THE UNIVERSITY OF SOUTH AUSTRALIA
SCHOOL OF ELECTRONIC ENGINEERING

CHANNEL CAPACITY CALCULATIONS FOR M–ARY N–DIMENSIONAL SIGNAL SETS

Philip Edward McIllree, B.Eng.

A thesis submitted in fulfilment of the requirement for the degree of Master of Engineering in Electronic Engineering.

February 1995

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SUMMARY

This thesis presents a technique for calculating channel capacity of M–ary digital modulation over an additive white Gaussian noise channel. A general channel capacity formula for N–dimensional signal sets has been developed and requires the signal constellation and noise variance for computation.

Channel capacity calculation involves integration in N dimensions. A numerical integration technique based on repeated Gaussian quadrature allows direct computation with an acceptable degree of precision. Accurate capacity calculation can be achieved for $N \leq 4$.

Capacity data is presented for well known signal sets and new results are obtained for a hybrid scheme. The signal sets examined are pulse amplitude modulation (PAM), quadrature amplitude modulation (QAM), phase shift keying (PSK), and M–ary orthogonal signalling. The hybrid scheme is L–orthogonal signalling. Bandwidth efficiency data is also presented and is calculated by normalising capacity with respect to occupied bandwidth.

DECLARATION

I declare that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of knowledge it does not contain any materials previously published or written by another person except where due reference is made in the text.

Philip E. McIllree

June 1994

ACKNOWLEDGMENT

The opportunity to undertake this study was made possible by Professor Michael J. Miller, Professor of Telecommunications, School of Electronic Engineering. I am indebted to Professor Miller for providing the inspiration and challenge which became the centre of study for this thesis.

My supervisor for the last two years, Dr. Steven S. Pietrobon, was an enormous help for guiding me in the technical aspects of the study and for reviewing draft versions of this thesis. I would like to thank Dr. Richard Wyrwas for involving me as a member of the Mobile Communications Research Centre (MCRC). The results of this work were presented in Singapore at the IEEE SICON/ICIE '93 conference and this would not have been possible without sponsorship from the MCRC.

I would like to give a special thanks to Mr. Len Colgan, Head of the Dept. of Mathematics. I consulted with him many times throughout the duration of this work. Mr. Colgan was especially helpful in my endeavours to compute complicated integrals. Thanks are also due to Dr. Ross Frick for his assistance in tackling multivariate integrals. I wish to thank Professor Ian Sloan, Head of School of Mathematics, University of New South Wales, for his reassurances that I was applying the most suitable integration technique. Thanks also goes to Dr. Frank Stenger, Dept. of Mathematics, University of Utah, and Ms. Karen McConaghy of the American Mathematical Society for providing additional data.

I am grateful for the time and patience of Mr. Bill Cooper and Mr. Peter Asenstorfer for helping me with writing computer code and using applications software. Many thanks goes to my colleagues at the Digital Communications Group for their support.

Finally I would like to thank my family, relatives, and friends for their encouragement.

1 INTRODUCTION

A digital communications system is generally characterised by bit error ratio (BER) performance. The maximum rate of information transmission is an additional characteristic obtained by channel capacity analysis [1–3]. Normalising capacity with respect to occupied bandwidth yields another useful parameter, bandwidth efficiency. This thesis presents the development of a systematic technique for calculating the channel capacity and bandwidth efficiency of M–ary digital modulation signal sets.

The capacity analysis of a particular modulation scheme can help assess the trade–off between the potential coding gain achieved by signal set expansion and required receiver complexity. This analysis is a first step in the design of trellis codes using M–ary signal sets [4, 5].

The communication system model for analysis of uncoded modulation is shown in Figure 1.1 . The information source outputs a k–bit message block, d_m. The modulator performs a one to one mapping of d_m to a transmitted waveform, $x_m(t)$. There are a total of $M = 2^k$ possible transmitted waveforms. The received waveform, r(t), is the transmitted waveform disturbed by additive noise, n(t). The demodulator processes the received waveform to provide an estimate of the transmitted message, $\hat{d}_m$, to the destination. A *waveform* will also be referred to as a *symbol* or *signal* throughout this study to suit a particular discussion or common phraseology.

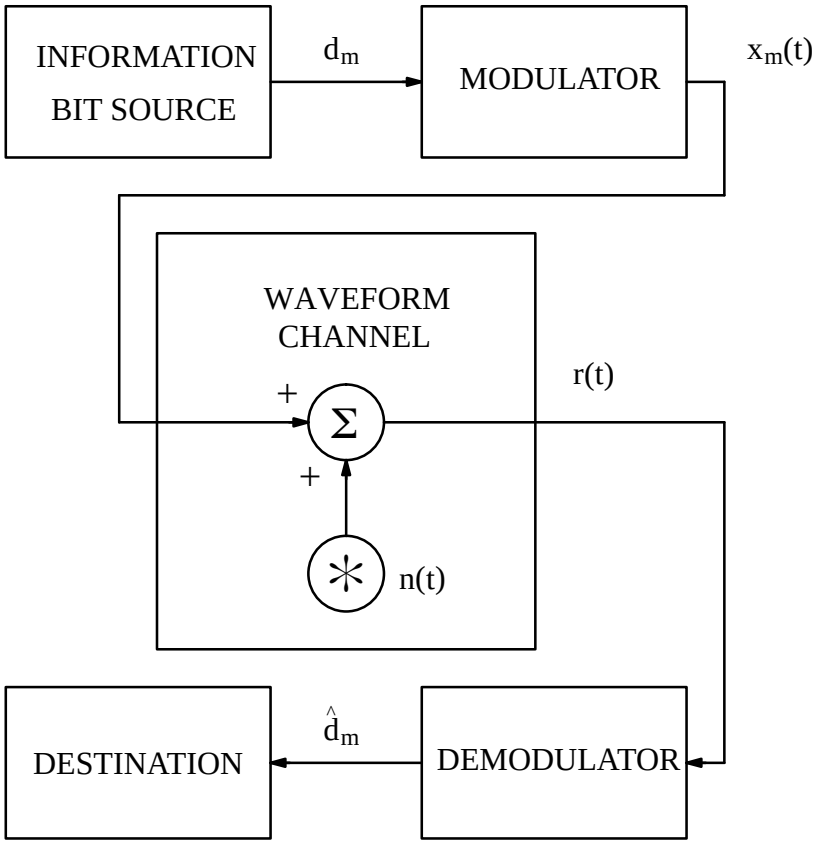


Figure 1.1 Uncoded M–ary waveform communications

In this study capacity analysis is restricted to sets of transmitted waveforms which are finite–energy bandlimited waveforms. Waveforms of this type can be described as vectors in N–dimensional space. The analysis of M–ary signal sets is made systematic by requiring only the signal constellation.

