the terminals. The authors addressed some aspects of the synchronization issues in the application of the physical-layer network coding in [30]. Most of the existing research, including the chapters of this dissertation, still remains in the theoretical investigation level. Only a few works actually completed the implementation of their proposed schemes. For instance, as previously mentioned, the authors evaluated their coding protocol in the software defined radios in [28, 35, 91]. The technical issues need to be addressed before the application can be further promoted. This is the future research direction after this dissertation.

BIBLIOGRAPHY

[1] N. Xu, S. Fu, and Y. Huang, "On the two-way relay channels using binary permutation modulations," in *IEEE GLOBECOM '11*, under review.

[2] K. Subbu, N. Xu, and R. Dantu, "*i*know where you are," in *IEEE Int'l Conf. Comput. Sci. Eng. '09*, vol. 4, Aug. 2009, pp. 469–474.

[3] J. Kim, "Performance analysis of physical layer network coding," Ph.D. dissertation, University of Michigan, 2009. [Online]. Available: http://hdl.handle.net/2027.42/64791

[4] B. Rohani and K. Homayounfar, "Physical layer network coding for wireless applications: A survey," IEICE, Tech. Rep. 445, Mar. 2009.

[5] N. Xu and S. Fu, "Performance analysis of space-time codes over two-way relay channels," in *IEEE MILCOM'09*, Boston, MA, Oct. 2009.

[6] ——, "On the performance of two-way relay channels using space-time codes," *International Journal of Communication Systems*, 2011, accepted. [Online]. Available: http://dx.doi.org/10.1002/dac.1205

[7] ——, "Block relaying for physical-layer network coding," *IEEE Trans. Commun.*, under review.

[8] N. Xu, S. Fu, and Y. Huang, "Binary permutation modulations for physical-layer network coding," *IEEE Trans. Wireless Commun.*, under review.

[9] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Prentice Hall, 2001.

[10] N. Xu, K. Subbu, and S. Tang, "Confounded factor effects on battery life in wireless sensor networks," in *ICDIM '09*, Nov. 2009, pp. 1–6.

[11] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Stat. Probab.*, vol. I, 1961, pp. 611–644.

[12] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, no. 1, pp. 120–154, 1971. [Online]. Available: http://www.jstor.org/stable/1426331

[13] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.

[14] J. Laneman and G. Wornell, "Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks," in *IEEE GLOBECOM '02*, vol. 1, 2002, pp. 77–81.

[15] ——, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[16] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[17] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. Part I. System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[18] ——, "User cooperation diversity. Part II. Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[19] W. Su, A. K. Sadek, and K. J. R. Liu, "SER performance analysis and optimum power allocation for decode-and-forward cooperation protocol in wireless networks," in *IEEE Wireless Commun. Netw. Conf.*, vol. 2, Mar. 2005, pp. 984–989.

[20] ——, "Cooperative communication protocols in wireless networks: Performance analysis and optimum power allocation," *Wireless Pers. Commun.*, vol. 44, no. 2, pp. 181–217, 2008.

[21] S. Ahmadzadeh, S. Motahari, and A. Khandani, "Signal space cooperative communication," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1266–1271, Apr. 2010.

[22] J. Boutros and E. Viterbo, "Signal space diversity: a power- and bandwidth-efficient diversity technique for the rayleigh fading channel," *IEEE Trans. Inform. Theory*, vol. 44, no. 4, pp. 1453–1467, July 1998.

[23] T. Cui, F. Gao, T. Ho, and A. Nallanathan, "Distributed space-time coding for two-way wireless relay networks," in *IEEE Int. Conf. Commun.*, May 2008, pp. 3888–3892.

[24] T. Cui, T. Ho, and J. Kliewer, "Some results on relay strategies for memoryless two-way relay channels," in *Inform. Theory Appl. Workshop*, 2008, pp. 158–164.

[25] T. Cui and J. Kliewer, "Memoryless relay strategies for two-way relay channels: Performance analysis and optimization," in *IEEE Int. Conf Commun.*, May 2008, pp. 1139–1143.

