UNIVERSITY OF CALIFORNIA, SAN DIEGO

Joint Optimization of Physical and Application Layers for Wireless Multimedia Communications

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Electrical Engineering
(Communication Theory and Systems)

by

Seok-Ho Chang

Committee in charge:

Professor Pamela C. Cosman, Co-Chair
Professor Laurence B. Milstein, Co-Chair
Professor Patrick J. Fitzsimmons
Professor William S. Hodgkiss
Professor Bhaskar Rao

2010

The dissertation of Seok-Ho Chang is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
Co-Chair

_____
Co-Chair

University of California, San Diego

2010

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisors Prof. Laurence Milstein and Prof. Pamela Cosman for their valuable advice, encouragement and constant support during the course of my doctoral studies. Prof. Milstein, for his dedication to the profession, expertise in the filed of telecommunications and passionate life, has always been an ideal model of scholar to me. Prof. Cosman's remarkable insight at each important step of the research enabled me to conduct the research with broad perspectives in both fields of wireless communications and image/video applications.

I am sincerely grateful to my dissertation committee members, Prof. Bhaskar Rao, Prof. William Hodgkiss, and Prof. Patrick Fitzsimmons, for their valuable comments and suggestions on this work. I should thank Prof. Rao for his presentation of the course Digital Signal Processing, Prof. Hodgkiss for his teaching of the course Detection Theory, and Prof. Fitzsimmons for his treatment of the course Probability Theory. I also sincerely thank Prof. Ruth Williams, one of my committee members in the qualifying examination.

There are a number of people who helped me during my stay at the University of California, San Diego. First, I would like to thank our lab colleagues and visiting scholars, Prof. Minjoong Rim, Prof. Sunyong Kim, Jiwoong, Hobin, Sheu-Sheu, Laura, Ramesh, Mayank, Ting-Lan, Suk Ryool, Jihyun, Yuan, Qihang, and Zhuwei. I convey my special thanks to Mr. Myoungbo Kwak, Mr. Seong-Ho Hur, Prof. Wonha Kim, Haryoon, Ms. Maureen Nichol, Intae, Yen-Lin, Dr. Jinseong Jeong, Young-Pyo, Mr. Hwan Joon Kwon, Dr. Jaewook Shim, and Mr. Kwangok Jeong for their much help. Outside the campus, I also express thanks to Mr. Hee Jun Lee, Oh-Soon, Dr. Sung-Moon Yang, Dr. Bongyong Song, Heechoon, Mr. Hee Gul Park, Taesang, Jun Won, Hyoung-Il, and Jinho for their kind advice.

Finally, I would like to present all my love and thanks to my parents for supporting me with great love. Without them, this work would never have come into existence.

The material in Chapter 2 of this dissertation appeared in "Optimized Unequal Error Protection using Multiplexed Hierarchical Modulation," which has been submitted for publication in *IEEE Transactions on Information Theory*. The material in Chapter 3 of this dissertation appeared in "Superposition MIMO Coding for the Broadcast of Layered Sources," which has been submitted for publication in *IEEE Transactions on*

Communications. The material in Chapter 4 of this dissertation appeared in "Performance Analysis of n-Channel Symmetric FEC-Based Multiple Description Coding for OFDM Networks," which has been accepted for publication in *IEEE Transactions on Image Processing*.

VITA

1997	Bachelor of Science, Electrical Engineering, Seoul National University, Korea
1997–1999	Graduate Research Assistant, Electrical Engineering, Seoul National University, Korea
1999	Master of Science, Electrical Engineering, Seoul National University, Korea
1999–2005	Senior Engineer, Mobile Communication Research Lab. LG Electronics, Korea
2006	Senior Engineer, Portable Internet (mobile WiMAX) Lab. POSCO ICT, Korea
2007–2010	Graduate Research Assistant, Electrical Engineering, University of California, San Diego
2010	Doctor of Philosophy, Electrical Engineering, University of California, San Diego

PUBLICATIONS

S.-H. Chang , P. C. Cosman and L. B. Milstein, "Chernoff-Type Bounds for the Gaussian Error Function," to appear in *IEEE Trans. Commun.*

S.-H. Chang , M. Rim, P. C. Cosman and L. B. Milstein, "Superposition MIMO Coding for the Broadcast of Layered Sources," in revision for *IEEE Trans. Commun.*

S.-H. Chang , P. C. Cosman and L. B. Milstein, "Performance Analysis of n-Channel Symmetric FEC-Based Multiple Description Coding for OFDM Networks," to appear in *IEEE Trans. Image Process.*

S.-H. Chang , M. Rim, P. C. Cosman and L. B. Milstein, "Optimized Unequal Error Protection Using Multiplexed Hierarchical Modulation," submitted to *IEEE Trans. Inform. Theory.*

S.-H Chang and Y.-H. Lee, "LMS Adaptive Phase Locked Loops for Carrier Recovery," in *Proc. IEEE ISPACS'1999*, Phuket, Thailand, Dec. 1999.

S.-H Chang and H. J. Lee, "Low-Biased Doppler Frequency Estimation Scheme Employing Variable Prefilter and Sampling Rate," in *Proc. IEEE VTC'2003 (Spring)*, Jeju, Korea, Apr. 2003.

S.-H. Chang , M. Rim, P. C. Cosman and L. B. Milstein, "Optimal Multiplexed Hierarchical Modulation for Unequal Error Protection of Progressive Bit Streams," in *Proc. IEEE GLOBECOM'2009*, Honolulu, Hawaii, USA, Nov.–Dec. 2009.

ABSTRACT OF THE DISSERTATION

Joint Optimization of Physical and Application Layers for Wireless Multimedia Communications

by

Seok-Ho Chang

Doctor of Philosophy in Electrical Engineering
(Communication Theory and Systems)

University of California, San Diego, 2010

Professor Pamela C. Cosman, Co-Chair
Professor Laurence B. Milstein, Co-Chair

One of the challenges in the next generation wireless communication systems is to provide high quality multimedia services. To address this issue, the approach we take in this dissertation is cross-layer design of communication systems.

In the first part of the dissertation, we address transmission of progressive images or scalable video using hierarchical modulation. Thanks to the progressive features, progressive bitstreams enable each user in a multi-user network to decode the source at the rate that their channel allows. These progressive sources have the feature that they have gradual differences of importance in the bitstreams. One would like to have gradual differences in unequal error protection (UEP) to correspond to the gradual differences

in importance. However, hierarchical modulation, which is often used for UEP and is currently employed in the Digital Video Broadcasting (DVB) standard, provides only a limited number of UEP levels. By multiplexing hierarchical modulation, we propose a high performance multilevel UEP system for the transmission of progressive sources.

In the second part, we consider the transmission of a layered source in a multiple-input multiple-output (MIMO) system for broadcast scenarios. We first analyze the tradeoff between two different MIMO approaches, Alamouti coding and spatial multiplexing, having the same transmission rate. For analytical tractability, we consider high SNR approximate (minimum distance) bit error rates (BER) for both MIMO approaches. Based on this, we propose superposition MIMO coding, where two different MIMO approaches are hierarchically combined such that low-rate high priority components of the source are Alamouti coded, high-rate low priority components are spatially multiplexed, and the two different components are superposed. It is demonstrated that in broadcast scenarios, the proposed MIMO coding maximizes the performance of a layered source which has different data rates for its components.

In the third part of the dissertation, we analyze the performance of n-channel symmetric FEC-based multiple description coding for a progressive mode of transmission. Multiple description source coding has recently emerged as an attractive framework for robust multimedia transmission over packet erasure channels. In the analysis, we consider transmission over orthogonal frequency division multiplexing (OFDM) networks in a frequency-selective, slowly-varying, Rayleigh faded environment.

Chapter 1

Introduction

Fourth generation (4G) wireless systems are expected to provide high quality multimedia services with improved reliability [1]. In achieving this goal, one of the challenging issues is to design efficient communication protocols. For this reason, the validity of the traditional layered protocol architecture is now being examined. Cross-layer design has emerged as an attractive system design approach, where a protocol stack can be designed by utilizing the dependencies among different layers in order to maximize the performance [2]– [4].

A typical communication protocol architecture is shown in Fig. 1.1. Instead of a traditional OSI (open systems interconnections) with seven layers [5], only five layers are shown for simplicity. Some exemplary functions for each layer are also shown. As an example of cross-layer design, we consider the so-called joint source and channel coding (JSCC), which actively exploits the dependencies between the link layer and the application layer. Assuming the system bandwidth is constrained, there could be an issue of what portions of the bandwidth should be allocated to the source coding of the application layer, and what proportion should be allocated to the channel coding of the link layer. If more bits are allocated to the source coding, the transmitted signals would be more vulnerable to channel errors. On the other hand, if more bits are assigned to the channel coding, the source would experience more distortion. Hence, the allocation of bit rates between different layers is an issue in designing reliable wireless systems [6]– [8]. In a similar manner, there exist cross-layer design issues between the physical layer and the application layer to maximize the system performance gains. This dissertation primarily focuses on the design of optimal modulation schemes (i.e., the physical layer)

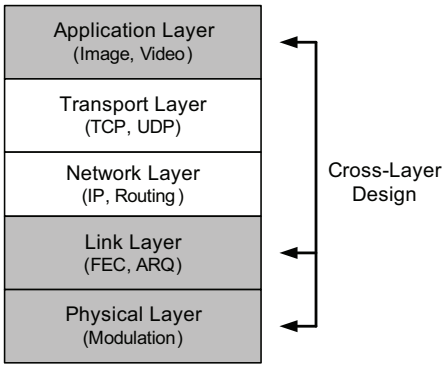


Figure 1.1: Cross-layer design of wireless multimedia communications.

to maximize the performance for the transmission of sources such as images or video.

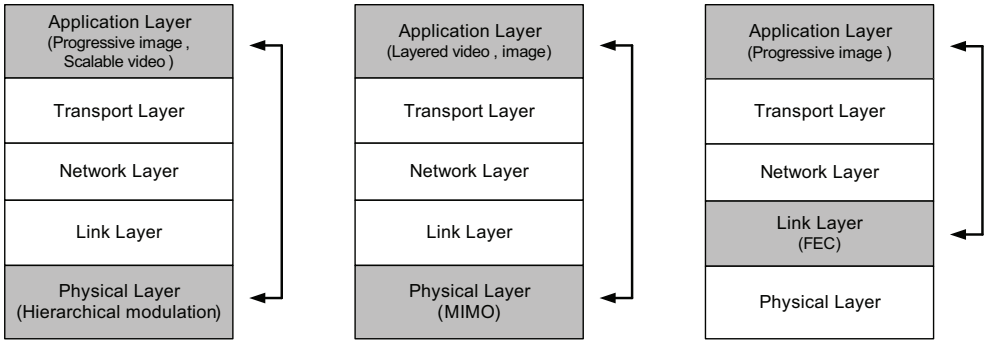
We first provide an overview of progressive source coding (e.g., progressive image and scalable video) which are the primary application sources in this dissertation. In a multi-service multi-user network, different users have different quality-of-service (QoS) requirements, due to various applications and scenarios coexisting within a network. As an example, the quality required for a real-time video played on a cellular phone is lower than that required on a laptop. Further, depending on each user's channel condition in terms of the received signal strength and the level of interference and noise power, the achievable quality for each user might be different. Progressive source coding [9]–[14], allowing partial decoding at various resolutions and quality levels from a single compressed bitstream, is a promising technology for multimedia communications in this environment. Thanks to the progressive features, in a multi-user network, each user decodes the same transmitted bitstream at the rate that their channel allows. Though a single progressive bitstream can be decoded at different data rates and provide a multitude of decoded quality levels, an error in the bitstream would make the subsequent bits useless. In other words, progressive source coders are usually extremely sensitive to channel impairments, and hence it is required to design a reliable transmission scheme for these sources. Progressive sources have an important feature in that they have steadily decreasing importance for bits later in the stream. Hence, unequal error protection (UEP) is a natural way to ensure reliable transmission, given the limited system resources such as system bandwidth and transmit power.

We next briefly describe superposition (or hierarchical) transmission. Theoretical investigation of efficient communication from a single source to multiple receivers

established the fundamental idea that optimal broadcast transmission could be achieved by a superposition or hierarchical transmission scheme [15]– [17]. Since the theoretical and conceptual basis for UEP was initiated by Cover [15], much of the work has shown that one practical method of achieving UEP is based on a constellation of nonuniformly spaced signal points [18]– [21], which is called a hierarchical, embedded, or multi-resolution constellation. In this constellation, more important bits in a symbol have a larger minimum Euclidian distance than less important bits. Hierarchical modulation was intensively studied for digital broadcasting systems [18] [20] [21] and multimedia transmission [22] [23]. Moreover, the Digital Video Broadcasting (DVB-T) standard [24], which is now commercially available, incorporated hierarchical QAM for layered video data transmission, since it provides enhanced system-level capacity and coverage in a wireless environment [25] [26]. However, hierarchical modulation can achieve only a limited number of UEP levels for a given constellation size, whereas progressive sources such as scalable video have gradual differences of importance in their bitstreams. Hence, if scalable video is to be incorporated in a digital video broadcasting system, the standard hierarchical modulation may not meet the system needs. In this dissertation, we address this problem and propose a high performance multilevel UEP system based on multiplexed hierarchical modulation.

Recently, multiple-input multiple-output (MIMO) systems have received a great deal of attention, since they can improve capacity and reliability relative to the single-input single-output (SISO) systems. Two popular techniques for MIMO systems are space-time coding [27]– [30] and spatial multiplexing [31]– [34]. Space-time coding is an approach where information is spread across multiple transmit antennas to maximize spatial diversity in fading channels. Spatial multiplexing is an approach whereby independent information is transmitted on each antenna, and thus the transmit data rate is increased without additional system bandwidth. In this dissertation, we study the broadcast of layered sources in MIMO systems. We first analyze the tradeoff between Alamouti coding and spatial multiplexing having the same maximum data rate. Based on the analysis, superposition MIMO coding is proposed for the broadcast of layered sources, where the low-rate important component is Alamouti coded, the high-rate, less important component is spatially multiplexed, and then the two unequally important components are superposed.

Another popular method to achieve UEP is based on channel coding. More



(a) Chapter 2 (b) Chapter 3 (c) Chapter 4

Figure 1.2: Cross-layer study on which this dissertation focuses.

powerful error-correction coding is applied to the more important component. The UEP methods based on channel coding have been widely used for layered video or image transmission [35]– [38]. However, channel coding becomes less effective in a slow fading channel, where prolonged deep fades often result in the erasure of the whole packet [36]. Multiple description source coding has emerged as an attractive framework for robust multimedia transmission over packet erasure channels [39]. Due to the individually decodable nature of the descriptions, the loss of some of them will not jeopardize the decoding of correctly received descriptions, and when more than one description is available at the decoder, they can be synergistically combined to enhance the quality [40]. In this dissertation, we mathematically analyze the performance of n-channel symmetric FEC-based multiple description coding for a progressive mode of transmission over orthogonal frequency division multiplexing (OFDM) networks in a frequency-selective, slowly-varying, Rayleigh faded environment.

The main contributions of this dissertation are contained in Chapters 2–4. In Chapters 2 and 3, we focus on the design of the optimal modulation schemes for progressive or layered sources, whereas in Chapter 4, our performance analysis is related to a cross-layer design between the link layer and the application layer. These are conceptually shown in Fig. 1.2. The rest of this dissertation is organized as follows.

In Chapter 2, a multilevel UEP scheme using multiplexed hierarchical modulation is proposed for the transmission of progressive sources. We propose a way of multiplexing hierarchical quadrature amplitude modulation (QAM) constellations, and

prove that arbitrarily large numbers of UEP levels are achieved by the proposed method. In addition, when the BER is dominated by the minimum Euclidian distance, an optimal multiplexing approach which minimizes both the average and peak powers is derived. While the suggested methods achieve multilevel UEP, the peak-to-average power ratio (PAPR) typically will be increased when constellations having distinct minimum distances are time-multiplexed. To mitigate this effect, an asymmetric hierarchical QAM constellation, which reduces the PAPR without performance loss, is designed. We also consider the case where multiplexed constellations need to have constant power. Numerical results show that the performance of progressive transmission over Rayleigh fading channels is significantly enhanced by the proposed multiplexing methods, without any additional system bandwidth or transmit power.

In Chapter 3, we propose superposition MIMO coding for the transmission of layered sources in a point-to-multipoint system. First, the tradeoff between Alamouti coding and spatial multiplexing is analyzed in terms of the average bit error rate (BER). As a way to compare both MIMO schemes fairly in broadcast systems, the maximum data rates of both are set to be equal. The results show that, for a given target bit error rate, Alamouti coding is preferable for a low data rate, and spatial multiplexing is preferable for a high data rate. In layered sources such as scalable video, the more important component typically has a low data rate, and the less important component typically has a high data rate. Based on these, we construct a superposition MIMO scheme where two different MIMO techniques are hierarchically combined. A successive decoding algorithm for the proposed scheme is also provided. Performance evaluation in a broadcasting scenario shows that the proposed superposition MIMO coding significantly outperforms the conventional MIMO coding schemes.

In Chapter 4, we analyze the performance of n-channel symmetric FEC-based multiple description coding for a progressive mode of transmission over OFDM networks in a frequency-selective slowly-varying Rayleigh faded environment. The expressions for the bounds of the throughput and distortion performance of the system are derived in an explicit closed form, while the exact performance is given by an expression in the form of a single integration. Based on this analysis, the performance of the system can be numerically evaluated. Our results show that at low SNR, the multiple description encoder should attempt to fine-tune the optimization parameters of the system. It is also shown that, despite the bursty nature of the errors in a slow fading environment, FEC-

based multiple description coding without temporal coding provides a greater advantage for smaller description sizes (i.e., shorter packets).

Finally, in Chapter 5, we summarize our contributions, and outline the possible extensions of our work.

Chapter 2

Optimized Unequal Error Protection Using Multiplexed Hierarchical Modulation

2.1 Introduction

When a communication system transmits messages over mobile radio channels, they are subject to errors, in part because mobile channels typically exhibit time-variant channel-quality fluctuations. For two-way communication links, these effects can be mitigated using adaptive methods [41]– [43]. However, the adaptive schemes require a reliable feedback link from the receiver to the transmitter. Moreover, for a one-way broadcast system, those schemes are not appropriate because of the nature of broadcasting. When adaptive schemes cannot be used, the way to ensure communications is to classify the data into multiple classes with unequal error protection (UEP). The most important class should be recovered by the receiver even under poor receiving conditions. Hence, strong error protection is used for the important data all of the time, even though sometimes there is no need for it. Less important data is always protected less even though sometimes it cannot be recovered successfully.

Since the theoretical and conceptual basis for UEP was initiated by Cover [15], much of the work has shown that one practical method of achieving UEP is based on a constellation of nonuniformly spaced signal points [18]– [21], which is called a hierarchical, embedded, or multi-resolution constellation. Hierarchical constellations

were previously considered in [44], and intensively studied for digital broadcasting systems [18] [20] [21]. Ramchandran *et al.* [18] designed an overall multiresolution digital HDTV broadcast system using hierarchical modulation under a joint source-channel coding (JSCC) framework. Calderbank and Seshadri [20] considered the use of hierarchical quadrature amplitude modulation (QAM) as the adaptive constellations for digital video broadcasting. Moreover, the Digital Video Broadcasting (DVB-T) standard [24] incorporated hierarchical QAM for layered video data transmission, since it provides enhanced system-level capacity and coverage in a wireless environment [25] [26]. Pursley and Shea [22] [23] also proposed communication systems based on hierarchical modulation which support multimedia transmission by simultaneously delivering different types of traffic, each with its own required quality of service.

Another well known and obvious method to achieve UEP is based on channel coding: more powerful error-correction coding is applied to a more important data class. Block codes for providing UEP were studied by Masnick and Wolf [45], and Suda and Miki [46]. The use of rate-compatible punctured convolutional (RCPC) codes to achieve UEP was suggested by Cox *et al.* [47]. These UEP methods based on error-correction coding have been widely used for layered video or image transmission [35]– [38]. Sometimes, UEP approaches based on hierarchical modulation and error-correction coding were jointly employed in a system [19] [20] [24] [22] [38]. For example, in the DVB-T standard [24], two different layers of video data are channel encoded with corresponding coding rates, and then they are mapped to hierarchical 16 or 64 QAM constellations. Pei and Modestino [38] showed that when an error-correction coding approach for UEP and hierarchical modulation are jointly used, more efficient and flexible UEP is achieved. Hierarchical modulation has other desirable properties in addition to performance considerations. The amount of UEP can be adjusted in a continuous manner by modifying the spacing between signal points of the constellation [19], and different levels of protection are achieved without an increase in bandwidth compared to channel coding [48].

Progressive image or scalable video encoders [9]– [14] employ a mode of transmission such that as more bits are received, the source can be reconstructed with better quality at the receiver. In other words, the decoder can use each additional received bit to improve the quality of the previously reconstructed images. Since these progressive transmissions have gradual differences of importance in their bitstreams, multiple levels

of error protection are required. However, unlike channel coding for UEP, hierarchical modulation can achieve only a limited number of UEP levels for a given constellation size. For example, hierarchical 16 QAM provides two levels of UEP, and hierarchical 64 QAM yields at most three levels [49]. In the DVB-T standard, video data encoded by MPEG-2 consists of two different layers, and thus the use of hierarchical 16 or 64 QAM meets the required number of UEP levels. However, if scalable video is to be incorporated in a digital video broadcasting system, hierarchical 16 or 64 QAM may not meet the system needs. Most of the work about hierarchical modulation up to now has been restricted to consideration of two layered source coding, and methods of achieving a large number of levels of UEP for progressive mode of transmission have rarely been studied.

In this chapter, we propose a multilevel UEP system using multiplexed hierarchical modulation for progressive transmission over mobile radio channels. We propose a way of multiplexing hierarchical QAM constellations, and show that arbitrarily large number of UEP levels are achieved by the proposed method. These results are presented in Section 2.2. When the BER is dominated by the minimum Euclidian distance, we derive an an optimal multiplexing approach which minimizes both the average and peak powers, which is presented in Section 2.3. While the suggested methods achieve multilevel UEP, the PAPR typically will be increased when constellations having distinct minimum distances are time-multiplexed. To mitigate this effect, an asymmetric hierarchical QAM constellation, which reduces the PAPR without performance loss, is designed in Section 2.4. In Section 2.5, we consider the case where multiplexed constellations need to have constant power, either due to the limited capability of a power amplifier, or for the ease of cochannel interference control. In Section 2.6, the performance of the suggested UEP system for the transmission of progressive images is analyzed in terms of the expected distortion, and Section 2.7 presents numerical results of performance analysis.

2.2 Multilevel UEP Based on Multiplexing Hierarchical QAM Constellations

2.2.1 Hierarchical 16 QAM Constellation

First, we analyze hierarchical 16 QAM as a special case. Fig. 2.1 shows a hierarchical 16 QAM constellation with Gray coded bit mapping [24]. The 16 signal

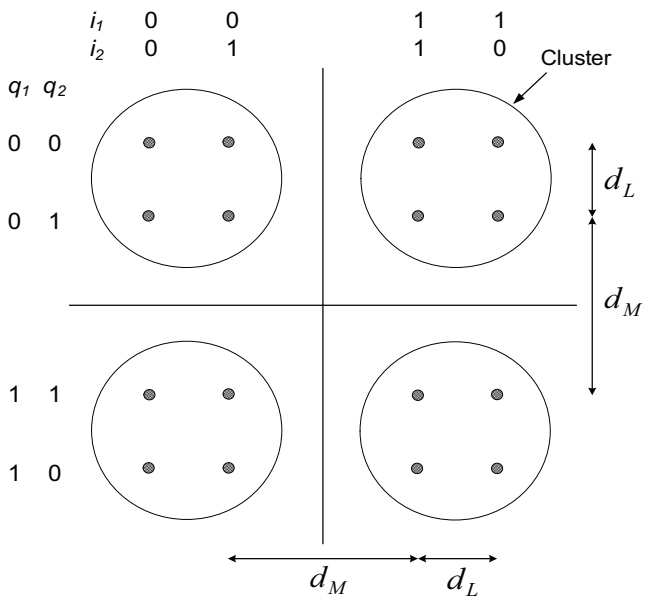


Figure 2.1: Hierarchical 16 QAM constellation.

points are divided into four clusters and each cluster consists of four signal points. The two most significant bits (MSBs), i_1 and q_1, determine one of the four clusters, and their minimum Euclidian distance is d_M. The two least significant bits (LSBs), i_2 and q_2, determine which of the four signal points within the cluster is chosen, and their minimum Euclidian distance is d_L. The distance ratio $\alpha = d_M/d_L$ (> 1) determines how much more the MSBs are protected against errors than are the LSBs. Hierarchical 16 QAM has one embedded QPSK subconstellation consisting of four clusters, and thus is denoted by 4/16 QAM.

We consider multiplexing N hierarchical 16 QAM constellations, all of which have distinct minimum distances. The average power per symbol of all the multiplexed constellations, S_{avg}, is given by

$$S_{avg} = \frac{1}{N}\sum_{i=1}^{N} S_{avg,i} \tag{2.1}$$

where $S_{avg,i}$ is the average power per symbol of constellation i. For hierarchical 16 QAM, $S_{avg,i}$ is given by

$$S_{avg,i} = \left(\frac{d_{M,i}}{2}\right)^2 + \left(\frac{d_{M,i}}{2} + d_{L,i}\right)^2 = \frac{d_{M,i}^2}{2} + d_{M,i}d_{L,i} + d_{L,i}^2 \tag{2.2}$$

where $d_{M,i}$ and $d_{L,i}$ are minimum distances for the MSBs and LSBs of constellation i, respectively. The BERs of the MSBs and LSBs of hierarchical 16 QAM constellation i, denoted by $P_{M,i}$ and $P_{L,i}$, respectively, are given by [49]

$$\begin{aligned} P_{M,i} &= \frac{1}{2}Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2}Q\left(\left(\frac{d_{M,i}}{2}+d_{L,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ P_{L,i} &= Q\left(\frac{d_{L,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2}Q\left(\left(d_{M,i}+\frac{d_{L,i}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ &\quad - \frac{1}{2}Q\left(\left(d_{M,i}+\frac{3d_{L,i}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \end{aligned} \tag{2.3}$$

where S_{avg} is given by (2.1) and (2.2), γ_s is the signal-to-noise ratio (SNR) per symbol, and $Q(x) = 1/\sqrt{2\pi}\int_x^\infty e^{-y^2/2}dy$.

The following theorem states that $2N$ levels of UEP can be achieved by multiplexing N hierarchical 16 QAM constellations.

Theorem 1: For N hierarchical 16 QAM constellations, $P_{M,i}$ and $P_{L,i}$, given by (2.3), satisfy

$$P_{M,1} < P_{M,2} < \cdots < P_{M,N} < P_{L,1} < P_{L,2} < \cdots < P_{L,N} \tag{2.4}$$

for all SNR if

$$d_{M,1} > d_{M,2} > \cdots > d_{M,N} > d_{L,1} > d_{L,2} > \cdots > d_{L,N}. \tag{2.5}$$

Proof: We will first show that, for $1 \le i, j \le N$,

$$P_{M,i} < P_{L,j} \quad \text{if} \quad d_{M,i} > d_{L,j}. \tag{2.6}$$

Since $Q(x)$ is a monotonically decreasing function, from (2.3), we have

$$P_{M,i} < \frac{1}{2}Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2}Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) = Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \tag{2.7}$$

If $d_{M,i} > d_{L,j}$, from (2.3) and (2.7), we have

$$P_{M,i} < Q\left(\frac{d_{L,j}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) < P_{L,j}. \tag{2.8}$$

We next show that, for $d_{M,1} > d_{M,2} > \cdots > d_{M,N}$ and $d_{L,1} > d_{L,2} > \cdots > d_{L,N}$,

$$P_{M,1} < P_{M,2} < \cdots < P_{M,N}. \tag{2.9}$$

Consider two constellations i and $i+1$ among N hierarchical constellations ($1 \le i \le N-1$). From (2.3), we have $P_{M,i} < P_{M,i+1}$ if $d_{M,i} > d_{M,i+1}$ and $d_{L,i} > d_{L,i+1}$.

Lastly, we show that for $d_{M,1} > d_{M,2} > \cdots > d_{M,N}$ and $d_{L,1} > d_{L,2} > \cdots > d_{L,N}$,

$$P_{L,1} < P_{L,2} < \cdots < P_{L,N}. \tag{2.10}$$

We define a function $f(x,y)$ as

$$f(x,y) = Q\left(\frac{y}{2}\right) + \frac{1}{2}Q\left(x+\frac{y}{2}\right) - \frac{1}{2}Q\left(x+\frac{3y}{2}\right). \tag{2.11}$$

$f(x,y)$ is a monotonically decreasing function of $x > 0$ and $y > 0$, since

$$\begin{aligned}
\frac{\partial f(x,y)}{\partial x} &= \frac{-1}{2\sqrt{2\pi}}\left[e^{-\frac{1}{2}\left(x+\frac{y}{2}\right)^2} - e^{-\frac{1}{2}\left(x+\frac{3y}{2}\right)^2}\right] < 0, \quad \text{and} \\
\frac{\partial f(x,y)}{\partial y} &= \frac{-1}{2\sqrt{2\pi}}\left[e^{-\frac{1}{2}\left(\frac{y}{2}\right)^2} - e^{-\frac{1}{2}\left(x+\frac{3y}{2}\right)^2} + \frac{1}{2}\left\{e^{-\frac{1}{2}\left(x+\frac{y}{2}\right)^2} - e^{-\frac{1}{2}\left(x+\frac{3y}{2}\right)^2}\right\}\right] \\
&< 0.
\end{aligned} \tag{2.12}$$

From (2.3) and (2.11), it is seen that $P_{L,i} = f\left(d_{M,i}\sqrt{2\gamma_s/S_{avg}},\ d_{L,i}\sqrt{2\gamma_s/S_{avg}}\right)$. Hence, from (2.12), we have

$$P_{L,i} < P_{L,i+1} \quad \text{if} \quad d_{M,i} > d_{M,i+1} \ \text{ and } \ d_{L,i} > d_{L,i+1}. \tag{2.13}$$

Finally, (2.4) and (2.5) are derived from (2.6), (2.9) and (2.10).

□

Theorem 1 tells us that $2N$ levels of UEP are achieved by multiplexing N hierarchical 16 QAM constellations having the minimum distances satisfying (2.5).

Corollary 2: Suppose that there are $2N$ unequally important data classes to be transmitted, and class i is more important than class $i+1$ for $1 \le i \le 2N-1$. Let P_i denote the BER of data class i. Then,

$$P_1 < P_2 < \cdots < P_{2N} \tag{2.14}$$

is satisfied for all SNR if the following conditions hold:

i) Class i and class $N+i$ are mapped to the MSBs and LSBs of constellation i, respectively, ($1 \le i \le N$).

ii) The minimum Euclidian distances of the constellations satisfy (2.5).

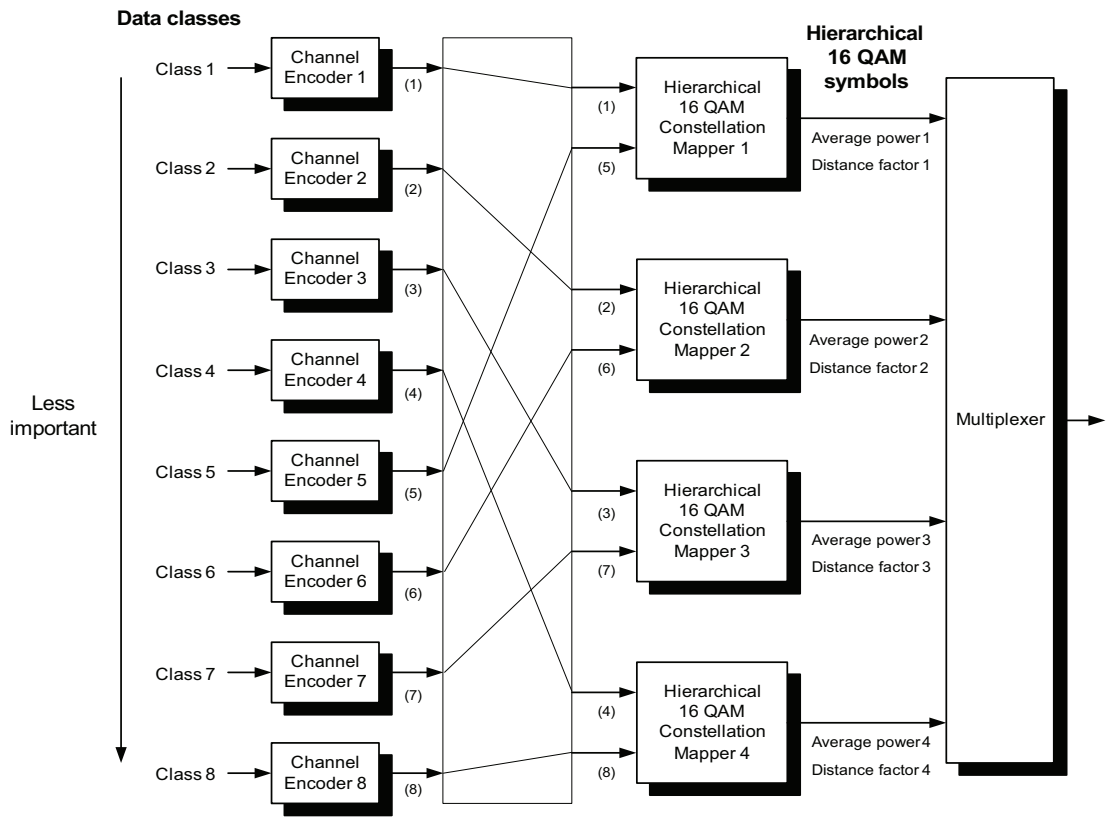


Figure 2.2: The multilevel UEP system using multiplexed hierarchical 16 QAM constellations based on Corollary 2 .

Proof: If i) is satisfied, P_i is given by

$$P_i = P_{M,i} \quad \text{and} \quad P_{N+i} = P_{L,i} \quad (1 \le i \le N). \tag{2.15}$$

If ii) is satisfied, we have $P_{M,1} < P_{M,2} < \cdots < P_{M,N} < P_{L,1} < P_{L,2} < \cdots < P_{L,N}$ from Theorem 1 .

□

Fig. 2.2 depicts the multilevel UEP system using multiplexed hierarchical 16 QAM constellations based on Corollary 2 for eight data classes ($N = 4$).

2.2.2 Hierarchical 2^{2K} $(K \ge 3)$ QAM Constellation

Next, we consider multiplexing hierarchical 2^{2K} $(K \ge 3)$ QAM constellations. As an example, Fig. 2.3 depicts a hierarchical 64 QAM constellation ($K = 3$). The two MSBs i_1 and q_1 determine the quadrant of the first cluster, and their minimum Euclidian

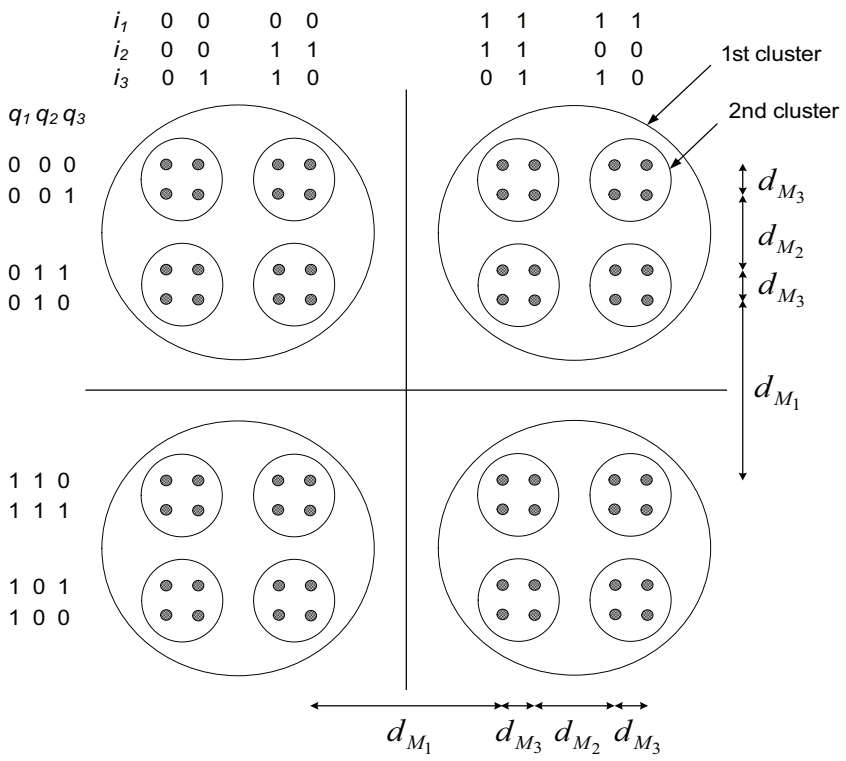


Figure 2.3: Hierarchical 64 QAM constellation.

distance is d_{M1}. The second two MSBs i_2 and q_2 determine the quadrant within the first cluster, and their minimum distance is d_{M2}. Lastly, the third two MSBs (or LSBs) i_3 and q_3 determine the symbol within the second cluster, and their minimum distance is d_{M3}. Hierarchical 64 QAM has two embedded subconstellations, and thus is denoted by 4/16/64 QAM. The hierarchical 64 QAM operates as QPSK when channel conditions are poor, and it operates as 16 or 64 QAM when channel quality gets better. The BER of hierarchical 2^{2K} QAM, P_{M_n}, is given by a recursive expression in [49].

In the following theorem, the BERs of hierarchical 2^{2K} QAM are derived under some assumption based on the fact that for hierarchical constellations, minimum distance for more important bits is greater than that for less important bits.

Theorem 3: Let d_{M_n} denote the minimum distance for the nth MSBs ($1 \leq n \leq K$). Note that the distance ratio of the hierarchical constellation, $d_{M_{n-1}}/d_{M_n}$, is greater than unity ($2 \leq n \leq K$). If the SNR of interest for the nth MSBs is sufficiently large so that the probability of the noise exceeding the Euclidian distance of $d_{M_{n-1}} + \frac{1}{2}d_{M_n}$ is insignificant compared to that of the noise exceeding $\frac{1}{2}d_{M_n}$, the BER of the nth MSBs

$(2 \le n \le K)$, P_{M_n}, becomes

$$P_{M_n}^{app} = \begin{cases} \sum_{p=0}^{2^{K-n}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), \\ \text{for } 2 \le n \le K-1 \\ Q\left(\frac{d_{M_K}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2}Q\left(\left(d_{M_{K-1}} + \frac{d_{M_K}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), \\ \text{for } n = K \end{cases} \tag{2.16}$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x, and $S_{avg} = \sum_{u=1}^{K}\sum_{v=u}^{K} \mu_{uv} d_{M_u} d_{M_v}$ is the average power of a hierarchical 2^{2K} QAM, where the μ_{uv} are constants. Note that for the MSBs (i.e., $n = 1$), the top line of (2.16) is the exact BER expression when n is set to unity (i.e., $P_{M_1}^{app} = P_{M_1}$).

Proof: See Appendix A.

□

$P_{M_n}^{app}$ is numerically evaluated for hierarchical 64 and 256 QAM in Appendix B as an example. For both constellations, $P_{M_n}^{app}$ $(2 \le n \le K)$ is shown to be close to the exact BER within 0.001 dB for BER ≤ 0.1 even at the lower bound of the distance ratio (i.e., $d_{M_{n-1}}/d_{M_n} = 1$). Note that for reference, the distance ratio of hierarchial 16 and 64 QAM in the DVB-T standard [24] is 2 or 4.

For N multiplexed hierarchical 2^{2K} QAM constellations, the average power per symbol of constellation i is given by

$$S_{avg,i} = \sum_{u=1}^{K}\sum_{v=u}^{K} \mu_{uv} d_{M_u,i} d_{M_v,i} \tag{2.17}$$

where $d_{M_n,i}$ $(1 \le n \le K)$ is the minimum distance for the nth MSBs of constellation i $(1 \le i \le N)$, and the μ_{uv} are constants. When the condition of Theorem 3 is satisfied, from (2.1), (2.16) and (2.17), the BER of the nth MSBs $(2 \le n \le K)$ of a hierarchical 2^{2K} QAM constellation i, $P_{M_n,i}$, becomes

$$P_{M_n,i}^{app} = \begin{cases} \sum_{p=0}^{2^{K-n}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), \\ \text{for } \ 2 \le n \le K-1 \\ Q\left(\frac{d_{M_k,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2}Q\left(\left(d_{M_{K-1},i} + \frac{d_{M_K,i}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), \\ \text{for } \ n = K. \end{cases} \tag{2.18}$$

Note that the top line of (2.18) is the exact BER expression when n is set to unity (i.e., $P_{M_1,i}^{app} = P_{M_1,i}$).

Theorem 4: For N hierarchical 2^{2K} QAM constellations, $P_{M_n,i}^{app}$, given by (2.18), satisfy

$$P_{M_1,1}^{app} < \cdots < P_{M_1,N}^{app} < P_{M_2,1}^{app} < \cdots < P_{M_2,N}^{app} < \cdots < P_{M_K,1}^{app} < \cdots < P_{M_K,N}^{app} \tag{2.19}$$

$$\text{if} \quad d_{M_1,1} > \cdots > d_{M_1,N} > d_{M_2,1} > \cdots > d_{M_2,N} > \cdots > d_{M_K,1} > \cdots > d_{M_K,N}. \tag{2.20}$$

Proof: We will first show that, for $1 \le i, j \le N$,

$$\begin{aligned} &P_{M_1,i}^{app} < P_{M_2,j}^{app},\ P_{M_2,i}^{app} < P_{M_3,j}^{app}, \cdots, P_{M_{K-1},i}^{app} < P_{M_K,j}^{app} \\ \text{if} \quad &d_{M_1,i} > d_{M_2,j},\ d_{M_2,i} > d_{M_3,j}, \cdots, d_{M_{K-1},i} > d_{M_K,j}. \end{aligned} \tag{2.21}$$

From (2.18), $P_{M_n,i}^{app}$ $(1 \le n \le K-2)$ can be expressed as

$$\begin{aligned} P_{M_n,i}^{app} = &\sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{2r+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ &+ \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{2r+1+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \end{aligned} \tag{2.22}$$

Eq. (2.22) can be rewritten as

$$\begin{aligned} P_{M_n,i}^{app} = &\sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{r+2^{K-q-1}}{2^{K-q}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ &+ \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{r+2^{-1}+2^{K-q-1}}{2^{K-q}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \end{aligned} \tag{2.23}$$

From (2.23), since $r + 2^{K-q-1}$ and 2^{K-q} are integers for $q \le K-1$, we have

$$\left\lfloor \frac{r+2^{-1}+2^{K-q-1}}{2^{K-q}} \right\rfloor = \left\lfloor \frac{r+2^{K-q-1}}{2^{K-q}} \right\rfloor \quad \text{for } q \le K-1. \tag{2.24}$$

From (2.23), for $q = K$, we have

$$\left\lfloor \frac{r+2^{K-q-1}}{2^{K-q}} \right\rfloor = r \ \text{ and } \ \left\lfloor \frac{r+2^{-1}+2^{K-q-1}}{2^{K-q}} \right\rfloor = r+1. \tag{2.25}$$

From (2.24) and (2.25), (2.23) can be rewritten as

$$P_{M_n,i}^{app} = \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K-1} \left\lfloor \frac{r+2^{K-q-1}}{2^{K-q}} \right\rfloor d_{M_q,i} + r d_{M_K,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K-1} \left\lfloor \frac{r+2^{K-q-1}}{2^{K-q}} \right\rfloor d_{M_q,i} + (r+1) d_{M_K,i}\right) \times\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \tag{2.26}$$

Setting $t = q+1$, $P_{M_n,i}^{app}$ $(1 \le n \le K-2)$, given by (2.26), can be expressed as

$$P_{M_n,i}^{app} = \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{t=n+2}^{K} \left\lfloor \frac{r+2^{K-t}}{2^{K-t+1}} \right\rfloor d_{M_{t-1},i} + r d_{M_K,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \sum_{r=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{t=n+2}^{K} \left\lfloor \frac{r+2^{K-t}}{2^{K-t+1}} \right\rfloor d_{M_{t-1},i} + (r+1) d_{M_K,i}\right) \times\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \tag{2.27}$$

From (2.18), $P_{M_{n+1},j}^{app}$ $(1 \le n \le K-2)$ can be rewritten as

$$P_{M_{n+1},j}^{app} = \sum_{p=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_{n+1},j}}{2} + \sum_{q=n+2}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,j}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \sum_{p=0}^{2^{K-n-1}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_{n+1},j}}{2} + \sum_{q=n+2}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,j}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \tag{2.28}$$

From (2.27) and (2.28), for $1 \le n \le K-2$, we have

$$P_{M_n,i}^{app} < P_{M_{n+1},j}^{app} \text{ if } d_{M_n,i} > d_{M_{n+1},j},\ d_{M_{n+1},i} > d_{M_{n+2},j}, \cdots, d_{M_{K-1},i} > d_{M_K,j}. \tag{2.29}$$

From (2.18), $P_{M_{K-1},i}^{app}$ is given by

$$P_{M_{K-1},i}^{app} = \frac{1}{2} Q\left(\frac{d_{M_{K-1},i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) + \frac{1}{2} Q\left(\left(\frac{d_{M_{K-1},i}}{2} + d_{M_K,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \tag{2.30}$$

From (2.18) and (2.30), we have

$$P^{app}_{M_{K-1},i} < P^{app}_{M_K,j} \quad \text{if} \quad d_{M_{K-1},i} > d_{M_K,j}. \tag{2.31}$$

From (2.29) and (2.31), (2.21) is derived.