The additional restrictions for capacity analysis are as follows:

1. The noise process is additive white Gaussian noise (AWGN).
2. All channel models are memoryless.
3. Received waveforms are coherently detected.
4. All transmitted waveforms are of equal symbol period, T.

This work came about while studying coding and modulation techniques for extremely noisy channels. Some modulation schemes sacrifice bandwidth efficiency to obtain reliable transmission [6]. Such bandwidth inefficient schemes have large dimensionality, $N > 2$. An investigation of the capacity at low signal to noise power ratio (SNR) was required but no literature was immediately available to perform analysis for large N.

The geometric interpretation of finite energy, bandlimited signals and the capacity bound for the bandlimited Gaussian channel were first presented by Shannon in the late 1940's [1–3]. The study of capacity for single channel models with M–ary waveform inputs has appeared in limited publications. The capacity for $N \leq 2$ signal sets was first published by [4] and [7 pp. 272–279]. The capacity for $N \geq 2$ signal sets has been treated recently by [5] and [8]. These results however are not directly applicable to this study. The capacity analysis in [5] and [8] was performed for constant envelope signals and non–coherent detection of orthogonal signals respectively. The aim of this study is to combine the analytical procedure of these publications to present a concise development of capacity for the N–dimensional channel model.

Channel capacity calculation requires integration in N dimensions. Monte Carlo simulation is used in [4] and [5] and the method is not stated in [7]. This study presents an alternative numerical integration technique which performs direct computation using compact computer code.

The integration in the capacity formula is performed using a standard numerical quadrature method [9, pp. 23–28]. The region of integration and weight function are identified as belonging to the Gauss–Hermite quadrature. Repeated application of this quadrature rule is the basis for multiple integration. The technique is applied for $N \leq 4$ with an acceptable degree of precision. Integration for N=8 is also attempted.

A general capacity formula for M–ary signal sets described by N–dimensional vectors is developed. Several well known modulation techniques are studied and families of curves for capacity and bandwidth efficiency are presented. New results for a hybrid modulation scheme is presented. Results for coding gain by signal set expansion are examined.

All calculations are implemented with double precision ANSI C programs on a modern workstation. A search for other quadrature rules has been carried out and a large reduction in the number of integration points is possible for the same degree of precision.

Chapter 2 presents the development of the general capacity formula for M–ary signal sets. Numerical results are presented in Chapter 3. Chapter 4 presents the technique for multiple integration. The conclusions for this work are presented in Chapter 5.

2 CAPACITY FOR BANDLIMITED INPUT ADDITIVE WHITE GAUSSIAN NOISE CHANNELS

2.1 Introduction

The channel capacity for certain bandlimited input AWGN channels are presented in this Chapter. The capacity formulae are derived for the N–dimensional signal space to enable analysis of M–ary digital signal sets in Chapter 3.

The channel models in this Chapter are discrete–time memoryless channels [10, pp. 71–72], [11, pp. 121–122]. The channel input and output are described as discrete–time sequences of letters belonging to finite or infinite alphabets. The memoryless condition holds for when an "output letter at a given time depends statistically only on the corresponding input letter" [10, p. 71]. The discrete memoryless channel (DMC) is presented in Section 2.2. Here, the input and output alphabets are finite and the formula for channel capacity is given. This capacity formula is then modified for the Gaussian channels.

To determine the capacity for M–ary waveform communications we remodel the waveform channel from Figure 1.1 as a discrete–time memoryless channel. The input and output waveforms are continuous functions of time. By restricting our analysis to finite–energy bandlimited waveforms we can use vector representation for the channel input and output. The discrete–time condition is met by using waveforms of equal symbol period, T. The memoryless condition holds for when an output waveform over each T second interval depends statistically only on the corresponding input waveform.

Variations of the DMC are extended to vector channel models in Section 2.3. Section 2.3 includes a review of describing finite–energy bandlimited waveforms as N–dimensional vectors. These channel models are used to obtain the Shannon bound in Section 2.4 and capacity for M–ary waveform signalling in Section 2.5. The conversion of channel capacity to bandwidth efficiency is presented in Section 2.6. The procedures for calculating capacity and efficiency are summarised in Section 2.7.

2.2 Capacity for Discrete Memoryless Channel

The capacity for the DMC is developed in this Section. This Section is a review of the work by [10, pp.13–27] and [12, pp. 67–74].

An example DMC is shown in Figure 2.1 . Let the input and output be the discrete random variables X and Y, respectively. X can be one of K values and Y one of J values. The assignment of $x = a_k$ and $y = b_j$ are called *events*. The sets of input and output values are

$$x \in \{a_k ; k = 1, \ldots, K\} ,$$
$$y \in \{b_j ; j = 1, \ldots, J\}. \tag{2.1}$$

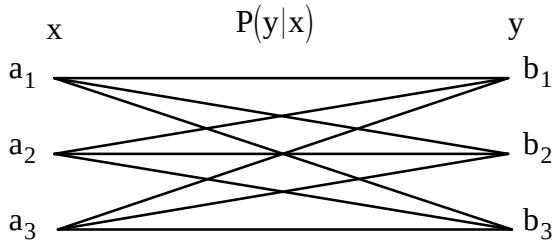


Figure 2.1 DMC for K=J=3

The probability of each event and transition probability are denoted

$$P(x = a_k) ,$$
$$P(y = b_j) ,$$
$$P(y = b_j | x = a_k). \tag{2.2}$$

Each line segment in Figure 2.1 represents $P(y = b_j | x = a_k)$.