[26] T. Cui, F. Gao, and C. Tellambura, "Differential modulation for two-way wireless communications: a perspective of differential network coding at the physical layer," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 2977–2987, Oct. 2009.

[27] M. A. Karim, T. Yang, J. Yuan, Z. Chen, and I. Land, "A novel soft forwarding technique for memoryless relay channels based on symbol-wise mutual information," *IEEE Commun. Lett.*, vol. 14, no. 10, pp. 927–929, Oct. 2010.

[28] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: practical wireless network coding," *SIGCOMM Comput. Commun. Rev.*, vol. 36, pp. 243–254, Aug. 2006. [Online]. Available: http://doi.acm.org/10.1145/1151659.1159942

[29] P. Hu and M. Ibnkahla, "A survey of physical-layer network coding in wireless networks," in *25th Biennial Symp. Commun.*, May 2010, pp. 311–314.

[30] S. Zhang, S.-C. Liew, and P. Lam, "On the synchronization of physical-layer network coding," in *IEEE Inform. Theory Workshop '06*, Oct. 2006, pp. 404–408.

[31] S. Zhang, S. C. Liew, and H. Wang, "Synchronization analysis in physical layer network coding," *CoRR*, vol. abs/1001.0069, 2010.

[32] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: physical-layer network coding," in *Proc. 12th Annu. Int. Conf. MobiCom.* New York, NY, USA: ACM, 2006, pp. 358–365.

[33] S. Zhang and S.-C. Liew, "Channel coding and decoding in a relay system operated with physical-layer network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 788–796, June 2009.

[34] S. Zhang, S. C. Liew, and L. Lu, "Physical layer network coding schemes over finite and infinite fields," in *IEEE GLOBECOM '08*, Dec. 2008.

[35] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: analog network coding," in *Proc. Conf. Appl. Technol. Arch. Protocols Comput. Commun.* New York, NY, USA: ACM, 2007, pp. 397–408.

[36] J. Liu, M. Tao, Y. Xu, and X. Wang, "Superimposed XOR: A new physical layer network coding scheme for two-way relay channels," in *IEEE GLOBECOM '09*, Nov. 2009, pp. 1 –6.

[37] D. To and J. Choi, "Convolutional codes in two-way relay networks with physical-layer network coding," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2724–2729, Sept. 2010.

[38] T. Cui, T. Ho, and J. Kliewer, "Memoryless relay strategies for two-way relay channels," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3132–3143, Oct. 2009.

[39] B. Vucetic and J. Yuan, *Space-Time Coding*. John Wiley & Sons, Inc., 2003.

[40] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[41] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," in *IEEE WCNC. '99*, vol. 3, Sept. 1999, pp. 1043–1047.

[42] ——, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1169–1174, July 2000.

[43] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[44] R. Louie, Y. Li, and B. Vucetic, "Practical physical layer network coding for two-way relay channels: performance analysis and comparison," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 764–777, Feb. 2010.

[45] K. Lu, S. Fu, and Y. Qian, "Capacity of random wireless networks: Impact of physical-layer network coding," in *IEEE Int. Conf. Commun. '08*, May 2008, pp. 3903–3907.

[46] Y. Fan and J. Thompson, "MIMO configurations for relay channels: Theory and practice," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1774–1786, May 2007.

[47] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[48] W. Nam, S.-Y. Chung, and Y. Lee, "Capacity bounds for two-way relay channels," in *IEEE Int. Zurich Seminar Commun. '08*, Mar. 2008, pp. 144–147.

[49] ——, "Capacity of the gaussian two-way relay channel to within bit," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.

[50] M. Wilson, K. Narayanan, H. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5641–5654, Nov. 2010.

[51] K. Lu, S. Fu, Y. Qian, and T. Zhang, "On the security performance of physical-layer network coding," in *ICC Int. Conf. Commun. '09*, June 2009, pp. 1–5.

[52] Y. Ji, J. Ge, J. Li, and J. Fang, "Design and performance analysis of a space-time cooperative physical-layer network coding scheme," in *Int. Conf. Comput. Appl. Sys. Model '10*, vol. 9, Oct. 2010, pp. 300–304.