We next show that

$$\begin{aligned} &P^{app}_{M_1,1} < \cdots < P^{app}_{M_1,N},\ P^{app}_{M_2,1} < \cdots < P^{app}_{M_2,N}, \cdots, P^{app}_{M_K,1} < \cdots < P^{app}_{M_K,N} \\ \text{if} \quad &d_{M_1,1} > \cdots > d_{M_1,N},\ d_{M_2,1} > \cdots > d_{M_2,N}, \cdots, d_{M_K,1} > \cdots < d_{M_K,N}. \end{aligned} \tag{2.32}$$

We define a function $f(x_n, x_{n+1}, \cdots, x_K)$ as

$$f(x_n, x_{n+1}, \cdots, x_K) = \sum_{p=0}^{2^{K-n}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{x_n}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor x_q\right)\right). \tag{2.33}$$

The $f(x_n, x_{n+1}, \cdots, x_K)$ is a monotonically decreasing function of $x_n > 0, x_{n+1} > 0, \cdots, x_K > 0$, since

$$\begin{aligned} \frac{\partial f(x_n, x_{n+1}, \cdots, x_K)}{\partial x_n} &= \frac{-1}{2\sqrt{2\pi}} \sum_{p=0}^{2^{K-n}-1} \frac{1}{2^{K-n}} e^{-\frac{1}{2}\left(\frac{x_n}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor x_q\right)^2} < 0, \quad \text{and} \\ \frac{\partial f(x_n, x_{n+1}, \cdots, x_K)}{\partial x_{n+m}} &= \frac{-1}{\sqrt{2\pi}} \sum_{p=0}^{2^{K-n}-1} \frac{1}{2^{K-n}} e^{-\frac{1}{2}\left(\frac{x_n}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor x_q\right)^2} \\ &\quad \times \left\lfloor \frac{p+2^{K-n-m}}{2^{K-n-m+1}} \right\rfloor < 0 \end{aligned} \tag{2.34}$$

for $m = 1, \cdots, K-n$ (i.e., for $x_{n+1}, \cdots, x_K$). From (2.18) and (2.33), it is seen that for $1 \le n \le K-1$,

$$P^{app}_{M_n,i} = f\left(d_{M_n,i}\sqrt{\frac{2\gamma_s}{S_{avg}}},\ d_{M_{n+1},i}\sqrt{\frac{2\gamma_s}{S_{avg}}}, \cdots, d_{M_K,i}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \tag{2.35}$$

From (2.34) and (2.35), for $1 \le n \le K-1$, we have

$$P^{app}_{M_n,i} < P^{app}_{M_n,i+1} \ \text{ if } \ d_{M_n,i} > d_{M_n,i+1},\ d_{M_{n+1},i} > d_{M_{n+1},i+1}, \cdots, d_{M_K,i} > d_{M_K,i+1}. \tag{2.36}$$

From (2.18), for $n = K$, we have

$$P^{app}_{M_K,i} < P^{app}_{M_K,i+1} \ \text{ if } \ d_{M_{K-1},i} > d_{M_{K-1},i+1} \quad \text{and} \quad d_{M_K,i} > d_{M_K,i+1}. \tag{2.37}$$

From (2.36) and (2.37), the following is derived.

$$\begin{aligned} &P^{app}_{M_1,i} < P^{app}_{M_1,i+1},\ P^{app}_{M_2,i} < P^{app}_{M_2,i+1}, \cdots, P^{app}_{M_K,i} < P^{app}_{M_K,i+1} \\ \text{if} \quad &d_{M_1,i} > d_{M_1,i+1},\ d_{M_2,i} > d_{M_2,i+1}, \cdots, d_{M_K,i} > d_{M_K,i+1}. \end{aligned} \tag{2.38}$$

With $i = 1, \cdots, N-1$, (2.38) leads to (2.32). Finally, from (2.21) and (2.32), (2.19) and (2.20) are derived.

□

Theorem 4 tells us that, by multiplexing N hierarchical 2^{2K} $(K \geq 3)$ QAM constellations having the minimum distances satisfying (2.20), KN levels of UEP are achieved under the assumption that the SNR of interest for the nth MSBs $(2 \leq n \leq K)$ is reasonably large so that the condition of Theorem 3 is satisfied.

2.3 Optimal Multiplexing of Hierarchical QAM Constellations for High SNR

In this section, we define high SNR as an SNR which is sufficiently large so that the BER is dominated by the error function term having the minimum Euclidian distance.

2.3.1 Hierarchical $2^{2J}/2^{2K}$ $(K > J \geq 1)$ QAM Constellation

Hierarchical $2^{2J}/2^{2K}$ QAM refers to a specific kind of hierarchical constellations which provide two levels of UEP. Typical examples are hierarchical 4/16 QAM (i.e., hierarchical 16 QAM) and 4/64 QAM which are employed in DVB-T standard. Similar to Section 2.2, we first analyze a hierarchical 16 QAM as a simple example. For high SNR, from (2.3), the BERs of a hierarchical 16 QAM constellation i $(1 \leq i \leq N)$ are given by

$$P_{M,i} \approx \frac{1}{2} Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \quad \text{and} \quad P_{L,i} \approx Q\left(\frac{d_{L,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \tag{2.39}$$

Theorem 5: Suppose that there are N multiplexed hierarchical 16 QAM constellations, and the minimum distances satisfying (2.5) are given. Also suppose the given minimum distances can be permuted such that $d_{M,1}, \cdots, d_{M,N}$ for the MSBs can be arbitrarily combined with $d_{L,1}, \cdots, d_{L,N}$ for the LSBs. After the distances are permuted, the resultant minimum distances for the MSBs and LSBs of constellation i, denoted by $\tilde{d}_{M,i}$ and $\tilde{d}_{L,i}$, respectively, can be expressed as

$$\tilde{d}_{M,i} = d_{M,i} \quad \text{and} \quad \tilde{d}_{L,\pi(i)} = d_{L,i} \tag{2.40}$$

where $\pi(i)$ is the index of the constellation to which $d_{L,i}$ is permuted. Then, with the

permuted distances given by (2.40), the BERs of the data classes satisfy

$$P_1 < P_2 < \cdots < P_{2N} \tag{2.41}$$

for high SNR if class i and class $N+i$ are mapped to the MSBs of constellation i and the LSBs of constellation $\pi(i)$, respectively $(1 \le i \le N)$.

Proof: After distances are permuted, from (2.39), (2.40) and the mapping condition below (2.41), the BERs of data classes are given by

$$P_i \approx \frac{1}{2} Q\left(\frac{d_{M,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \quad \text{and} \quad P_{N+i} \approx Q\left(\frac{d_{L,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \quad (1 \le i \le N). \tag{2.42}$$

Since $d_{M,N} > d_{L,1}$ from (2.5), and from (2.42), we have $P_N < P_{N+1}$. Since $d_{M,i} > d_{M,i+1}$ and $d_{L,i} > d_{L,i+1}$ $(1 \le i \le N-1)$ from (2.5), and from (2.42), we have

$$P_i < P_{i+1} \quad \text{and} \quad P_{N+i} < P_{N+1+i} \quad (1 \le i \le N-1). \tag{2.43}$$

Since $P_N < P_{N+1}$, and from (2.43), it follows that $P_1 < \cdots < P_N < P_{N+1} < \cdots < P_{2N}$.

□

In contrast to Theorem 1 and Corollary 2 , Theorem 5 tells us that $2N$ levels of UEP are achieved for high SNR even after the minimum distances satisfying (2.5) are arbitrarily permuted.

Corollary 6: From Theorem 5 , when the minimum distances $d_{M,1}, \cdots, d_{M,N}$ and $d_{L,1}, \cdots, d_{L,N}$ are permuted for high SNR, the BERs of the data classes, $P_1, \cdots, P_{2N}$, are unchanged.

Proof: From (2.42), it is seen that P_i $(1 \le i \le 2N)$ is not dependent on the choice of $\pi(i)$.

□

Theorem 7: After the distances are permuted as described in Theorem 5 , the average power of all the multiplexed hierarchical 16 QAM constellations, S_{avg}, given by

$$S_{avg} = \frac{1}{N}\sum_{i=1}^{N} S_{avg,i} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\tilde{d}_{M,i}^2}{2} + \tilde{d}_{M,i}\tilde{d}_{L,i} + \tilde{d}_{L,i}^2\right) \tag{2.44}$$

is minimized if and only if distances are permuted such that $d_{M,i}$ is combined with $d_{L,N+1-i}$ in the same constellation. That is,

$$\tilde{d}_{M,i} = d_{M,i} \quad \text{and} \quad \tilde{d}_{L,i} = d_{L,N+1-i} \quad (1 \le i \le N). \tag{2.45}$$

Proof: We will prove the following by induction on the number of hierarchical constellations: For given distances $d_{M,1} > \cdots > d_{M,N}$ and $d_{L,1} > \cdots > d_{L,N}$,

$$f_N^* = \sum_{i=1}^{N} \left(\frac{d_{M,i}^2}{2} + d_{M,i} d_{L,N+1-i} + d_{L,N+1-i}^2 \right) \tag{2.46}$$

is the minimum of $f_N = \sum_{i=1}^{N} \left(\tilde{d}_{M,i}^2/2 + \tilde{d}_{M,i} \tilde{d}_{L,i} + \tilde{d}_{L,i}^2 \right)$.

Consider two constellations (i.e., $N = 2$). For given $d_{M,1} > d_{M,2}$ and $d_{L,1} > d_{L,2}$, the distances can be permuted such that $d_{M,1}$ is combined with either $d_{L,1}$ or $d_{L,2}$. The two possible values of f_2 are given by

$$\begin{aligned} f_{2,\sharp 1} &= \frac{d_{M,1}^2}{2} + d_{M,1} d_{L,1} + d_{L,1}^2 + \frac{d_{M,2}^2}{2} + d_{M,2} d_{L,2} + d_{L,2}^2 \\ f_{2,\sharp 2} &= \frac{d_{M,1}^2}{2} + d_{M,1} d_{L,2} + d_{L,2}^2 + \frac{d_{M,2}^2}{2} + d_{M,2} d_{L,1} + d_{L,1}^2. \end{aligned} \tag{2.47}$$

The difference between $f_{2,\sharp 1}$ and $f_{2,\sharp 2}$ is given by

$$f_{2,\sharp 1} - f_{2,\sharp 2} = (d_{M,1} - d_{M,2})(d_{L,1} - d_{L,2}) > 0 \tag{2.48}$$

because $d_{M,1} > d_{M,2}$ and $d_{L,1} > d_{L,2}$. From (2.48), it is seen that $f_{2,\sharp 2}$ is the minimum. For $N = 2$, f_2^* given by (2.46) is equal to $f_{2,\sharp 2}$.

Suppose that (2.46) holds when there are l constellations (i.e., $N = l$). In other words, for given $d_{M,1} > \cdots > d_{M,l}$ and $d_{L,1} > \cdots > d_{L,l}$, $f_l^* = \sum_{i=1}^{l} (d_{M,i}^2/2 + d_{M,i} d_{L,l+1-i} + d_{L,l+1-i}^2)$ is the minimum of f_l. Consider $l+1$ constellations (i.e., $N = l+1$). For given $d_{M,1} > \cdots > d_{M,l+1}$ and $d_{L,1} > \cdots > d_{L,l+1}$, we will prove that if f_{l+1} is minimized, $d_{M,1}$ should be combined with $d_{L,l+1}$ in the same constellation by contradicting the following assumption: f_{l+1} is minimized with $d_{M,1}$ and $d_{L,l+1}$ not being combined. By the assumption, $d_{M,1}$ and $d_{L,j}$ (for some j in the range of $1 \le j < l+1$) are combined in some specific constellation, and $d_{M,k}$ and $d_{L,l+1}$ (for some k in the range of $1 < k \le l+1$) are combined in another constellation. The corresponding f_{l+1}, denoted by $f_{l+1,\sharp 1}$, is given by

$$\begin{aligned} f_{l+1,\sharp 1} = &\left(\frac{d_{M,1}^2}{2} + d_{M,1} d_{L,j} + d_{L,j}^2 \right) + \left(\frac{d_{M,k}^2}{2} + d_{M,k} d_{L,l+1} + d_{L,l+1}^2 \right) \\ &+ \sum_{\substack{i=2 \\ i \neq k}}^{l+1} \left(\frac{\tilde{d}_{M,i}^2}{2} + \tilde{d}_{M,i} \tilde{d}_{L,i} + \tilde{d}_{L,i}^2 \right) \end{aligned} \tag{2.49}$$

where the other minimum distances, except $d_{M,1}, d_{M,k}, d_{L,j}$, and $d_{L,l+1}$, are arbitrarily combined. We modify $f_{l+1,\sharp 1}$ such that $d_{M,1}$ and $d_{L,l+1}$ are combined, and $d_{M,k}$ and $d_{L,j}$

are combined. The modified f_{l+1} is denoted by $f_{l+1,\sharp 2}$:

$$f_{l+1,\sharp 2} = \left(\frac{d_{M,1}^2}{2} + d_{M,1}d_{L,l+1} + d_{L,l+1}^2\right) + \left(\frac{d_{M,k}^2}{2} + d_{M,k}d_{L,j} + d_{L,j}^2\right) + \sum_{\substack{i=2 \\ i\neq k}}^{l+1} \left(\frac{\tilde{d}_{M,i}^2}{2} + \tilde{d}_{M,i}\tilde{d}_{L,i} + \tilde{d}_{L,i}^2\right). \tag{2.50}$$

The difference between $f_{l+1,\sharp 1}$ and $f_{l+1,\sharp 2}$ is given by

$$f_{l+1,\sharp 1} - f_{l+1,\sharp 2} = (d_{M,1} - d_{M,k})(d_{L,j} - d_{L,l+1}) > 0 \tag{2.51}$$

because $d_{M,1} > d_{M,k}$ and $d_{L,j} > d_{L,l+1}$. From (2.51), $f_{l+1,\sharp 1}$, given by (2.49), cannot be the minimum of f_{l+1}, and thus the above assumption is false. We have thus showed that the largest distance for the MSBs, $d_{M,1}$ should be combined with the smallest distance for the LSBs, $d_{L,l+1}$. The other minimum distances, except $d_{M,1}$ and $d_{L,l+1}$, are given by

$$d_{M,2} > d_{M,3} > \cdots > d_{M,l+1} \quad \text{and} \quad d_{L,1} > d_{L,2} > \cdots > d_{L,l}. \tag{2.52}$$

By the induction hypothesis, the following is the minimum for $2l$ distances given by (2.52):

$$\sum_{i=1}^{l} \left(\frac{d_{M,i+1}^2}{2} + d_{M,i+1}d_{L,l+1-i} + d_{L,l+1-i}^2\right). \tag{2.53}$$

Thus, the minimum of f_{l+1} is given by

$$\frac{d_{M,1}^2}{2} + d_{M,1}d_{L,l+1} + d_{L,l+1}^2 + \sum_{i=1}^{l} \left(\frac{d_{M,i+1}^2}{2} + d_{M,i+1}d_{L,l+1-i} + d_{L,l+1-i}^2\right) = \sum_{i=1}^{l+1} \left(\frac{d_{M,i}^2}{2} + d_{M,i}d_{L,l+2-i} + d_{L,l+2-i}^2\right). \tag{2.54}$$

Setting $N = l + 1$ in (2.46), we obtain $f_{l+1}^* = \sum_{i=1}^{l+1}(d_{M,i}^2/2 + d_{M,i}d_{L,l+2-i} + d_{L,l+2-i}^2)$, and this is identical to (2.54). Hence, (2.46) holds for $N = l + 1$.

□

Corollary 6 and Theorem 7 indicate that the average power of all the multiplexed constellations is minimized by permuting distances according to (2.45), while the BERs are unchanged for high SNR.

Next, we consider the peak signal power of the multiplexed hierarchical constellations. If we assume that all the hierarchical constellations are time-multiplexed, the peak power of all the multiplexed constellations, S_{peak}, is given by

$$S_{peak} = \max\left[\left\{ S_{peak,i} \middle| 1 \le i \le N \right\}\right] \tag{2.55}$$

where $\max[X]$ denotes the maximum element of the set X, and $S_{peak,i}$ is the peak power of a hierarchical constellation i. For hierarchical 16 QAM, $S_{peak,i}$ is given by

$$S_{peak,i} = 2\left(\frac{d_{M,i}}{2} + d_{L,i}\right)^2 = \frac{d_{M,i}^2}{2} + 2d_{M,i}d_{L,i} + 2d_{L,i}^2. \tag{2.56}$$

Theorem 8: After the distances are permuted as described in Theorem 5 , the peak power of all the multiplexed hierarchical 16 QAM constellations, S_{peak}, given by

$$S_{peak} = \max\left[\left\{ S_{peak,i} \middle| 1 \le i \le N \right\}\right] = \max\left[\left\{ \frac{\tilde{d}_{M,i}^2}{2} + 2\tilde{d}_{M,i}\tilde{d}_{L,i} + 2\tilde{d}_{L,i}^2 \middle| 1 \le i \le N \right\}\right] \tag{2.57}$$

is minimized if the distances are permuted according to (2.45) of Theorem 7

Proof: When (2.45) is satisfied, the corresponding S_{peak}, denoted by $S_{peak,\sharp 1}$, is given by

$$\begin{aligned} S_{peak,\sharp 1} &= \max\left[\left\{ \frac{d_{M,i}^2}{2} + 2d_{M,i}d_{L,N+1-i} + 2d_{L,N+1-i}^2 \middle| 1 \le i \le N \right\}\right] \\ &= \frac{d_{M,j}^2}{2} + 2d_{M,j}d_{L,N+1-j} + 2d_{L,N+1-j}^2, \end{aligned} \tag{2.58}$$

for some j in the range of $1 \le j \le N$. We will contradict the following assumption: When distances are permuted in some way other than (2.45), the corresponding S_{peak}, denoted by $S_{peak,\sharp 2}$, is smaller than $S_{peak,\sharp 1}$. Let $d_{L,k}$ be the distance with which $d_{M,j}$ is combined (for some k in the range of $1 \le k \le N$) when the distances are permuted in a different manner from (2.45). The possible values of k can be classified into

$$1 \le k < N+1-j, \quad k = N+1-j, \quad \text{and} \quad N+1-j < k \le N. \tag{2.59}$$

i) For $1 \le k < N+1-j$, $S_{peak,\sharp 2} > S_{peak,\sharp 1}$. To see this, note that

$$\begin{aligned} S_{peak,\sharp 2} &\ge \frac{1}{2}d_{M,j}^2 + 2d_{M,j}d_{L,k} + 2d_{L,k}^2 \\ &> \frac{1}{2}d_{M,j}^2 + 2d_{M,j}d_{L,N+1-j} + 2d_{L,N+1-j}^2 \\ &= S_{peak,\sharp 1} \end{aligned} \tag{2.60}$$

where the strict inequality follows from $d_{L,k} > d_{L,N+1-j}$ (since $k < N+1-j$).
ii) For $k = N+1-j$, $S_{peak,\sharp 2} \geq S_{peak,\sharp 1}$ since $S_{peak,\sharp 2} \geq \frac{1}{2}d_{M,j}^2 + 2d_{M,j}d_{L,N+1-j} + 2d_{L,N+1-j}^2 = S_{peak,\sharp 1}$.
iii) For $N+1-j < k \leq N$, $S_{peak,\sharp 2} > S_{peak,\sharp 1}$. This is proved as follows: Since $d_{M,j}$ is combined with $d_{L,k}$, other distances $\{d_{M,i} | 1 \leq i \leq N,\ i \neq j\}$ should be combined with $\{d_{L,i} | 1 \leq i \leq N,\ i \neq k\}$. Note that

$$\left| \left\{ d_{M,i} \middle| 1 \leq i < j \right\} \right| = j-1 \ \text{ and } \ \left| \left\{ d_{L,i} \middle| N+1-j < i \leq N,\ i \neq k \right\} \right| = j-2 \quad (2.61)$$

where $|X|$ denotes the cardinality of the set X, and the equality of the second expression follows from $N+1-j < k \leq N$. Since $j-1 > j-2$ in (2.61), at least one element of $\{d_{M,i} | 1 \leq i < j\}$ should be combined with one element of $\{d_{L,i} | 1 \leq i \leq N+1-j\}$. Suppose that $d_{M,p}$ is combined with $d_{L,q}$ for some $p \in \{1, \cdots, j-1\}$ and $q \in \{1, \cdots, N+1-j\}$. Then, we have

$$\begin{aligned} S_{peak,\sharp 2} &\geq \frac{1}{2}d_{M,p}^2 + 2d_{M,p}d_{L,q} + 2d_{L,q}^2 \\ &> \frac{1}{2}d_{M,j}^2 + 2d_{M,j}d_{L,N+1-j} + 2d_{L,N+1-j}^2 \\ &= S_{peak,\sharp 1} \end{aligned} \quad (2.62)$$

where the strict inequality follows from the fact that $d_{M,p} > d_{M,j}$ and $d_{L,q} \geq d_{L,N+1-j}$ (since $p < j$ and $q \leq N+1-j$). From i), ii), and iii), it is seen that there is no possible way of permuting distances which makes $S_{peak,\sharp 2}$ smaller than $S_{peak,\sharp 1}$. Therefore, the assumption below (2.58) is false.

□

Theorems 7 and 8 tell us that the permutation of the distances that minimizes the average power of all the multiplexed hierarchical constellations also, coincidentally, minimizes the peak power. Note that from (2.5) and (2.45), these optimally permuted distances satisfy

$$\tilde{d}_{M,1} > \cdots > \tilde{d}_{M,N} > \tilde{d}_{L,N} > \cdots > \tilde{d}_{L,1}. \quad (2.63)$$

Corollary 9: When the distances are optimally permuted according to (2.45) of Theorem 7 , the BERs of the data classes satisfy $P_1 < P_2 < \cdots < P_{2N}$ for high SNR if class i and class $2N+1-i$ are mapped to the MSBs and LSBs of constellation i, respectively ($1 \leq i \leq N$).

Proof: The proof is similar to the proof of Corollary 2 .

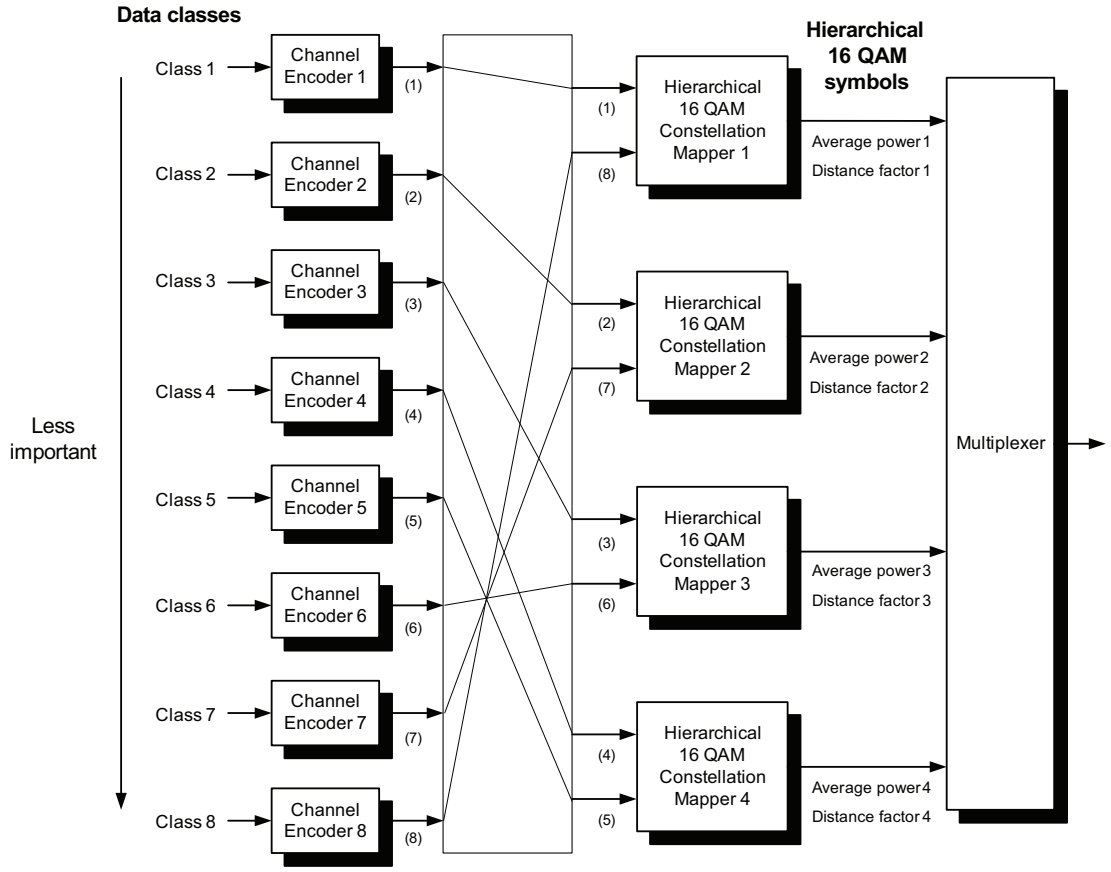


Figure 2.4: The optimal multilevel UEP system using multiplexed hierarchical 16 QAM constellations for high SNR based on Corollary 9 .

□

Fig. 2.4 depicts the multilevel UEP system using multiplexed hierarchical 16 QAM constellations based on Corollary 9 for eight data classes ($N = 4$).

Next, we generalize to hierarchical $2^{2J}/2^{2K}$ ($K > J \geq 1$) QAM constellations. Recall that $d_{M_n,i}$ denotes the minimum distance for the nth MSBs ($1 \leq n \leq K$) of a hierarchical 2^{2K} QAM constellation i. Hierarchical $2^{2J}/2^{2K}$ QAM has two distinct minimum Euclidian distances such that [49]

$$d_{M_n,i} = \begin{cases} d_{M_J,i}, & \text{for } 1 \leq n \leq J \\ d_{M_K,i}, & \text{for } J+1 \leq n \leq K. \end{cases} \tag{2.64}$$

The average power of a hierarchical $2^{2J}/2^{2K}$ QAM constellation i ($1 \leq i \leq N$) can be

expressed, from (2.17) and (2.64), as the following:

$$S_{avg,i} = \sum_{u=1}^{J}\sum_{v=u}^{J} \mu_{uv} d_{M_J,i}^2 + \sum_{u=1}^{J}\sum_{v=J+1}^{K} \mu_{uv} d_{M_J,i} d_{M_K,i} + \sum_{u=J+1}^{K}\sum_{v=u}^{K} \mu_{uv} d_{M_K,i}^2. \quad (2.65)$$

Lemma 10: For high SNR, the BERs of a hierarchical $2^{2J}/2^{2K}$ QAM constellation i $(1 \le i \le N)$ are given by

$$P_{M_n,i} \approx \begin{cases} \frac{1}{2^{K-n}} Q\left(\frac{d_{M_J,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), & \text{for } 1 \le n \le J \\ \frac{1}{2^{K-n}} Q\left(\frac{d_{M_K,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right), & \text{for } J+1 \le n \le K \end{cases} \quad (2.66)$$

where S_{avg} is given by (2.1) and (2.65).

Proof: The BERs of a hierarchical 2^{2K} QAM constellation i, $P_{M_n,i}^{app}$ $(1 \le n \le K-1)$, given by (2.18), can be rewritten as

$$\begin{aligned} P_{M_n,i}^{app} &= \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ &\quad + \sum_{p=1}^{2^{K-n}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ &= \frac{1}{2^{K-n}} Q\left(\frac{d_{M_n,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \\ &\quad + \sum_{p=1}^{2^{K-n}-1} \frac{1}{2^{K-n}} Q\left(\left(\frac{d_{M_n,i}}{2} + \sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i}\right)\sqrt{\frac{2\gamma_s}{S_{avg}}}\right). \end{aligned} \quad (2.67)$$

From (2.67), we have

$$\sum_{q=n+1}^{K} \left\lfloor \frac{p+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i} \ge \sum_{q=n+1}^{K} \left\lfloor \frac{1+2^{K-q}}{2^{K-q+1}} \right\rfloor d_{M_q,i} \ge \left\lfloor \frac{1+2^0}{2^1} \right\rfloor d_{M_K,i} = d_{M_K,i} \quad (2.68)$$

where the first inequality follows from $p \ge 1$ in (2.67). From (2.67) and (2.68), it is clear that the first error function term of (2.67) is the only term having the minimum distance of $d_{M_n,i}$ for the nth MSBs $(1 \le n \le K-1)$. Also, for $P_{M_K,i}^{app}$ (i.e., $n = K$) given by (2.18), it is clear that the first error function term is the only term having the minimum distance of $d_{M_K,i}$. From the condition of approximation described in Theorem 3 , it follows that the error function term having the minimum distance in $P_{M_K,i}^{app}$, given by (2.18), is the same as that in $P_{M_n,i}$, the exact BER. Therefore, from (2.64) and (2.67), (2.66) is derived.

□

From (2.66), the average BER for $n = 1, \cdots J$th MSBs of constellation i , denoted by $(P_{M_J,i})_{avg}$, is given by

$$(P_{M_J,i})_{avg} = \frac{1}{J}\sum_{n=1}^{J} P_{M_n,i} \approx A_J Q\left(\frac{d_{M_J,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \tag{2.69}$$

where $A_J = \frac{1}{J}\sum_{n=1}^{J} 1/2^{K-n}$. Likewise, the average BER for $n = J+1, \cdots, K$th MSBs of constellation i, denoted by $(P_{M_K,i})_{avg}$, is given by

$$(P_{M_K,i})_{avg} = \frac{1}{K-J}\sum_{n=J+1}^{K} P_{M_n,i} \approx A_K Q\left(\frac{d_{M_K,i}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}}}\right) \tag{2.70}$$

where $A_K = \frac{1}{K-J}\sum_{n=J+1}^{K} 1/2^{K-n}$. Similar to the average power given by (2.65), the peak power of a hierarchical $2^{2J}/2^{2K}$ QAM constellation i $(1 \le i \le N)$ can be expressed as

$$\begin{aligned} S_{peak,i} &= \sum_{u=1}^{K}\sum_{v=u}^{K} \lambda_{uv} d_{M_u,i} d_{M_v,i} \\ &= \sum_{u=1}^{J}\sum_{v=u}^{J} \lambda_{uv} d_{M_J,i}^2 + \sum_{u=1}^{J}\sum_{v=J+1}^{K} \lambda_{uv} d_{M_J,i} d_{M_K,i} + \sum_{u=J+1}^{K}\sum_{v=u}^{K} \lambda_{uv} d_{M_K,i}^2 \end{aligned} \tag{2.71}$$

where the λ_{uv} are constants.

Theorem 11: Theorems 5 , 7 and 8 , and Corollary 6 hold for hierarchical $2^{2J}/2^{2K}$ QAM when

i) $d_{M,i}$ and $d_{L,i}$ are replaced by $d_{M_J,i}$ and $d_{M_K,i}$, respectively, and $P_{M,i}$ and $P_{L,i}$ are replaced by $(P_{M_J,i})_{avg}$ and $(P_{M_K,i})_{avg}$, respectively.

ii) Eq. (2.2) and (2.56) are replaced by (2.65) and (2.71), respectively.

Proof: From (2.69) and (2.70), $A_J < A_K$ since

$$A_J = \frac{1}{J}\sum_{n=1}^{J}\frac{1}{2^{K-n}} < \frac{1}{2^{K-J}} \quad \text{and} \quad A_K = \frac{1}{K-J}\sum_{n=J+1}^{K}\frac{1}{2^{K-n}} > \frac{1}{2^{K-J-1}}. \tag{2.72}$$

Hence, Theorem 5 and Corollary 6 hold for hierarchical $2^{2J}/2^{2K}$ QAM.

Since $\sum_{u=1}^{J}\sum_{v=u}^{J}\mu_{uv}$, $\sum_{u=1}^{J}\sum_{v=J+1}^{K}\mu_{uv}$ and $\sum_{u=J+1}^{K}\sum_{v=u}^{K}\mu_{uv}$ of (2.65) are constants just as 1/2, 1, and 1 of (2.2) are constants, Theorem 7 holds for hierarchical $2^{2J}/2^{2K}$ QAM. Likewise, $\sum_{u=1}^{J}\sum_{v=u}^{J}\lambda_{uv}$, $\sum_{u=1}^{J}\sum_{v=J+1}^{K}\lambda_{uv}$, and $\sum_{u=J+1}^{K}\sum_{v=u}^{K}\lambda_{uv}$ of (2.71) are constants as 1/2, 2, and 2 of (2.56) are constants, and thus Theorem 8 holds for hierarchical $2^{2J}/2^{2K}$ QAM.

□

2.4 Asymmetric Hierarchical QAM Constellation

While the proposed methods provide a large number of levels of UEP, the peak-to-average power ratio (PAPR) typically will be increased when hierarchical constellations having distinct minimum distances are time-multiplexed. To mitigate this effect, we design an asymmetric hierarchical QAM which reduces the PAPR without performance loss. From here onwards, we refer to conventional hierarchical QAM, which has been presented in Sections 2.2 and 2.3, as symmetric hierarchical QAM, in order to distinguish it from asymmetric hierarchical QAM.

2.4.1 Asymmetric Hierarchical 2^{2K} $(K \geq 2)$ QAM Constellation

For an asymmetric hierarchical 2^{2K} QAM, the minimum distances for the inphase and quadrature components are different from each other. Similar to the previous sections, we first present asymmetric hierarchical 16 QAM, depicted in Fig. 2.5, as a simple example. The MSB i_1 for the inphase component determines the first cluster, and its minimum distance is $d_M^{A,I}$. The MSB q_1 for the quadrature component determines the second cluster within the first cluster that i_1 determined, and its minimum distance is $d_M^{A,Q}$. The LSB i_2 for the inphase component determines the third cluster, and its minimum distance is $d_L^{A,I}$, and the LSB q_2 for the quadrature component determines the specific signal point within the third cluster, and has minimum distance $d_L^{A,Q}$. Asymmetric hierarchical 16 QAM has three embedded subconstellations, and it provides four levels of UEP if $d_M^{A,I} > d_M^{A,Q} > d_L^{A,I} > d_L^{A,Q}$, which will be shown below in Corollary 13 .

In order to provide $2N$ levels of UEP, we consider multiplexing $N/2$ (N is assumed to be even) asymmetric hierarchical 16 QAM constellations instead of N symmetric hierarchical 16 QAM constellations. The average power per symbol of all the multiplexed asymmetric constellations, S_{avg}^A, is given by

$$S_{avg}^A = \frac{1}{N/2} \sum_{i=1}^{N/2} S_{avg,i}^A \tag{2.73}$$

where $S_{avg,i}^A$ is the average power per symbol of asymmetric constellation i. For

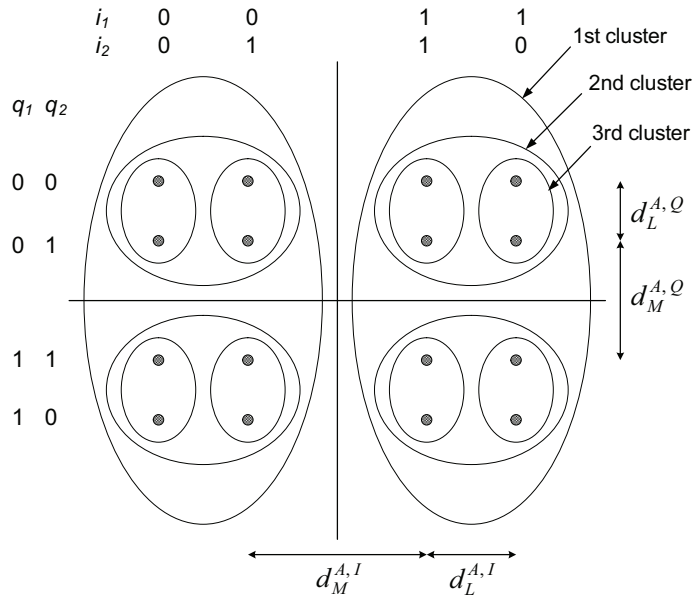


Figure 2.5: Asymmetric hierarchical 16 QAM constellation.

asymmetric hierarchical 16 QAM, $S_{avg,i}^{A}$ is given by

$$\begin{aligned} S_{avg,i}^{A} &= S_{avg,i}^{A,I} + S_{avg,i}^{A,Q} \\ &= \frac{1}{2}\left(\left(\frac{d_{M,i}^{A,I}}{2}\right)^2 + \left(\frac{d_{M,i}^{A,I}}{2} + d_{L,i}^{A,I}\right)^2\right) + \frac{1}{2}\left(\left(\frac{d_{M,i}^{A,Q}}{2}\right)^2 + \left(\frac{d_{M,i}^{A,Q}}{2} + d_{L,i}^{A,Q}\right)^2\right) \end{aligned} \tag{2.74}$$

where $S_{avg,i}^{A,I}$ and $S_{avg,i}^{A,Q}$ are the average powers per symbol for the inphase and quadrature components of asymmetric constellation i, respectively, and $d_{M,i}^{A,I}$, $d_{L,i}^{A,I}$, $d_{M,i}^{A,Q}$, and $d_{L,i}^{A,Q}$ are the minimum distances for the inphase MSB and LSB, and quadrature MSB and LSB, respectively. Note that the BERs of rectangular QAM are derived from those of the corresponding PAMs since the inphase and quadrature components are separated at the demodulator [49] [50]. Let $P_{M,i}^{A,I}$, $P_{L,i}^{A,I}$, $P_{M,i}^{A,Q}$, and $P_{L,i}^{A,Q}$ denote the BERs for the inphase MSB and LSB, and quadrature MSB and LSB of asymmetric hierarchical constellation i, respectively ($1 \leq i \leq N/2$). From (2.3), (2.73), and (2.74), they are

derived as

$$\begin{aligned}
P_{M,i}^{A,I} &= \frac{1}{2}Q\left(\frac{d_{M,i}^{A,I}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}^A}}\right)+\frac{1}{2}Q\left(\left(\frac{d_{M,i}^{A,I}}{2}+d_{L,i}^{A,I}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^A}}\right)\\
P_{L,i}^{A,I} &= Q\left(\frac{d_{L,i}^{A,I}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}^A}}\right)+\frac{1}{2}Q\left(\left(d_{M,i}^{A,I}+\frac{d_{L,i}^{A,I}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^A}}\right)\\
&\quad -\frac{1}{2}Q\left(\left(d_{M,i}^{A,I}+\frac{3d_{L,i}^{A,I}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^A}}\right)\\
P_{M,i}^{A,Q} &= \frac{1}{2}Q\left(\frac{d_{M,i}^{A,Q}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}^A}}\right)+\frac{1}{2}Q\left(\left(\frac{d_{M,i}^{A,Q}}{2}+d_{L,i}^{A,Q}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^A}}\right)\\
P_{L,i}^{A,Q} &= Q\left(\frac{d_{L,i}^{A,Q}}{2}\sqrt{\frac{2\gamma_s}{S_{avg}^A}}\right)+\frac{1}{2}Q\left(\left(d_{M,i}^{A,Q}+\frac{d_{L,i}^{A,Q}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^A}}\right)\\
&\quad -\frac{1}{2}Q\left(\left(d_{M,i}^{A,Q}+\frac{3d_{L,i}^{A,Q}}{2}\right)\sqrt{\frac{2\gamma_s}{S_{avg}^A}}\right).
\end{aligned} \tag{2.75}$$

Theorem 12: Suppose there are N multiplexed symmetric hierarchical 16 QAM constellations whose minimum distances are given by $d_{M,1},\cdots,d_{M,N}$ and $d_{L,1},\cdots,d_{L,N}$. Also suppose there are $N/2$ asymmetric hierarchical 16 QAM constellations, and the minimum distances for the inphase and quadrature components of asymmetric hierarchical constellation i are the same as those of two distinct symmetric hierarchical constellations $x(i)$ and $y(i)$, respectively ($1\le i\le N/2$). In other words,

$$d_{M,i}^{A,I}=d_{M,x(i)},\ d_{L,i}^{A,I}=d_{L,x(i)},\ d_{M,i}^{A,Q}=d_{M,y(i)},\text{ and } d_{L,i}^{A,Q}=d_{L,y(i)}\ \ (1\le i\le N/2) \tag{2.76}$$

where $x(i)$ and $y(i)$ satisfy

$$x(i),y(i)\in\{1,\cdots,N\}\ \text{ and }\ \{x(i),y(i)|1\le i\le N/2\}=\{1,\cdots,N\}. \tag{2.77}$$

With the minimum distances given by (2.76), the average power and BERs of $N/2$ multiplexed asymmetric hierarchical 16 QAM constellations are the same as those of N multiplexed symmetric hierarchical 16 QAM constellations, regardless of the choice of $x(i)$ and $y(i)$ satisfying (2.77).

Proof: From (2.74) and (2.76), $S_{avg,i}^A$ can be expressed as

$$\begin{aligned}
S_{avg,i}^A &= \frac{1}{2}\left(\left(\frac{d_{M,x(i)}}{2}\right)^2+\left(\frac{d_{M,x(i)}}{2}+d_{L,x(i)}\right)^2\right)\\
&\quad +\frac{1}{2}\left(\left(\frac{d_{M,y(i)}}{2}\right)^2+\left(\frac{d_{M,y(i)}}{2}+d_{L,y(i)}\right)^2\right)\\
&= \frac{1}{2}S_{avg,x(i)}+\frac{1}{2}S_{avg,y(i)},
\end{aligned} \tag{2.78}$$

where the second equality follows from (2.2). From (2.73) and (2.78), S_{avg}^{A} is given by

$$S_{avg}^{A} = \frac{1}{N/2}\sum_{i=1}^{N/2}\left(\frac{1}{2}S_{avg,x(i)} + \frac{1}{2}S_{avg,y(i)}\right) = \frac{1}{N}\sum_{i=1}^{N/2}\left(S_{avg,x(i)} + S_{avg,y(i)}\right). \quad (2.79)$$

From (2.77), (2.79) can be rewritten as

$$S_{avg}^{A} = \frac{1}{N}\sum_{i=1}^{N} S_{avg,i} = S_{avg} \quad (2.80)$$

where the second equality follows from (2.1). We next compare the BERs of asymmetric and symmetric constellations. From (2.3), (2.75) and (2.76), we have

$$P_{M,i}^{A,I} = P_{M,x(i)},\ P_{L,i}^{A,I} = P_{L,x(i)},\ P_{M,i}^{A,Q} = P_{M,y(i)},\ \text{and}\ P_{L,i}^{A,Q} = P_{L,y(i)} \quad (1 \le i \le N/2). \quad (2.81)$$

From (2.77) and (2.81), a set of $2N$ BERs for $N/2$ multiplexed asymmetric constellations satisfy

$$\left\{P_{M,i}^{A,I}, P_{L,i}^{A,I}, P_{M,i}^{A,Q}, P_{L,i}^{A,Q}\middle| 1 \le i \le N/2\right\} = \left\{P_{M,x(i)}, P_{L,x(i)}, P_{M,y(i)}, P_{L,y(i)}\middle| 1 \le i \le N/2\right\}$$
$$= \left\{P_{M,i}, P_{L,i}\middle| 1 \le i \le N\right\}. \quad (2.82)$$

Hence, a set of $2N$ BERs for $N/2$ multiplexed asymmetric constellations is the same as that for N multiplexed symmetric constellations.