Mutual information, $I_{X;Y}(a_k; b_j)$, is defined as

$$I_{X;Y}(a_k; b_j) = \log_2 \left[\frac{P(x = a_k | y = b_j)}{P(x = a_k)} \right]. \tag{2.3}$$

The base of the logarithm is 2 and the units of mutual information are bits. $I_{X;Y}(a_k; b_j)$ is a measure of "information provided about the event $x=a_k$ by the occurrence of the event $y=b_j$" [10, p. 16].

Average mutual information, I(X; Y), is the expectation of $I_{X;Y}(a_k; b_j)$

$$I(X; Y) = \sum_{k=1}^{K} \sum_{j=1}^{J} P(x = a_k, y = b_j) \log_2 \left[\frac{P(x = a_k | y = b_j)}{P(x = a_k)} \right] \text{ [bit/sym]}. \tag{2.4}$$

The unit, bit/sym, is used to indicate the number of bits per transmitted symbol. Using the abbreviation

$$P(x = a_k) = p(k) ,$$
$$P(y = b_j) = p(j) \tag{2.5}$$

and using the probability identities [13, p. 142]

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} ,$$

$$p(x, y) = p(y|x)p(x) \tag{2.6}$$

average mutual information is rewritten as

$$I(X; Y) = \sum_{k=1}^{K} \sum_{j=1}^{J} p(j|k)p(k) \log_2 \left[\frac{p(j|k)}{p(j)} \right] \text{ [bit/sym]}. \tag{2.7}$$

Channel capacity, C_{DMC}, is defined as the largest average mutual information. C_{DMC} is obtained by finding the set of input probability assignments, $\{p(k) \ ; \ k = 1, \ldots, K\}$, which maximises $I(X; Y)$. C_{DMC} is written as [10, p. 74]

$$C_{DMC} = \max_{\{p(k) \ ; \ k = 1, \ldots, K\}} \sum_{k=1}^{K} \sum_{j=1}^{J} p(j|k)p(k) \log_2 \left[\frac{p(j|k)}{p(j)} \right] \text{ [bit/sym]} \ . \tag{2.8}$$

The channel capacity formula, (2.8), represents the largest average mutual information for a set of possible channel inputs and outputs. The capacity can be expressed in bits per second to indicate the maximum bit rate for that channel. If a symbol enters the channel at a rate of 1/T sym/s, the capacity per second is denoted [12, p. 131]

$$C^*_{DMC} = \frac{C_{DMC}}{T} \text{ [bit/s]} \ . \tag{2.9}$$

2.3 Vector Channel Model

2.3.1 Introduction

This Section describes two variations of the DMC to represent the waveform channel of Chapter 1. The first channel type is the continuous input continuous output memoryless channel (CCMC) and will be referred to as a *continuous* channel. The second channel is the discrete input continuous output memoryless channel (DCMC).

Sections 2.3.2 and 2.3.3 describe vector channel models for the continuous channel and DCMC. Section 2.3.4 reviews the vector representation of a finite–energy bandlimited waveform.

2.3.2 Continuous Channel

A continuous channel is shown in Figure 2.2 [11, pp. 141–148], [12, pp. 74–75]. The input, x, is continuous valued over the interval $[-\infty, +\infty]$. The output, y, is a continuous random variable which is the input disturbed by additive noise, n.

If the channel input and noise process can be described using N–dimensional real vectors the channel model is shown in Figure 2.3 . Here the coefficients of $\underline{x}$ and $\underline{y}$ are continuous valued

$$x_n \in [-\infty, +\infty],$$

$$y_n \in [-\infty, +\infty], \quad n = 1, 2, \ldots, N. \tag{2.10}$$

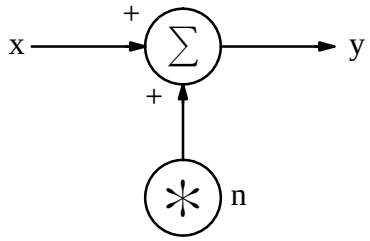


Figure 2.2 Continuous channel with single variable input and output

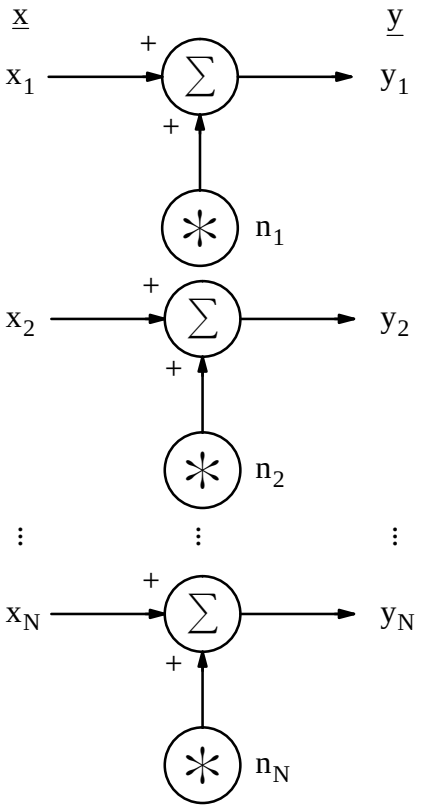


Figure 2.3 Continuous channel with vector input and output

2.3.3 Discrete Input Continuous Output Memoryless Channel

A DCMC is shown in Figure 2.4 [12, pp. 75–76, p. 132]. The input, x_m, belongs to the discrete set of M values

$$\{x_m; m = 1, 2, \ldots, M\}. \tag{2.11}$$

The output, y, is a continuous random variable which is the input disturbed by additive noise, n.

