

Coded Modulation and Equalization for Wireless Infrared Communications

A THESIS
Presented to
The Academic Faculty
By
Hyuncheol Park

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering



School of Electrical and Computer Engineering

Georgia Institute of Technology

September, 1997

Copyright ©1997 by Hyuncheol Park

Coded Modulation and Equalization for Wireless Infrared Communications

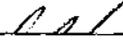
Approved:



Professor John R. Barry, Chairman



Professor Steven W. McLaughlin



Professor Gordon L. Stüber

Date approved by Chairman 8-20-97

Acknowledgments

I would like to thank to my mother, my wife Kyungeun, my son Junyong, my daughter Jiwon, and my sister Hyunjoo for their encouragement, endurance, and sacrifice for my late study. Without them, I would not have been where I am.

My great thanks go to professor John Barry. He has done an excellent job as a research advisor, and I could not make a better choice. I sincerely appreciate for his advice, suggestions, and patience during the last four years which leads to this dissertation.

I would like to thank Professor Gordon Stüber for serving as the chairman of my qualifying examination, and the reading committee of both my proposal and defense examinations. I also would like to thank Professor Steven McLaughlin for serving as the chairman of my proposal examination, and the reading committee of my defense examination. I appreciate Professor Douglas Williams and Professor Xingxing Yu for serving on my defense examination committee.

I am thankful for my colleagues; Anuj Batra, Ricky Causey, and Chen-chu Yeh. Anuj had been a my officemate for the first two years. I learned a lot from Ricky about the Star Trek physics. In particular, Chen-chu has been a my officemate and lunchmate for the last three years, and I really enjoyed the times that he shared with me.

Finally, I would like to thank to my Korean friends in communication and signal processing groups Jinsoup, Yongsub, Sangyoun, Ho, Jeongwook, and Dukhyun for their encouragements.

Not that I have already obtained all this, or have already been made perfect, but I press on to take hold of that for which Christ Jesus took hold of me (Philippians 3:12).

Table of Contents

List of Tables	vii
List of Figures	viii
Summary	xii
1 Introduction	1
1.1 WIRELESS INFRARED COMMUNICATIONS.....	2
1.2 MODELING OF WIRELESS INFRARED CHANNEL.....	5
1.3 THESIS OUTLINE.....	9
2 Performance of Modulation Schemes	11
2.1 POWER EFFICIENCY AND BANDWIDTH EFFICIENCY.....	12
2.1.1 Definitions.....	12
2.1.2 Pulse-position Modulation (PPM).....	14
2.1.3 Multiple-Pulse Position Modulation (MPPM).....	15
2.1.4 Overlapping Pulse-Position Modulation (OPPM).....	18
2.1.5 Pulse Amplitude and Position Modulation (PAPM).....	20
2.1.6 Discussions	22
2.2 ACCURATE BANDWIDTH CALCULATION.....	22
2.3 MPPM BOUND.....	26
2.4 SUMMARY AND CONCLUSIONS.....	27
3 Equalization	29
3.1 INTRODUCTION.....	29

3.1.1	Maximum-Likelihood Sequence Detection (MLSD)	30
3.1.2	Symbol-by-symbol Equalization	30
3.1.3	Precoding	32
3.1.4	Outline	33
3.2	SYSTEM MODEL	33
3.3	UNEQUALIZED RECEIVER	35
3.3.1	Error Probability for MPPM and OPPM	35
3.3.2	Extension to On-Off Keying	38
3.3.3	Simplifications when the Channel is Ideal	39
3.3.4	Probability of Error Approximation for PPM	40
3.4	MAXIMUM-LIKELIHOOD SEQUENCE DETECTION (MLSD)	44
3.5	SYMBOL-BY-SYMBOL EQUALIZATION	50
3.5.1	Zero-forcing Block Decision Feedback Equalization (ZF-BDFE)	51
3.5.2	Zero-forcing Linear Equalization (ZF-LE)	53
3.5.3	Maximum a Posteriori Probability (MAP) Detection	55
3.6	PARTIAL-RESPONSE PRECODING SCHEME	59
3.7	NUMERICAL RESULTS	64
3.8	SUMMARY AND CONCLUSIONS	65
4	Achievable Information Rate and Cutoff Rate	68
4.1	INTRODUCTION	69
4.2	MEMORYLESS CHANNEL	70
4.2.1	Achievable Information Rate	70
4.2.2	Cutoff Rate	71
4.3	ISI CHANNEL	72
4.3.1	Lower Bound of the Information Rate: I_L	73
4.3.2	Upper Bound of the Information Rate: I_U	75
4.3.3	Cutoff Rate	76
4.4	NUMERICAL RESULTS	77

4.5	SUMMARY AND CONCLUSIONS	80
5	Coded Modulation	82
5.1	INTRODUCTION	82
5.2	CONVOLUTIONAL CODED PPM	85
5.2.1	Rate $1/\log_2 L$ Convolutional Coded L -PPM	85
5.2.2	Rate $(\log_2 L - 1)/\log_2 L$ Convolutional Coded L -PPM	87
5.3	TRELLIS CODED OPPM	91
5.4	TRELLIS CODED MPPM.....	93
5.4.1	Design of Trellis Coded MPPM	94
5.4.2	An Approximation for the d_{min} of Trellis Coded MPPM	101
5.5	PERFORMANCE OF CODED MODULATION SCHEMES	102
5.6	SUMMARY AND CONCLUSION	104
6	Coded Modulation on a Multipath Channel	105
6.1	INTRODUCTION	105
6.2	SYSTEM MODEL	108
6.3	SUPERSTATE MLSD.....	109
6.4	PARALLEL DECISION FEEDBACK DETECTION (PDFD)	111
6.5	PRECODING	111
6.6	NUMERICAL RESULTS	113
6.7	SUMMARY AND CONCLUSIONS	116
7	Conclusions and Future Work	118
7.1	CONCLUSIONS	118
7.2	FUTURE WORK	120
7.2.1	Channel Capacity of Wireless Infrared Channel	120
7.2.2	Concatenated Codes.....	122
7.2.3	Turbo Codes.....	124
7.2.4	Synchronization	124

7.2.5 System Implementation	126
References	127
Vita	135

List of Tables

1-1	Examples of infrared links.	5
3-1	Fractional energy ratio (%) of impulse response for various R_b / W	42
5-1	The distance weight profile of trellis coded 8-MPPM.	98
5-2	Parity check coefficients for trellis coded 128-MPPM in octal form.	100
7-1	The bandwidth-efficiency ratio between two serial concatenated schemes.	123

List of Figures

1-1	Two most typical wireless infrared links. (a) Directed, LOS link, (b) Nondirected, non-LOS, or diffuse link.	3
1-2	(a) Transmission and reception in an infrared link with intensity modulation and direct detection, (b) modeling link as a baseband linear, time-invariant system having impulse response $h(t)$ with additive noise $n(t)$	6
2-1	Four pulse-position modulation (PPM).	14
2-2	The $(\frac{4}{2})$ multiple-pulse position modulation.	16
2-3	Choose overlapping pulse-position modulation (OPPM) codewords from MPPM codewords.	19
2-4	The 2-4-pulse amplitude and position modulation.	21
2-5	Power and bandwidth efficiency of uncoded modulation schemes on an ideal channel.	23
2-6	Power spectral density of modulation schemes (a) PPM, (b) MPPM, (c) OPPM. ..	25
2-7	Binary entropy function.	27
2-8	Power efficiency and bandwidth efficiency of $(\frac{n}{w})$ -MPPM and MPPM bound.	28
3-1	(a) Block diagram of MPPM system, (b) equivalent vector channel.	33
3-2	Decision device for the unequalized receiver.	36
3-3	Gaussian approximation method and the exact method for the unequalized receiver.	43
3-4	The required power versus bit rate on an ISI channel (a) unequalized system, (b) with MLSD.	48
3-5	Power efficiency versus bandwidth efficiency of PPM over ISI channel.	49

3-6	Normalized power efficiency versus bit-rate-to-bandwidth ratio on an multipath channel with MLSD for (a) PPM, (b) MPPM.	50
3-7	Block diagram of ZF-BDFE.	51
3-8	Block diagram of chip-rate linear equalizer.	53
3-9	Bit error rate of equalization schemes and their upper bounds for 8-PPM at bit-rate-to-bandwidth ratio of 1: (a) DFE with suboptimal decision device (DFE-COM) and DFE with ML decision device, (b) LE.	55
3-10	Normalized power requirement of equalization schemes for 16-PPM.	56
3-11	Block diagram of the MAP-DFE.	58
3-12	Precoding operation (a) moving DFE to the transmitter, (b) the proposed precoding scheme, (c) equivalent block diagram for (b).	60
3-13	Two types of receiver for precoding scheme; the upper one is an inverse precoder and lower one is for partial-response precoding.	62
3-14	MAP-DFE at the receiver when transmitter uses partial-response precoding.	64
3-15	Performance of equalization schemes for 4-PPM at bit-rate-to-bandwidth is equal to 1.	66
3-16	Complexity vs. performance for each equalization scheme using 4-PPM at $R_b / W = 1$; x -axis represents the number of floating-point operations (flops) per bit and y -axis represents the required SNR to achieve 10^{-3} bit error rate.	67
4-1	Approximate information rate, $I_{i.i.d.}$, and cutoff rate, R_0 , for 4-PPM as a function of SNR at $R_b / W = 0.01$ and $R_b / W = 0.5$	78
4-2	Bounds on the information rate for 2-PPM over an ISI channel as a function of SNR; information rate for the memoryless channel I_{ideal} , lower bound I_L , upper bound I_U , and approximate information rate $I_{i.i.d.}$	79
4-3	The required power to achieve a cutoff rate of $(\log_2 L - \epsilon)$ bits / codeword ($\epsilon = 10^{-3}$) normalized by P_{OOK} , versus bit rate on an ISI channel.	80
5-1	Block diagram of coded modulation.	83

5-2	A convolutional coded 4-PPM and its trellis.	86
5-3	An error event with Hamming distance $2(v + 1)$	87
5-4	The calculated minimum distance (denoted by circles) of convolutional coded PPM with its upper bounds (5-2) and (5-7) (denoted by solid line), and simplex bound (5-8) (denoted by dashed line).	90
5-5	Convolutional coded 16-PPM with cutoff rate bound.	91
5-6	The 8-OPPM signal set and its set partitioning.	94
5-7	Constellations for $(\frac{n}{2})$ -MPPM; the shaded circles represent the chosen L codewords and unshaded circles represent the unused codewords. The x -axis represents the position of first pulse and the y -axis represents the position of second pulse. (a) $L = 8$ and $n = 5$, (b) $L = 16$ and $n = 7$, (c) $L = 32$ and $n = 9$, (d) $L = 64$ and $n = 13$, (e) $L = 128$ and $n = 17$	95
5-8	Set partitioning of 8-MPPM.	97
5-9	Systematic feedback encoder for trellis coded 128-MPPM with constraint length v and parity check coefficients h_{mn} , $m = 1, 2, \dots, v - 1$ and $n = 0, 1, \dots, 6$	99
5-10	Power efficiency and bandwidth efficiency of coded modulation.	103
6-1	Block diagram of coded modulation on an ISI channel.	109
6-2	Block diagram of DFE with tentative decision and MLSD.	112
6-3	Power penalties of trellis coded 128-MPPM vs. trellis coded 16-PPM.	114
6-4	Performance of equalizer for the convolutional coded 8-PPM at $R_b / W = 1$	115
6-5	Performance of equalizer for the trellis coded 128-PPM at $R_b / W = 1$	116
6-6	Complexity vs. performance for each equalization scheme using convolutional coded 8-PPM at $R_b / W = 1$; x -axis represents the number of floating-point operations per bit and y -axis represents the required SNR to achieve 10^{-3} bit error rate.	117
7-1	Serial concatenated codes: (a) convolutional coded PPM,	

(b) PPM with convolutional code, (c) a systematic feedforward convolutional encoder for configuration (b).	122
7-2 Effects of timing error for OOK when (a) current bit = 0 and next bit = 0, (b) current bit = 0 and next bit = 1, (c) current bit = 1 and next bit = 0, (d) current bit = 1 and next bit = 1.	125

Summary

The channel model for wireless infrared communication is unique: it combines the intensity-modulation constraints of the Poisson photon-counting channel with the multipath dispersion, limited bandwidth, and Gaussian noise of the conventional radio channel. The objective of this thesis is to develop and analyze new power-efficient modulation, coding, and equalization schemes that are well-suited to the wireless infrared channel. Our strategy is to combine new trellis codes with precoding, so that the coding gain of the trellis code is obtained in combination with the equalization performance of a decision-feedback equalizer.

We investigate the performance of several candidate uncoded modulation techniques in terms of error probability, bandwidth expansion, information rate, cutoff rate, and sensitivity to multipath dispersion. To mitigate intersymbol interference, we propose a partial-response precoding scheme that is compatible with the infrared channel. For each modulation scheme, we compare the performance of symbol-by-symbol equalizers such as the linear equalizer, the decision feedback equalizer, the maximum a posteriori detector, and the proposed precoder to the optimum maximum-likelihood sequence detector.

We design new trellis codes based on multiple pulse-position modulation that offer high power efficiency. We randomly search for the optimal code that produces the trellis code with the largest minimum Euclidean distance. To verify our results, we derive an approximation for minimum distance. We combine partial-response precoding with parallel decision-feedback detection to equalize and to decode the trellis codes. The performance of the proposed scheme is compared to linear equalization, decision feedback equalization, parallel decision-feedback detection, and super-state maximum-likelihood sequence detection. Together, the proposed trellis codes and precoding schemes are an effective solution to the signaling design problem, especially in the face of severe multipath dispersion.

CHAPTER 1

INTRODUCTION

The rapid growth of the laptop and handheld computer industries has elevated the importance of indoor wireless communications and wireless local area networks. There are several options for the transmission medium in indoor wireless communication: radio wave, microwave, millimeter wave, and infrared radiation. In both research and commercial products, radio and microwave with frequencies less than 30 GHz are the most commonly used. Millimeter wave (30 ~ 300 GHz) lies between the microwave and far infrared region. There is currently an interest in the millimeter wave near 60 GHz because of its high attenuation due to oxygen absorption. This frequency band is useful for applications requiring a high attenuation beyond the normal service area to reduce co-channel interference [1]. However, devices operating at this frequency are still very expensive.

An attractive alternative is infrared radiation with wavelengths in the 750-1000 nm range. As a medium for indoor wireless networks, nondirected infrared radiation offers several advantages over radio, microwave, and millimeter wave as follows:

- Infrared offers an immense window of unregulated bandwidth. On the contrary, the spectral regions for radio, microwave, and millimeter wave are strictly regulated to use and almost scarce.
- Infrared radiation cannot penetrate walls and the transmitted signal remains in the same room where it originates. This prevents eavesdropping and interferences from neighboring rooms. (But, the signal confinement in a room is also a drawback since it limits the range covered).
- The dimension of infrared detector is much larger than the operating wavelength. This leads to an equivalent spatial diversity which prevents the multipath fading.

1.1 WIRELESS INFRARED COMMUNICATIONS

At present, most infrared links are directed, *line-of-sight* (LOS) transmission in the 850-950 nm range as shown in Fig. 1-1-a. In directed, LOS links, a directional transmitter and receiver are used and must be aimed to establish a link. This type of link depends on the existence of a LOS path between the transmitter and receiver to maximize the power efficiency and to minimize multipath distortion. For example, JOLT announced a 125 Mb/s LOS system that employs a 1° transmitter beam and a 6° receiver field of view. British Telecommunication Laboratory reported a 155 Mb/s directed, LOS link using on-off keying (OOK) [2]. In 1993, a consortia of over 150 companies formed the Infrared Data Association (IrDA) to set and to support hardware and software standards that create infrared communication links [3]. IrDA developed the standard for short range, low cost,

and low power LOS links operating at a bit rate of 4 Mb/s using 4 pulse-position modulation (PPM). In 1995, Microsoft announced support for IrDA connectivity to Windows'95, enabling wireless connectivity between Windows'95 based PCs and peripheral devices. IrDA-compliant IR ports are now an integral feature of most laptop computers, printers, and electronic organizers, and they will be incorporated into cellular phones, pagers, watches, and automatic teller machine [4].

Consider next the nondirected, non-LOS, or *diffuse* configuration as shown in Fig. 1-1-b. In a diffuse link, the transmitter and receiver are not aligned, and the link design does not depend on the LOS path, but depends on reflections from ceiling, walls, and other reflectors. A diffuse link is the most convenient and easy-to-use, but it has a higher path loss than the LOS configuration. For example, Spectrix Corporation has developed a 4

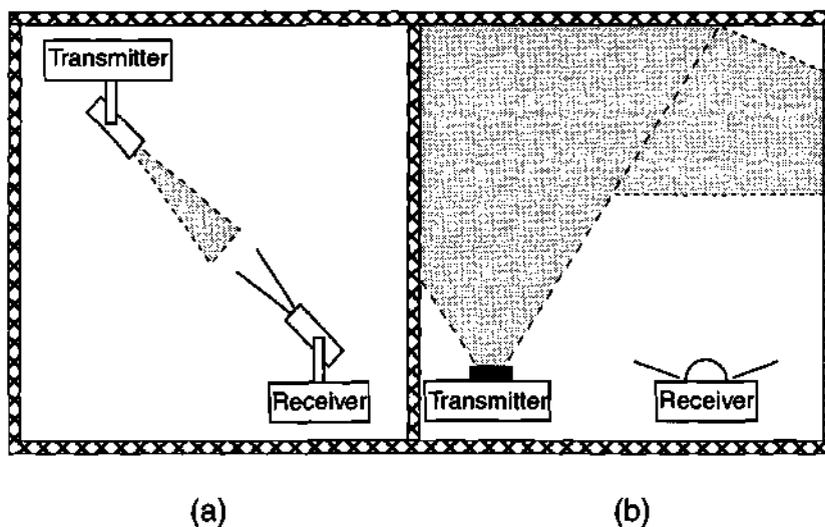


Fig. 1-1. Two most typical wireless infrared links. (a) Directed, LOS link (b) Nondirected, non-LOS, or diffuse link [9].

Mb/s wireless LAN employing OOK with an operating range of 15 m [5]. This configuration is designed to communicate between a portable transceiver and a host computer in large open areas such as offices, factories, or the trading floors of stock markets. Photonics and IBM developed a diffuse infrared *ad hoc* LAN operating at 1 Mb/s using 16-PPM within a 10 m² room. This type of link is employed to achieve direct, peer-to-peer communication between a number of portables and fixed terminals. Finally, the performance of a 50 Mb/s diffuse link using OOK was demonstrated by an experiment at UC Berkeley [6]. Table 1-1 summarizes the key features of several infrared links. All systems listed, except the first prototype by IBM, use OOK or pulse-position modulation (PPM). Compared to the diffuse links, the directed, LOS links achieve higher data rates.

Since Gfeller and Bapst's pioneering work [7] in diffuse infrared communication in 1979, there has been a growth in research activity and product development in this area. Kotzin [8] described the design of an experimental prototype for a portable telephone system. Barry made contributions in modulation analysis [9], channel modeling [10], and link design [11]. Since 1994, Kahn and his students at UC Berkeley have made numerous contributions, including experimental characterization of the nondirected indoor infrared channel [12][13], performance evaluation of modulation, coding, and equalization based on measured channels [14][15][16][17], and channel reuse strategies. Kavehrad and his students [18] at the University of Ottawa have worked on the measurement of the channel frequency response, the design of a diffuser, and diversity techniques. Researchers at Aveiro University [19] have worked on modeling and simulation for the indoor infrared channel.

TABLE 1-1: Examples of infrared links.

Affiliation	Year	Type	Modulation	Bit rate
IBM	1979	diffuse	BPSK	125 Kb/s
ATS	1985	directed, LOS	OOK	10 Mb/s
Spectrix	1987	diffuse	OOK	4 Mb/s
IrDA	1993	directed, LOS	4-PPM	4 Mb/s
IBM/Photonics	1993	diffuse	16-PPM	1 Mb/s
Apple	1994	directed, LOS	OOK	38.4 Kb/s
UC Berkeley	1994	diffuse	OOK	50 Mb/s
BT Lab	1994	directed, LOS	OOK	155 Mb/s

1.2 MODELING OF WIRELESS INFRARED CHANNEL

In a wireless optical system, the most practical modulation technique is intensity modulation, in which the information modulates the instantaneous power of the carrier. The most practical detection technique is direct detection, in which a photodetector produces a current proportional to the received instantaneous power. The model of the infrared channel with intensity modulation / direct detection (IM / DD) is illustrated in Fig. 1-2 [11]. The transmitted signal $x(t)$ is the instantaneous optical power of the infrared transmitter. The received signal $y(t)$ is the instantaneous current in the receiving photodetector, which is proportional to the integral of received optical power over the photodetector surface.

The appropriate channel model for wireless optical communications systems using IM / DD depends on the intensity of the background light. In low background light, it is common to model the received signal as a Poisson process with rate $\lambda_s(t) + \lambda_n$, where

$\lambda_s(t)$ is proportional to the instantaneous optical power of the received signal, and λ_n is proportional to the power of the background light. When λ_n is zero, the channel is quantum limited. However, the background light in typical indoor environments is very intense; even after a narrow-band (10 nm) optical filter, λ_n will be between 10^{11} and 10^{14} photons/s, depending on the proximity to a window [7]. Therefore, the photodetector shot noise is accurately modeled as an additive white Gaussian noise (AWGN) plus a d.c. offset [20]. Also, as illustrated in Fig. 1-2, typical detector areas are millions of square wavelengths, leading to spatial diversity that prevents multipath fading. Furthermore, because

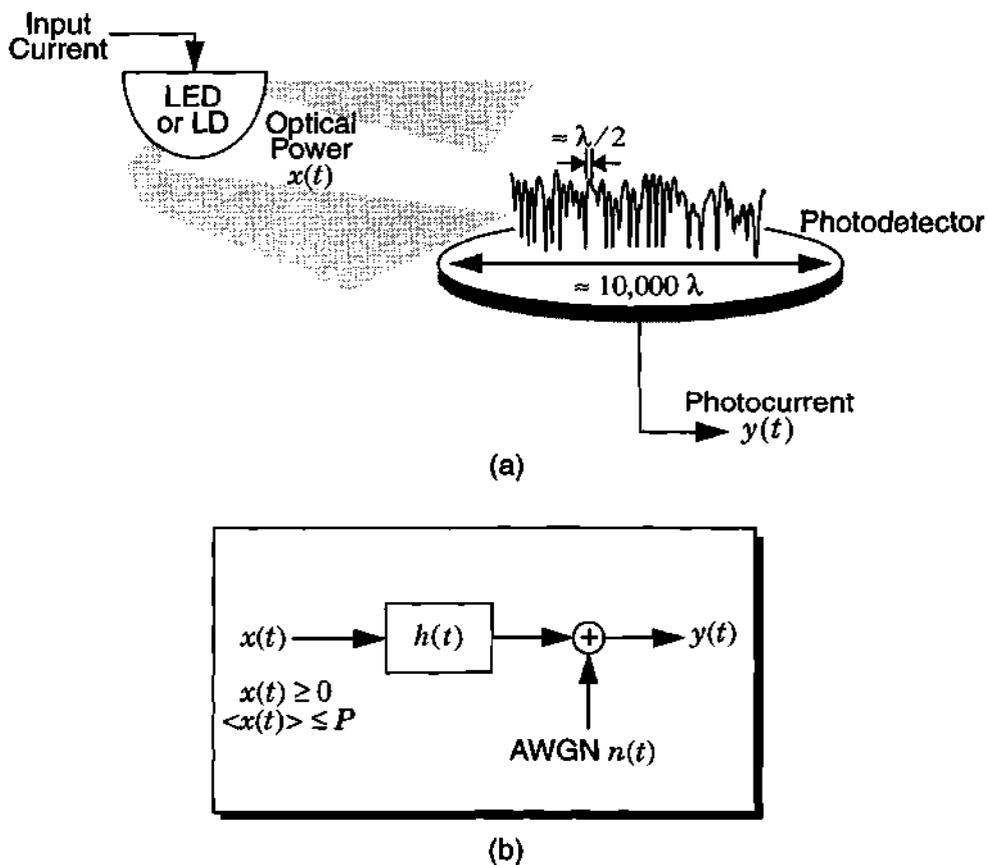


Fig. 1-2. (a) Transmission and reception in an infrared link with intensity modulation and direct detection, (b) modeling link as a baseband linear, time-invariant system having impulse response $h(t)$ with additive noise $n(t)$. [11]

the multipath propagation destroys spatial coherence, the effects of multipath propagation can be characterized by a baseband linear filter [12][13]. This leads to the following equivalent baseband channel model, the additive white Gaussian noise (AWGN) model, for wireless infrared communications using intensity modulation and direct detection as shown in Fig. 1-2-b [9]:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau) d\tau + n(t), \quad (1-1)$$

where $x(t)$ represents the instantaneous optical power of the transmitted signal, $y(t)$ represents the instantaneous current of the receiving photodetector, $h(t)$ represents the multipath-induced temporal dispersion, and $n(t)$ is white Gaussian noise. The same model (1-1) is used to represent conventional radio channels, where the input $x(t)$ represents amplitude of transmitted signal and must satisfy $\langle x^2(t) \rangle \leq P_o$, where P_o is the average power constraint of the radio transmitter.¹ However, $x(t)$ represents optical *power* in our application, so it must satisfy:

$$x(t) \geq 0 \text{ and } \langle x(t) \rangle \leq P, \quad (1-2)$$

where P is the average optical power constraint of the transmitter. These constraints dramatically alter the choice of modulation schemes.

In a diffuse infrared channel, the transmitted signal arrives at the receiver after multiple reflections from the ceiling, walls, and other reflecting objects. Since each reflecting surface has a reflectivity of less than one, the received signal undergoing multiple reflec-

1. $\langle \cdot \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (\cdot) dt$

tions has a smaller power than that of the transmitted signal. A natural model for the impulse response of the multipath channel is an exponential decay:

$$h(t) = We^{-Wt}u(t), \quad (1-3)$$

where W is the 3-dB bandwidth and $u(t)$ is unit-step function. Note that the channel has unity D.C gain. Carruthers *et al.* [13] derived the ceiling-bounce model for the wireless infrared channel. In the ceiling-bounce model, the impulse response due to diffuse reflection from a single infinite-plane reflector such as a large ceiling is:

$$h(t) = \frac{6a^6}{(t+a)^7} u(t), \quad (1-4)$$

where $a = 2H / c$, H is the height of the ceiling above the transmitter and the receiver, and c is the speed of light. Compared to (1-3), (1-4) provides a slightly better fit to the measured channel [13]. But, ceiling-bounce model assumes the transmitter and receiver are colocated. The exponential decay model is more general because it does not depend on the geometry of transmitter and receiver, and is characterized by one parameter, W . For example, the measured value of W for an empty conference room having dimension ($7.5 m \times 5.5 m \times 3.5 m$) is 34 MHz [9]. Hence, we will use this model from here on.

The unique characteristics of this channel model have motivated recent research in search of power-efficient and bandwidth-efficient modulation and coding techniques that are well-suited to the channel [4][6][8][9][14]-[17][21][28][30][63]. Many of the conventional digital communication results do not hold for this channel. For example, quadrature amplitude modulation (QAM), which is frequently used in conventional channels, is not suitable for the diffuse infrared channel because of its poor power efficiency. Similarly,

coded modulation schemes designed for the conventional channel may not perform well on the infrared channel. Instead, as will be shown in chapter 2, the constraint (1-2) favors modulation schemes with low duty cycle, such as pulse-position modulation.

1.3 THESIS OUTLINE

The goal of this thesis is to develop and analyze new power-efficient modulation, coding, and equalization schemes that are compatible with the infrared channel.

In chapter 2, we try to find some efficient uncoded modulation schemes on an ideal channel. First, we compare the performance of several candidate uncoded modulation techniques by calculating the required bandwidth and the average optical power required at a given bit rate to achieve a desired bit error rate. We also calculate the exact bandwidth of each modulation scheme from its power spectrum density. We derive a bound for the power and bandwidth efficiency of multiple-pulse position modulation.

In chapter 3, we take into account intersymbol interference (ISI) due to the multipath distortion. We compare the performance of several equalization techniques for the uncoded modulation schemes considered in chapter 2. First, we derive a vector channel model for MPPM on an ISI channel. We calculate the exact error probability when the receiver does not use equalization, and derive the Gaussian approximation for the special case of PPM. We then derive an upper bound for the error probability of the optimum maximum-likelihood sequence detection (MLSD) over a vector channel. To mitigate the effect of ISI with reduced complexity, we propose a partial-response precoding scheme and compare it with several symbol-by-symbol equalization schemes, including zero-

forcing block decision feedback equalization (ZF-BDFE), zero-forcing linear equalization (ZF-LE), and maximum *a posteriori* (MAP) detection.

In chapter 4, we consider the achievable information rate and cutoff rate for modulation schemes as a limit of performance. We present expressions for these parameters on an ideal channel when the input codewords are independent, identically uniform-distributed. We then derive lower and upper bounds for the information rate over an ISI channel. We also calculate the required optical power to achieve a specific cutoff rate.

In chapter 5, we design a new trellis codes so as to improve the overall power efficiency. First, we consider convolutional coded PPM and derive an upper bound for the minimum Euclidean distance. We then present an expression for the coding gain of trellis-coded OPPM. Finally, we present the design procedure and computer search results for new trellis codes based on MPPM. We derive an approximation for the minimum Euclidean distance of trellis coded MPPM and compare it to the well-known simplex bound.

In chapter 6, we evaluate the performance of the coded modulation schemes developed in chapter 5 on a multipath channel by employing superstate MLSD, LE with MLSD, DFE with MLSD, parallel decision feedback detection (PDFD), and the proposed precoding scheme with PDFD.

In chapter 7, we conclude our study with some interesting topics for future research.

CHAPTER 2

PERFORMANCE OF MODULATION SCHEMES

As we indicated in section 1.2, conventional modulation schemes do not work well under the constraints (1-1) and (1-2). As we will see, the probability of error depends on $x(t)^2$ rather than $x(t)$, and a signal set whose peak-power-to-average-power ratio is large offers better performance in terms of power efficiency. Therefore, in this chapter we concentrate on modulation schemes having a low duty-cycle. Most modulation schemes we consider here have been used for the photon-counting channel. However, due to the unique nature of the diffuse infrared channel, we cannot apply the results of photon-counting channel to our case directly.

In section 2.1, we calculate the power efficiency and bandwidth efficiency of several uncoded modulation schemes. In 2.2, we calculate the accurate bandwidth of each modulation scheme by calculating the power spectrum density. In 2.3, we derive the MPPM bound as a limit of performance of MPPM when the codeword length is arbitrarily large.

2.1 POWER EFFICIENCY AND BANDWIDTH EFFICIENCY

When we evaluate the performance of modulation schemes, the two most important criteria are power efficiency and bandwidth efficiency. The power efficiency corresponds to the required power to achieve a specific bit error rate, and the bandwidth efficiency corresponds to the bandwidth occupied to achieve a specific bit rate. In this section, we introduce several candidate modulation schemes and compare these schemes based on the two criteria.

2.1.1 Definitions

We first review the classic problem of determining the error probability for an L -ary modulation scheme in the presence of additive white Gaussian noise, assuming maximum-likelihood (ML) detection, and neglecting intersymbol interference [22][23]. The transmitter conveys information at a rate of R_b bits/second by transmitting one of L nonnegative signals $\{x_1(t), x_2(t), \dots, x_L(t)\}$ every $T = \log_2 L / R_b$ seconds, and the channel adds white Gaussian noise with power spectrum N_0 . To prevent intersymbol interference, each signal is confined to the interval $[0, T)$. The signal set satisfies (1-2) with equality, so that the average signal power is $\frac{1}{L} \sum_i \langle x_i(t) \rangle = P$. For example, an on-off-keying (OOK) transmitter emits a rectangular pulse of duration $1/R_b$ and of intensity $2P$ to signify a one bit,

and no pulse to signify a zero bit. The bandwidth required by OOK is roughly R_b , the inverse of the pulse width.

To simplify analysis, we make the high-SNR assumption that the probability of bit error is dominated by the two nearest signals, so that:

$$Pr[\text{bit error}] \approx Q(d_{min}/2\sqrt{N_0}), \quad (2-1)$$

where d_{min} is the minimum Euclidean distance between any pair of valid signals:

$$d_{min}^2 = \min_{i \neq j} \int (x_i(t) - x_j(t))^2 dt. \quad (2-2)$$

In fact, (2-1) is exact for OOK (and any time $L = 2$); the minimum distance between the two signals in the OOK signal set is:

$$d_{OOK} = \frac{2P}{\sqrt{R_b}}, \quad (2-3)$$

and the probability of bit error, assuming ML detection, is

$$Pr[\text{bit error}] = Q\left(\frac{P}{\sqrt{N_0 R_b}}\right). \quad (2-4)$$

We will use OOK as a benchmark to compare the power efficiencies of various modulation schemes. The power required by OOK to achieve a given bit error rate (BER) is $P_{OOK} \equiv \sqrt{N_0 R_b} Q^{-1}(BER)$. The power required by any other modulation scheme to achieve the same BER is approximately $P = (d_{OOK}/d_{min})P_{OOK}$, assuming the SNR is high enough that (2-1) is accurate. Therefore, in the remainder of the chapter we will use the distance ratio d_{OOK}/d_{min} to characterize the power requirement of any modulation scheme.

2.1.2 Pulse-position Modulation (PPM)

Pierce first proposed communication using *pulse-position modulation* (PPM) with direct detection by photon counting at the receiver [24]. In a PPM scheme, each symbol interval of duration $T = \log_2 L / R_b$ is partitioned into L sub-intervals, or chips, each of duration T/L , and the transmitter sends an optical pulse during one and only one of these chips. Fig. 2-1 shows an example of 4-PPM. PPM is similar to L -ary FSK, in that all signals are orthogonal and have equal energy. PPM can be viewed as the rate- $\log_2 L / L$ block code consisting of all binary L -tuples having unity Hamming weight. A PPM signal satisfying (1-2) is:

$$x(t) = LP \sum_{k=0}^{L-1} c_k p(t - kT/L), \quad (2-5)$$

where $[c_0, c_1, \dots, c_{L-1}]$ is the PPM codeword, and where $p(t)$ is a rectangular pulse of duration T/L and unity height. All of the signals are equidistant, with:

$$d_{min}^2 = \min_{i \neq j} \int (x_i(t) - x_j(t))^2 dt = 2LP^2 \log_2 L / R_b. \quad (2-6)$$

Therefore, the average power requirement is approximately [9]:

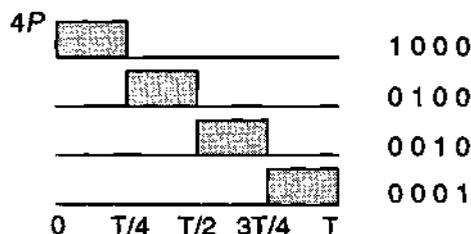


Fig. 2-1. Four pulse-position modulation (PPM).

$$P_{PPM}/P_{OOK} \approx d_{OOK}/d_{min} = \sqrt{\frac{2}{L \log_2 L}}. \quad (2-7)$$

From (2-7) we see that, for any L greater than 2, PPM requires less optical power than OOK. In principle, the optical power requirement can be made arbitrarily small by making L suitably large, at the expense of increased bandwidth; the bandwidth required by PPM to achieve a bit rate of R_b is approximately the inverse of one chip duration, $B = L/T = LR_b/\log_2 L$:

$$B_{PPM}/R_b = \frac{L}{\log_2 L}. \quad (2-8)$$

2.1.3 Multiple-Pulse Position Modulation (MPPM)

One generalization of PPM is *multiple-pulse position modulation* (MPPM), suggested by Sugiyama and Nosu [25]. In MPPM, each symbol interval of duration $T = \log_2 L/R_b$ is partitioned into n chips, each of duration T/n , and the transmitter sends an optical pulse during w of these chips. Fig. 2-2 shows an example of $\binom{4}{2}$ -MPPM. The transmitted signal is given by:

$$x(t) = a \sum_{k=0}^{n-1} c_k \phi(t - kT/n), \quad (2-9)$$

where $[c_0, c_1, \dots, c_{n-1}]$ is a binary n -tuple of weight w , where $\phi(t) = \sqrt{n/T} p(t)$ is a unit-energy rectangular pulse of duration T/n , and where the constant a is chosen so that the average optical power is P : $a = (P/w) \sqrt{nT} = d_{OOK} \sqrt{n \log_2 L} / 2w$. There are $\binom{n}{w}$ binary n -tuples of weight w , but it may be desirable to use only a fraction L of these; for example,

we may choose the codewords to have a large minimum Hamming distance d . That is, we may restrict attention to an (n, d, w) constant weight code [26][27], which is a set of binary n -tuples having weight w and minimum Hamming distance d .

For a given n , d , and w , let $L \leq \binom{n}{w}$ be the number of valid codewords. We must have $d \geq 2$, because it is impossible for two binary n -tuples of weight w to differ in only one position. If we admit all binary n -tuples of weight w , then $L = \binom{n}{w}$ and $d = 2$. The bandwidth is roughly n/T , the inverse of the chip duration, so that [28]:

$$B_{MPPM}/R_b = \frac{n}{\log_2 L}. \quad (2-10)$$

Because $\{\phi(t - kT/n)\}$ is an orthonormal set, (2-9) implies that the Euclidean distance between any two MPPM waveforms $x_i(t)$ and $x_j(t)$ is $a\sqrt{d_{ij}}$, where d_{ij} is the Hamming distance between the corresponding binary n -tuples. Thus, the minimum distance is $d_{min} = a\sqrt{d}$, where d is the minimum Hamming distance and $a = d_{OOK} \sqrt{n \log_2 L} / 2w$. The ratio of d_{OOK} to d_{min} gives the average power requirement:

$$P_{MPPM}/P_{OOK} = \frac{2w}{\sqrt{nd \log_2 L}}. \quad (2-11)$$

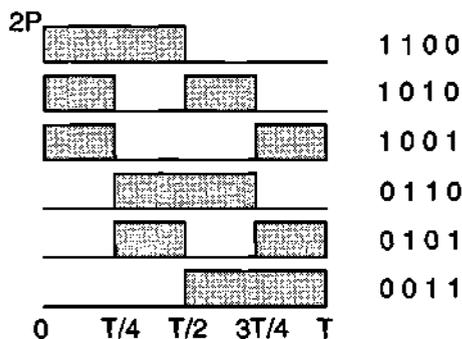


Fig. 2-2. The $\binom{4}{2}$ multiple-pulse position modulation.