□

Theorem 13: Suppose that the minimum distances of the N multiplexed symmetric hierarchical 16 QAM constellations satisfy (2.5) of Theorem 1 . Then, with the minimum distances given by (2.76), $N/2$ multiplexed asymmetric hierarchical 16 QAM constellations also provide $2N$ levels of UEP.

Proof: Since $d_{M,i}$ and $d_{L,i}$ satisfy (2.5), $P_{M,i}$ and $P_{L,i}$ satisfy (2.4) by Theorem 1 . From (2.82), it follows that $N/2$ multiplexed asymmetric hierarchical 16 QAM constellations also provide $2N$ levels of UEP.

□

As an example, suppose that there is single asymmetric hierarchical 16 QAM (i.e., $N = 2$), and $x(i)$ and $y(i)$ satisfying (2.77) are chosen as $x(1) = 1$ and $y(1) = 2$. From (2.76) and (2.81), (2.4) and (2.5) of Theorem 1 lead to the following:

$$P_{M,1}^{A,I} < P_{M,1}^{A,Q} < P_{L,1}^{A,I} < P_{L,1}^{A,Q} \quad \text{if} \quad d_{M,1}^{A,I} > d_{M,1}^{A,Q} > d_{L,1}^{A,I} > d_{L,1}^{A,Q}. \quad (2.83)$$

Next, we consider the peak power of all the multiplexed asymmetric hierarchical constellations, S^A_{peak}, which is given by

$$S^A_{peak} = \max\left[\left\{ S^A_{peak,i} \middle| 1 \le i \le N/2 \right\}\right] \tag{2.84}$$

where $S^A_{peak,i}$ is the peak power of an asymmetric hierarchical constellation i. For asymmetric hierarchical 16 QAM, $S^A_{peak,i}$ is given by

$$S^A_{peak,i} = S^{A,I}_{peak,i} + S^{A,Q}_{peak,i} = \left(\frac{d^{A,I}_{M,i}}{2} + d^{A,I}_{L,i}\right)^2 + \left(\frac{d^{A,Q}_{M,i}}{2} + d^{A,Q}_{L,i}\right)^2 \tag{2.85}$$

where $S^{A,I}_{peak,i}$ and $S^{A,Q}_{peak,i}$ are the peak powers of the inphase and quadrature components of asymmetric hierarchical constellation i, respectively.

Theorem 14: Suppose that the minimum distances of the N multiplexed symmetric hierarchical 16 QAM satisfy (2.5) of Theorem 1 . With the minimum distances given by (2.76), the peak power of all $N/2$ multiplexed asymmetric hierarchical 16 QAM constellations, S^A_{peak} , given by (2.84) and (2.85), is less than that of all N multiplexed symmetric hierarchical 16 QAM, S_{peak}, given by (2.55) and (2.56), regardless of the choice of $x(i)$ and $y(i)$ satisfying (2.77).

Proof: From (2.76) and (2.85), $S^A_{peak,i}$ is given by

$$S^A_{peak,i} = \left(\frac{d_{M,x(i)}}{2} + d_{L,x(i)}\right)^2 + \left(\frac{d_{M,y(i)}}{2} + d_{L,y(i)}\right)^2 = \frac{1}{2}S_{peak,x(i)} + \frac{1}{2}S_{peak,y(i)} \tag{2.86}$$

where the second equality follows from (2.56). From (2.84) and (2.86), S^A_{peak} is given by

$$\begin{aligned} S^A_{peak} &= \max\left[\left\{ \frac{1}{2}S_{peak,x(i)} + \frac{1}{2}S_{peak,y(i)} \middle| 1 \le i \le N/2 \right\}\right] \\ &= \frac{1}{2}S_{peak,x(j)} + \frac{1}{2}S_{peak,y(j)}, \end{aligned} \tag{2.87}$$

for some j in the range of $1 \le j \le N/2$. Since $x(i), y(i) \in \{1, \cdots, N\}$ from (2.77), we have

$$S_{peak,x(j)} \le \max\left[\left\{ S_{peak,i} \middle| 1 \le i \le N \right\}\right] = S_{peak}, \text{ and } \; S_{peak,y(j)} \le S_{peak}, \tag{2.88}$$

where the second equality of the first expression follows from (2.55). From (2.5) and (2.56), the peak powers of each symmetric hierarchical 16 QAM constellation satisfy

$$S_{peak,1} > S_{peak,2} > \cdots > S_{peak,N}. \tag{2.89}$$

From (2.77), (2.88) and (2.89), $S_{peak,x(j)}$ and $S_{peak,y(j)}$ satisfy either of the following:

$$S_{peak,x(j)} < S_{peak,y(j)} \leq S_{peak} \quad \text{or} \quad S_{peak,y(j)} < S_{peak,x(j)} \leq S_{peak}. \tag{2.90}$$

From (2.87) and (2.90), we have

$$S^A_{peak} = \frac{1}{2} S_{peak,x(j)} + \frac{1}{2} S_{peak,y(j)} < S_{peak}. \tag{2.91}$$

□

Theorems 12 and 14 tell us that when asymmetric hierarchical 16 QAM is used instead of symmetric hierarchical 16 QAM, the PAPR is reduced without performance loss.

The following theorem states how to choose $x(i)$ and $y(i)$ $(1 \leq i \leq N/2)$ satisfying (2.77) to minimize the PAPR of all the multiplexed asymmetric hierarchical constellations.

Theorem 15: Suppose that the minimum distances of the N multiplexed symmetric hierarchical 16 QAM satisfy (2.5) of Theorem 1 . Also suppose the minimum distances of $N/2$ multiplexed asymmetric hierarchical 16 QAM are given by (2.76). Then, from (2.84) and (2.86), S^A_{peak} is given by

$$S^A_{peak} = \max \left[\left\{ \frac{1}{2} S_{peak,x(i)} + \frac{1}{2} S_{peak,y(i)} \middle| 1 \leq i \leq N/2 \right\} \right] \tag{2.92}$$

and this is minimized if $x(i)$ and $y(i)$ satisfying (2.77) are chosen as

$$x(i) = i \quad \text{and} \quad y(i) = N + 1 - i \quad (1 \leq i \leq N/2). \tag{2.93}$$

Proof: The proof is similar to the proof of Theorem 8 .

□

Next, we generalize to asymmetric hierarchical 2^{2K} $(K \geq 2)$ QAM. Let $d^{A,I}_{M_n,i}$ and $d^{A,Q}_{M_n,i}$ denote the minimum distances of the nth MSB $(1 \leq n \leq K)$ for the inphase and quadrature components of asymmetric hierarchical 2^{2K} QAM constellation i $(1 \leq i \leq N/2)$. From (2.17), the average power of asymmetric hierarchical 2^{2K} QAM constellation i, $S^A_{avg,i}$, can be expressed as

$$S^A_{avg,i} = S^{A,I}_{avg,i} + S^{A,Q}_{avg,i} = \sum_{u=1}^{K} \sum_{v=u}^{K} \frac{\mu_{uv}}{2} d^{A,I}_{M_u,i} d^{A,I}_{M_v,i} + \sum_{u=1}^{K} \sum_{v=u}^{K} \frac{\mu_{uv}}{2} d^{A,Q}_{M_u,i} d^{A,Q}_{M_v,i} \tag{2.94}$$

where $S^{A,I}_{avg,i}$ and $S^{A,Q}_{avg,i}$ are the average powers for the inphase and quadrature components of asymmetric constellation i.

Let $P_{M_n,i}^{A,I}$ and $P_{M_n,i}^{A,Q}$ denote the BERs of the nth MSB ($1 \le n \le K$) for the inphase and quadrature components of asymmetric hierarchical 2^{2K} QAM constellation i ($1 \le i \le N/2$). Recall that $P_{M_n,i}$ denotes the BER of the nth MSBs ($1 \le n \le K$) of symmetric hierarchical 2^{2K} QAM constellation i ($1 \le i \le N$).

Theorem 16: Suppose that there are N multiplexed symmetric hierarchical 2^{2K} QAM whose minimum distances are given by $d_{M_n,1}, \cdots, d_{M_n,N}$ ($1 \le n \le K$). Also suppose that the minimum distances of $N/2$ multiplexed asymmetric hierarchical 2^{2K} QAM satisfy

$$d_{M_n,i}^{A,I} = d_{M_n,x(i)} \quad \text{and} \quad d_{M_n,i}^{A,Q} = d_{M_n,y(i)} \quad (1 \le n \le K,\ 1 \le i \le N/2) \tag{2.95}$$

where $x(i)$ and $y(i)$ satisfy (2.77). Theorem 12 holds for asymmetric hierarchical 2^{2K} QAM when

i) $d_{M,i}^{A,I}$ and $d_{L,i}^{A,I}$ are replaced by $d_{M_n,i}^{A,I}$ ($1 \le n \le K$); $d_{M,i}^{A,Q}$ and $d_{L,i}^{A,Q}$ are replaced by $d_{M_n,i}^{A,Q}$ ($1 \le n \le K$); $d_{M,i}$ and $d_{L,i}$ are replaced by $d_{M_n,i}$ ($1 \le n \le K$).

ii) $P_{M,i}^{A,I}$ and $P_{L,i}^{A,I}$ are replaced by $P_{M_n,i}^{A,I}$ ($1 \le n \le K$); $P_{M,i}^{A,Q}$ and $P_{L,i}^{A,Q}$ are replaced by $P_{M_n,i}^{A,Q}$ ($1 \le n \le K$); $P_{M,i}$ and $P_{L,i}$ are replaced by $P_{M_n,i}$ ($1 \le n \le K$).

iii) Eq. (2.76) is replaced by (2.95).

Proof: For asymmetric hierarchical 2^{2K} QAM, from (2.94) and (2.95), $S_{avg,i}^{A}$ is given by

$$\begin{aligned} S_{avg,i}^{A} &= \sum_{u=1}^{K}\sum_{v=u}^{K} \frac{\mu_{uv}}{2} d_{M_u,x(i)} d_{M_v,x(i)} + \sum_{u=1}^{K}\sum_{v=u}^{K} \frac{\mu_{uv}}{2} d_{M_u,y(i)} d_{M_v,y(i)} \\ &= \frac{1}{2} S_{avg,x(i)} + \frac{1}{2} S_{avg,y(i)}, \end{aligned} \tag{2.96}$$

where the second equality follows from (2.17). From (2.95), we have

$$P_{M_n,i}^{A,I} = P_{M_n,x(i)} \quad \text{and} \quad P_{M_n,i}^{A,Q} = P_{M_n,y(i)} \quad (1 \le n \le K,\ 1 \le i \le N/2) \tag{2.97}$$

From (2.78), (2.81), (2.96) and (2.97), it follows that Theorem 12 holds for asymmetric hierarchical 2^{2K} QAM.

□

We next consider the peak power for asymmetric hierarchical 2^{2K} QAM. In the following, we rewrite the peak power of symmetric hierarchical 2^{2K} QAM constellation i ($1 \le i \le N$), $S_{peak,i}$, given by (2.71):

$$S_{peak,i} = \sum_{u=1}^{K}\sum_{v=u}^{K} \lambda_{uv} d_{M_u,i} d_{M_v,i}. \tag{2.98}$$

From (2.98), the peak power of asymmetric hierarchical 2^{2K} QAM constellation i, $S^A_{peak,i}$, can be expressed as

$$S^A_{peak,i} = S^{A,I}_{peak,i} + S^{A,Q}_{peak,i} = \sum_{u=1}^{K}\sum_{v=u}^{K} \frac{\lambda_{uv}}{2} d^{A,I}_{M_u,i} d^{A,I}_{M_v,i} + \sum_{u=1}^{K}\sum_{v=u}^{K} \frac{\lambda_{uv}}{2} d^{A,Q}_{M_u,i} d^{A,Q}_{M_v,i} \quad (2.99)$$

where $S^{A,I}_{peak,i}$ and $S^{A,Q}_{peak,i}$ are the peak powers for the inphase and quadrature components of asymmetric constellation i.

Theorem 17: Theorems 14 and 15 hold for asymmetric hierarchical 2^{2K} QAM when

i) $d^{A,I}_{M,i}$ and $d^{A,I}_{L,i}$ are replaced by $d^{A,I}_{M_n,i}$ $(1 \le n \le K)$; $d^{A,Q}_{M,i}$ and $d^{A,Q}_{L,i}$ are replaced by $d^{A,Q}_{M_n,i}$ $(1 \le n \le K)$; $d_{M,i}$ and $d_{L,i}$ are replaced by $d_{M_n,i}$ $(1 \le n \le K)$.

ii) $S_{peak,i}$ given by (2.56) is replaced by (2.98).

iii) $S^A_{peak,i}$ given by (2.85) is replaced by (2.99).

iv) Eq. (2.5) of Theorem 1 is replaced by (2.20) of Theorem 4 .

v) Eq. (2.76) is replaced by (2.95).

Proof: For asymmetric hierarchical 2^{2K} QAM, from (2.95) and (2.99), $S^A_{peak,i}$ is given by

$$\begin{aligned} S^A_{peak,i} &= \sum_{u=1}^{K}\sum_{v=u}^{K} \frac{\lambda_{uv}}{2} d_{M_u,x(i)} d_{M_v,x(i)} + \sum_{u=1}^{K}\sum_{v=u}^{K} \frac{\lambda_{uv}}{2} d_{M_u,y(i)} d_{M_v,y(i)} \\ &= \frac{1}{2} S_{peak,x(i)} + \frac{1}{2} S_{peak,y(i)}, \end{aligned} \quad (2.100)$$

where the second equality follows from (2.98). From (2.20) and (2.98), the peak powers of each symmetric hierarchical 2^{2K} QAM constellation satisfy

$$S_{peak,1} > S_{peak,2} > \cdots > S_{peak,N}. \quad (2.101)$$

From (2.86), (2.89), (2.100), and (2.101), it follows that Theorems 14 and 15 hold for asymmetric hierarchical 2^{2K} QAM.

□

2.5 Multilevel UEP Based on Multiplexing Hierarchical QAM Constellations Having Constant Power

In this section, we consider the case where it is desirable for the multiplexed hierarchical QAM constellations to have the same average power (i.e., constant power),

either due to the limited capability of a power amplifier, or for cochannel interference control.

2.5.1 Symmetric Hierarchical $2^{2J}/2^{2K}$ $(K > J \geq 1)$ QAM Constellation

Theorem 18: When N multiplexed symmetric hierarchical 16 QAM constellations are required to have constant power, there exist minimum distances satisfying

$$d_{M,1} > d_{M,2} > \cdots > d_{M,N} > d_{L,N} > d_{L,N-1} > \cdots > d_{L,1}. \tag{2.102}$$

Proof: Since all the multiplexed constellations have the same average power, we have

$$S_{avg,1} = S_{avg,2} = \cdots = S_{avg,N}. \tag{2.103}$$

From (2.2) and (2.103), it is clear that there exist distances $d_{M,1}, d_{M,2}, \cdots, d_{M,N}$ and $d_{L,1}, d_{L,2}, \cdots, d_{L,N}$ satisfying (2.102).

□

From (2.63) and (2.102), it is seen that even if symmetric hierarchical 16 QAM constellations have constant power, the suggested UEP system, depicted in Fig. 2.4, can provide $2N$ levels of UEP for high SNR.

Theorem 18 holds for symmetric hierarchical $2^{2J}/2^{2K}$ $(K > J \geq 1)$ QAM, when $d_{M,i}$ and $d_{L,i}$ are replaced by $d_{M_J,i}$ and $d_{M_K,i}$, respectively.

2.5.2 Asymmetric Hierarchical 2^{2K} $(K \geq 2)$ QAM Constellation

Theorem 19: Suppose that $N/2$ multiplexed asymmetric hierarchical 16 QAM constellations are required to have constant power, and their minimum distances are given by (2.76). If $x(i)$ and $y(i)$ are chosen according to (2.93) of Theorem 15 , there exist minimum distances satisfying both (2.5) of Theorem 1 and (2.76).

Proof: Since all the multiplexed constellations have the same average power, we have

$$S^A_{avg,1} = S^A_{avg,2} = \cdots = S^A_{avg,N/2}. \tag{2.104}$$

Recall that, with the minimum distances given by (2.76), $S^A_{avg,i}$ can be expressed as the combination of $S_{avg,x(i)}$ and $S_{avg,y(i)}$, as given by (2.78). Using (2.78), (2.104) can be rewritten as

$$S_{avg,x(1)} + S_{avg,y(1)} = S_{avg,x(2)} + S_{avg,y(2)} = \cdots = S_{avg,x(N/2)} + S_{avg,y(N/2)}. \tag{2.105}$$

If we let $x(i) = i$ and $y(i) = N + 1 - i$ according to (2.93), we have

$$S_{avg,1} + S_{avg,N} = S_{avg,2} + S_{avg,N-1} = \cdots = S_{avg,N/2} + S_{avg,N/2+1}. \quad (2.106)$$

From (2.106), it is clear that there exist $S_{avg,1}, S_{avg,2}, \cdots, S_{avg,N}$ satisfying

$$S_{avg,1} > S_{avg,2} > \cdots > S_{avg,N}. \quad (2.107)$$

From (2.2) and (2.107), it is clear that there exist minimum distances $d_{M,1}, d_{M,2}, \cdots, d_{M,N}$ and $d_{L,1}, d_{L,2}, \cdots, d_{L,N}$ satisfying $d_{M,1} > d_{M,2} > \cdots > d_{M,N} > d_{L,1} > d_{L,2} > \cdots > d_{L,N}$ (i.e., (2.5) of Theorem 1 .

□

From Corollary 13 and Theorem 19 , it follows that even if asymmetric hierarchical 16 QAM constellations have constant power, $2N$ levels of UEP can be achieved.

Theorem 19 holds for asymmetric hierarchical 2^{2K} ($K \geq 3$) QAM, when

i) $d_{M,i}$ and $d_{L,i}$ are replaced by $d_{M_n,i}$ ($1 \leq n \leq K$).

ii) Eq. (2.76) is replaced by (2.95).

iii) Eq. (2.5) of Theorem 1 is replaced by (2.20) of Theorem 4 .

Theorem 20: Suppose that $N/2$ multiplexed asymmetric hierarchical 2^{2K} ($K \geq 2$) QAM constellations are required to have constant power. Then the performance of the system stays the same or degrades compared to the case where multiplexed constellations are not required to have constant power.

Proof: When $N/2$ multiplexed asymmetric hierarchical 2^{2K} ($K \geq 2$) QAM constellations have constant power, we have

$$S^A_{avg,1} = S^A_{avg,2} = \cdots = S^A_{avg,N/2} = C \quad (2.108)$$

where C is a constant. On the other hand, when multiplexed constellations can have variable powers under the constraint that average power of all of them is C, we have

$$S^A_{avg} = \frac{1}{N/2} \sum_{i=1}^{N/2} S^A_{avg,i} = C \quad \text{or} \quad S^A_{avg,1} + S^A_{avg,2} + \cdots + S^A_{avg,N/2} = \frac{N}{2} C. \quad (2.109)$$

From (2.94), the set of values which $d^{A,I}_{M_n,i}$ and $d^{A,Q}_{M_n,i}$ ($1 \leq n \leq K$, $1 \leq i \leq N/2$) can have under the constraint of (2.108) is a subset of that under the constraint of (2.109). Note that the BERs of the multiplexed constellations are functions of the minimum distances,

$d^{A,I}_{M_n,i}$ and $d^{A,Q}_{M_n,i}$, and that the performance of the system is a function of the BERs. Hence, the set of values which the minimum distances can have is the domain of the system objective function. Since the domain under the constraint of (2.108) is a subset of that under the constraint of (2.109), the range of the system objective function under (2.108) is also a subset of that under (2.109).

□

2.6 The Performance of the Proposed UEP System for Progressive Bitstream Transmission

In this section, we analyze the performance of the proposed UEP system for progressive image source transmission over Rayleigh fading channels. We first consider the UEP system depicted in Fig. 2.2. The system takes successive blocks (data classes) of the compressed progressive bitstream, and transforms them into a sequence of channel codewords of fixed length l_c [37] with error detection and correction capability. Then, the coded classes are mapped to the multiplexed symmetric hierarchical 16 QAM constellations. At the receiver, if a received class is correctly decoded, then the next class is considered by the decoder. Otherwise, the decoding is stopped and the image is reconstructed from the correctly decoded classes. We assume that all decoding errors can be detected.

Let r_i be an error correction code rate for class i $(1 \leq i \leq 2N)$, and $\underline{d}_i = (d_{M,c(i)},\ d_{L,c(i)})$ be a pair of minimum distances of some specific constellation $c(i)$ $(1 \leq c(i) \leq N)$ to which class i $(1 \leq i \leq 2N)$ is mapped. From Corollary 2 , $\underline{d}_i$ $(1 \leq i \leq 2N)$ is given by

$$\underline{d}_i = \begin{cases} (d_{M,i},\ d_{L,i}), & \text{for } 1 \leq i \leq N \\ (d_{M,i-N},\ d_{L,i-N}), & \text{for } N+1 \leq i \leq 2N \end{cases} \tag{2.110}$$

where $d_{M,1}, \cdots, d_{M,N}$ and $d_{L,1}, \cdots, d_{L,N}$ satisfy (2.5) of Theorem 1 to achieve $2N$ levels of UEP. Let $p(r_i, \underline{d}_i, \gamma_s)$ denote the probability of a decoding error of class i. Then, the probability that no decoding errors occur in the first i classes with an error in the next one, $P_{c,i}$ is given by

$$P_{c,i} = p(r_{i+1}, \underline{d}_{i+1}, \gamma_s) \prod_{j=1}^{i} \big(1 - p(r_j, \underline{d}_j, \gamma_s)\big) \quad \text{for } 1 \leq i \leq 2N-1. \tag{2.111}$$

Note that $P_{c,0} = p(r_1, \underline{d}_1, \gamma_s)$ is the probability of an error in the first class, and $P_{c,2N} = \prod_{j=1}^{2N} \left(1 - p(r_j, \underline{d}_j, \gamma_s)\right)$ is the probability that all $2N$ classes are correctly decoded. The end-to-end performance can be measured by the expected distortion, $E[D]$, given by

$$E[D] = \sum_{i=0}^{2N} P_{c,i} D_i \tag{2.112}$$

where D_i is the reconstruction error using the first i classes ($1 \leq i \leq 2N$), and D_0 is a constant. For the case of an uncoded system, D_i is given by $D_i = V(il_c)$, where $V(x)$ denotes the operational rate-distortion function of the source coder. Also, for the uncoded system, the probability of a decoding error of class i, $p(r_i, \underline{d}_i, \gamma_s) = p(\underline{d}_i, \gamma_s)$, can be obtained analytically:

$$p(\underline{d}_i, \gamma_s) = 1 - \{1 - P_i(\underline{d}_i, \gamma_s)\}^{l_c}. \tag{2.113}$$

Recall that P_i, a function of $\underline{d}_i$ and γ_s, is the BER of data class i. P_i ($1 \leq i \leq 2N$) is given by (2.3) and (2.15) of Corollary 2 . We define a frame as a group of constellation symbols to which one image bitstream is mapped. We assume the channel experiences slow Rayleigh fading such that the fading coefficients are nearly constant over a frame. With this channel model, from (2.111)–(2.113), the expected distortion for the uncoded system is given by

$$\begin{aligned} E[D] = \int_0^\infty \Bigg\{ & \left(1 - \left\{1 - P_1(\underline{d}_1, h^2\gamma_s)\right\}^{l_c}\right) V(0) \\ & + \sum_{i=1}^{2N-1} \left[\left(1 - \left\{1 - P_{i+1}(\underline{d}_{i+1}, h^2\gamma_s)\right\}^{l_c}\right) \prod_{j=1}^{i} \left\{1 - P_j(\underline{d}_j, h^2\gamma_s)\right\}^{l_c} \right] V(il_c) \\ & + \prod_{j=1}^{2N} \left\{1 - P_j(\underline{d}_j, h^2\gamma_s)\right\}^{l_c} V(2Nl_c) \Bigg\} f(h) dh \end{aligned} \tag{2.114}$$

where h is the Rayleigh-distributed envelope of complex channel coefficients and $f(h)$ is the Rayleigh-distributed probability density function of h. Note that for a given SNR of γ_s, $E[D]$ is the conditional expected distortion. In situations when exact SNR information is not available at the transmitter, one can find the minimum distances, $\underline{d}_1, \cdots, \underline{d}_{2N}$ (or $d_{M,1}, \cdots, d_{M,N}$ and $d_{L,1}, \cdots, d_{L,N}$), which minimize the expected distortion over a

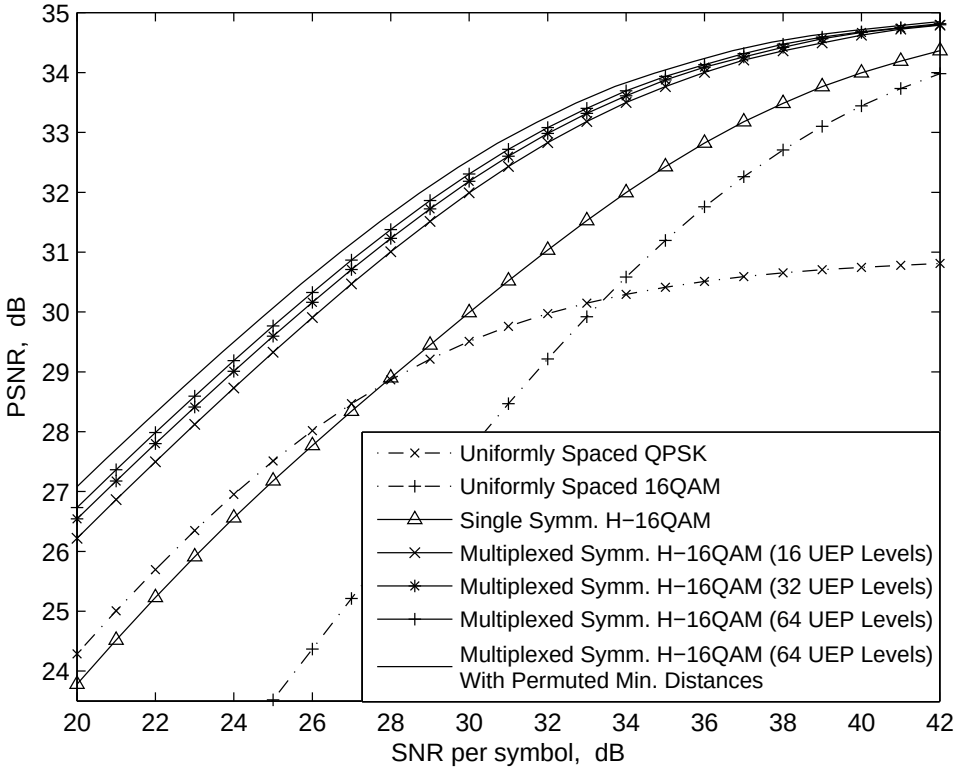


Figure 2.6: PSNR performance of UEP system using multiplexed symmetric hierarchical 16 QAM (H-16QAM denotes hierarchical 16 QAM).

range of expected SNRs using the weighted cost function

$$\underset{\underline{d}_1,\cdots,\underline{d}_{2N}}{\arg\min} \frac{\int_0^\infty \omega(\gamma_s) E[D] d\gamma_s}{\int_0^\infty \omega(\gamma_s) d\gamma_s} \tag{2.115}$$

where $\omega(\gamma_s)$ in [0, 1] is the weight function. For example, $\omega(\gamma_s)$ can be given by

$$\omega(\gamma_s) = \begin{cases} 1, & \text{for } \gamma_s^a \le \gamma_s \le \gamma_s^b \\ 0, & \text{otherwise.} \end{cases} \tag{2.116}$$

2.7 Numerical Results

We evaluate the performance of the proposed UEP system using multiplexed hierarchical 16 QAM constellations for the progressive source coder SPIHT [10] as an example. We provide the results for the standard 8 bits per pixel (bpp) 512×512 Lena

Table 2.1: PAPR of Multiplexed Symmetric or Asymmetric Hierarchical 16 QAM

	PAPR (dB)		
Number of UEP levels	4	16	64
Multiplexed symmetric hierarchical 16 QAM	3.31	6.87	9.43
Multiplexed symmetric hierarchical 16 QAM with permuted min. distances	2.82	5.84	8.32
Multiplexed asymmetric hierarchical 16 QAM	1.11	4.18	6.60
Multiplexed asymmetric hierarchical 16 QAM having constant power	1.11	1.43	1.46

Table 2.2: PAPR of Uniformly Spaced 16 QAM and Single Symmetric Hierarchical 16 QAM

	PAPR (dB)
Uniformly spaced 16 QAM	2.55
Single symmetric hierarchical 16 QAM	0.90

image with a transmission rate of 0.375 bpp. To compare the image quality, we use peak-signal-to-noise ratio (PSNR) defined as

$$\text{PSNR} = 10\log\frac{255^2}{E[D]} \quad \text{(dB)} \tag{2.117}$$

where 255 is due to the 8 bpp image, and $E[D]$ is given by (2.114).

We present the PSNR performance for the uncoded case by numerically evaluating (2.114)–(2.117) as follows: We first compute (2.115) for the block Rayleigh fading channel using the expected distortion, $E[D]$, given by (2.114), and the weight function, $\omega(\gamma_s)$, given by (2.116). Next, with $\underline{d}_1, \cdots, \underline{d}_{2N}$ (or $d_{M,1}, \cdots, d_{M,N}$ and $d_{L,1}, \cdots, d_{L,N}$) obtained from (2.115), we evaluate PSNR using (2.114) and (2.117) over a range of expected SNRs given by (2.116).

Fig. 2.6 shows the PSNR performance of the multiplexed symmetric hierarchical 16 QAM constellations. For reference, it also shows PSNRs for single symmetric hierarchical 16 QAM, as well as uniformly spaced QPSK and 16 QAM constellations. The PSNR of single symmetric hierarchical constellation is evaluated in the same way as that for multiplexed symmetric hierarchical constellations. From Fig. 2.6, it is seen

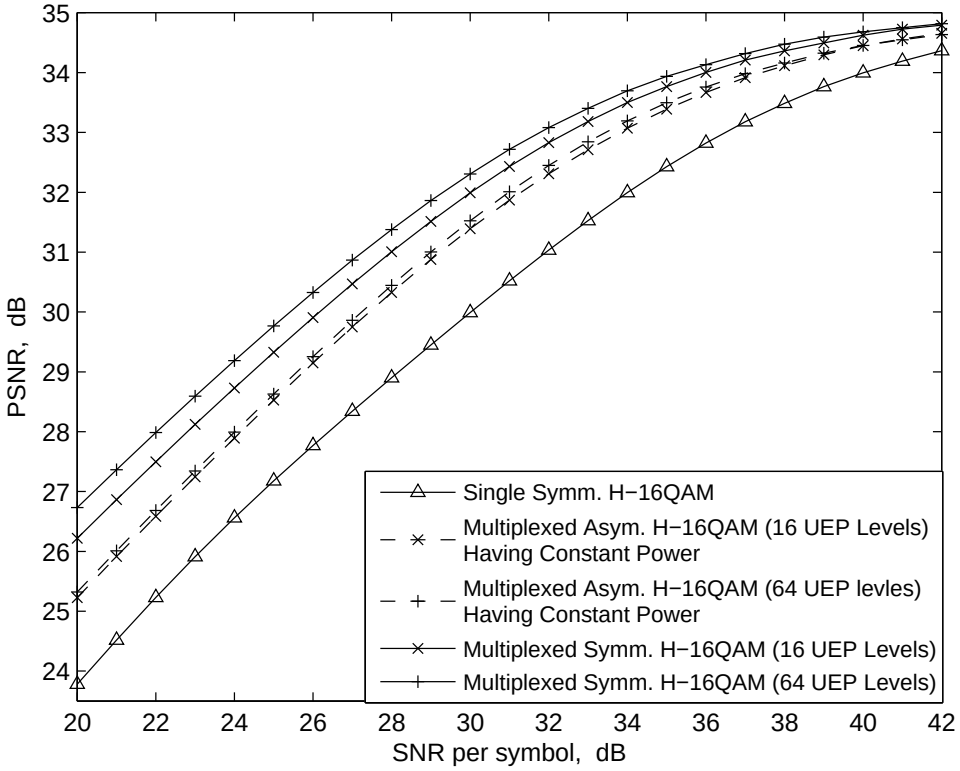


Figure 2.7: PSNR performance of UEP system using multiplexed asymmetric hierarchical 16 QAM having constant power (H-16QAM denotes hierarchical 16 QAM).

that multiplexed symmetric hierarchical constellations improve the performance more than does single symmetric hierarchical constellation. It is also seen that 32 multiplexed symmetric hierarchical 16 QAM constellations, which provide 64 levels of UEP, have almost saturated performance in this evaluation. However, by optimally permuting the minimum distances according to Theorem 7 , an additional SNR gain of more than 0.5 dB is achieved. Note that the performance of $N/2$ multiplexed asymmetric hierarchical constellations is the same as that of N multiplexed symmetric hierarchical constellations (N=8,16,32) as stated by Theorem 12 , though the former is not depicted here.

Table I shows the PAPRs of the multiplexed symmetric or asymmetric hierarchical 16 QAM constellations. For reference, the PAPRs of single symmetric hierarchical 16 QAM and uniformly spaced 16 QAM constellations are also listed in Table II. From Tables I and II, it is seen that when symmetric hierarchical 16 QAM constellations are time-multiplexed, they have larger PAPR than does uniformly spaced 16 QAM as well

as single symmetric hierarchical 16 QAM constellation. Table I also shows that PAPR is reduced when asymmetric hierarchical constellation is used, as stated by Theorem 14

Fig. 2.7 shows the PSNR performance of the multiplexed asymmetric hierarchical 16 QAM constellations having constant power. It is shown that the performance is degraded when constellations are required to have constant power, which is consistent with Theorem 20. However, as seen from Table I, this scheme provides PAPR smaller than uniformly-spaced QAM, and a high PAPR problem is solved.

2.8 Conclusion

Progressive image or scalable video encoders employ progressive transmission, so that encoded data have gradual differences of importance in their bitstreams, which necessitates multiple levels of UEP. Though hierarchical modulation has been intensively studied for digital broadcasting or multimedia transmission, methods of achieving a large number of levels of UEP for progressive mode of transmission have rarely been studied.

In this chapter, we proposed a multilevel UEP system using multiplexed hierarchical modulation for progressive transmission over mobile radio channels. Specifically, we proposed a way of multiplexing N hierarchical 2^{2K} QAM constellations ($K \geq 2$) and proved that KN levels of UEP are achieved, under the assumption that the SNR of interest for the nth most important bits is reasonably large so that the probability of noise exceeding the Euclidian distance of $d_{M_{n-1}} + \frac{1}{2}d_{M_n}$ is insignificant compared to that of noise exceeding $\frac{1}{2}d_{M_n}$, where d_{M_n} and $d_{M_{n-1}}$ are the minimum distances for the nth and $n-1$th important bits, respectively ($2 \leq n \leq K$). This assumption is based on the fact that for hierarchical constellations, the minimum distance for more important bits is greater than that for less important bits (i.e., $d_{M_{n-1}} > d_{M_n}$). As a special case, for hierarchical 16 QAM ($K = 2$), we showed that $2N$ levels of UEP are achieved without the assumption.

When the BER is dominated by the error function term having the minimum Euclidian distance, we derived an optimal multiplexing approach which minimizes both the average and peak powers for hierarchical $2^{2J}/2^{2K}$ QAM ($K > J \geq 1$) constellations (typical examples are 4/16 QAM and 4/64 QAM which are employed in the DVB-T standard). While the suggested methods achieve multiple levels of UEP, the PAPR typically will be increased when constellations having distinct minimum distances are time-multiplexed. To mitigate this effect, an asymmetric hierarchical QAM constellation,

which reduces the PAPR without performance loss, was proposed. We also considered the case where multiplexed constellations need to have constant power, and showed that multilevel UEP can be achieved while the performance stays the same or degrades in this case. Numerical results showed that the proposed multilevel UEP system based on multiplexed modulation significantly enhances the performance for progressive transmission over Rayleigh fading channels without any additional system bandwidth or transmit power.

2.9 Appendix A: Proof of Theorem 3

2.9.1 Gray coded bit mapping vector for hierarchical 2^K PAM

For a hierarchical 2^K PAM constellation, let $g_{n,i}$ denote the Gray code for the nth MSB ($1 \le n \le K$) assigned to the ith signal point ($1 \le i \le 2^K$) from the left. Then, it can be shown that the 2^K-tuple Gray coded bit mapping vector, $\mathbf{g}_n = \left[g_{n,1}\ g_{n,2} \cdots g_{n,2^K}\right]$, for the nth MSB is given by

$$\mathbf{g}_n = \begin{cases} \left[\mathbf{0}_{2^{K-1}}\ \mathbf{1}_{2^{K-1}}\right], & \text{for } n = 1 \\ \left[\mathbf{0}_{2^{K-n}}\ \mathbf{1}_{2^{K-n}}\ \mathbf{1}_{2^{K-n}}\ \mathbf{0}_{2^{K-n}} \cdots \mathbf{0}_{2^{K-n}}\ \mathbf{1}_{2^{K-n}}\ \mathbf{1}_{2^{K-n}}\ \mathbf{0}_{2^{K-n}}\right], & \text{for } 2 \le n \le K \end{cases} \tag{2.118}$$

where $\mathbf{0}_l$ is a l-tuple all zero vector, and $\mathbf{1}_l$ is a l-tuple all one vector.

2.9.2 Euclidian distance between adjacent signal points for hierarchical 2^K PAM

Let $S_i^{(K)}$ ($1 \le i \le 2^K$) and $S_i^{(K+1)}$ ($1 \le i \le 2^{K+1}$) denote the ith signal point from the left for hierarchical 2^K and 2^{K+1} PAM constellations, respectively. Also, let $d_{M_n}^{(K)}$ ($1 \le n \le K$) and $d_{M_n}^{(K+1)}$ ($1 \le n \le K+1$) denote minimum distances for the nth MSB of hierarchical 2^K and 2^{K+1} PAM constellations, respectively. Fig. 2.8 shows how hierarchical 2^{K+1} PAM is constructed from hierarchical 2^K PAM. There are two rules with regard to the construction of hierarchical 2^{K+1} PAM from hierarchical 2^K PAM:

i) The ith signal point for 2^K PAM, $S_i^{(K)}$, is replaced by the $2i-1$th and $2i$th signal points for 2^{K+1} PAM, $S_{2i-1}^{(K+1)}$ and $S_{2i}^{(K+1)}$, which satisfy

$$d\left(S_{2i-1}^{(K+1)}, S_{2i}^{(K+1)}\right) = d_{M_{K+1}}^{(K+1)} \quad \text{for } 1 \le i \le 2^K \tag{2.119}$$

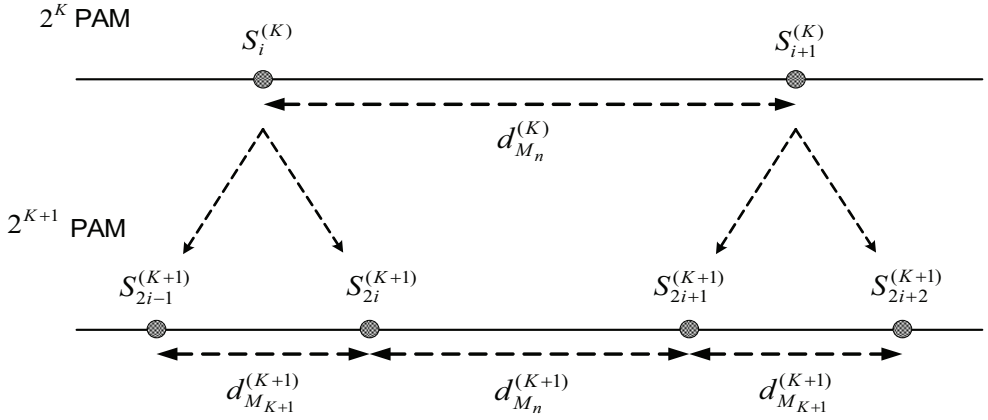


Figure 2.8: The construction of hierarchical 2^{K+1} PAM from hierarchical 2^K PAM.

where $d(X, Y)$ is the Euclidian distance between two signal points, X and Y.

ii) If the distance between $S_i^{(K)}$ and $S_{i+1}^{(K)}$ for 2^K PAM is $d_{M_n}^{(K)}$, then the distance between $S_{2i}^{(K+1)}$ and $S_{2i+1}^{(K+1)}$ for 2^{K+1} PAM is $d_{M_n}^{(K+1)}$. That is, for $1 \le i \le 2^K - 1$ and $1 \le n \le K$,

$$d\Big(S_{2i}^{(K+1)}, S_{2i+1}^{(K+1)}\Big) = d_{M_n}^{(K+1)} \quad \text{if} \quad d\Big(S_i^{(K)}, S_{i+1}^{(K)}\Big) = d_{M_n}^{(K)}. \tag{2.120}$$

As an example, Fig. 2.9 depicts hierarchical 4 and 8 PAM constellations.