If the channel input and noise process can be described using N–dimensional real vectors the channel model is shown in Figure 2.5 . Here the coefficients of $\underline{x}_m$ are discrete valued and $\underline{y}$ continuous valued

$$\underline{x}_m = \left(x_{m1}, x_{m2}, \ldots, x_{mN}\right), \quad m = 1, 2, \ldots, M,$$

$$y_n \in [-\infty, +\infty], \quad n = 1, 2, \ldots, N \ . \tag{2.12}$$

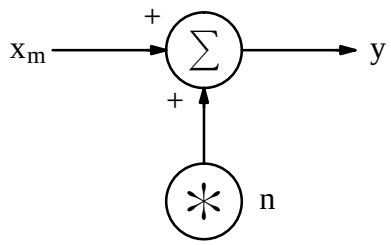


Figure 2.4 DCMC with single variable input and output

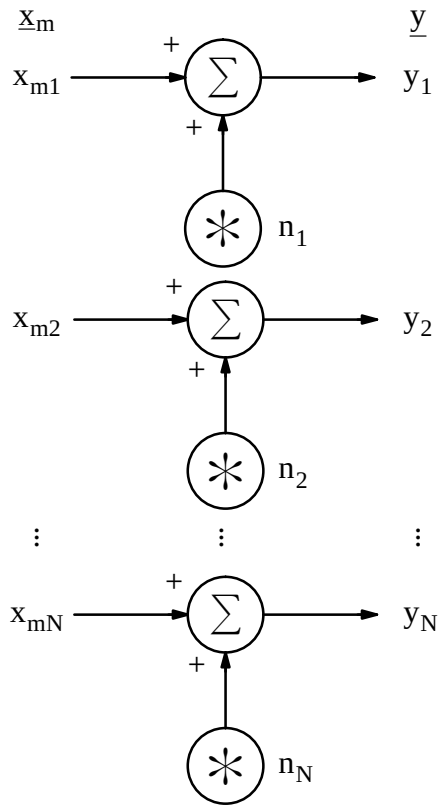


Figure 2.5 DCMC with vector input and output

2.3.4 Representation of Finite–Energy Bandlimited Signals as Vectors

In this Section we review methods for describing finite–energy bandlimited signals as N–dimensional vectors. The following assumptions are made.

1. The finite–energy waveform is strictly limited to the time interval, $0 \le t \le T$.
2. The waveform is constrained to an ideal lowpass or bandpass bandwidth, W.

Assumptions 1 and 2 cannot be fullfilled simultaneously. However, if a signal is strictly time–limited it can be shown that it is possible to keep the spectrum very small outside the band, W [14, pp. 348–350]. Similarly, if a signal is strictly band–limited then the signal can be kept very small outside the interval, T [3]. The derivation of the Shannon

capacity bound using physically consistent mathematical models for the band–limited Gaussian channel can be found in [15].

We express a waveform, x(t), as a linear combination of orthonormal functions, $\{\phi_n(t)\}$,

$$x(t) = \sum_{n=1}^{N} x_n \phi_n(t) \ . \tag{2.13}$$

The coefficients, x_n, are obtained by

$$x_n = \int_0^T x(t)\phi_n(t)\ dt \tag{2.14}$$

for all n, and

$$N = 2WT \tag{2.15}$$

where N is the dimensionality of the signal space. The vector representation of x(t) is the vector of coefficients, $\underline{x}$,

$$\underline{x} = \left(x_1, x_2, \ldots, x_N\right) \ . \tag{2.16}$$

The sampling theorem justifies the vector representation of finite–energy bandlimited waveforms. A waveform can be uniquely represented by samples taken at the Nyquist rate, 2W samples per second. We denote the n'th sample

$$x_n = x\left(\frac{n}{2W}\right) \ . \tag{2.17}$$

The waveform can be reconstructed using the interpolation formula [12, p. 53]

$$x(t) = \sum_{n=-\infty}^{\infty} x_n \frac{\sin 2\pi W\left(t-\frac{n}{2W}\right)}{2\pi W\left(t-\frac{n}{2W}\right)} \ . \tag{2.18}$$

Hence an orthonormal function is described by

$$\phi_n(t) = \frac{\sin 2\pi W\left(t-\frac{n}{2W}\right)}{2\pi W\left(t-\frac{n}{2W}\right)} \ . \tag{2.19}$$

If the function, x(t), is zero outside the interval, T, then the function is fully defined by N=2WT samples. All other samples will be zero. The spectrum of the orthonormal function, $\phi_n(t)$, is constant in the band, W, and zero outside [3].

There are other methods for generating a set of orthonormal functions. One method is to write a Fourier series and another is the application of the Gram–Schmidt procedure.

If the time limited waveform is extended to repeat every T seconds then a Fourier series can be determined for the periodic signal [13, pp. 211–215]. Under the restriction that the signal is strictly bandlimited then all coefficients of sine and cosine frequency terms outside the band, W, are zero. The number of sine and cosine terms is WT hence the

waveform is uniquely specified by N=2WT coefficients. The sine and cosine frequency terms form the set of orthonormal functions.

If a set of M waveforms are known explicitly as a function of time then a set of $N \leq M$ orthonormal functions can be obtained by application of the Gram–Schmidt orthogonalisation procedure [14, pp. 266–273]. We assume the set of M signals hence the set of orthonormal functions satisfy the restriction of being time and bandlimited. With this assumption the dimensionality of the signal space remains N=2WT [14, pp. 348–351].

A number of methods to generate a set of orthonormal functions have been presented. The dimensionality of the discrete signal space is always N=2WT. The geometric configuration (or signal constellation) remains the same regardless the choice of orthonormal functions [14, p. 273].

2.4 Shannon's Theorem for Capacity of a Bandlimited, Average Power Constrained, Continuous Channel

This Section presents Shannon's famous bound for the bandlimited AWGN channel [2, 3]. The modulation techniques in Chapter 3 are compared against this bound.

The Shannon bound is obtained by finding the capacity of a continuous AWGN channel with bandlimited Gaussian noise at the input as shown in Figure 2.6 . Ideal bandlimiting filters are included to meet the input bandwidth restriction [11, pp. 157–163] and Nyquist sampling requirement.