Note that PPM is a special case of MPPM with $n = L$, $d = 2$, and $w = 1$, and that (2-11) reduces to (2-7) in this case.

$\binom{n}{w}$ -MPPM can be viewed as a binary *permutation modulation*, introduced by Slepian [29]. All codewords are generated by permuting the initial vector \mathbf{c}_0 , which can be represented in general form:

$$\mathbf{c}_0 = \left[\underbrace{000\dots 0}_{n-w} \quad \underbrace{111\dots 1}_w \right]^T. \quad (2-12)$$

The other codewords are obtained by permuting the order in all possible ways, and the number of codewords is:

$$L = \binom{n}{w} = \frac{n!}{w!(n-w)!}. \quad (2-13)$$

Since all codewords have the same length and weight, they lie on a sphere with a radius of square-root of signal energy. For example, $\binom{4}{2}$ -MPPM codewords can be obtained by permuting the initial codeword $\mathbf{c}_0 = [0011]^T$:

$$\{[0011]^T, [0101]^T, [0110]^T, [1001]^T, [1010]^T, [1100]^T\}. \quad (2-14)$$

The receiver on an ideal channel decides on the codeword \mathbf{c}_l that maximizes the correlation [29][30]:

$$\Lambda_l = \mathbf{c}_l^T \mathbf{y}_k \quad \text{for } l = 0, \dots, L-1. \quad (2-15)$$

Note that PPM is also a permutation modulation with $w = 1$.

Permutation modulation is a special case of a *group code* [31]. A code is defined as a group code if its codewords are generated by multiplication of an initial vector \mathbf{c}_0 with $n \times n$ orthogonal matrices $\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_g$, and this collection of matrices forms a group Γ

under matrix multiplication. In other words, there is a member in Γ that will map any given codeword \mathbf{c}_i into any other codeword \mathbf{c}_j . In general, the order g of the group is greater than the number of codewords, M , and $g = n!$ for a permutation modulation.

A group code is a set of vectors with complete symmetry, that is, all codewords have (a) the same error probability, (b) the same set of distances to the other codewords, and (c) the same energy. In general, the error probability of MPPM is given by [30] (see section 3.3.3):

$$Pr[\text{error}] \leq \sum_{k=1}^w N_k Q\left(\sqrt{\frac{ks^2}{2N_0}}\right), \quad (2-16)$$

where $N_k = \binom{w}{k} \binom{n-k}{k}^w$ is the number of codewords with mutual distance $2k$ and $s = (P/w) \sqrt{n \log_2 L / R_b}$. Note that MPPM satisfies all the properties of a group code.

PPM is also a *cyclic* group code because group Γ is the set of matrices whose elements are the powers of a generator matrix \mathbf{O} . In other words, codewords are obtained by:

$$\mathbf{c}_i = \mathbf{O}^i \mathbf{c}_0 \text{ for } i = 0, 1, \dots, L-1, \quad (2-17)$$

where $\mathbf{O} = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}$.

2.1.4 Overlapping Pulse-Position Modulation (OPPM)

We define $\binom{n}{w}$ *overlapping pulse-position modulation* (OPPM) code as a subset of MPPM code, where the w ones are constrained to be consecutive. In other words, each

symbol interval of duration $T = \log_2 L / R_b$ is divided into n chips, each of duration T/n , and a rectangular pulse spanning w chips is transmitted, beginning at any of the first $L = n - w + 1$ chips. For example, among the six $\binom{4}{2}$ -MPPM codewords, we choose the three OPPM codewords in which the two ones are consecutive, as shown in Fig. 2-3. The motivation for constraining the w ones to be consecutive is the decreased bandwidth that results; unfortunately, this benefit is offset by the reduced alphabet size, because L drops from $\binom{n}{w}$ to $n - w + 1$. Note that this definition of OPPM is slightly more general than the usual definition [32], because it allows the possibility that n/w is not an integer. We refer to the ratio $\alpha = w/n$ as the duty cycle. Note also that specifying L does not uniquely specify n and w ; for example, 4-OPPM can arise from $\binom{5}{2}$, $\binom{6}{3}$, $\binom{7}{4}$, etc. Thus, it takes two parameters to specify OPPM, either n and w or L and α .

The bandwidth of OPPM is $n/(wT)$, where $T = \log_2 L / R_b$, so that:

$$B_{\text{OPPM}}/R_b = \frac{n/w}{\log_2(n-w+1)}, \quad (2-18)$$

which is clearly smaller than that of PPM, since n/w is less than L . The minimum Hamming distance between OPPM codewords is 2, so that the minimum Euclidean distance

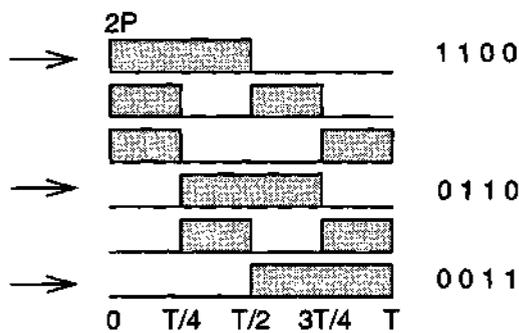


Fig. 2-3. Choose overlapping pulse-position modulation (OPPM) codewords from MPPM codewords.

between received signals is $d_{min} = \sqrt{2} a = (P/w)\sqrt{2nT}$. Dividing d_{OOK} by d_{min} yields the average power requirement for OPPM:

$$P_{OPPM}/P_{OOK} = \frac{2w}{\sqrt{2n \log_2(n-w+1)}}. \quad (2-19)$$

With $w = 1$, n becomes L , and this equation reduces to (2-7).

2.1.5 Pulse Amplitude and Position Modulation (PAPM)

Another way of increasing the throughput of PPM is by amplitude-modulating $p(t)$ in (2-5). The information is conveyed by the amplitude as well as the position of pulse. The advantage of this modulation is to increase the number of information bits without increasing the bandwidth and to maintain the low duty cycle property of PPM. This leads to a combination of pulse amplitude modulation (PAM) and PPM. Since PAM is a bandwidth efficient modulation [9] and PPM is a power efficient one, the combination of PAM and PPM enjoys the advantages of both schemes. We define n as number of slots and M as number of levels. The M - n -pulse amplitude and position modulation (M - n -PAPM) signal set is given by:

$$x(t) = nAP \sum_{k=0}^{n-1} c_k p(t - kT/n), \quad (2-20)$$

where $A \in \{1/M, 3/M, \dots, (2M-1)/M\}$, and $[c_0, c_1, \dots, c_{L-1}]$ is the PPM codeword. Fig. 2-4 shows an example of 2-4-PAPM. For each pulse slot of 4-PPM, there is 2 levels, and the number of codewords is 8. In general, the number of codewords L for n - M -PAPM

is equal to nM . The bandwidth of n - M -PAPM is same as that of n -PPM, so that the normalized bandwidth requirement:

$$B_{PAPM} / R_b = \frac{n}{\log_2 n M}, \quad (2-21)$$

which is smaller than that of PPM. The minimum Euclidean distance of the received signal is $d_{min} = \sqrt{2nT}/M$, where $T = \log_2 L / R_b$. Dividing d_{OOK} by d_{min} provides the normalized power requirement for PAPM:

$$P_{PAPM} / P_{OOK} = \sqrt{\frac{2M^2}{n \log_2 n M}}. \quad (2-22)$$

Note that when $M = 1$, n becomes L , and (2-22) reduces to (2-7).

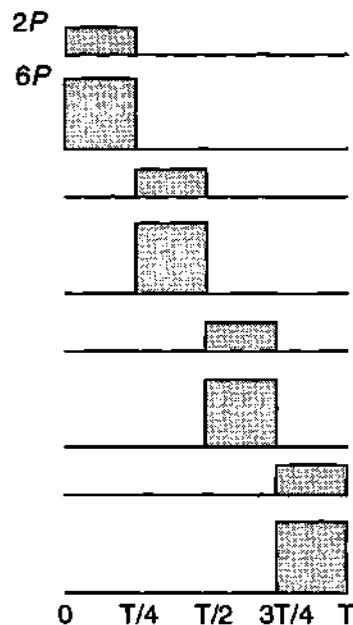


Fig. 2-4. The 2-4-pulse amplitude and position modulation.

2.1.6 Discussion

The bandwidth and power efficiency for various modulation schemes on the AWGN channel are shown in Fig. 2-5. The benchmark modulation OOK is marked with the symbol 'x'. PPM requires less power as L increases, but its bandwidth increases as well. MPPM with weight 2 outperforms PPM both in terms of bandwidth efficiency and power efficiency. MPPM with weight 8 is even more bandwidth efficient, but it requires a large number of chips n to achieve good power efficiency. OPPM with a duty cycle of $\alpha = 1/2$ is extremely bandwidth efficient. Decreasing the duty cycle to $\alpha = 1/4$ increases the power efficiency at the expense of bandwidth. $2-n$ -PAPM is more bandwidth-efficient but less power-efficient than L -PPM.

2.2 ACCURATE BANDWIDTH CALCULATION

There is no single universal definition of bandwidth [33]. The bandwidth of an MPPM modulation scheme can be roughly approximated by the inverse of the shortest pulse width that corresponds to the width of the main spectral lobe. This definition is a simple and popular measure of signal bandwidth. Thus, the bandwidth required by OOK is roughly the bit rate R_b , and the bandwidth required by MPPM, OPPM and PPM to achieve a bit rate R_b is the inverse of the duration of the shortest pulse, or $nR_b/\log_2 L$, $(n/w)R_b/\log_2 L$, and $LR_b/\log_2 L$, respectively. For example, $\binom{5}{2}$ -MPPM, $\binom{8}{2}$ -OPPM and 8-PPM require 1.5, 1.4, and 2.7 times more bandwidth than OOK, respectively. More accurately, it is common to specify the bandwidth B_x that includes $x\%$ of the signal power [34][35]. Under the assumption that the codewords are chosen independently and with equal probability, a general expression for the power spectral density (PSD) of any L -ary modulation scheme is given in [35]:

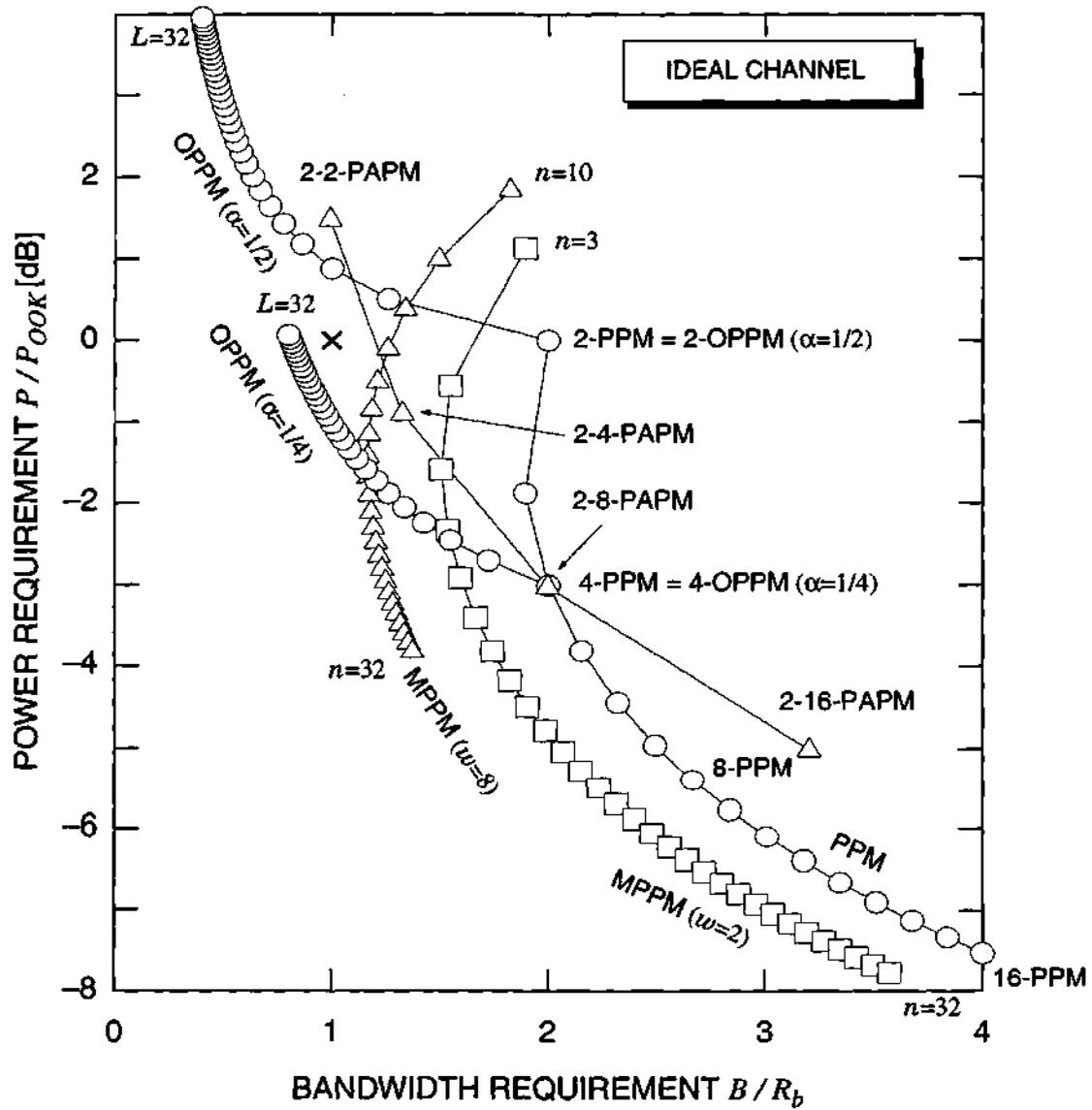


Fig. 2-5. Power and bandwidth efficiency of uncoded modulation schemes on an ideal channel.

$$S_X(f) = \frac{1}{LT} \sum_{l=0}^{L-1} |P_l(f)|^2 - \frac{1}{L^2T} \left| \sum_{l=0}^{L-1} P_l(f) \right|^2 \left[1 - \frac{1}{T} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T}\right) \right], \quad (2-23)$$

where $P_l(f)$ is the Fourier transform of the signal corresponding to the l -th codeword.

Applying this result to L -PPM, $\binom{n}{2}$ -MPPM, and $\binom{n}{w}$ -OPPM yields the following expressions for the continuous part of the PSD, respectively:

$$S_{PPM}(f) = \frac{|P(f)|^2}{T} \left[1 - \left| \frac{1}{L} \frac{\sin(\pi f T)}{\sin(\pi f T / L)} \right|^2 \right], \quad (2-24)$$

$$S_{MPPM}(f) = \frac{|P(f)|^2}{T} \left[\frac{1}{L} \sum_{k=1}^{n-1} (n-k) \left| 1 + e^{-j2\pi k f T / n} \right|^2 - \left| \frac{n-1}{L} \frac{\sin(\pi f T)}{\sin(\pi f T / L)} \right|^2 \right], \quad (2-25)$$

$$S_{OPPM}(f) = \frac{|P(f)|^2}{T} \left(1 - \left| \frac{1}{L} \frac{\sin(L\pi f T / n)}{\sin(\pi f T / n)} \right|^2 \right), \quad (2-26)$$

where $P(f)$ is Fourier transform of transmitter filter $p(t)$.

In Fig. 2-6, we plot the continuous part of the PSD of 4, 8-PPM, $\binom{12}{2}$, $\binom{17}{2}$ -MPPM, and $\binom{5}{2}$, $\binom{6}{3}$ -OPPM, using the results of (2-24) - (2-26). The results for the PPM are shown in Fig. 2-6-a. where the first null of the spectrum, which corresponds to our approximation of the inverse of shortest pulse-width, is close to B_{90} . For 4-PPM, the first null is at $4/T$ and $B_{90} = 4.3/T$, and for 8-PPM, the first null and B_{90} are both at $8/T$. For both $\binom{12}{2}$ -MPPM and $\binom{17}{2}$ -MPPM, B_{90} nearly coincides with the first null of the PSD as shown in Fig. 2-6-b. However, the first-null approximation is less accurate for OPPM as shown in Fig. 2-6-c. For $\binom{5}{2}$ -OPPM, the first null is at $2.5/T$ whereas $B_{90} = 3.5/T$, and for $\binom{6}{3}$ -OPPM, the first null is at $2/T$ whereas $B_{90} = 3.2/T$. Thus, the bandwidth of OPPM is not accurately approximated by the inverse of the pulse duration especially at high duty cycles.

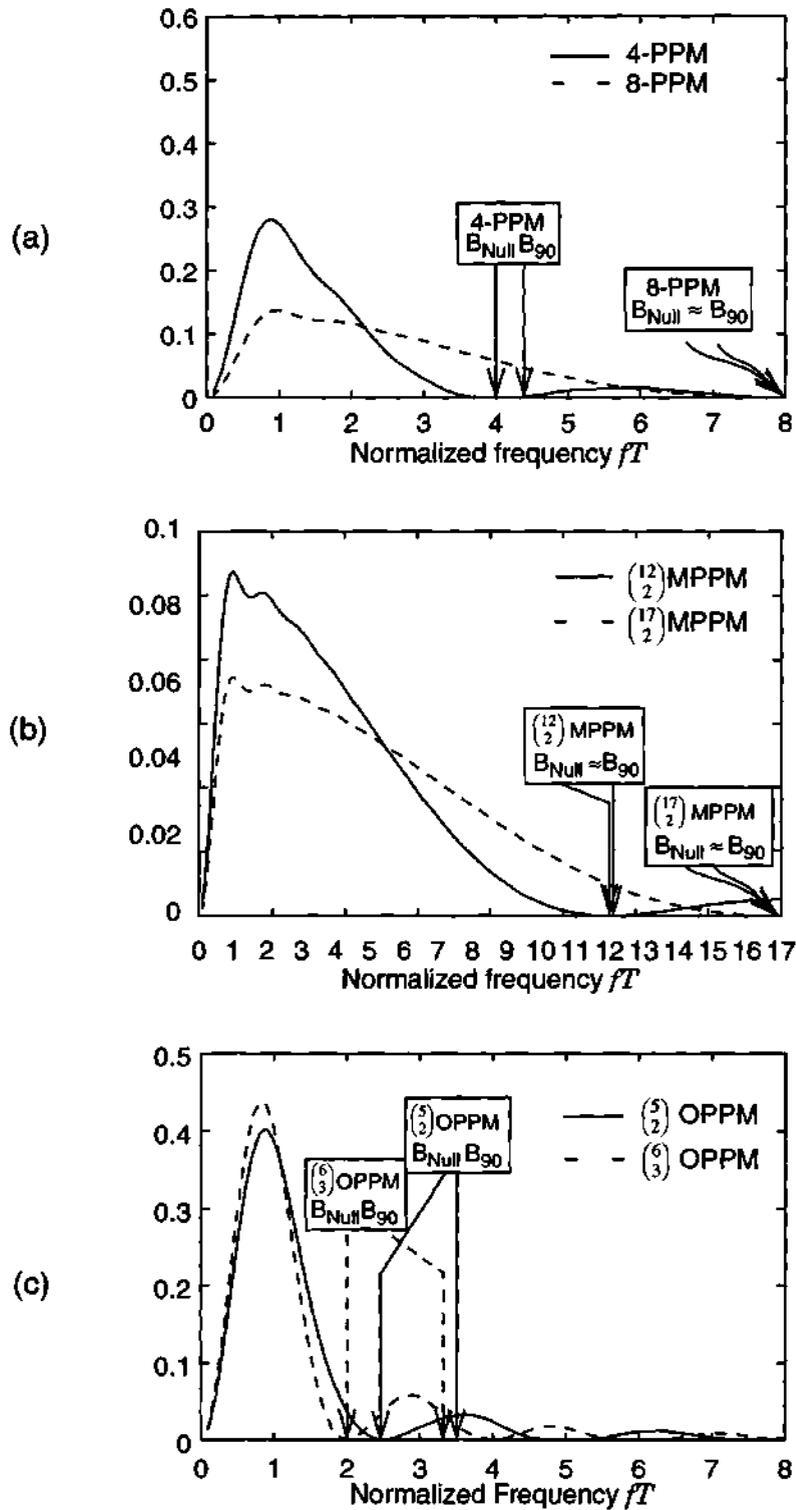


Fig. 2-6. Power spectral density of modulation schemes (a) PPM, (b) MPPM, (c) OPPM.

Observe that the sidelobe of the PSD for $(\frac{6}{3})$ -OPPM is larger than that for $(\frac{5}{2})$ -OPPM. To illustrate this result, consider a $(\frac{6}{3})$ -OPPM signal when the codeword (000111) is followed by the codeword (011100), producing a chip sequence of (000111011100). The isolated 0 chip in position 7 requires more bandwidth than that predicted by the inverse of the pulse width.

2.3 MPPM BOUND

In Fig. 2-5, the performance of modulation schemes using the finite length of codeword was shown. In this section, we calculate the performance of MPPM when we increase the codeword length arbitrarily large. We use Stirling's approximation [36]:

$$\binom{n}{w} \rightarrow 2^{nh\left(\frac{w}{n}\right)} \text{ as } n \rightarrow \infty, \quad (2-27)$$

where $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function shown in Fig. 2-7.

We can find a bound on the power and bandwidth efficiency of MPPM by applying (2-27) to (2-10) and (2-11):

$$P_{MPPM}/P_{OOK} \rightarrow \frac{\alpha}{\sqrt{h(\alpha)/2}}, \quad (2-28)$$

$$B_{MPPM}/R_b \rightarrow \frac{1}{h(\alpha)}, \quad (2-29)$$

as $n \rightarrow \infty$ and $\alpha = w/n$. We call (2-28) and (2-29) the MPPM bound, because these equations represent the limit in the performance of power efficiency and bandwidth efficiency for all $(\frac{n}{w})$ -MPPM. We plot the power and bandwidth efficiency of various $(\frac{n}{w})$ -MPPM with the MPPM bound in Fig. 2-8. Note that MPPM with weight 2 outperforms PPM both

in terms of both bandwidth efficiency and power efficiency. MPPM with more weight is even more bandwidth efficient, but it requires a large number of chips n to achieve good power efficiency. The MPPM bound shows that as $\alpha \rightarrow 1/2$, $P_{MPPM} / P_{OOK} \rightarrow 1/2$ (-1.5 dB) and $B_{MPPM} / R_b \rightarrow 1$. As $\alpha \rightarrow 0$, $P_{MPPM} / P_{OOK} \rightarrow 0$ and $B_{MPPM} / R_b \rightarrow \infty$.

2.4 SUMMARY AND CONCLUSIONS

We have examined the performance of uncoded modulation schemes by calculating their power efficiency and bandwidth efficiency. We have shown that MPPM and PPM are power-efficient modulations and OOK and OPPM are bandwidth-efficient. We calculated the accurate bandwidth of each modulation scheme by calculating the power spectrum density, and compared this result with our first-null approximation. Our approximation is quite accurate for most modulation schemes considered except for high-duty cycle OPPM. To understand the behavior of MPPM, we derived a bound for the power and bandwidth efficiency of MPPM as the codeword length $n \rightarrow \infty$. The MPPM bound implies that using

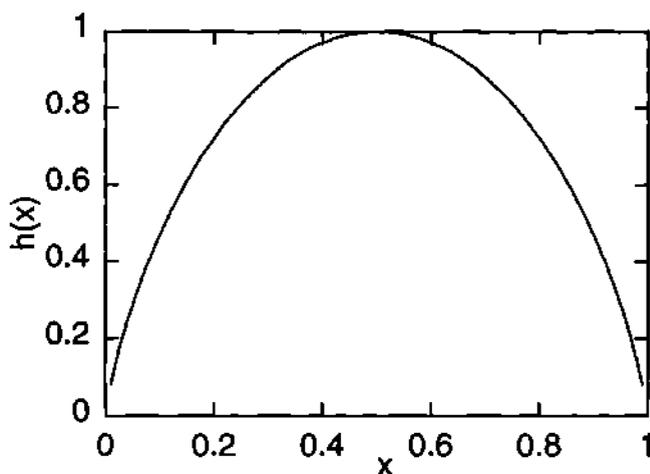


Fig. 2-7. Binary entropy function.

certain MPPM with infinite codeword length, at least theoretically we can achieve the same bandwidth as OOK with 1.5 dB less power, or achieve an arbitrarily low power requirement at the expense of infinite bandwidth.

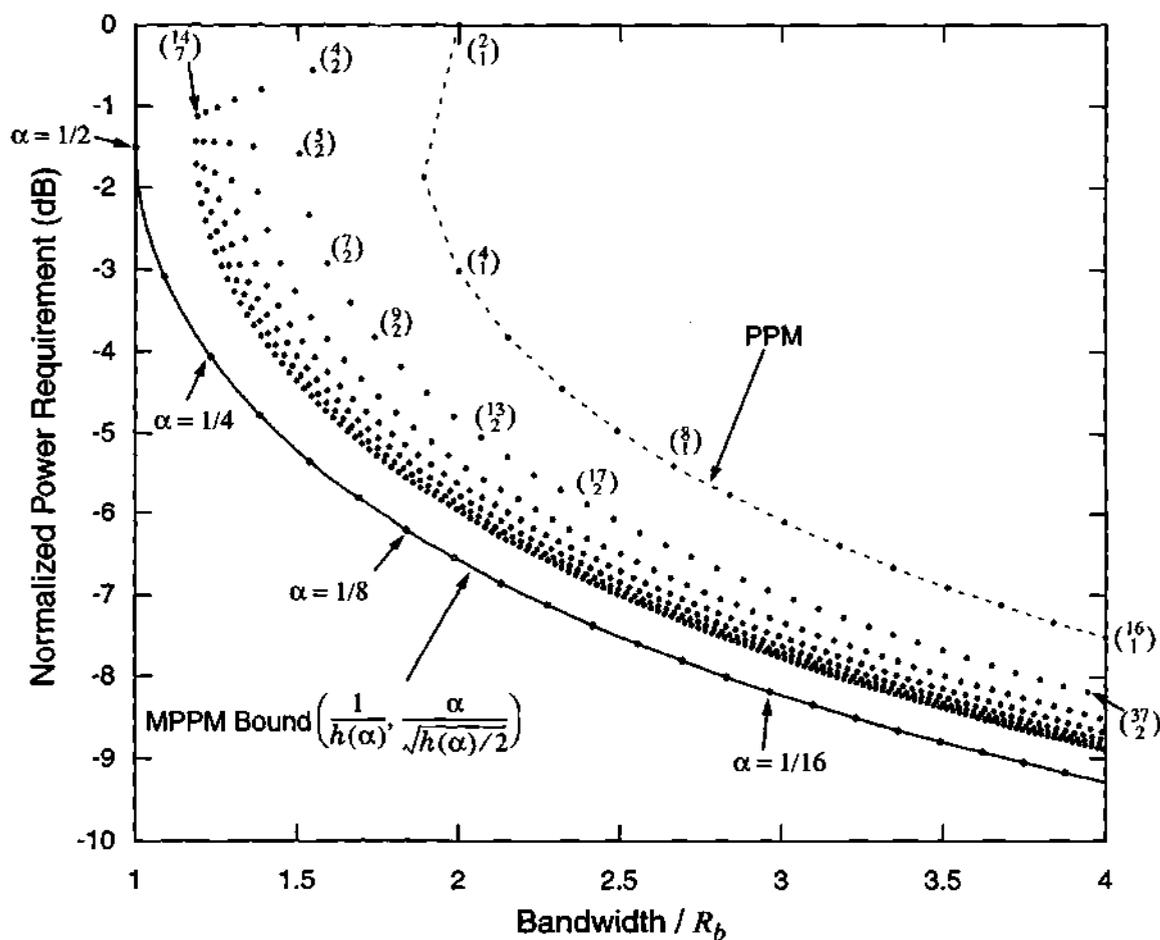


Fig. 2-8. Power efficiency and bandwidth efficiency of $\binom{n}{w}$ -MPPM and MPPM bound; $\alpha = w/n$.

CHAPTER 3

EQUALIZATION

In the previous chapter, we examined the performance of uncoded modulation schemes on an ideal channel, where additive white Gaussian noise (AWGN) is the primary source of performance degradation. In this chapter, we will include the effects of multipath distortion which degrade the performance further. We will consider several strategies to mitigate the effect of this multipath distortion.

3.1 INTRODUCTION

In the diffuse link, from transmitter to receiver the signal experiences temporal dispersion due to reflections. Usually, the received pulse duration is much larger than the symbol duration, especially at high bit rates, and there is a cross-talk or spill-over from one

symbol to another. This intersymbol interference (ISI) degrades the performance of the receiver. Equalization is the mitigation of ISI through signal processing. We consider three equalization techniques: maximum-likelihood sequence detection (MLSD), symbol-by-symbol equalization, and precoding.

3.1.1 Maximum-Likelihood Sequence Detection (MLSD)

Forney [37][38] showed that the optimum detector for linear PAM signals in the presence of ISI consists of a whitened-matched filter (WMF) followed by a Viterbi detector. This method is referred to as maximum-likelihood sequence detection (MLSD). The complexity of the MLSD grows exponentially with the channel memory. In reduced state sequence detection (RSSD) [39][40], the complexity of MLSD can be reduced by forming groups of states and tracking only one surviving path per group of states in the Viterbi algorithm. In delayed decision-feedback equalization (DDFE) [41], the complexity of MLSD is reduced by considering a few states of the channel. The ISI due to the remainder of the states is estimated using a feedback detection analogous to that of decision feedback equalizer (DFE). As in the DFE, error propagation affects this algorithm. The price paid for these methods for reduction in complexity is a performance degradation due to the reduction of minimum distance. Barry [21] and Audeh *et al.* [15] used MLSD to detect PPM on a multipath channel.

3.1.2 Symbol-by-symbol Equalization

Symbol-by-symbol equalization is a suboptimal strategy for detection in the presence of intersymbol interference. Its primary advantage over maximum-likelihood sequence detection is a reduction in complexity. Two basic approaches to symbol-by-symbol equal-

ization are used: linear equalization (LE) and decision feedback equalization (DFE). The equalization structure can be derived under the zero-forcing (ZF) criterion, which completely removes the ISI, or under the minimum mean-square error (MMSE) criterion, which weighs both ISI and noise. In ZF-LE, we choose an equalizer after a sampled matched filter as the inverse of the folded spectrum of the channel. If the equalizer simply inverts the channel, the receiver would enhance the noise over frequency regions where the channel has a null. This phenomenon is called noise enhancement. The ZF-DFE recreates the postcursor ISI from its data decision and subtracts the result from the incoming signal. One obvious potential problem with the DFE is that any decision error will cause a corrupted estimate of the ISI. The result is that a single error causes a reduction in the margin against noise for a number of future decisions. This phenomenon is called error propagation. Price [42] first observed that the SNR gap to capacity at high signal-to-noise ratio is the same for channels with ISI as it is for ideal channels, as long as ZF-DFE is used at the receiver. In MMSE-LE, we are willing to accept more ISI after equalization to reduce the noise enhancement. While a ZF-DFE forces the ISI to zero at the slicer input, a MMSE-DFE minimizes the variance of the slicer error. Barry [21] and Audeh *et al.* [16] proposed symbol-rate and chip-rate ZF-DFE for PPM.

Another type of symbol-by-symbol equalization is based on maximizing the *a posteriori* probabilities (MAP detection) developed by Abend and Fritchman [43]. MAP detection is optimum in the sense of minimizing the probability of symbol error. Simulation results indicate that the performance of MAP detector is superior to that of DFE and comparable to that of MLSD [22]. But unlike other equalization schemes, MAP detection requires the knowledge of the statistics of the noise. It has found little application in prac-

tice due to its high complexity. Recently, Williamson *et al.* [44] developed a MAP detector based on DFE (MAP-DFE). Although more complex than conventional DFE, MAP-DFE outperforms DFE.

3.1.3 Precoding

Precoding is a technique similar to DFE that eliminates error propagation by moving cancellation of the postcursor ISI from the receiver to the transmitter. However, this requires knowledge of the channel response at the transmitter, a requirement that is compatible only with channels that are stationary or slowly time-varying. Simply performing linear equalization at the transmitter is not practical because it increases both average and peak power of the transmitter signal. It also violates the nonnegative constraint of (1-2). Tomlinson and Harashima [45][46] used a modulo operation to reduce these power penalties (TH precoding). Recently Laroia, Tretter, and Farvadin (LTF) [47][48] introduced a new precoding scheme referred to as LTF precoding. In LTF precoding, the postcursor ISI is quantized to the nearest point in predetermined constellations, and only the difference between the data signal and the quantization error is transmitted. The advantage of LTF precoding scheme is that it supports constellation shaping on ISI channels.

In fact, with uncoded systems, precoding has not received much attention in practice. Its performance is no better than that of ZF-DFE under the ideal decision assumption. As a result, DFE has generally been preferred for uncoded systems over precoding because it does not require information about the channel at the transmitter. We remark that the TH precoding and LTF precoding are incompatible with the constraint (1-2).

3.1.4 Outline

The remainder of this chapter is organized as follows. In section 3.2, we develop a vector channel model for MPPM on an ISI as a generalization of [21]. In section 3.3, we analyze the performance of an unequalized receiver, and in section 3.4, we analyze the performance of maximum likelihood sequence detection. In section 3.5, we consider three symbol-by-symbol equalization strategies: zero-forcing block decision equalizer, zero-forcing linear equalizer, and maximum *a posteriori* detector. In section 3.6, we propose a partial-response precoding scheme to be compatible with the infrared channel as a generalization of LTF precoding without considering shaping gain.

3.2 SYSTEM MODEL

Consider the system model shown in Fig. 3-1-a. Information bits with rate R_b (b/s) enter the encoder, which maps each block of $\log_2 L$ bits into one of L -MPPM codewords $c_0 \dots c_{L-1}$. We consider only binary codewords of length n and Hamming weight w . The

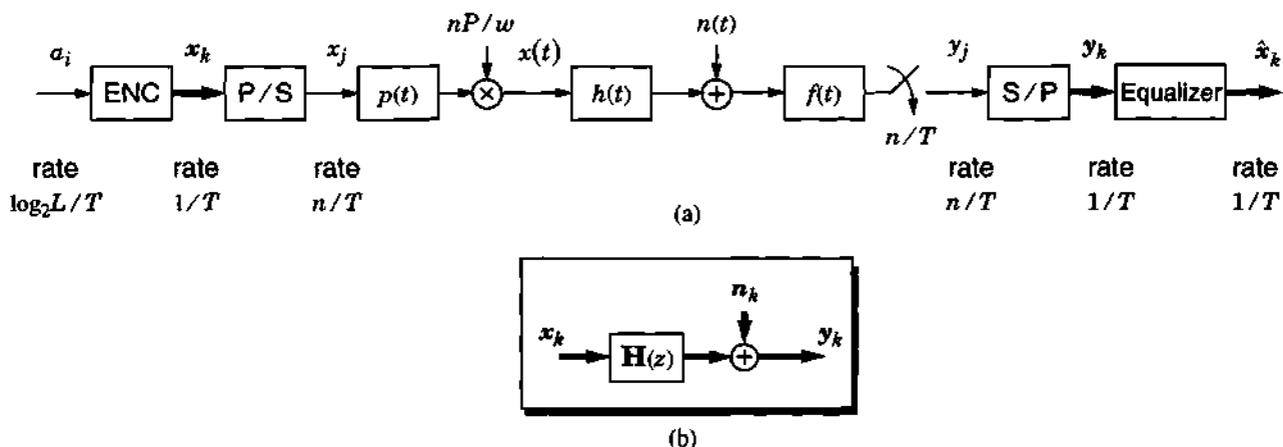


Fig. 3-1. (a) Block diagram of MPPM system, (b) equivalent vector channel.

output of the encoder is a sequence of codewords $\{\mathbf{x}_k\}$ with rate $1/T = R_b/\log_2 L$. This sequence is serialized to produce the binary chip sequence $\{x_j\}$ with rate n/T , where $\mathbf{x}_k = [x_{kn}, x_{kn+1}, \dots, x_{kn+n-1}]^T$. The binary chip sequence drives a transmitter filter with a rectangular pulse shape $p(t)$ of duration T/n and unity height. To satisfy the power constraint of (1-2), the filter output is multiplied by (nP/w) before the signal is sent across the channel.

As we explained in section 1.2, the channel model for diffuse link without fading is an exponential decay model given by:

$$h(t) = We^{-Wt}u(t), \quad (3-1)$$

where W is a 3-dB bandwidth and $u(t)$ is the unit-step function.

As shown in Fig. 3-1-a, the receiver uses a unit-energy filter $f(t)$ and samples the output at the chip rate n/T producing y_j . The receiver groups the samples y_j into blocks of length n , producing a sequence of observation vectors $\{\mathbf{y}_k\}$, where $\mathbf{y}_k = [y_{kn}, y_{kn+1}, \dots, y_{kn+n-1}]^T$. The receiver passes each observation vector through an equalizer to form an estimate $\hat{\mathbf{x}}_k$ of \mathbf{x}_k . For the symbol-by-symbol equalizer, the decision device is memoryless and has no decoding delay, however, for the maximum-likelihood sequence detector, the decision device will have memory and a decoding delay.

The equivalent discrete-time channel between transmitted and received chips is:

$$y_j = \sum_{i=-\infty}^{\infty} h_i x_{j-i} + n_j = s_j + n_j, \quad (3-2)$$

where h_j is the equivalent chip-rate impulse response:

$$h_j = \frac{nP}{w} \left(p(t) * h(t) * f(t) \right) \Big|_{t=jT/n}, \quad (3-3)$$

and where s_j is defined by (3-2).

We assume that $f(t)$ has unit energy and is the whitened-matched filter. In this case, the noise samples $\{n_j\}$ will be independent zero-mean Gaussian random variables with variance N_0 . As shown in Fig. 3-1-b, the equivalent vector channel between transmitted codewords \mathbf{x}_k and observation vectors \mathbf{y}_k is given by:

$$\mathbf{y}_k = \sum_{j=0}^{\infty} \mathbf{H}_j \mathbf{x}_{k-j} + \mathbf{n}_k = \mathbf{s}_k + \mathbf{n}_k, \quad (3-4)$$

where the channel impulse response is a Toeplitz sequence \mathbf{H}_k , with $[\mathbf{H}_k]_{ij} = h_{kn+i-j}$, the signal component is $\mathbf{s}_k = [s_{kn}, s_{kn+1}, \dots, s_{kn+n-1}]^T$, the noise component is $\mathbf{n}_k = [n_{kn}, n_{kn+1}, \dots, n_{kn+n-1}]^T$. To simplify the analysis, we will consider only μ nonzero terms in the impulse response $\{\mathbf{H}_k\}$ in (3-4).