We will prove the following by induction: For hierarchical 2^K PAM ($K \ge 2$), the Euclidian distance between adjacent signal points is given by

$$d\Big(S_{(2i-1)2^{K-n}}^{(K)}, S_{(2i-1)2^{K-n}+1}^{(K)}\Big) = d_{M_n}^{(K)} \quad \text{for } 1 \le i \le 2^{n-1} \text{ and } 1 \le n \le K. \tag{2.121}$$

Consider hierarchical 4 PAM. From Fig. 2.9, it is seen that

$$d\Big(S_2^{(2)}, S_3^{(2)}\Big) = d_{M_1}^{(2)} \quad \text{and} \quad d\Big(S_1^{(2)}, S_2^{(2)}\Big) = d\Big(S_3^{(2)}, S_4^{(2)}\Big) = d_{M_2}^{(2)}. \tag{2.122}$$

If we let $K = 2$ in (2.121), we have

$$d\Big(S_{(2i-1)2^{2-n}}^{(2)}, S_{(2i-1)2^{2-n}+1}^{(2)}\Big) = d_{M_n}^{(2)} \quad \text{for } 1 \le i \le 2^{n-1} \text{ and } 1 \le n \le 2. \tag{2.123}$$

From (2.123), for $n = 1$, we have

$$d\Big(S_{(2i-1)2}^{(2)}, S_{(2i-1)2+1}^{(2)}\Big) = d_{M_1}^{(2)} \quad \text{for } i = 1 \;\Leftrightarrow\; d\Big(S_2^{(2)}, S_3^{(2)}\Big) = d_{M_1}^{(2)} \tag{2.124}$$

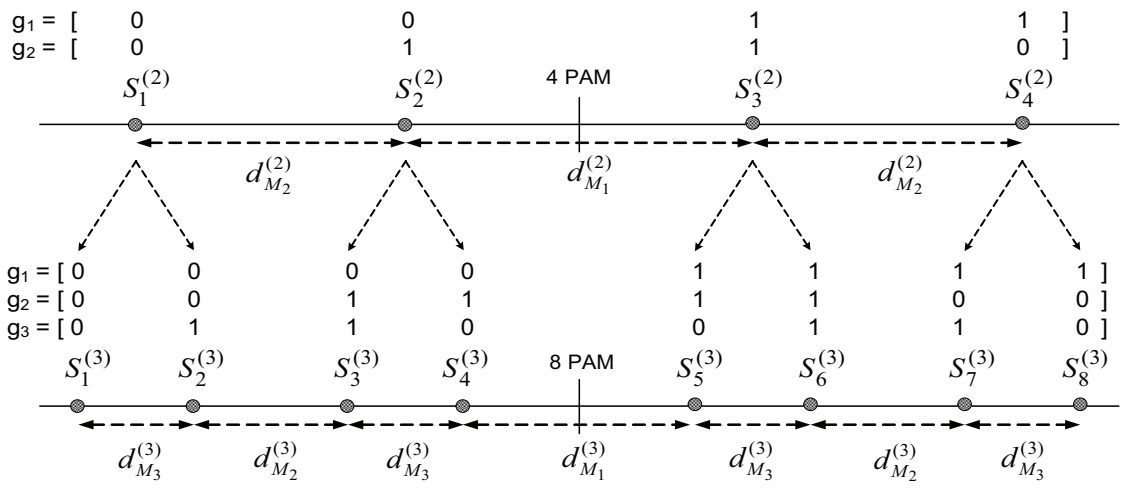


Figure 2.9: Hierarchical 4 and 8 PAM constellations.

where $A \Leftrightarrow B$ denotes A and B are identical. From (2.123), for $n = 2$, we have

$$d\Big(S_{2i-1}^{(2)}, S_{2i}^{(2)}\Big) = d_{M_2}^{(2)} \text{ for } i = 1, 2 \;\Leftrightarrow\; d\Big(S_1^{(2)}, S_2^{(2)}\Big) = d\Big(S_3^{(2)}, S_4^{(2)}\Big) = d_{M_2}^{(2)}. \quad (2.125)$$

It is seen that (2.124) and (2.125) are identical to (2.122). Suppose that (2.121) holds for 2^l PAM. That is,

$$d\Big(S_{(2i-1)2^{l-n}}^{(l)}, S_{(2i-1)2^{l-n}+1}^{(l)}\Big) = d_{M_n}^{(l)} \quad \text{for } 1 \le i \le 2^{n-1} \text{ and } 1 \le n \le l. \quad (2.126)$$

Consider hierarchical 2^{l+1} PAM. Eq. (2.120) can be rewritten as

$$d\Big(S_{2i}^{(l+1)}, S_{2i+1}^{(l+1)}\Big) = d_{M_n}^{(l+1)} \quad \text{if} \quad d\Big(S_i^{(l)}, S_{i+1}^{(l)}\Big) = d_{M_n}^{(l)}, \quad (2.127)$$

for $1 \le i \le 2^l - 1$ and $1 \le n \le l$. From (2.126) and (2.127), it can be shown that

$$d\Big(S_{(2i-1)2^{l+1-n}}^{(l+1)}, S_{(2i-1)2^{l+1-n}+1}^{(l+1)}\Big) = d_{M_n}^{(l+1)} \quad \text{for } 1 \le i \le 2^{n-1} \text{ and } 1 \le n \le l. \quad (2.128)$$

Eq. (2.119) can be rewritten as

$$d\Big(S_{2i-1}^{(l+1)}, S_{2i}^{(l+1)}\Big) = d_{M_{l+1}}^{(l+1)} \quad \text{for } 1 \le i \le 2^l. \quad (2.129)$$

From (2.129), (2.128) can be extended to the case $n = l + 1$. That is,

$$d\Big(S_{(2i-1)2^{l+1-n}}^{(l+1)}, S_{(2i-1)2^{l+1-n}+1}^{(l+1)}\Big) = d_{M_n}^{(l+1)} \quad \text{for } 1 \le i \le 2^{n-1} \text{ and } 1 \le n \le l+1. \quad (2.130)$$

If we let $K = l + 1$ in (2.121), it is identical to (2.130). Hence, (2.121) holds for hierarchical 2^{l+1} PAM.

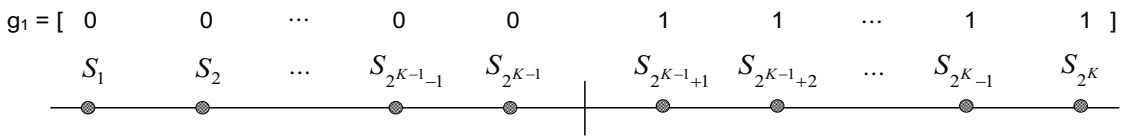


Figure 2.10: Hierarchical 2^K PAM constellation with the bit mapping vector $\mathbf{g}_1$ for the MSB.

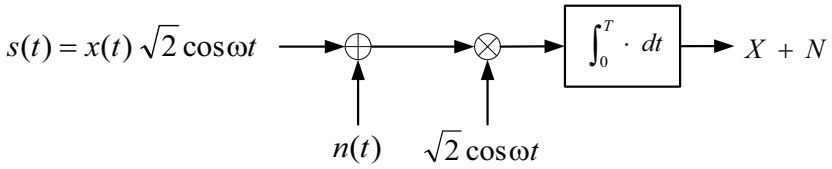


Figure 2.11: System model for hierarchical PAM.

For convenience, from here onwards, we use S_i and d_{M_n} instead of $S_i^{(K)}$ and $d_{M_n}^{(K)}$ for hierarchical 2^K PAM. For integers j, n in the range of $1 \le j \le 2^K - 1$ and $1 \le n \le K$, we define a function $f_n(j)$ as

$$f_n(j) = \begin{cases} 1, & \text{for } \; j = (2 \cdot 1 - 1)2^{K-n}, (2 \cdot 2 - 1)2^{K-n}, \cdots, (2 \cdot 2^{n-1} - 1)2^{K-n} \\ 0, & \text{otherwise.} \end{cases} \tag{2.131}$$

From (2.131), it can be shown that (2.121) is expressed as

$$d\Big(S_j, S_{j+1}\Big) = \sum_{n=1}^{K} f_n(j) d_{M_n} \quad \text{for } 1 \le j \le 2^K - 1. \tag{2.132}$$

2.9.3 BER of the MSB for hierarchical 2^K PAM

Fig. 2.10 depicts a hierarchical 2^K PAM constellation with the bit mapping vector $\mathbf{g}_1$ for the MSB given by (2.118). The system model for hierarchical 2^K PAM is shown in Fig. 2.11. The transmitted signal is given by

$$\begin{aligned} s(t) &= x(t)\sqrt{2}\cos\omega t \\ &= \text{sgn}(i - 2^{K-1} - 0.5) d\Big(0, S_i\Big) P_T(t)\sqrt{2}\cos\omega t \quad \text{for } 1 \le i \le 2^K \end{aligned} \tag{2.133}$$

where $\text{sgn}(\cdot)$ denotes the sign of the real number, $d\Big(0, S_i\Big)$ is the Euclidian distance between the origin and ith signal point S_i $(1 \le i \le 2^K)$, and $P_T(t)$ is the transmit pulse

defined as

$$P_T(t) = \begin{cases} 1, & 0 \le t \le T \\ 0, & \text{elsewhere} \end{cases} \tag{2.134}$$

where T is the symbol duration. $n(t)$ is zero-mean additive white Gaussian noise having a power spectral density of $N_0/2$. At the receiver, the decision statistic is given by

$$X = \text{sgn}(i - 2^{K-1} - 0.5) d\Big(0, S_i\Big) T \;\text{ and }\; N = \int_0^T n(t)\sqrt{2}\cos\omega t \; dt \tag{2.135}$$

where the standard deviation of N is $\sqrt{N_0 T/2}$. From Fig. 2.10, since the decision boundary for bits 0 and 1 is the origin, the probability of correct decision for a signal point assigned for bit 1, S_i ($i > 2^{K-1} + 1$), is given by

$$\begin{aligned} P_{c,S_i} &= \Pr\left[0 \le d\Big(0, S_i\Big) T + N < \infty\right] \\ &= 1 - Q\left(\frac{d\Big(0, S_i\Big) T}{\sqrt{N_0 T/2}}\right) = 1 - Q\left(d\Big(0, S_i\Big)\sqrt{\frac{2T}{N_0}}\right). \end{aligned} \tag{2.136}$$

From (2.136), the probability of correct decision for the MSB is given by

$$P_c = \frac{1}{2^{K-1}} \sum_{i=2^{K-1}+1}^{2^K} P_{c,S_i} = 1 - \frac{1}{2^{K-1}} \sum_{i=2^{K-1}+1}^{2^K} Q\left(d\Big(0, S_i\Big)\sqrt{\frac{2T}{N_0}}\right) \tag{2.137}$$

and the BER for the MSB, P_{M_1}, is given by

$$P_{M_1} = 1 - P_c = \frac{1}{2^{K-1}} \sum_{i=2^{K-1}+1}^{2^K} Q\left(d\Big(0, S_i\Big)\sqrt{\frac{2T}{N_0}}\right). \tag{2.138}$$

From (2.121), for $n = 1$, we have

$$d\Big(S_{(2i-1)2^{K-1}}, S_{(2i-1)2^{K-1}+1}\Big) = d_{M_1} \;\text{ for } i = 1 \;\Leftrightarrow\; d\Big(S_{2^{K-1}}, S_{2^{K-1}+1}\Big) = d_{M_1}. \tag{2.139}$$

Since the hierarchical PAM constellation is symmetric with respect to the origin, from (2.139), we have

$$d\Big(0, S_{2^{K-1}+1}\Big) = \frac{1}{2} d\Big(S_{2^{K-1}}, S_{2^{K-1}+1}\Big) = \frac{d_{M_1}}{2}. \tag{2.140}$$

For $i \geq 2^{K-1}+2$, $d\Big(0, S_i\Big)$ can be expressed as

$$d\Big(0, S_i\Big) = d\Big(0, S_{2^{K-1}+1}\Big) + \sum_{j=2^{K-1}+1}^{i-1} d\Big(S_j, S_{j+1}\Big) = \frac{d_{M_1}}{2} + \sum_{j=2^{K-1}+1}^{i-1} d\Big(S_j, S_{j+1}\Big) \tag{2.141}$$

where the second equality follows from (2.140). From (2.140) and (2.141), the BER of the MSB, given by (2.138), can be rewritten as

$$\begin{aligned} P_{M_1} = & \frac{1}{2^{K-1}} Q\left(\frac{d_{M_1}}{2}\sqrt{\frac{2T}{N_0}}\right) \\ & + \frac{1}{2^{K-1}} \sum_{i=2^{K-1}+2}^{2^K} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{j=2^{K-1}+1}^{i-1} d\Big(S_j, S_{j+1}\Big)\right)\sqrt{\frac{2T}{N_0}}\right). \end{aligned} \tag{2.142}$$

From (2.132), $\sum_{j=2^{K-1}+1}^{i-1} d\Big(S_j, S_{j+1}\Big)$ in (2.142) can be rewritten as

$$\sum_{j=2^{K-1}+1}^{i-1} d\Big(S_j, S_{j+1}\Big) = \sum_{j=2^{K-1}+1}^{i-1} \sum_{n=1}^{K} f_n(j) d_{M_n} = \sum_{n=1}^{K} d_{M_n} \sum_{j=2^{K-1}+1}^{i-1} f_n(j). \tag{2.143}$$

From (2.131), it can be shown that $\sum_{j=1}^{l} f_n(j)$ is expressed as

$$\sum_{j=1}^{l} f_n(j) = \left\lfloor \frac{l + 2^{K-n}}{2^{K-n+1}} \right\rfloor \quad \text{for } 1 \leq l \leq 2^K - 1 \text{ and } 1 \leq n \leq K. \tag{2.144}$$

From (2.144), (2.143) can be rewritten as

$$\sum_{j=2^{K-1}+1}^{i-1} d\Big(S_j, S_{j+1}\Big) = \sum_{n=1}^{K} d_{M_n} \left(\left\lfloor \frac{i-1+2^{K-n}}{2^{K-n+1}} \right\rfloor - \left\lfloor \frac{2^{K-1}+2^{K-n}}{2^{K-n+1}} \right\rfloor \right). \tag{2.145}$$

From (2.145), the second term of P_{M_1} given by (2.142) can be expressed as

$$\frac{1}{2^{K-1}} \sum_{i=2^{K-1}+2}^{2^K} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{n=1}^{K} d_{M_n} \left\{ \left\lfloor \frac{i-1+2^{K-n}}{2^{K-n+1}} \right\rfloor - \left\lfloor \frac{2^{K-1}+2^{K-n}}{2^{K-n+1}} \right\rfloor \right\}\right)\sqrt{\frac{2T}{N_0}}\right). \tag{2.146}$$

Let $p = i - 2^{K-1} - 1$. Then (2.146) can be rewritten as

$$\frac{1}{2^{K-1}} \sum_{p=1}^{2^{K-1}-1} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{n=1}^{K} d_{M_n} \left\{ \left\lfloor \frac{p+2^{K-n}}{2^{K-n+1}} + 2^{n-2} \right\rfloor - \lfloor 2^{n-2} + 2^{-1} \rfloor \right\}\right)\sqrt{\frac{2T}{N_0}}\right). \tag{2.147}$$

For $n \geq 2$, we have

$$\left\lfloor \frac{p+2^{K-n}}{2^{K-n+1}} + 2^{n-2} \right\rfloor = \left\lfloor \frac{p+2^{K-n}}{2^{K-n+1}} \right\rfloor + 2^{n-2} \quad \text{and} \quad \left\lfloor 2^{n-2} + 2^{-1} \right\rfloor = 2^{n-2}. \tag{2.148}$$

For $n = 1$, we have

$$\left\lfloor \frac{p+2^{K-n}}{2^{K-n+1}} + 2^{n-2} \right\rfloor = \left\lfloor \frac{p}{2^K} + 1 \right\rfloor = 1 \quad \text{and} \quad \left\lfloor 2^{n-2} + 2^{-1} \right\rfloor = 1 \tag{2.149}$$

where the second equality of the first expression follows from $1 \leq p \leq 2^{K-1}-1$ in (2.147). From (2.148) and (2.149), the second term of P_{M_1}, given by (2.147), can be rewritten as

$$\frac{1}{2^{K-1}} \sum_{p=1}^{2^{K-1}-1} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{n=2}^{K} d_{M_n} \left\lfloor \frac{p+2^{K-n}}{2^{K-n+1}} \right\rfloor \right) \sqrt{\frac{2T}{N_0}} \right). \tag{2.150}$$

Since $\sum_{n=2}^{K} d_{M_n} \left\lfloor \frac{p+2^{K-n}}{2^{K-n+1}} \right\rfloor = 0$ for $p = 0$, from (2.150), the BER of the MSB given by (2.142) can be expressed as

$$P_{M_1} = \frac{1}{2^{K-1}} \sum_{p=0}^{2^{K-1}-1} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{n=2}^{K} d_{M_n} \left\lfloor \frac{p+2^{K-n}}{2^{K-n+1}} \right\rfloor \right) \sqrt{\frac{2T}{N_0}} \right). \tag{2.151}$$

Note that (2.151) is the exact BER expression for the MSB of hierarchical 2^K PAM.

2.9.4 BER of the n_0th MSB $(2 \leq n_0 \leq K-1)$ for hierarchical 2^K PAM

A. Classification of 2^K signal points into 2^{n_0-1} mutually exclusive groups

We first find every pair of adjacent signal points which are separated by a Euclidian distance greater than $d_{M_{n_0}}$ (i.e., $d_{M_{n_0-1}}, d_{M_{n_0}-2}, \cdots, d_{M_1}$): For given n_0 in the range of $2 \leq n_0 \leq K-1$, let $n = n_0 - m$ $(1 \leq m \leq n_0 - 1)$ in (2.121). Then, we have

$$\begin{aligned} d\Big(S_{(2i-1)2^{K-n_0+m}}, S_{(2i-1)2^{K-n_0+m}+1}\Big) &= d_{M_{n_0-m}} \\ &\text{for } 1 \leq i \leq 2^{n_0-m-1} \text{ and } 1 \leq m \leq n_0 - 1. \end{aligned} \tag{2.152}$$

It can be shown that $\left\{(2i-1)2^{m-1} \mid 1 \leq i \leq 2^{n_0-m-1} \text{ and } 1 \leq m \leq n_0-1\right\}$ is identical to $\left\{j \mid 1 \leq j \leq 2^{n_0-1}-1\right\}$. Hence, every pair of adjacent signal points which are separated by a Euclidian distance greater than $d_{M_{n_0}}$, given by (2.152), can be expressed as

$$S_{j\cdot 2^{K+1-n_0}},\ S_{j\cdot 2^{K+1-n_0}+1} \qquad \text{for } 1 \leq j \leq 2^{n_0-1}-1. \tag{2.153}$$

Next, we classify 2^K signal points into 2^{n_0-1} mutually exclusive groups such that the Euclidian distance between adjacent signal points of the same group is smaller than or equal to $d_{M_{n_0}}$. From (2.153), the signal points of the jth group can be derived as

$$S_{(j-1)2^{K+1-n_0}+1}, S_{(j-1)2^{K+1-n_0}+2}, \cdots, S_{j\cdot 2^{K+1-n_0}} \quad \text{for } 1 \le j \le 2^{n_0-1}. \tag{2.154}$$

We rewrite (2.121) in the following: For hierarchical 2^K PAM ($K \ge 2$), the Euclidian distance between adjacent signal points is given by

$$d\Big(S_{(2i-1)2^{K-n}}, S_{(2i-1)2^{K-n}+1}\Big) = d_{M_n} \quad \text{for } 1 \le i \le 2^{n-1} \text{ and } 1 \le n \le K. \tag{2.155}$$

From (2.154) and (2.155), it can be shown that the Euclidian distance between adjacent signal points of the jth group is given by

$$d\Big(S_{(2i-1)2^{K-n}}, S_{(2i-1)2^{K-n}+1}\Big) = d_{M_n}$$
$$\text{for } (j-1)2^{n-n_0}+1 \le i \le j\cdot 2^{n-n_0},\ n_0 \le n \le K, \text{ and } 1 \le j \le 2^{n_0-1}. \tag{2.156}$$

Let $p = i-(j-1)2^{n-n_0}$. Then, (2.156) can be rewritten as

$$d\Big(S_{(2p-1)2^{K-n}+(j-1)2^{K+1-n_0}}, S_{(2p-1)2^{K-n}+(j-1)2^{K+1-n_0}+1}\Big) = d_{M_n}$$
$$\text{for } 1 \le p \le 2^{n-n_0},\ n_0 \le n \le K, \text{ and } 1 \le j \le 2^{n_0-1}. \tag{2.157}$$

Notation change: Let $S_i^{(j)}$ denote $S_{(j-1)2^{K+1-n_0}+i}$ for convenience. Then, every pair of adjacent signal points which are separated by a Euclidian distance greater than $d_{M_{n_0}}$ (i.e., $d_{M_{n_0-1}}$, $d_{M_{n_0}-2}$, $\cdots$, d_{M_1}), given by (2.153), can be rewritten as

$$S_{2^{K+1-n_0}}^{(j)},\ S_1^{(j+1)} \quad \text{for } 1 \le j \le 2^{n_0-1}-1. \tag{2.158}$$

The signal points of the jth group, given by (2.154), can be expressed as

$$S_1^{(j)}, S_2^{(j)}, \cdots, S_{2^{K+1-n_0}}^{(j)} \quad \text{for } 1 \le j \le 2^{n_0-1}. \tag{2.159}$$

Lastly, the Euclidian distance between adjacent signal points of the jth group, given by (2.157), can be rewritten as

$$d\Big(S_{(2p-1)2^{K-n}}^{(j)}, S_{(2p-1)2^{K-n}+1}^{(j)}\Big) = d_{M_n}$$
$$\text{for } 1 \le p \le 2^{n-n_0},\ n_0 \le n \le K, \text{ and } 1 \le j \le 2^{n_0-1}. \tag{2.160}$$

B. Probability of correct decision for signal points of the jth group

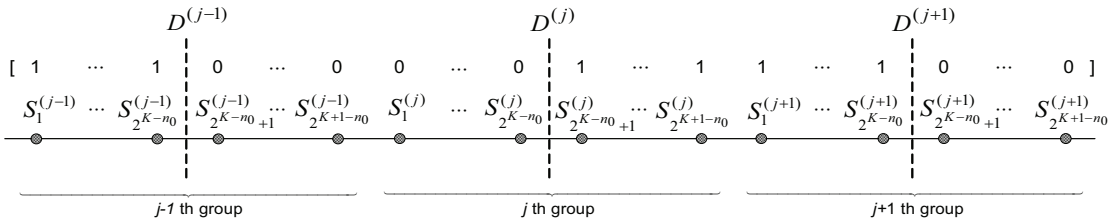


Figure 2.12: The $j-1$, j and $j+1$th groups with the bit mapping vector for $j =$ odd.

From (2.159), for $2 \le j \le 2^{n_0-1}-1$, the signal points of the $j-1$, j, and $j+1$th groups are given by

$$\underbrace{S_1^{(j-1)}, S_2^{(j-1)}, \cdots, S_{2^{K+1-n_0}}^{(j-1)}}_{j-1\text{th group}}, \underbrace{S_1^{(j)}, S_2^{(j)}, \cdots, S_{2^{K+1-n_0}}^{(j)}}_{j\text{th group}}, \underbrace{S_1^{(j+1)}, S_2^{(j+1)}, \cdots, S_{2^{K+1-n_0}}^{(j+1)}}_{j+1\text{th group}}. \tag{2.161}$$

From (2.118), the bit mapping vector for the n_0th MSB ($2 \le n_0 \le K-1$) of the $j-1$, j and $j+1$th groups is derived as

$$\begin{cases} \left[\mathbf{0}_{2^{K-n_0}}\ \mathbf{1}_{2^{K-n_0}}\ \mathbf{1}_{2^{K-n_0}}\ \mathbf{0}_{2^{K-n_0}}\ \mathbf{0}_{2^{K-n_0}}\ \mathbf{1}_{2^{K-n_0}}\right], & \text{for } j = \text{even} \\ \left[\mathbf{1}_{2^{K-n_0}}\ \mathbf{0}_{2^{K-n_0}}\ \mathbf{0}_{2^{K-n_0}}\ \mathbf{1}_{2^{K-n_0}}\ \mathbf{1}_{2^{K-n_0}}\ \mathbf{0}_{2^{K-n_0}}\right], & \text{for } j = \text{odd}. \end{cases} \tag{2.162}$$

From (2.161) and (2.162), $j-1$, j, and $j+1$th groups with the bit mapping vector for $j =$ odd are shown in Fig. 2.12, where $D^{(j-1)}$, $D^{(j)}$, and $D^{(j+1)}$ denote the decision boundaries for bits 0 and 1 in the $j-1$, j, and $j+1$th groups, respectively. In the following, we will derive the probability of correct decision for signal points of the jth group ($1 \le j \le 2^{n_0-1}$):

i) Signal points assigned for bit 0 when j is odd in the range of $2 \le j \le 2^{n_0-1}-1$

We here assume that for $S_i^{(j)}$ ($1 \le i \le 2^{K-n_0}$), a signal point of the jth group which is assigned for bit 0, the probability of correct decision can be calculated without considering the other groups except for the $j-1$, j, and $j+1$ th groups (we will later show that the assumption is correct if the SNR condition of this theorem is satisfied). Fig. 2.13 shows the correct decision area for $S_i^{(j)}$ ($1 \le i \le 2^{K-n_0}$) under the above assumption. From Fig. 2.13, it follows that the probability of correct decision for $S_i^{(j)}$

$(1 \le i \le 2^{K-n_0})$ based on the system model depicted in Fig. 2.11 is given by

$$\begin{aligned} P_c^{\text{bit0}} &= \Pr\Big[- d\Big(D^{(j-1)}, S_i^{(j)}\Big)T < N < d\Big(S_i^{(j)}, D^{(j)}\Big)T\Big] \\ &+ \Pr\Big[d\Big(S_i^{(j)}, D^{(j+1)}\Big)T < N < d\Big(S_i^{(j)}, S_{2^{K+1-n_0}}^{(j+1)}\Big)T\Big] \\ &= \Pr\Big[- \Big(d\Big(D^{(j-1)}, S_1^{(j)}\Big) + d\Big(S_1^{(j)}, S_i^{(j)}\Big)\Big)\, T < \\ &\qquad N < \Big(d\Big(S_i^{(j)}, S_{2^{K-n_0}}^{(j)}\Big) + d\Big(S_{2^{K-n_0}}^{(j)}, D^{(j)}\Big)\Big)\, T\Big] \\ &+ \Pr\Big[\Big(d\Big(S_i^{(j)}, S_{2^{K-n_0}}^{(j)}\Big) + d\Big(S_{2^{K-n_0}}^{(j)}, D^{(j+1)}\Big)\Big)\, T < \\ &\qquad N < \Big(S_i^{(j)}, S_{2^{K+1-n_0}}^{(j+1)}\Big)T\Big] \end{aligned} \tag{2.163}$$

where the first and second terms follow from the correct decision areas #1 and #2 shown in Fig. 2.13, respectively. Eq. (2.163) can be rewritten as

$$\begin{aligned} P_c^{\text{bit0}} &= 1 - \Pr\Big[N > \Big(d\Big(S_i^{(j)}, S_{2^{K-n_0}}^{(j)}\Big) + d\Big(S_{2^{K-n_0}}^{(j)}, D^{(j)}\Big)\Big)\, T\Big] \\ &\quad -\Pr\Big[N > \Big(d\Big(D^{(j-1)}, S_1^{(j)}\Big) + d\Big(S_1^{(j)}, S_i^{(j)}\Big)\Big)\, T\Big] \\ &\quad +\Pr\Big[\Big(d\Big(S_i^{(j)}, S_{2^{K-n_0}}^{(j)}\Big) + d\Big(S_{2^{K-n_0}}^{(j)}, D^{(j+1)}\Big)\Big)\, T < N < \Big(S_i^{(j)}, S_{2^{K+1-n_0}}^{(j+1)}\Big)T\Big]. \end{aligned} \tag{2.164}$$

From Fig. 2.13, $d\Big(D^{(j-1)}, S_1^{(j)}\Big)$ in the second term of (2.164) can be expressed as

$$\begin{aligned} d\Big(D^{(j-1)}, S_1^{(j)}\Big) &= d\Big(D^{(j-1)}, S_{2^{K-n_0}+1}^{(j-1)}\Big) + d\Big(S_{2^{K-n_0}+1}^{(j-1)}, S_{2^{K+1-n_0}}^{(j-1)}\Big) \\ &\quad + d\Big(S_{2^{K+1-n_0}}^{(j-1)}, S_1^{(j)}\Big). \end{aligned} \tag{2.165}$$

From (2.160), for $n = n_0$, we have

$$d\Big(S_{2^{K-n_0}}^{(j)}, S_{2^{K-n_0}+1}^{(j)}\Big) = d_{M_{n_0}} \quad \text{for } 1 \le j \le 2^{n_0-1}. \tag{2.166}$$

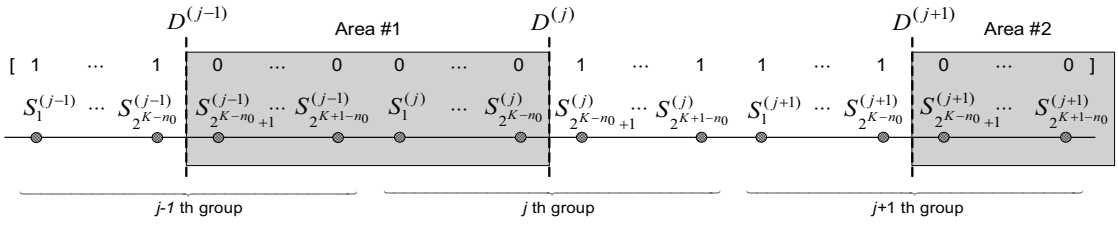


Figure 2.13: The correct decision area for $S_i^{(j)}$ $(1 \le i \le 2^{K-n_0})$ when j = odd.

From the fact that $d\Big(D^{(j-1)}, S^{(j-1)}_{2^{K-n_0}+1}\Big) = \frac{1}{2}d\Big(S^{(j-1)}_{2^{K-n_0}}, S^{(j-1)}_{2^{K-n_0}+1}\Big)$ and (2.166), (2.165) can be rewritten as

$$d\Big(D^{(j-1)}, S^{(j)}_1\Big) = \frac{1}{2}d_{M_{n_0}} + d\Big(S^{(j-1)}_{2^{K-n_0}+1}, S^{(j-1)}_{2^{K+1-n_0}}\Big) + d\Big(S^{(j-1)}_{2^{K+1-n_0}}, S^{(j)}_1\Big). \quad (2.167)$$

From (2.158), we have

$$d\Big(S^{(j)}_{2^{K+1-n_0}}, S^{(j+1)}_1\Big) \geq d_{M_{n_0-1}} \quad \text{for } 1 \leq j \leq 2^{n_0-1} - 1. \quad (2.168)$$

From (2.168), $d\Big(D^{(j-1)}, S^{(j)}_1\Big)$, given by (2.167), satisfies

$$d\Big(D^{(j-1)}, S^{(j)}_1\Big) \geq \frac{1}{2}d_{M_{n_0}} + d_{M_{n_0-1}} + d\Big(S^{(j-1)}_{2^{K-n_0}+1}, S^{(j-1)}_{2^{K+1-n_0}}\Big). \quad (2.169)$$

Since $d\Big(S^{(j-1)}_{2^{K-n_0}+1}, S^{(j-1)}_{2^{K+1-n_0}}\Big) > 0$ for $n_0 \leq K-1$, we have

$$d\Big(D^{(j-1)}, S^{(j)}_1\Big) > \frac{1}{2}d_{M_{n_0}} + d_{M_{n_0-1}}. \quad (2.170)$$

Likewise, from Fig. 2.13, $d\Big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\Big)$ in the third term of (2.164) can be expressed as

$$\begin{aligned} d\Big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\Big) = {} & d\Big(S^{(j)}_{2^{K-n_0}}, S^{(j)}_{2^{K-n_0}+1}\Big) + d\Big(S^{(j)}_{2^{K-n_0}+1}, S^{(j)}_{2^{K+1-n_0}}\Big) \\ & + d\Big(S^{(j)}_{2^{K+1-n_0}}, S^{(j+1)}_1\Big) + d\Big(S^{(j+1)}_1, S^{(j+1)}_{2^{K-n_0}}\Big) + d\Big(S^{(j+1)}_{2^{K-n_0}}, D^{(j+1)}\Big). \end{aligned} \quad (2.171)$$

Since $d\Big(S^{(j+1)}_{2^{K-n_0}}, D^{(j+1)}\Big) = \frac{1}{2}d\Big(S^{(j+1)}_{2^{K-n_0}}, S^{(j+1)}_{2^{K-n_0}+1}\Big)$ and from (2.166), (2.171) can be rewritten as

$$\begin{aligned} d\Big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\Big) = {} & \frac{3}{2}d_{M_{n_0}} + d\Big(S^{(j)}_{2^{K-n_0}+1}, S^{(j)}_{2^{K+1-n_0}}\Big) + d\Big(S^{(j)}_{2^{K+1-n_0}}, S^{(j+1)}_1\Big) \\ & + d\Big(S^{(j+1)}_1, S^{(j+1)}_{2^{K-n_0}}\Big). \end{aligned} \quad (2.172)$$

From (2.168), $d\Big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\Big)$ satisfies

$$\begin{aligned} d\Big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\Big) = {} & \frac{3}{2}d_{M_{n_0}} + d_{M_{n_0-1}} + d\Big(S^{(j)}_{2^{K-n_0}+1}, S^{(j)}_{2^{K+1-n_0}}\Big) \\ & + d\Big(S^{(j+1)}_1, S^{(j+1)}_{2^{K-n_0}}\Big). \end{aligned} \quad (2.173)$$

We have $d\Big(S^{(j)}_{2^{K-n_0}+1}, S^{(j)}_{2^{K+1-n_0}}\Big) > 0$ and $d\Big(S^{(j+1)}_1, S^{(j+1)}_{2^{K-n_0}}\Big) > 0$ for $n_0 \leq K-1$. Hence, $d\Big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\Big)$ satisfies

$$d\Big(S^{(j)}_{2^{K-n_0}}, D^{(j+1)}\Big) > \frac{3}{2}d_{M_{n_0}} + d_{M_{n_0-1}}. \quad (2.174)$$

From (2.170) and (2.174), it follows that the second and third terms of P_c^{bit0}, given by (2.164), are insignificant when the condition of this theorem is satisfied. Since $S_i^{(j)}$, $S_{2^{K-n_0}}^{(j)}$, and $D^{(j)}$ belong to the jth group, $d\Big(S_i^{(j)}, S_{2^{K-n_0}}^{(j)}\Big) + d\Big(S_{2^{K-n_0}}^{(j)}, D^{(j)}\Big)$ in the first term of P_c^{bit0} is the combination of $d_{M_{n_0}}, d_{M_{n_0+1}}, \cdots, d_{M_K}$ from (2.160), and thus the first term is not affected by the condition of this theorem. Hence, if the condition of this theorem is satisfied, P_c^{bit0}, given by (2.164), becomes

$$P_c^{\text{bit0}} \approx 1 - \Pr\left[N > \left(d\Big(S_i^{(j)}, S_{2^{K-n_0}}^{(j)}\Big) + d\Big(S_{2^{K-n_0}}^{(j)}, D^{(j)}\Big)\right) T\right], \tag{2.175}$$

which is identical to the probability of correct decision calculated only by considering 2^{K+1-n_0} signal points of the isolated jth group. Since the $j-1$ and $j+1$th groups have no effect on the correct decision probability for signal points of the jth group due to the condition of this theorem, the other groups (i.e., $1, \cdots, j-2, j+2, \cdots, 2^{n_0-1}$th groups), which are separated by larger Euclidian distances from the jth group than are the $j-1$ and $j+1$th groups, also have no effect. Hence, the assumption above (2.163) is correct.
ii) Signal points assigned for bit 1 when j is odd in the range of $2 \le j \le 2^{n_0-1} - 1$

It can be shown that the probability of correct decision for $S_i^{(j)}$ $(2^{K-n_0}+1 \le i \le 2^{K+1-n_0})$ based on the system model depicted in Fig. 2.11 is given by

$$\begin{aligned} P_c^{\text{bit1}} = 1 &- \Pr\left[N > \left(d\Big(D^{(j)}, S_{2^{K-n_0+1}}^{(j)}\Big) + d\Big(S_{2^{K-n_0+1}}^{(j)}, S_i^{(j)}\Big)\right) T\right] \\ &- \Pr\left[N > \left(d\Big(S_i^{(j)}, S_{2^{K+1-n_0}}^{(j)}\Big) + d\Big(S_{2^{K+1-n_0}}^{(j)}, D^{(j+1)}\Big)\right) T\right] \\ &+ \Pr\left[\left(d\Big(D^{(j-1)}, S_{2^{K-n_0+1}}^{(j)}\Big) + d\Big(S_{2^{K-n_0+1}}^{(j)}, S_i^{(j)}\Big)\right) T < N \right. \\ &\qquad \left. < d\Big(S_1^{(j-1)}, S_i^{(j)}\Big) T\right], \end{aligned} \tag{2.176}$$

where $d\Big(S_{2^{K+1-n_0}}^{(j)}, D^{(j+1)}\Big)$in the second term of (2.176) satisfies [1]

$$d\Big(S_{2^{K+1-n_0}}^{(j)}, D^{(j+1)}\Big) > \frac{1}{2} d_{M_{n_0}} + d_{M_{n_0-1}}, \tag{2.177}$$

and $d\Big(D^{(j-1)}, S_{2^{K-n_0+1}}^{(j)}\Big)$ in the third term of (2.176) satisfies

$$d\Big(D^{(j-1)}, S_{2^{K-n_0+1}}^{(j)}\Big) > \frac{3}{2} d_{M_{n_0}} + d_{M_{n_0-1}}. \tag{2.178}$$

[1] Since the analysis of ii) is similar to that of i), we omit the detailed steps.

From (2.176)–(2.178), if the condition of this theorem is satisfied, P_c^{bit1}, given by (2.176), becomes

$$P_c^{\text{bit1}} \approx 1 - \Pr\left[N > \left(d\left(D^{(j)}, S^{(j)}_{2^{K-n_0}+1}\right) + d\left(S^{(j)}_{2^{K-n_0}+1}, S_i^{(j)}\right)\right) T\right], \tag{2.179}$$

which is identical to the probability of correct decision calculated only by considering 2^{K+1-n_0} signal points of the isolated jth group.

iii) Signal points assigned for bit 0/1 when j is even in the range of $2 \le j \le 2^{n_0-1} - 1$

From (2.162), the bit mapping vector for j = even is just the complement of that for j = odd. Hence, for j = even, P_c^{bit0} and P_c^{bit1} are given by (2.176) and (2.164), respectively, and the results of i) and ii) hold for the case j is even.

iv) Signal points assigned for bit 0/1 when $j = 1$ (odd)

From Fig. 2.13, it follows that P_c^{bit0} for $j = 1$ is given by

$$\begin{aligned} P_c^{\text{bit0}} = 1 &- \Pr\left[N > \left(d\left(S_i^{(1)}, S^{(1)}_{2^{K-n_0}}\right) + d\left(S^{(1)}_{2^{K-n_0}}, D^{(1)}\right)\right) T\right] \\ &+ \Pr\left[\left(d\left(S_i^{(1)}, S^{(1)}_{2^{K-n_0}}\right) + d\left(S^{(1)}_{2^{K-n_0}}, D^{(2)}\right)\right) T < N < d\left(S_i^{(1)}, S^{(2)}_{2^{K+1-n_0}}\right) T\right]. \end{aligned} \tag{2.180}$$

The only difference between (2.164) and (2.180) is that (2.180) does not have the second term of (2.164), and thus the result of i) holds for the case $j = 1$. In a similar way, it can be shown that for bit 1, the result of ii) holds for the case $j = 1$.

v) Signal points assigned for bit 0/1 when $j = 2^{n_0-1}$ (even)

In a similar way to iv), it can be shown that the result of iii) holds for the case $j = 2^{n_0-1}$.

From i)–v), it is seen that if the SNR condition of this theorem is satisfied, the BER of the n_0th MSB can be calculated only by considering 2^{K+1-n_0} signal points of the isolated jth group given by (2.159).

C. BER of the n_0th MSB $(2 \le n_0 \le K-1)$ ***for the isolated jth group***

We derive the BER of the n_0th MSB for the isolated jth group of 2^K PAM from that of the MSB for 2^{K+1-n_0} PAM.

i) For hierarchical 2^{K+1-n_0} PAM, from (2.155), the Euclidian distance between adjacent signal points is given by

$$\begin{aligned} d\left(S_{(2i-1)2^{K+1-n_0-n}}, S_{(2i-1)2^{K+1-n_0-n}+1}\right) &= d_{M_n} \\ \text{for } 1 \le i \le 2^{n-1} &\text{ and } 1 \le n \le K+1-n_0. \end{aligned} \tag{2.181}$$

Let $r = n + n_0 - 1$ and $p = i$. Then, (2.181) can be rewritten as

$$d\Big(S_{(2p-1)2^{K-r}}, S_{(2p-1)2^{K-r}+1}\Big) = d_{M_{r+1-n_0}} \quad \text{for } 1 \le p \le 2^{r-n_0} \text{ and } n_0 \le r \le K. \tag{2.182}$$

From (2.160) and (2.182), it is seen that, if $d_{M_{r+1-n_0}}$ in (2.182) is set equal to d_{M_r}, the Euclidian distance between adjacent signal points for 2^{K+1-n_0} PAM is the same as that for the jth group of 2^K PAM.

ii) For hierarchical 2^{K+1-n_0} PAM, from (2.118), the bit mapping vector for the MSB is given by

$$\mathbf{g}_1 = \Big[\mathbf{0}_{2^{(K+1-n_0)-1}}\ \mathbf{1}_{2^{(K+1-n_0)-1}}\Big] = \Big[\mathbf{0}_{2^{K-n_0}}\ \mathbf{1}_{2^{K-n_0}}\Big]. \tag{2.183}$$

For the jth group of hierarchical 2^K PAM, from (2.162), the bit mapping vector for the n_0th MSB is given by

$$\begin{cases} \Big[\mathbf{1}_{2^{K-n_0}}\ \mathbf{0}_{2^{K-n_0}}\Big], & \text{for } j = \text{even} \\ \Big[\mathbf{0}_{2^{K-n_0}}\ \mathbf{1}_{2^{K-n_0}}\Big], & \text{for } j = \text{odd}. \end{cases} \tag{2.184}$$

It is seen that (2.184) is the same as or the complement of (2.183).

From i) and ii), it follows that the BER of the MSB for 2^{K+1-n_0} PAM is the same as that of the n_0th MSB for the isolated jth group of 2^K PAM, if $d_{M_{r+1-n_0}}$ for 2^{K+1-n_0} PAM is set equal to d_{M_r} (i.e., d_{M_x} is set equal to $d_{M_{x+n_0-1}}$). From (2.151), the BER of the MSB for hierarchical 2^{K+1-n_0} PAM ($2 \le n_0 \le K-1$) is derived as

$$\frac{1}{2^{K-n_0}} \sum_{p=0}^{2^{K-n_0}-1} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{n=2}^{K+1-n_0} \left\lfloor \frac{p + 2^{K+1-n_0-n}}{2^{K+2-n_0-n}} \right\rfloor d_{M_n}\right)\sqrt{\frac{2T}{N_0}}\right). \tag{2.185}$$

Let $r = n_0 - 1 + n$. Then (2.185) can be rewritten as

$$\frac{1}{2^{K-n_0}} \sum_{p=0}^{2^{K-n_0}-1} Q\left(\left(\frac{d_{M_1}}{2} + \sum_{r=n_0+1}^{K} \left\lfloor \frac{p + 2^{K-r}}{2^{K-r+1}} \right\rfloor d_{M_{r-n_0+1}}\right)\sqrt{\frac{2T}{N_0}}\right). \tag{2.186}$$

As stated above (2.185), by setting d_{M_x} equal to $d_{M_{x+n_0-1}}$ in (2.186), the BER for the n_0th MSB ($2 \le n_0 \le K-1$) of the isolated jth group can be derived as

$$\frac{1}{2^{K-n_0}} \sum_{p=0}^{2^{K-n_0}-1} Q\left(\left(\frac{d_{M_{n_0}}}{2} + \sum_{r=n_0+1}^{K} \left\lfloor \frac{p + 2^{K-r}}{2^{K-r+1}} \right\rfloor d_{M_r}\right)\sqrt{\frac{2T}{N_0}}\right). \tag{2.187}$$

Note that the BER expression for the n_0th MSB ($2 \le n_0 \le K-1$) of hierarchial 2^K PAM, given by (2.187), holds if the condition of this theorem is satisfied.