We assume the two noise sources, x(t) and n(t), are Gaussian. We assume samples of the noise sources are taken at the Nyquist rate (after bandlimiting), and are independent identically distributed (*iid*) Gaussian random variables with zero mean and variance, σ^2, for x(t) and, $N_0/2$, for n(t).

Vector representation of the channel input and output is obtained by sampling at the Nyquist rate. By defining a symbol period, T, the time–continuous waveforms can be described as vectors of N=2WT samples. The equivalent vector channel is shown in Figure 2.7 .

The channel capacity of the DMC can be extended for the continuous channel

$$C_{CCMC} = \max_{p(\underline{x})} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{\text{2N–fold}} p(\underline{y}|\underline{x}) p(\underline{x}) \log_2 \left[\frac{p(\underline{y}|\underline{x})}{p(\underline{y})} \right] d\underline{x} d\underline{y} \text{ [bit/sym]} . \tag{2.20}$$

Knowing that $p(\underline{x})$, $p(\underline{y})$, and $p(\underline{y}|\underline{x})$, are Gaussian the Shannon bound is derived in Appendix A. The Shannon bound is denoted C

$$C = WT \log_2(1 + SNR) \text{ [bit/sym]} . \tag{2.21}$$

Denote $C^* = C/T$ [11, p. 161] then

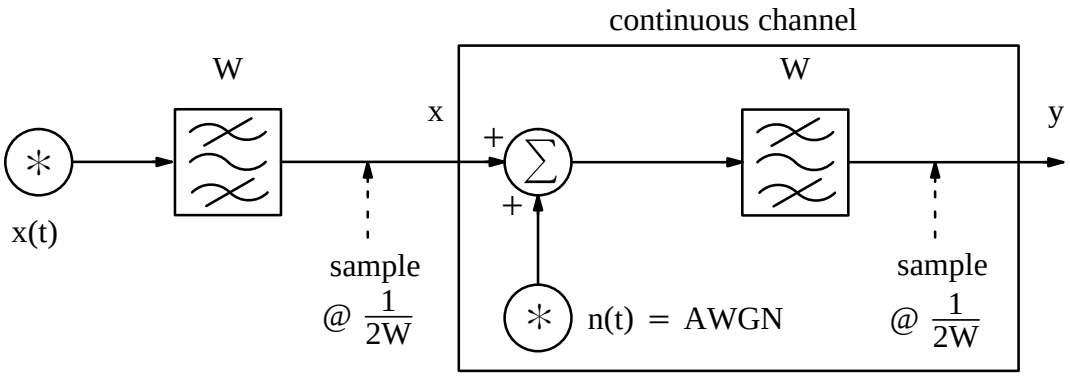


Figure 2.6 Channel model for calculating the Shannon bound

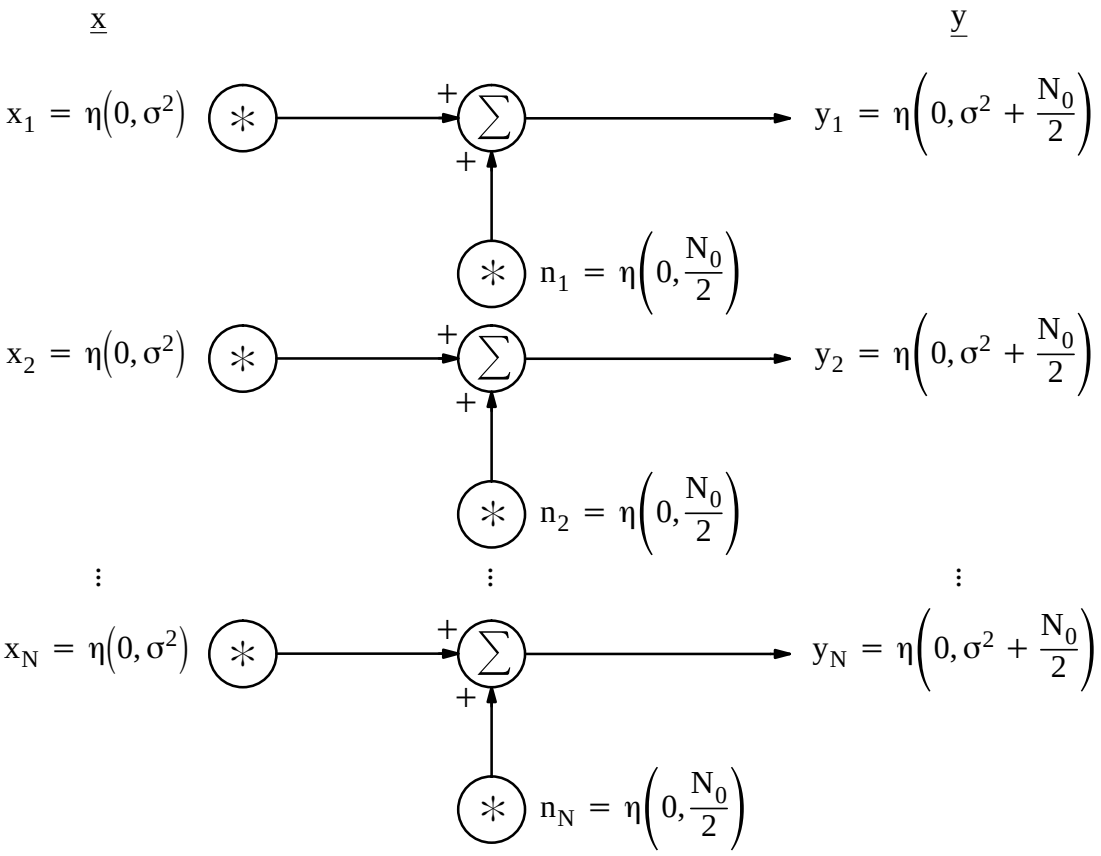


Figure 2.7 Vector channel model for calculating the Shannon bound

$$C^* = W \log_2(1 + SNR) \text{ [bit/s]} . \tag{2.22}$$

To compare the Shannon bound against signal sets of different dimensionality we rewrite (2.21) and (2.22) as a function of N using (2.15)

$$C = \frac{N}{2} \log_2(1 + SNR) \text{ [bit/sym]} ,$$

$$C^* = \frac{N}{2T} \log_2(1 + SNR) \text{ [bit/s]} . \tag{2.23}$$

The Shannon bound is the maximum rate for transmitting binary digits over a bandlimited AWGN channel [3]. By operating signalling schemes at a rate less than C^* it is

possible to keep the probability of error arbitrarily small. When signalling at a rate greater than C^* the error probability is close to unity for any set of M–ary signals [14, p.321].