3.3 UNEQUALIZED RECEIVER

3.3.1 Error Probability for MPPM and OPPM

By definition, the unequalized receiver uses the decision device that would be optimal were there no ISI. In other words, it decides on the codeword \mathbf{c}_l that maximizes the correlation:

$$\Lambda_l = \mathbf{c}_l^T \mathbf{y}_k \quad \text{for } l = 0, \dots, L-1. \quad (3-5)$$

This unequalized decision rule is illustrated in Fig. 3-2. If we knew that $\mathbf{x}_k = \mathbf{c}_i$, and if we knew all past codewords $\mathbf{X}' = \{\dots, \mathbf{x}_{k-2}, \mathbf{x}_{k-1}\}$, then the probability of error $\hat{\mathbf{x}}_k \neq \mathbf{x}_k$ could be bounded using the union bound:

$$\begin{aligned} Pr[\text{error} \mid \mathbf{x}_k = \mathbf{c}_i, \mathbf{X}'] &\leq \sum_{j=0, j \neq i}^{L-1} Pr[\Lambda_i < \Lambda_j \mid \mathbf{x}_k = \mathbf{c}_i, \mathbf{X}'], \\ &= \sum_{j=0, j \neq i}^{L-1} Pr[(\mathbf{c}_i - \mathbf{c}_j)^T \mathbf{y}_k < 0 \mid \mathbf{x}_k = \mathbf{c}_i, \mathbf{X}'], \\ &= \sum_{j=0, j \neq i}^{L-1} Pr[(\mathbf{c}_i - \mathbf{c}_j)^T \mathbf{n}_k > (\mathbf{c}_i - \mathbf{c}_j)^T \mathbf{s}_k \mid \mathbf{x}_k = \mathbf{c}_i, \mathbf{X}'], \end{aligned} \quad (3-6)$$

where the last equality follows from (3-4). But $(\mathbf{c}_i - \mathbf{c}_j)^T \mathbf{n}_k$ is a zero-mean Gaussian random variable with variance $d_{ij}N_0$, where $d_{ij} = d_H(\mathbf{c}_i, \mathbf{c}_j)$ is the Hamming distance between codewords \mathbf{c}_i and \mathbf{c}_j . Therefore, (3-6) reduces to:

$$Pr[\text{error} \mid \mathbf{x}_k = \mathbf{c}_i, \mathbf{X}'] \leq \sum_{j=0, j \neq i}^{L-1} Q\left(\frac{(\mathbf{c}_i - \mathbf{c}_j)^T \mathbf{s}_k}{\sqrt{d_{ij}N_0}}\right). \quad (3-7)$$

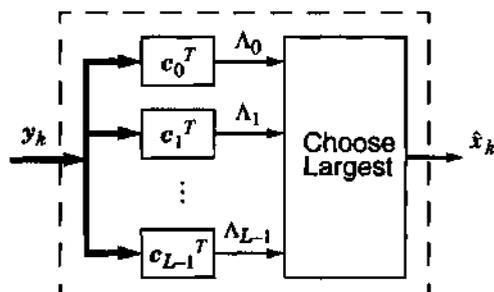


Fig. 3-2. Decision device for the unequalized receiver.

Averaging overall possible codeword sequences gives:

$$Pr[error] \leq \frac{1}{L^{\mu+1}} \sum_{\mathbf{X}'} \sum_{i=0}^{L-1} \sum_{j=0, j \neq i}^{L-1} Q\left(\frac{(\mathbf{c}_i - \mathbf{c}_j)^T \mathbf{s}_k}{\sqrt{d_{ij} N_0}}\right), \quad (3-8)$$

where the first summation is over all $\mathbf{X}' \in C^\mu$, where C is the set of L valid codewords and μ is the number of nonzero terms in the impulse response $\{\mathbf{H}_k\}$.

Finally, following [26], the bit-error probability is:

$$Pr[bit\ error] = 1 - (1 - Pr[error])^{1/\log_2 L}. \quad (3-9)$$

For example, the 6 codewords of $(\frac{4}{2})$ -MPPM are:

$$\mathbf{c}_0 = [1\ 1\ 0\ 0]^T, \mathbf{c}_1 = [1\ 0\ 1\ 0]^T, \mathbf{c}_2 = [1\ 0\ 0\ 1]^T,$$

$$\mathbf{c}_3 = [0\ 1\ 1\ 0]^T, \mathbf{c}_4 = [0\ 1\ 0\ 1]^T, \text{ and } \mathbf{c}_5 = [0\ 0\ 1\ 1]^T.$$

Thus, from (3-7), the symbol error probability for MPPM given that \mathbf{c}_0 is transmitted is:

$$Pr[error \mid \mathbf{x}_k = \mathbf{c}_0, \mathbf{X}'] \leq Q\left(\frac{s_1 - s_2}{\sqrt{2N_0}}\right) + Q\left(\frac{s_1 - s_3}{\sqrt{2N_0}}\right) + Q\left(\frac{s_0 - s_2}{\sqrt{2N_0}}\right) + Q\left(\frac{s_0 - s_3}{\sqrt{2N_0}}\right) + Q\left(\frac{s_0 + s_1 - s_2 - s_3}{\sqrt{4N_0}}\right), \quad (3-10)$$

where s_i , $i = 0, \dots, 3$, represents the chip-value of signal vector \mathbf{s} at the output of channel.

On an ideal channel without ISI, $s_0 = s_1 = s = (P/w)\sqrt{n \log_2 L / R_b}$, and $s_2 = s_3 = 0$, so that

(3-10) simplifies to $4Q(s/\sqrt{2N_0}) + Q(s/\sqrt{N_0})$, independent of the transmitted codeword

\mathbf{x}_k .

As another example, $(\frac{6}{3})$ -OPPM has $L = 4$ codewords:

$$\mathbf{c}_0 = [1\ 1\ 1\ 0\ 0\ 0]^T, \mathbf{c}_1 = [0\ 1\ 1\ 1\ 0\ 0]^T, \mathbf{c}_2 = [0\ 0\ 1\ 1\ 1\ 0]^T, \text{ and } \mathbf{c}_3 = [0\ 0\ 0\ 1\ 1\ 1]^T.$$

Following the previous reasoning, we have:

$$\Pr[\text{error} | \mathbf{x}_k = \mathbf{c}_0, \mathbf{X}] \leq Q\left(\frac{s_0 - s_3}{\sqrt{2N_0}}\right) + Q\left(\frac{s_0 + s_1 - s_3 - s_4}{\sqrt{4N_0}}\right) + Q\left(\frac{s_0 + s_1 + s_2 - s_3 - s_4 - s_5}{\sqrt{6N_0}}\right). \quad (3-11)$$

If there is no ISI, the symbol error probability reduces to the following, independent of the transmitted codeword:

$$\Pr[\text{error}] \leq \frac{1}{4}[6Q(s/\sqrt{2N_0}) + 4Q(s/\sqrt{N_0}) + 2Q(\sqrt{3}s/\sqrt{2N_0})]. \quad (3-12)$$

3.3.2 Extension to On-Off Keying

The model of Fig. 3-1-a can also be used for on-off keying (OOK) by setting the block length to $n = 1$ and by setting the weight parameter to $w = 1/2$. In this case, the bit stream, symbol stream, and chip stream are all one in the same. If all previous bits $\mathbf{x}' = \{\dots, x_{j-2}, x_{j-1}\}$ were known, then the bit error probability would be [9]:

$$\Pr[\text{error} | \mathbf{x}'] = \frac{1}{2}Q\left(\frac{h_0/2 - \sum' x_n h_{j-n}}{\sqrt{N_0}}\right) + \frac{1}{2}Q\left(\frac{h_0/2 + \sum' x_n h_{j-n}}{\sqrt{N_0}}\right), \quad (3-13)$$

where the prime indicates that the term $n = 0$ is to be omitted from the summation. By averaging over all possible bit streams $\{\mathbf{x}'\}$, the total bit error probability is:

$$\Pr[\text{bit error}] = \frac{1}{2^\mu} \sum_{\mathbf{x}'} Q\left(\frac{h_0/2 - \sum' x_n h_{j-n}}{\sqrt{N_0}}\right), \quad (3-14)$$

where the summation is overall all binary μ -tuples $\{\mathbf{x}'\}$.

3.3.3 Simplifications when the Channel is Ideal

In this section we present simplified expressions for the symbol error probability for the special case of an ideal channel, without ISI.

First, consider MPPM: when the channel has no ISI, the expression (3-8) simplifies to:

$$Pr[\text{error}] \leq \sum_{k=1}^w N_k Q\left(\sqrt{\frac{ks^2}{2N_0}}\right), \quad (3-15)$$

where $N_k = \binom{w}{k} \binom{n-k}{k}$ is the number of codewords with mutual distance $2k$, and where $s = (P/w) \sqrt{n \log_2 L / R_b}$. This expression follows from the ISI-free results for the photon counting channel of [49]. When $w = 1$, (i.e., for PPM), (3-15) simplifies further to:

$$Pr[\text{error}] \leq (L-1) Q\left(\frac{s}{\sqrt{2N_0}}\right), \quad (3-16)$$

where $s = P \sqrt{L \log_2 L / R_b}$ and this agrees with [17].

Next, consider $\binom{n}{w}$ -OPPM: when the channel has no ISI, the expression (3-8) simplifies to:

$$Pr[\text{error}] \leq \frac{1}{L} \sum_{k=1}^w M_k Q\left(\sqrt{\frac{ks^2}{2N_0}}\right), \text{ where } M_k = \begin{cases} 2(L-k) & k=1, 2, \dots, w-1 \\ (L-w)(L-w+1) & k=w \end{cases}, \quad (3-17)$$

where again $s = (P/w) \sqrt{n \log_2 L / R_b}$. This result uses the photon-counting results of [49].

Finally, consider OOK: when the channel has no ISI, then h_j from (3-3) reduces to $(2P/\sqrt{R_b})\delta_j$, and so the error probability from (3-14) simplifies to $Q(P/\sqrt{N_0 R_b})$ which agrees with the result of chapter 2.

3.3.4 Probability of Error Approximation for PPM

Exact evaluation of the probability of error in (3-8) is computationally infeasible as μ becomes large. We now consider a simpler estimate for PPM based on a Gaussian approximation [50].

To illustrate the proposed technique, consider first OOK. Since $\{x_j\}$ for OOK are i.i.d. random variables, the ISI term $\sum' h_n x_{j-n}$ is the sum of independent random variables and can be treated roughly as a Gaussian random variable with mean $0.5 \sum' h_n$ and variance $0.25 \sum' h_n^2$. This Gaussian approximation leads to the following estimate of the bit error rate for OOK [9]:

$$Pr[\text{bit error}] \approx Q\left(\frac{h_0/2}{\sqrt{N_0 + 0.25 \sum' h_n^2}}\right). \quad (3-18)$$

We can use a similar technique to approximate the error probability of PPM. From (3-4), the input to the PPM decision device can be written as the sum of signal plus ISI plus noise:

$$\mathbf{y}_k = \mathbf{H}_0 \mathbf{x}_k + \mathbf{z} + \mathbf{n}_k, \quad (3-19)$$

where \mathbf{z} is the ISI contribution:

$$\mathbf{z} = \sum' \mathbf{H}_l \mathbf{x}_{k-l}, \quad (3-20)$$

where again the prime indicates that the $l = 0$ term should be omitted from the summation. From (3-6), the probability that the k -th decision is erroneous given the past codewords is:

$$Pr[\text{error} | \mathbf{X}] \leq \sum_{j=0, j \neq i}^{L-1} Pr[(\mathbf{c}_i - \mathbf{c}_j)^T (\mathbf{n}_k + \mathbf{z}) < -(\mathbf{c}_i - \mathbf{c}_j)^T \mathbf{H}_0 \mathbf{c}_i | \mathbf{x}_k = \mathbf{c}_i, \mathbf{X}]. \quad (3-21)$$

Note that the mean μ_{ij} and variance σ_{ij}^2 of $(\mathbf{c}_j - \mathbf{c}_i)^T \mathbf{z}$ are:

$$\mu_{ij} = \frac{1}{L} \sum (\mathbf{c}_i - \mathbf{c}_j)^T \mathbf{H}_l \mathbf{1}, \quad (3-22)$$

$$\sigma_{ij}^2 = \frac{1}{L} \sum \|\mathbf{H}_l^T (\mathbf{c}_i - \mathbf{c}_j)\|^2 + \frac{1}{L^2} \sum \left((\mathbf{c}_i - \mathbf{c}_j)^T \mathbf{H}_l \mathbf{1} \right)^2, \quad (3-23)$$

where $\mathbf{1} = [1, 1 \dots 1]^T$. If we approximate $(\mathbf{c}_j - \mathbf{c}_i)^T \mathbf{z}$ as a Gaussian random variable with mean μ_{ij} and variance σ_{ij}^2 , then (3-21) reduces to:

$$Pr[\text{error}] \approx \frac{1}{L} \sum_{i=0}^{L-1} \sum_{j=0, j \neq i}^{L-1} Q \left(\frac{(\mathbf{c}_i - \mathbf{c}_j)^T \mathbf{H}_0 \mathbf{c}_i + \mu_{ij}}{\sqrt{\|\mathbf{c}_i - \mathbf{c}_j\|^2 N_0 + \sigma_{ij}^2}} \right). \quad (3-24)$$

As we will see the above approximation is accurate for small R_b/W . Note that on the ideal channel, the above equation reduces to (3-16).

To reduce computational complexity, we truncate the vector channel of (3-4) to four terms, so that $\mathbf{y}_k = \sum_{l=0}^3 \mathbf{H}_l \mathbf{x}_{k-l} + \mathbf{n}_k$. This truncation will have no appreciable effect when n is large or when R_b/W is small, although it may not be accurate for small n and large R_b/W . To validate our reasoning, we calculate the ratio of the fractional energy of $h(t)$ contained outside the truncation length to the total energy of $h(t)$;

$$\epsilon = \frac{\int_0^{\infty} h^2(t) dt}{\int_0^{\infty} h^2(t) dt}, \quad (3-25)$$

where T_{tr} represents the truncation time. We truncate up to $4n$ chip samples, so that $T_{tr} = 4T$. We list the fractional energy ratio for various R_b/W in Table 3-1. Except OOK and 2-

PPM at $R_b/W = 1$ and 2, more than 99.999% of energy is contained within the truncated samples for the modulation schemes we consider. Hence the channel energy we are dis-

TABLE 3-1: Fractional energy ratio (%) of impulse response for various R_b / W . Note that “-” means $< 0.001\%$.

	$R_b/W = 0.01$	$R_b/W = 0.1$	$R_b/W = 1$	$R_b/W = 2$
OOK	-	-	0.3%	2%
2-PPM	-	-	0.3%	2%
4-PPM	-	-	-	0.03%
8-PPM	-	-	-	0.006%
16-PPM	-	-	-	-
32-PPM	-	-	-	-
$\binom{4}{2}$ -MPPM	-	-	-	0.003%
$\binom{5}{2}$ -MPPM	-	-	-	-
$\binom{7}{2}$ -MPPM	-	-	-	-
$\binom{9}{2}$ -MPPM	-	-	-	-
$\binom{12}{2}$ -MPPM	-	-	-	-
$\binom{17}{2}$ -MPPM	-	-	-	-
$\binom{6}{3}$ -OPPM	-	-	-	-
$\binom{8}{4}$ -OPPM	-	-	-	-

carding by truncating the channel is negligible. We calculate the optical power required to achieve a 10^{-6} bit-error-rate for 4-PPM over this ISI channel using the exact method (3-8) and the Gaussian approximation method (3-24). The results are summarized in Fig. 3-3, where the normalized power requirement is plotted versus the bit-rate-to-bandwidth ratio, R_b/W . The power requirements are normalized by $P_{OOK} = \sqrt{N_0 R_b} Q^{-1}(10^{-6})$, the power required by OOK in the ideal case ($W = \infty$) to achieve a 10^{-6} bit error rate. We can see that when R_b/W is small, the Gaussian approximation method agrees well with the exact method. However, when $R_b/W > 0.3$ the approximation is not accurate. For example the difference between the two methods is about 3 dB when $R_b/W = 0.5$. We can conclude

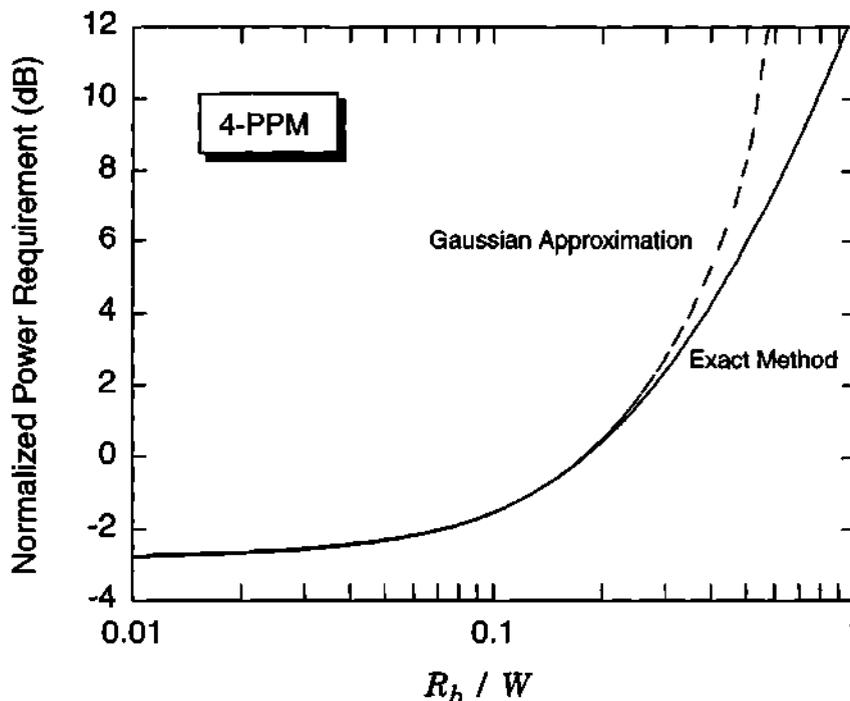


Fig. 3-3. Gaussian approximation method (3-24) and the exact method (3-8) for the unequalized receiver.

that when ISI is small, the Gaussian assumption provides a reasonable approximation for the error probability.

3.4 MAXIMUM-LIKELIHOOD SEQUENCE DETECTION (MLSD)

Using an argument similar to the scalar case [37], we can derive an upper bound for the symbol error probability for the maximum-likelihood sequence detector. The MLSD for PPM is derived in [21], and it easily generalizes to MPPM and OPPM. The MLSD chooses the vector sequence $\hat{\mathbf{x}}_k$ that minimizes the metric:

$$\lambda_{\mathbf{x}} = \sum_k \|\mathbf{y}_k - \sum_j \mathbf{H}_j \hat{\mathbf{x}}_{k-j}\|^2. \quad (3-26)$$

Because the input symbol \mathbf{x}_k is chosen uniformly and independently from the set of L valid MPPM codewords, every symbol sequence $\{\mathbf{x}_k\}$ is equally likely. We define the state of trellis at time k as:

$$\mathbf{S}_k = [\mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \dots, \mathbf{x}_{k-\mu}], \quad (3-27)$$

and there are L^μ states in the trellis. If the estimated codewords from the Viterbi detector are $\{\hat{\mathbf{x}}_k\}$, then the corresponding estimated state at time k is:

$$\hat{\mathbf{S}}_k = [\hat{\mathbf{x}}_{k-1}, \hat{\mathbf{x}}_{k-2}, \dots, \hat{\mathbf{x}}_{k-\mu}]. \quad (3-28)$$

Suppose the estimated path diverges from the correct path at time k and remerges with the correct path at time $k+l+1$. Thus, $\hat{\mathbf{S}}_k = \mathbf{S}_k$ and $\hat{\mathbf{S}}_{k+l+1} = \mathbf{S}_{k+l+1}$, but $\hat{\mathbf{S}}_i \neq \mathbf{S}_i$ for $k+1 \leq i \leq k+l$. We call this an error event. We define an error vector \mathbf{e} of length $l-\mu+1$ corresponding to the error event as:

$$\mathbf{e} = [\hat{\mathbf{x}}_k - \mathbf{x}_k, \hat{\mathbf{x}}_{k+1} - \mathbf{x}_{k+1}, \dots, \hat{\mathbf{x}}_{k+l-\mu} - \mathbf{x}_{k+l-\mu}]. \quad (3-29)$$

Let E be the set of all nonzero error events starting at time k and let $w(\mathbf{e})$ be the number of decision errors due to error event. The probability of symbol error for the MLSD at time k is upper-bounded by:

$$Pr[\text{error}] \leq \sum_{\mathbf{e} \in E} w(\mathbf{e}) Pr[\mathbf{e}]. \quad (3-30)$$

Let $N(\mathbf{e})$ be the event that an addition of the transmitted sequence $\{\mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_{k+l-\mu}\}$ to the error sequence $\{\mathbf{e}_k, \mathbf{e}_{k-1}, \dots, \mathbf{e}_{k+l-\mu}\}$ results in the valid sequence. Then, $Pr[\mathbf{e}]$ in (3-30) is upper bounded by [37][50]:

$$Pr[\mathbf{e}] \leq Pr[\lambda_{\mathbf{x}+\mathbf{e}} > \lambda_{\mathbf{x}} | N(\mathbf{e})] Pr[N(\mathbf{e})], \quad (3-31)$$

where $Pr[\lambda_{\mathbf{x}+\mathbf{e}} > \lambda_{\mathbf{x}} | N(\mathbf{e})]$ is the probability that the sum of the branch metrics of the estimated path exceeds the sum of the branch metrics of the correct path;

$$Pr[\lambda_{\mathbf{x}+\mathbf{e}} > \lambda_{\mathbf{x}} | N(\mathbf{e})] = Pr\left[\sum_{i=k}^{k+l-\mu} \|\mathbf{y}_i - \sum_{j=0}^{\mu} \mathbf{H}_j \hat{\mathbf{x}}_{i-j}\|^2 < \sum_{i=k}^{k+l-\mu} \|\mathbf{y}_i - \sum_{j=0}^{\mu} \mathbf{H}_j \mathbf{x}_{i-j}\|^2 \right], \quad (3-32)$$

$$= Pr\left[\sum_{i=k}^{k+l-\mu} \left\| \sum_{j=0}^{\mu} \mathbf{H}_j (\mathbf{x}_{i-j} - \hat{\mathbf{x}}_{i-j}) + \mathbf{n}_i \right\|^2 < \sum_{i=k}^{k+l-\mu} \|\mathbf{n}_i\|^2 \right], \quad (3-33)$$

$$= Pr\left[\sum_{i=k}^{k+l-\mu} \left\| \sum_{j=0}^{\mu} \mathbf{H}_j \mathbf{e}_{i-j} + \mathbf{n}_i \right\|^2 < \sum_{i=k}^{k+l-\mu} \|\mathbf{n}_i\|^2 \right], \quad (3-34)$$

$$= Pr\left[\sum_{i=k}^{k+l-\mu} 2\mathbf{s}_i^T \mathbf{n}_i > \sum_{i=k}^{k+l-\mu} \|\mathbf{s}_i\|^2 \right], \quad (3-35)$$

where $\mathbf{s}_i = \sum_{j=0}^{\mu} \mathbf{H}_j \mathbf{e}_{i-j}$. But $\sum_{i=k}^{k+l-\mu} 2\mathbf{s}_i^T \mathbf{n}_i$ is a zero-mean Gaussian random variable with variance $4N_0 \sum_{i=k}^{k+l-\mu} \|\mathbf{s}_i\|^2$. Then (3-35) reduces to

$$Pr[\lambda_{x+e} < \lambda_x \mid N(\mathbf{e})] = Q\left(\frac{\sqrt{\sum_{i=k}^{k+l-\mu} \|\mathbf{s}_i\|^2}}{2\sqrt{N_0}}\right) = Q\left(\frac{\sqrt{\sum_{i=k}^{k+l-\mu} \left\|\sum_{j=0}^{\mu} \mathbf{H}_j \mathbf{e}_{i-j}\right\|^2}}{2\sqrt{N_0}}\right). \quad (3-36)$$

The $Pr[N(\mathbf{e})]$ depends only on the statistical properties of the input sequence. Because the input symbol \mathbf{x}_k is chosen uniformly and independently, then for PPM:

$$Pr[N(\mathbf{e})] = \prod_i L^{-w_H(\mathbf{e}_i)}, \quad (3-37)$$

where $w_H(\cdot)$ is the Hamming weight. Using (3-30), (3-31), (3-36), and (3-37), we can calculate the upper bound of symbol error:

$$Pr[\text{error}] \leq \sum_{\mathbf{e} \in E} w(\mathbf{e}) \prod_i L^{-w_H(\mathbf{e}_i)} Q\left(\frac{\sqrt{\sum_{i=k}^{k+l-\mu} \left\|\sum_{j=0}^{\mu} \mathbf{H}_j \mathbf{e}_{i-j}\right\|^2}}{2\sqrt{N_0}}\right). \quad (3-38)$$

Let E_{min} be the set of error events corresponding to d_{min} , the minimum distance between received sequences:

$$d_{min}^2 = \sum_{i=k}^{k+l-\mu} \left\| \sum_{j=0}^{\mu} \mathbf{H}_j \mathbf{e}_{i-j} \right\|^2. \quad (3-39)$$

Then, the upper bound for probability of symbol error (3-38) can be approximated by summing over E_{min} rather than E :

$$Pr[\text{error}] \leq \sum_{\mathbf{e} \in E_{min}} w(\mathbf{e}) \prod_i L^{-w_H(\mathbf{e}_i)} Q\left(\frac{d_{min}}{2\sqrt{N_0}}\right). \quad (3-40)$$

At high SNR, (3-40) is well approximated by [51]:

$$Pr[error] \approx Q\left(\frac{d_{min}}{2\sqrt{N_0}}\right). \quad (3-41)$$

In Fig. 3-4-a, we plot power requirement versus R_b/W using (3-8) when equalization is not used. We use the same truncated vector channel considered in the last section. We see that some modulation schemes are more sensitive to ISI than others. At large bandwidth ($R_b/W < 0.1$), the ISI penalties are small. At one extreme is OOK, with a power requirement increasing slowly with decreasing bandwidth. At the other extreme is OPPM, for which the power requirement grows rapidly with decreasing bandwidth. With $R_b/W = 0.5$, the ISI penalties, as compared to the ideal channel for OOK, $(\frac{4}{2})$ -MPPM, 4-PPM, and $(\frac{6}{3})$ -OPPM, are 4.8 dB, 7.8 dB, 8.8 dB, and greater than 12 dB, respectively. It is thus highly desirable to use signal processing at the receiver to mitigate ISI, either symbol-by-symbol equalization or maximum-likelihood sequence detection.

In contrast to the unequalized results of Fig. 3-4-a, the results of Fig. 3-4-b are based on the maximum-likelihood sequence detector (MLSD). Comparing Fig. 3-4-a and Fig. 3-4-b, we see that MLSD offers significant improvement. The power requirements do not grow as rapidly as in the unequalized case, and the normalized power requirement is always less than 12 dB, even when $R_b/W = 1$. For example, when the bit rate is equal to the bandwidth, $R_b/W = 1$, the ISI penalties for OOK, $(\frac{4}{2})$ -MPPM, $(\frac{6}{3})$ -OPPM, and 4-PPM are 4.6 dB, 8.1 dB, 9.3 dB, and 9.5 dB, respectively. We note that MLSD is far more effective in reducing the power requirement for OOK than for other modulation schemes.

The bandwidth efficiency and power efficiency for PPM both using the unequalized receiver and the MLSD are summarized in Fig. 3-5. For a small R_b/W , the difference in power requirement of both receivers is negligible for all L , but for large R_b/W , the difference is appreciable. Also, when MLSD is employed each PPM-based scheme suffers approximately the same penalty due to ISI. For example, at $R_b/W = 0.1$ and $R_b/W = 0.5$, each PPM-based scheme requires 1-2 dB and 6-7 dB more power, respectively, to achieve the 10^{-6} BER than is required at $R_b/W = 0.01$.

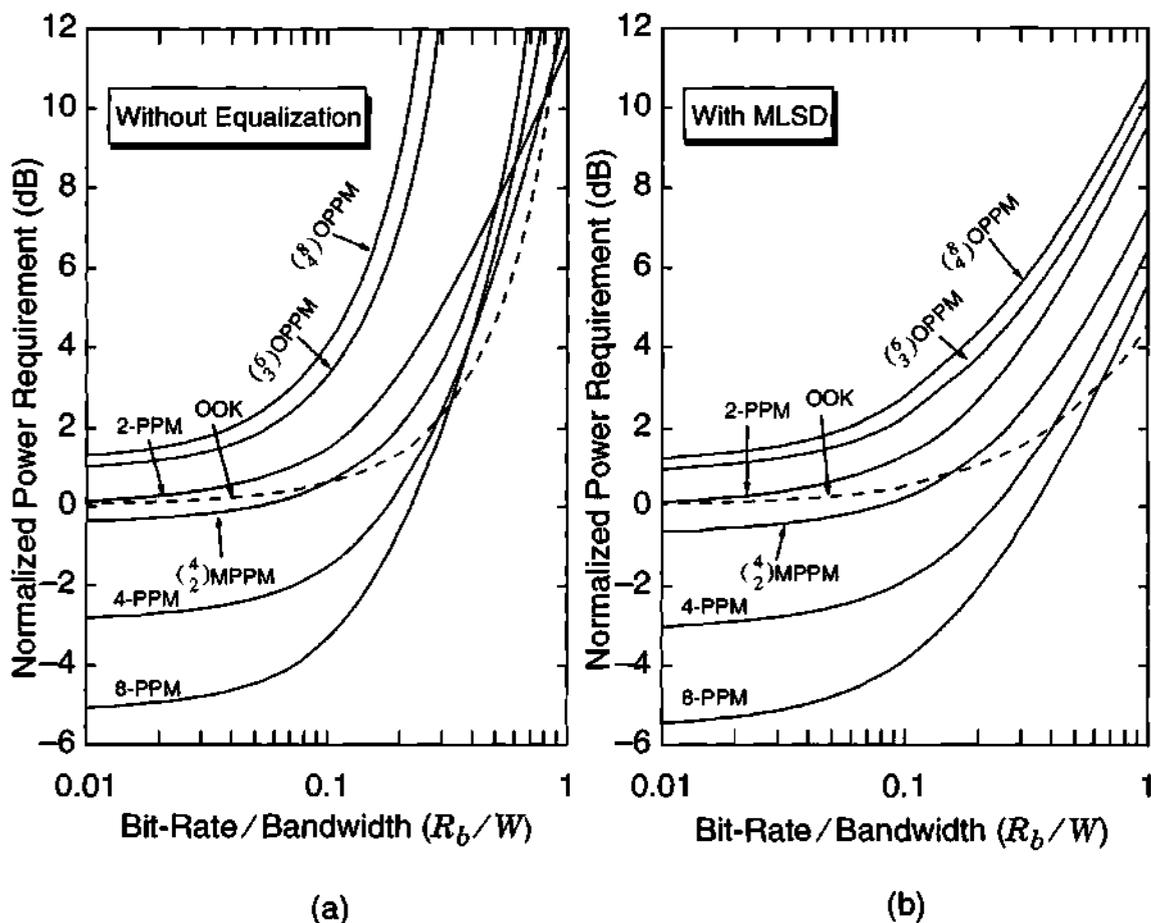


Fig. 3-4. The required power versus bit rate on an ISI channel (a) unequalized system, (b) with MLSD.

We also compare the performance of PPM and MPPM when the receiver uses MLSD. The results are summarized in Fig. 3-6. We see that MPPM is less sensitive to ISI than PPM. For example, power penalties for 2, 4, 8, 16, and 32-PPM at $R_b/W = 1$ are 9.5, 9.5, 11, 13, and 15 dB respectively. Even 32-PPM starts to require more power than 16-PPM as the R_b/W approaches 0.2. Power penalties for $\binom{4}{2}$, $\binom{5}{2}$, $\binom{7}{2}$, $\binom{9}{2}$, $\binom{12}{2}$, and $\binom{17}{2}$ -MPPM at $R_b/W = 1$ are 8, 8, 8, 8.5, 9, and 10 dB, respectively. From Fig. 2-8, we can also expect that MPPM is less susceptible to ISI than PPM because MPPM has better bandwidth efficiency than PPM for a given length of codeword n . This fact motivates us to develop trellis coded MPPM for the multipath channel (see the section 5.4).

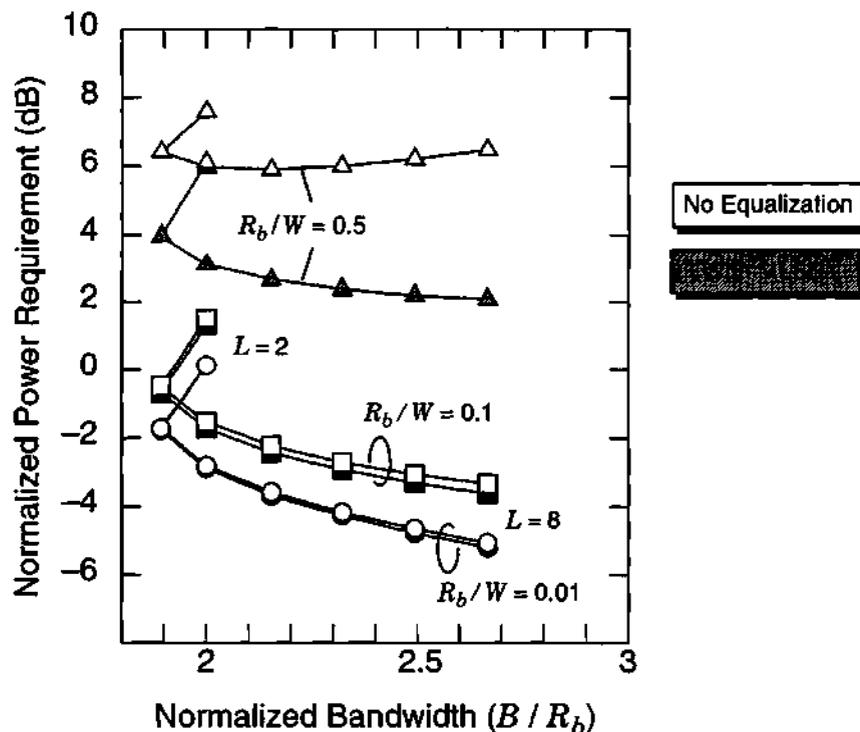


Fig. 3-5. Power efficiency versus bandwidth efficiency of PPM over ISI channel.

3.5 SYMBOL-BY-SYMBOL EQUALIZATION

In the previous section, we examined the performance of MLSD for MPPM. The MLSD employs the Viterbi algorithm so that its complexity grows exponentially with the channel memory, and so that it has a decoding delay. In this section, we will consider several symbol-by-symbol equalization techniques to mitigate the effect of ISI with reduced complexity, as well as with small or no decoding delay.

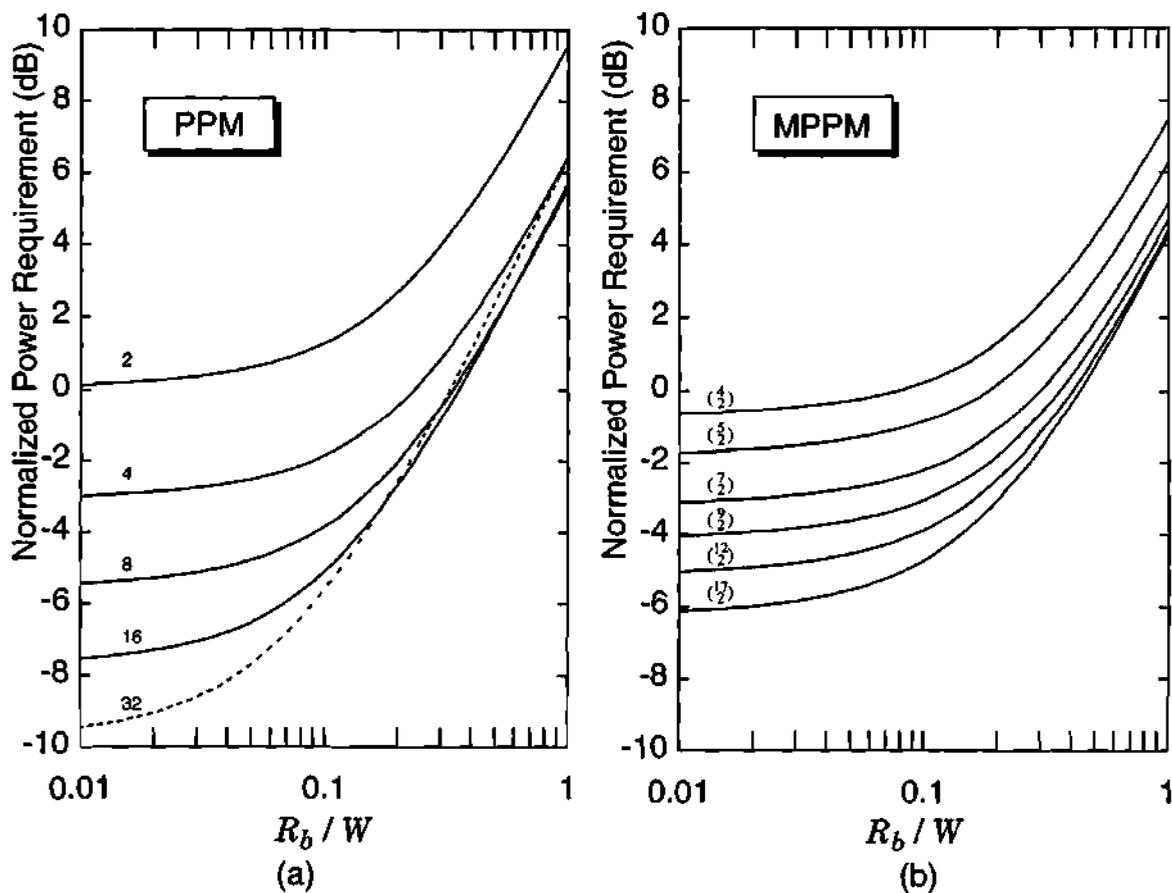


Fig. 3-6. Normalized power efficiency versus bit-rate-to-bandwidth ratio on an multipath channel with MLSD for (a) PPM, (b) MPPM.