2.9.5 BER of the Kth MSB (or LSB) for hierarchical 2^K PAM

A. Classification of 2^K signal points into 2^{K-2} mutually exclusive groups

For the Kth MSB (or LSB), we define the signal points of the jth group as

$$S_1^{(j)}, S_2^{(j)}, S_3^{(j)}, S_4^{(j)} \quad \text{for } 1 \le j \le 2^{K-2}, \tag{2.188}$$

which is identical to (2.159) with $n_0 = K-1$. If we let $n_0 = K-1$ in (2.158), every pair of adjacent signal points which are separated by Euclidian distances greater than $d_{M_{K-1}}$ (i.e., $d_{M_{K-2}}$, $d_{M_{K-3}}$, $\cdots$, d_{M_1}) is given by

$$S_4^{(j)},\ S_1^{(j+1)} \quad \text{for } 1 \le j \le 2^{K-2}-1. \tag{2.189}$$

Also let $n_0 = K-1$ in (2.160). Then, for $1 \le j \le 2^{K-2}$, the Euclidian distance between adjacent signal points of the jth group can be derived as

$$d\Big(S_2^{(j)}, S_3^{(j)}\Big) = d_{M_{K-1}} \ \text{ and } \ d\Big(S_1^{(j)}, S_2^{(j)}\Big) = d\Big(S_3^{(j)}, S_4^{(j)}\Big) = d_{M_K}. \tag{2.190}$$

B. Probability of correct decision for signal points of the jth group

From (2.188), for $2 \le j \le 2^{K-2}-1$, the signal points of the $j-1$, j, and $j+1$th groups are given by

$$\underbrace{S_1^{(j-1)}, S_2^{(j-1)}, S_3^{(j-1)}, S_4^{(j-1)}}_{j-1\text{th group}}, \underbrace{S_1^{(j)}, S_2^{(j)}, S_3^{(j)}, S_4^{(j)}}_{j\text{th group}}, \underbrace{S_1^{(j+1)}, S_2^{(j+1)}, S_3^{(j+1)}, S_4^{(j+1)}}_{j+1\text{th group}}. \tag{2.191}$$

From (2.118), the bit mapping vector for the Kth MSB of the $j-1$, j and $j+1$th groups is given by

$$\Big[0\ \ 1\ \ 1\ \ 0\ \ 0\ \ 1\ \ 1\ \ 0\ \ 0\ \ 1\ \ 1\ \ 0\Big]. \tag{2.192}$$

From (2.191) and (2.192), $j-1$, j, and $j+1$th groups with the bit mapping vector are shown in Fig. 2.14, where $D_i^{(j-1)}$, $D_i^{(j)}$, and $D_i^{(j+1)}$ $(i = 1, 2)$ denote the decision boundaries for bits 0 and 1 in the $j-1$, j, and $j+1$th groups, respectively. In the following, we will derive the probability of correct decision for signal points of the jth group $(1 \le j \le 2^{K-2})$:

i) Signal points assigned for bit 0 when $2 \le j \le 2^{K-2}-1$

We here assume that for $S_i^{(j)}$ $(i = 1, 4)$, a signal point of the jth group which is assigned for bit 0, the probability of correct decision can be calculated without

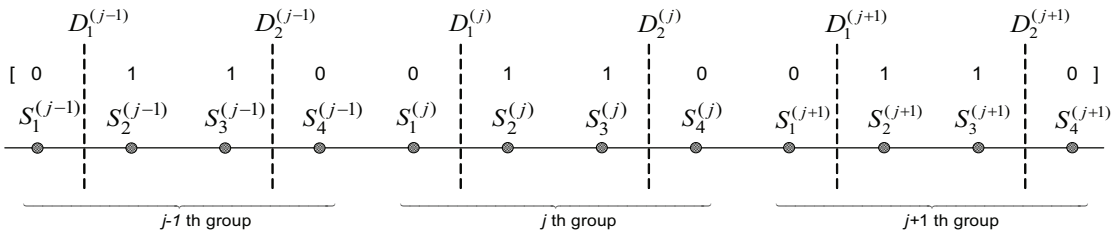


Figure 2.14: The $j-1$, j and $j+1$th groups with the bit mapping vector.

considering the other groups except for the $j-1$, j, and $j+1$ th groups (we will later show that the assumption is correct if the SNR condition of this theorem is satisfied). Fig. 2.15 shows the correct decision area for $S_i^{(j)}$ $(i=1,4)$ under the above assumption. From Fig. 2.15, it follows that the probability of correct decision for $S_1^{(j)}$, denoted by P_{c,S_1}^{bit0}, based on the system model depicted in Fig. 2.11 is given by

$$\begin{aligned} P_{c,S_1}^{\text{bit0}} = &\Pr\Big[N < -d\Big(D_1^{(j-1)}, S_1^{(j)}\Big)T\Big] + \Pr\Big[-d\Big(D_2^{(j-1)}, S_1^{(j)}\Big)T < N < d\Big(S_1^{(j)}, D_1^{(j)}\Big)T\Big] \\ &+\Pr\Big[d\Big(S_1^{(j)}, D_2^{(j)}\Big)T < N < d\Big(S_1^{(j)}, D_1^{(j+1)}\Big)T\Big] \\ &+\Pr\Big[N > d\Big(S_1^{(j)}, D_2^{(j+1)}\Big)T\Big] \end{aligned} \tag{2.193}$$

where the first, second, third and fourth terms follow from the correct decision areas #1, #2, #3, and #4 shown in Fig. 2.15, respectively. Eq. (2.193) can be rewritten as

$$\begin{aligned} P_{c,S_1}^{\text{bit0}} = &1 + \Pr\Big[N > d\Big(D_1^{(j-1)}, S_1^{(j)}\Big)T\Big] - \Pr\Big[N > d\Big(S_1^{(j)}, D_1^{(j)}\Big)T\Big] \\ &-\Pr\Big[N > d\Big(D_2^{(j-1)}, S_1^{(j)}\Big)T\Big] + \Pr\Big[d\Big(S_1^{(j)}, D_2^{(j)}\Big)T < N < d\Big(S_1^{(j)}, D_1^{(j+1)}\Big)T\Big] \\ &+\Pr\Big[N > d\Big(S_1^{(j)}, D_2^{(j+1)}\Big)T\Big]. \end{aligned} \tag{2.194}$$

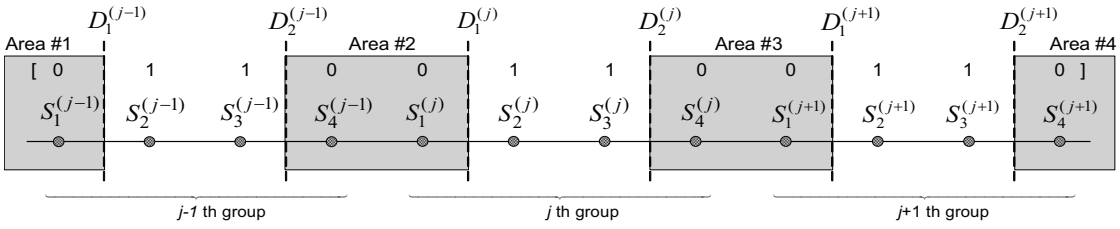


Figure 2.15: The correct decision area for $S_i^{(j)}$ $(i=1,4)$.

From (2.189), (2.190), and Fig. 2.15, it can be shown that

$$\begin{aligned} d\Big(D_1^{(j-1)}, S_1^{(j)}\Big) &\geq \frac{3}{2}d_{M_K} + d_{M_{K-1}} + d_{M_{K-2}} \\ d\Big(D_2^{(j-1)}, S_1^{(j)}\Big) &\geq \frac{1}{2}d_{M_K} + d_{M_{K-2}} \\ d\Big(S_1^{(j)}, D_2^{(j)}\Big) &= \frac{3}{2}d_{M_K} + d_{M_{K-1}} \\ d\Big(S_1^{(j)}, D_2^{(j+1)}\Big) &\geq \frac{7}{2}d_{M_K} + 2d_{M_{K-1}} + d_{M_{K-2}} \\ d\Big(S_1^{(j)}, D_1^{(j)}\Big) &= \frac{1}{2}d_{M_K}. \end{aligned} \tag{2.195}$$

Note that the first four distances in (2.195) are greater than $\frac{1}{2}d_{M_K} + d_{M_{K-1}}$, whereas the last distance is smaller than $\frac{1}{2}d_{M_K} + d_{M_{K-1}}$. Hence, the second, fourth, fifth and sixth terms of P_{c,S_1}^{bit0}, given by (2.194), are insignificant when the condition of this theorem is satisfied, whereas the third term of P_{c,S_1}^{bit0} is not affected by the condition of this theorem. Therefore, if the condition of this theorem is satisfied, P_{c,S_1}^{bit0}, given by (2.194), becomes

$$P_{c,S_1}^{\text{bit0}} \approx 1 - \Pr\left[N > d\Big(S_1^{(j)}, D_1^{(j)}\Big)T\right]. \tag{2.196}$$

In a similar manner, it can be shown that under the condition of this theorem, the probability of correct decision for $S_4^{(j)}$, denoted by P_{c,S_4}^{bit0}, becomes

$$P_{c,S_4}^{\text{bit0}} \approx 1 - \Pr\left[N > d\Big(D_2^{(j)}, S_4^{(j)}\Big)T\right]. \tag{2.197}$$

From (2.196) and (2.197), it is clear that the $j-1$ and $j+1$th groups have no effect on the correct decision probability for signal points of the jth group, due to the condition of this theorem. Hence, the other groups (i.e., $1, \cdots, j-2, j+2, \cdots, 2^{K-2}$th groups), which are separated by larger Euclidian distances from the jth group than are the $j-1$ and $j+1$th groups, also have no effect. Therefore, the assumption above (2.193) is correct.

Since $d\Big(S_1^{(j)}, D_1^{(j)}\Big) = d\Big(D_2^{(j)}, S_4^{(j)}\Big) = \frac{1}{2}d_{M_K}$, and from (2.196) and (2.197), it can be shown that the probability of correct decision for $S_i^{(j)}$ $(i = 1, 4)$, P_c^{bit0}, under the condition of this theorem is given by

$$P_c^{\text{bit0}} = \frac{1}{2}\Big(P_{c,S_1}^{\text{bit0}} + P_{c,S_4}^{\text{bit0}}\Big) \approx 1 - Q\left(\frac{d_{M_K}}{2}\sqrt{\frac{2T}{N_0}}\right). \tag{2.198}$$

ii) Signal points assigned for bit 1 when $2 \leq j \leq 2^{K-2} - 1$

It can be shown that the probability of correct decision for $S_2^{(j)}$, denoted by P_{c,S_2}^{bit1}, is given by

$$\begin{aligned} P_{c,S_2}^{\text{bit1}} = \Pr\Big[- d\Big(D_1^{(j-1)}, S_2^{(j)}\Big)T < N < -d\Big(D_2^{(j-1)}, S_2^{(j)}\Big)T\Big] \\ +\Pr\Big[- d\Big(D_1^{(j)}, S_2^{(j)}\Big)T < N < d\Big(S_2^{(j)}, D_2^{(j)}\Big)T\Big] \\ +\Pr\Big[d\Big(S_2^{(j)}, D_1^{(j+1)}\Big)T < N < d\Big(S_2^{(j)}, D_2^{(j+1)}\Big)T\Big]. \end{aligned} \tag{2.199}$$

Eq. (2.199) can be rewritten as

$$\begin{aligned} P_{c,S_2}^{\text{bit1}} = 1 + \Pr\Big[d\Big(D_2^{(j-1)}, S_2^{(j)}\Big)T < N < d\Big(D_1^{(j-1)}, S_2^{(j)}\Big)T\Big] \\ -\Pr\Big[N > d\Big(S_2^{(j)}, D_2^{(j)}\Big)T\Big] - \Pr\Big[N > d\Big(D_1^{(j)}, S_2^{(j)}\Big)T\Big] \\ +\Pr\Big[d\Big(S_2^{(j)}, D_1^{(j+1)}\Big)T < N < d\Big(S_2^{(j)}, D_2^{(j+1)}\Big)T\Big]. \end{aligned} \tag{2.200}$$

From (2.189), (2.190), and Fig. 2.15, it can be shown that

$$\begin{aligned} d\Big(D_2^{(j-1)}, S_2^{(j)}\Big) &\geq \frac{3}{2}d_{M_K} + d_{M_{K-2}} \\ d\Big(S_2^{(j)}, D_1^{(j+1)}\Big) &\geq \frac{3}{2}d_{M_K} + d_{M_{K-1}} + d_{M_{K-2}} \\ d\Big(S_2^{(j)}, D_2^{(j)}\Big) &= \frac{1}{2}d_{M_K} + d_{M_{K-1}} \\ d\Big(D_1^{(j)}, S_2^{(j)}\Big) &= \frac{1}{2}d_{M_K}. \end{aligned} \tag{2.201}$$

Note that the first two distances in (2.201) are greater than $\frac{1}{2}d_{M_K} + d_{M_{K-1}}$, whereas the last distance is smaller than $\frac{1}{2}d_{M_K} + d_{M_{K-1}}$. Hence, the second and fifth terms of P_{c,S_2}^{bit1}, given by (2.200), are insignificant when the condition of this theorem is satisfied, whereas the fourth term of P_{c,S_2}^{bit1} is not affected by the condition of this theorem. Due to the fact that the third distance in (2.201) is the exactly the same as $\frac{1}{2}d_{M_K} + d_{M_{K-1}}$, the third term of P_{c,S_2}^{bit1} can be either discarded or preserved when the the condition of this theorem is satisfied. If we do not discard the third term, under the condition of this theorem, P_{c,S_2}^{bit1}, given by (2.200), becomes

$$P_{c,S_2}^{\text{bit1}} \approx 1 - \Pr\Big[N > d\Big(S_2^{(j)}, D_2^{(j)}\Big)T\Big] - \Pr\Big[N > d\Big(D_1^{(j)}, S_2^{(j)}\Big)T\Big]. \tag{2.202}$$

In a similar manner, it can be shown that if the condition of this theorem is satisfied, the probability of correct decision for $S_3^{(j)}$, denoted by P_{c,S_3}^{bit1}, becomes

$$P_{c,S_3}^{\text{bit1}} \approx 1 - \Pr\Big[N > d\Big(S_3^{(j)}, D_2^{(j)}\Big)T\Big] - \Pr\Big[N > d\Big(D_1^{(j)}, S_3^{(j)}\Big)T\Big]. \tag{2.203}$$

Since $d\Big(S_2^{(j)}, D_2^{(j)}\Big)$=$d\Big(D_1^{(j)}, S_3^{(j)}\Big)$ =$\frac{1}{2}d_{M_K} + d_{M_{K-1}}$ and $d\Big(D_1^{(j)}, S_2^{(j)}\Big)$ =$d\Big(S_3^{(j)}, D_2^{(j)}\Big)$= $\frac{1}{2}d_{M_K}$, and from (2.202) and (2.203), it can be shown that the probability of correct decision for $S_i^{(j)}$ $(i = 2, 3)$, P_c^{bit1}, under the condition of this theorem is given by

$$\begin{aligned} P_c^{\text{bit1}} &= \frac{1}{2}\Big(P_{c,S_2}^{\text{bit1}} + P_{c,S_3}^{\text{bit1}}\Big) \\ &\approx 1 - Q\left(\left(d_{M_{K-1}} + \frac{d_{M_K}}{2}\right)\sqrt{\frac{2T}{N_0}}\right) - Q\left(\frac{d_{M_K}}{2}\sqrt{\frac{2T}{N_0}}\right). \end{aligned} \tag{2.204}$$

iii) Signal points assigned for bit 0/1 when $j = 1$ or $j = 2^{K-2}$

It can be shown that the results of i) and ii) hold for the case $j = 1$ or $j = 2^{K-2}$.

From i)–iii), it follows that if the SNR condition of this theorem is satisfied, the BER of the Kth MSB (or LSB), P_{M_K}, is derived as

$$\begin{aligned} P_{M_K} &= 1 - \frac{1}{2}\Big(P_c^{bit0} + P_c^{bit1}\Big) \\ &\approx Q\left(\frac{d_{M_K}}{2}\sqrt{\frac{2T}{N_0}}\right) + \frac{1}{2}Q\left(\left(d_{M_{K-1}} + \frac{d_{M_K}}{2}\right)\sqrt{\frac{2T}{N_0}}\right) \end{aligned} \tag{2.205}$$

where P_c^{bit0} and P_c^{bit1} are given by (2.198) and (2.204), respectively.

From (2.151), (2.187), and (2.205), the BER of the n_0th MSB $(1 \le n_0 \le K)$ for hierarchical 2^K PAM can be expressed as

$$\begin{cases} \sum_{p=0}^{2^{K-n_0}-1} \frac{1}{2^{K-n_0}} Q\left(\left(\frac{d_{M_{n_0}}}{2} + \sum_{r=n_0+1}^{K} \left\lfloor \frac{p+2^{K-r}}{2^{K-r+1}} \right\rfloor d_{M_r}\right)\sqrt{\frac{2T}{N_0}}\right), \\ \text{for } 1 \le n_0 \le K-1 \\ Q\left(\frac{d_{M_K}}{2}\sqrt{\frac{2T}{N_0}}\right) + \frac{1}{2}Q\left(\left(d_{M_{K-1}} + \frac{d_{M_K}}{2}\right)\sqrt{\frac{2T}{N_0}}\right), \\ \text{for } n_0 = K. \end{cases} \tag{2.206}$$

Note that (2.206) is the exact BER expression for the MSB, but for $2 \le n_0 \le K$ th MSB, (2.206) holds if the condition of this theorem is satisfied. Lastly, it can be shown that the BER of the inphase or quadrature components for hierarchical 2^{2K} QAM is the same as that for hierarchical 2^K PAM. For hierarchical 2^{2K} QAM, let $E_s = P_{avg}T$ denote the average energy of the transmitted signal. Setting $2T/N_0 = 2E_s/N_0P_{avg} = 2\gamma_s/P_{avg}$ in (2.206), (2.16) is derived.

2.10 Appendix B: Numerical Evaluation of the BER Expression (2.16)

Figs. 2.16–2.19 show the numerical evaluation of the BER expression given by (2.16) for hierarchical 64 and 256 QAM when the distance ratio is 1 or 2.

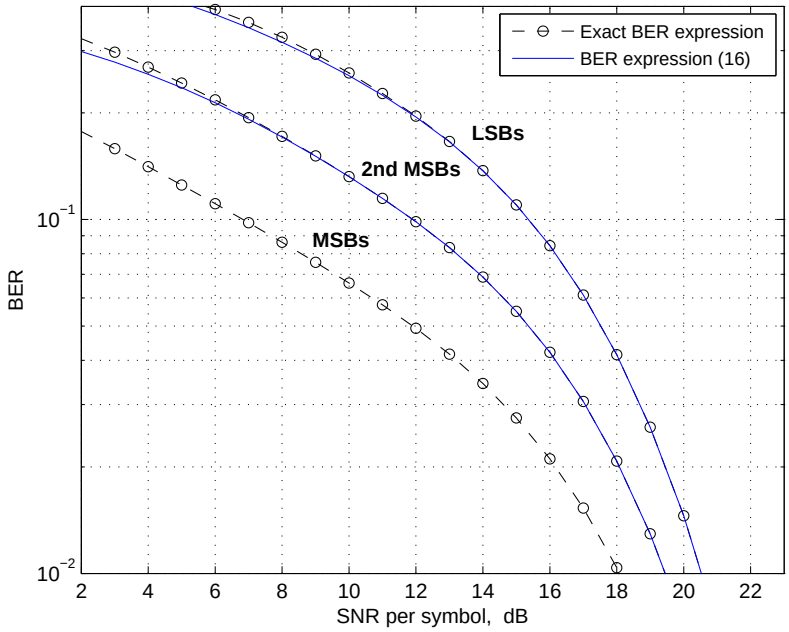


Figure 2.16: BER for hierarchical 64 QAM (the distance ratio $d_{M_{n-1}}/d_{M_n} = 1$ i.e., the lower bound).

2.11 Acknowledgements

The text of this chapter, in full, is a reprint of "Optimized Unequal Error Protection using Multiplexed Hierarchical Modulation," which has been submitted for publication in *IEEE Transactions on Information Theory*. The dissertation author was the primary investigator of this paper.

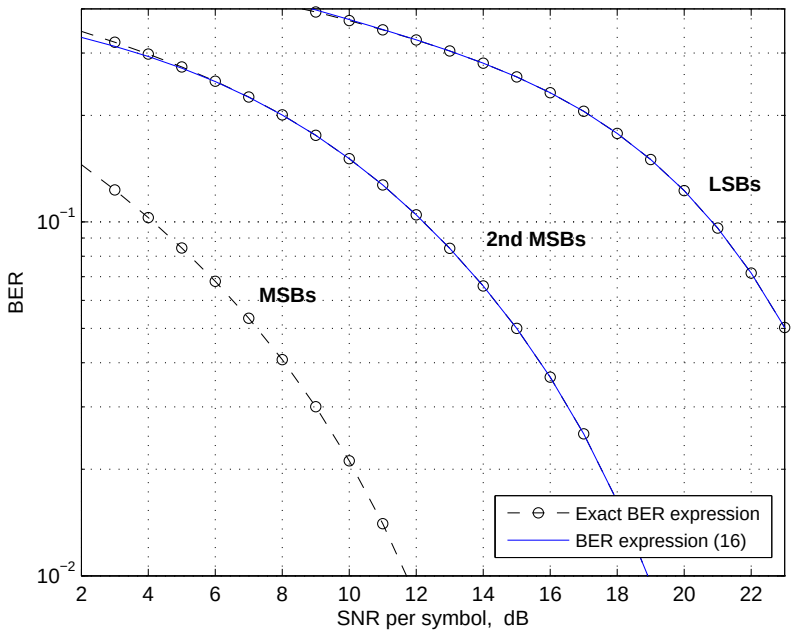


Figure 2.17: BER for hierarchical 64 QAM (the distance ratio $d_{M_{n-1}}/d_{M_n} = 2$).

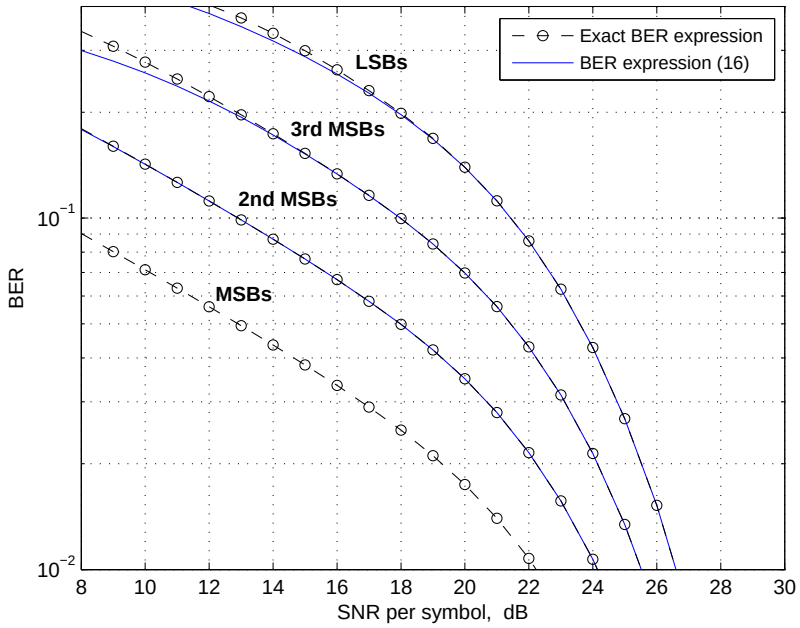


Figure 2.18: BER for hierarchical 256 QAM (the distance ratio $d_{M_{n-1}}/d_{M_n} = 1$ i.e., the lower bound).

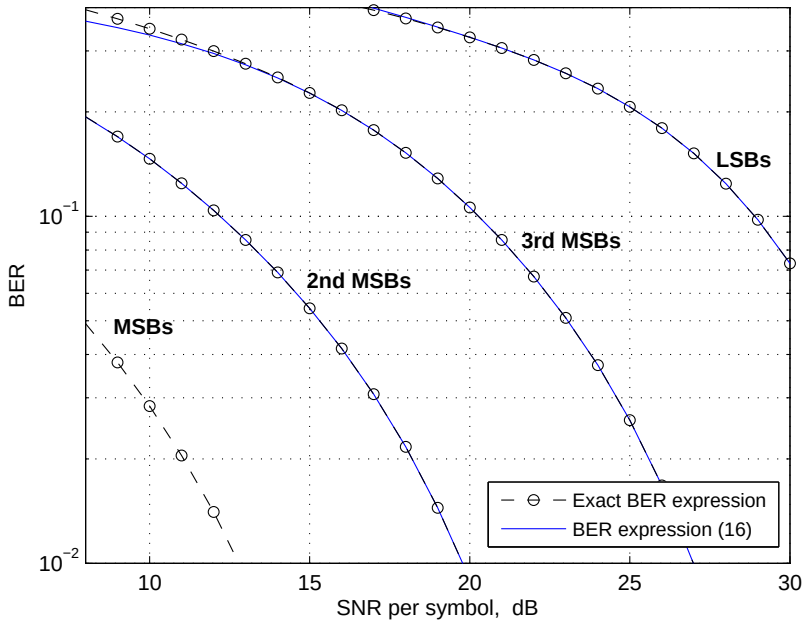


Figure 2.19: BER for hierarchical 256 QAM (the distance ratio $d_{M_{n-1}}/d_{M_n} = 2$).

Chapter 3

Superposition MIMO Coding for the Broadcast of Layered Sources

3.1 Introduction

Recently, multiple-input multiple-output (MIMO) systems have received a great deal of attention, since they can improve capacity and reliability relative to single-input single-output (SISO) systems. Two popular techniques for MIMO systems are space-time coding [27]– [30] and spatial multiplexing [31]– [34]. Space-time coding is an approach where information is spread across multiple transmit antennas to maximize spatial diversity in fading channels. Spatial multiplexing is an approach whereby independent information is transmitted on each antenna, and thus the transmit data rate is increased without additional system bandwidth.

In this chapter, we propose superposition MIMO coding for the transmission of layered sources such as progressive images or scalable video in a point-to-multipoint system. In a point-to-multipoint transmission such as broadcasting or multicasting, a single source transmits an encoded signal to multiple receivers. Depending on its location, each receiver experiences different channel conditions in terms of the received signal strength and the level of interference and noise power. Even if channel state information for each receiver were available at the transmitter based on feedback from the receivers, the modulation alphabet size and the MIMO mode cannot be adapted at each receiver due to the nature of broadcast transmission. We first analyze the tradeoff between space-time coding and spatial multiplexing under the constraint that

the two systems have the same maximum data rate. In this analysis, the modulation alphabet sizes for space-time coding and spatial multiplexing are chosen to be different such that the maximum data rates for both schemes are the same. The results show that for a given target bit error rate (BER), space-time coding is preferable to spatial multiplexing for a low data rate (i.e., small alphabet size), and vice versa for a high data rate (i.e., large alphabet size). In a hierarchical transmission [18]– [21] [24], important components and less important components, which are delivered by the basic subconstellation and the secondary subconstellation, respectively, do not necessarily have the same data rate. A typical example is scalable video. The base layer, which is more important, has a smaller number of bits than does the enhancement layer. In other words, the more important component has lower data rate than does the less important component. Therefore, when a layered source is transmitted hierarchically in MIMO systems, a tradeoff between two different MIMO approaches, space-time coding and spatial multiplexing, should be considered for each component of the layered source. Based on this fact and the analysis of the tradeoff between space-time coding and spatial multiplexing, we propose a layered source broadcasting system where two different MIMO techniques are hierarchically combined. More specifically, Alamouti coding is applied for the more important component which has lower data rate, in order to maximize the performance for the receivers with poor channel qualities. Spatial multiplexing is applied for the less important component having higher data rate, which is decoded only by receivers having good channel conditions. Superposition of two different MIMO approaches is embodied in a way that basic subconstellation symbols are encoded with Alamouti coding, secondary subconstellation streams are spatially multiplexed, and then two subconstellation symbols are superposed to construct the final transmit symbols.

The rest of the chapter is organized as follows. In Section 3.2, we analyze the tradeoff between Alamouti coding [27] and spatial multiplexing having the same maximum data rate. Based on the analysis, in Section 3.3 we propose a superposition of different MIMO approaches. In Section 3.4, numerical results are provided, and we conclude our work in Section 3.5.

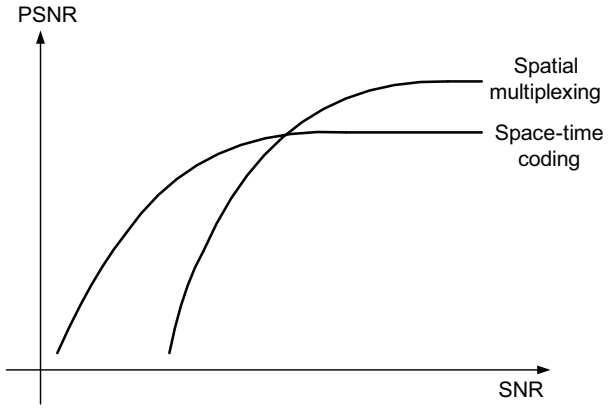


Figure 3.1: The PSNR performance of spatial multiplexing and space-time coding for the same modulation alphabet size.

3.2 Comparison of Alamouti Coding and Spatial Multiplexing Having the Same Maximum Data Rate

When the same modulation alphabet size is employed for both spatial multiplexing and space-time coding schemes, the former achieves better peak-signal-to-noise ratio (PSNR) performance at high SNR due to a increased data rate, whereas the latter retains robustness at low SNR. This is qualitatively depicted in Fig. 3.1. Since we are considering a point-to-multipoint transmission where a single source transmits data to multiple receivers having various channel qualities, we cannot optimally switch between two different MIMO modes (i.e, both modulation and MIMO mode are fixed). For this case, as a way to compare spatial multiplexing and space-time coding schemes fairly, the maximum data rates of both are set to be equal.

3.2.1 Channel Model

The MIMO system is equipped with N_T transmit and $N_R(\geq N_T)$ receive antennas. The propagation channel from the transmitter to the receiver is characterized by an $N_R \times N_T$ channel matrix $\mathbf{H}$ whose element h_{jk} at the jth row and kth column is the channel gain from the kth transmit antenna to the jth receive antenna, and the h_{jk}'s are assumed to be independent and identically distributed (i.i.d) complex Gaussian random variables with zero mean and unit variance. The N_R-dimensional received signal

vector $\mathbf{y} = [y_1\ y_2 \cdots y_{N_R}]^T$ after matched filtering and sampling can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n} \tag{3.1}$$

where $\mathbf{s} = [s_1\ \ s_2 \cdots s_{N_T}]^T$ denotes the N_T-dimensional transmit symbol vector, and $\mathbf{n} = [n_1\ n_2 \cdots n_{N_R}]^T$ is the N_R-dimensional noise vector whose elements are assumed to be i.i.d. complex Gaussian random variables with zero mean and variance of σ_n^2. It is assumed that transmission bandwidth is much less than the coherence bandwidth of the channel, and that the symbol time period is much less than the coherence time.

3.2.2 Average BER

We first express the average BER of the Alamouti scheme for an M-ary square QAM constellation. A closed-form expression for the average BER of such a constellation for SISO systems in an AWGN channel is given by [51, eq. (14)]

$$P_{b,SISO} = \frac{4}{\sqrt{M}\log_2 M} \sum_{k=1}^{\log_2\sqrt{M}} \sum_{i=0}^{(1-2^{-k})\sqrt{M}-1} \left\{ (-1)^{\left\lfloor \frac{i\cdot 2^{k-1}}{\sqrt{M}} \right\rfloor} \left(2^{k-1} - \left\lfloor \frac{i\cdot 2^{k-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor \right) \times Q\left((2i+1)\sqrt{\frac{3\gamma_s}{M-1}} \right) \right\} \tag{3.2}$$

where γ_s is the signal-to-noise ratio (SNR) per symbol. We define the average SNR per symbol in a MIMO system as

$$\gamma_s \triangleq \frac{E\left[|s_k|^2\right]}{\sigma_n^2}, \qquad k = 1, 2, \cdots, N_T \tag{3.3}$$

where s_k is the kth component of the transmit symbol vector $\mathbf{s}$ (i.e., a constellation symbol transmitted from the kth antenna). For the Alamouti scheme, the instantaneous post-processing SNR per symbol at the receiver end, denoted by η, can be expressed as ([52, eq. (5.40)])

$$\eta = \gamma_s \|\mathbf{H}\|_F^2 \tag{3.4}$$

where $\|\mathbf{H}\|_F^2$ is the squared Frobenius norm of $\mathbf{H}$. The η has a chi-square probability density function (PDF) with $2N_TN_R$ degrees of freedom [52, eq. (3.43)]:

$$f_\eta(x) = \frac{x^{N_TN_R-1}}{(N_TN_R-1)!\,\gamma_s^{N_TN_R}}\, e^{-\frac{x}{\gamma_s}}, \qquad x > 0. \tag{3.5}$$

From [53, eq. (3.37)], it can be shown that

$$\int_0^\infty Q\left(\sqrt{\alpha x}\right) \frac{x^{N_T N_R - 1}}{(N_T N_R - 1)!\, \gamma_s^{N_T N_R}}\, e^{-\frac{x}{\gamma_s}}\, dx$$
$$= \left(\frac{1-\beta}{2}\right)^{N_T N_R} \sum_{j=0}^{N_T N_R - 1} \left\{ \binom{N_T N_R - 1 + j}{j} \times \left(\frac{1+\beta}{2}\right)^j \right\} \tag{3.6}$$

where $\beta = \sqrt{\frac{\alpha\,\gamma_s}{2+\alpha\,\gamma_s}}$. In the Alamouti scheme, an identical constellation symbol, s_k, is transmitted twice during two symbol time periods ([52, eqs. (5.34) and (5.35)]), [27], and thus for an M-ary QAM constellation, the SNR per bit, γ_b, is derived from γ_s using a factor of two:

$$\gamma_b = 2 \times \frac{\gamma_s}{\log_2 M}. \tag{3.7}$$

From (3.2), (3.6) and (3.7), the exact BER of the Alamouti scheme for an M-ary square QAM constellation can be expressed as

$$P_{b,Alamouti} = \frac{4}{\sqrt{M}\log_2 M} \sum_{k=1}^{\log_2\sqrt{M}} \sum_{i=0}^{(1-2^{-k})\sqrt{M}-1} \left[(-1)^{\left\lfloor \frac{i\cdot 2^{k-1}}{\sqrt{M}} \right\rfloor} \left(2^{k-1} - \left\lfloor \frac{i\cdot 2^{k-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor \right) \right.$$
$$\left. \times \left(\frac{1-\mu(i)}{2}\right)^{N_T N_R} \sum_{j=0}^{N_T N_R - 1} \left\{ \binom{N_T N_R - 1 + j}{j} \left(\frac{1+\mu(i)}{2}\right)^j \right\} \right] \tag{3.8}$$

where

$$\mu(i) = \sqrt{\frac{3(2i+1)^2 \left(\log_2 M\right)\gamma_b}{4(M-1) + 3(2i+1)^2 \left(\log_2 M\right)\gamma_b}}.$$

We next derive the average BER of the spatial multiplexing with a zero forcing (ZF) receiver for M-ary square QAM. It has been shown that for the ZF receiver, the instantaneous post-processing SNR on each independent transmit stream is a chi-squared random variable with $2(N_R - N_T + 1)$ degrees of freedom [54] [55], and thus the exact BER expression is achievable. The PDF of the instantaneous post-processing SNR on each transmit stream is given by

$$f_\eta(x) = \frac{x^{N_R - N_T}}{(N_R - N_T)!\, \gamma_s^{N_R - N_T + 1}}\, e^{-\frac{x}{\gamma_s}}, \qquad x > 0 \tag{3.9}$$

where γ_s is defined in (3.3). It can be shown that

$$\int_0^\infty Q\left(\sqrt{\alpha x}\right) \frac{x^{N_R-N_T}}{(N_R-N_T)!\,\gamma_s^{N_R-N_T+1}}\, e^{-\frac{x}{\gamma_s}}\, dx$$
$$= \left(\frac{1-\beta}{2}\right)^{N_R-N_T+1} \sum_{j=0}^{N_R-N_T} \left\{ \binom{N_R-N_T+j}{j} \times \left(\frac{1+\beta}{2}\right)^j \right\} \tag{3.10}$$

where $\beta = \sqrt{\frac{\alpha\,\gamma_s}{2+\alpha\,\gamma_s}}$. For spatial multiplexing, the SNR per bit, γ_b, is given by

$$\gamma_b = \frac{\gamma_s}{\log_2 M}. \tag{3.11}$$

Note that (3.7) and (3.11) differ by a factor of two. From (3.2), (3.10) and (3.11), the exact BER of the spatial multiplexing with ZF receiver for an M-ary square QAM constellation can be expressed as

$$P_{b,SM-ZF} = \frac{4}{\sqrt{M}\log_2 M} \sum_{k=1}^{\log_2\sqrt{M}} \sum_{i=0}^{(1-2^{-k})\sqrt{M}-1} \left[(-1)^{\left\lfloor \frac{i\cdot 2^{k-1}}{\sqrt{M}} \right\rfloor} \left(2^{k-1} - \left\lfloor \frac{i\cdot 2^{k-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor \right) \right.$$
$$\left. \times \left(\frac{1-\mu(i)}{2}\right)^{N_R-N_T+1} \sum_{j=0}^{N_R-N_T} \left\{ \binom{N_R-N_T+j}{j} \left(\frac{1+\mu(i)}{2}\right)^j \right\} \right] \tag{3.12}$$

where

$$\mu(i) = \sqrt{\frac{3(2i+1)^2\left(\log_2 M\right)\gamma_b}{2(M-1)+3(2i+1)^2\left(\log_2 M\right)\gamma_b}}.$$

3.2.3 High SNR Approximation (Minimum Euclidian Distance Approximation) of the Average BER

For high SNR, the BER is approximated by the Q-function term having the minimum Euclidian distance. If we discard the terms having non-minimum Euclidian distances in (3.2) (i.e., only consider $i = 0$ in (3.2)), we have

$$P_{b,SISO} \approx \frac{4}{\sqrt{M}\log_2 M} \sum_{k=1}^{\log_2\sqrt{M}} 2^{k-1}\, Q\left(\sqrt{\frac{3\gamma_s}{M-1}}\right)$$
$$= \frac{4(\sqrt{M}-1)}{\sqrt{M}\log_2 M} Q\left(\sqrt{\frac{3\gamma_s}{M-1}}\right) \tag{3.13}$$

where the second equality follows from $\sum_{k=1}^{\log_2\sqrt{M}} 2^{k-1} = \sqrt{M}-1$. From (3.6) and (3.13), we have

$$P_{b,Alamouti} \approx \frac{4(\sqrt{M}-1)}{\sqrt{M}\log_2 M}\left(\frac{1-\mu}{2}\right)^{N_TN_R}\sum_{j=0}^{N_TN_R-1}\binom{N_TN_R-1+j}{j}\left(\frac{1+\mu}{2}\right)^{j} \quad (3.14)$$

where

$$\mu = \sqrt{\frac{3(\log_2 M)\gamma_b}{4(M-1)+3(\log_2 M)\gamma_b}}. \quad (3.15)$$

Further, for high SNR, (3.15) can be approximated as

$$\mu \approx 1 - \frac{2(M-1)}{3(\log_2 M)\gamma_b} \quad (3.16)$$

where we have used $\sqrt{\frac{x}{1+x}} \approx 1 - \frac{1}{2x}$ for $x >> 1$. From (3.16), we have

$$\frac{1-\mu}{2} \approx \frac{M-1}{3(\log_2 M)\gamma_b} \quad \text{and} \quad \frac{1+\mu}{2} \approx 1 - \frac{M-1}{3(\log_2 M)\gamma_b} \approx 1. \quad (3.17)$$

Using (3.17), it can be shown that (3.14) can be rewritten as

$$P_{b,Alamouti} \approx P^{app}_{b,Alamouti} = \binom{2N_TN_R-1}{N_TN_R}\frac{4(\sqrt{M}-1)}{\sqrt{M}\log_2 M}\left(\frac{M-1}{3\log_2 M}\right)^{N_TN_R}\left(\frac{1}{\gamma_b}\right)^{N_TN_R} \quad (3.18)$$

In the same way, it can be shown that for high SNR, $P_{b,SM-ZF}$, given by (3.12), can be approximated as

$$P_{b,SM-ZF} \approx P^{app}_{b,SM-ZF} = \binom{2(N_R-N_T)-1}{N_R-N_T}\frac{4(\sqrt{M}-1)}{\sqrt{M}\log_2 M}\left(\frac{M-1}{6\log_2 M}\right)^{N_R-N_T+1} \times \left(\frac{1}{\gamma_b}\right)^{N_R-N_T+1}. \quad (3.19)$$

3.2.4 Comparison of High SNR BERs of Alamouti scheme and Spatial Multiplexing for the Same Maximum Data Rate

1) Crossover point for SNR: We compare the high SNR approximate BERs of the Alamouti scheme and the spatial multiplexing scheme having the same maximum data rate. To do this, we employ m-ary QAM for spatial multiplexing, and m^2-ary QAM for the Alamouti scheme. It is assumed that m is greater than or equal to 4 (i.e., QPSK). Note that $N_R \geq N_T = 2$. If we let $M = m^2$ in (3.18), we have

$$P^{app}_{b,Alamouti} = \binom{2N_TN_R-1}{N_TN_R}\frac{2(m-1)}{m\log_2 m}\left(\frac{m^2-1}{6\log_2 M}\right)^{N_TN_R}\left(\frac{1}{\gamma_b}\right)^{N_TN_R}. \quad (3.20)$$

If we let $M = m$ in (3.19), we have

$$P_{b,SM-ZF}^{app} = \binom{2(N_R - N_T) - 1}{N_R - N_T} \frac{4(\sqrt{m} - 1)}{\sqrt{m} \log_2 m} \left(\frac{m-1}{6 \log_2 m} \right)^{N_R - N_T + 1} \left(\frac{1}{\gamma_b} \right)^{N_R - N_T + 1} \tag{3.21}$$

We find the SNR, γ_b^*, for which (3.20) and (3.21) are the same. That is,

$$\begin{aligned} &\binom{2N_T N_R - 1}{N_T N_R} \frac{2(m-1)}{m \log_2 m} \left(\frac{m^2 - 1}{6 \log_2 M} \right)^{N_T N_R} \left(\frac{1}{\gamma_b^*} \right)^{N_T N_R} \\ &= \binom{2(N_R - N_T) - 1}{N_R - N_T} \frac{4(\sqrt{m} - 1)}{\sqrt{m} \log_2 m} \left(\frac{m-1}{6 \log_2 m} \right)^{N_R - N_T + 1} \left(\frac{1}{\gamma_b^*} \right)^{N_R - N_T + 1}. \end{aligned} \tag{3.22}$$

It can be shown that the γ_b^* satisfying (3.22) is given by

$$\begin{aligned} \gamma_b^* = \Bigg\{ &\frac{\binom{2N_T N_R - 1}{N_T N_R}}{2\binom{2(N_R - N_T) - 1}{N_R - N_T} 6^{(N_R+1)(N_T-1)}} \\ &\times \frac{(\sqrt{m} + 1)(m-1)^{(N_R+1)(N_T-1)} (m+1)^{N_T N_R}}{\sqrt{m} (\log_2 m)^{(N_R+1)(N_T-1)}} \Bigg\}^{\frac{1}{(N_R+1)(N_T-1)}}. \end{aligned} \tag{3.23}$$

We define the function $f(m)$ as

$$f(m) = (\sqrt{m} + 1) \frac{m+1}{\sqrt{m}} \left(\frac{m-1}{\log_2 m} \right)^{(N_R+1)(N_T-1)} (m+1)^{N_T N_R - 1}. \tag{3.24}$$

Also, let $f_1(m) = (m+1)/\sqrt{m}$, and let $f_2(m) = (m-1)/\log_2 m$. Note that for $m \geq 4$, we have

$$\frac{df_1(m)}{dm} = \frac{m-1}{2m\sqrt{m}} > 0 \quad \text{and} \quad \frac{df_2(m)}{dm} = \frac{\log_2 m + \frac{1}{m} - 1}{(\log_2 m)^2} > 0. \tag{3.25}$$

From (3.24) and (3.25), it is clear that $f(m)$ is a strictly increasing function in m. From (3.23) and (3.24), it is seen that as the alphabet size, m, increases, γ_b^* strictly increases, regardless of the number of receive antennas.