2.5 Capacity for Discrete Input Continuous Output Memoryless Channel

This Section presents the capacity for a DCMC with AWGN. The capacity of the modulation techniques in Chapter 3 are calculated using this result.

The DCMC is drawn in Figure 2.8 resembling the digital communications system of Chapter 1. A message block, d_m, is mapped into a transmitted waveform, $x_m(t)$. The received waveform, $r_m(t)$, is the transmitted waveform disturbed by AWGN, n(t). The received signal is coherently detected and processed to provide an estimate of the transmitted message, $\hat{d}_m$.

We assume the set of M–ary symbols, $\{x_m(t)\ ; m = 1, \ldots, M\}$, are finite–energy bandlimited waveforms with equal symbol period, T. Application of the Gram–Schmidt

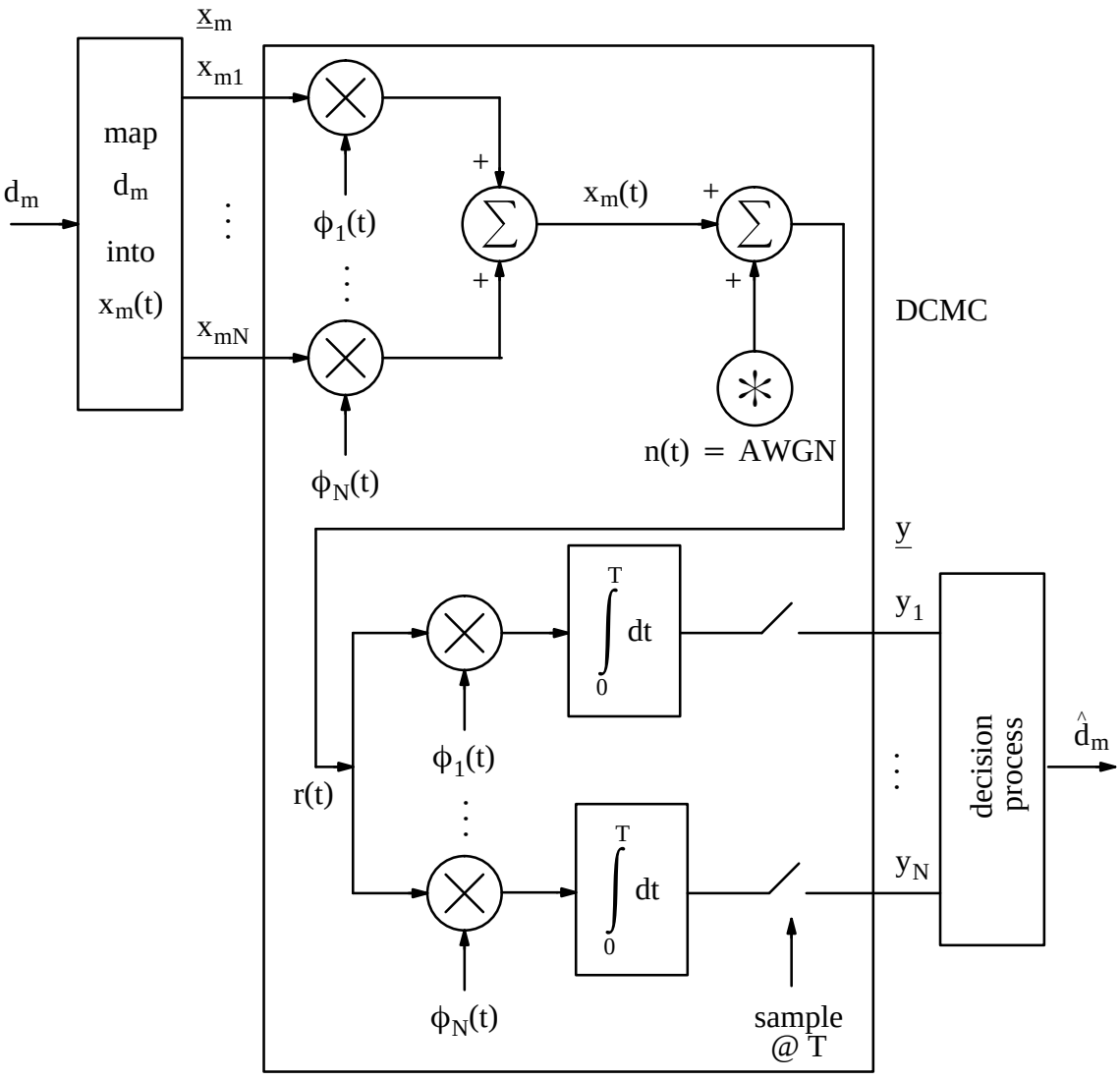


Figure 2.8 DCMC model for M–ary waveform system

orthogonalisation procedure yields a set of orthonormal functions, $\{\phi_n(t)\ ; n = 1, \ldots, N\}$. Each waveform can be expressed as a linear combination [16, pp. 47–54]

$$x_m(t) = \sum_{n=1}^{N} x_{mn}\phi_n(t)\ ;\ m = 1, 2, .., M \tag{2.24}$$

where the coefficients, x_{mn}, are obtained by

$$x_{mn} = \int_0^T x_m(t)\phi_n(t)\ dt \tag{2.25}$$

for each m and n. The vector representation of $x_m(t)$ is the vector of coefficients, $\underline{x}_m$,

$$\underline{x}_m = \left(x_{m1}, \ldots, x_{mN}\right)\ ; m = 1, \ldots, M. \tag{2.26}$$

The synthesis of $x_m(t)$ is detailed in Figure 2.8 .