[53] J. Wu, Y. R. Zheng, A. Gumaste, and C. Xiao, "Error performance of double space time transmit diversity system," in *IEEE Int. Conf. Commun. '06*, vol. 11, June 2006, pp. 4859–4864.

[54] ——, "Error performance of double space time transmit diversity system," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3191–3196, Sept. 2007.

[55] T. Zhang, K. Lu, A. Jafari, S. Fu, and Y. Qian, "On the capacity bounds of large-scale wireless network with physical-layer network coding under the generalized physical model," in *IEEE Int. Conf. Commun. '10*, May 2010, pp. 1–5.

[56] Z. Chen, J. Yuan, and B. Vucetic, "Analysis of transmit antenna selection/maximal-ratio combining in rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 54, no. 4, pp. 1312–1321, July 2005.

[57] H. A. David, *Order Statistics*. John Wiley & Sons, Inc, 1970.

[58] Z. Chen, J. Yuan, B. Vucetic, and Z. Zhou, "Performance of alamouti scheme with transmit antenna selection," in *IEEE PIMRC '04*, vol. 2, Sept. 2004, pp. 1135–1141.

[59] ——, "Performance of alamouti scheme with transmit antenna selection," *Electron. Lett.*, vol. 39, no. 23, pp. 1666–1668, Nov. 2003.

[60] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[61] T. Koike-Akino, P. Popovski, and V. Tarokh, "Optimized constellations for two-way wireless relaying with physical network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 773–787, June 2009.

[62] ——, "Denoising maps and constellations for wireless network coding in two-way relaying systems," in *IEEE GLOBECOM '08*, Nov. 2008, pp. 1–5.

[63] H. J. Yang, Y. Choi, and J. Chun, "Modified high-order pams for binary coded physical-layer network coding," *IEEE Commun. Lett.*, vol. 14, no. 8, pp. 689–691, Aug. 2010.

[64] S. Yun, H. Kim, and K. Tan, "Towards zero retransmission overhead: A symbol level network coding approach to retransmission," *IEEE Trans. Mobile Comput.*, no. 99, pp. 1–14, Dec. 2010.

[65] C. Pan and J. Zheng, "Mapping codebook-based physical network coding for asymmetric two-way relay channels," in *IEEE ICC '10*, May 2010, pp. 1–5.

[66] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074–1080, July 2002.

[67] J. Lassing, E. Strom, E. Agrell, and T. Ottosson, "Bit error probability of coherent M-ary PSK over flat rayleigh fading channels," *Electron. Lett.*, vol. 41, no. 21, pp. 1186–1187, Oct. 2005.

[68] P. Popovski and H. Yomo, "Wireless network coding by amplify-and-forward for bi-directional traffic flows," *IEEE Commun. Lett.*, vol. 11, no. 1, pp. 16–18, Jan. 2007.

[69] K. Lu, S. Fu, Y. Qian, and H.-H. Chen, "SER performance analysis for physical layer network coding over AWGN channels," in *IEEE GLOBECOM '09*, Nov. 2009, pp. 1–6.

[70] J. Lassing, E. Strom, T. Ottosson, and E. Agrell, "The exact symbol and bit error probabilities of coherent M-ary PSK," in *Proc. IEEE Int. Symp. Inform. Theory '03*,

29 June-4 July 2003, p. 11.

[71] I. S. Gradsteyn and I. M. Ryzhik, *Table of Integrals, Series and Products.*, 7th ed. Academic Press, 2007.

[72] M.-S. A. M. K. Simon, *Digital Communication over Fading Channels.*, 2nd ed. Wiley-IEEE Press, 2004.

[73] M. Dai and C. Sung, "A distributed on-off amplify-and-forward protocol for the fading parallel relay channel," *IEEE Commun. Lett.*, vol. 13, no. 9, pp. 643 –645, Sept. 2009.