3.5.1 Zero-forcing Block Decision Feedback Equalization (ZF-BDFE)

A zero-forcing block decision feedback equalizer (ZF-BDFE) is shown in Fig. 3-7 [21]. The received signal is passed through a forward filter of \mathbf{H}_0^{-1} . Decisions $\hat{\mathbf{x}}_k$ are fed through a feedback filter $\mathbf{H}_0^{-1} \mathbf{H}(z) - \mathbf{I}$, and the result is subtracted from the input of a vector slicer. Assuming correct decisions, $\hat{\mathbf{x}}_k = \mathbf{x}_k$, the slicer input, is then:

$$\mathbf{w}_k = \mathbf{H}_0^{-1} \mathbf{y}_k - \mathbf{f}_k = \sum_{j \geq 0} \mathbf{H}_0^{-1} \mathbf{H}_j \mathbf{x}_{k-j} + \mathbf{H}_0^{-1} \mathbf{n}_k - \sum_{j \geq 1} \mathbf{H}_0^{-1} \mathbf{H}_j \hat{\mathbf{x}}_{k-j} = \mathbf{x}_k + \mathbf{H}_0^{-1} \mathbf{n}_k. \quad (3-42)$$

There are two decision devices that operate on $\mathbf{w}_k = \mathbf{x}_k + \mathbf{H}_0^{-1} \mathbf{n}_k$: the suboptimal comparator that minimizes $\|\mathbf{w}_k - \hat{\mathbf{x}}_k\|^2$ and can be implemented using a simple “choose max” device for PPM, and the ML slicer, which minimizes $\|\mathbf{H}_0 \mathbf{w}_k - \mathbf{H}_0 \hat{\mathbf{x}}_k\|^2$.

A suboptimal decision device chooses $\hat{\mathbf{x}}_k$ to minimize $\|\hat{\mathbf{x}}_k - \mathbf{w}_k\|^2$ or to maximize the correlation:

$$\hat{\mathbf{x}}_k^T (\mathbf{x}_k + \mathbf{H}_0^{-1} \mathbf{n}_k). \quad (3-43)$$

This device is suboptimal because $\mathbf{H}_0^{-1} \mathbf{n}_k$ is not white.

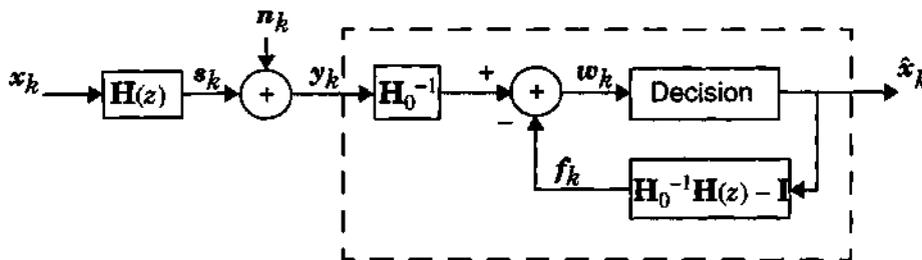


Fig. 3-7. Block diagram of ZF-BDFE [21].

If we knew that $\mathbf{x}_k = \mathbf{c}_i$, then the probability of error $\hat{\mathbf{x}}_k \neq \mathbf{x}_k$ can be bounded using the union bound:

$$\begin{aligned} Pr[\text{error} | \mathbf{x}_k = \mathbf{c}_i] &\leq \sum_{j=0, j \neq i}^{L-1} Pr[\mathbf{c}_j^T(\mathbf{c}_i + \mathbf{H}_0^{-1} \mathbf{n}_k) \geq \mathbf{c}_i^T(\mathbf{c}_i + \mathbf{H}_0^{-1} \mathbf{n}_k) | \mathbf{x}_k = \mathbf{c}_i], \\ &= \sum_{j=0, j \neq i}^{L-1} Pr[(\mathbf{c}_j - \mathbf{c}_i)^T \mathbf{H}_0^{-1} \mathbf{n}_k \geq \|\mathbf{c}_i\|^2 - \mathbf{c}_j^T \mathbf{c}_i | \mathbf{x}_k = \mathbf{c}_i], \end{aligned} \quad (3-44)$$

and $(\mathbf{c}_j - \mathbf{c}_i)^T \mathbf{H}_0^{-1} \mathbf{n}_k$ is a zero-mean Gaussian random variable with variance $N_0 \|(\mathbf{c}_j - \mathbf{c}_i)^T \mathbf{H}_0^{-1}\|^2$. Therefore, (3-44) reduces to:

$$Pr[\text{error} | \mathbf{x}_k = \mathbf{c}_i] \leq \sum_{j=0, j \neq i}^{L-1} Q\left(\frac{\|\mathbf{c}_i\|^2 - \mathbf{c}_j^T \mathbf{c}_i}{\sqrt{N_0} \|(\mathbf{c}_j - \mathbf{c}_i)^T \mathbf{H}_0^{-1}\|}\right). \quad (3-45)$$

Averaging over all possible codewords gives:

$$Pr[\text{error}] \leq \frac{1}{L} \sum_{i=0}^{L-1} \sum_{j=0, j \neq i}^{L-1} Q\left(\frac{\|\mathbf{c}_i\|^2 - \mathbf{c}_j^T \mathbf{c}_i}{\sqrt{N_0} \|(\mathbf{c}_j - \mathbf{c}_i)^T \mathbf{H}_0^{-1}\|}\right). \quad (3-46)$$

The optimal (maximum likelihood) slicer chooses $\hat{\mathbf{x}}_k$ so as to minimize $\|\mathbf{H}_0 \mathbf{w}_k - \mathbf{H}_0 \hat{\mathbf{x}}_k\|^2$. Since the noise \mathbf{n}_k is white, the probability of error depends mainly on the Euclidean distance $d^2 = \|\mathbf{H}_0 \mathbf{x}_k - \mathbf{H}_0 \hat{\mathbf{x}}_k\|^2$. We can derive an upper bound for the ML slicer considering all the error events [16]:

$$Pr[\text{error}] \leq \frac{1}{L} \sum_{i=0}^{L-1} \sum_{j=0, j \neq i}^{L-1} Q\left(\frac{\|\mathbf{H}_0(\mathbf{c}_j - \mathbf{c}_i)\|}{2\sqrt{N_0}}\right). \quad (3-47)$$

3.5.2 Zero-forcing Linear Equalization (ZF-LE)

A chip-rate zero-forcing linear equalizer (ZF-LE) uses the inverse filter $g(z) = h(z)^{-1}$ where $h(z)$ is the equivalent chip-rate impulse response as shown in Fig. 3-8. The equivalent channel between the transmitted chips and output of equalizer is:

$$r_j = x_j + z_j, \quad (3-48)$$

where z_j is the Gaussian random variable with zero-mean and variance $\sigma_z^2 = \frac{N_0}{2\pi} \int_{-\pi}^{\pi} |h^{-2}(e^{j\theta})| d\theta$. Because $h(z)$ is a causal, minimum phase filter, $h(z)^{-1}$ is stable, and σ_z^2 is finite. We group the chip samples into blocks of length n , and the equivalent vector channel is:

$$\mathbf{r}_k = \mathbf{x}_k + \mathbf{z}_k. \quad (3-49)$$

As we did in the previous section, the decision device chooses \hat{x}_k to minimize $\|\hat{x}_k - r_k\|^2$ or to maximize the correlation: $\hat{x}_k^T (\mathbf{x}_k + \mathbf{z}_k)$.

If we knew that $\mathbf{x}_k = \mathbf{e}_i$, then the probability of error $\hat{x}_k \neq \mathbf{x}_k$ could be bounded using the union bound:

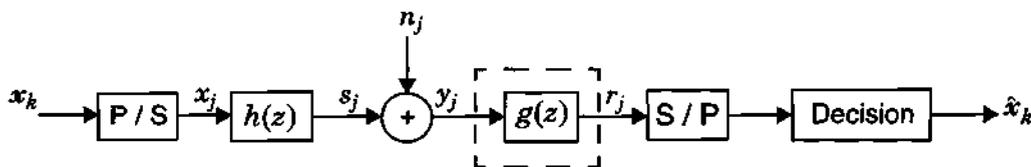


Fig. 3-8. Block diagram of chip-rate linear equalizer.

$$\begin{aligned}
Pr[\text{error} \mid \mathbf{x}_k = \mathbf{c}_i] &\leq \sum_{j=0, j \neq i}^{L-1} Pr[\mathbf{c}_j^T(\mathbf{c}_i + \mathbf{z}_k) \geq \mathbf{c}_i^T(\mathbf{c}_i + \mathbf{z}_k) \mid \mathbf{x}_k = \mathbf{c}_i], \\
&= \sum_{j=0, j \neq i}^{L-1} Pr[(\mathbf{c}_j - \mathbf{c}_i)^T \mathbf{z}_k \geq \|\mathbf{c}_i\|^2 - \mathbf{c}_j^T \mathbf{c}_i \mid \mathbf{x}_k = \mathbf{c}_i], \quad (3-50)
\end{aligned}$$

and $(\mathbf{c}_j - \mathbf{c}_i)^T \mathbf{z}_k$ is a zero-mean Gaussian random variable with variance $(\mathbf{c}_j - \mathbf{c}_i)^T \Sigma (\mathbf{c}_j - \mathbf{c}_i)$, where $\Sigma = E[\mathbf{z}_k \mathbf{z}_k^T]$. The matrix Σ is a Toeplitz matrix with $[\Sigma]_{i,j} = N_0 \sum_l g_l g_{l+i-j}$.

Averaging over all possible codeword gives:

$$Pr[\text{error}] \leq \frac{1}{L} \sum_{i=0}^{L-1} \sum_{j=0, j \neq i}^{L-1} Q\left(\frac{\|\mathbf{c}_i\|^2 - \mathbf{c}_j^T \mathbf{c}_i}{\sqrt{N_0 (\mathbf{c}_j - \mathbf{c}_i)^T \Sigma (\mathbf{c}_j - \mathbf{c}_i)}}\right). \quad (3-51)$$

We calculate the BER of DFE with suboptimal slicer, DFE with ML decision device, and LE with their analytical upper bounds (3-46), (3-47), and (3-51), respectively. The results are summarized in Fig. 3-9 for 8-PPM when $R_b/W = 1$. All simulations for DFE include error propagation. The results of our simulation (denoted by circles) approach the calculated upper bounds (dashed line) as the SNR increases. Therefore, we can use the analytical upper bound to describe the performance of the DFE and LE at lower BER.

We compare the performance of the unequalized receiver, MLSD, DFE with ML decision, and chip-rate LE for 16-PPM. We calculate the required power to achieve the 10^{-6} BER using (3-8), (3-41), (3-47), and (3-51). The results are summarized in Fig. 3-10. When R_b/W is small, all receivers show almost the same performance. When R_b/W is large, the receiver employing equalizers shows better performance than the unequalized

receiver. DFE, especially, shows better performance than the LE. DFE requires 0.3 dB more than MLSD when $R_b/W = 1$.

3.5.3 Maximum *a Posteriori* Probability (MAP) Detection

We now examine the performance of a symbol-by-symbol detection method based on maximizing *a posteriori* probabilities (MAP) developed by Abend and Fritchman [43].

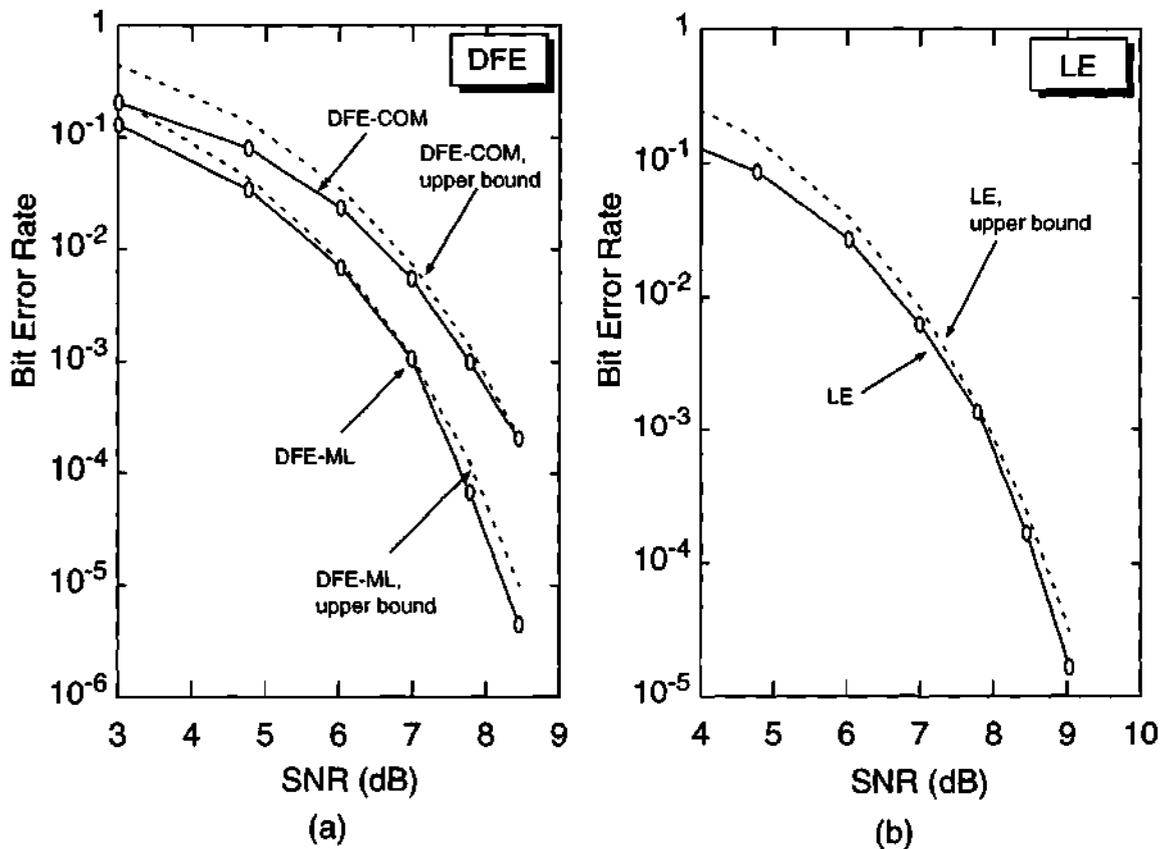


Fig. 3-9. Bit error rate of equalization schemes and their upper bounds for 8-PPM at bit-rate-to-bandwidth ratio of 1: (a) DFE with suboptimal decision device (DFE-COM) and DFE with ML decision device, (b) LE.

We want to decide the codeword \hat{x}_k given the received sequence $\{y_k, y_{k+1}, \dots, y_{k+D-1}\}$, where the delay D is chosen to satisfy $D \geq \mu$. Based on the received sequences, we decide on a \hat{x}_k that maximizes *a posteriori* probabilities (MAP) [43]:

$$\hat{x}_k = \max_{x_k} p(x_k | y_k, y_{k+1}, \dots, y_{k+D-1}). \quad (3-52)$$

By Bayes' rule:

$$p(x_k | y_k, y_{k+1}, \dots, y_{k+D-1}) = \frac{p(y_k, y_{k+1}, \dots, y_{k+D-1} | x_k) p(x_k)}{p(y_k, y_{k+1}, \dots, y_{k+D-1})}, \quad (3-53)$$

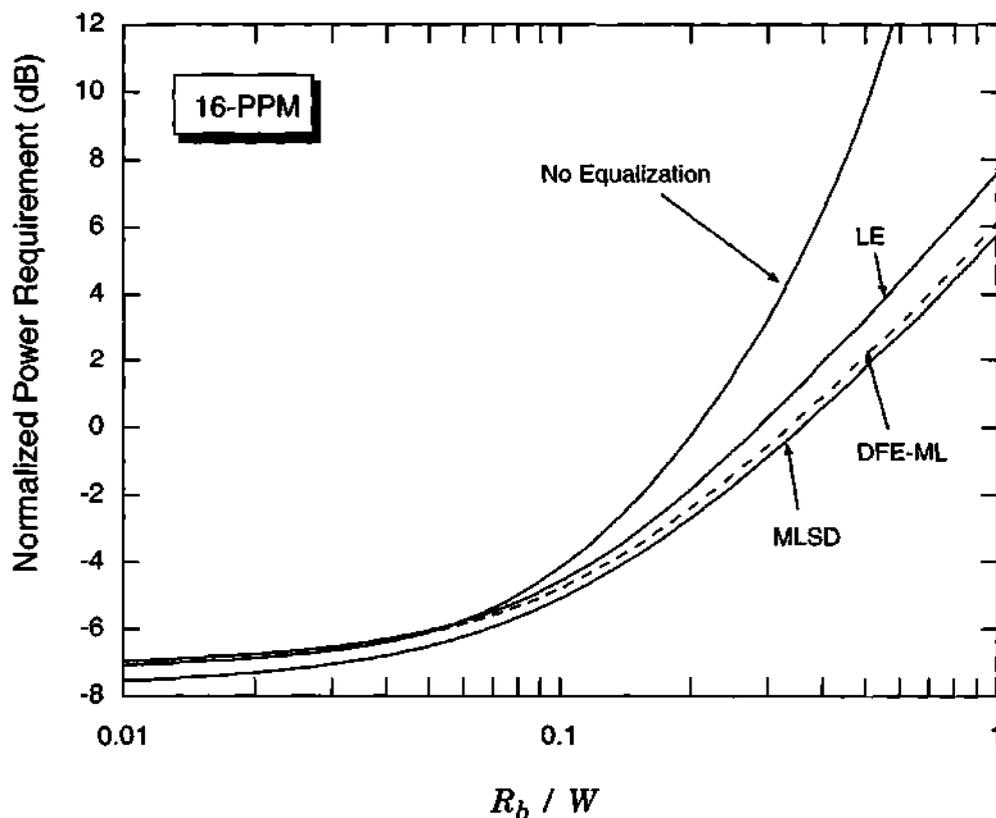


Fig. 3-10. Normalized power requirement of equalization schemes for 16-PPM.

where $p(\mathbf{y}_k, \mathbf{y}_{k+1}, \dots, \mathbf{y}_{k+D-1})$ is common to all, and the codewords are equally probable, $p(\mathbf{x}_k) = 1/L$. The MAP criterion is equivalent to choosing the $\hat{\mathbf{x}}_k$ that maximizes:

$$\hat{\mathbf{x}}_k = \max_{\mathbf{x}_k} p(\mathbf{y}_k, \mathbf{y}_{k+1}, \dots, \mathbf{y}_{k+D-1} | \mathbf{x}_k), \quad (3-54)$$

$$= \max_{\mathbf{x}_k} \sum_{\mathbf{x}_{k+1}} \sum_{\mathbf{x}_{k+2}} \dots \sum_{\mathbf{x}_{k+D-1}} p(\mathbf{y}_k, \mathbf{y}_{k+1}, \dots, \mathbf{y}_{k+D-1} | \mathbf{x}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_{k+D-1}), \quad (3-55)$$

where each summation is taken over the L possible codewords. We assume input codewords $\{\mathbf{x}_k\}$ are equally probable and independent of the noise. Equation (3-55) can be calculated recursively beginning with $k = 1$:

$$\begin{aligned} \hat{\mathbf{x}}_1 &= \max_{\mathbf{x}_1} \sum_{\mathbf{x}_2} \sum_{\mathbf{x}_3} \dots \sum_{\mathbf{x}_D} p(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_D | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_D), \\ &= \max_{\mathbf{x}_1} \sum_{\mathbf{x}_2} \sum_{\mathbf{x}_3} \dots \sum_{\mathbf{x}_D} p(\mathbf{y}_D | \mathbf{x}_D, \dots, \mathbf{x}_{D-\mu}) \dots p(\mathbf{y}_2 | \mathbf{x}_2, \mathbf{x}_1) p(\mathbf{y}_1 | \mathbf{x}_1), \end{aligned} \quad (3-56)$$

where $p(\mathbf{y}_D | \mathbf{x}_D, \dots, \mathbf{x}_{D-\mu}) = (2\pi N_0)^{-n/2} \exp(-\|\mathbf{y}_D - \sum_{m=0}^{\mu} \mathbf{H}_m \mathbf{x}_{D-m}\|^2 / 2N_0)$, and we assume that $\mathbf{x}_k = \mathbf{0}$ for $k < 0$. Unlike the other equalizer, the MAP detector must know the variance of the noise. Observe that (3-56) involves a large number of computations, including the summation of exponential factors, thus making it difficult to implement the MAP detector in real applications, especially when L is large.

Williamson *et al.* [44] developed a MAP-like detector based on the decision feedback equalizer structure when $D = \mu + 1$, as shown in Fig. 3-11. The received codeword \mathbf{y}_k is buffered to generate a sequence of D signals, $\{\mathbf{y}_k, \mathbf{y}_{k+1}, \dots, \mathbf{y}_{k+D-1}\}$. An input signal in each DFE is subtracted by the output of feedback filter. The input vectors for the decision device are:

$$w_k^i = y_{k+i-1} - \sum_{j=i}^{\mu} H_j \hat{x}_{k-j+i-1} \quad i = 1, 2, \dots, D. \quad (3-57)$$

The MAP decision device chooses \hat{x}_k based on the $\{w_k^1, w_k^2, \dots, w_k^D\}$. We can rewrite the MAP decision rule (3-52) as [44]:

$$\hat{x}_k = \max_{x_k} p(x_k | y_k, y_{k+1}, \dots, y_{k+D-1}) = \max_{x_k} p(x_k | w_k^1, w_k^2, \dots, w_k^D, e_k = 0), \quad (3-58)$$

where $e_k = x_k - \hat{x}_k$, and where the second equality (3-58) assumes no error propagation.

Using the same argument used in (3-54)–(3-56),

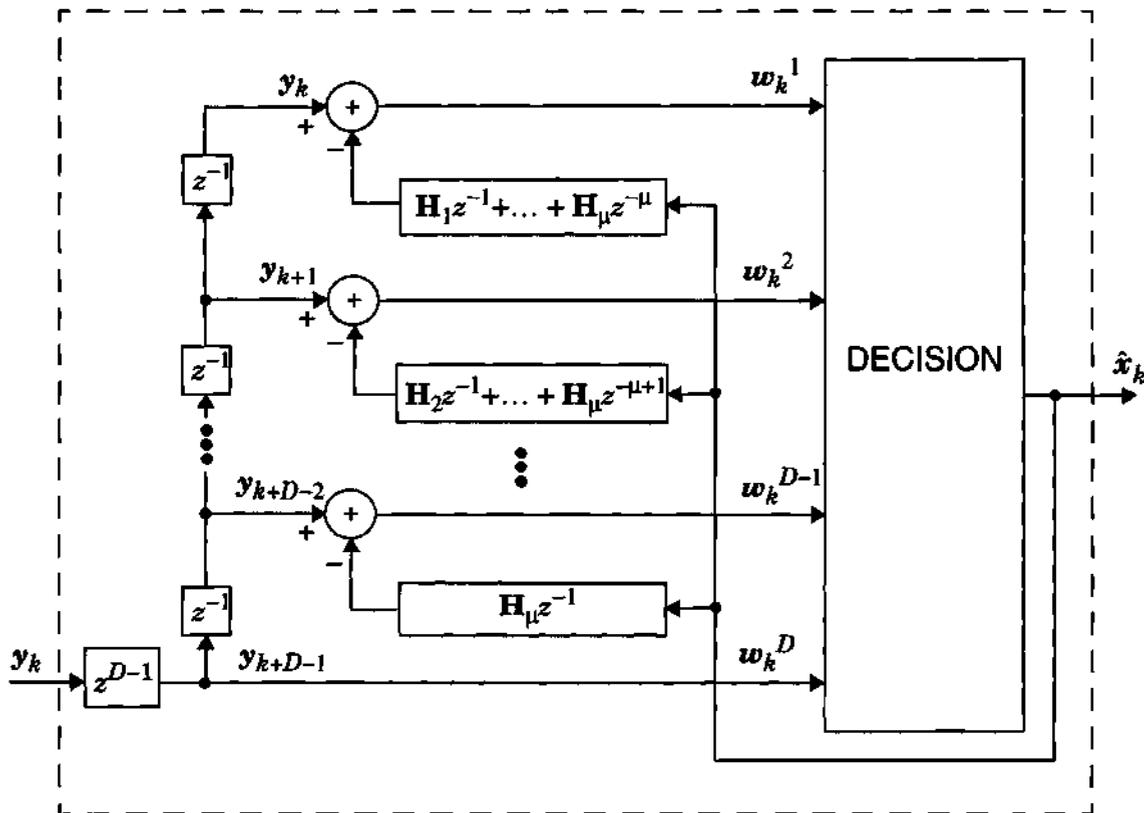


Fig. 3-11. Block diagram of the MAP-DFE.

$$\hat{\mathbf{x}}_k = \max_{\mathbf{x}_k} \sum_{\mathbf{x}_{k+1}} \sum_{\mathbf{x}_{k+2}} \cdots \sum_{\mathbf{x}_{k+D-1}} p(\mathbf{w}_k^1, \mathbf{w}_k^2, \dots, \mathbf{w}_k^D | \mathbf{x}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_{k+D-1}, \mathbf{e}_k = 0). \quad (3-59)$$

We can interpret (3-59) as a structure equivalent to symbol-by-symbol detector (3-55) coupled with decision feedback equalizer. We can see that the complexity of MAP-DFE (3-59) is much smaller than the original MAP detection (3-55) because \mathbf{w}_k^i contains less ISI terms than \mathbf{y}_{k+i-1} to average out assuming correct decisions. Williamson *et al.* [44] showed that the performance of MAP-DFE range from conventional DFE to the MLSD depending on D .

3.6 PARTIAL-RESPONSE PRECODING SCHEME

Precoding is a technique similar to DFE that eliminates error propagation by moving cancellation of the postcursor ISI from the receiver to the transmitter. One advantage of using precoding over DFE in our applications is the reduction of complexity at the receiver. Usually, the battery powered portable receiver is required to consume little power. Therefore, equalization at the transmitter can reduce the power consumption of the receiver. We move the feedback filter in Fig. 3-7 to the transmitter as shown in Fig. 3-12-a. However, simply moving the equalizer to the transmitter is not practical because of the following reasons.

- The transmitted signal $\mathbf{x}_k = \mathbf{b}_k - \mathbf{f}_k$ violates the average power and nonnegative constraint of (1-2).
- If we add a positive constant vector \mathbf{p} to compensate all negative components of \mathbf{x}_k and satisfy the constraint (1-2), which decreases the distances between the transmitted codewords.
- The resulting \mathbf{x}_k loses the low duty-cycle property.

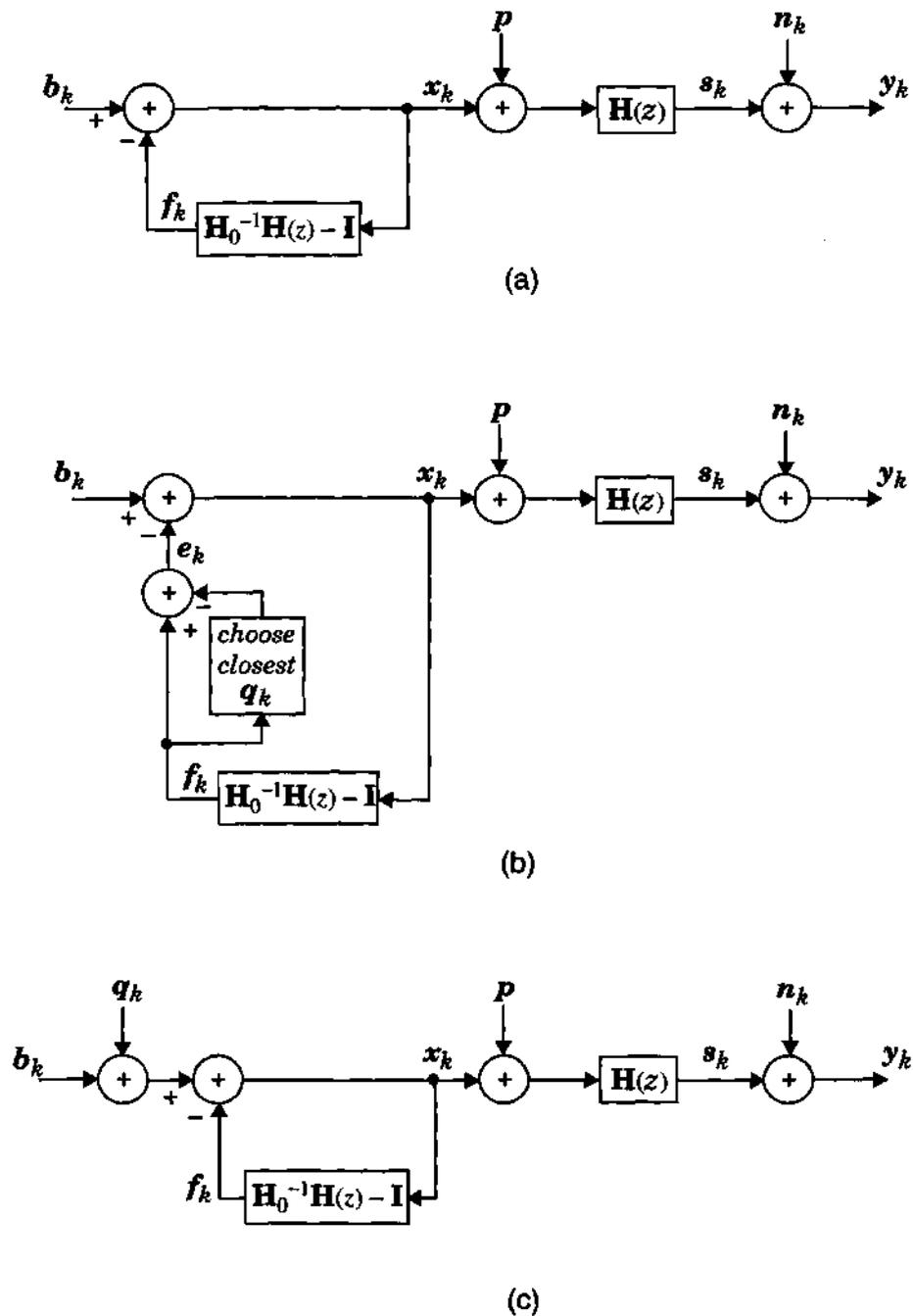


Fig. 3-12. Precoding operation (a) moving DFE to the transmitter, (b) the proposed precoding scheme, (c) equivalent block diagram for (b).

To overcome the above problems we stated, we now propose a new precoding scheme as a modification of LTF precoding [47][48]. Let $C = \{\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_{L-1}\}$ is a set of L valid MPPM codewords, and Q is a set of L codewords defined later. In the proposed precoding scheme as shown in Fig. 3-12-b, the output of feedback filter $\mathbf{f}_k = \sum_{j \geq 1} \mathbf{H}_0^{-1} \mathbf{H}_j \mathbf{x}_{k-j}$ is not subtracted directly from the input signal \mathbf{b}_k , but is quantized to the closest codeword \mathbf{q}_k in Q , and only the quantization error $\mathbf{e}_k = \mathbf{f}_k - \mathbf{q}_k$ is subtracted. The output of the precoder is:

$$\mathbf{x}_k = \mathbf{b}_k - \mathbf{e}_k, \quad (3-60)$$

$$= \mathbf{b}_k + \mathbf{q}_k - \sum_{j \geq 1} \mathbf{H}_0^{-1} \mathbf{H}_j \mathbf{x}_{k-j}. \quad (3-61)$$

From (3-61), the equivalent operation of quantizer in Fig. 3-12-b is obtained by adding \mathbf{q}_k to the input \mathbf{c}_k as shown in Fig. 3-12-c. To satisfy the (1-2), we add a positive constant vector \mathbf{p} to \mathbf{x}_k . At the output of the channel, after removing the constant:

$$\tilde{\mathbf{y}}_k = \mathbf{y}_k - \sum_{j \geq 0} \mathbf{H}_j \mathbf{p}, \quad (3-62)$$

$$= \mathbf{H}_0 \mathbf{x}_k + \sum_{j \geq 1} \mathbf{H}_j \mathbf{x}_{k-j} + \mathbf{n}_k, \quad (3-63)$$

$$= \mathbf{H}_0(\mathbf{b}_k + \mathbf{q}_k) + \mathbf{n}_k. \quad (3-64)$$

Equation (3-64) is obtained by applying (3-61) to (3-63).

The first type of receiver to recover $\hat{\mathbf{b}}_k$ from $\tilde{\mathbf{y}}_k$ uses inverse precoder as shown in Fig. 3-13. The decision device chooses $\hat{\mathbf{b}}_k + \hat{\mathbf{q}}_k$, which minimizes the Euclidean distance $\|\tilde{\mathbf{y}}_k - \mathbf{H}_0(\hat{\mathbf{b}}_k + \hat{\mathbf{q}}_k)\|^2$. In order to recover $\hat{\mathbf{b}}_k$ from $\hat{\mathbf{b}}_k + \hat{\mathbf{q}}_k$, we invert the operation of the precoder [52]. From (3-60), (3-61), we can recover $\hat{\mathbf{b}}_k - \hat{\mathbf{e}}_k$ by subtracting the feed-

back signal f_k from $\hat{b}_k + \hat{q}_k$. We then obtain the codeword \hat{b}_k by adding \hat{e}_k to the $\hat{b}_k - \hat{e}_k$.

To make this precoding effective, we should design Q to ensure that (1) it minimize the average energy of e_k , in order to minimize the increment of power at the transmitter, and (2) ensure that $\{b_k + q_k\}$ are distinctive, in order to decode effectively. This design of Q is the most challenging problem of our precoding scheme. Since $e_k = f_k - q_k$, q_k should be selected to be close to the f_k . We assume that the equivalent discrete channel coefficients h_j are a monotonically decreasing sequence, so that the feedback signal $f_k = \sum_{j \geq 1} \mathbf{H}_0^{-1} \mathbf{H}_j x_{k-j}$ depends mainly on $\mathbf{H}_0^{-1} \mathbf{H}_1 x_{k-1} = \mathbf{H}_0^{-1} \mathbf{H}_1 (b_{k-1} - e_{k-1})$. We choose the L codewords, Q , according to the value of b_{k-1} :

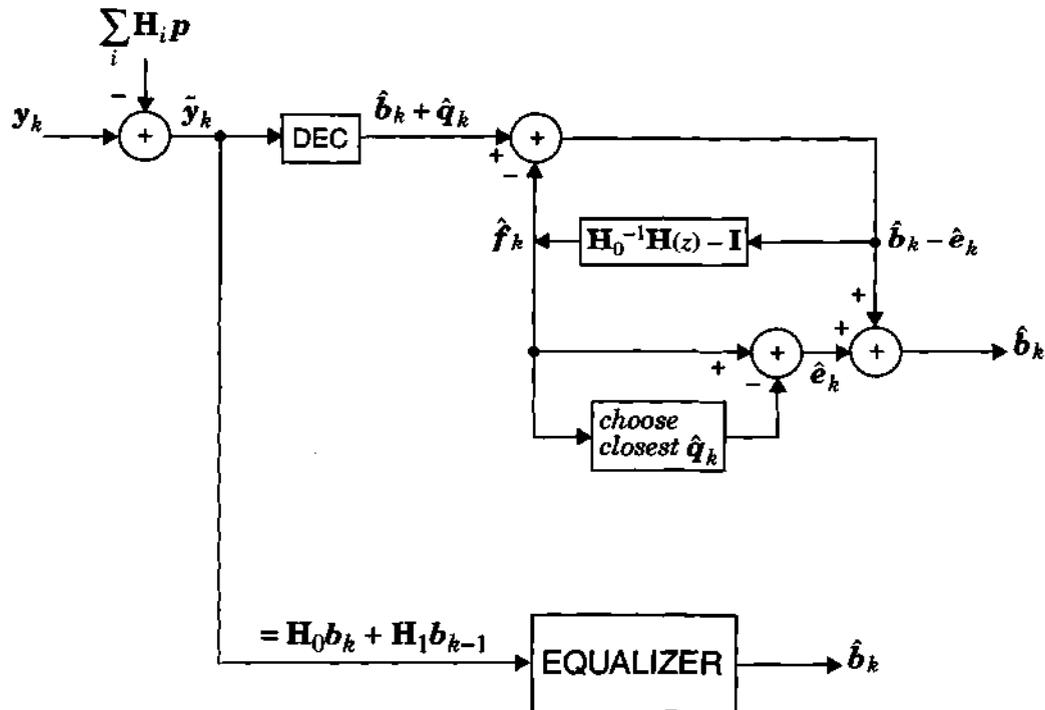


Fig. 3-13. Two types of receiver for precoding scheme; the upper one is an inverse precoder and lower one is for partial-response precoding.

$$\mathbf{q}_k \in Q = \{ \mathbf{H}_0^{-1} \mathbf{H}_1 \mathbf{c}_0, \mathbf{H}_0^{-1} \mathbf{H}_1 \mathbf{c}_1, \dots, \mathbf{H}_0^{-1} \mathbf{H}_1 \mathbf{c}_{L-1} \}, \quad (3-65)$$

where \mathbf{c}_i is a valid MPPM codeword. Therefore, $\{\mathbf{b}_k + \mathbf{q}_k\}$ are not valid MPPM codewords and the number of codewords is L^2 . We can expect the performance of this precoding to be inferior to the ideal ZF-DFE because of the reduced minimum distance among the signal set for $\{\mathbf{b}_k + \mathbf{q}_k\}$ compared to $\{\mathbf{b}_k\}$.