2) Crossover point for BER: It can be shown that if we substitute γ_b^*, given by (3.23), into (3.20), the corresponding BER, P_b^*, is given by

$$\begin{aligned} P_b^* = 2\binom{2N_T N_R - 1}{N_T N_R} &\left\{ \frac{2\binom{2(N_R - N_T) - 1}{N_R - N_T}}{\binom{2N_T N_R - 1}{N_T N_R}} \right\}^{\frac{N_T N_R}{(N_R+1)(N_T-1)}} \frac{m-1}{m \log_2 m} \left(\frac{\sqrt{m}}{\sqrt{m}+1} \right)^{\frac{N_T N_R}{(N_R+1)(N_T-1)}} \\ &\times \left(\frac{1}{m+1} \right)^{\frac{N_T N_R (N_R - N_T + 1)}{(N_R+1)(N_T-1)}}. \end{aligned} \tag{3.26}$$

Let

$$g(m) = \frac{m-1}{m\log_2 m}\left(\frac{\sqrt{m}}{\sqrt{m}+1}\right)^{\frac{N_T N_R}{(N_R+1)(N_T-1)}}\left(\frac{1}{m+1}\right)^{\frac{N_T N_R(N_R-N_T+1)}{(N_R+1)(N_T-1)}}$$
$$= \frac{m-1}{m\log_2 m}\left(\frac{1}{\sqrt{m}+1}\right)^{\frac{N_T N_R}{(N_R+1)(N_T-1)}}\left(\frac{\sqrt{m}}{m+1}\right)^{\frac{N_T N_R}{(N_R+1)(N_T-1)}}$$
$$\times\left(\frac{1}{m+1}\right)^{\frac{N_T N_R(N_R-N_T)}{(N_R+1)(N_T-1)}}. \tag{3.27}$$

We let $g_1(m) = \sqrt{m}/(m+1)$, and $g_2(m) = (m-1)/(m\log_2 m)$. Then, for $m \geq 4$, we have

$$\frac{dg_1(m)}{dm} = \frac{1-m}{2\sqrt{m}(m+1)^2} < 0. \tag{3.28}$$

We also have

$$\frac{dg_2(m)}{dm} = \frac{\log_2 m - \frac{m}{\ln 2} + \frac{1}{\ln 2}}{(m\log_2 m)^2} < 0. \tag{3.29}$$

where the inequality is derived from the following:

i) Let $h(m)$ be the numerator of $dg_2(m)/dm$.

ii) For $m \geq 4$, we have $dh(m)/dm = (1-m)/(m\ln 2) < 0$. Further, $h(4) = 2 - 3/\ln 2 < 0$.

iii) Hence, $h(m) < 0$ for $m \geq 4$

From (3.27)–(3.29), $g(m)$ is a strictly decreasing function in m. From (3.26) and (3.27), as the alphabet size, m, increases, P_b^* strictly decreases, regardless of the number of receive antennas.

3) Comparison of the BERs: From (3.20) and (3.21), the ratio of $P_{b,Alamouti}^{app}$ to $P_{b,SM-ZF}^{app}$, denoted by $R(\gamma_b)$, is given by

$$R(\gamma_b) = \frac{\binom{2N_T N_R - 1}{N_T N_R}(\sqrt{m}+1)\left(\frac{m^2-1}{6\log_2 m}\right)^{N_T N_R}}{2\binom{2(N_R-N_T)-1}{N_R-N_T}\sqrt{m}\left(\frac{m-1}{6\log_2 m}\right)^{N_R-N_T+1}}\left(\frac{1}{\gamma_b}\right)^{(N_R+1)(N_T-1)}. \tag{3.30}$$

From (3.30), it is seen that $R(\gamma_b)$ is a strictly decreasing function in γ_b, and thus from $R(\gamma_b^*) = 1$, we have

$$P_{b,Alamouti}^{app} < P_{b,SM-ZF}^{app} \quad \text{for} \quad \gamma_b > \gamma_b^*$$
$$P_{b,Alamouti}^{app} > P_{b,SM-ZF}^{app} \quad \text{for} \quad \gamma_b < \gamma_b^*. \tag{3.31}$$

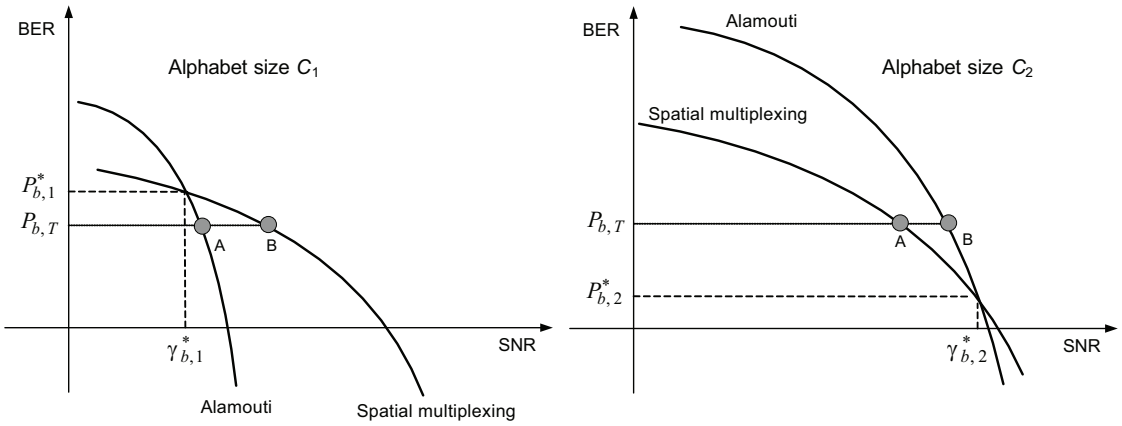


Figure 3.2: High SNR approximate BERs of the Alamouti and spatial multiplexing schemes for the same maximum data rate. For alphabet size $C_1 < C_2$, these BERs have the following properties: i) $\gamma^*_{b,1} < \gamma^*_{b,2}$ ii) $P^*_{b,1} > P^*_{b,2}$ iii) $P^{app}_{b,i,Alamouti} < P^{app}_{b,i,SM-ZF}$ for $\gamma_{b,i} > \gamma^*_{b,i}$, and $P^{app}_{b,i,Alamouti} > P^{app}_{b,i,SM-ZF}$ for $\gamma_{b,i} < \gamma^*_{b,i}$ $(i = 1, 2)$.

Let $P^*_{b,1}$ and $\gamma^*_{b,1}$ denote the crossover point when a modulation alphabet size $m = C_1$ is employed, and $P^*_{b,2}$ and $\gamma^*_{b,2}$ denote the crossover point when an alphabet size $m = C_2$ is used. Suppose that $C_1 < C_2$. Then, from the results below (3.25) and (3.29), we have

$$\gamma^*_{b,1} < \gamma^*_{b,2} \quad \text{and} \quad P^*_{b,1} > P^*_{b,2}. \tag{3.32}$$

Based on (3.31) and (3.32), the high SNR approximate BERs of the Alamouti scheme and spatial multiplexing with ZF receiver for the same maximum data rate are qualitatively depicted in Fig. 3.2. Suppose that a target bit error rate, $P_{b,T}$, is smaller than $P^*_{b,1}$ but greater than $P^*_{b,2}$. Then, from Fig. 3.2, it is seen that the Alamouti scheme is preferable to spatial multiplexing for an alphabet size C_1, whereas spatial multiplexing is preferable for an alphabet size C_2. From here onwards, for a given target bit error rate of $P_{b,T}$, we refer to an alphabet size which satisfies $P^*_b > P_{b,T}$ as a small alphabet size (i.e., low data rate), and refer to an alphabet size which satisfies $P^*_b < P_{b,T}$ as a large alphabet size (i.e., high data rate). Suppose that a single source transmits data to multiple receivers having various average SNRs in a broadcast system. If we properly employ the Alamouti scheme for a small alphabet size, and spatial multiplexing for a large alphabet size, receivers having average SNRs between points **A** and **B** in Fig. 3.2 would achieve the target bit error rate. Otherwise, only receivers having average SNRs greater than point **B** would accomplish the target bit error rate.

Using (3.8) and (3.12), we evaluate the exact BERs of the Alamouti scheme and

the spatial multiplexing with ZF receiver having the same maximum data rate. The BERs are evaluated for 2×2 MIMO systems for various maximum data rates (i.e., alphabet sizes), and the results are shown in Fig. 3.3 (a)–(c). The BER performance of spatial multiplexing with the optimal maximum likelihood (ML) receiver is also shown in Fig. 3.3 (a)–(c). We note that since the exact BER of the ML receiver is not analyzable, the curve is obtained from the simulation. From Fig. 3.3 (a)–(c), it is observed that as alphabet size increases, the crossover point for the exact BERs of the Alamouti scheme and spatial multiplexing with ZF receiver, γ_b^* and P_b^*, moves in a way predicted by the analysis based on the high SNR approximate BER expressions (see Fig. 3.2). Further, it is observed that as the alphabet size increases, the crossover point for the Alamouti and the spatial multiplexing scheme with ML receiver moves in the same way as that for the Alamouti and spatial multiplexing with ZF receiver. If we focus on a BER of 10^{-4}, it is seen that the Alamouti scheme outperforms spatial multiplexing with the ML receiver for the maximum data rate of 4 bits per symbol period (Fig. 3.3 (a)), whereas the latter outperforms the former for the maximum data rate of 10 bits per symbol period (Fig. 3.3 (c)). The two schemes perform roughly the same for the maximum data rate of 8 bits per symbol period (Fig. 3.3 (b)). This indicates that if the target bit error rate is 10^{-4}, the Alamouti scheme is preferable for maximum data rates less than 8 bits per symbol period, and spatial multiplexing with the ML receiver is preferable otherwise. We again note that this preference depends on the target bit error rate of an application. For example, if the target bit error rate is 10^{-6}, the Alamouti scheme outperforms the spatial multiplexing with ML receiver even for the maximum data rate of 10 bits per symbol period.

3.3 Superposition of Alamouti Coding and Spatial Multiplexing

In layered sources, important components and less important components, which are delivered by the basic and secondary subconstellations of hierarchical modulation, respectively, do not necessarily have the same data rate. For example, in scalable video, the base layer, which is more important, has a smaller number of bits than does the enhancement layer. In Section 3.2, we showed that, given the same maximum data rate, the Alamouti scheme is preferable to spatial multiplexing for a small alphabet size (i.e.,

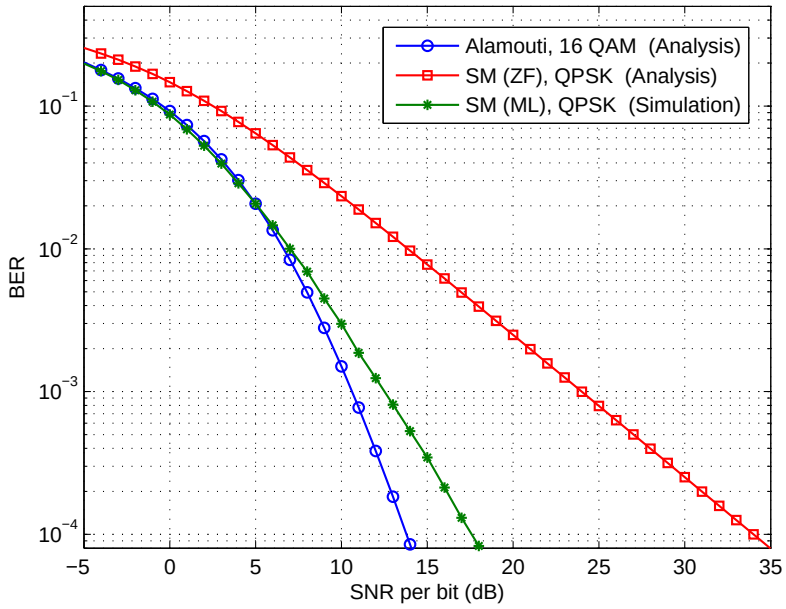


Fig. 3.3 (a) The maximum data rate: 4 bits per symbol period
(Alamouti coding for 16 QAM and spatial multiplexing for QPSK).

low data rate), and vice versa for a large alphabet size (i.e., high data rate). Therefore, when a hierarchical transmission is used in conjunction with MIMO systems for the broadcast of a layered source, we are motivated to apply different MIMO approaches for each component of the layered source. In particular, we consider the case where the more important component consists of a smaller number of bits than does the less important component. For this case, we apply the Alamouti scheme for the basic subconstellation, and spatial multiplexing for the secondary subconstellation.

The proposed superposition MIMO scheme is depicted in Fig. 3.4. We denote the basic subconstellation symbols for the more important component by $x_1[n]$ ($n = 0, 1, \cdots, 2L-1$), and the secondary subconstellation symbols for the less important component by $x_2[n]$ ($n = 0, 1, \cdots, 4L-1$). We demultiplex $x_2[n]$ into two transmit antenna streams, $x_{2A}[n]$ and $x_{2B}[n]$, such that

$$x_{2A}[n] = x_2[2n] \quad \text{and} \quad x_{2B}[n] = x_2[2n+1], \quad n = 0, 1, \cdots, 2L-1. \tag{3.33}$$

Then, for $n = 0, 1, \cdots, L-1$, the final transmit constellation symbols of the proposed

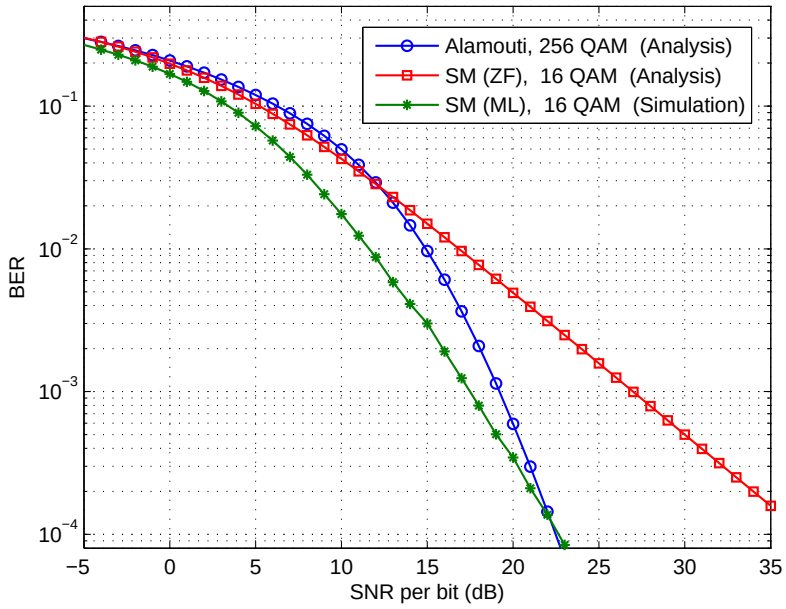


Fig. 3.3 (b) The maximum data rate: 8 bits per symbol period
(Alamouti coding for 256 QAM and spatial multiplexing for 16 QAM).

scheme are given by

$$\begin{bmatrix} S_A[2n] & S_A[2n+1] \\ S_B[2n] & S_B[2n+1] \end{bmatrix} = \begin{bmatrix} x_1[2n] + x_{2A}[2n] & -x_1^*[2n+1] + x_{2A}[2n+1] \\ x_1[2n+1] + x_{2B}[2n] & x_1^*[2n] + x_{2B}[2n+1] \end{bmatrix} \quad (3.34)$$

where each row corresponds to a transmit antenna, each column corresponds to a time symbol, $S_A[2n]$ and $S_A[2n+1]$ are hierarchical constellation symbols transmitted on antenna A, and $S_B[2n]$ and $S_B[2n+1]$ are hierarchical constellation symbols transmitted on antenna B. For example, Fig. 3.5 shows the proposed superposition MIMO coding where the Alamouti coding is applied to a QPSK basic subconstellation, and spatial multiplexing is applied to a 16 QAM secondary subconstellation.

In the following, we describe the decoding of the proposed scheme at the receiver. We first consider the optimal ML decoding of the proposed scheme. For ML decoding, the basic and secondary subconstellations are decoded at the same time. Since Alamouti coding is applied to the basic subconstellation in the proposed scheme (see eq. (3.34)), this indicates that ML decoding should be performed during two symbol time periods for both the basic and the secondary subconstellations (i.e., for an entire hierarchical constellation). This requires a complex receiver. Thus, we also consider the successive

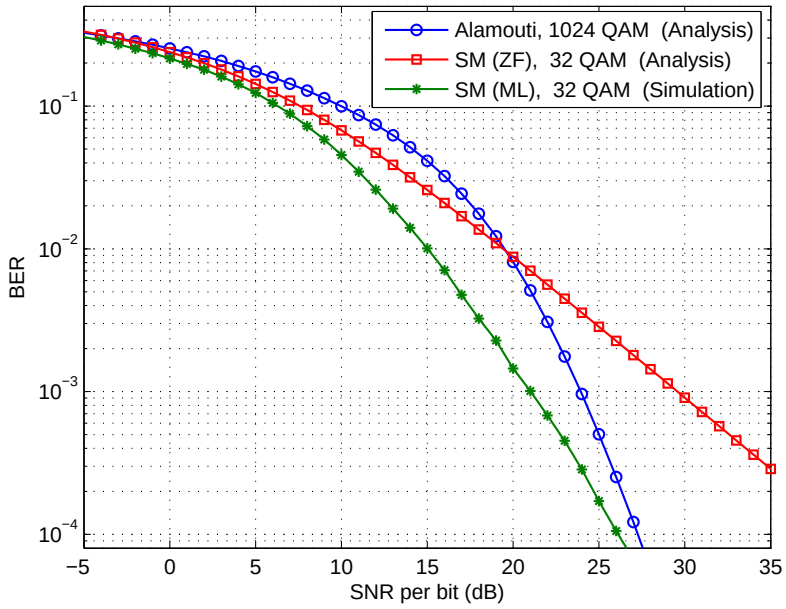


Fig. 3.3 (c) The maximum data rate: 10 bits per symbol period

(Alamouti coding for 1024 QAM and spatial multiplexing for 32 QAM).

Figure 3.3: The exact BERs of Alamouti scheme and spatial multiplexing for various alphabet sizes (i.e., for various maximum data rates) in 2×2 MIMO systems (SM denotes spatial multiplexing).

decoding of the proposed scheme. The successive decoding has the following steps:

1) Alamouti decoding is performed for the basic subconstellation.

2) The decoded basic subconstellation is subtracted from the received signal.

3) Spatial demultiplexing (such as ML, MMSE (minimum mean square error), or ZF decoding) is performed for the secondary subconstellation.

Note that in step 3, spatial demultiplexing such as ML decoding can be performed during only one symbol time period, since the basic subconstellation encoded with Alamouti coding has already been subtracted in step 2. This implies that even if ML decoding is used for spatial demultiplexing in step 3, the complexity of successive decoding is much less than that of the optimal ML decoding of an entire hierarchical constellation. From (3.34), it is seen that when the basic subconstellation symbols, $x_1[2n]$ and $x_1[2n+1]$, are Alamouti decoded in step 1, the secondary subconstellation symbols, $x_{2A}[2n]$, $x_{2A}[2n+1]$,

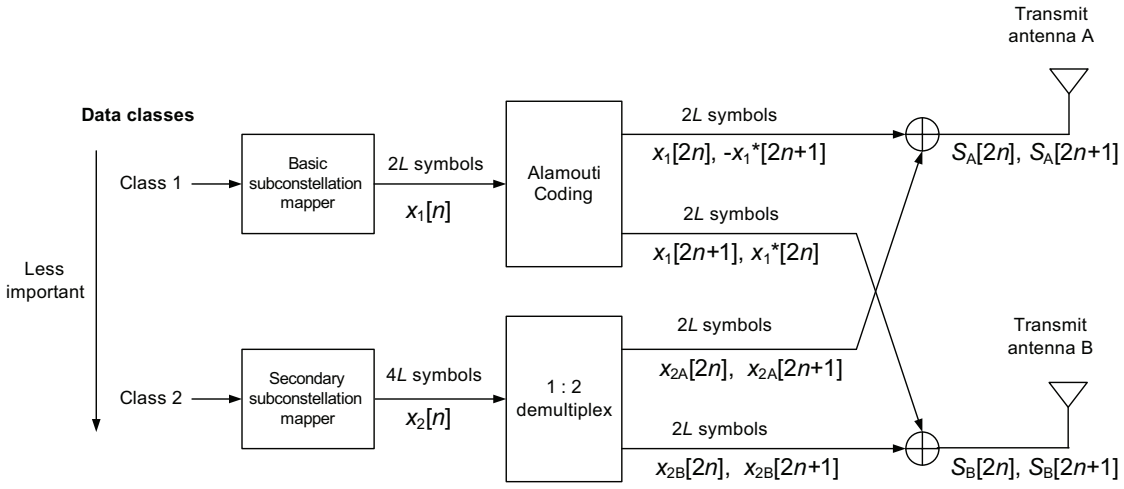
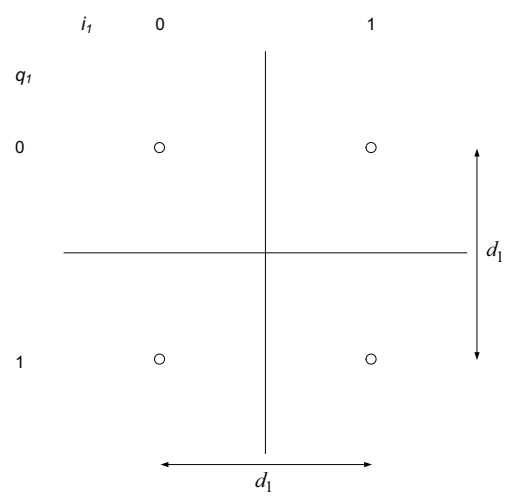


Figure 3.4: Superposition MIMO coding: two different MIMO codes are hierarchically combined such that Alamouti coding is applied for the more important data, and spatial multiplexing is applied for the less important data.

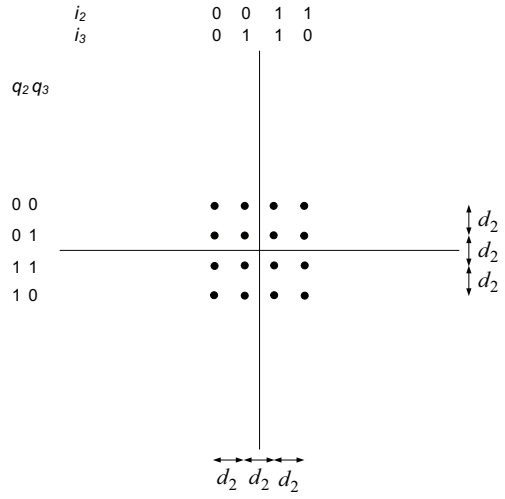
$x_{2B}[2n]$, and $x_{2B}[2n+1]$, act as interference. Therefore, the performance of the successive decoding depends on the distance ratio, α, of the hierarchical constellation, since the distance ratio is related to the energies of the basic and secondary subconstellations (note that the distance ratio is defined as the ratio of the minimum Euclidian distance of the basic subconstellation to that of the secondary subconstellation). The BER performance of the optimal ML decoding and successive decoding are compared in 2×2 MIMO systems for hierarchical 4/16 QAM with distance ratios of 2.0 and 4.0 (these are typical ratios for a hierarchical QAM constellation [24]). In successive decoding, ML decoding is performed for spatial demultiplexing in step 3. The results are shown in Fig. 3.6 (a)(b). It is seen that for a distance ratio of $\alpha = 2.0$, the performance of successive decoding is worse than that of the optimal ML decoding by SNR of 0.3 dB at BER of 10^{-3}, and for a distance ratio of $\alpha = 4.0$, the performance of successive decoding is nearly identical to that of the optimal ML decoding (SNR gap is 0.03 dB at BER of 10^{-3}).

3.4 Performance Evaluation

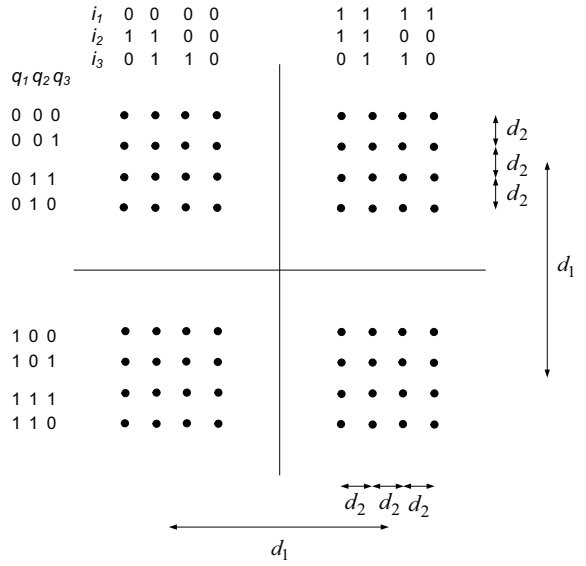
We evaluate the PSNR performance of the proposed superposition MIMO coding. In this evaluation, we consider hierarchical 4/64 QAM where the Alamouti code is applied for the QPSK basic subconstellation, and spatial multiplexing is applied for the 16 QAM



(a) $x_1[n]$: QPSK (Alamouti)



(b) $x_{2A}[n]$, $x_{2B}[n]$: 16 QAM (spatial multiplexing)



(c) $S_A[n]$, $S_B[n]$: hierarchical 4/64 QAM (Alamouti and spatial multiplexing are superposed).

Figure 3.5: Construction of superposition MIMO code.

secondary subconstellation. Note that for one symbol time period and two transmit antennas, the proposed scheme transmits 2 bits for the more important component (2 bits (QPSK) $\times$ 1 (Alamouti code)), and 8 bits for the less important component (4 bits (16 QAM) $\times$ 2 (spatial multiplexing)). For comparison purposes, we also evaluate the performance of the transmission schemes where different MIMO approaches are

not hierarchically applied. The first scheme is hierarchical 2/32 QAM with spatial multiplexing, and the second scheme is hierarchical 4/1024 QAM with Alamouti coding.[1] We refer to these as pure spatial multiplexing and pure Alamouti schemes. We note that the above three schemes under consideration have the same maximum data rate for the more important component (2 bits per symbol period) and the less important component (8 bits per symbol period). Regarding decoding methods at the receiver, we use ML decoding for the pure spatial multiplexing scheme, and Alamouti decoding for the pure Alamouti scheme. For the proposed scheme, we use successive decoding where ML decoding is performed for the secondary subconstellation. Note that the optimal decoding is used for both the pure spatial multiplexing and pure Alamouti schemes, whereas the successive decoding, which is suboptimal, is used for the proposed scheme. In our evaluation, error correction coding is not considered.

We evaluate the performance for 2×2 MIMO systems using the progressive source coder SPIHT [10] as an example, and provide the results for the standard 8 bits per pixel (bpp) 512×512 Lena image with a transmission rate of 0.375 bpp. We assume that the transmitted signal experiences a slow fading channel such that channel coefficients are nearly constant over an image, and that channel estimation at the receiver is perfect. Let $D(x)$ denote the operational rate-distortion curve. $D(0)$ refers to the distortion when the decoder must reconstruct the source without being able to use any of the transmitted information. For a still image, this means reconstructing the entire image at the mean pixel value, so the image is worthless. For a video, on the other hand, reconstruction of a frame without using any transmitted information for that frame might mean that the decoder will hold over the previous frame. For low motion videos, $D(0)$ might not be large. It can be shown that the end-to-end performance can be measured by the expected distortion of the image or video frame, $E[D]$, given by

$$E[D] = D(0)Pr(0) + \sum_{n=1}^{N} \left\{ D\left(\sum_{i=1}^{n} r_i\right) Pr\left(\sum_{i=1}^{n} r_i\right) \right\} \tag{3.35}$$

where N is the number of transmitted packets of the compressed progressive bitstream, r_i is the number of bits of the ith important packet, and $Pr(R)$ is the probability that the throughput of the progressive bitstream is R bits. Note that $E[D]$ is a function of the channel SNR and the distance ratio of the hierarchical constellation. For a performance

[1] Hierarchical 2/32 QAM consists of a BPSK basic subconstellation and a 16 QAM secondary subconstellation. Hierarchical 4/1024 QAM consists of a QPSK basic subconstellation and a 256 QAM secondary subconstellation.

comparison, we find the optimal distance ratio of a hierarchical constellation for each scheme which minimizes the expected distortion over a range of average SNRs using the weighted cost function

$$\arg\min_{\alpha} \frac{\int_0^\infty \omega(\gamma_b) E[D] d\gamma_b}{\int_0^\infty \omega(\gamma_b) d\gamma_b} \tag{3.36}$$

where α is a distance ratio, and $w(\gamma_b)$ in [0, 1] is the weight function. For example, $w(\gamma_b)$ is given by

$$\omega(\gamma_b) = \begin{cases} 1, & \text{for } \gamma_b^A \le \gamma_b \le \gamma_b^B \\ 0, & \text{otherwise.} \end{cases} \tag{3.37}$$

Note that in broadcast systems, the weight function given by (3.37) indicates that average SNRs of multiple receivers are uniformly distributed in the range of $\gamma_b^A \le \gamma_b \le \gamma_b^B$. Eq. (3.36) indicates that α is chosen such that the sum of the expected distortion of the receivers distributed in the range of $\gamma_b^A \le \gamma_b \le \gamma_b^B$ is minimized. To compare the image quality, we use PSNR which is defined in (2.117) as

$$\text{PSNR} = 10\log\frac{255^2}{E[D]} \quad \text{(dB)} \tag{3.38}$$

where $E[D]$ is given by (3.35). We present the PSNR performance of each scheme by evaluating (3.35)–(3.38) as follows: We first compute (3.36) using both the expected distortion $E[D]$, given by (3.35), which is calculated by simulation, and the weight function $w(\gamma_b)$, given by (3.37). Next, with the distance ratio of α obtained from (3.36), we evaluate the PSNR using (3.35) and (3.38) over a range of average SNRs given by (3.37).

In Fig. 3.7 (a)–(c), for various SNR ranges of weight function in (3.37), the PSNR performance of the proposed superposition MIMO coding is compared with those of the pure Alamouti and the pure spatial multiplexing schemes. From Fig. 3.7 (a)–(c), it is seen that, on the average, the proposed scheme has a channel SNR gain of about 2dB, and a PSNR gain of about 1dB, compared to the two pure MIMO schemes. This is because the Alamouti scheme outperforms spatial multiplexing for the basic subconstellation supporting a low data rate, and spatial multiplexing outperforms the Alamouti scheme for the secondary subconstellation supporting a high data rate, as indicated by the results in Section 3.2. From 3.7 (a)–(c), it is also seen that as a SNR range covers lower SNRs, the performance at low SNRs is improved, but the performance at high SNRs is

degraded. This is because the optimal distance ratio satisfying (3.36) becomes higher to protect the basic subconstellation much more than the secondary subconstellation, as the SNR range of our interest is focused on lower SNRs.

3.5 Conclusion

In this chapter, we proposed superposition MIMO coding for the transmission of unequally important sources in a point-to-multipoint system. We first analyzed the tradeoff between Alamouti coding and spatial multiplexing having the same maximum data rate in terms of the average bit error rate. As a way to compare both schemes fairly in broadcast system, the maximum data rates of both were set to be equal. The results showed that for a given target bit error rate, the Alamouti coding is preferable to spatial multiplexing for a low data rate, and the spatial multiplexing is preferable for a high data rate. In layered sources, the important component and the less important component do not necessarily have the same data rate. In particular, we considered the broadcast of layered sources where the more important component consists of a smaller number of bits than the less important component (a typical example is scalable video). As a result, in the proposed scheme, two different MIMO techniques are hierarchically combined such that important data is Alamouti encoded, less important data is spatially multiplexed, and then two differently encoded data symbols are superposed. A successive decoding algorithm for the proposed scheme was provided, and for a sufficiently large distance ratio of hierarchical modulation, this decoding was shown to have performance nearly identical to that of the complex optimal ML decoding. Performance evaluation in a broadcasting scenario showed that the proposed superposition MIMO coding significantly outperforms the pure Alamouti coding and pure spatial multiplexing scheme.

Lastly, we note that the tradeoff between the Alamouti coding and spatial multiplexing analyzed in this chapter also can be used for the design of the time-multiplexing broadcasting transmission of layered sources. In other words, when unequally important packets with different data rates (i.e., different alphabet sizes) are transmitted in a time-multiplexed way, the optimal MIMO approach for each packet can be determined based on our analysis.

3.6 Acknowledgements

The text of this chapter, in full, is a reprint of "Superposition MIMO Coding for the Broadcast of Layered Sources," which has been submitted for publication in *IEEE Transactions on Communications*. The dissertation author was the primary investigator of this paper.

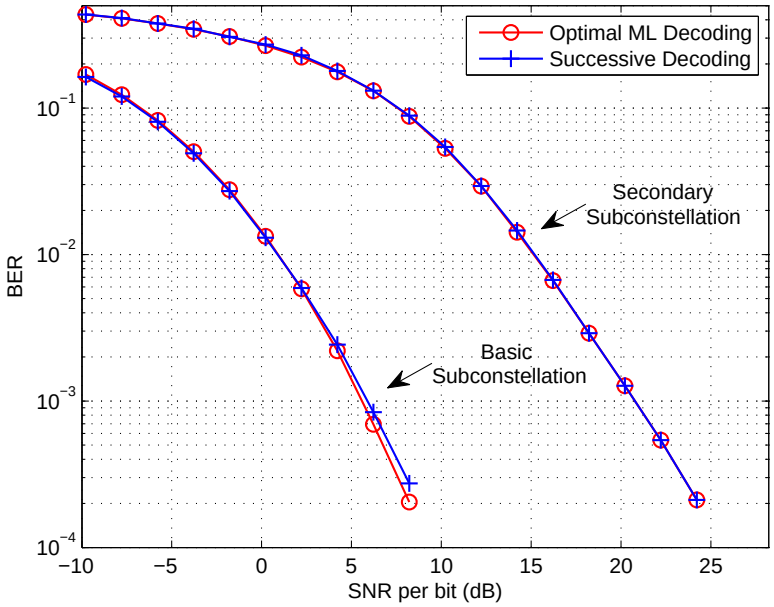


Fig. 3.6 (a) Hierarchical 4/16 QAM with a distance ratio of 2.0.

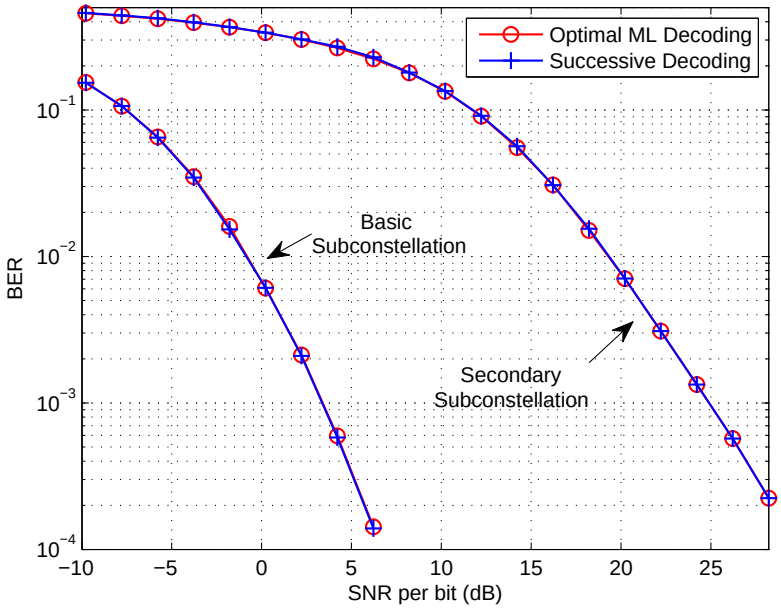


Fig. 3.6 (b) Hierarchical 4/16 QAM with a distance ratio of 4.0.

Figure 3.6: The BER performance of the optimal ML decoding and successive decoding of the proposed scheme in 2×2 MIMO systems. For successive decoding, ML decoding is performed for spatial demultiplexing of the secondary subconstellation.

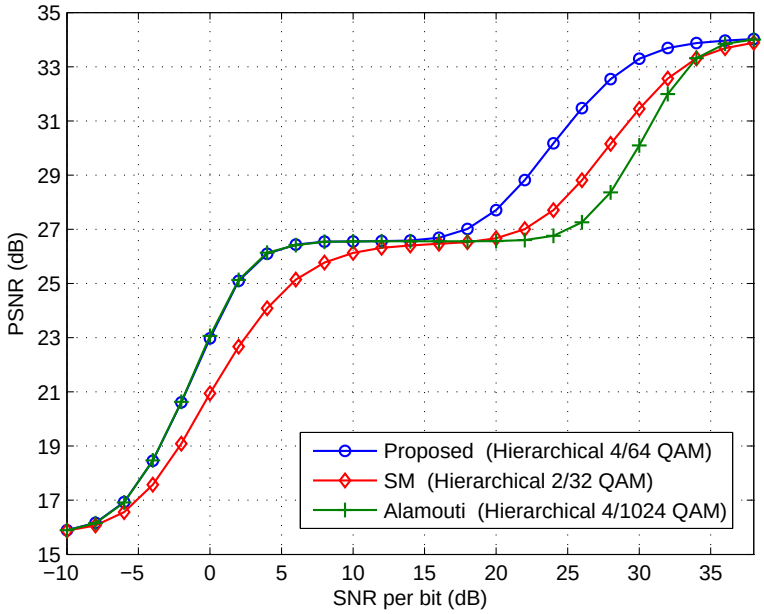


Fig. 3.7 (a) SNR range of weight function: 0–38 dB (i.e., $\gamma_b^A = 0$ dB and $\gamma_b^B = 38$ dB in (3.37)).

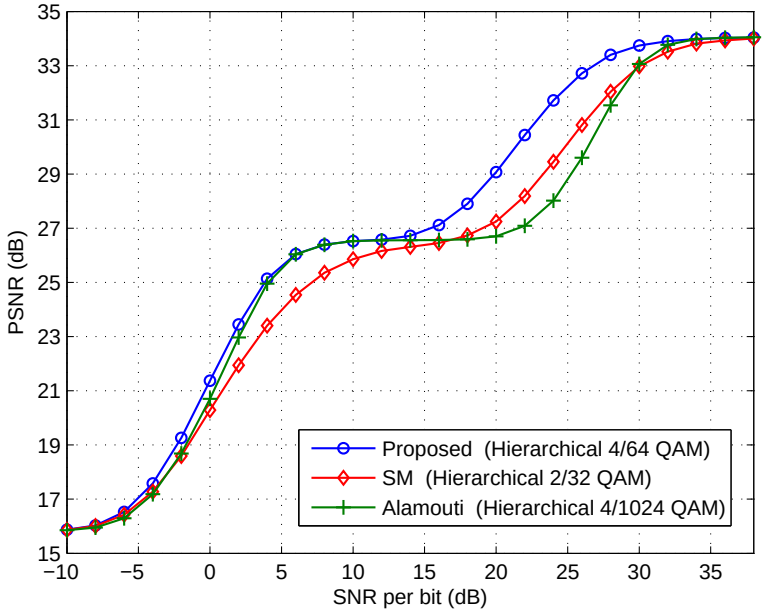


Fig. 3.7 (b) SNR range of weight function: 4–38 dB (i.e., $\gamma_b^A = 4$ dB and $\gamma_b^B = 38$ dB in (3.37)).

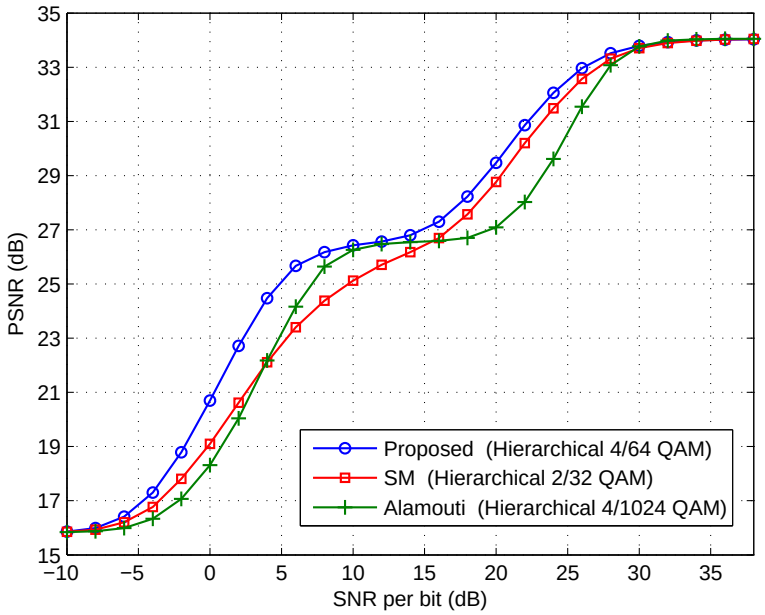


Fig. 3.7 (c) SNR range of weight function: 8–38 dB (i.e., $\gamma_b^A = 8$ dB and $\gamma_b^B = 38$ dB in (3.37)).

Figure 3.7: The PSNR performance of the proposed superposition MIMO, pure Alamouti, and pure spatial multiplexing schemes in 2×2 MIMO systems. For the proposed scheme, successive decoding with ML decoding for spatial demultiplexing is used.

Chapter 4

Performance Analysis of n-Channel Symmetric FEC-Based Multiple Description Coding for OFDM Networks

4.1 Introduction

In recent years, there has been significant interest in the transmission of multimedia services over wireless channels, and it has invoked intense research for cross-layer optimization design [56] [57], which is particularly important for the transmission over mobile radio channels exhibiting time-variant channel-quality fluctuations.

Progressive image or scalable video coders [10]– [14] employ a progressive mode of transmission such that as more bits are received, the source can be reconstructed with better quality at the receiver. Such coders are usually sensitive to channel impairments. Early studies [35] [36] considered the transmission of a progressively compressed bitstream using rate-compatible punctured convolutional codes. However, channel coding becomes less effective in a slow fading channel where prolonged deep fades often result in the erasure of the whole packet [36].

Multiple description source coding has recently emerged as an attractive framework for robust multimedia transmission over packet erasure channels [39]. The basic idea is to generate multiple descriptions of the source such that each independently de-

scribes the source with a certain fidelity. When more than one description is available at the decoder, they can be synergistically combined to enhance the quality [40]. Due to the individually decodable nature of the descriptions, the loss of some of them will not jeopardize the decoding of correctly received descriptions. Earlier studies of multiple description source coding concentrated on information-theoretic bounds for specific input source models [58]– [60]. Recently, practical implementation of multiple description source coding has received attention [39]. For progressive bitstream under deep fades in a mobile channel, n-channel symmetric FEC-based multiple description coding [61]– [64] is employed [61] [65]– [67]. In this scheme, contiguous information symbols from the progressive bitstreams are spread across multiple packets (i.e., descriptions) instead of being packed in the same packets [35] [68]. The information symbols are then protected against channel errors using systematic maximum distance separable (MDS) erasure codes, and the level of error protection depends on the relative importance of the information symbols. This FEC-based multiple description coding has become popular [66] [67] [69] [70] since it is flexible in generating arbitrary numbers of descriptions from a progressive bitstream.

Orthogonal frequency division multiplexing (OFDM) is being considered in a large number of current applications. OFDM differentiates itself from single carrier systems in many ways such as the robustness against frequency-selective fading. The use of FEC-based multiple description coding over OFDM systems was considered in [69] [71] for progressive images and scalable video, in a frequency-selective slow Rayleigh fading channel. It was demonstrated that the multiple description coding in [69] using the SPIHT image coder provides superior performance over the approach in [72] which does not use multiple description coding over OFDM systems. The FEC-based multiple description coding in OFDM systems since then has been of much interest [73]– [76].

In this chapter, we mathematically analyze the performance of n-channel symmetric FEC-based multiple description coding for a progressive mode of transmission over OFDM networks in a frequency-selective slowly Rayleigh fading channel. Based on this analysis, the performance of the system can be numerically evaluated, and system parameters such as channel code rates can be determined without a Monte-Carlo simulation. The rest of this chapter is organized as follows. In Section 4.2, we provide some technical preliminaries, and the system model is described in Section 4.3. The analysis of the throughput and distortion performance is derived in Section 4.4. In Section 4.5,

numerical results and discussions are provided, and we conclude our work in Section 4.6.

4.2 Preliminaries

4.2.1 Orthogonal Frequency Division Multiplexing (OFDM)

OFDM splits a high-rate data stream into a number of lower-rate data streams that are transmitted over orthogonal subcarriers. Based on the frequency-selectivity of a channel, frequency diversity can be exploited by adding redundancy across the subcarriers. Generally, the maximum achievable diversity gain of an OFDM system is proportional to the number of independent fading channels, N. Note that $N = 1$ corresponds to a flat-fading environment, while $N > 1$ corresponds to a frequency-selective environment.