Let the noise process be written as

$$n(t) = \sum_{n=1}^{N} n_n\phi_n(t) + \hat{n}(t)\ . \tag{2.27}$$

The coefficients of the noise process, $\{n_n\ ; n = 1, \ldots, N\}$, are *iid* Gaussian random variables with zero mean and variance, $N_0/2$. The probability distribution function of each noise coefficient is written as

$$p(n_n) = \eta\left(0, \frac{N_0}{2}\right)\ . \tag{2.28}$$

The second portion, $\hat{n}(t)$, is that noise which is orthogonal to the signal space.

Demodulation of the received signal involves correlation detection as shown in Figure 2.8 . The correlator outputs are sampled at time, T, to provide the output sample, $\underline{y}$, which is processed to yield $\hat{d}_m$. The correlator detection removes $\hat{n}(t)$.

The equivalent vector channel is shown in Figure 2.9 . The statistic of the channel is represented by

$$\begin{aligned} p\left(\underline{y}|\underline{x}_m\right) &= \prod_{n=1}^{N} p\left(y_n|x_{mn}\right) \\ &= \prod_{n=1}^{N} \frac{1}{\sqrt{\pi N_0}} \exp\left[\frac{-\left(y_n - x_{mn}\right)^2}{N_0}\right]\ . \end{aligned} \tag{2.29}$$

By rewriting the capacity for the DMC the capacity for the DCMC is

$$C_{DCMC} = \max_{p\left(\underline{x}_1\right) \cdots\ p\left(\underline{x}_M\right)} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{\text{N–fold}} p\left(\underline{y}|\underline{x}_m\right) p\left(\underline{x}_m\right) \log_2\left[\frac{p\left(\underline{y}|\underline{x}_m\right)}{p\left(\underline{y}\right)}\right] d\underline{y}\ \text{[bit/sym]} \tag{2.30}$$

Assuming equally likely inputs and channel statistic given by (2.29) the capacity is derived in Appendix B. The channel capacity can be rewritten as

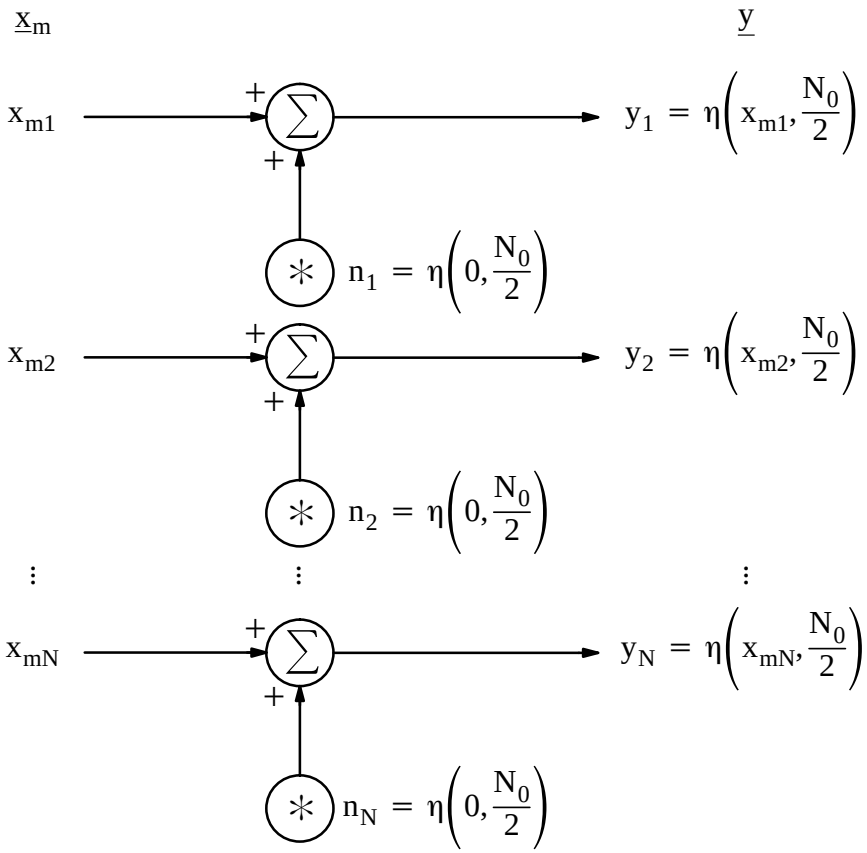


Figure 2.9 Vector channel model for DCMC

$$C_{DCMC} = \log(M) - \frac{1}{M\left(\sqrt{\pi}\right)^N} \cdot$$

$$\sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{N\text{–fold}} \exp\left(-|\underline{t}|^2\right) \log_2\left[\sum_{i=1}^{M} \exp\left(-2\underline{t} \cdot \underline{d}_{mi} - |\underline{d}_{mi}|^2\right)\right] d\underline{t} \quad [\text{bit/sym}] \tag{2.31}$$

where $\underline{d}_{mi} = \left(\underline{x}_m - \underline{x}_i\right)/\sqrt{N_0}$ and with new variable of integration $\underline{t} = \left(t_1, t_2, \ldots, t_N\right)$. We denote capacity per second in a similar manner as for the DMC

$$C^*_{DCMC} = \frac{C_{DCMC}}{T} \ [\text{bit/s}] \ . \tag{2.32}$$

The average SNR is expressed [4]

$$SNR = \frac{E\left\{x_m^2(t)\right\}}{E\left\{n^2(t)\right\}} \ . \tag{2.33}$$

We denote average symbol energy, E_s ,

$$\begin{aligned} E_s &= E\left\{x_m^2(t)\right\} \\ &= \frac{1}{M} \sum_{m=1}^{M} |\underline{x}_m|^2 \end{aligned} \tag{2.34}$$

and average noise energy, $E\{n^2(t)\}$, is expressed

$$E\{n^2(t)\} = \sum_{n=1}^{N} E\{n_n^2\}$$

$$= N\frac{N_0}{2} \ . \tag{2.35}$$

The SNR can be written

$$SNR = \frac{2E_s}{NN_0} \ . \tag{2.36}$$

The SNR (2.36) is a power ratio and is equivalent to the definition of SNR for the Shannon bound (A18) using (2.15)

$$SNR = \frac{E_s}{TWN_0} \ . \tag{2.37}$$

We observe for high SNR that N_0 is very small and $\underline{d}_{mi}$ becomes very large. The exponential terms in the logarithm of (2.31) become very small forcing the logarithm to become very small. Therefore at high SNR

$$C_{DCMC} = \log_2(M) \text{ [bit/sym]} \ . \tag{2.38}$$

The presence of a logarithm of a sum in (2.31) does not allow any reduction of the dimensionality of the N–fold integral. This integral is evaluated numerically using a technique outlined in Chapter 4.