In the first type of receiver, we considered \mathbf{q}_k as a valid codeword in Q . However, in the second type of receiver we can think from a different point of view. We quantize the feedback signal $\mathbf{f}_k = \sum_{j \geq 1} \mathbf{H}_0^{-1} \mathbf{H}_j \mathbf{x}_{k-j}$ to $\mathbf{q}_k = \mathbf{H}_0^{-1} \mathbf{H}_1 \mathbf{b}_{k-1}$, and the quantization error is $\mathbf{e}_k = \sum_{j \geq 2} \mathbf{H}_0^{-1} \mathbf{H}_j \mathbf{x}_{k-j} - \mathbf{H}_0^{-1} \mathbf{H}_1 \mathbf{e}_{k-1}$. Then the output of channel is:

$$\bar{\mathbf{y}}_k = \sum_{j \geq 0} \mathbf{H}_j \mathbf{x}_{k-j} + \mathbf{n}_k, \quad (3-66)$$

$$= \mathbf{H}_0(\mathbf{b}_k + \mathbf{q}_k) + \mathbf{n}_k, \quad (3-67)$$

$$= \mathbf{H}_0(\mathbf{b}_k + \mathbf{H}_0^{-1} \mathbf{H}_1 \mathbf{b}_{k-1}) + \mathbf{n}_k, \quad (3-68)$$

$$= \mathbf{H}_0 \mathbf{b}_k + \mathbf{H}_1 \mathbf{b}_{k-1} + \mathbf{n}_k. \quad (3-69)$$

Therefore, the precoding operation truncates the ISI channel from $\mu + 1$ taps to 2 taps. We call this scheme *partial-response precoding*². The effect of ISI caused by the truncated channel can be mitigated using any of the equalization schemes, with reduced complexity, that we considered in this chapter. For example, the MAP detector and its DFE implementation of Fig. 3-11 are complex. However, after partial-response precoding, we can implement the MAP equalizer at the receiver with reduced complexity, as shown in Fig. 3-14.

2. The term partial response is used in connection with precoding since the response to an input symbol is spread over two symbol period.

3.7 NUMERICAL RESULTS

Fig. 3-15 shows the bit-error-rate performance of all equalization schemes we have considered at $R_b/W = 1$ assuming 4-PPM is used. We can see that all equalization schemes are reasonably effective at mitigating the effects of ISI. The best performance is achieved by the MAP detector using the two-stage DFE with $D = 2$; it is about 0.2 dB from the optimum ML sequence detector. If we use more stages, the performance will be comparable with MLSD [44]. As we derived in the previous section, when we use partial-response precoding at the transmitter, its operation approximately truncates the ISI channel memory length to 1. We can use several effective receiver equalization methods with reduced complexity. For example, partial-response precoding with MAP-DFE (Fig. 3-14) shows better performance than conventional DFE with ML decision. However, partial-response precoding at the transmitter with DFE at the receiver to remove the 1 memory ISI, or precoding with inverse precoder, does not show better performance than

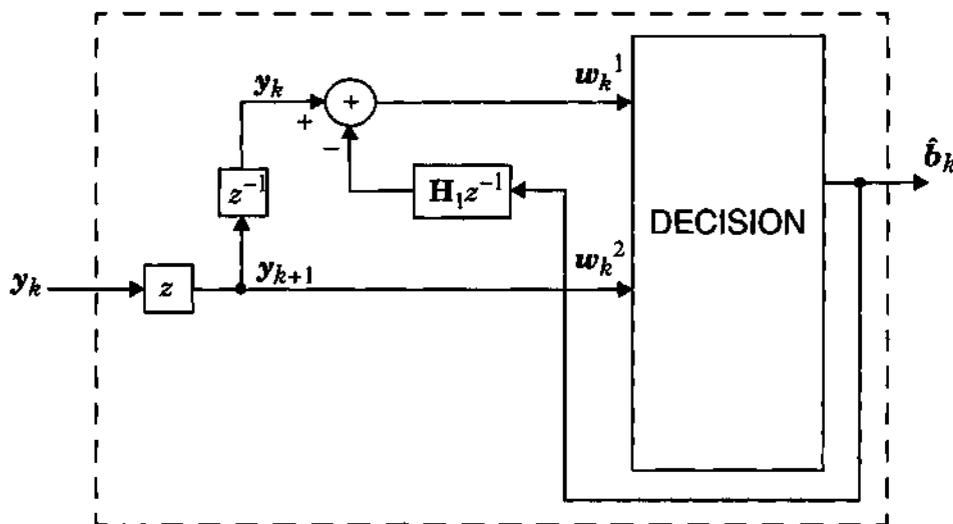


Fig. 3-14. MAP-DFE at the receiver when transmitter uses partial-response precoding.

DFE. The performance of the DFE with comparator is almost the same as the LE. Note that if the receiver does not employ equalization, the BER performance does not improve significantly, even at high SNR.

Fig. 3-16 shows the performance of equalization schemes in terms of receiver complexity and performance. The x -axis represents the number of floating-point operations (flops) per bit and the y -axis represents the required SNR to achieve 10^{-3} bit-error-rate. In calculating the flops, we use the `flops` command in MATLAB. MLSD achieves the best performance with largest complexity. The complexity of the MAP-DFE and the precoding with MAP-DFE is smaller than MLSD, but it is much larger than those of the other symbol-by-symbol equalization schemes. LE requires the least complexity but its performance is poor. DFE and precoding with DFE shows good performance and small complexity. Especially in our applications, precoding at the transmitter and DFE at the receiver is a good choice because of its relatively simple receiver structure.

3.8 SUMMARY AND CONCLUSIONS

We examined the performance of MPPM and its variants on ISI channels using several equalization schemes at the receiver. Compared to conventional LE and DFE, MAP-DFE performs well, and its performance is close to the optimum MLSD. However, its complexity is large and does not seem to be compatible with coding. The DFE with ML decision device is effective at mitigating ISI. We developed a new precoding scheme called partial-response precoding as a modification of LTF precoding. Accounting for both complexity and performance, partial-response precoding with DFE is a good choice for our application.

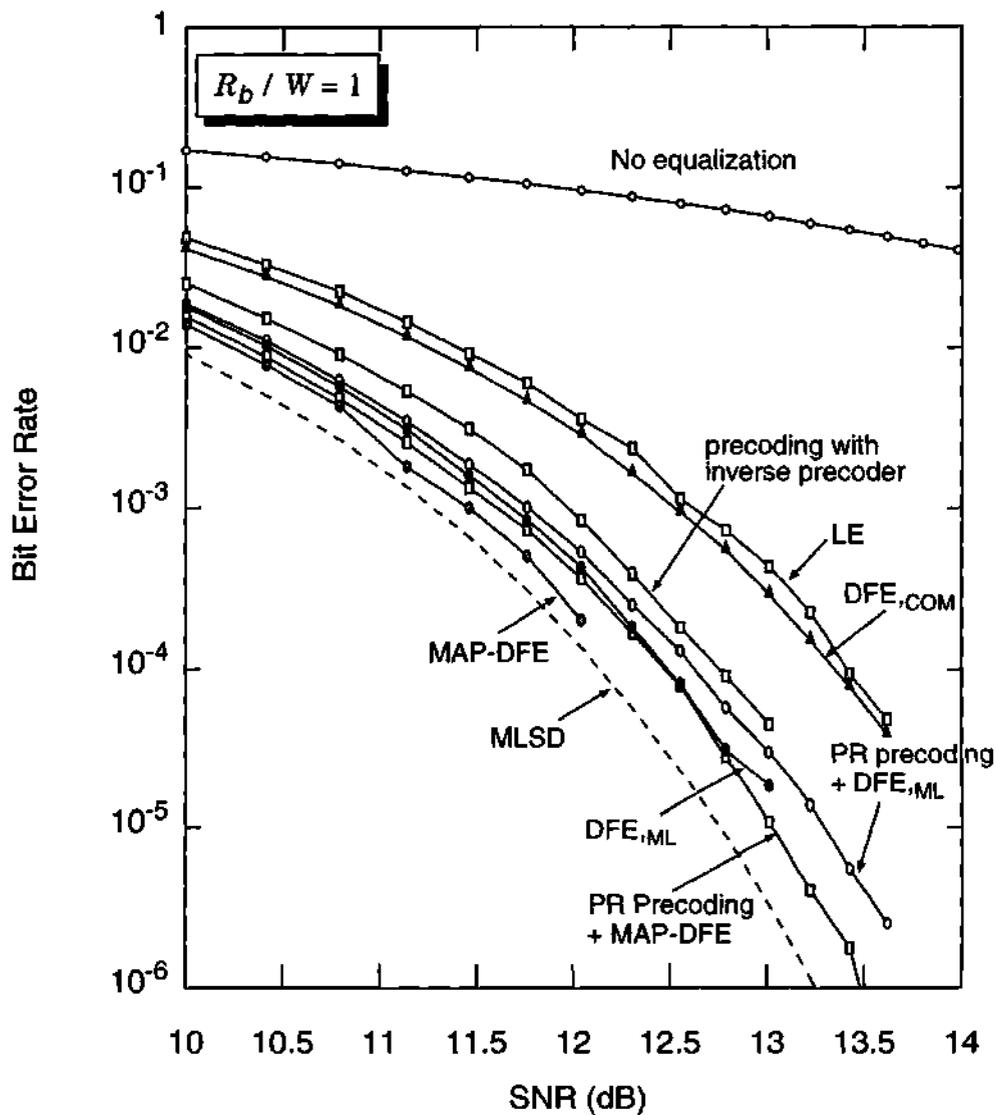


Fig. 3-15. Performance of equalization schemes for 4-PPM at bit-rate-to-bandwidth is equal to 1.

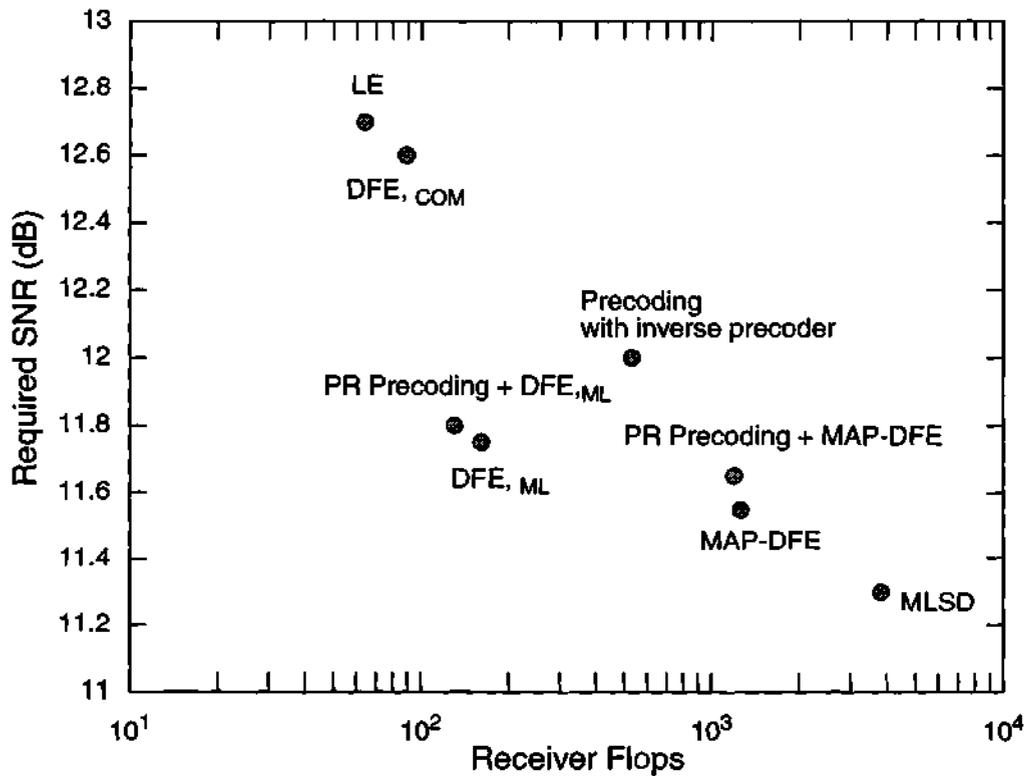


Fig. 3-16. Complexity vs. performance for each equalization scheme using 4-PPM at $R_b / W = 1$; x -axis represents the number of floating-point operations (flops) per bit and y -axis represents the required SNR to achieve 10^{-3} bit error rate.

CHAPTER 4

ACHIEVABLE INFORMATION RATE AND CUTOFF RATE

Although bit error rate analysis in the previous chapters provides us with valuable information about the performance of uncoded modulation schemes, it does not give us information about the limit of performance. In this chapter, we will consider more fundamental questions of information transfer such as:

- What is an achievable rate at SNR of 10 dB using a specific modulation scheme?
- How much SNR is required to achieve a cutoff rate of 2 bits / channel symbol using 4-OPPM?

4.1 INTRODUCTION

The *channel capacity* is the highest rate in bits per channel use at which information can be sent with arbitrarily low probability of error [53]. The capacity of a discrete-time memoryless channel subject to various input constraints has been studied since the advent of information theory by Shannon. The most common input constraint is the average-power, and the associated capacity of a Gaussian channel has already been determined by Shannon. Channel capacity for an average-power and peak-power limited input is achieved by discrete random variables taking on a finite number of values [54]. Capacity for average amplitude with positive input constraint such as the wireless indoor infrared channel (see (1-2)) has not been reported (see section 7.2.1).

Tsybakov [55] derived the channel capacity for discrete-time ISI channels with an average input power constraint. In that case, the capacity is achieved by letting the input sequence be correlated Gaussian random variables. Hirt [56] approximated the information rate with an independent, identically distributed (i.i.d.) input, using the Monte Carlo method. Shamai [57] derived lower and upper bounds on the information rate for a scalar Gaussian channel with i.i.d. input. The bounds are represented in terms of the average mutual information of a memoryless Gaussian channel with scaled i.i.d. input symbols.

The *cutoff rate* is another measure of communication efficiency. While the capacity of a channel is the rate beyond which it is impossible to communicate over the channel, the cut-off rate is widely believed to be the rate beyond which it is very expensive to communicate over the channel [58]. This parameter was initially considered in conjunction with sequential decoding of convolutional codes. Massey [59] recognized the cutoff rate as a

figure of merit for modulation schemes. Biglieri [60] derived the cutoff rate of an ISI channel using random coding bounds.

Georghiades [49] investigated channel capacity and cutoff rate for MPPM, OPPM, and PPM over a photon counting channel without ISI.

In section 4.2, we present expressions for achievable information rate and cutoff rate on the ideal channel when the input codewords are independent and identically distributed with a uniform distribution. In section 4.3, we derive lower and upper bounds to the information rate over ISI channels and compare it to a Monte Carlo approximation. We also present cutoff rate, and calculate the power to achieve $\log_2 L$ bits per channel symbol. In section 4.4, we present numerical results for our channel.

4.2 MEMORYLESS CHANNEL

In this section, we will consider two fundamental limits of information transfer, achievable information rate and cutoff rate. We assume an ideal AWGN channel with independent and identically distributed uniform input.

4.2.1 Achievable Information Rate

In a memoryless channel, $\mathbf{H}_0 = \mathbf{I}$, $\mathbf{H}_1 = \mathbf{0}$, ... in (3-4). The *information rate* (bits per channel symbol) under the i.i.d. with a uniform distribution $p(x_k) = 1 / L$ over the AWGN channel is represented by the mutual information between input and output vectors:

$$I_{i.i.d.} = I(\mathbf{x}; \mathbf{x} + \mathbf{n}) =$$

$$\frac{1}{L} \sum_{l=0}^{L-1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi N_0)^{L/2}} e^{-\|\mathbf{y} - \mathbf{x}_l\|^2 / 2N_0} \log_2 \left(\frac{e^{-\|\mathbf{y} - \mathbf{x}_l\|^2 / 2N_0}}{\frac{1}{L} \sum_{m=0}^{L-1} e^{-\|\mathbf{y} - \mathbf{x}_m\|^2 / 2N_0}} \right) d\mathbf{y}. \quad (4-1)$$

As for one-dimensional signals [58], we substitute $\mathbf{z} = (\mathbf{y} - \mathbf{x}_l) / \sqrt{N_0}$ and $\mathbf{v}_m = \mathbf{x}_m / \sqrt{N_0}$.

The above equation then becomes:

$$I_{i.i.d.} = \log_2 L - \frac{1}{L} \sum_{l=0}^{L-1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{2\pi^{L/2}} e^{-\|\mathbf{z}\|^2 / 2} \log_2 \left(\sum_{m=0}^{L-1} e^{-(\mathbf{v}_l - \mathbf{v}_m)\mathbf{z}} e^{-\|\mathbf{v}_l - \mathbf{v}_m\|^2 / 2} \right) d\mathbf{z}. \quad (4-2)$$

The above equation contains an n dimensional integral and has no simple closed form solution. As a consequence, the Monte Carlo method is used to estimate $I_{i.i.d.}$.

4.2.2 Cutoff Rate

The *cutoff rate* R_0 is defined when the input codewords are i.i.d. with a uniform distribution $p(\mathbf{x}_k) = 1/L$ as [59]:

$$R_0 = -\log_2 \left(\frac{1}{L^2} \sum_{l=0}^{L-1} \sum_{m=0}^{L-1} e^{-\|\mathbf{v}_l - \mathbf{v}_m\|^2 / 8} \right) \text{ bits / channel symbol}, \quad (4-3)$$

where $\mathbf{v} = \mathbf{x} / \sqrt{N_0}$.

We can simplify for the special case of MPPM using the results for the photon counting channel [49]:

$$R_{0, MPPM} = \log_2 L - \log_2 \left(\sum_{k=0}^w N_k e^{-ks^2/4} \right), \quad (4-4)$$

where $s = \frac{P}{w} \sqrt{n \log_2 L / R_b N_0}$, and $N_k = \binom{w}{k} \binom{n-w}{k}$ is the number of codewords with mutual distance $2k$.

Recall that PPM is a special case of MPPM with $n = L$ and $w = 1$. Because PPM codewords have a unity weight and equal Hamming distance of 2, the cutoff rate can be simplified to:

$$R_{0, PPM} = \log_2 L - \log_2(1 + (L-1)e^{-s^2/4}), \quad (4-5)$$

where $s = P \sqrt{L \log_2 L / R_b N_0}$.

For OPPM, we use the results for the photon counting channel [49]:

$$R_{0, OPPM} = \log_2 L - \log_2 \left(1 + \frac{1}{L} \sum_{k=1}^w M_k e^{-ks^2/4} \right), \quad (4-6)$$

where $M_k = \begin{cases} 2(L-k) & k = 1, 2, \dots, w-1 \\ (L-w)(L-w+1) & k = w \end{cases}$, and $s = \frac{P}{w} \sqrt{n \log_2 L / R_b N_0}$.

For OOK:

$$R_{0, OOK} = 1 - \log_2(1 + e^{-s^2/8}), \text{ where } s = 2P / \sqrt{R_b N_0}. \quad (4-7)$$

4.3 ISI CHANNEL

Following [57], we can also represent the channel using matrix notation:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N} = \mathbf{S} + \mathbf{N}, \quad (4-8)$$

where $\mathbf{Y} = [\mathbf{y}_0^T, \mathbf{y}_1^T, \dots, \mathbf{y}_{N-1}^T]^T$, $\mathbf{X} = [\mathbf{x}_0^T, \mathbf{x}_1^T, \dots, \mathbf{x}_{N-1}^T]^T$, $\mathbf{S} = [\mathbf{s}_0^T, \mathbf{s}_1^T, \dots, \mathbf{s}_{N-1}^T]^T$, and $\mathbf{N} = [\mathbf{n}_0^T, \mathbf{n}_1^T, \dots, \mathbf{n}_{N-1}^T]^T$ are all $Nn \times 1$ column vectors. The two equations (3-4) and (4-8) are equivalent as $N \rightarrow \infty$, and the rows of \mathbf{H} are specified by circular shifts of $\{\mathbf{H}_i\}$:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{H}_1 & \mathbf{H}_0 & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_\mu & \dots & \mathbf{H}_0 \end{bmatrix}. \quad (4-9)$$

The achievable information rate for under the i.i.d. constraint with a uniform distribution $p(\mathbf{x}_k) = 1/L$ is [53][56]:

$$I_{i.i.d.} = \lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{Y}; \mathbf{X}) = \lim_{N \rightarrow \infty} \frac{1}{N} (h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X})), \quad (4-10)$$

$$= \log_2 L - \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{l=0}^{L^N-1} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{1}{L} \frac{1}{2\pi^{Nn/2}} e^{-|\mathbf{z}|^2/2} \log_2 \left(\sum_{m=0}^{L^N-1} e^{-(\mathbf{s}_l - \mathbf{s}_m)\mathbf{z} - |\mathbf{s}_l - \mathbf{s}_m|^2/2} \right) d\mathbf{z}. \quad (4-11)$$

Exact evaluation of this expression is not possible, so we resort to upper and lower bounds. In the following discussion, we present lower and upper bounds for the information rate under the i.i.d. constraint. Our presentation is a straightforward generalization of the scalar results of Shamai [57] applied to the vector channel (3-4).

4.3.1 Lower Bound of the Information Rate: I_L

Following [57], we represent the entropy of the output vector by using the chain rule:

$$h(\mathbf{Y}) = \sum_{l=0}^{N-1} h(\mathbf{y}_l | \mathbf{y}_{l-1}, \mathbf{y}_{l-2}, \dots, \mathbf{y}_0). \quad (4-12)$$

Because conditioning decreases the entropy:

$$\begin{aligned}
h(\mathbf{Y}) &\geq \sum_{l=0}^{N-1} h(\mathbf{s}_l + \mathbf{n}_l | \mathbf{s}_{l-1}, \dots, \mathbf{s}_0, \mathbf{n}_{l-1}, \dots, \mathbf{n}_0), \\
&= \sum_{l=0}^{N-1} h(\mathbf{s}_l + \mathbf{n}_l | \mathbf{s}_{l-1}, \dots, \mathbf{s}_0),
\end{aligned} \tag{4-13}$$

where the last equality follows because $\mathbf{s}_l, \mathbf{n}_l, \mathbf{n}_{l-1}, \dots, \mathbf{n}_0$ are independent. Since $\mathbf{s}_l = \sum_j \mathbf{H}_{l-j} \mathbf{x}_j$, (4-13) reduces to:

$$h(\mathbf{Y}) \geq \sum_{l=0}^{N-1} h(\mathbf{H}_0 \mathbf{x}_l + \mathbf{n}_l). \tag{4-14}$$

Therefore, the mutual information between the input and output vectors is:

$$I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{N}) \geq \sum_{l=0}^{N-1} \{h(\mathbf{H}_0 \mathbf{x}_l + \mathbf{n}_l) - h(\mathbf{n}_l)\}, \tag{4-15}$$

$$= \sum_{l=0}^{N-1} I(\mathbf{H}_0 \mathbf{x}_l + \mathbf{n}_l; \mathbf{x}_l). \tag{4-16}$$

This equation leads to the following lower bound I_L for the information rate under the i.i.d. constraint:

$$\lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{X}; \mathbf{Y}) \geq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{l=0}^{N-1} I(\mathbf{H}_0 \mathbf{x}_l + \mathbf{n}_l; \mathbf{x}_l) = I(\mathbf{H}_0 \mathbf{x} + \mathbf{n}; \mathbf{x}) = I_L. \tag{4-17}$$

We can evaluate the above equation by replacing \mathbf{x} with $\mathbf{H}_0 \mathbf{x}$ in (4-2). As we can see, the lower bound of the information rate is equivalent to the mutual information between the input and the output of an error-free zero-forcing block decision feedback equalizer (ZF-BDFE) with ML-decision (see section 3.5.1) [21].

4.3.2 Upper Bound of the Information Rate: I_U

We now present an upper bound I_U for the information rate $I_{i.i.d.}$ of (4-11) as an extension of [57]. By using the chain rule, we can represent the mutual information between the input and output of (4-8) as:

$$\begin{aligned} I(\mathbf{Y}; \mathbf{X}) &= \sum_{l=0}^{N-1} I(\mathbf{Y}; \mathbf{x}_l | \mathbf{x}_{l-1}, \mathbf{x}_{l-2}, \dots, \mathbf{x}_0), \\ &= \sum_{l=0}^{N-1} [h(\mathbf{x}_l | \mathbf{x}_{l-1}, \mathbf{x}_{l-2}, \dots, \mathbf{x}_0) - h(\mathbf{x}_l | \mathbf{x}_{l-1}, \mathbf{x}_{l-2}, \dots, \mathbf{x}_0, \mathbf{Y})]. \end{aligned} \quad (4-18)$$

Because $\{\mathbf{x}_k\}$ are i.i.d. and conditioning decreases the entropy:

$$\begin{aligned} I(\mathbf{Y}; \mathbf{X}) &\leq \\ &\sum_{l=0}^{N-1} [h(\mathbf{x}_l | \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{l-1}, \mathbf{x}_{l+1}, \mathbf{x}_{l+2}, \dots, \mathbf{x}_N) - \\ &\quad h(\mathbf{x}_l | \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{l-1}, \mathbf{x}_{l+1}, \mathbf{x}_{l+2}, \dots, \mathbf{x}_N, \mathbf{Y})], \\ &= \sum_{l=0}^{N-1} I(\mathbf{x}_l; \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{l-1}, \mathbf{x}_{l+1}, \mathbf{x}_{l+2}, \dots, \mathbf{x}_N, \mathbf{Y}), \\ &= \sum_{l=0}^{N-1} I(\mathbf{x}_l; \mathbf{H}_0 \mathbf{x}_l + \mathbf{n}_0, \mathbf{H}_1 \mathbf{x}_l + \mathbf{n}_1, \dots, \mathbf{H}_\mu \mathbf{x}_l + \mathbf{n}_\mu), \\ &= \sum_{l=0}^{N-1} I(\mathbf{x}_l; \tilde{\mathbf{Y}}), \end{aligned} \quad (4-19)$$

where $\tilde{\mathbf{Y}} = \tilde{\mathbf{H}} \mathbf{x}_l + \tilde{\mathbf{N}}$, and where $\tilde{\mathbf{H}} = [\mathbf{H}_0^T, \mathbf{H}_1^T, \dots, \mathbf{H}_\mu^T]^T$, and $\tilde{\mathbf{N}} = [\mathbf{n}_0^T, \mathbf{n}_1^T, \dots, \mathbf{n}_\mu^T]^T$.

Let $\mathbf{V} = \tilde{\mathbf{H}}^T \tilde{\mathbf{Y}} = \mathbf{B} \mathbf{x}_l + \mathbf{w}$, where $\tilde{\mathbf{H}}^T$ is a matched filter, so that $\mathbf{B} = \sum_{m=0}^{\mu} \mathbf{H}_m^T \mathbf{H}_m$, and \mathbf{w} is a zero-mean Gaussian vector with correlation matrix $\mathbf{E}[\mathbf{w} \mathbf{w}^T] = N_0 \mathbf{B}$. Because \mathbf{B} is positive definite, it can be factored into $\mathbf{B} = \Gamma \Gamma^T$ for some matrix Γ [61]. We can whiten the noise by applying Γ^{-1} to \mathbf{V} , yielding $\mathbf{Z} = \Gamma^{-1} \mathbf{V} = \Gamma^T \mathbf{x}_l + \mathbf{n}$, where \mathbf{n} has the same distribution as \mathbf{n}_l , a zero-mean Gaussian vector with correlation

matrix $N_0\mathbf{I}$. Since both the matched filter and noise whitener are information lossless, we have $I(\mathbf{x}_l; \tilde{\mathbf{Y}}) = I(\mathbf{x}_l; \mathbf{V}) = I(\mathbf{x}_l; \mathbf{Z})$. Therefore, (4-19) reduces to:

$$I(\mathbf{Y}; \mathbf{X}) \leq \sum_{l=0}^{N-1} I(\mathbf{x}_l; \Gamma^T \mathbf{x}_l + \mathbf{n}_l). \quad (4-20)$$

Finally, taking the limit as $N \rightarrow \infty$ yields our upper bound I_U of the information rate under the i.i.d constraint:

$$\lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{X}; \mathbf{Y}) \leq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{l=0}^{N-1} I(\mathbf{x}_l + \mathbf{n}_l; \mathbf{x}_l) = I(\Gamma^T \mathbf{x} + \mathbf{n}; \mathbf{x}) = I_U. \quad (4-21)$$

We can evaluate (4-21) by replacing \mathbf{x} with $\Gamma^T \mathbf{x}$ in (4-2). Note that, for the scalar case, the matrix Γ^T reduces to $\sqrt{\sum_m |h_m|^2}$, and the upper bound (4-21) reduces to the matched filter bound [57].

4.3.3 Cutoff Rate

Following the discussion of [60], we can generalize the results of the scalar channel to the vector case. Let $S = (\mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \dots, \mathbf{x}_{k-\mu}; \tilde{\mathbf{x}}_{k-1}, \tilde{\mathbf{x}}_{k-2}, \dots, \tilde{\mathbf{x}}_{k-\mu})$, and let $\Xi = (\mathbf{x}_k; \tilde{\mathbf{x}}_k)$ denote a state pair and a input pair respectively, and define:

$$\Phi(S, \Xi) = \frac{1}{L^2} \exp \left(- \left\| \mathbf{H}_0(\mathbf{x}_k - \tilde{\mathbf{x}}_k) + \sum_{m=1}^{\mu} \mathbf{H}_m(\mathbf{x}_{k-m} - \tilde{\mathbf{x}}_{k-m}) \right\|^2 / 8N_0 \right). \quad (4-22)$$

We can derive the cutoff rate of the vector ISI channel using a random coding bound [62]:

$$R_0 = -\log_2 \lambda_{max}, \quad (4-23)$$

where λ_{max} is the maximum eigenvalue of a $L^{2\mu} \times L^{2\mu}$ matrix A , where $A_{ij} = \Phi(S_j, \Xi)$, if a state pair S_i can be reached from S_j for some input pair Ξ and $A_{ij} = 0$ otherwise.

For example, if $L = 2$ (with codewords \mathbf{c}_0 and \mathbf{c}_1) and $\mu = 1$, then the matrix A is given by:

$$A = \begin{bmatrix} a & b & b & a \\ c & d & e & c \\ c & e & d & c \\ a & b & b & a \end{bmatrix}, \quad (4-24)$$

where $a = 1$,

$$b = \frac{1}{4}e^{-\|\mathbf{H}_1(\mathbf{c}_0 - \mathbf{c}_1)\|^2/8N_0},$$

$$c = \frac{1}{4}e^{-\|\mathbf{H}_0(\mathbf{c}_0 - \mathbf{c}_1)\|^2/8N_0},$$

$$d = \frac{1}{4}e^{-\|\mathbf{H}_0(\mathbf{c}_0 - \mathbf{c}_1) + \mathbf{H}_1(\mathbf{c}_0 - \mathbf{c}_1)\|^2/8N_0}, \text{ and}$$

$$e = \frac{1}{4}e^{-\|\mathbf{H}_0(\mathbf{c}_0 - \mathbf{c}_1) + \mathbf{H}_1(\mathbf{c}_1 - \mathbf{c}_0)\|^2/8N_0}.$$

The cutoff rate can be calculated by finding the maximum eigenvalue of A and can be approximated at high SNR as:

$$R_0 \approx 1 - \log_2(1 + e^{-(\|\mathbf{H}_0(\mathbf{c}_0 - \mathbf{c}_1)\|^2 + \|\mathbf{H}_1(\mathbf{c}_0 - \mathbf{c}_1)\|^2)/8N_0}). \quad (4-25)$$

4.4 NUMERICAL RESULTS

Because there is no simple closed form for (4-2), we use Monte Carlo methods to estimate the multiple integral. Fig. 4-1 shows the approximate information rate using the Monte Carlo method and cutoff rate for 4-PPM at $R_b / W = 0.01$ and $R_b / W = 0.5$. To

evaluate the information rate, we use 1000 Gaussian sample vectors for each estimate [63]. To achieve 1 bit / channel symbol, there are about 1.1 dB difference between the information rate and the cutoff rate for the 4 PPM at a bit rate $R_b / W = 0.5$.

In Fig. 4-2, I_{ideal} (4-2), I_L (4-17), and I_U (4-21) are shown and compared to the approximate information rate. Our results show that I_L and I_U are 0.5 dB apart at moderate SNR at $R_b / W = 0.5$. When the channel is ideal, at SNR = 3.3 dB, the information rate is 0.95 bits/channel symbol; however the information rate is only 0.18 bits/channel symbol when $R_b/W = 0.5$. To achieve a information rate of 0.95 bits/channel symbol using 2-PPM at $R_b/W = 0.5$, the required SNR is 9.5 dB. In contrast, we see that an uncoded 2-PPM system with MLSD requires an additional 3.3 dB, or 12.8 dB SNR, to achieve 10^{-6} BER at $R_b/W = 0.5$ (see Fig. 3-4-b). Thus, in this case, the coding gain for a code based

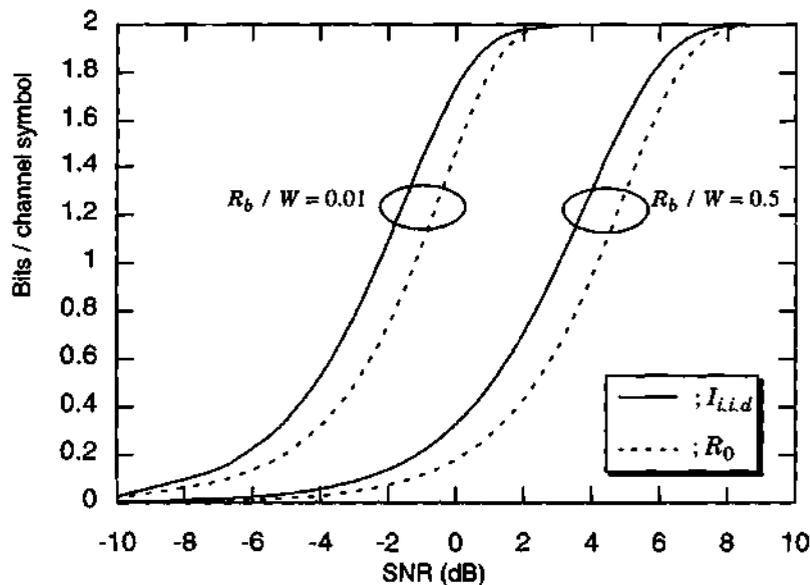


Fig. 4-1. Approximate information rate, $I_{i.i.d.}$, and cutoff rate, R_0 , for 4-PPM as a function of SNR at $R_b / W = 0.01$ and $R_b / W = 0.5$.

on 2-PPM can be at most 3.3 dB. Higher coding gains are possible for higher-order alphabets.

For each modulation scheme, we have calculated the optical power required to achieve a cutoff rate of $(\log_2 L - \epsilon)$ bits / channel symbol over this ISI channel, where ϵ is an arbitrary small number. The power requirements are normalized by $P_{OOK} = \sqrt{N_0 R_b} \sqrt{\log\left(\frac{1}{2^\epsilon - 1}\right)^2}$, the power required by OOK in the ideal case ($W = \infty$) to achieve a $(1 - \epsilon)$ bits / channel symbol. The results are summarized in Fig. 4-3 for $\epsilon = 10^{-3}$, where the normalized power requirement is plotted versus the bit-rate-to-bandwidth ratio, R_b/W . The normalized power requirement is always less than 12 dB, even when $R_b/W = 1$. For

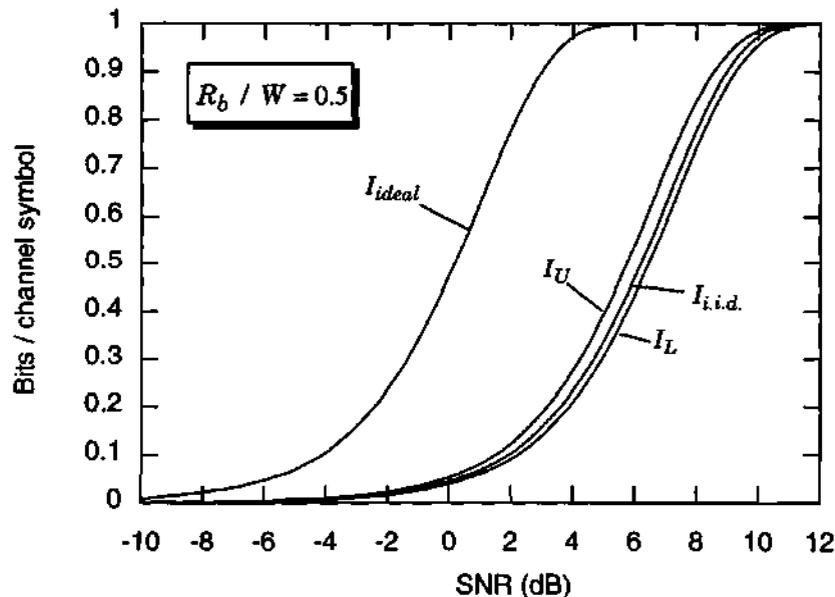


Fig. 4-2. Bounds on the information rate for 2-PPM over an ISI channel as a function of $SNR = P / \sqrt{R_b N_0}$; information rate for the memoryless channel I_{ideal} , lower bound I_L , upper bound I_U , and approximate information rate $I_{i.i.d.}$.

example, when the bit rate is equal to the bandwidth, $R_b/W = 1$, the ISI penalties for OOK, $(\frac{6}{3})$ -OPPM, $(\frac{3}{2})$ -MPPM, and 4-PPM are 4.5 dB, 9 dB, 9.1 dB, and 9.2 dB, respectively.

4.5 SUMMARY AND CONCLUSIONS

We have examined the achievable information rate and cutoff rate of MPPM and its variants PPM, OPPM, and OOK on an ISI channel with additive white Gaussian noise. We have presented expressions for the information rate under the i.i.d. constraint and cutoff rate on an ideal channel for each modulation scheme. When input codewords are i.i.d. with uniform distribution, we have derived bounds for the information rate on an ISI that can be

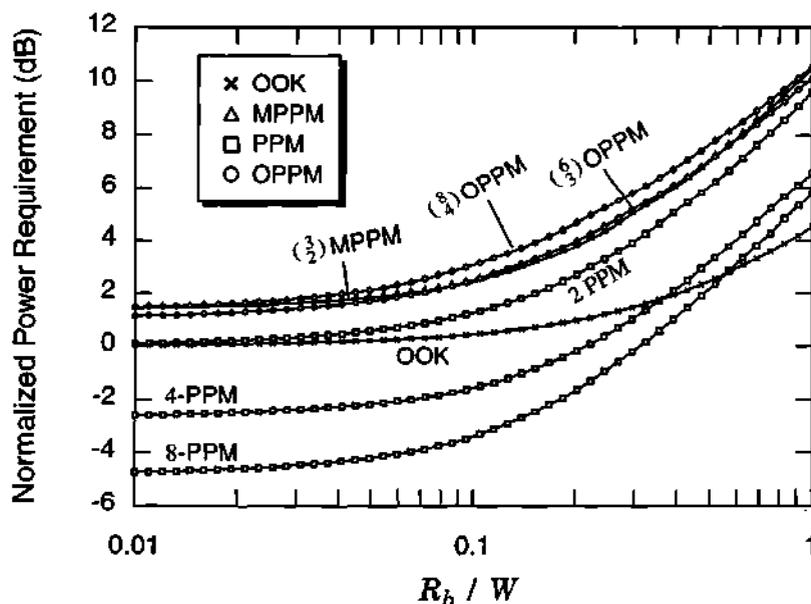


Fig. 4-3. The required power to achieve a cutoff rate of $(\log_2 L - \epsilon)$ bits / channel symbol ($\epsilon = 10^{-3}$) normalized by P_{OOK} , versus bit rate on an ISI channel.

expressed as the information rate of an ideal channel. We show that the lower bound of the information rate is equivalent to the mutual information between the input and the output in the block zero-forcing decision feedback equalizer (ZF-DFE) with ML decision, and that upper bound of the information rate is equal to the vector version of matched filter bound. We also calculate the optical power to achieve a cutoff rate of $(\log_2 L - \epsilon)$ bits / channel symbol. Our results quantify the unavoidable penalties due to ISI. Among all the modulation schemes considered here, OOK is much less sensitive to ISI than any of the PPM-based modulation schemes, thus agreeing with the results in section 3.4.