4.2.2 FEC-Based Multiple Description Coding

We provide a brief overview of the FEC-based multiple description coding [61] [65] [66] in which MDS erasure codes are used. An (n, k) erasure code with minimum distance d_{min} refers to a construction where k information symbols are encoded into n channel symbols such that the reception of any $(n - d_{min} + 1)$ of the n channel symbols enables k information symbols to be recovered. Channel codes with $d_{min} = n - k + 1$ are referred to as MDS codes, which implies that the k information symbols can be recovered if any k channel symbols are correctly received. Reed-Solomon (RS) codes have this property. Fig. 4.1 shows a typical progressive bitstream, in which the source can be reconstructed progressively from the prefixes of the bitstream, while an error generally renders the subsequent bits useless. Fig. 4.2 illustrates a practical realization of n-channel symmetric FEC-based multiple description coding [61] [62] [65] by applying unequal FEC to different parts of a progressive bitstream. A progressive bitstream from a source encoder is converted into multiple descriptions in which contiguous information symbols are spread across the multiple descriptions. The information symbols are protected against channel errors using systematic $(n = 4, k)$ MDS codes, with the level of protection depending on the relative importance of the information symbols. If any g out of n descriptions are received, those codewords with information symbols less than or equal to g can be decoded. As a result, decoding is guaranteed at least up to distortion $D(R_g)$ which is the distortion achieved with R_g information symbols. For example, in

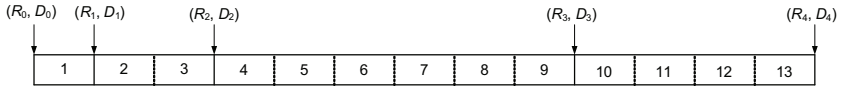


Figure 4.1: A progressive description from the source coder partitioned into five quality levels of rate R_g and distortion $D(R_g) = D_g$ $(g = 0, 1, \cdots, 4)$.

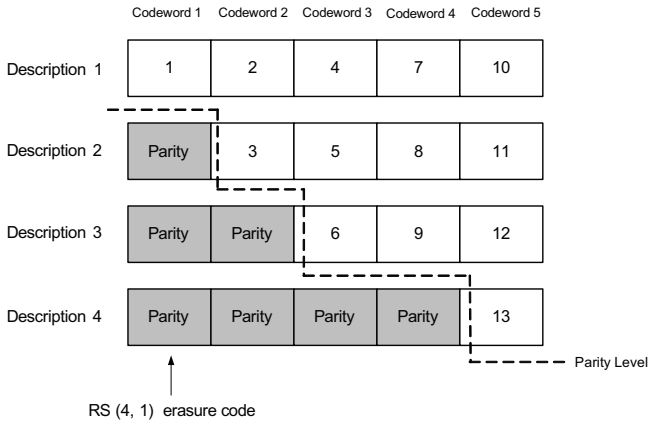


Figure 4.2: n-channel symmetric FEC-based multiple description coding technique for a progressive bitstream.

codeword 1, the erasure of any three descriptions still allows the decoder to reconstruct information symbol 1 and achieve a delivered quality equal to $D(R_1)$.

4.3 System Model

We briefly describe the system model considered in [69] [71]. The total number of subcarriers of the OFDM system is denoted by N_t. A frequency-selective fading environment has N independent fading channels and each of the N channels consists of M highly correlated subcarriers ($N_t = NM$). Let $s[n, u, v]$ be the vth input modulated symbol of a description at the uth subcarrier in the nth channel. Let V denote the block length of a description in terms of the modulated symbols. At the receiver, the output signal $r[n, u, v]$ can be expressed as

$$r[n, u, v] = \alpha[n, u, v]s[n, u, v] + w[n, u, v], \quad \text{for } 1 \le n \le N,\ 1 \le u \le M,\ 1 \le v \le V \tag{4.1}$$

where $w[n,u,v]$ is a zero-mean complex Gaussian random variable. It is assumed that $w[n,u,v]$ is independent for different n's, u's and v's. Due to the highly correlated nature of the subcarriers within a channel, we have

$$\alpha[n,u,v] \approx \alpha[n,v] \tag{4.2}$$

where α is a zero-mean complex valued Gaussian random variable with Rayleigh-distributed envelope. This corresponds to the widely used block fading channel model [77]– [80] in the frequency domain. In the time domain, the channel is assumed to experience slow Rayleigh fading (i.e., the channel symbol duration is much smaller than the coherence time) such that the fading coefficients are nearly constant over a description, and hence we have $\alpha[n,u,v] \approx \alpha[n]$.

Fig. 4.3 shows n-channel symmetric FEC-based multiple description coding for a progressive bitstream transmission over OFDM systems. The bitstream is converted into NM descriptions using the FEC-based multiple description encoder [61] [62] [65]. Due to the assumption of slow Rayleigh fading, channel coding plus interleaving in the time domain is not considered [81] [82]. Each RS code symbol consists of eight bits (four QPSK symbols). Cyclic redundancy check (CRC) codes are appended to each description for error detection. The NM independent descriptions are mapped to $N_t = NM$ subcarriers and transmitted through the OFDM system. The description size in terms of code symbols is denoted by L. Since each RS code symbol contains four QPSK symbols, the description size in terms of modulated symbols is $V = 4L$.

4.4 Throughput and Distortion Analysis

We first derive the average throughput and distortion in terms of the probability of n description errors in an OFDM frame ($0 \le n \le N_t$). N_t is the total number of descriptions in an OFDM frame. Let c_l denote the number of RS code symbols assigned to information symbols for codeword l ($1 \le l \le L$). As the compressed bitstream from an image/video encoder has different sensitivities toward channel errors, the overall system performance is improved by employing unequal error protection (UEP). Error protection decreases for the codewords on the right (i.e., $c_1 \le c_2 \le \cdots \le c_L$) [69], as shown in Fig. 4.4.

Let $E[R]$, $E[D]$, and $P_f(n)$ denote the average throughput, average distortion, and the probability of n description errors in a frame. $D(x)$ denotes the operational rate-

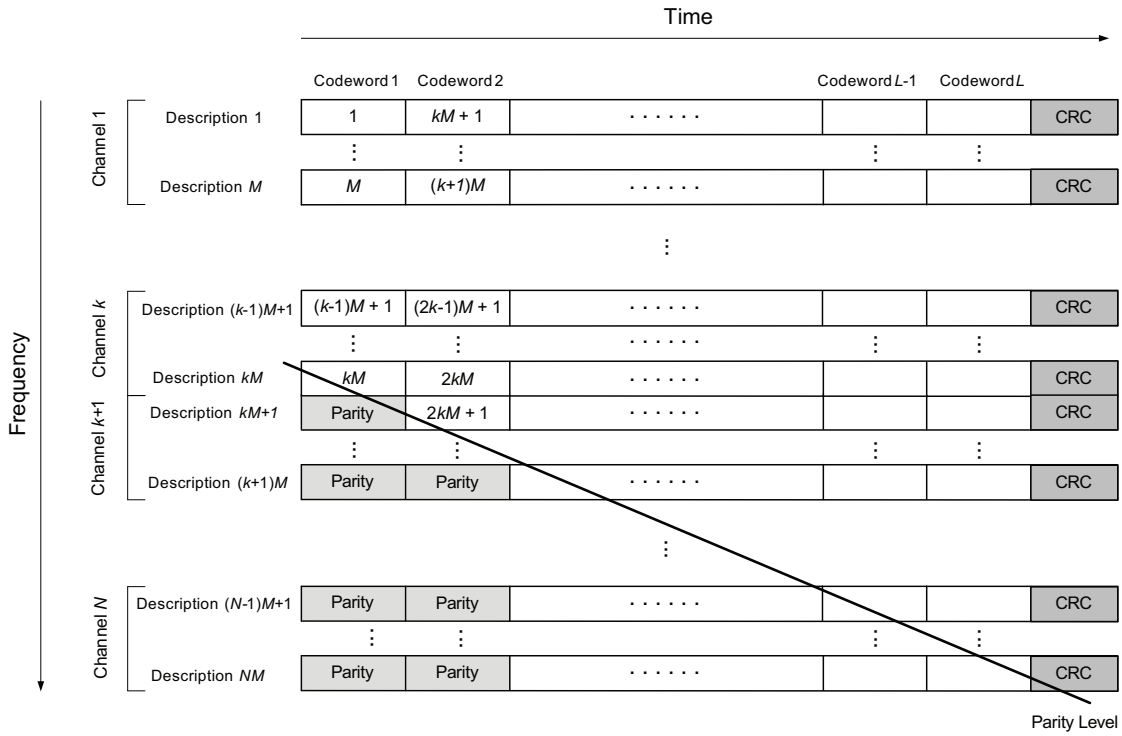


Figure 4.3: n-channel symmetric FEC-based multiple description coding for a progressive bitstream transmission over an OFDM system.

distortion curve, and $D(0)$ is the distortion when the decoder reconstructs the source without any transmitted information. Since an RS code is used for each codeword, k information symbols can be recovered if any k channel symbols are correctly received. It can be shown that the average throughput in terms of the number of bits is

$$E[R] = \sum_{l=1}^{L} \left(8c_l \sum_{n=0}^{N_t - c_l} P_f(n) \right). \tag{4.3}$$

It can be shown that the average distortion is given by

$$E[D] = D(0)Pr(0) + \sum_{l=1}^{L} \left\{ D\left(\sum_{k=1}^{l} 8c_k \right) Pr\left(\sum_{k=1}^{l} 8c_k \right) \right\} \tag{4.4}$$

where $Pr(R)$ is the probability that the throughput is R bits, and $Pr(0)$ is the probability that no information bits are successfully decoded at the receiver. It can be shown that

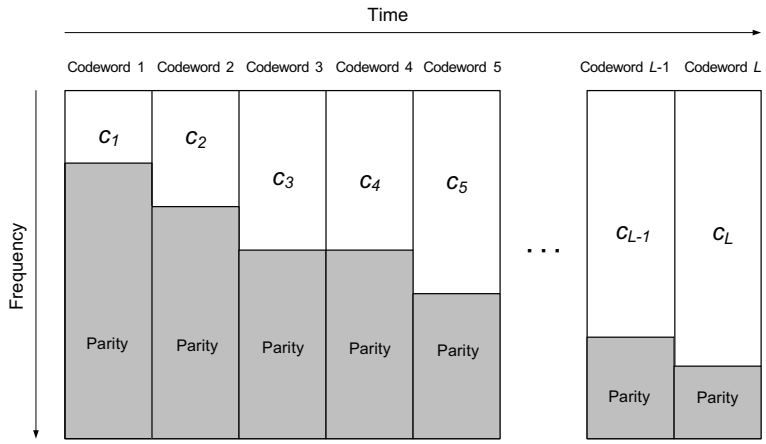


Figure 4.4: UEP techniques employing decreasing level of error protection for the codewords ($c_1 \leq c_2 \leq \cdots \leq c_L$).

$E[D]$ is expressed as

$$\begin{aligned} E[D] &= D(0) \sum_{n=N_t-c_1+1}^{N_t} P_f(n) + D\left(\sum_{k=1}^{1} 8c_k\right) \sum_{n=N_t-c_2+1}^{N_t-c_1} P_f(n) \\ &+ D\left(\sum_{k=1}^{2} 8c_k\right) \sum_{n=N_t-c_3+1}^{N_t-c_2} P_f(n) \\ &+ \cdots + D\left(\sum_{k=1}^{L-1} 8c_k\right) \sum_{n=N_t-c_L+1}^{N_t-c_{L-1}} P_f(n) + D\left(\sum_{k=1}^{L} 8c_k\right) \sum_{n=0}^{N_t-c_L} P_f(n) \\ &= D(0) \sum_{n=N_t-c_1+1}^{N_t} P_f(n) + \sum_{l=1}^{L} \left\{ D\left(\sum_{k=1}^{l} 8c_k\right) \sum_{n=N_t-c_{l+1}+1}^{N_t-c_l} P_f(n) \right\} \end{aligned} \tag{4.5}$$

where $\sum_{i=I_1}^{I_2} f(i) \triangleq 0$ for $I_2 < I_1$, and $c_{L+1} \triangleq N_t + 1$.

We next derive the probability of n description errors in a group of M subcarriers which have the same fading coefficients. Since we assume slow Rayleigh fading, the conditional probability of a description error for a given Rayleigh fading coefficient h, denoted by $P_d(e|h)$, is

$$P_d(e|h) = 1 - (1 - P_{cs}(e|h))^L \tag{4.6}$$

where $P_{cs}(e|h)$ is the conditional probability of an RS code symbol error given h. It equals

$$P_{cs}(e|h) = 1 - (1 - P_s(e|h))^4 = 1 - (1 - Q\left(h\sqrt{\gamma_s}\right))^8 \tag{4.7}$$

where $P_s(e|h) = 2Q(h\sqrt{\gamma_s}) - Q^2(h\sqrt{\gamma_s})$ is the conditional probability of a modulated QPSK symbol error for a given h, and γ_s is the signal-to-noise ratio (SNR) per modulated QPSK symbol. We define a group to be M correlated subcarriers with the same fading coefficients. The conditional probability of n description errors in a group ($0 \leq n \leq M$) for a given h, $P_g(n|h)$, is given by

$$P_g(n|h) = \binom{M}{n} P_d^n(e|h) \left(1 - P_d(e|h)\right)^{M-n}. \tag{4.8}$$

From (4.6) and (4.7), for $0 \leq n \leq M$, $P_g(n|h)$ can be expressed as

$$P_g(n|h) = \binom{M}{n} \left\{1 - \left(1 - Q\left(h\sqrt{\gamma_s}\right)\right)^{8L}\right\}^n \left(1 - Q\left(h\sqrt{\gamma_s}\right)\right)^{8L(M-n)}. \tag{4.9}$$

The probability of n description errors in a group, $P_g(n)$, can be calculated by taking the expectation of $P_g(n|h)$ with regard to h representing a Rayleigh probability distribution:

$$P_R(h) = \frac{2h}{\Omega} \exp\left(-\frac{h^2}{\Omega}\right), \quad h \geq 0 \tag{4.10}$$

where Ω is the second moment of h. From (4.9) and (4.10), for $0 \leq n \leq M$, $P_g(n)$ is given by

$$\begin{aligned}
P_g(n) &= \int_0^\infty \binom{M}{n} \left\{1 - \left(1 - Q\left(h\sqrt{\gamma_s}\right)\right)^{8L}\right\}^n \left(1 - Q\left(h\sqrt{\gamma_s}\right)\right)^{8L(M-n)} P_R(h)dh \\
&= \int_0^\infty \binom{M}{n} \sum_{p=0}^{n} \left[\binom{n}{p} (-1)^p \left(1 - Q\left(h\sqrt{\gamma_s}\right)\right)^{8L(p+M-n)}\right] P_R(h)dh \\
&= \binom{M}{n} \sum_{p=0}^{n} \left[\binom{n}{p} (-1)^p \sum_{q=0}^{8L(p+M-n)} \left\{(-1)^q \binom{8L(p+M-n)}{q}\right.\right. \\
&\quad \left.\left. \times \int_0^\infty Q^q\left(h\sqrt{\gamma_s}\right) P_R(h)dh\right\}\right].
\end{aligned} \tag{4.11}$$

Note that for an integer of $q \geq 5$, there is no closed-form expression for the integral in (4.11) [83]. To avoid numerical integration in $P_g(n)$, instead, we can use exponential-type upper and lower bounds on $Q(x)$ [84]. From [84, eqs. (8), (9), and (26)], the upper and lower bounds on $Q(x)$ for $x \geq 0$, denoted by $f_u(x)$ and $f_l(x)$, respectively, are given by

$$f_u(x) = \sum_{i=1}^{B} a_i \exp\left(-b_i x^2\right) \quad \text{and} \quad f_l(x) = \sum_{i=2}^{B} a_i \exp\left(-b_{i-1} x^2\right) \tag{4.12}$$

where $a_i = (\theta_i - \theta_{i-1})/\pi$, $b_i = \csc^2\theta_i/2$, and $\theta_0, \theta_1, \cdots, \theta_{B-1}$ are arbitrary values satisfying $0 = \theta_0 \leq \theta_1 \leq \cdots \leq \theta_{B-1} \leq \theta_B = \pi/2$ [84, eq. (8)]. By increasing B, the bounds $f_u(x)$ and $f_l(x)$ converge to the exact $Q(x)$ [84]. Using (4.12), the upper bound of the integration in (4.11) is

$$\begin{aligned}\int_0^\infty Q^q\left(h\sqrt{\gamma_s}\right) P_R(h)dh &\leq \int_0^\infty f_u^q\left(h\sqrt{\gamma_s}\right) P_R(h)dh \\ &= \int_0^\infty \left(\sum_{i=1}^{B} a_i \exp\left(-b_i\gamma_s h^2\right)\right)^q \frac{2h}{\Omega}\exp\left(-\frac{h^2}{\Omega}\right) dh. \quad (4.13)\end{aligned}$$

We have

$$\begin{aligned}\left(\sum_{i=1}^{B} a_i \exp\left(-b_i\gamma_s h^2\right)\right)^q = \sum_{k_1,k_2,\cdots,k_B} \Bigg\{ &\binom{q}{k_1, k_2, \cdots, k_B} a_1^{k_1} a_2^{k_2} \cdots a_B^{k_B} \\ &\times \exp\left(-\left(b_1k_1 + b_2k_2 + \cdots + b_Bk_B\right)\gamma_s h^2\right) \Bigg\} \quad (4.14)\end{aligned}$$

where the summation is taken over all sequences of nonnegative indices $k_1, k_2, \cdots, k_B$ such that the sum of all $k_1, k_2, \cdots, k_B$ is equal to q, and $\binom{q}{k_1,k_2,\cdots,k_B} = \frac{q!}{k_1!k_2!\cdots k_B!}$ are the multinomial coefficients. It can be readily shown that

$$\int_0^\infty \exp\left(-\alpha h^2\right) \frac{2h}{\Omega}\exp\left(-\frac{h^2}{\Omega}\right) dh = \frac{1}{\alpha\Omega + 1}. \quad (4.15)$$

From (4.14) and (4.15), (4.13) can be rewritten as

$$\begin{aligned}\int_0^\infty Q^q\left(h\sqrt{\gamma_s}\right) P_R(h)dh \leq \sum_{k_1,k_2,\cdots,k_B} \Bigg\{ &\binom{q}{k_1, k_2, \cdots, k_B} \\ &\times \frac{a_1^{k_1} a_2^{k_2} \cdots a_B^{k_B}}{(b_1k_1 + b_2k_2 + \cdots + b_Bk_B)\gamma_s\Omega + 1} \Bigg\}. \quad (4.16)\end{aligned}$$

Therefore, $P_g(n)$, given by (4.11), is upper bounded as

$$\begin{aligned}P_g(n) \leq \binom{M}{n} \sum_{p=0}^{n} \Bigg[\binom{n}{p}(-1)^p \sum_{q=0}^{8L(p+M-n)} \Bigg\{ &(-1)^q \binom{8L(p+M-n)}{q} \\ \times \sum_{k_1,k_2,\cdots,k_B} &\left(\binom{q}{k_1, k_2, \cdots, k_B} \frac{a_1^{k_1} a_2^{k_2} \cdots a_B^{k_B}}{(b_1k_1 + b_2k_2 + \cdots + b_Bk_B)\gamma_s\Omega + 1} \right) \Bigg\} \Bigg]. \quad (4.17)\end{aligned}$$

In a similar way, from (4.12), it can be shown that $P_g(n)$, given by (4.11), is lower bounded as

$$P_g(n) \geq \binom{M}{n} \sum_{p=0}^{n} \left[\binom{n}{p} (-1)^p \sum_{q=0}^{8L(p+M-n)} \left\{ (-1)^q \binom{8L(p+M-n)}{q} \right.\right.$$
$$\left.\left. \times \sum_{k_1,k_2,\cdots,k_{B-1}} \left(\binom{q}{k_1, k_2, \cdots, k_{B-1}} \frac{a_2^{k_1} a_3^{k_2} \cdots a_B^{k_{B-1}}}{(b_1 k_1 + b_2 k_2 + \cdots + b_{B-1} k_{B-1}) \gamma_s \Omega + 1} \right) \right\} \right]. \tag{4.18}$$

We next derive the probability of n description errors in an OFDM frame. An erroneous group is defined as a group which has at least one description error. For n description errors in a frame ($0 \leq n \leq N_t = NM$), the number of erroneous groups, m, is in the range of

$$\left\lceil \frac{n}{M} \right\rceil \leq m \leq \min(N, n). \tag{4.19}$$

To see this, note that

i) The total number of descriptions in all erroneous groups, mM, should be larger than or equal to the number of description errors in a frame, that is $mM \geq n$. Since m is an integer, the infimum of m is given by $\lceil n/M \rceil$.

ii) The number of erroneous groups, m, should be smaller than or equal to both the number of description errors in a frame, n, and the number of groups in a frame, N. Hence, the supremum of m is given by $\min(N, n)$.

Next, we will show that for n description errors in a frame ($0 \leq n \leq NM$), the probability of m erroneous groups, $P_f(n, m)$, is given by

$$P_f(n,m) = \begin{cases} \sum_{k_m=\lceil \frac{n}{m} \rceil}^{\min(M,\ n-(m-1))} \sum_{k_{m-1}=\lceil \frac{n-k_m}{m-1} \rceil}^{\min(k_m,\ n-k_m-(m-2))} \cdots \sum_{k_2=\left\lceil \frac{n-k_m-k_{m-1}\cdots-k_3}{2} \right\rceil}^{\min(k_3,\ n-k_m-k_{m-1}\cdots-k_3-1)} \\ \quad P_g(k_m) P_g(k_{m-1}) \cdots P_g(k_2) P_g(n - \sum_{i=2}^{m} k_i) P_g^{N-m}(0), \quad \text{for} \quad m \geq 2 \\ P_g^m(n) P_g^{N-m}(0), \quad \text{for} \quad m = 0, 1 \end{cases} \tag{4.20}$$

where $\lceil x \rceil$ denotes the smallest integer which is greater than or equal to x, $P_g(l)$ is given by (4.11), and it is assumed that for $i = 1, 2, \cdots, m$, the ith group from the top in a

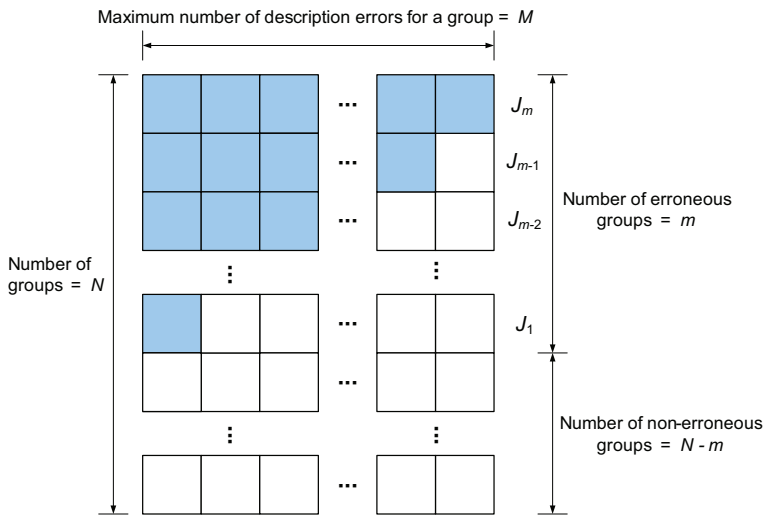


Figure 4.5: An OFDM frame with m erroneous groups.

frame has the ith largest number of description errors among m erroneous groups. After proving (4.20), we will generalize (4.20) without the assumption.

Proof: Note that the maximum number of description errors which one group can have is M. Let J_i denote the $m+1-i$th largest number of description errors which a group among m erroneous groups has ($i = 1, 2, \cdots, m$). That is,

$$M \geq J_m \geq J_{m-1} \geq \cdots \geq J_1 \geq 1 \quad \text{and} \quad J_m + J_{m-1} + \cdots + J_1 = n. \tag{4.21}$$

By the assumption below (4.20), the ith group from the top in a frame has J_{m+1-i} description errors ($i = 1, 2, \cdots, m$). Fig. 4.5 shows an OFDM frame with m erroneous groups, where each subcarrier (i.e., description) is denoted by a small square box. Dark and bright boxes indicate erroneous and non-erroneous descriptions, respectively.

1) The case where $m \geq 2$: First, we will prove (4.20) for $m \geq 2$ by induction on the number of erroneous groups in a frame. Consider two erroneous groups (i.e., $m = 2$), shown in Fig. 4.6. If we let $m = 2$ in (4.21), we have

$$M \geq J_2 \geq J_1 \geq 1 \quad \text{and} \quad J_2 + J_1 = n. \tag{4.22}$$

From (4.22), it follows that $J_2 \geq J_1 = n - J_2$ or $J_2 \geq n/2$. Since J_2 is an integer, we have

$$J_2 \geq \left\lceil \frac{n}{2} \right\rceil. \tag{4.23}$$

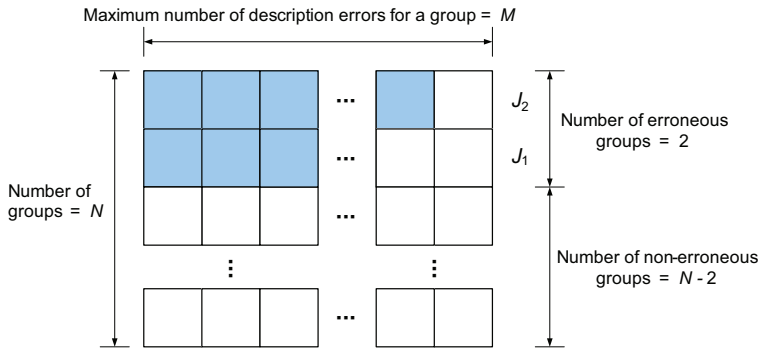


Figure 4.6: The case where there are two erroneous groups (i.e., $m = 2$) for n description errors in a frame.

From (4.22), it follows that

$$J_1 = n - J_2 \geq 1 \;\Leftrightarrow\; J_2 \leq n - 1. \tag{4.24}$$

From (4.22)–(4.24), we have

$$\left\lceil \frac{n}{2} \right\rceil \leq J_2 \leq \min(M, n-1). \tag{4.25}$$

From (4.25) and Fig. 4.6, it follows that

$$\begin{aligned} P_f(n,2) &= \sum_{J_2=\lceil \frac{n}{2} \rceil}^{\min(M,n-1)} P_g(J_2) P_g(J_1) P_g^{N-2}(0) \\ &= \sum_{J_2=\lceil \frac{n}{2} \rceil}^{\min(M,n-1)} P_g(J_2) P_g(n-J_2) P_g^{N-2}(0) \end{aligned} \tag{4.26}$$

where the second equality follows from (4.22). If we let $m = 2$ in (4.20), we have

$$P_f(n,2) = \sum_{k_2=\lceil \frac{n}{2} \rceil}^{\min(M,n-1)} P_g(k_2) P_g(n-k_2) P_g^{N-2}(0). \tag{4.27}$$

It is seen that (4.27) is identical to (4.26).

Consider the case where there are r erroneous groups (r is some integer in the range of $2 \leq r \leq N-1$). If we let $m = r$ in (4.21), we have

$$M \geq J_r \geq J_{r-1} \geq \cdots \geq J_1 \geq 1 \;\text{ and }\; J_r + J_{r-1} + \cdots + J_1 = n. \tag{4.28}$$

By the assumption below (4.20), the ith group from the top in a frame has J_{r+1-i} description errors ($i = 1, 2, \cdots, r$), which is shown in Fig. 4.7. Suppose that (4.20) holds

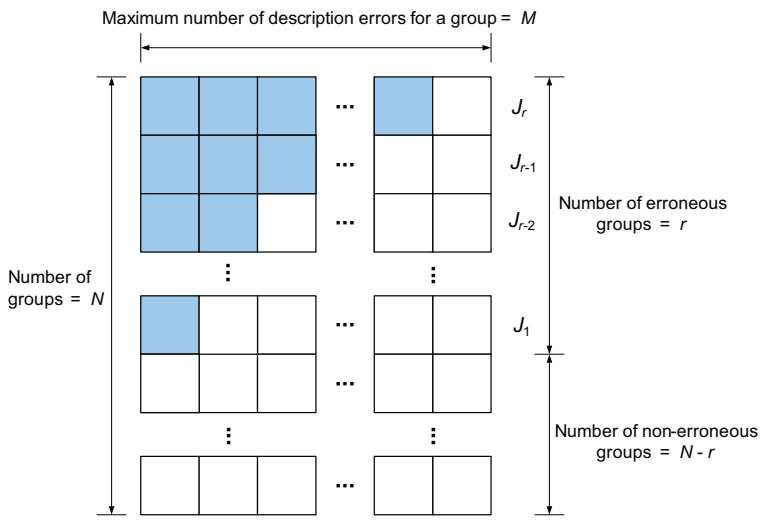


Figure 4.7: The case where there are r erroneous groups for n description errors in a frame.

for this case. That is,

$$P_f(n,r) = \sum_{k_r=\lceil \frac{n}{r} \rceil}^{\min(M,\ n-(r-1))} \sum_{k_{r-1}=\lceil \frac{n-k_r}{r-1} \rceil}^{\min(k_r,\ n-k_r-(r-2))} \cdots \sum_{k_2=\left\lceil \frac{n-k_r-k_{r-1}\cdots-k_3}{2} \right\rceil}^{\min(k_3,\ n-k_r-k_{r-1}\cdots-k_3-1)} P_g(k_r)P_g(k_{r-1})\cdots P_g(k_2)P_g(n-\sum_{i=2}^{r} k_i)P_g^{N-r}(0). \tag{4.29}$$

Consider the case where there are $r+1$ erroneous groups (r is some integer in the range of $2 \le r \le N-1$). If we let $m = r+1$ in (4.21), we have

$$M \ge J_{r+1} \ge J_r \ge \cdots \ge J_1 \ge 1 \quad \text{and} \quad J_{r+1} + J_r + \cdots + J_1 = n. \tag{4.30}$$

By the assumption below (4.20), the ith group from the top in a frame has J_{r+2-i} description errors ($i = 1, 2, \cdots, r+1$), which is shown in Fig. 4.8.

The range of the largest number of description errors which a group among $r+1$ erroneous groups can have is given by

$$\left\lceil \frac{n}{r+1} \right\rceil \le J_{r+1} \le \min(M, n-r). \tag{4.31}$$

To see this, note that

i) Since J_{r+1} is the largest number of description errors in an erroneous group, J_{r+1} should be larger than or equal to $n/(r+1)$, which is the average number

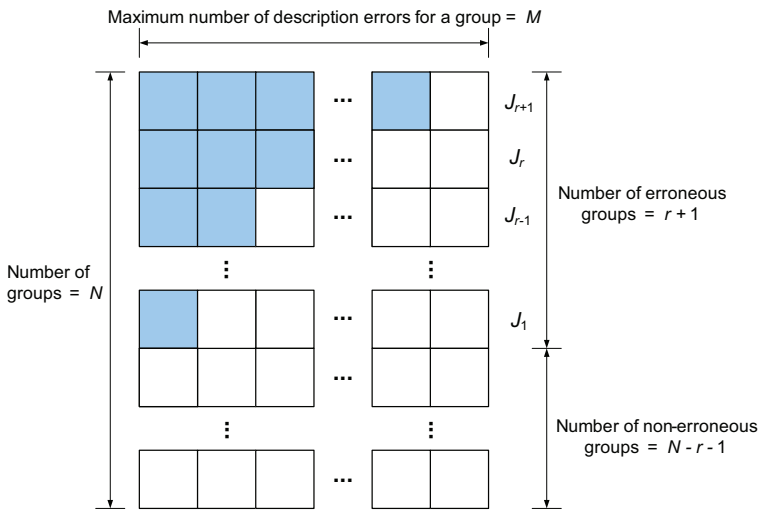


Figure 4.8: The case where there are $r+1$ erroneous groups for n description errors in a frame.

of description errors in an erroneous group. Since J_{r+1} is an integer, we have $J_{r+1} \geq \lceil n/(r+1) \rceil$.

ii) J_{r+1} should be less than or equal to $n-r$ since each of the other r erroneous groups should have at least one description error.

iii) From (4.30), we have $J_{r+1} \leq M$.

From (4.31), the probability of $r+1$ erroneous groups in a frame can be expressed as

$$P_f(n, r+1) = \sum_{J_{r+1}=\lceil \frac{n}{r+1} \rceil}^{\min(M, n-r)} P_g(J_{r+1}) P_{(N-1)g}(n - J_{r+1}, r) \tag{4.32}$$

where $P_g(J_{r+1})$ is the probability of J_{r+1} description errors in the first group from the top in a frame, and $P_{(N-1)g}(n - J_{r+1}, r)$ is the probability that for $n - J_{r+1}$ description errors in a frame except the first group, there are r erroneous groups. Fig. 4.9 shows $N-1$ groups except the first group having J_{r+1} description errors. From Figs. 4.7 and 4.9, note that

i) The maximum number of description errors which one group can have is J_{r+1} instead of M.

ii) The number of groups is $N-1$ instead of N.

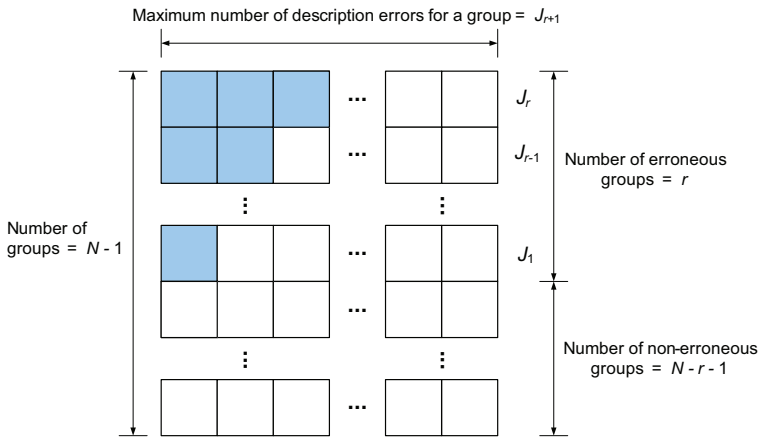


Figure 4.9: The case where there are r erroneous groups for $n - J_{r+1}$ description errors in a frame except the first group which has J_{r+1} description errors.

iii) There are $n - J_{r+1}$ description errors instead of n description errors.

Therefore, $P_{(N-1)g}(n - J_{r+1}, r)$ in (4.32) can be derived from the induction hypothesis given by (4.29) if M, N, and n are replaced by J_{r+1}, $N-1$, and $n - J_{r+1}$, respectively. That is,

$$P_{(N-1)g}(n - J_{r+1}, r) =$$
$$\sum_{k_r=\left\lceil \frac{n-J_{r+1}}{r} \right\rceil}^{\min(J_{r+1},\ n-J_{r+1}-(r-1))} \sum_{k_{r-1}=\left\lceil \frac{n-J_{r+1}-k_r}{r-1} \right\rceil}^{\min(k_r,\ n-J_{r+1}-k_r-(r-2))} \cdots \sum_{k_2=\left\lceil \frac{n-J_{r+1}-k_r-k_{r-1}\cdots-k_3}{2} \right\rceil}^{\min(k_3,\ n-J_{r+1}-k_r\cdots-k_3-1)}$$
$$P_g(k_r)P_g(k_{r-1})\cdots P_g(k_2)P_g(n - J_{r+1} - \sum_{i=2}^{r} k_i)P_g^{N-1-r}(0). \tag{4.33}$$

From (4.33), $P_f(n, r+1)$, given by (4.32), can be expressed as

$$P_f(n, r+1) =$$
$$\sum_{J_{r+1}=\left\lceil \frac{n}{r+1} \right\rceil}^{\min(M,\ n-r)} \sum_{k_r=\left\lceil \frac{n-J_{r+1}}{r} \right\rceil}^{\min(J_{r+1},\ n-J_{r+1}-(r-1))} \sum_{k_{r-1}=\left\lceil \frac{n-J_{r+1}-K_r}{r-1} \right\rceil}^{\min(k_r,\ n-J_{r+1}-k_r-(r-2))} \cdots \sum_{k_2=\left\lceil \frac{n-J_{r+1}-k_r-k_{r-1}\cdots-k_3}{2} \right\rceil}^{\min(k_3,\ n-J_{r+1}-k_r\cdots-k_3-1)}$$
$$P_g(J_{r+1})P_g(k_r)P_g(k_{r-1})\cdots P_g(k_2)P_g(n - J_{r+1} - \sum_{i=2}^{r} k_i)P_g^{N-1-r}(0). \tag{4.34}$$

If we let $m = r + 1$ in (4.20), we have

$$P_f(n, r+1) = \sum_{k_{r+1}=\lceil \frac{n}{r+1} \rceil}^{\min(M,\ n-r)} \sum_{k_r=\lceil \frac{n-k_{r+1}}{r} \rceil}^{\min(k_{r+1},\ n-k_{r+1}-(r-1))} \cdots \sum_{k_2=\lceil \frac{n-k_{r+1}-k_r\cdots-k_3}{2} \rceil}^{\min(k_3,\ n-k_{r+1}-k_r\cdots-k_3-1)} P_g(k_{r+1})P_g(k_r)\cdots P_g(k_2)P_g(n-\sum_{i=2}^{r+1} k_i)P_g^{N-r-1}(0). \tag{4.35}$$

It is seen that (4.34) and (4.35) are the same. We have proved (4.20) for $m \geq 2$ by induction.

2) The case where $m = 1$ or 0: Next, we will prove (4.20) for $m = 1$, which is the case where there is only one erroneous group. If we let $m = 1$ in (4.21), we have

$$M \geq J_1 \geq 1 \ \text{ and } \ J_1 = n. \tag{4.36}$$

By the assumption below (4.20), the first group from the top in a frame has $J_1 = n$ description errors, and thus

$$P_f(n, 1) = P_g(n)P_g^{N-1}(0), \tag{4.37}$$

which is identical to (4.20) for $m = 1$. Lastly, we will prove (4.20) for $m = 0$, which is the case where there is no erroneous group (i.e., $n = 0$). It is obvious that for this case, $P_f(0, 0) = P_g^N(0)$, which is identical to (4.20) for $m = 0$.

□

We have proved (4.20) under the assumption that for $i = 1, 2, \cdots, m$, the ith group from the top in a frame has the ith largest number of description errors. From (4.20), it follows that k_i is the ith largest number of description errors which a group has since

$$k_m \geq k_{m-1} \geq \cdots \geq k_2 \geq k_1 \tag{4.38}$$

where $k_1 \triangleq n - \sum_{i=2}^{m} k_i$. Next, we will generalize (4.20) without the assumption. We will show that the number of ways of assigning $k_m, k_{m-1}, \cdots, k_1$ satisfying (4.38) to N groups in a frame is given by

$$C(m) = \frac{N!}{\left(1 + \sum_{i=1}^{m-1} \delta(k_m - k_{m-i})\right)!} \times \frac{1}{\prod_{j=1}^{m-2} \left\{1 + (1 - \delta(k_{m-j} - k_{m+1-j})) \sum_{i=j+1}^{m-1} \delta(k_{m-j} - k_{m-i})\right\}! (N-m)!} \tag{4.39}$$

where $m \geq 1$, $\delta(x)$ is the Kronecker delta function and

$$\sum_{i=I_1}^{I_2} f(i) \triangleq 0 \quad \text{and} \quad \prod_{i=I_1}^{I_2} f(i) \triangleq 1 \quad \text{for} \quad I_1 > I_2. \tag{4.40}$$

Proof: We define the following:

i) B_i: a set of erroneous groups each of which has the same number of non-zero description errors ($i = 1, 2, \cdots, l$, where l is the number of distinct sets). We assume that for $1 \leq j < k \leq l$, the number of description errors which an erroneous group in a set B_k has is greater than that in a set B_j. As a result, for $k_m \geq k_{m-1} \geq \cdots \geq k_1$ given by (4.38), B_l is a set of erroneous groups each of which has k_m description errors.

ii) a_i: the cardinality of set B_i ($i = 1, 2, \cdots, l$). Since the total number of erroneous groups in a frame is m, we have

$$a_l + a_{l-1} + \cdots + a_1 = m. \tag{4.41}$$

From i) and ii), the number of ways of assigning $k_m, k_{m-1}, \cdots, k_1$ to N groups in a frame, $C(m)$, is given by

$$C(m) = \frac{N!}{a_l!\, a_{l-1}! \cdots a_1!\, (N-m)!}. \tag{4.42}$$

For example, when $N = 16$, $m = 7$, and $k_7 = k_6 > k_5 > k_4 = k_3 = k_2 > k_1$, we have

$$l = 4,\ a_4 = 2,\ a_3 = 1,\ a_2 = 3,\ a_1 = 1 \tag{4.43}$$

and (4.42) becomes $\frac{16!}{2!\,1!\,3!\,1!\,9!}$.

Note that for given $k_m \geq k_{m-1} \geq \cdots \geq k_1$ in (4.38), a_l can be expressed as

$$a_l = 1 + \sum_{s=1}^{m-1} \delta(k_m - k_{m-s}). \tag{4.44}$$

Fig. 4.10 (a) shows an OFDM frame with m erroneous groups, where each group is denoted by a rectangular box. Dark and bright rectangular boxes indicate erroneous and non-erroneous groups.

1) The case where $N \geq 2$: First, we will prove (4.39) for $N \geq 2$. Consider the case where there is one erroneous group (i.e., $m = 1$). From (4.41), we have $l = 1$ and $a_1 = 1$. From (4.42), the number of ways of assigning k_1 to N groups is given by

$$C(1) = \frac{N!}{1!\,(N-1)!} = N. \tag{4.45}$$

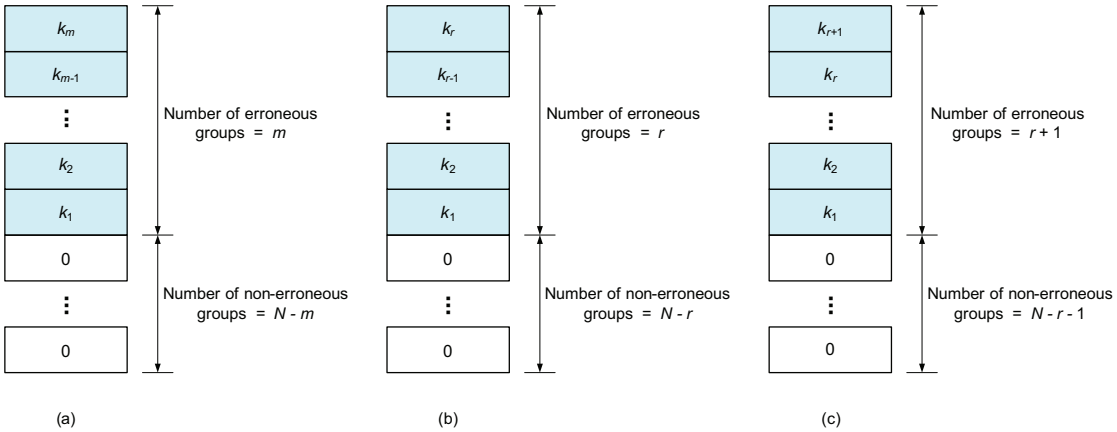


Figure 4.10: (a) An OFDM frame with m erroneous groups (b) The case where there are r erroneous groups (c) The case where there are $r + 1$ erroneous groups.

If we let $m = 1$ in (4.39), we have

$$\begin{aligned} C(1) &= \frac{N!}{\left(1 + \sum_{i=1}^{0} \delta(k_1 - k_{1-i})\right)!} \\ &\quad \times \frac{1}{\prod_{j=1}^{-1} \left\{1 + (1 - \delta(k_{1-j} - k_{2-j})) \sum_{i=j+1}^{0} \delta(k_{1-j} - k_{1-i})\right\}! \, (N-1)!} \\ &= \frac{N!}{(1+0)! \, 1 \, (N-1)!} = N \end{aligned} \tag{4.46}$$

where the second equality follows from (4.40). It is seen that (4.46) is identical to (4.45).