[74] M. Hasna and M.-S. Alouini, "Optimal power allocation for relayed transmissions over rayleigh fading channels," in *IEEE VTC '03*, vol. 4, Apr. 2003, pp. 2461–2465.

[75] ——, "Optimal power allocation for relayed transmissions over rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1999–2004, Nov. 2004.

[76] W. Shin, N. Lee, J. B. Lim, and C. Shin, "An optimal transmit power allocation for the two-way relay channel using physical-layer network coding," in *IEEE ICC '09*, June 2009, pp. 1–6.

[77] Y. Zhang, Y. Ma, and R. Tafazolli, "Power allocation for bidirectional AF relaying over Rayleigh fading channels," *IEEE Commun. Lett.*, vol. 14, no. 2, pp. 145–147, Feb. 2010.

[78] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.

[79] M. Ju, H.-K. Song, and I.-M. Kim, "Exact ber analysis of distributed alamouti's code for cooperative diversity networks," *IEEE Trans. Commun.*, vol. 57, no. 8, pp. 2380–2390, Aug. 2009.

[80] T. Duong, D.-B. Ha, H.-A. Tran, and N.-S. Vo, "Symbol error probability of distributed-alamouti scheme in wireless relay networks," in *Proc. IEEE VTC*, Nov. 2008, pp. 648–652.

[81] F. Gao, R. Zhang, and Y.-C. Liang, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024–3033, Oct. 2009.

[82] D. G. Brennan, "Linear diversity combining techniques," *Proc. IEEE*, vol. 91, no. 2, pp. 331 – 356, Feb. 2003.

[83] M. K. Simon and M. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. IEEE*, vol. 86, no. 9, pp. 1860 –1877, Sept. 1998.

[84] D. Slepian, "Permutation modulation," *Proc. IEEE*, vol. 53, no. 3, pp. 228–236, Mar. 1965.

[85] T. Ericson, "Permutation codes," Rapport de Recherche INRIA 2109, Nov. 1993.

[86] J. Karlof, "Permutation codes for the Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 35, no. 4, pp. 726–732, July 1989.

[87] I. Ingemarsson, "Optimized permutation modulation," *IEEE Trans. Inform. Theory*, vol. 36, no. 5, pp. 1098–1100, Sept. 1990.

[88] N. Gaarder, "Probability of error for binary permutation modulation on a fading gaussian channel," *IEEE Trans. Inform. Theory*, vol. 17, no. 4, pp. 412– 418, July 1971.

[89] A. Nordio and E. Viterbo, "Permutation modulation for fading channels," in *ICT '03*, vol. 2, Feb.–Mar. 2003, pp. 1177–1183.

[90] D. Silva and W. Finamore, "Vector permutation modulation," *IEEE Commun. Lett.*, vol. 9, no. 8, pp. 673–675, Aug. 2005.

[91] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: practical wireless network coding," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 36, no. 4, pp. 243–254, June 2006.

[92] Y. Wu, P. Chou, and S. Kung, "Information exchange in wireless networks with network coding and physical-layer broadcast," Microsoft Research, Redmond WA, Tech. Rep. MSRTR-2004-78, Aug. 2004.

[93] M. Chen and A. Yener, "Multiuser two-way relaying: detection and interference management strategies," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4296–4305, Aug. 2009.

[94] K. Gomadam and S. Jafar, "The effect of noise correlation in amplify-and-forward relay networks," *IEEE Trans. Inform. Theory*, vol. 55, no. 2, pp. 731–745, Feb. 2009.

[95] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Comparison of bi-directional relaying protocols," in *IEEE Sarnoff Symp.*, Apr. 2008, pp. 1–5.

[96] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[97] H. Min, S. Lee, K. Kwak, and D. Hong, "Effect of multiple antennas at the source on outage probability for amplify-and-forward relaying systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 633–637, Feb. 2009.

[98] J. G. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2000.

[99] C. Gao, A. Haimovich, and D. Lao, "Bit error probability for space-time block code with coherent and differential detection," in *Proc. 56th IEEE Veh. Technol. Conf.*, vol. 1, 2002, pp. 410–414.