CHAPTER 5

C O D E D M O D U L A T I O N

So far, we have examined the performance of uncoded modulation schemes on both ideal and ISI channels. In this chapter, we will combine coding with modulation to improve the error rate performance. Here we consider only convolutional code and trellis code because we prefer to use the Viterbi algorithm at the receiver. We assume ideal channels (no ISI), and the effects of ISI will be considered in chapter 6.

5.1 INTRODUCTION

The performance of modulation schemes can be improved by combining them with coding techniques. Conventional coding operates by inserting redundant symbols into the

data to allow the receiver to correct errors and erasures that occur in transmission. Because the signal bandwidth is proportional to the symbol rate, this increases the bandwidth. Rather than treating the coding and modulation as separate processes, coded modulation techniques combine them by *matching* the coding to the modulation scheme. The basis of coded modulation is to treat coding and modulation as one entity; coded modulation is a single process that converts the data stream directly into a suitable waveform for transmission over the channel. Fig. 5-1 shows a block diagram of a coded modulation scheme. The coded modulation performs two mapping operations. A block of $(\log_2 L - 1)$ information bits, \mathbf{a} , is mapped onto a $\log_2 L$ bits label vector \mathbf{b} by a rate $(\log_2 L - 1) / \log_2 L$ convolutional encoder. The label vector \mathbf{b} is then mapped onto a codeword \mathbf{x} , and this mapping is denoted by $f(\mathbf{b}) = \mathbf{x}$. In general, the mapping operation is memoryless and nonlinear.

In 1982, Ungerboeck [64] suggested *trellis coded modulation* (TCM) to improve the performance without bandwidth expansion using set partitioning. He showed that a coding gain of 6 dB relative to an uncoded system could be achieved using a trellis code with 128 states and without bandwidth expansion. Calderbank and Sloane [65] observed that the signal constellation may be regarded as a finite set of points taken from an infinite lattice, and the partitioning of the constellation into subsets corresponds to the partitioning of a lattice into a sublattice and its cosets.

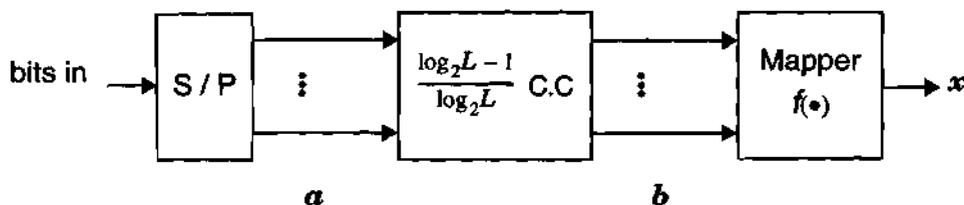


Fig. 5-1. Block diagram of coded modulation.

Because the error probability of a trellis code at high SNR primarily depends on the minimum Euclidean distance between sequences, d_{min} , the calculation of this parameter is the most important task. In general, in order to evaluate the d_{min} of a trellis code, all pairs of encoded sequences must be considered. Large complexity saving in calculating d_{min} may be achieved if we may assume that the all-zero path may serve as the reference path. This has already been observed by Ungerboeck and subsequently by Zehavi and Wolf [66]. Here, we consider a condition of symmetry for the signal set developed by Zehavi and Wolf, the Z.W. condition for a rate $(\log_2 L - 1) / \log_2 L$ trellis code. The Z.W. condition requires that when the signal set is partitioned into two subsets, the distance weight profiles of the two subsets be identical. If a trellis code satisfies this Z.W. condition, d_{min} can be calculated from the all-zero path. Benedetto [67] summarized the performance evaluation issue in detail, analyzing different symmetry properties of trellis code design.

Optimal codes may be selected by performing an exhaustive search [64]. For a given constraint length, we consider every possible generator polynomial, and choose the polynomial that gives the largest d_{min} . No other method is known that produces the optimal code. However, as the code complexity increases, an exhaustive search become impractical. In such cases, limited searches are necessary, even if they do not always provide the optimal code. One simple limited search algorithm is to choose a code at random. This natural method [68] has been reported recently. More detailed works and references on TCM can be found in [69].

For a photon counting channel, Forestieri *et al.* [70] suggested several convolutional coded PPM schemes, Georghiades [49][71] considered trellis coded OPPM and trellis

coded MPPM, and Pottie [72] applied the Z.W. condition to find a good trellis coded OPPM based on the Georghiades metric.

In section 5.2, we consider convolutional coded PPM and derive an upper bound for the d_{min} of convolutional coded PPM. In section 5.3, we calculate the coding gain of trellis-coded OPPM. In section 5.4, we present a design procedure and computer search results for new trellis codes based on MPPM, and derive an approximate upper bound for the d_{min} of trellis-coded MPPM.

5.2 CONVOLUTIONAL CODED PPM

In this section, we will combine convolutional encoder with PPM. Uncoded PPM is a power-efficient modulation, and its error rate performance will be improved further using coding, at the cost of increased bandwidth.

5.2.1 Rate $1/\log_2 L$ Convolutional Coded L -PPM

One method for combining a convolutional code and PPM is a rate $1/\log_2 L$ convolutional code followed by a L -PPM encoder [70]. For every information bit coming in, a single pulse is transmitted. Thus, the symbol rate and bit rate are identical, and the required bandwidth is roughly L/T , so the normalized bandwidth requirement is:

$$B_{CC, 1/\log_2 L} / R_b = L. \quad (5-1)$$

The bandwidth increases linearly with L .

A simple convolutional coded 4-PPM and one stage of its trellis diagram are shown in Fig. 5-2. Associated with each transition on the trellis diagram is a 4-PPM codeword.

Consider a rate $1/\log_2 L$ convolutional code with constraint length ν (i.e., with ν memory elements). We will use a similar argument as in [73]. An information sequence of length K bits results in a sequence of $K + \nu$ PPM symbols. Because of the encoder memory, two information sequences that differ by only a single bit will result in two trellis paths that disagree in exactly $\nu + 1$ consecutive transitions. For example, the trellis in Fig. 5-3 shows the two paths corresponding to the all-zero sequence and a single one bit. Two paths can differ in more than $\nu + 1$ consecutive transitions, but never less. As illustrated in Fig. 5-3, the Hamming distance between any two branches is either zero or 2. Therefore, an upper bound on the squared minimum Euclidean distance for convolutional coded PPM is:

$$d_{min}^2 \leq 2(\nu + 1). \quad (5-2)$$

If we design the convolutional encoder so that the distance in each of the $\nu + 1$ transitions is always 2, then the Hamming distance due to a one-bit error is $2(\nu + 1)$. In particular, the squared minimum Euclidean distance for the convolutional encoded PPM system of Fig. 5-2 is $d_{min}^2 = 6$, so it satisfying with equality.

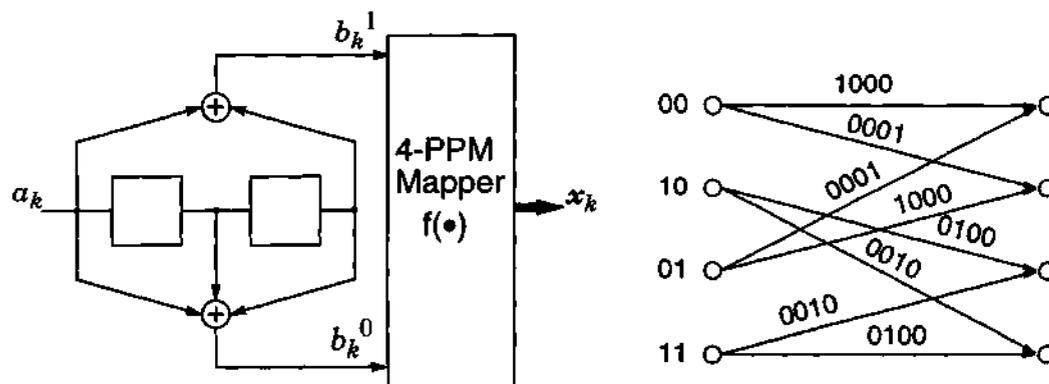


Fig. 5-2. A convolutional coded 4-PPM and its trellis.

The PPM waveform during each symbol period has the form of (2-9), where the constant a is chosen so that the total average power is P ; since there are L chips per symbol, $a = P\sqrt{L/R_b}$. The minimum Euclidean distance between coded PPM sequences is scaled by a . Therefore, the average power requirement for convolutional coded PPM is:

$$P_{CC, 1/\log_2 L} / P_{OOK} = \frac{2}{\sqrt{Ld_{min}^2}} = \sqrt{\frac{2}{L(v+1)}}, \quad (5-3)$$

the last approximation being valid when (5-2) is approximately satisfied with equality.

5.2.2 Rate $(\log_2 L - 1)/\log_2 L$ Convolutional Coded L -PPM

A rate $(\log_2 L - 1)/\log_2 L$ convolutional code can also be combined with PPM. The $(\log_2 L - 1)$ bits of information are shifted into the encoder at each symbol period and the $\log_2 L$ bits of the encoder output are converted to a L -PPM symbol. The chip duration is T/L and the bit rate is $R_b = (\log_2 L - 1)/T$, so the bandwidth expansion factor is:

$$\frac{B_{CC, (\log_2 L - 1)/(\log_2 L)}}{R_b} = \frac{L}{\log_2 L - 1}. \quad (5-4)$$

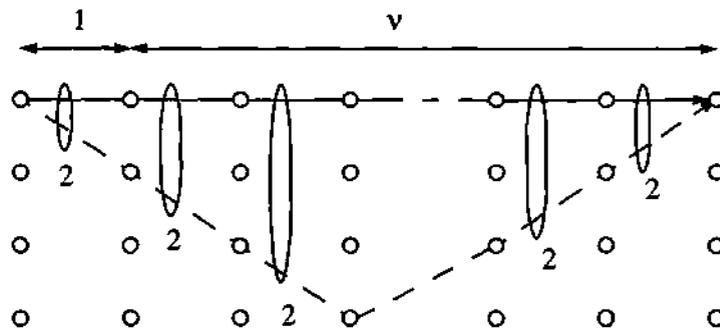


Fig. 5-3. An error event with Hamming distance $2(v+1)$.

Following the same argument used in section 5.2.1, we can derive the upper bound for the minimum distance. The minimum distance of convolutional coded PPM is the smallest among the distances of pairs of sequences arising from an error event. Each trellis path of length l involves l PPM codewords. We define the trellis path vector $\mathbf{X} = [\mathbf{x}_0^T, \mathbf{x}_1^T, \dots, \mathbf{x}_{l-1}^T]^T$ of dimension lL , where \mathbf{x}_i is an L -PPM codeword. If S_a and S_b are two trellis states that are connected by a path of length l , then the number of trellis paths is $M(l) = 2^{\kappa l - \nu}$ if $l \geq \nu$ [74] where $\kappa = (\log_2 L - 1)$. The minimum distance is less than an average distance:

$$d_{min}^2 = \min_{\substack{i,j \\ i \neq j}} \{\|\mathbf{X}^i - \mathbf{X}^j\|^2\} \leq d_{avg}^2 = \frac{1}{M(M-1)} \sum_{i=0}^{M-1} \sum_{\substack{j=0 \\ j \neq i}}^{M-1} \|\mathbf{X}^i - \mathbf{X}^j\|^2 \quad (5-5)$$

$$= \frac{1}{M(M-1)} \sum_{i=0}^{M-1} \sum_{\substack{j=0 \\ j \neq i}}^{M-1} [\|\mathbf{X}^i\|^2 + \|\mathbf{X}^j\|^2 - 2(\mathbf{X}^i, \mathbf{X}^j)]. \quad (5-6)$$

When two trellis paths, \mathbf{X}^i and \mathbf{X}^j , are orthogonal to each other, $(\mathbf{X}^i, \mathbf{X}^j) = 0$, (5-6) is upper-bounded:

$$d_{min}^2 \leq \frac{2}{M} \sum_{i=0}^{M-1} \|\mathbf{X}^i\|^2 = 2l = 2 \left(1 + \left\lfloor \frac{\nu}{\kappa} \right\rfloor \right), \quad (5-7)$$

where the first equality follows from fact the that every trellis \mathbf{X}^i contains l weights, κ is equal to $(\log_2 L - 1)$, and $\lfloor \cdot \rfloor$ takes the integer part of its argument. We can see that for $\kappa = 1$ ($L = 4$), the upper bound (5-7) reduces to (5-2).

In comparison, the simplex bound is [74]:

$$d_{min}^2 \leq \min_{t \geq 1} \frac{2^{\kappa t + 1}}{2^{\kappa t} - 1} \left(1 + \left\lfloor \frac{\nu}{\kappa} \right\rfloor \right). \quad (5-8)$$

Fig. 5-4 shows the calculated minimum distance of convolutional coded PPM with its upper bounds (5-2) and (5-7), and simplex bound (5-8). We can see that our upper bound is tighter than the simplex bound. To calculate the d_{min} of rate 1 / 3 convolutional coded 8-PPM and rate 1 / 4 convolutional coded 16-PPM, we use the generator polynomial listed in Table 11.1 of [75]. For rate 1 / 3 convolutional coded 8-PPM and rate 1 / 4 convolutional coded 16-PPM, the calculated d_{min} achieves the upper bound up to $v = 10$ and $v = 12$ respectively. For rate 2 / 3 convolutional coded 8-PPM and rate 3 / 4 convolutional coded 16-PPM, we use the search results listed in [14]. In these cases, the search results of d_{min} achieve the upper bound for every v we consider. As in (2-9), the constant α is chosen so that the total average power is P ; unlike (2-9), however, there are L chips per symbol, not $\log_2 L$, and the symbol period is $T = (\log_2 L - 1) / R_b$, not $1 / R_b$, so that $\alpha = P \sqrt{L(\log_2 L - 1) / R_b}$. Therefore, d_{OOK} / d_{min} gives the power requirement as:

$$P_{CC, (\log_2 L - 1) / (\log_2 L)} / P_{OOK} = \frac{2}{\sqrt{L\kappa d_{min}^2}} \approx \sqrt{\frac{2}{L\kappa \left(1 + \left\lfloor \frac{v}{\kappa} \right\rfloor\right)}}, \quad (5-9)$$

the last approximation being valid when $d_{min}^2 \approx 2 \left(1 + \left\lfloor \frac{v}{\kappa} \right\rfloor\right)$, where $\kappa = (\log_2 L - 1)$. From (5-4) and (5-9), we see that, with respect to bandwidth efficiency and power efficiency, the rate $(\log_2 L - 1) / \log_2 L$ encoders are better than the rate $1 / \log_2 L$ encoders.

We also compare the performance of convolutional coded 16-PPM with the cutoff rate bound. The cutoff rate bound for a coded system is defined as the SNR at which the cutoff rate, R_0 , is equal to the same number of information bits per symbol period [68]. We cal-

calculate the R_0 bound using the equation (4-5). For convolutional coded 16-PPM, 3 information bits are transmitted per symbol period. The result is shown in Fig. 5-5. For

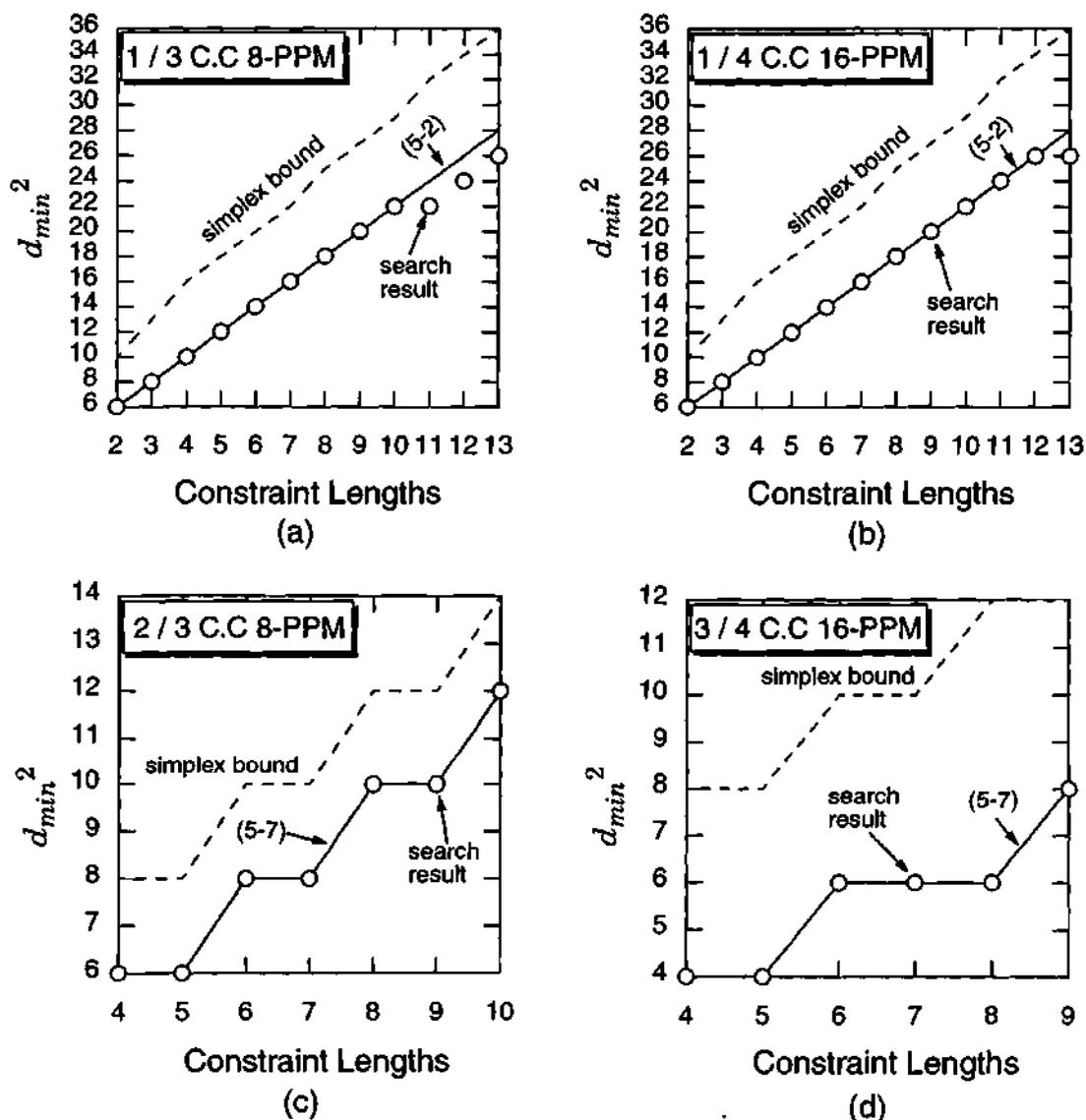


Fig. 5-4. The calculated minimum distance (denoted by circles) of convolutional coded PPM with its upper bounds (5-2) and (5-7) (denoted by solid line), and simplex bound (5-8) (denoted by dashed line).

convolutional coded 16-PPM with $v = 6$, the SNR needed to achieve 10^{-6} BER is 0.6 dB more than that of the cutoff rate bound. However, uncoded 8-PPM, which has the same number of information bits per symbol period as coded 16-PPM, requires 3.7 dB more than the cutoff rate bound.

5.3 TRELLIS CODED OPPM

If we use a convolutional code to reduce the probability of error, there is an inevitable increase in bandwidth, as derived in the section 5.2. It is well known that trellis-coded

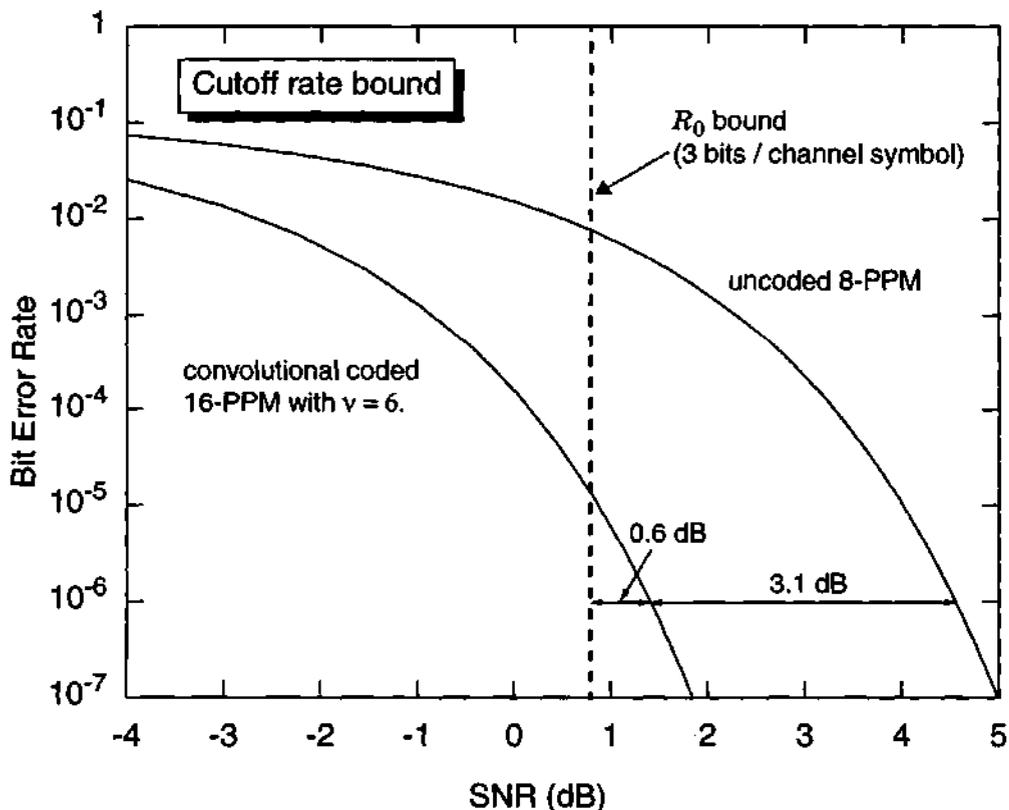


Fig. 5-5. Convolutional coded 16-PPM with cutoff rate bound.

modulation (TCM) is a technique that improves performance without increasing the bandwidth. Because PPM has the same Hamming distance between any two codewords, no gains can be made through set partitioning. OPPM is an attractive alternative since it has a low duty cycle and equal energy signals [49][71]. But doubling the number of $(L/2)$ -OPPM symbols without increasing bandwidth, as Ungerboeck suggested [64], requires that the overall duty cycle $\alpha = w/n$ remain fixed, and so the number of slots in each baud interval must increase from $n_u = (L/2 - 1)/(1 - \alpha)$ to $n_c = (L - 1)/(1 - \alpha)$, which, in turn, decreases the minimum distance. Ideally, the coding gain achieved through set partitioning will be large enough to compensate for the decreased minimum distance. Trellis-coded L -OPPM has a bit rate of $R_b = \log_2(L/2) / T$ and a bandwidth expansion factor of:

$$B_{TC, L-OPPM} / R_b = \frac{n/w}{\log_2(L/2)}, \quad (5-10)$$

which is the same as $L/2$ -OPPM. However, now the required power is:

$$P_{TC, L-OPPM} / P_{OOK} = \sqrt{\frac{4(L-1)\alpha^2}{d_{min,c}^2 \cdot (1-\alpha)\log_2(L/2)}}. \quad (5-11)$$

In contrast, the requirement for uncoded $(L/2)$ -OPPM is:

$$P_{OPPM} / P_{OOK} = \sqrt{\frac{4(L/2-1)\alpha^2}{d_{min,u}^2 \cdot (1-\alpha)\log_2(L/2)}}, \quad (5-12)$$

where $d_{min,c}$ and $d_{min,u} = \sqrt{2}$ are the minimum Euclidean distance for the coded and uncoded systems respectively. The asymptotic coding gain is [28]:

$$\text{Asymptotic Coding Gain} = 10\log_{10}\left(\sqrt{\left(\frac{L/2-1}{L-1}\right) \cdot \frac{d_{min,c}^2}{2}}\right) \approx 10\log_{10}\left(\frac{d_{min,c}}{2}\right), \quad (5-13)$$

where the first term represents loss due to expansion in size of constellations and the second term is gain due to coding. To get improved performance using TCM, the minimum Euclidean distance must be greater than about 2.

The signal sets for 8-OPPM ($L = 8$, $n_c = 14$, $w_c = 7$) and its set partitioning are shown in Fig. 5-6. If we use a 4-state TCM, the minimum Euclidean distance is $\sqrt{8}$, which is the distance for a parallel transition resulting from the set partitioning. From (5-13), therefore, the asymptotic coding gain relative to uncoded 4-OPPM (with $n_u = 6$, $w_u = 3$) is 1.2 dB.

We have seen that trellis coded OPPM improves the power efficiency without bandwidth expansion. But, the improvement is not appreciable due to its relatively small number of codewords given the length of codeword.

5.4 TRELLIS CODED MPPM

In the previous sections, we combined coding with PPM or OPPM. Convolutional coded PPM improves the power efficiency with a large sacrifice in bandwidth efficiency. Trellis coded OPPM achieves better power efficiency than the uncoded OPPM without bandwidth expansion. But, as we indicated before, this scheme does not provide us with good power efficiency compared to the other coded modulation schemes.

In this section, we will develop new trellis codes based on MPPM. Since MPPM is both power-efficient and bandwidth-efficient, trellis-coded MPPM improve the power-efficiency without significantly increasing the bandwidth.

5.4.1 Design of Trellis Coded MPPM

In general, $\log_2\binom{n}{w}$ is not an integer, therefore MPPM can not be used straightforwardly at the output of convolutional encoder. The design rule is to choose n and w so that $\binom{n}{w}$ is close to the power of 2 in order to minimize the loss of throughput. For $w > 2$, n must be too large to get good power efficiency. Therefore, we choose $w = 2$ as a good starting point. The MPPM codewords with $w = 2$, can be mapped into a plane. Fig. 5-7 shows $\binom{n}{2}$ -MPPM codewords for $n = 5, 7, 9, 13$, and 17. This idea is similar to [76], where an optical pulse is mapped into a plane by assigning the starting time of a pulse to the x -

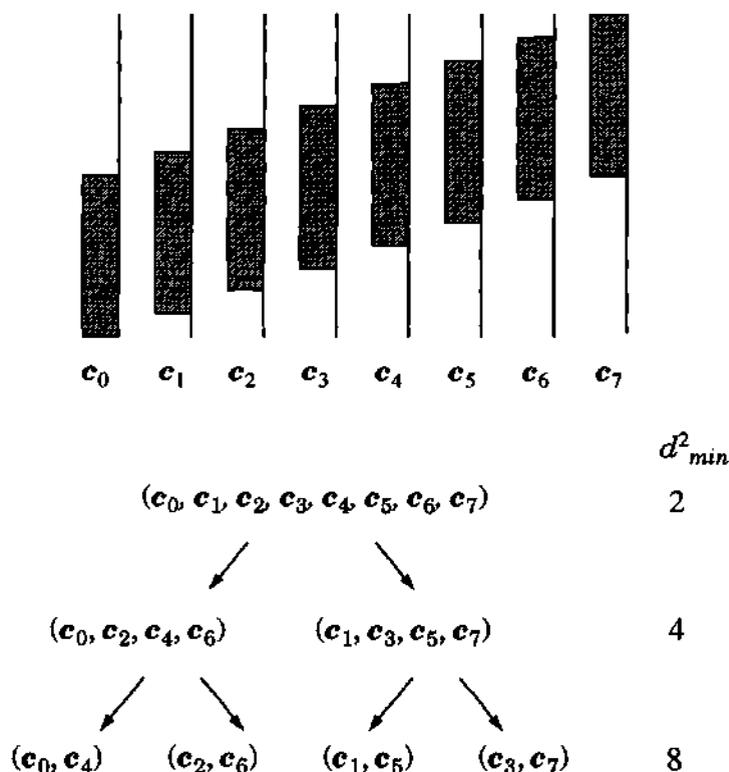


Fig. 5-6. The 8-OPPM signal set and its set partitioning [71].

axis and the ending time to the y -axis. In our case, the x -axis represents the position of first pulse, and the y -axis represents the position of second pulse. For example, codeword $[1\ 1\ 0\ 0\ 0]^T$ is mapped into $(1, 2)$, $[0\ 0\ 1\ 0\ 1]^T$ is mapped into $(3, 5)$, and so on. Note that Ham-

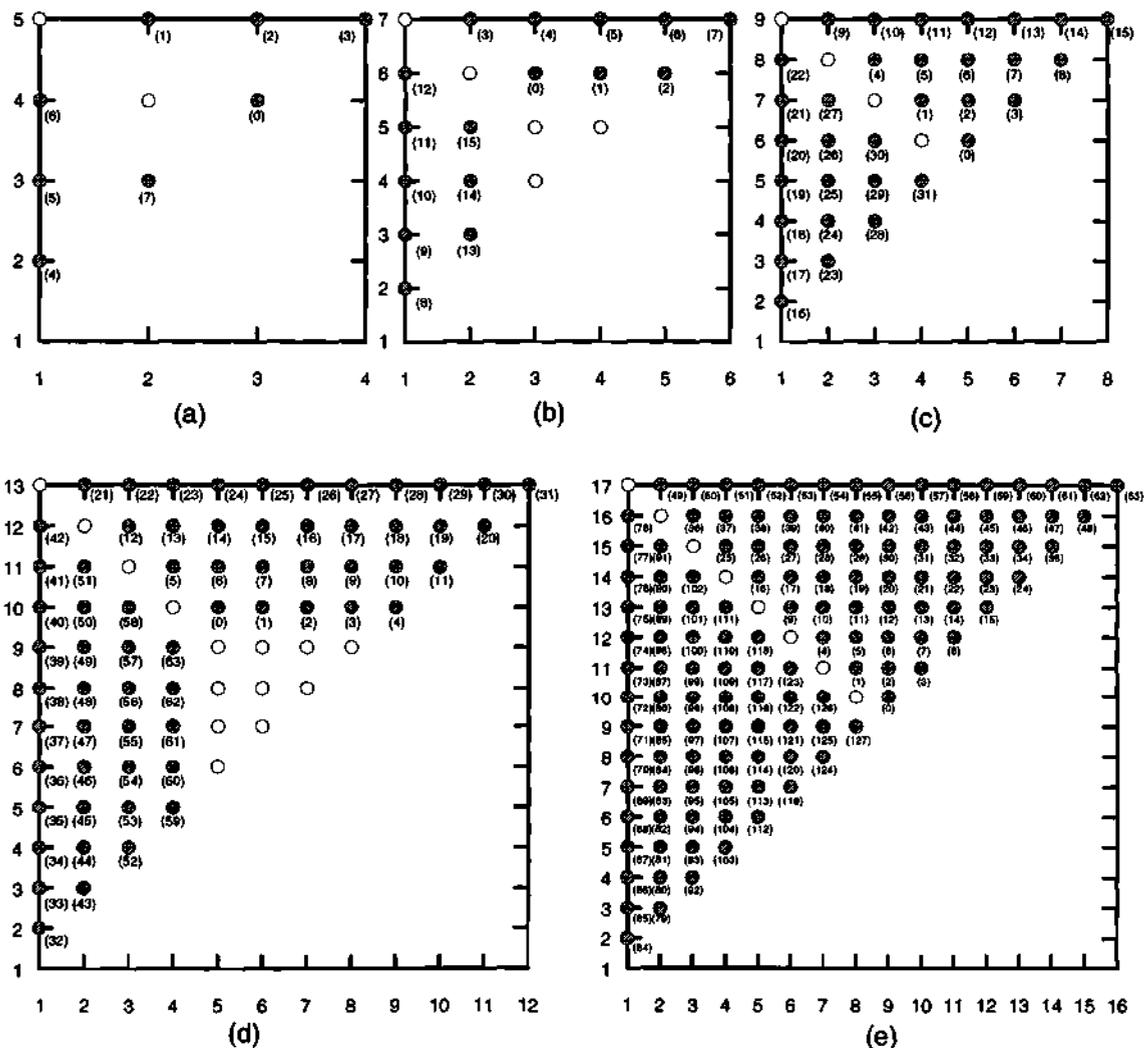


Fig. 5-7. Constellations for $\binom{n}{2}$ -MPPM; the shaded circles represent the chosen L codewords and unshaded circles represent the unused codewords. The x -axis represents the position of first pulse and the y -axis represents the position of second pulse. (a) $L = 8$ and $n = 5$, (b) $L = 16$ and $n = 7$, (c) $L = 32$ and $n = 9$, (d) $L = 64$ and $n = 13$, (e) $L = 128$ and $n = 17$.

ming distance between the two codewords having the same x -component or y -component is two, and Hamming distance between the two codewords having different x -component and y -component is four.

The Zehavi and Wolf condition requires that when the signal set is partitioned into two subsets, $S^{(1)}$ and $S^{(2)}$, the distance weight profiles of the two subsets be identical. The distance weight profiles of a subset with respect to $\log_2 L$ -tuple binary error vector \mathbf{e} is defined as [66][67]:

$$F(S^{(i)}, \mathbf{e}, Z) = \sum_d N_d Z^{d_H(f(\mathbf{b}), f(\mathbf{b} \oplus \mathbf{e}))} \quad \text{for } i = 1, 2, \quad (5-14)$$

where N_d is the number of codewords having Hamming distance d between the codeword $f(\mathbf{b})$ and codeword $f(\mathbf{b} \oplus \mathbf{e})$, and the summation is taken over all the possible d . If trellis code satisfies this Z.W. condition, then d_{min} can be calculated from the all-zero path. In Fig. 5-7, the shaded circles represent the L selected codewords, and the number below each constellation point represents the label of the codeword. Fig. 5-8 shows set partitioning of 8-MPPM. The 8-MPPM signal set is partitioned into two subsets, $S^{(1)} = \{0, 2, 4, 6\}$, and $S^{(2)} = \{1, 3, 5, 7\}$. The distance weight profile of these subsets are listed in Table 5-1, and they are identical. Therefore the resulting 8-MPPM satisfies the Z.W. condition. By calculating the distance weight profile (5-14), we can prove that the other MPPM signal sets, shown in Fig. 5-7 also satisfy the Z.W. condition.

We choose $\binom{17}{2}$ -MPPM as an uncoded modulation scheme upon which to design a trellis code. From $\binom{17}{2} = 136$ MPPM codewords, we choose 128 codewords to combine with the rate $6/7$ convolutional encoder, as shown in Fig. 5-7-e. For trellis coded 128-

MPPM, a random search is executed over systematic feedback convolutional encoder as shown in Fig. 5-9. The rate 6/7 convolutional encoder operates on 6 bits, $\alpha_k = [a_k^1, a_k^2, a_k^3, a_k^4, a_k^5, a_k^6]^T$, and produces 7 encoded bits, a labeling vector $b_k = [b_k^0, b_k^1, b_k^2, b_k^3, b_k^4, b_k^5, b_k^6]^T$. The coded bits are mapped into one of the 128-MPPM code-words according to the mapping rule $f(\bullet)$. We choose a systematic feedback configuration, because it has a smaller number of coefficients to search than that of feedforward configuration, also it is free from catastrophic condition [64]. For 128-MPPM with constraint length v , the number of coefficients to search is $2^{7(v-1)}$, making it almost impossible to perform an exhaustive search for a large constraint length.

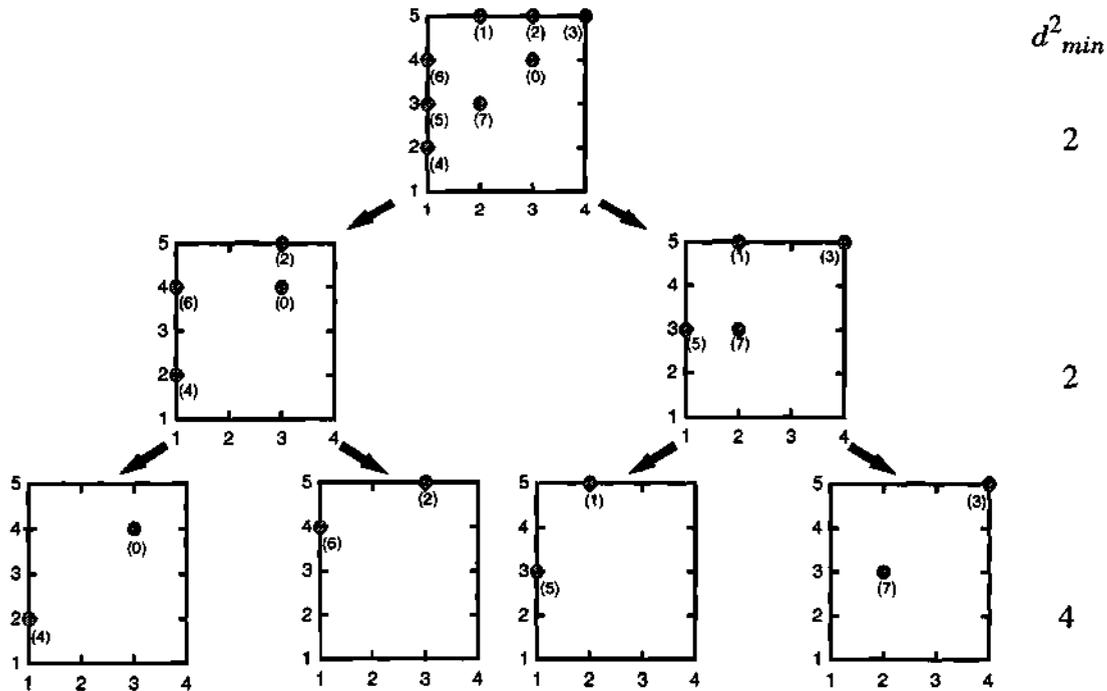


Fig. 5-8. Set partitioning of 8-MPPM.