Consider the case where there are $m = r$ erroneous groups (r is some integer in the range of $1 \leq r \leq N - 1$), which is shown in Fig. 4.10 (b). From (4.41), we have $a_l + a_{l-1} + \cdots + a_1 = r$, and from (4.42), the number of ways of assigning $k_r, k_{r-1}, \cdots, k_1$ satisfying $k_r \geq k_{r-1} \geq \cdots \geq k_1$ to N groups is given by

$$C(r) = \frac{N!}{a_l! \, a_{l-1}! \cdots a_1! \, (N-r)!}. \tag{4.47}$$

From (4.44), for given $k_r \geq k_{r-1} \geq \cdots \geq k_1$, the cardinality of a set of erroneous groups with k_r description errors, a_l, can be expressed as

$$a_l = 1 + \sum_{s=1}^{r-1} \delta(k_r - k_{r-s}). \tag{4.48}$$

Suppose that (4.39) holds for this case. That is, $C(r)$ is given by

$$C(r) = \frac{N!}{\left(1+\sum_{i=1}^{r-1}\delta(k_r-k_{r-i})\right)!} \times \frac{1}{\prod_{j=1}^{r-2}\left\{1+(1-\delta(k_{r-j}-k_{r+1-j}))\sum_{i=j+1}^{r-1}\delta(k_{r-j}-k_{r-i})\right\}!(N-r)!}. \tag{4.49}$$

Consider the case where there are $m = r+1$ erroneous groups (r is some integer in the range of $1 \le r \le N-1$), which is the case shown in Fig. 4.10 (c). From (4.38), we have

$$k_{r+1} \ge k_r \ge \cdots \ge k_1. \tag{4.50}$$

i) For $k_{r+1} > k_r$: From (4.47), and Fig. 4.10 (b) and (c), it follows that $C(r+1)$ can be expressed as

$$\begin{aligned} C(r+1) &= \frac{N!}{a_{l+1}!\,a_l!\,a_{l-1}!\cdots a_1!\,(N-r-1)!} \\ &= \frac{N!}{1!\,a_l!\,a_{l-1}!\cdots a_1!(N-r-1)!} \end{aligned} \tag{4.51}$$

where a_{l+1} is the cardinality of B_{l+1}, a set of erroneous groups having k_{r+1} description errors, and it is obvious that $a_{l+1} = 1$ due to $k_{r+1} > k_r$. From (4.47) and (4.51), $C(r+1)$ can be expressed in terms of $C(r)$.

$$C(r+1) = \frac{(N-r)!}{(N-r-1)!}\,C(r). \tag{4.52}$$

ii) For $k_{r+1} = k_r$: From (4.47), and Fig. 4.10 (b) and (c), it follows that $C(r+1)$ can be expressed as

$$C(r+1) = \frac{N!}{(a_l+1)!\,a_{l-1}!\cdots a_1!\,(N-r-1)!}. \tag{4.53}$$

From (4.47) and (4.53), $C(r+1)$ can be expressed in terms of $C(r)$:

$$C(r+1) = \frac{a_l!}{(a_l+1)!}\cdot\frac{(N-r)!}{(N-r-1)!}\,C(r). \tag{4.54}$$

Using (4.48), $C(r+1)$, given by (4.54), can be rewritten as

$$C(r+1) = \frac{\left(1+\sum_{s=1}^{r-1}\delta(k_r-k_{r-s})\right)!}{\left(2+\sum_{s=1}^{r-1}\delta(k_r-k_{r-s})\right)!}\cdot\frac{(N-r)!}{(N-r-1)!}\,C(r). \tag{4.55}$$

From the results of i) and ii) (i.e., (4.52) and (4.55)), $C(r+1)$ can be expressed as

$$C(r+1) = \frac{\left(1+\sum_{s=1}^{r-1}\delta(k_r-k_{r-s})\right)!}{\left(1+\delta(k_{r+1}-k_r)+\sum_{s=1}^{r-1}\delta(k_r-k_{r-s})\right)!}\cdot\frac{(N-r)!}{(N-r-1)!}\,C(r) \quad (4.56)$$

for $k_{r+1}\geq k_r$. Using (4.49), $C(r+1)$, given by (4.56), can be rewritten as

$$\begin{aligned} C(r+1) = {} & \frac{N!}{\left(1+\delta(k_{r+1}-k_r)+\sum_{s=1}^{r-1}\delta(k_r-k_{r-s})\right)!} \\ & \times \frac{1}{\prod_{j=1}^{r-2}\left\{1+(1-\delta(k_{r-j}-k_{r+1-j}))\sum_{i=j+1}^{r-1}\delta(k_{r-j}-k_{r-i})\right\}!\,(N-r-1)!}. \end{aligned} \quad (4.57)$$

Let $p=j+1$ and $q=i+1$. Then, $C(r+1)$ can be expressed as

$$\begin{aligned} C(r+1) = {} & \frac{N!}{\left(1+\delta(k_{r+1}-k_r)+\sum_{s=1}^{r-1}\delta(k_r-k_{r-s})\right)!} \\ & \times \frac{1}{\prod_{p=2}^{r-1}\left\{1+(1-\delta(k_{r+1-p}-k_{r+2-p}))\sum_{q=p+1}^{r}\delta(k_{r+1-p}-k_{r+1-q})\right\}!} \\ & \times \frac{1}{(N-r-1)!}. \end{aligned} \quad (4.58)$$

If we let $m=r+1$ in (4.39), we have

$$\begin{aligned} C(r+1) = {} & \frac{N!}{(1+\sum_{i=1}^{r}\delta(k_{r+1}-k_{r+1-i}))!} \\ & \times \frac{1}{\prod_{j=1}^{r-1}\left\{1+(1-\delta(k_{r+1-j}-k_{r+2-j}))\sum_{i=j+1}^{r}\delta(k_{r+1-j}-k_{r+1-i})\right\}!} \\ & \times \frac{1}{(N-r-1)!}. \end{aligned} \quad (4.59)$$

It can be shown that the ratio of (4.59) to (4.58), γ, is given by

$$\gamma = \frac{\left(1+\delta(k_{r+1}-k_r)+\sum_{s=1}^{r-1}\delta(k_r-k_{r-s})\right)!}{(1+\sum_{i=1}^{r}\delta(k_{r+1}-k_{r+1-i}))!\ \{1+(1-\delta(k_r-k_{r+1}))\sum_{i=2}^{r}\delta(k_r-k_{r+1-i})\}!}. \quad (4.60)$$

It is clear that

$$\left(1+\sum_{i=1}^{r}\delta(k_{r+1}-k_{r+1-i})\right)! = \left(1+\delta(k_{r+1}-k_r)+\sum_{i=1}^{r-1}\delta(k_{r+1}-k_{r-i})\right)! \quad (4.61)$$

and

$$\sum_{i=2}^{r} \delta(k_r - k_{r+1-i}) = \sum_{i=1}^{r-1} \delta(k_r - k_{r-i}). \tag{4.62}$$

Using (4.61) and (4.62), γ, given by (4.60), can be rewritten as

$$\begin{aligned}\gamma = & \frac{\Big(1 + \delta(k_{r+1} - k_r) + \sum_{s=1}^{r-1} \delta(k_r - k_{r-s})\Big)!}{\Big(1 + \delta(k_{r+1} - k_r) + \sum_{i=1}^{r-1} \delta(k_{r+1} - k_{r-i})\Big)!} \\ & \times \frac{1}{\Big\{1 + (1 - \delta(k_r - k_{r+1})) \sum_{i=1}^{r-1} \delta(k_r - k_{r-i})\Big\}!}.\end{aligned} \tag{4.63}$$

i) For $k_{r+1} = k_r$, γ is given by

$$\gamma = \frac{\Big(2 + \sum_{s=1}^{r-1} \delta(k_r - k_{r-s})\Big)!}{\Big(2 + \sum_{i=1}^{r-1} \delta(k_{r+1} - k_{r-i})\Big)!} = \frac{\Big(2 + \sum_{s=1}^{r-1} \delta(k_r - k_{r-s})\Big)!}{\Big(2 + \sum_{i=1}^{r-1} \delta(k_r - k_{r-i})\Big)!} = 1 \tag{4.64}$$

where the second equality follows from the fact that $k_{r+1} = k_r$.

ii) For $k_{r+1} > k_r$, γ is given by

$$\begin{aligned}\gamma &= \frac{\Big(1 + \sum_{s=1}^{r-1} \delta(k_r - k_{r-s})\Big)!}{\Big(1 + \sum_{i=1}^{r-1} \delta(k_{r+1} - k_{r-i})\Big)! \ \Big(1 + \sum_{i=1}^{r-1} \delta(k_r - k_{r-i})\Big)!} \\ &= \frac{1}{\Big(1 + \sum_{i=1}^{r-1} \delta(k_{r+1} - k_{r-i})\Big)!} = 1\end{aligned} \tag{4.65}$$

where the third equality follows from the fact that $k_{r+1} > k_r \geq k_{r-1} \geq \cdots \geq k_1$ which is derived from both (4.50) and $k_{r+1} > k_r$.

From (4.64) and (4.65), it is seen that γ is always 1, and thus (4.58) and (4.59) are the same. We have proved (4.39) for $N \geq 2$.

2) The case where $N = 1$: For this case, we have $m = 1$, and thus the number of ways of assigning k_1 to a group is one (i.e., $C(1) = 1$). If we let $m = 1$ in (4.39), we have

$$C(1) = \frac{N!}{(1+0)!\, 1\, (N-1)!} = 1 \tag{4.66}$$

where the first equality follows from (4.40). Therefore, (4.39) holds for $N = 1$.

□

Note that $C(m)$, given by (4.39), is defined for $m \geq 1$. If we let $m = 0$ in (4.39), however, $C(0)$ yields

$$C(0) = \frac{N!}{(1+0)!\,1\,(N-0)!} = 1 \tag{4.67}$$

where the first equality follows from (4.40). From (4.19), (4.20), (4.39) and (4.67), the probability of n $(0 \leq n \leq NM)$ description errors in a frame, $P_f(n)$, is given by the following:

$$P_f(n) = \sum_{m=\lceil \frac{n}{N} \rceil}^{\min(N,n)} P'_f(n,m) \tag{4.68}$$

where

$$P'_f(n,m) = \begin{cases} \sum_{k_m=\lceil \frac{n}{m} \rceil}^{\min(M,\ n-(m-1))} \sum_{k_{m-1}=\lceil \frac{n-k_m}{m-1} \rceil}^{\min(k_m,\ n-k_m-(m-2))} \cdots \sum_{k_2=\lceil \frac{n-k_m-k_{m-1}\cdots-k_3}{2} \rceil}^{\min(k_3,\ n-k_m-k_{m-1}\cdots-k_3-1)} \\ \quad C(m)\, P_g(k_m) P_g(k_{m-1}) \cdots P_g(k_2) P_g(n - \sum_{i=2}^{m} k_i) P_g^{N-m}(0), \quad \text{for} \quad m \geq 2 \\ C(m)\, P_g^m(n) P_g^{N-m}(0), \quad \text{for} \quad m = 0, 1 \end{cases}$$

where $P_g(l)$ is given by (4.11), and

$$C(m) = \frac{N!}{\left(1 + \sum_{i=1}^{m-1} \delta(k_m - k_{m-i})\right)!} \times \frac{1}{\prod_{j=1}^{m-2} \left\{1 + (1 - \delta(k_{m-j} - k_{m+1-j})) \sum_{i=j+1}^{m-1} \delta(k_{m-j} - k_{m-i})\right\}! (N-m)!}.$$

Note that the expression of $P_f(n)$, given by (4.68), holds to the case $n = 0$ since we have $C(0) = 1$ from (4.67).

Finally, from (4.3), (4.5), and (4.68), the average throughput, $E[R]$, and the average distortion, $E[D]$, are obtained in explicit expressions for given parameters such as the number of descriptions in a frame (N_t), the size of a description (L), the number of groups (N), the number of information symbols for codeword l $(c_l,\ l = 1, \ldots, L)$, SNR per modulated symbol (γ_s), and the operational rate-distortion curve $(D(x))$. Note that the expressions of $E[R]$ and $E[D]$ are not closed forms due to a single integration in $P_g(l)$ given by (4.11), while the upper and lower bounds on the average throughput or distortion are expressed in closed forms using (4.18) and (4.17), respectively. Note that these bounds can become arbitrarily close to the exact average throughput or distortion by increasing the number of terms, B, in (4.17) and (4.18) [84].

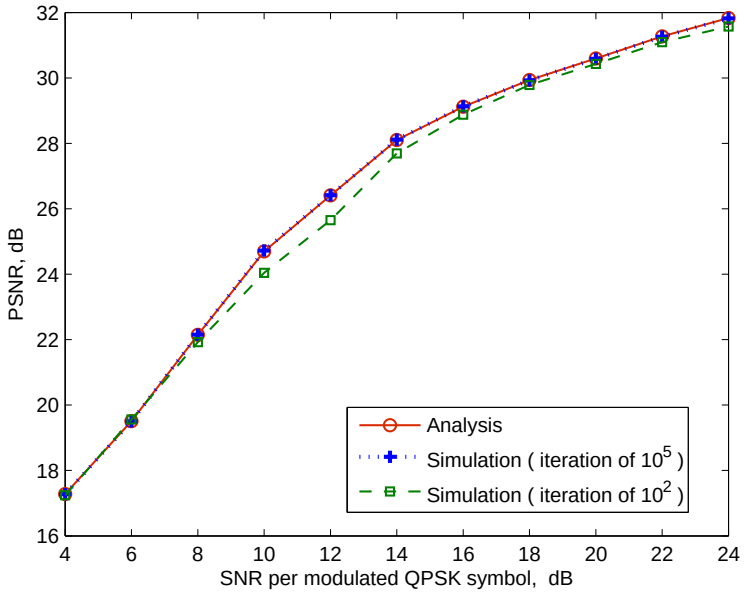


Figure 4.11: PSNR of the FEC-based multiple description coding for image transmission over OFDM system.

4.5 Numerical Evaluation and Discussion

Based on the analysis in the previous section, we evaluate the peak-signal-to-noise ratio (PSNR) performance of the FEC-based multiple description coding technique for progressive image transmission in OFDM systems. The performance is evaluated for the standard 8 bits per pixel (bpp) 512×512 Lena image with a transmission rate of 0.25 bpp using the progressive source coder SPIHT [10]. Using (4.5) and (4.68), we evaluate the average distortion, $E[D]$, for $N_t = 128$, $N = 8$, and $L = 64$ as an example, and convert $E[D]$ into PSNR using the relation of $\text{PSNR} = 10 \log \left(255^2 / E[D]\right)$. The number of parity symbols for codeword l $(1 \leq l \leq L)$, $N_t - c_l$, is optimized using the approach in [61]. The resultant PSNR for various channel SNRs is depicted in Fig. 4.13. The optimal parity symbol allocation for RS codewods at a specific SNR of 14 dB is depicted in Fig. 4.12. Monte-Carlo simulation results are also depicted in Figs. 4.11 and 4.12, and its results are almost the same as the analytical results.

For a high SNR of 18 dB, in Figs. 4.13 and 4.14, we show the probability of n description errors in a frame, $P_f(n)$, and the probability of m erroneous groups, $P'_f(n, m)$, which are given by (4.68). Note that for the block fading channel model in the frequency

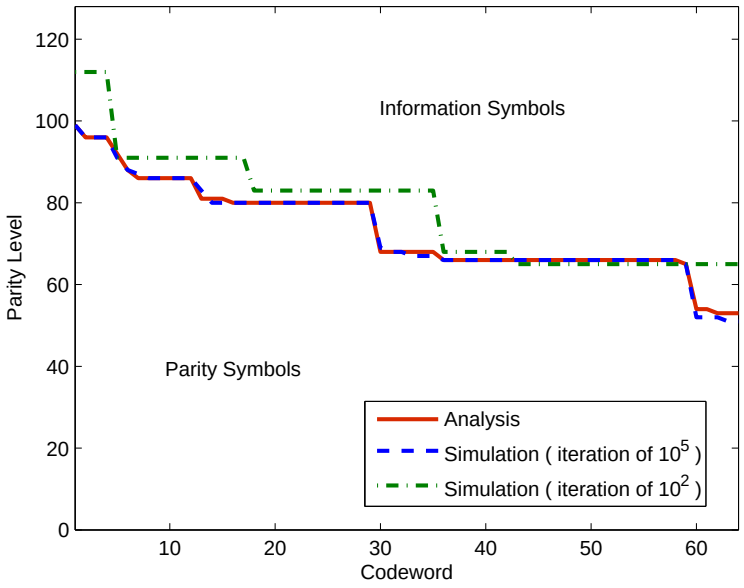


Figure 4.12: The optimal allocation of parity symbols for RS codewords at an SNR of 14 dB.

domain, 16 subcarriers within a group (i.e., $M = N_t/N = 16$) are highly correlated as shown in (4.2). Therefore, it is expected that most description errors occur as a bundle of $16, 32, 48, \cdots, 128$ (i.e., kM, $k = 1, 2, \cdots, N$), as shown in Fig. 4.13. Moreover, for $n = kM$ $(k = 1, 2, \cdots, N)$ description errors in a frame, the numbers of erroneous groups are the most likely to be k, as verified in Fig. 4.14. This indicates that as described in [71] for the transmission of fine granular scalable (FGS) video, the encoder should decide whether to include M more or M fewer descriptions in the motion-compensated prediction (MCP) loop, instead of attempting to fine-tune by including one more or one fewer description, because the correlated nature of the subcarriers makes it unlikely that only one more or one fewer description would be received. For a low SNR of 8 dB, we also evaluate $P_f(n)$ and $P'_f(n, m)$ in Figs. 4.15 and 4.16. In Fig. 4.15, the probability of kM $(k = 1, 2, \cdots, N)$ description errors in a frame is not very dominant compared to a high SNR case. Moreover, Fig. 4.16 shows that for $n = kM$ description errors in a frame, the probability of k erroneous groups is not dominant either. This result implies that at low SNR, multiple description errors do not occur in a highly correlated manner. For low SNR, the probability of error is not dominated by the fading channel effect of $\alpha[n, u, v]$ in (4.1) which is highly correlated for different u's, but is very affected

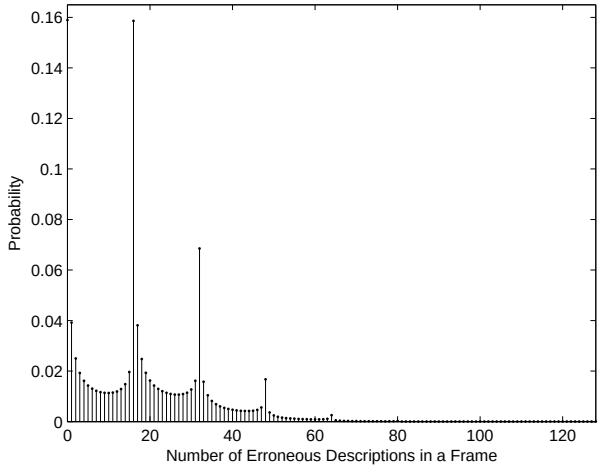
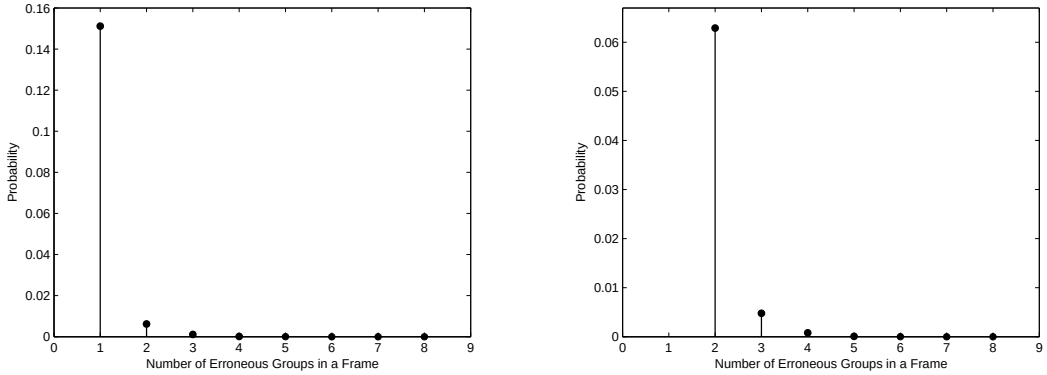


Figure 4.13: The probability of n description errors in a frame, $P_f(n)$, given by (4.68) at SNR = 18 dB.

by the additive Gaussian noise of $w[n, u, v]$ which is independent for different n's and u's. Hence, we note that at low SNR, the encoder should decide whether to include one more or one fewer description in the MCP loop of the FGS video. Likewise, when deciding the number of parity symbols of RS codewords, the encoder should attempt to fine-tune by including one more or one fewer description for error protection levels.



(a) $P'_f(n = 16, m)$ (b) $P'_f(n = 32, m)$

Figure 4.14: The probability of m erroneous groups, $P'_f(n, m)$, given by (4.68) for n description errors in a frame at SNR = 18 dB.

We next observe the probability of a code symbol error of RS codewords for various description sizes. From (4.6) and (4.7), the probability of a description error,

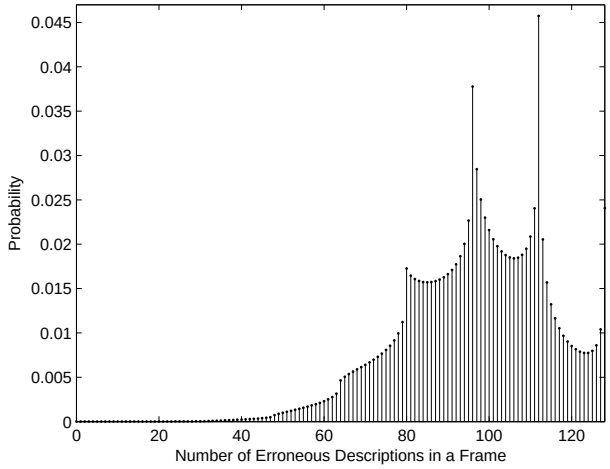
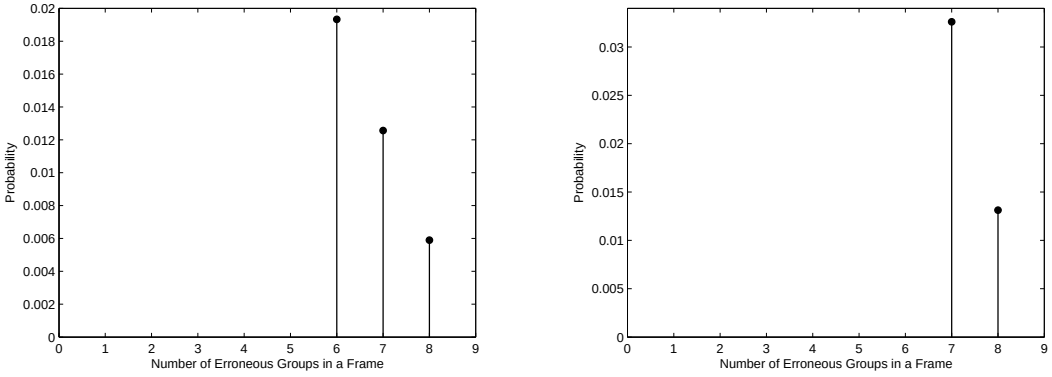


Figure 4.15: The probability of n description errors in a frame, $P_f(n)$, given by (4.68) at SNR = 8 dB.



(a) $P'_f(n = 96, m)$ (b) $P'_f(n = 112, m)$

Figure 4.16: The probability of m erroneous groups, $P'_f(n, m)$, given by (4.68) for n description errors in a frame at SNR = 8 dB.

$P_d(e)$, is derived as

$$\begin{aligned} P_d(e) &= \int_0^\infty \left(1 - (1 - P_{cs}(e|h))^L\right) P_R(h) dh \\ &= 1 - \int_0^\infty (1 - Q(h\sqrt{\gamma_s}))^{8L} P_R(h) dh \\ &= 1 - \sum_{q=0}^{8L} \left\{ (-1)^q \binom{8L}{q} \int_0^\infty Q^q(h\sqrt{\gamma_s}) P_R(h) dh \right\} \end{aligned} \tag{4.69}$$

where $P_R(h)$ is given by (4.10). From (4.7), the probability of a RS code symbol error,

$P_{cs}(e)$, is expressed as

$$P_{cs}(e) = 1 - \int_0^\infty \left(1 - Q(h\sqrt{\gamma_s})\right)^8 P_R(h)dh$$
$$= 1 - \sum_{q=0}^{8} \left\{ (-1)^q \binom{8}{q} \int_0^\infty Q^q(h\sqrt{\gamma_s}) P_R(h)dh \right\}. \quad (4.70)$$

Note that from Fig. 4.2, codewords are encoded by RS codes, and the probability of a channel symbol error for RS codewords is the same as the probability of a description error given by (4.69). On the other hand, consider the case where multiple description coding is not employed, but modulated QPSK symbols are directly encoded by RS codes in the same way as in Fig. 4.2. For this case, there is no information about whether a channel symbol is erroneous or not, and thus we do not have erasure channels for RS codes. Note that (n, k) RS codes can correct up to $n - k$ channel symbol erasures, while it can correct only up to $\lfloor (n-k)/2 \rfloor$ channel symbol errors. However, if multiple description coding is not employed, the probability of a channel symbol error for RS codewords becomes lower since it is the same as $P_{cs}(e)$ given by (4.70). To see this, in Fig. 4.17, we evaluated the probability of a description error given by (4.69) and the probability of a code symbol error given by (4.70). Fig. 4.17 shows that the probabilities of a description error for description sizes $L = 4$, 16, 64, 256 are about 2, 3, 4, 5 times greater than the probability of a code symbol error in almost all SNRs. From this, it is seen that despite the bursty nature of the errors associated with a slow fading environment, FEC-based multiple description coding without temporal coding has more advantage for smaller description sizes.

4.6 Conclusions

Multiple description source coding has emerged as an attractive framework for robust multimedia transmission over packet erasure channels. In this chapter, we mathematically analyzed the performance of n-channel symmetric FEC-based multiple description coding for transmission of progressive bitstream over OFDM networks in a frequency-selective slowly-varying Rayleigh faded environment. Using induction, we derived the average throughput and distortion performance in an explicit expression for given parameters such as the number of descriptions in an OFDM frame, the size of a description, and the channel conditions. While these exact expressions are in the form of a single integration, the upper and lower bounds of the performance were derived in a

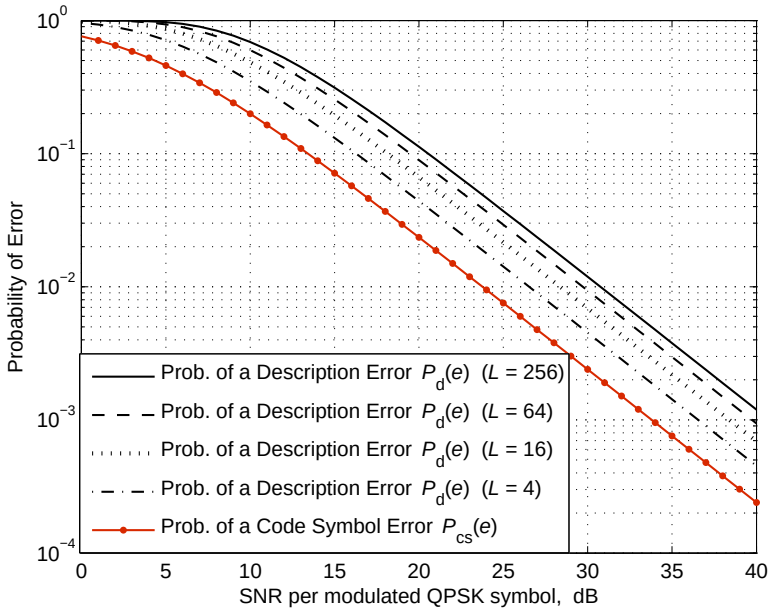


Figure 4.17: The probabilities of a description error for various description sizes (L = 4, 16, 64, 256) and the probability of a RS code symbol error.

closed-form expression. Using the derived analysis, the performance for the transmission of progressive image was numerically evaluated, and it was shown to be the same as the computationally intensive simulation results. We also numerically evaluated the number of description errors and the number of erroneous groups in a frame. This showed that at low SNR, the multiple description encoder should attempt to fine-tune the system optimization parameters such as the error protection level for RS codewords and the number of descriptions included in the motion-compensated prediction loop of fine granular scalable video. We also evaluated the probability of a description error, which showed that despite the bursty nature of the errors in a slow fading environment, FEC-based multiple description coding without temporal coding in a wireless environment has more advantage for smaller description sizes.

4.7 Acknowledgements

The text of this chapter, in full, is a reprint of "Performance Analysis of n-Channel Symmetric FEC-Based Multiple Description Coding for OFDM Networks," which has been accepted for publication in *IEEE Transactions on Image Processing*.

The dissertation author was the primary investigator of this paper.

Chapter 5

Conclusion

This dissertation presented a cross-layer study of wireless multimedia communications. The following are the main contributions of this dissertation. Some possible future work is also outlined.

5.1 Chapter 2

To enhance the performance for the hierarchical transmission of progressive sources, we proposed a way of achieving an arbitrarily large number of UEP levels. We proved that multiple levels of UEP can be achieved by multiplexing hierarchical modulation. We next derived an optimal multiplexing approach which minimizes both the average and peak powers when the BER is dominated by the Gaussian Q-function term having the minimum Euclidian distance. To mitigate the high PAPR effect, an asymmetric hierarchical QAM constellation was proposed. We also considered the case where multiplexed constellations need to have constant power, and showed that multilevel UEP can be achieved for this case. Numerical results showed that the proposed multilevel UEP system significantly enhances the performance for progressive transmission over Rayleigh fading channels without an increase in system bandwidth or transmit power.

5.2 Chapter 3

For the broadcast of layered sources in MIMO systems, we proposed superposition MIMO coding. First, the tradeoff between Alamouti coding and spatial multiplexing having the same transmission rate is analyzed in terms of the average bit error rate. The

results showed that, for a given target bit error rate, the Alamouti coding is preferable to spatial multiplexing for a low data rate source, and vice versa for a high data rate source. In particular, we considered layered sources in which the more important component consists of a smaller number of bits than the less important component (a typical example is scalable video). In the proposed scheme, two different MIMO techniques are hierarchically combined such that low-rate important data is Alamouti encoded, high-rate less important data is spatially multiplexed, and then the two differently encoded data symbols are superposed. A successive decoding algorithm for the proposed scheme was also provided. Performance evaluation in a broadcasting scenario showed that the proposed superposition MIMO coding significantly outperforms the conventional MIMO coding.

In the analysis of the tradeoff between space-time coding and spatial multiplexing, we assumed that the receiver has access to perfect channel status information (CSI). Effects of imperfect CSI on the tradoff is an interesting research area. In addition, we only considered the Alamouti coding for space-time coding schemes.

5.3 Chapter 4

We analyzed the performance of n-channel symmetric FEC-based multiple description coding for transmission of progressive bitstreams over OFDM networks, where a channel is assumed to experience frequency-selective, slowly-varying, Rayleigh fading. We derived the average throughput and distortion as a function of parameters such as the number
of descriptions in an OFDM frame, the size of a description, and the channel conditions. Using the analysis, the performance for the transmission of a progressive image was numerically evaluated, and it was shown to yield the same result as does a computationally intensive simulation. We also numerically evaluated the number of description errors and the number of erroneous groups in a frame, which showed that at low SNR, the multiple description encoder should attempt to fine-tune the system optimization parameters. For example, in the transmission of fine granular scalable video, the encoder should decide whether to include one more or one fewer description in the motion-compensated prediction loop.

Bibliography

[1] Y. K. Kim and R. Prasad, *4G Roadmap and Emerging Communication Technologies*, Artech House, 2006.

[2] S. Shakkottai, T. S. Rappaport, and P. C. Karlsson, "Cross-layer design for wireless networks," *IEEE Commun. Magazine*, vol. 41, no. 10, pp. 74–80, Oct. 2003.

[3] V. Kawadia and P. R. Kumar, "A cautionary perspective on cross-layer design," *IEEE Wireless Commun.*, vol. 12, no. 1, pp. 3–11, Feb. 2005.

[4] M. V. D. Schaar and S. N. Shankar, "Cross-layer wireless multimedia transmission: Challenges, principles, and new paradigms," *IEEE Wireless Commun.*, vol. 12, no. 4, pp. 50–58, Aug. 2005.

[5] A. S. Tanenbaum, *Computer Networks*, 2nd ed., Printice-Hall, Englewood Cliffs, NJ, 1989.

[6] B. Hochwald and K. Zeger, "Tradeoff between source and channel coding," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1412–24, Sept. 1997.

[7] B. Hochwald, "Tradeoff between source and channel coding on a Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 3044-3055, Nov. 1998.

[8] A. Mehes and K. Zeger, "Source and channel rate allocation for channel codes satisfying the Gilbert-Varshamov or Tsfasman-Vladut-Zink bounds," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 2133–2151, Sept. 2000.

[9] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3445–3462, Dec. 1993.

[10] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 243–249, June 1996.

[11] D. Taubman and M. Marcellin, JPEG2000: Image Compression Fundamentals, Standards, and Practice. Norwell, MA: Kluwer, 2001.

[12] J. Reichel, H. Schwarz, and M. Wien (eds.), "Scalable Video Coding -Working Draft 1," Joint Video Team of ITU-T VCEG and ISO/IEC MPEG, Doc. JVT-N020, Hong Kong, CN, Jan. 2005.

[13] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sept. 2007.

[14] F. Wu, S. Li, and Y.-Q. Zhang, "A framework for efficient progressive fine granularity scalable video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 282–300, Mar. 2001.

[15] T. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[16] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inform. Theory*, vol. 19, no. 2, pp. 197–207, Mar. 1973.

[17] E. C. van der Meulen, "A survey of multi-way channels in information theory: 1961–1976," *IEEE Trans. Inform. Theory*, vol. 23, pp. 1–37, Jan. 1977.

[18] K. Ramchandran, A. Ortega, K.M. Uz, and M.Vetterli, "Multiresolution broadcast for digital HDTV using joint source/channel coding," *IEEE J. Select. Areas Commun.*, vol. 11, no. 1, pp. 6–23, Jan. 1993.

[19] L.-F. Wei, "Coded modulation with unequal error protection," *IEEE Trans. Commun.*, vol. 41, no. 10, pp. 1439.1449, Oct. 1993.

[20] A. R. Calderbank and N. Seshadri, "Multilevel codes for unequal error protection," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1234–1248, July 1993.

[21] M. Morimoto, H. Harada, M. Okada, and S. Komaki, "A study on power assignment of hierarchical modulation schemes for digital broadcasting," *IEICE Trans. Commun.*, vol. E77-B, pp. 1495–1500, Dec. 1994.

[22] M. B. Pursley and J. M. Shea, "Nonuniform phase-shift-key modulation for multimedia multicast transmission in mobile wireless networks," *IEEE J. Select. Areas Commun.*, vol. 5, pp. 774–783, May 1999.

[23] M. B. Pursley and J. M. Shea, "Adaptive nonuniform phase-shift-key modulation for multimedia traffic in wireless networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1394–1407, Aug. 2000.

[24] "Digital Video Broadcasting (DVB); Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television," ETSI EN 300 744 V1.5.1, Nov. 2004.

[25] C. Weck, "Coverage aspects of digital terrestrial television broadcasting," *EBU Techn. Rev.*, pp. 19–30, Winter 1996.

[26] A. Schertz and C.Weck, "Hierarchical modulation-the transmission of two independent DVB-T multiplexes on a single frequency," *EBU Techn. Rev.*, Apr. 2003.

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

[28] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

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

[30] G. Ganesan and P. Stoica, "Space-time block codes: A maximum SNR approach," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1650–1656, May 2001.

[31] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs. Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[32] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Int. Symp. Signals, Systems, Electronics*, Pisa, Italy, Sept.-Oct. 1998, pp. 295–300.

[33] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[34] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–15, 1999.

[35] P. G. Sherwood and K. Zeger, "Progressive image coding for noisy channels," *IEEE Signal Process. Lett.*, vol. 4, no. 7, pp. 191–198, July 1997.

[36] P. Cosman, J. Rogers, P. G. Sherwood, and K. Zeger, "Combined forward error control and packetized zerotree wavelet encoding for transmission of images over varying channels," *IEEE Trans. Image Process.*, vol. 9, no. 6, pp. 132–140, June 2000.

[37] V. Stankovic, R. Hamzaoui, Y. Charfi, and Z. Xiong, "Real-time unequal error protection algorithms for progressive image transmission," *IEEE J. Select. Areas Commun.*, vol. 21, no. 10, Dec. 2003.

[38] Y. Pei and J. W. Modestino, "Cross-Layer Design for Video Transmission over Wireless Rician Slow-Fading Channels Using an Adaptive Multiresolution Modulation and Coding Scheme," *EURASIP J. Advances Signal Process.*, vol. 2007, article ID 86915.

[39] V. K. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 74–93, Sept. 2001.

[40] S. S. Pradhan, R. Puri, and K. Ramchandran, "n-channel symmetric multiple descriptions-Part I: (n, k) source channel erasure codes," *IEEE Trans. Inform. Theory*, vol. 50, no. 1, pp. 47–61, Jan. 2004.

[41] J. F. Hayes, "Adaptive feedback communications," *IEEE Trans. Commun. Technol.*, vol. 16, no. 1, pp. 29–34, Feb. 1968.

[42] W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Trans. Commun.*, vol. 43, no. 7, pp. 2223–2230, July 1995.

[43] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

[44] C.-E. W. Sundberg, W. C. Wong, and R. Steele, "Logarithmic PCM weighted QAM transmission over Gaussian and Rayleigh fading channels," *Proc. Inst. Elec. Eng.*, vol. 134, pp. 557–570, Oct. 1987.

[45] B. Masnick and J. Wolf, "On linear unequal error protection codes," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 600–607, Oct. 1967.

[46] H. Suda and T. Miki, "An error protected 16 kb/s voice transmission for land mobile radio channel," *IEEE J. Select. Areas Commun.*, vol. 6, pp. 346–352, Feb. 1988.

[47] R. V. Cox, J. Hagenauer, N. Seshadri, and C.-E. W. Sundberg, "Subband speech coding and matched convolutional channel coding for mobile radio channels," *IEEE Trans. Signal Processing*, vol. 39, pp. 1717–1731, Aug. 1991.

[48] B. Barmada, M.M. Ghandi, E.V. Jones, and M. Ghanbari, "Prioritized transmission of data partitioned H.264 video with hierarchical QAM," *IEEE Signal Process. Lett.*, vol. 12, pp. 577–580, Aug. 2005.

[49] P. K. Vitthaladevuni and M.-S. Alouini, "A recursive algorithm for the exact BER computation of generalized hierarchical QAM constellations," *IEEE Trans. Inform. Theory*, vol. 49, pp. 297–307, Jan. 2003.

[50] J. G. Proakis, *Digital Communication*, 3rd ed. New York: McGraw-Hill, 1995.

[51] 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.

[52] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge, U.K.: Cambridge Univ. Press, 2003.

[53] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge, U.K.: Cambridge Univ. Press, 2005.

[54] J. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42. no. 234, pp. 1740–1751, Feb./Mar./Apr. 1994.

[55] D. Gore, R. W. Heath Jr., and A. Paulraj, "On performance of the zero forcing receiver in presence of transmit correlation," in *Proc. Int. Symp. Inf. Theory*, 2002, p. 159.

[56] A. J. Goldsmith and S. B.Wicker, "Design challenges for energy-constrained Ad Hoc wireless networks," *IEEE Wireless Commun.*, vol. 9, no. 4, pp. 8–27, Aug. 2002.

[57] Y. Shen, P. C. Cosman, and L. B. Milstein, "Video coding with fixed length packetization for a tandem channel," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 273–288, Feb. 2006.

[58] L. Ozarow, "On a source-coding problem with two channels and three receivers," *Bell Syst. Tech. J.*, vol. 59, no. 10, pp. 1909–1921, Dec. 1980.

[59] A. El Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inform. Theory*, vol. 28, no. 6, pp. 851–857, Nov. 1982.

[60] J. K.Wolf, A. D.Wyner, and J. Ziv, "Source coding for multiple descriptions," *Bell Syst. Tech. J.*, vol. 59, no. 8, pp. 1417–1426, Oct. 1980.

[61] A. Mohr, E. Riskin, and R. Ladner, "Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction," *IEEE J. Select. Areas Commun.*, vol. 18, no. 6, pp. 819–828, June 2000.

[62] R. Puri and K. Ramchandran, "Multiple description source coding using forward error correction codes," in *Proc. 33rd Asilomar Conf. Signals, Systems and Computer*, Pacific Grove, CA, Oct. 1999, pp. 342–346.

[63] L. Cheng, W. Zhang, and L. Chen, "Rate-distortion optimized unequal loss protection for FGS compressed video," *IEEE Trans. Broadcasting*, vol. 50, no. 2, pp. 126–131, June 2004.

[64] M. van der Schaar and H. Radha, "Unequal packet loss resilience for fine-granular-scalability video," *IEEE Trans. Multimedia*, vol. 3, no. 4, pp. 381–394, Dec. 2001.

[65] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoded transmission," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 1737–1744, Nov. 1996.

[66] R. Puri, K.-W. Lee, K. Ramchandran, and V. Bharghavan, "An integrated source transcoding and congestion control paradigm for video streaming in the Internet," *IEEE Trans. Multimedia*, vol. 3, no. 1, pp. 18–32, Mar. 2001.

[67] D. G. Sachs, R. Anand, and K. Ramchandran, "Wireless image transmission using multiple-description based concatenated codes," in *Proc. SPIE*, San Jose, CA, Jan. 2000, vol. 3974, pp. 300–311.

[68] P. G. Sherwood and K. Zeger, "Error protection for progressive image transmission over memoryless and fading channels," *IEEE Trans. Commun.*, vol. 46, no. 12, pp. 1555–1559, Dec. 1998.

[69] Y. S. Chan, P. C. Cosman, and L. B. Milstein, "A cross-layer diversity technique for multi-carrier OFDM multimedia networks," *IEEE Trans. Image Process.*, vol. 15, no. 4, pp. 833–847, Apr. 2006.

[70] Y. Sun, Z. Xiong, and X. Wang, "Scalable image transmission over differentially space-time coded OFDM systems," in *Proc. IEEE Globecom 2002*, Taipei, Taiwan, R.O.C., Nov. 2002, vol. 1, pp. 379–383.

[71] Y. S. Chan, P. C. Cosman, and L. B. Milstein, "A multiple description coding and delivery scheme for motion-compensated fine granularity scalable video," *IEEE Trans. Image Process.*, vol. 17, no. 8, pp. 1353–1367, Aug. 2008.

[72] J. Song and K. J. R. Liu, "Robust progressive image transmission over OFDM systems using space-time block code," *IEEE Trans. Multimedia*, vol. 4, no. 3, pp. 394–406, Sept. 2002.

[73] L. Toni, Y. S. Chan, P. C. Cosman, and L. B. Milstein, "Channel coding for progressive images in a 2-D time-frequency OFDM block with channel estimation errors," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2476–2490, Nov. 2009.

[74] S.-S. Tan, M. Rim, P. C. Cosman, and L. B. Milstein, "Adaptive modulation for OFDM-based multiple description progressive image transmission," in *Proc. IEEE Globecom 2008*, New Orleans, LA, Nov./Dec. 2008.

[75] Lingling Pu, Michael W. Marcellin, Ivan Djordjevic, Bane Vasic, and Ali Bilgin, "Joint source-channel rate allocation in parallel channels," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2016–2022, Aug. 2007.

[76] S. K. Bandyopadhyay, G. Partasides, L. P. Kondi, "Cross-layer optimization for video transmission over multirate GMC-CDMA wireless links," *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 1020–1024, June 2008.

[77] R. J. McEliece and W. E. Stark, "Channels with block interference," *IEEE Trans. Inform. Theory*, vol. 30, no. 1, pp. 44–53, Jan. 1984.

[78] M. Medard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 840–852, Apr. 2002.

[79] E. Malkamaki and H. Leib, "Performance of truncated type-II hybrid ARQ schemes with noisy feedback over block fading channels," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1477–1487, Sept. 2000.

[80] R. Knopp and P. A. Humblet, "On coding for block fading channels," *IEEE Trans. Inform. Theory*, vol. 46, no. 1, pp. 189–205, Jan. 2000.

[81] M. Zorzi, R. R. Rao, and L. B. Milstein, "Error statistics in data transmission over fading channels," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1468–1477, Nov. 1998.

[82] H. Bischl and E. Lutz, "Packet error rate in the non-interleaved Rayleigh channel," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 1375–1382, Feb./Mar./Apr. 1995.

[83] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels*, 2nd ed. New York: Wiley, 2005.

[84] M. Chiani, D. Dardari, and M. K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 840–845, July 2003.