TABLE 5-1: The distance weight profile of trellis coded 8-MPPM.

Error vector e	$S^{(1)} = \{0, 2, 4, 6\}$	$S^{(2)} = \{1, 3, 5, 7\}$
000	4	4
001	$2Z^2 + 2Z^4$	$2Z^2 + 2Z^4$
010	$4Z^2$	$4Z^2$
011	$4Z^2$	$4Z^2$
100	$4Z^4$	$4Z^4$
101	$4Z^2$	$4Z^2$
110	$2Z^2 + 2Z^4$	$2Z^2 + 2Z^4$
111	$2Z^2 + 2Z^4$	$2Z^2 + 2Z^4$

The random search results are shown in Table 5-2. We compare this trellis coded MPPM with the uncoded modulation that has same bandwidth efficiency. Trellis coded L -MPPM has a bit rate of $R_b = \log_2(L/2)/T$, and bandwidth and power efficiencies are:

$$\frac{B_{TC, L-MPPM}}{R_b} = \frac{n}{\log_2(L/2)}, \quad (5-15)$$

$$\frac{P_{TC, L-MPPM}}{P_{OOK}} = \frac{2w}{\sqrt{n \cdot d_{min,c}^2 \cdot \log_2(L/2)}}, \quad (5-16)$$

where $d_{min,c}$ is the minimum Euclidean distance of coded sequences. Define the integer \tilde{L} so that \tilde{L} -PPM has the same bandwidth efficiency as trellis coded L -MPPM. The

asymptotic coding gain over \tilde{L} -PPM, with the same bandwidth efficiency and minimum Euclidean distance $d_{min,u} = \sqrt{2}$, is:

$$\text{Asymptotic Coding Gain} = 10 \log_{10} \left(\sqrt{\frac{n \log_2(L/2)}{w^2 \tilde{L} \log_2 \tilde{L}} \cdot \frac{d_{min,c}^2}{2}} \right), \text{ (dB)} \quad (5-17)$$

where the first term represents loss due to expansion in size of constellations and the second term is gain due to coding. We find $\tilde{L} = 9$ for PPM since trellis coded 128-MPPM and 9-PPM have approximately the same bandwidth efficiency, $B / R_b \approx 2.8$ in both cases. Using the constraint lengths 4, 7, and 12, we achieve an asymptotic coding gain relative to uncoded 9-PPM of 1.4, 2.3, and 2.9, respectively.

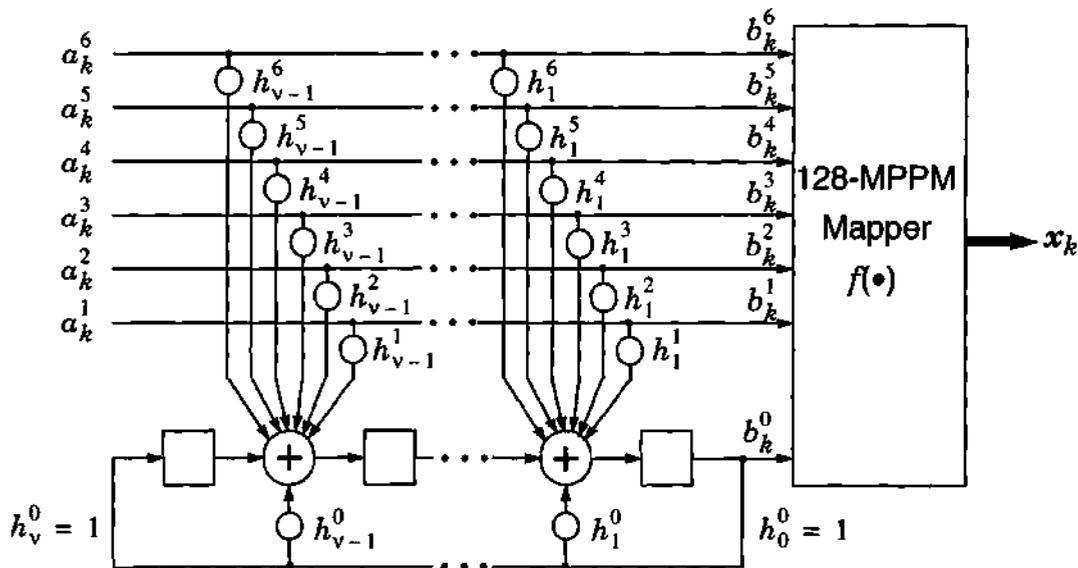


Fig. 5-9. Systematic feedback encoder for trellis coded 128-MPPM with constraint length v and parity check coefficients h_m^n , $m = 1, 2, \dots, v-1$ and $n = 0, 1, \dots, 6$.

TABLE 5-2: Parity check coefficients for trellis coded 128-MPPM in octal form. Coding gains are calculated as compared to the uncoded 9-PPM which has the same bandwidth efficiency as coded 128-MPPM. Random search is used to find the coefficients of trellis codes. To validate the search result, we compare this with the \tilde{d}_{avg}^2 and the simplex bound. Note that the performance of trellis codes for $v = 5$ and 6 are no better than $v = 4$, and trellis codes for $v = 8, 9, 10$, and 11 are no better than $v = 7$.

v	Parity Check Coefficients							d_{min}^2	P_{req} (dB)	Cod- ing Gain [dB]	\tilde{d}_{avg}^2 (Sim- plex)
	h^0	h^1	h^2	h^3	h^4	h^5	h^6				
4	23	10	06	14	16	00	04	4	-7.0	1.4	4 (4)
5	65	14	34	22	12	20	16	4	-7.0	1.4	4 (4)
6	103	054	072	016	014	024	066	4	-7.0	1.4	7 (8)
7	357	144	014	024	040	140	102	6	-7.9	2.3	7 (8)
8	443	102	040	064	276	022	164	6	-7.9	2.3	7 (8)
9	1057	0516	0324	0546	0720	0604	0206	6	-7.9	2.3	7 (8)
10	3341	1406	1330	0176	1266	1746	1156	6	-7.9	2.3	7 (8)
11	7433	0736	1162	1316	2032	2272	3302	6	-7.9	2.3	7 (8)
12	10017	02102	05242	03314	01374	07550	01044	8	-8.5	2.9	11 (12)

5.4.2 An Approximation for the d_{min} of Trellis Coded MPPM

In this section, we derive an approximation for the minimum distance of trellis coded MPPM for a given constraint length. The minimum distance of a trellis code is the smallest among the distances of pairs of sequences arising from an error event. Each trellis path associated with the error event of length l involves l MPPM codewords. We define the trellis path vector, $\mathbf{X} = [\mathbf{x}_0^T, \mathbf{x}_1^T, \dots, \mathbf{x}_{l-1}^T]^T$ of dimension ln where \mathbf{x}_i is the $\binom{n}{w}$ -MPPM codeword corresponding to the i -th branch in the path. Observe that the trellis vector \mathbf{X} is a valid $\binom{N}{\Omega}$ -MPPM codeword with length $N = ln$, and weight $\Omega = lw$. As we indicated in section 2.1.3, any valid MPPM codeword has the same set of distance with respect to the other codewords, and the number of codewords for $\binom{N}{\Omega}$ -MPPM with mutual Hamming distance $2m$ is $N_m = \binom{\Omega}{m} \binom{N-\Omega}{m}$. We can calculate the average distance for $\binom{N}{\Omega}$ -MPPM as:

$$\tilde{d}_{avg}^2 = \frac{1}{L(L-1)} \sum_{i=0}^{L-1} \sum_{\substack{j=0 \\ j \neq i}}^{L-1} \|\mathbf{X}^i - \mathbf{X}^j\|^2, \quad (5-18)$$

$$= \frac{1}{L-1} \sum_{j=1}^{L-1} \|\mathbf{X}^0 - \mathbf{X}^j\|^2, \quad (5-19)$$

$$= \frac{1}{L-1} \sum_{m=1}^{\Omega} 2m N_m, \quad (5-20)$$

$$= \frac{2 \binom{\Omega}{1} \binom{N-\Omega}{1} + 4 \binom{\Omega}{2} \binom{N-\Omega}{2} + \dots + 2\Omega \binom{\Omega}{\Omega} \binom{N-\Omega}{\Omega}}{L-1}, \quad (5-21)$$

where $L = \binom{N}{\Omega}$ is the number of extended MPPM codeword. Since not all valid $\binom{N}{\Omega}$ -MPPM codewords are included in the set of $\{\mathbf{X}\}$, the \bar{d}_{avg}^2 in (5-18) is only an *approximation* for d_{min}^2 .

We also can apply this approximation method to trellis-coded PPM by treating the trellis-coded PPM sequences of length l as extended MPPM codewords with length $N = lL$ and weight $\Omega = l$, and then apply (5-21).

The approximation based on average distance are listed in Table 5-2. For comparison, we calculated the simplex bound:

$$d_{min}^2 \leq \min_{t \geq 1} \frac{2^{\kappa t + 1}}{2^{\kappa t} - 1} w \left(1 + \left\lfloor \frac{v}{\kappa} \right\rfloor \right), \quad (5-22)$$

where $\kappa = (\log_2 L - 1)$, and $\lfloor \bullet \rfloor$ takes the integer part of its argument. We can see that approximation method is tighter than the simplex bound.

5.5 PERFORMANCE OF CODED MODULATION SCHEMES

We calculate power efficiency and bandwidth efficiency of coded modulation schemes using (5-1), (5-3), (5-4), and (5-9) for convolutional coded PPM; (5-10), (5-11) for trellis coded OPPM; and (5-15), (5-16) for trellis coded MPPM. We assume each coded modulation has $d_{min}^2 = 4, 6, \dots, 16$. We ignore the complexity problem to achieve this, since this assumption may require a large constraint length for some coded modulation schemes. The result along with MPPM bound is shown in Fig. 5-10. The most power-efficient scheme is convolutional coded 16-PPM with bandwidth efficiency of $B / R_b = 5.3$. The most bandwidth-efficient scheme is trellis coded 8-OPPM, but it has poor power effi-

ciency. The next power-efficient scheme with moderate bandwidth efficiency of $B / R_b = 2.83$ is trellis coded 128-MPPM. Note that this scheme requires approximately a half ($2.83/5.3 = 0.53$) of the bandwidth required by the convolutional coded 16-PPM. There-

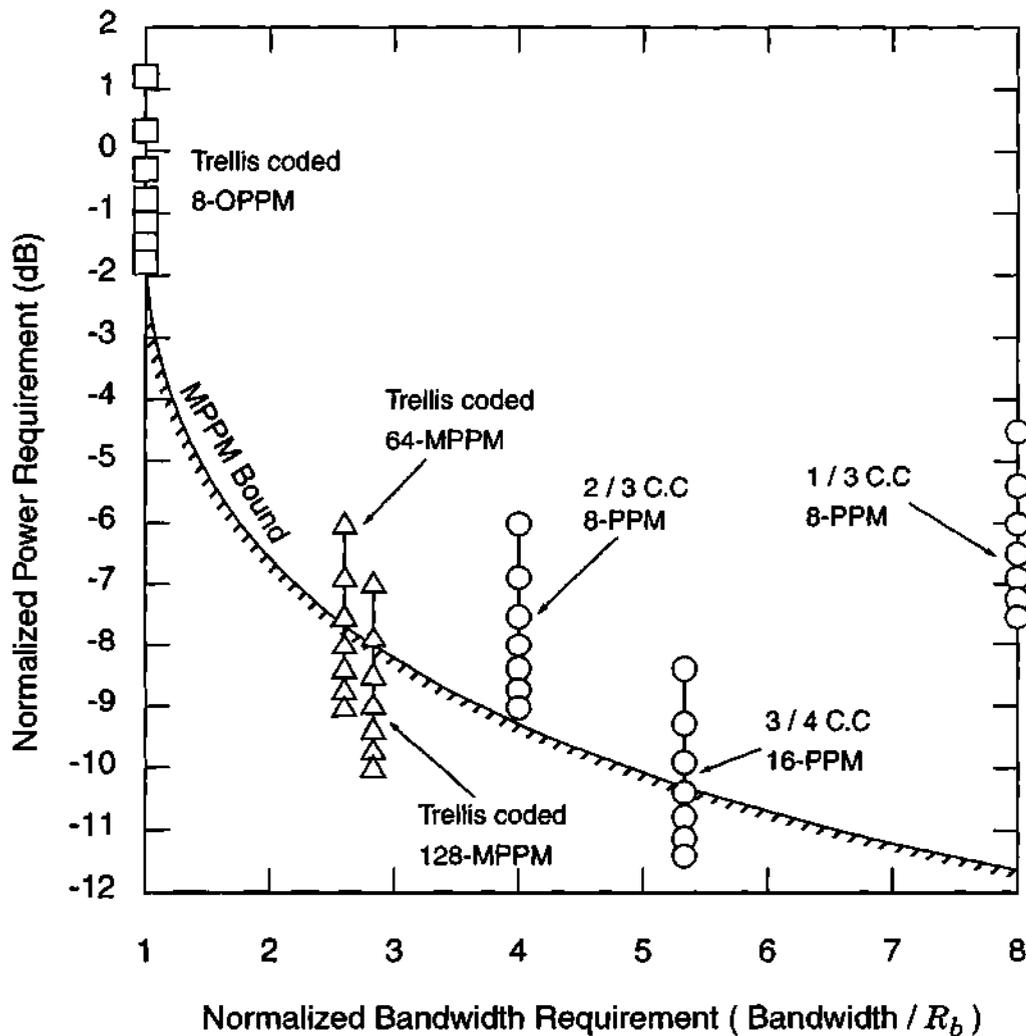


Fig. 5-10. Power efficiency and bandwidth efficiency of coded modulation.

fore, over a multipath channel trellis coded 128-MPPM is very promising, because it is less susceptible to ISI than convolutional coded 16-PPM.

5.6 SUMMARY AND CONCLUSION

We derived upper bounds for the minimum distance of convolutional coded PPM and an approximation for the minimum distance of trellis coded MPPM, and compared our results with the well-known simplex bound. Numerical results showed that our upper bounds are tighter than the simplex bound. We designed new trellis codes based on MPPM, and performed a code search for the parity check polynomials of trellis-coded MPPM which give a large minimum distance. We calculated the power efficiency and bandwidth efficiency for coded modulation schemes. Of all the schemes considered, trellis coded 128-MPPM, both power-efficient and bandwidth-efficient, is most suitable for our applications.

CHAPTER 6

C O D E D M O D U L A T I O N O N A M U L T I P A T H C H A N N E L

In this chapter, we will evaluate the performance of the coded modulation schemes developed in the previous chapter on a multipath channel using several equalization strategies developed in chapter 3.

6.1 INTRODUCTION

Decoding of TCM over an ISI channel via the MLSD is an optimal way to jointly process ISI and coding. TCM in the presence of ISI is modeled by a single finite state

machine consisting of a TCM encoder, the channel, and a WMF [77][78][79]. The complexity of this super-trellis, which results from the combination of TCM and ISI, is often very large. Reduced states sequence detection (RSSD) considered for uncoded sequence detection can also be used for coded sequences. The complexity of the RSSD trellis ranges between that of the encoder trellis and that of the super-state trellis. When we apply the Viterbi algorithm to the original trellis encoder, but perform decision-feedback equalization on each survivor path in the trellis based on the history of that path [80], we implement parallel decision feedback decoding (PDFD). If N states exist in the trellis, then N distinct postcursor equalization filters are used. Each filter uses the decisions from one of the N survivor paths to construct the next decoder input.

Recently, Lee *et al.* [14] suggested trellis coded PPM which applied the Ungerboeck set partitioning rule to the PPM signal set accounting for the effects of multipath dispersion.

Conventional LE or DFE can be used with the Viterbi decoder at the receiver to decode the trellis encoded sequences in the presence of ISI. Thapar [81] applied linear equalization (LE) for voiceband telephone modem, and Wong and McLane [82] used fractional spaced LE for ISI channels that contain an in-band spectral null, such as high frequency (HF) radio systems. However, on channels with severe attenuation, the performance of LE is poor because LE enhances and correlates the noise. By using DFE, noise enhancement can be substantially reduced. Also, ideal ZF-DFE results in an ideal AWGN channel. Price [42] first observed that as long as ideal zero-forcing decision feedback equalization (ZF-DFE) is used at the receiver, the SNR gap to capacity at high signal-to-noise ratio is the same for ISI channels as it is for ideal channels. Thus, channel coding, such as trellis coded modulation (TCM), has potentially the same benefit on channels with ISI as on

ideal channels. The problem using channel coding with DFE is that DFE requires immediate reliable decisions and cannot tolerate decoding delay. Precoding is a technique similar to DFE in that it eliminates error propagation by moving cancellation of the postcursor ISI from the receiver to the transmitter [45][46]. Ideally, using precoding, the coding gain of the trellis code is obtained in combination with the equalization performance of the ideal ZF-DFE [42]. Trellis precoding was invented by Eyuboglu and Forney [83] as a generalization of TH precoding. Recently, Laroia, Tretter, and Farvadin (LTF) [47][48] introduced a new precoding scheme which is comparable in performance to trellis precoding [52]. But, as we indicated in section 3.6, a precoding scheme based on MPPM cannot achieve the ideal performance of ZF-DFE. Instead, the result of precoding approximately truncates the channel memory length to one. We called this scheme partial-response precoding. We can use a simple equalizer at the receiver to remove this ISI and decode the trellis-coded signal.

In this chapter, we consider five equalization strategies:

- LE with VA
- DFE with VA
- superstate MLSD
- PDFD
- partial-response precoding with PDFD.

In section 6.2, we develop a channel model for coded modulation on an ISI channel. In section 6.3, we employ MLSD on the combined trellis formed by the convolutional encoder and channel ISI to perform equalization and decoding simultaneously. In section 6.4, we consider PDFD to reduce the complexity of superstate MLSD. In

section 6.5, we combine partial-response precoding with PDFD. In section 6.6, we compare the performance of the various equalization schemes.

6.2 SYSTEM MODEL

Consider the system model shown in Fig. 6-1. Information bits with rate R_b (b/s) enter the trellis encoder, which consists of a linear convolutional encoder and a signal mapper. The $(\log_2 L - 1) / \log_2 L$ convolutional encoder transforms the input bits into coded bits and the mapper converts each block of $\log_2 L$ bits into one of L codewords $\mathbf{c}_0 \dots \mathbf{c}_{L-1}$. The output of the encoder is a sequence of codewords $\{\mathbf{x}_k\}$ with rate $1/T = R_b / \log_2 L$. This sequence is serialized to produce the binary chip sequence $\{x_j\}$ with rate n/T , where $\mathbf{x}_k = [x_{kn}, x_{kn+1}, \dots, x_{kn+n-1}]^T$. The binary chip sequence drives a transmitter filter with a rectangular pulse shape $p(t)$ of duration T/n and unity height. To satisfy the power constraint of (1-2), the filter output is multiplied by (nP/w) before the signal is sent across the channel.

As shown in Fig. 6-1, the receiver uses a unit-energy whitened matched filter $f(t)$ and samples the output at the chip rate n/T producing y_j . The receiver groups the samples y_j into blocks of length n , producing a sequence of observation vectors $\{\mathbf{y}_k\}$, where $\mathbf{y}_k = [y_{kn}, y_{kn+1}, \dots, y_{kn+n-1}]^T$. As shown in Fig. 6-1, there are several options to process \mathbf{y}_k . The receiver passes each observation vector through a LE or DFE to eliminate the ISI. Then, we perform MLSD on the trellis formed by the convolutional encoder. Or we perform MLSD on the combined trellis formed by the convolutional encoder and channel ISI to perform equalization and decoding simultaneously. Yet another option is to perform PDFD on the trellis formed by the convolutional encoder in order to perform equalization

and decoding simultaneously. In the final option, we combine the precoding scheme with PDFD to reduce the complexity further.

6.3 SUPERSTATE MLSD

The trellis encoded signal is a function of the convolutional encoder state α_k and the information bits \mathbf{a}_k :

$$\mathbf{x}_k = f(\mathbf{a}_k, \alpha_k), \quad (6-1)$$

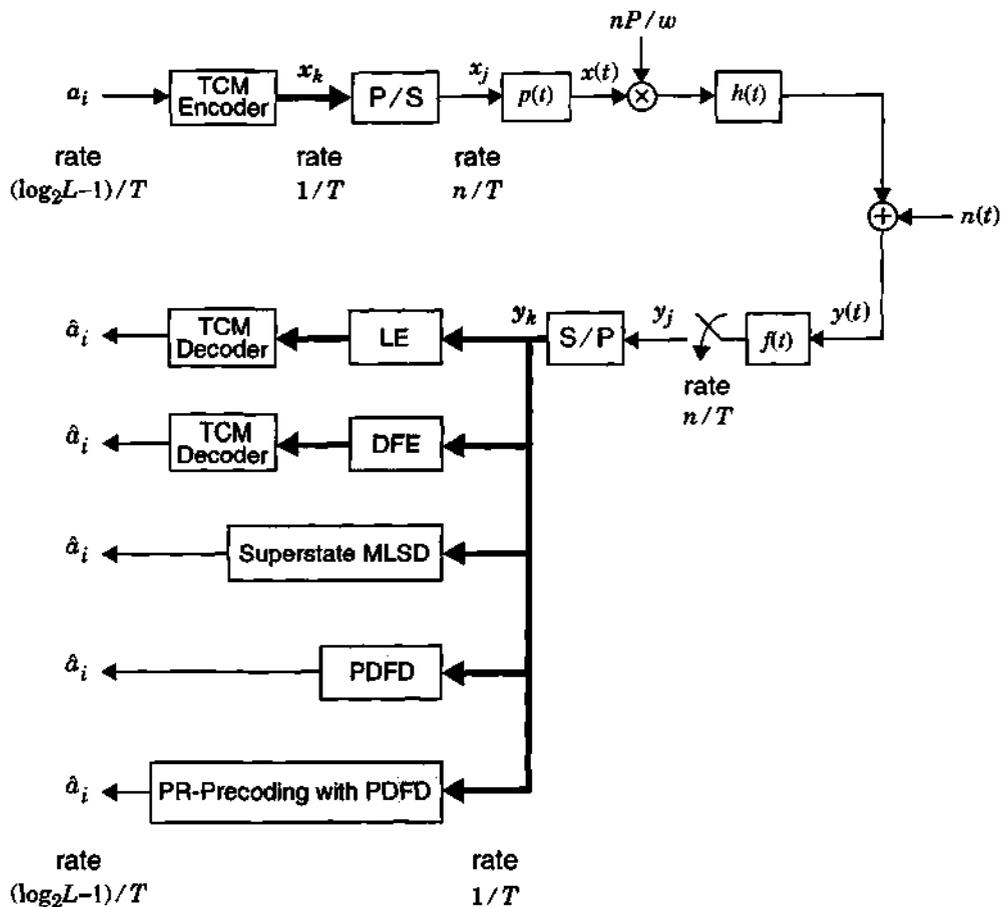


Fig. 6-1. Block diagram of coded modulation on an ISI channel.

and the state transition equation is as follows:

$$\alpha_{k+1} = g(\mathbf{a}_k, \alpha_k). \quad (6-2)$$

Trellis encoded signals in the presence of ISI are modeled as a single finite-state machine. For a rate of $(\log_2 L - 1) / \log_2 L$ trellis coded MPPM, there are $(L/2)^\mu$ ISI states associate with each encoder state. The states for the combined finite-state machine are:

$$\beta_k = (\alpha_{k-\mu}, \mathbf{a}_{k-\mu}, \mathbf{a}_{k-\mu+1}, \dots, \mathbf{a}_{k-1}). \quad (6-3)$$

If the convolutional encoder has 2^m states, the combined trellis has $L^\mu 2^{m-\mu}$ states with $L/2$ transitions emerging from each state.

We perform MLSD on the combined trellis formed by the convolutional encoder and channel ISI. In the presence of ISI and AWGN, the MLSD determines the trellis coded vector sequence, $\{\hat{\mathbf{x}}_k\}$, that minimizes the branch metric:

$$\| \mathbf{y}_k - \sum_{i=1}^{\mu} \mathbf{H}_i \hat{\mathbf{x}}_{k-i} - \mathbf{H}_0 \hat{\mathbf{x}}_k \|^2. \quad (6-4)$$

The complexity requirement of this optimum decoder grows exponentially with the channel memory μ . In practice, this makes it difficult to implement for large μ . The performance of superstate MLSD is approximated at high SNR:

$$Pr[\text{error}] \approx Q\left(\frac{d_{min}}{2\sqrt{N_0}}\right), \quad (6-5)$$

where d_{min} is the minimum Euclidean distance over the set of all possible error events in the super-trellis.

6.4 PARALLEL DECISION FEEDBACK DETECTION (PDFD)

Complexity of the ML sequence detector can be reduced by truncating the vector channel memory from μ to δ , where $0 \leq \delta \leq \mu$. A Viterbi decoder operates on a reduced number of combined encoder and ISI states. The number of states of this reduced state sequence detector (RSSD) is given by $L^\delta 2^{m-\delta}$. To compensate for the performance degradation, DFE is used in the branch metric computation. ISI terms not considered by the truncated states are estimated from the past codewords and subtracted. When $\delta = \mu$, RSSD becomes optimum superstate MLSD. When $\delta = 0$, RSSD reduces to parallel decision feedback detection (PDFD). PDFD applies the Viterbi algorithm to the original trellis encoder ($\beta_k = \alpha_k$), but performs decision-feedback equalization on each path survivor in the trellis based on the history of that path. Instead of using only one sequence of decisions in the feedback path of the DFE, equalization is accomplished by a unique sequence of decisions for each path of the trellis. The PDFD searches the trellis coded vector sequence $\{\hat{x}_k\}$ that minimizes the branch metric:

$$\|y_k - \sum_{i=1}^{\mu} \mathbf{H}_i \hat{x}_{k-i}(\alpha_k) - \mathbf{H}_0 \hat{x}_k\|^2, \quad (6-6)$$

where minimization is taken over all trellis branches originating from encoder state, $\{\alpha_k\}$, to the next state, $\{\alpha_{k+1}\}$. Since PDFD uses a reduced trellis, it is suboptimal, and it has a smaller minimum distance than the superstate MLSD.

6.5 PRECODING

Price results discussed in chapter 3 can be extended to coded modulation as follows. The SNR gap to channel capacity using the coded modulation is the same for ISI channels

at high SNR as it is for ideal channels, as long as the receiver uses ideal DFE followed by MLSD. However, to decode a trellis coded signal, DFE cannot be combined with MLSD in a straightforward way, because a DFE requires delay-free feedback decisions, which are in fact available only after long decoding delay. As shown in Fig. 6-2, we combine DFE and VA using tentative decisions: the DFE generates tentative decisions \hat{x}_k and then subtracts them from the incoming signal. The input to the MLSD is:

$$r_k = \mathbf{H}_0 \mathbf{x}_k + \sum_{j \geq 1}^{\mu} \mathbf{H}_j \mathbf{x}_{k-j} - \sum_{j \geq 1}^{\mu} \mathbf{H}_j \hat{\mathbf{x}}_{k-j} + n_k. \quad (6-7)$$

The tentative decisions can be unreliable, and it is expected that there will be error propagation which degrades performance.

When we assume correct decisions, $\hat{\mathbf{x}}_k = \mathbf{x}_k$, the trellis decoder chooses the coded vector sequence $\{\hat{\mathbf{x}}_k\}$ that minimizes the branch metric:

$$\| r_k - \mathbf{H}_0 \hat{\mathbf{x}}_k \|^2. \quad (6-8)$$

When the precoding scheme described in section 3.6 is combined with the trellis coded signal, this configuration is a general description of LTF precoding [47]. In LTF precoding with QAM signals, \mathbf{b}_k represents the signal point in the constellation of $\lambda = \mathbf{Z}^2$

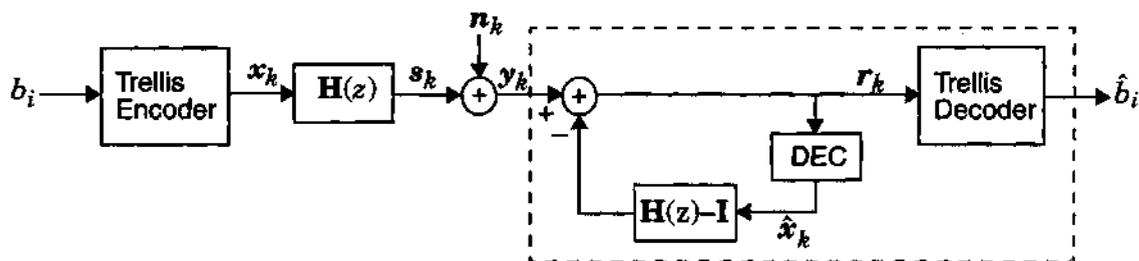


Fig. 6-2. Block diagram of DFE with tentative decision and MLSD.

+ (1 / 2, 1 / 2) and \mathbf{q}_k represents the signal point in the constellation of $\lambda_c = 2Z^2$. If a \mathbf{b}_k is valid trellis coded sequence, $\mathbf{b}_k + \mathbf{q}_k$ is also a valid trellis coded sequence.

But in our precoding scheme, $\mathbf{b}_k + \mathbf{q}_k$ at the output of the ISI channel is not a valid MPPM codeword. We can decide \mathbf{b}_k from the extended codeword $\{\mathbf{b}_k + \mathbf{q}_k\}$ using the inverse precoder. Or we treat $\mathbf{b}_k + \mathbf{q}_k$ as the output of the 1-memory ISI channel (partial-response precoding). We consider here the latter method. At the output of channel, removing the constant (3-69);

$$\tilde{\mathbf{y}}_k = \mathbf{H}_0 \mathbf{b}_k + \mathbf{H}_1 \mathbf{b}_{k-1} + \mathbf{n}_k. \quad (6-9)$$

We can use an equalization scheme to remove the truncated ISI, and to decode the trellis coded signal with reduced complexity.

6.6 NUMERICAL RESULTS

We show the performance of trellis coded 128-MPPM and 16-PPM on a multipath channel using superstate MLSD in Fig. 6-3. We use the coefficients listed in [14] for trellis coded 16-PPM. We use the same truncated vector channel considered for the uncoded case. As in the uncoded case, we calculated the optical power required to achieve a 10^{-6} bit-error rate over this ISI channel. Trellis coded 16-PPM shows better performance up to a bit-rate-to-bandwidth ratio of 0.15. But above that, trellis coded 128-MPPM outperforms trellis coded 16-PPM. At a bit-rate-to-bandwidth ratio of 1, the normalized powers for trellis coded 16-PPM with constraint lengths $\nu = 4$ and 7 are 3.4 dB and 2.9 dB, respectively. But the required normalized powers for trellis coded 128-MPPM, with $\nu = 4$ and 7, are 2 dB and 1.5 dB, respectively. Therefore, trellis coded 128-MPPM requires 1.4 dB less

power than the trellis coded 16-PPM when both schemes use the same constraint length at $R_b / W = 1$.

Fig. 6-4 shows the performance of convolutional coded 8-PPM with constraint length $\nu = 4$ at $R_b / W = 1$. The parity check coefficients for convolutional coded 8-PPM are $h^0 = 23$, $h^1 = 16$, and $h^2 = 12$. When we use superstate MLSD, the number of states in the super-trellis is 1024. The best performance is achieved by PDFD; it is about 1 dB from an optimum superstate ML sequence detector when bit error rate is equal to 10^{-3} . With pre-coding at the transmitter, we use PDFD at the receiver to remove the residual ISI, and to decode the coded signal. Its performance is almost the same as the PDFD and outperforms

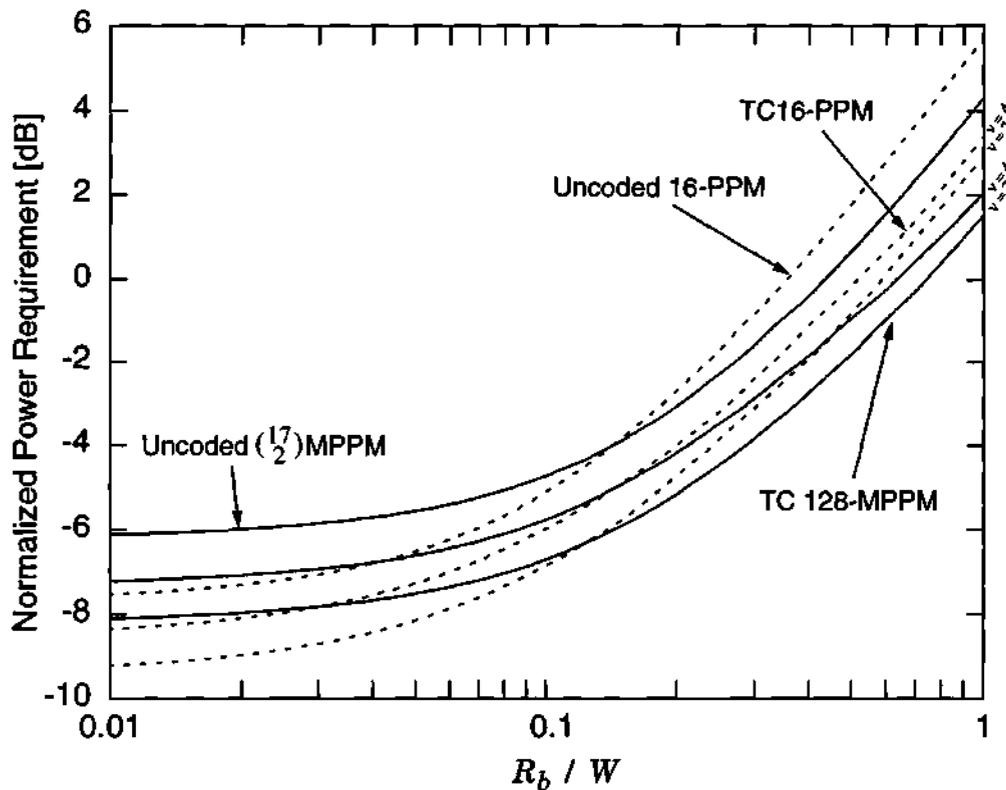


Fig. 6-3. Power penalties of trellis coded 128-MPPM vs. trellis coded 16-PPM

the DFE with VA and LE with VA. We also show the performance of trellis coded 128-PPM with constraint length $v = 7$ at $R_b / W = 1$ in Fig. 6-5. We can see that precoding with PDFD is also effective to remove the ISI, and to decode the coded signal for trellis coded MPPM.

Fig. 6-6 shows the performance of equalization schemes in terms of receiver complexity and performance. The x -axis represents the number of floating-point operations (flops) per bit and the y -axis represents the required SNR to achieve 10^{-3} bit-error-rate. Superstate MLSD achieves the best performance with largest complexity. PDFD and partial-response (PR) precoding with PDFD achieve the almost same performance, but partial-

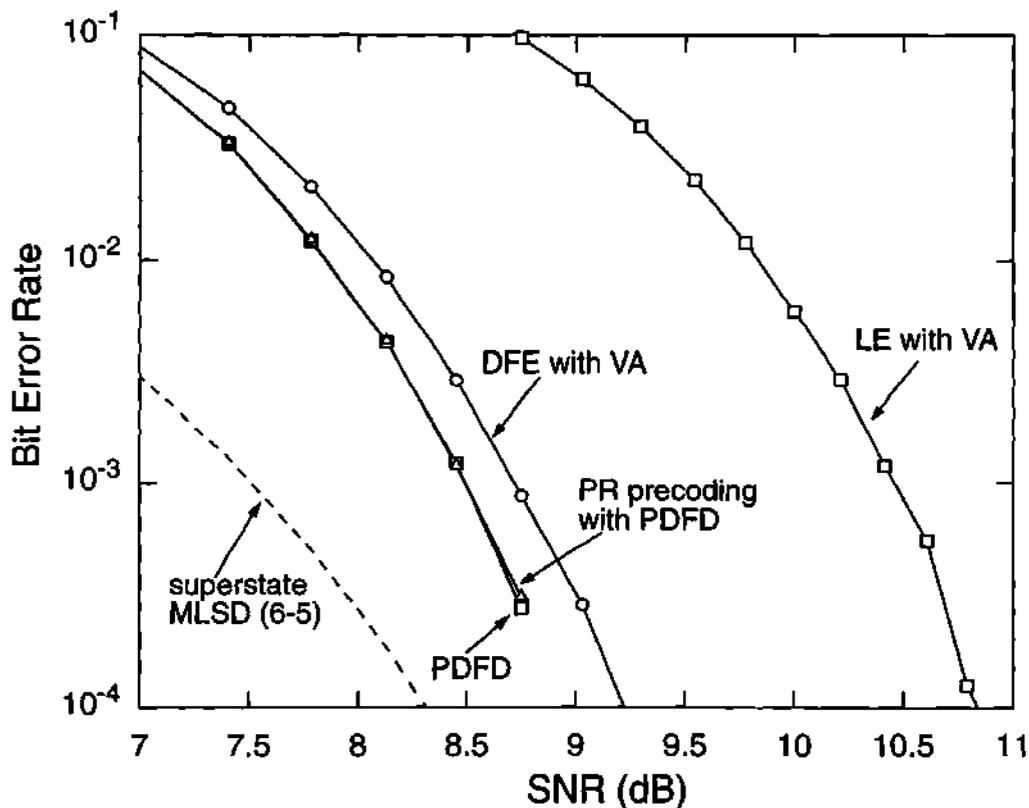


Fig. 6-4. Performance of equalizer for the convolutional coded 8-PPM at $R_b / W = 1$.

tial-response precoding with PDFD requires less complexity than that of PDFD. LE with VA requires the least complexity but its performance is poor. Especially in our applications, precoding at the transmitter and PDFD at the receiver is a good choice to perform equalization and decoding simultaneously.

6.7 SUMMARY AND CONCLUSIONS

Superstate MLSD is an optimum way to jointly perform equalization and decoding for coded signals in the presence of ISI. But its complexity is large even for short ISI channels. First we compared performance of trellis-coded 128-MPPM and trellis coded 16-

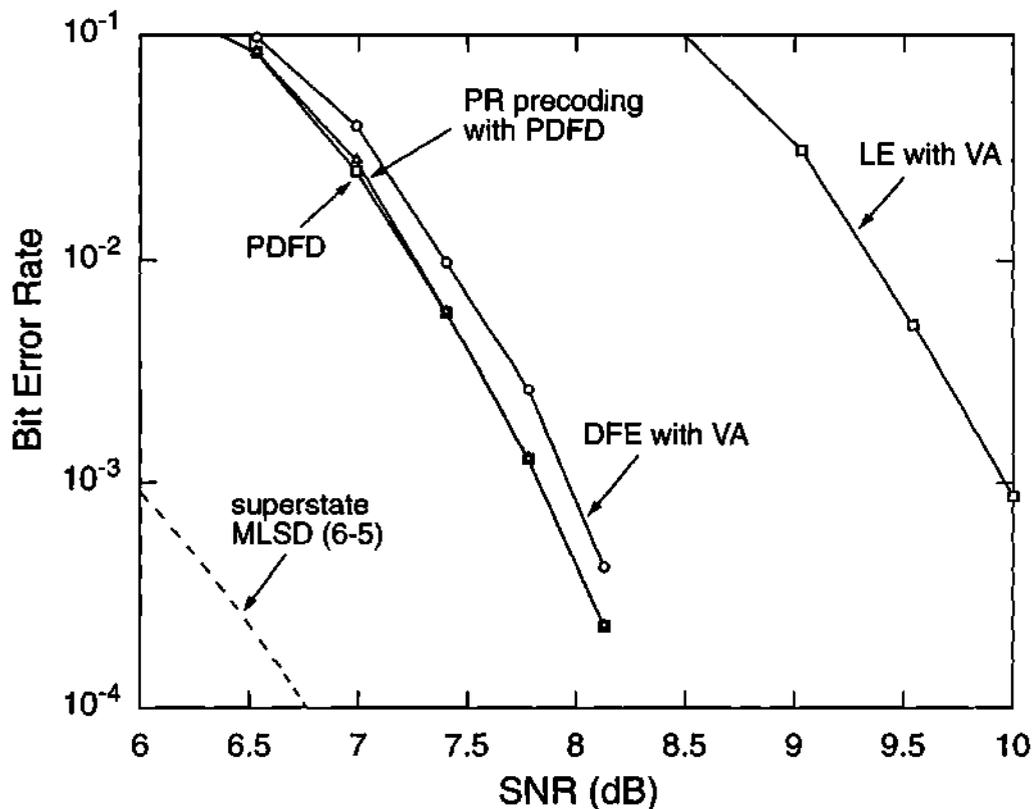


Fig. 6-5. Performance of equalizer for the trellis coded 128-MPPM at $R_b / W = 1$.

PPM using the superstate MLSD. The trellis coded 128-MPPM outperforms trellis coded 16-PPM when bit-rate-to-bandwidth ratio is large. We also compared the performance of various equalization schemes for the coded signal. Precoding at the transmitter resulted in a truncated ISI channel of memory one, which reduced the complexity of receiver significantly. Numerical result showed that partial-response precoding with PDFD is a good choice for our applications in terms of both error probability and complexity.

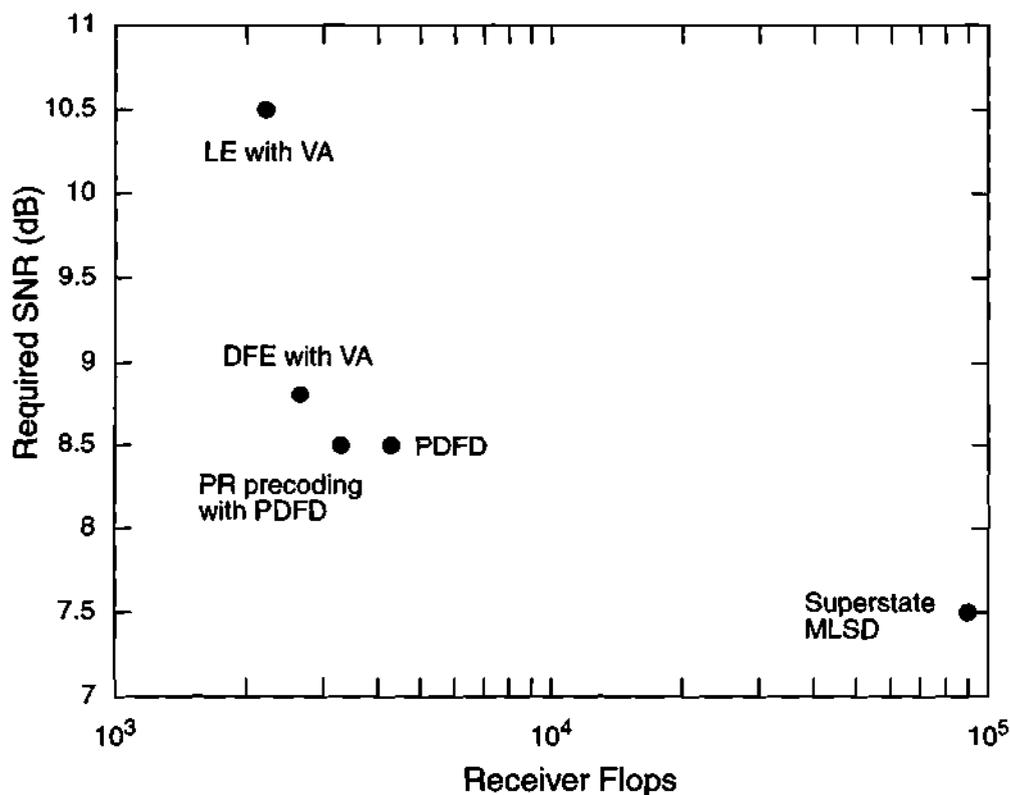


Fig. 6-6. Complexity vs. performance for each equalization scheme using convolutional coded 8-PPM at $R_b / W = 1$; x-axis represents the number of floating-point operations (flops) per bit and y-axis represents the required SNR to achieve 10^{-3} BER.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 CONCLUSIONS

This thesis has concentrated on the design of coded modulation and equalization for high-speed wireless infrared communications with severe ISI. We proposed new trellis codes based on MPPM and new equalization schemes.

Due to its unique channel characteristics, we had to reevaluate the relative performance of various modulation schemes. In chapter 2, we made comparisons of power efficiency and bandwidth efficiency for each candidate modulation scheme. It turned out that MPPM and PPM are power-efficient modulation, and OOK and OPPM are bandwidth-efficient. Calculating the power spectrum density confirmed our first-null approximation

for the bandwidth calculation. The MPPM bound provided the performance limit for bandwidth and power efficiency of uncoded MPPM and its variants. It was shown that certain MPPM can achieve the same bandwidth as OOK with 1.5 dB less power, or can achieve the arbitrarily low power requirement with infinite bandwidth.

The performance comparisons of MPPM and its variants on ISI channels using several equalization schemes were addressed in chapter 3. We proposed a new precoding scheme compatible with the infrared channel. When we used DFE at the receiver to remove the residual ISI, the proposed precoding scheme showed good performances in terms of both complexity and error probability.

We calculated the achievable information rate and cutoff rate of MPPM and its variants as theoretical performance limits in chapter 4.

When we combine coding with modulation, we can improve the overall power efficiency. We combined existing convolutional codes with PPM and developed a new trellis coded MPPM in chapter 5. We performed a computer search for the parity check polynomials of trellis coded MPPM that gives a large minimum distance. We calculated an approximation for the minimum distance and compared our results with the well-known simplex bound. We calculated the power efficiency and bandwidth efficiency for coded modulation schemes. Since poor bandwidth-efficient schemes are sensitive to multipath distortion, we chose both power-efficient and bandwidth-efficient coded schemes such as trellis coded 128-MPPM. We compared the performance of trellis coded 128-MPPM and trellis coded 16-PPM using the superstate MLSD. Trellis coded 128-MPPM outperforms trellis coded 16-PPM on a severe ISI channel. We showed that trellis coded 128-MPPM is

the most power-efficient scheme developed so far in indoor infrared channel, especially in the face of severe multipath dispersion.

Although superstate MLSD is an optimum way to jointly perform equalization and decoding for coded signals in the presence of ISI, its complexity is large even for short ISI channels. We proposed suboptimal methods of equalization and decoding to reduce the complexity in chapter 6. Conventional LE and DFE did not work well with the coded signal. We proposed to use partial-response precoding at the transmitter and to use PDFD at the receiver to equalize and to decode the coded signals, thus reducing the complexity of the receiver significantly. Numerical result showed that the proposed partial-response precoding scheme combined with PDFD is an effective solution for our application.

7.2 FUTURE WORK

7.2.1 Channel Capacity of Wireless Infrared Channel

An interesting problem is the calculation of the channel capacity of (1-1) under the constraint of (1-2). The solution of this problem may lead to useful insight into the signal set design problem. Consider an ideal AWGN channel with average amplitude input constraint:

$$y_k = x_k + n_k, \quad (7-1)$$

under the constraints $x_k \geq 0$ and $E\{x_k\} \leq P$, and n_k is white Gaussian noise with variance N_0 , and is assumed to be independent of x_k . The average mutual information between the channel input and output is [53]:

$$I(x_k; y_k) = H(y_k) - H(y_k | x_k) = H(y_k) - H(n_k), \quad (7-2)$$

where the conditional entropy $H(y_k | x_k)$ is equal to the noise entropy and independent of input distribution. Thus, the problem of finding the capacity for an additive noise channel reduces to maximize the output entropy $H(y_k)$ subject to the input constraint:

$$C = \max_{p_X(x_k)} I(x_k; y_k) = \max_{p_X(x_k)} H(y_k) - H(n_k). \quad (7-3)$$

Since $y_k = x_k + n_k$, $E\{x_k\} \leq P$ implies $E\{y_k\} \leq P$. The general strategy is to find $p_Y(y_k)$ with $\int_{-\infty}^{\infty} y_k p_Y(y_k) dy_k \leq P$ to maximize $H(y_k)$, and then try to find an input distribution $p_X(x_k)$ that yields $p_Y(y_k)$. Finding a channel capacity reduces the problem of maximizing entropy, $H(y_k) = \int_{-\infty}^{\infty} p_Y(y_k) \log[p_Y(y_k)] dy_k$, over all $p_Y(y_k)$ satisfying:

$$p_Y(y_k) \geq 0, \int_{-\infty}^{\infty} p_Y(y_k) dy_k = 1, \text{ and } \int_{-\infty}^{\infty} y_k p_Y(y_k) dy_k \leq P. \quad (7-4)$$

The maximum entropy principle [36] implies that the maximum entropy is infinite, and that there is no maximum entropy distribution. Therefore, the channel capacity under the average amplitude constraint is *infinite*. We can show that the capacity of the multipath channel is also infinite using the arguments similar to those in section 4.3. This result is not surprising since the capacity of the Poisson channel under the same constraints is also infinite [84].

There are several interesting questions arising from this result: (1) How can we achieve the infinite channel capacity? (2) What is the optimum signal set for this channel? (3) If it is impossible to achieve it, what are more practical constraints for the input? What is channel capacity under the modified constraints? These questions still remain open.

7.2.2 Concatenated Codes

In section 5.2, we combined the convolutional code and PPM as shown Fig. 7-1-a. Information bits with bit rate R_b enter the $(\log_2 L - 1) / \log_2 L$ convolutional encoder. The PPM encoder maps each block of $\log_2 L$ bits into one L -PPM codeword. The advantage of this type of scheme is the use of low duty-cycle modulation, such as PPM, at the output of the transmitter. But it has a poor bandwidth efficiency:

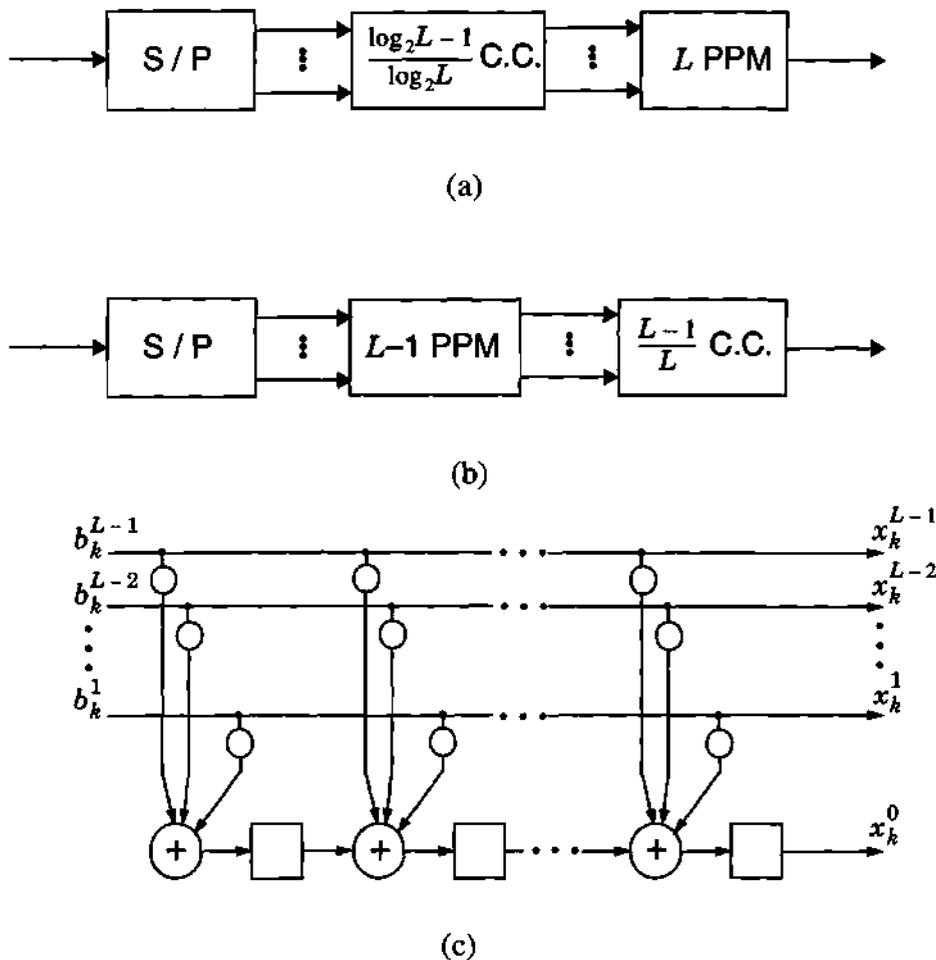


Fig. 7-1. Serial concatenated codes: (a) convolutional coded PPM, (b) PPM with convolutional code, (c) a systematic feedforward convolutional encoder for configuration (b).

$$\beta_1 = \frac{\text{Bandwidth}}{R_b} = \frac{L}{\log_2 L - 1} \quad (7-5)$$

One possible modification of this scheme is exchange the order of convolutional encoder and PPM mapper as shown in Fig. 7-1-b. Information bits with rate R_b enter the $(L - 1)$ PPM encoder, which maps each block of $\log_2(L - 1)$ bits into one of $L - 1$ PPM codeword. The $(L - 1) / L$ convolutional encoder transforms the $L - 1$ input bits into L coded bits. The bandwidth efficiency of this configuration is:

$$\beta_2 = \frac{\text{Bandwidth}}{R_b} = \frac{L}{\log_2(L - 1)} \quad (7-6)$$

We can see that this configuration is more bandwidth efficient since the ratio of (7-5) to (7-6) is:

$$\frac{\beta_2}{\beta_1} = \frac{\log_2 L - 1}{\log_2(L - 1)} < 1 \quad (7-7)$$

Table 7-1 lists the ratio β_1 / β_2 .

TABLE 7-1: The bandwidth-efficiency ratio between two serial concatenated schemes.

L	β_1 / β_2
4	0.63
8	0.71
16	0.77
32	0.81

To maintain the low duty-cycle property of PPM, we can use a systematic feedforward convolutional encoder as shown in Fig. 7-1-c. The weight of the output of the convolutional encoder is only one or two depending on the value of x_k^0 .

However, to handle a large number of alphabets, the $(L - 1) / L$ convolutional encoder should be high-rate. For example, the rates of convolutional encoder are $4/5$, $8/9$, and $16/17$ for 4, 8, 16-PPM, respectively, and the numbers of outgoing branches from each state are 16, 256, 65536, respectively. Therefore, the use of high-rate convolutional codes with Viterbi decoding becomes problematic. To avoid this complexity problem, we can use a punctured convolutional code [85].

7.2.3 Turbo Codes

In 1993, Berrou *et al.* [86] reported a new class of code called turbo codes, whose performance in terms of BER was a few tenths of a dB from the Shannon limit. Since then, many researchers have focused on the design and development of turbo codes. A turbo encoder is based on a parallel concatenation of two recursive systematic convolutional code and an interleaver. While TCM is a bandwidth-efficient scheme, turbo coding is a power-efficient scheme. Surely, turbo codes are good candidates for wireless infrared channel, where power efficiency is the most important parameter. However, the performance of turbo codes over a multipath channel is unknown.

7.2.4 Synchronization

In wireless infrared communications, highly peaked and narrow optical pulses are used to achieve high power efficiency, but accurate timing synchronization is also required for ideal detection. In this thesis, we have assumed perfect symbol synchronization. But

there exists a timing offset in any real system, and this offset influences the performance of the receiver [87]. Consider first the OOK system where a timing error Δ ($0 \leq \Delta \leq T/2$) exists, such that sampling occur at $T + \Delta$. In Fig. 7-2, we show the four possible cases for current symbol during interval $[0, T)$ and the next symbol during $[T, 2T)$. When a received bits arrive, the matched filter convolves with only portions of the current pulse and also with some portions of the next pulse, and sampling occurs at $kT + \Delta$. Following the analysis of [87], we can calculate the bit error rate for OOK considering all four cases in Fig. 7-2:

$$Pr[\text{bit error}] = \frac{1}{2}Q\left(\frac{P}{\sqrt{N_0 R_b}}\right) + \frac{1}{2}Q\left(\frac{P}{\sqrt{N_0 R_b}}\left(1 - \frac{2\Delta}{T}\right)\right). \quad (7-8)$$

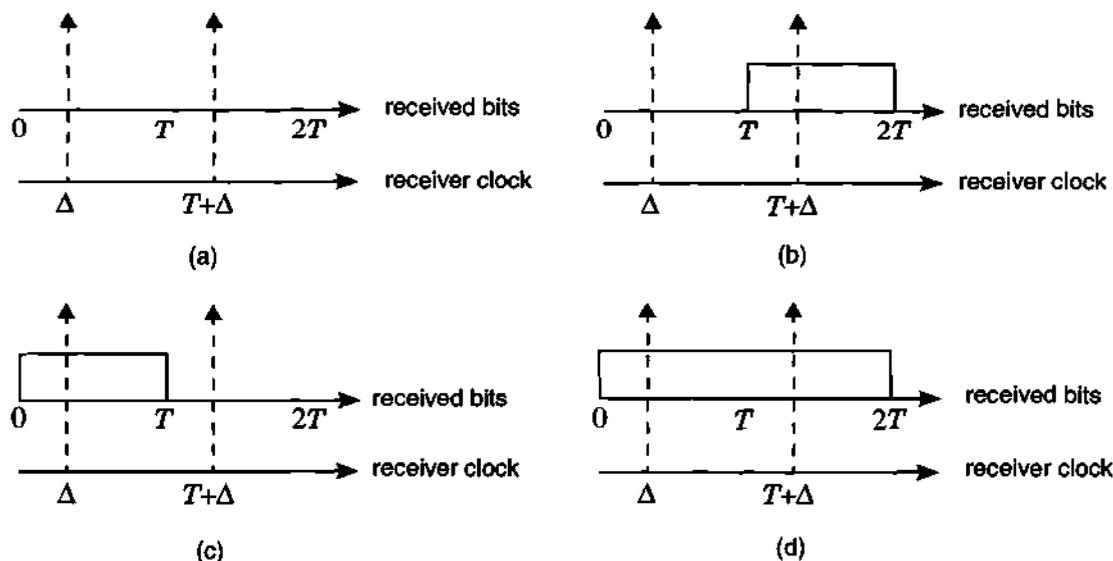


Fig. 7-2. Effects of timing error for OOK when (a) current bit = 0 and next bit = 0, (b) current bit = 0 and next bit = 1, (c) current bit = 1 and next bit = 0, (d) current bit = 1 and next bit = 1.

When $\Delta \rightarrow 0$, the above expression reduces to (2-4). When $\Delta \rightarrow T/2$, the above expression approaches to 1/2, and the system does not work at all.

Similarly, we can calculate the probability of symbol error for PPM when there exists a timing error Δ , using the result of [87] and section 3.3:

$$Pr[\text{error}] \leq Q\left(P \sqrt{\frac{L \log_2 L}{2N_0 R_b}} \left(1 - \frac{2\Delta}{T_c}\right)\right) + (L-2)Q\left(P \sqrt{\frac{L \log_2 L}{2N_0 R_b}} \left(1 - \frac{\Delta}{T_c}\right)\right), \quad (7-9)$$

where $T_c = T/L$ is the chip interval. We can derive a similar expression for the symbol error probabilities of MPPM. However, the loss of orthogonality in MPPM may make this calculation more difficult than PPM.

Georghiades and his students in Texas A & M university [88][89] have studied the frame and symbol synchronization for MPPM and its variants on a Poisson channel. Their results should be very helpful in a similar study for our channel.

7.2.5 System Implementation

As we indicated in section 1.1, most real systems have used OOK or PPM. But we showed that theoretically, MPPM is a very good choice for both uncoded and coded modulation schemes on a severe multipath channel. Both for academic and practical purposes, complete system design and implementation of high-speed wireless infrared communication using MPPM is very promising.

REFERENCES

- [1] A. R. Tharek, "Indoor Communication and Other Applications at Millimeter-wave Band," *IEEE Tencon' 90*, pp. 241-245, 1990.
- [2] M. J. McCullagh and D. R. Wisely, "155 Mb/s Optical Wireless Link Using a Bootstrapped Silicon APD Receiver," *Electronic Letter*, vol. 30, no. 5, pp. 430-432, 1994.
- [3] Infrared Data Association (<http://www.irda.org>).
- [4] J. M. Kahn and J. R. Barry, "Wireless Infrared Communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265-298, February 1997.
- [5] Spectrix Corporation (<http://www.spectrixcorp.com>).
- [6] G.W. Marsh and J. M. Kahn, "50-Mb/s Diffuse Infrared Free-Space Link Using On-Off Keying with Decision Feedback Equalization," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1086-1089, 1994.
- [7] F. R. Gfeller and U. Bapst, "Wireless In-House Data Communications via Diffuse Infrared Radiation," *Proceeding of IEEE*, vol. 67, no. 11, pp. 1474-1486, November 1979.
- [8] M. D. Kotzin, "Short-Range Communications Using Diffusely Scattered Infrared Radiation," Ph.D. Dissertation, Northwestern University, Evanston, IL, June 1981.
- [9] J. R. Barry, *Wireless Infrared Communications*, Kluwer Academic Publishers, 1994.
- [10] J. R. Barry, J. M. Kahn, W. J. Krause, E. A. Lee, and D. G. Messerschmitt, "Simulation of Multipath Impulse Response for Wireless Optical Channels," *IEEE*

- Journal of Selected Areas in Communication*, vol. 11, no. 3, pp. 367-379, April 1993.
- [11] J. M. Kahn, J. R. Barry, M. D. Audeh, J. B. Carruthers, W. Krause, and G. W. Marsh, "Non-Directed Infrared Links for High-Capacity Wireless LANs," *IEEE Personal Communications Magazine*, vol. 1, no. 2, pp. 12-25, Second Quarter 1994.
- [12] J. M. Kahn, W. J. Krause, and J. B. Carruthers, "Experimental Characterization of Non-Directed Indoor Infrared Links," *IEEE Transactions on Communications*, vol.43, no. 2-4, pp. 1613-1623, February-April 1995.
- [13] J. B. Carruthers and J. M. Kahn, "Modeling of Non-Directed Wireless Infrared Channels," *IEEE International Conference on Communications*, Dallas, Texas, pp. 1227-1231, June 1996.
- [14] D. C. Lee, M. D. Audeh, and J. M. Kahn, "Trellis-coded Pulse-Position Modulation for Indoor Wireless Infrared Communications," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 349-353, Taipei, Taiwan, October 1996.
- [15] M. D. Audeh, J. M. Kahn and J. R. Barry, "Performance of Pulse-Position Modulation on Measured Non-Directed Indoor Infrared Channels," *IEEE Transaction on Communications*, vol. 44, pp. 654-659, June 1996.
- [16] M. D. Audeh, J.M. Kahn and J.R. Barry, "Decision-Feedback Equalization of Pulse-Position Modulation on Measured Non-Directed Indoor Infrared Channels", submitted. to *IEEE Transaction on Communications*, June 1996.
- [17] M. D. Audeh, "Power-Efficient Modulation for High-Speed Non-Directed Infrared Communication," Ph.D. Dissertation, Univ. of California, Berkeley, 1995.
- [18] H. Hashemi, G. Yung, M. Kavehrad, R. Behbahani and P. A. Galko, "Indoor Propagation Measurements at Infrared Frequencies for Wireless Local Area Networks Applications," *IEEE Transaction on Vehicular Technology*, vol. 43, pp. 562-576, August 1994.
- [19] R. T. Valadas, "Interference Modeling for The simulation of IEEE 802.11 Infrared Local Area Networks," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 257-261, Taipei, October 1996.
- [20] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1984.

- [21] J. R. Barry, "Sequence Detection and Equalization for Pulse-Position Modulation," *IEEE International Conference on Communications*, New Orleans, LA, pp. 1561-1565, May 1994.
- [22] J. G. Proakis, *Digital Communications*, Third Edition, McGraw-Hill, 1995.
- [23] R. D. Gitlin, J. F. Hayes, and S. B. Weinstein, *Data Communication Principles*, Plenum Press, 1992.
- [24] J. R. Pierce, "Optical Channel: Practical limits with Photon Counting," *IEEE Transactions on Communications*, vol. 26, no. 12, pp. 1819-1821, December 1978.
- [25] H. Sugiyama and K. Nosu, "MPPM: A method for improving the Band-Utilization Efficiency in Optical PPM," *Journal of Lightwave Technology*, vol 7, no. 3 pp. 465-472, March 1989.
- [26] G. E. Atkin, K. S. Fung, "Coded Multipulse Modulation in Optical Communication System," *IEEE Transactions on Communications*, vol. 42, no. 2, pp. 574-582, February 1994.
- [27] A. E. Brouwer, J. B. Shearer, N. J. A. Sloane, and W. D. Smith, "A New Table of Constant Weight Codes," *IEEE Transactions on Information Theory*, vol. 36, no. 6, pp. 1334-1380, November 1990.
- [28] H. Park and J. R. Barry, "Modulation Analysis for Wireless Infrared Communication," *IEEE International Conference on Communications*, Seattle, WA, pp. 1182-1186, June 1995.
- [29] D. Slepian, "Permutation Modulation," *Proceedings of IEEE*, vol. 53, pp. 228-236, March 1965.
- [30] H. Park and J. R. Barry, "The Performance of Multiple Pulse-Position Modulation on Multipath Channels," *IEE Proceedings, Optoelectronics*, vol. 143, no. 6, pp. 360-364, December 1996.
- [31] D. Slepian, "Group Codes for the Gaussian Channel," *Bell Syst. Tech. J.*, vol. 47, pp. 575-602, April 1968.
- [32] I. Bar-David and G. Kaplan, "Information Rates of Photon-limited Overlapping Pulse Position Modulation Channel," *IEEE Transactions on Information Theory*, vol. 30, no. 3, pp. 455-463, May 1984.

- [33] F. Amoroso, "The Bandwidth of Digital Data Signals," *IEEE Communications Magazine*, vol. 18, pp. 13-24, November 1980.
- [34] S. G. Wilson, *Digital Modulation and Coding*, Prentice-Hall, Inc., 1996.
- [35] W. A. Gardner, *Statistical Spectral Analysis: A Non-probabilistic Theory*, Prentice-Hall, 1987.
- [36] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, 1991.
- [37] G. D. Forney, Jr., "Maximum-likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," *IEEE Transactions on Information Theory*, vol. 18, no. 3, pp. 363-377, May 1972.
- [38] G. D. Forney, Jr., "The Viterbi Algorithm," *Proceedings of IEEE*, vol. 61, no. 3, pp. 268-278, March 1973.
- [39] M. V. Eyuboglu and S. U. H. Qureshi, "Reduced-State Sequence Estimation with Set Partitioning and Decision Feedback," *IEEE Transactions on Communication*, vol. 36, no. 1, pp. 13-20, January 1988.
- [40] M. V. Eyuboglu and S. U. H. Qureshi, "Reduced-State Sequence Estimation for Coded Modulation on Intersymbol Interference Channels," *IEEE Journal of Selected Areas in Communication*, vol. 7, no. 6, pp. 989-995, August 1989.
- [41] A. Duel-Hallen and C Heegard, "Delayed Decision-Feedback Sequence Estimation," *IEEE Transactions on Communication*, vol. 37, no.5, pp. 428-436, May 1989.
- [42] R. Price, "Nonlinearly Feedback Equalized PAM versus Capacity for Noisy Channel Filters," *IEEE International Conference on Communications*, pp. 22.12-22.17, 1972.
- [43] K. Abend and B. D. Fritchman, "Statistical Detection for Communication Channels with Intersymbol Interference," *Proceeding of IEEE*, vol. 58, no. 5, pp. 799-785, May 1970.
- [44] D. Williamson, R. A. Kennedy, and G. W. Pulford, "Block Decision Feedback equalization," *IEEE Transactions on Communication*, vol.40, no. 2, pp. 255-264, February 1992.

- [45] M. Tomlinson, "New Automatic Equalizer Employing Modulo Arithmetic," *Electronic Letters*, vol. 7, no. 5/6, pp. 138-139, March 1971.
- [46] H. Harashima and H. Miyakawa, "Matched Transmission Technique for Channels with Intersymbol Interference," *IEEE Transactions on Communication*, vol. 20, no. 4, pp. 774-780, August 1972.
- [47] R. Laroia, S. A. Tretter, and N. Farvardin, "A Simple and Effective Precoding Scheme for Noise Whitening on Intersymbol Interference Channels," *IEEE Transactions on Communication*, vol. 41, no. 10, pp. 1460-1462, October 1993.
- [48] R. Laroia, "Coding for Intersymbol Interference Channels - Combined Coding and Precoding," *IEEE Transactions on Information Theory*, vol. 42, no. 4, pp. 1053-1061, July 1996.
- [49] C. N. Georghiades, "Modulation and Coding for Throughput-Efficient Optical Systems," *IEEE Transactions on Information Theory*, vol. 40, no.5, pp. 1313-1326, September 1994.
- [50] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1987.
- [51] E. A. Lee, D. G. Messerschmitt, *Digital Communication*, Second Edition, Kluwer Academic Publishers, 1994.
- [52] R. F. H. Fischer and J. B. Huber, "Comparison of Precoding Schemes for Digital Subscriber Lines," *IEEE Transactions on Communication*, vol. 45, no. 3, pp. 334-343, March 1997.
- [53] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, 1968.
- [54] J. G. Smith, "The Information Capacity of Amplitude and Variance-Constrained Scalar Gaussian Channels," *Information and Control*, vol. 18, pp. 203-219, 1971.
- [55] B. S. Tsybakov, "Capacity of a Discrete Time Gaussian Channel with a Filter," *Problemy Peredachi Informatsii*, vol. 6, no. 3, pp. 78-82, July-September 1970.
- [56] W. Hirt, "Capacity and Information Rates of Discrete-Time Channels with Memory," Ph.D. Dissertation, Swiss Federal Institute of Technology, Zürich, 1988.
- [57] S. Shamai (Shitz), L. H. Ozarow, and A. D. Wyner, "Information Rates for a Discrete-Time Gaussian Channel with Intersymbol Interference and Stationary

- Inputs," *IEEE Transactions on Information Theory*, vol. 37, no. 6, pp. 1527-1539, November 1991.
- [58] R. E. Blahut, *Principles and Practice of Information Theory*, Reading, MA, Addison-Wesley, 1987.
- [59] J. L. Massey, "Coding and Modulation in Digital Communication," *Proc. Int. Zürich Seminar on Digital Communications*, Switzerland, E2 1-4, March 1974.
- [60] E. Biglieri, "The Computational Cutoff Rate of Channels Having Memory," *IEEE Transactions on Information Theory*, Vol. 27, No. 3, pp. 352-357, May 1981.
- [61] R. A. Horn, and C. A. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
- [62] A. J. Viterbi, and J. K. Omura, *Principles of Digital Communication and Coding*, McGraw-Hill Book Company, New York, 1979.
- [63] H. Park and J. R. Barry, "Performance Analysis and Channel Capacity of Multiple Pulse-Position Modulation on Multipath Channels," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 247-251, Taipei, October 1996.
- [64] G. Ungerboeck, "Channel Coding with Multilevel / Phase Signals," *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 55-67, January 1982.
- [65] A. R. Calderbank and N. J. A. Sloane, "New Trellis Codes Based on Lattice and Cosets," *IEEE Transactions on Information Theory*, vol. 32, no. 2, pp. 177-195, March 1987.
- [66] E. Zehavi and J. K. Wolf, "On the Performance Evaluation of Trellis Codes," *IEEE Transactions on Information Theory*, vol. 33, no. 2, pp. 196-202, March 1987.
- [67] S. Benedetto, M. Mondin, and G. Montorsi, "Performance Evaluation of Trellis Coded Modulation Schemes," *Proceedings of IEEE*, vol. 82, no. 6, pp. 833-855, June 1994.
- [68] F. Wang and D. J. Costello, "Probabilistic Construction of Large Constraint Length Trellis Codes for Sequential Decoding," *IEEE Transactions on Communication*, vol. 43, no. 9, pp. 2439-2447, September 1995.

- [69] E. Biglieri, D. Divsalar, P. McLane, and M. Simon, *Introduction to Trellis-coded Modulation with Applications*, Mcmillan, 1991.
- [70] E. Forestieri, R. Gangopadhyay, and G. Prati, "Performance of Convolutional Codes in a Direct-Detection Optical PPM Channel," *IEEE Transactions on Communications*, vol. 37, no. 12, pp. 1303-13, December 1989.
- [71] C. N. Georghiades, "Some Implications of TCM for Optical Direct-Detection Channels," *IEEE Transactions on Communications*, vol. 37, no.5, pp. 481-487, May 1989.
- [72] G. J. Pottie, "Trellis Codes for the Optical Direct-Detection Channel," *IEEE Transactions on Communications*, vol. 39, no. 8, pp. 1182-1183, August 1991.
- [73] J. A. Heller, "Sequential Decoding: Sort Constraint Length Convolutional Code," *JPL Space Program Summary, 37-54*, vol 3, pp. 171-177, December 1967.
- [74] A. R. Calderbank, J. E. Mazo, and V. K. Wei, "Asymptotic Upper Bounds on the Minimum Distance of Trellis Codes," *IEEE Transactions on Communication*, Vol. 33, No. 4, pp. 305-309, April 1985.
- [75] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*, Prentice-Hall, 1983.
- [76] G. L. Bechtel and J. W. Modestino, "Pulsewidth-constrained Signaling and Trellis-coded Modulation on the Direct-detection Optical Channel," *IEEE Global Telecommunications Conference*, vol. 2, pp. 842-847, 1988.
- [77] P. R. Chevillat and E. Eleftheriou, "Decoding of Trellis-Encoded Signals in the Presence of Intersymbol Interference and Noise," *IEEE Transactions on Communication*, vol. 37, no. 7, pp. 669-676, July 1989.
- [78] I. Oka and E. Biglieri, "Error Probability Bounds for Trellis Coded Modulation over Sequence Dependent Channels," *The Transactions of IEICE*, vol. E72, no. 4, pp. 375-382, April 1989.
- [79] J. E. Porath, "On Trellis-coded Modulation for Gaussian and Band-limited Channels," *Tech. Rep. 221*, Chalmers University of Technology, Goteborg, Sweden, 1991.
- [80] K. Wesolowski, "Efficient Digital Receiver Structure for Trellis-coded Signals Transmitted through Channels with Intersymbol Interference," *Electronic Letters*, vol. 23, no. 24, pp. 1265-1267, November 1987.

- [81] H. K. Thapar, "Real-Time Application of Trellis Coding to High-Speed Voiceband Data Transmission," *IEEE Journal of Selected Areas in Communications*, vol. 2, no. 5, pp. 648-658, September 1984.
- [82] L. N. Wong and P. J. McLane, "Performance of Trellis Codes for a Class of Equalized ISI Channels," *IEEE Transactions on Communication*, vol. 36, no. 12, pp. 1330-1336, December 1988.
- [83] G. D. Forney Jr. and M. V. Eyuboglu, "Combined Equalization and Coding Using Precoding," *IEEE Communications Magazine*, pp. 25-34, December 1991.
- [84] M. H. A. Davis, "Capacity and Cutoff Rate for Poisson-type Channels," *IEEE Transactions on Information Theory*, vol. 26, no. 6, pp. 710-715, November 1980.
- [85] J. B. Cain, G. C. Clark, Jr., and J. M. Geist, "Punctured Convolutional Codes of Rate $(n-1)/n$ and Simplified Maximum Likelihood Decoding," *IEEE Transactions on Information Theory*, vol. 25, pp. 97-100, January 1979.
- [86] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon Limit Error-correcting Coding and Decoding: Turbo-codes," *IEEE International Conference on Communications*, Geneva, Switzerland, pp. 1064-1070, May 1993.
- [87] R. M. Gagliardi and S. Karp, *Optical Communications*, Second Edition, John Wiley & Sons, Inc., 1995.
- [88] S. Patarasen and C. N. Georghiades, "Maximum-Likelihood Symbol Synchronization and detection of OPPM Sequences," *IEEE Transactions on Communication*, vol. 42, no. 6, pp. 2282-2290, June 1994.
- [89] R. Velidi and C. N. Georghiades, "Frame Synchronization for Optical Multi-Pulse Pulse-Position Modulation," *IEEE Transactions on Communication*, vol. 43, no. 2-4, pp. 1838-1843, April 1995.

Vita

Hyuncheol Park was born in Pusan, Korea on June 7, 1960 and grew up in Seoul, Korea. He received B.S. and M.S. both in Electronics Engineering from Yonsei University, Seoul, Korea in 1983 and 1985, respectively, and Ph.D. in Electrical Engineering from Georgia Institute of Technology, Atlanta, USA in 1997. From 1985 to 1991, he was a research engineer in Samsung Electronics Co, Suwon, Korea where he developed the broadcast satellite receivers, microwave low-noise amplifiers, and automatic vehicle monitoring systems. From 1992 to 1997, he was a graduate research assistant in Georgia Institute of Technology, Atlanta, USA. His current research interests include coded modulations, wireless infrared communications, and information theory.