2.6 Conversion of Capacity Data to Bandwidth Efficiency

The capacity analysis of the continuous channel and DCMC yields the number of useful information bits per symbol as a function of SNR. The bandwidth efficiency is the capacity normalised by occupied bandwidth and has units, bit/s/Hz, and is plotted as a function of bit energy to noise density ratio, E_b/N_0 .

2.6.1 Shannon Bound

The Shannon bound can be expressed as bandwidth efficiency, $\eta = C^*/W$, by using (2.22)

$$\eta = \log_2(1 + SNR) \text{ [bit/s/Hz]} \tag{2.39}$$

From Appendix A

$$SNR = \frac{E_s}{WTN_0} \tag{2.40}$$

where E_s/T is the average transmitter power.
Let

$$E_s = CE_b$$

then

$$\mathrm{SNR} = \frac{C}{WT}\frac{E_b}{N_0} = \eta\frac{E_b}{N_0} \ . \tag{2.41}$$

We rewrite (2.39)

$$\eta = \log_2\left(1 + \eta\frac{E_b}{N_0}\right)$$

$$\frac{E_b}{N_0} = \frac{2^{\eta}-1}{\eta} \ . \tag{2.42}$$

We can evaluate E_b/N_0 for $\eta \to 0$ by using L'Hôpital's rule [17, p. 502]

$$\begin{aligned}\lim_{\eta \to 0} \frac{E_b}{N_0} &= \lim_{\eta \to 0} \frac{2^{\eta}-1}{\eta} \\ &= \lim_{\eta \to 0} 2^{\eta}\ln(2) \\ &= \ln(2) \\ &\approx -1.6 \text{ [dB]} \ .\end{aligned} \tag{2.43}$$

2.6.2 DCMC

The bandwidth efficiency for the DCMC is obtained by first calculating C_{DCMC} for a value of SNR. We then normalise C_{DCMC} with respect to WT and convert SNR to E_b/N_0.

Bandwidth efficiency for the DCMC is written

$$\eta_{DCMC} = \frac{C_{DCMC}}{WT} = \frac{C_{DCMC}}{(N/2)} \text{ [bit/s/Hz]} \tag{2.44}$$

and we obtain E_b/N_0 similarly to (2.41)

$$\mathrm{SNR} = \eta_{DCMC}\frac{E_b}{N_0} \ . \tag{2.45}$$

Using logarithmic values

$$\frac{E_b}{N_0} \text{ [dB]} = \mathrm{SNR} \text{ [dB]} - 10\log_{10}\left(\eta_{DCMC}\right) \ . \tag{2.46}$$

2.7 Conclusion

This Chapter presented capacity formulas for transmission of finite–energy bandlimited waveforms over an AWGN channel. By using vector representation, the waveform channel can be modelled as a discrete–time memoryless channel. The Shannon bound was derived using the continuous channel and capacity for M–ary waveform signalling was derived using the DCMC.

The Shannon bound is calculated using (2.23) and can be expressed as bandwidth efficiency using (2.39). The capacity for M–ary signalling is given by (2.31). The capacity formula, (2.31), is calculated using an integration technique from Chapter 4. This capacity formula requires the coefficients of the signal vectors and the noise variance. The variance is calculated using (2.34) and (2.36). The bandwidth efficiency for M–ary signalling is obtained using (2.44) and (2.46).

The formula for C_{DCMC} is written as an N–dimensional integral. Chapters 3 and 4 discuss the reduction in decimal place accuracy when calculating C_{DCMC} for $N \geq 8$. Further research may yield a bound on C_{DCMC} which does not require N–dimensional integration. The capacity of signal sets with large N could be analysed using this bound.

3 CHANNEL CAPACITY FOR M–ARY DIGITAL SIGNAL SETS

This Chapter studies the capacity of well known M–ary digital signal sets plus a hybrid scheme. For each signal set the signal constellation is described first followed by plots of channel capacity and bandwidth efficiency. The probability of symbol error, P_e, and bit error, P_b, at 10^{-5} are included. All capacity data is calculated for SNR over the range –10 dB to +30 dB in steps of 0.1 dB.

3.1 Pulse Amplitude Modulation

3.1.1 Description

Digital pulse amplitude modulation (PAM) is a scheme where the information is contained in the amplitude of the transmitted waveform. Let the set of waveforms be described by [12, pp. 272–273], [14, p. 312], and [18, p. 341]

$$x_m(t) = \begin{cases} (2m-1-M)\sqrt{\frac{2}{T}}\cos\omega_0 t\ , & 0 \le t \le T \\ 0 & , \text{ elswhere} \end{cases}$$

$$, \ m = 1, 2, \ldots, M\ . \tag{3.1}$$

The single orthonormal function is

$$\phi_1(t) = \sqrt{\frac{2}{T}}\cos\omega_0 t \tag{3.2}$$

and each signal vector is given by

$$\underline{x}_m = 2m - 1 - M\ ,\ m = 1, 2, ..., M. \tag{3.3}$$

The signal constellation for PAM is a set of evenly spaced points on the real number line. Figure 3.1 shows constellations for M=2, 4, and 8. The symbol, O, represents the origin.

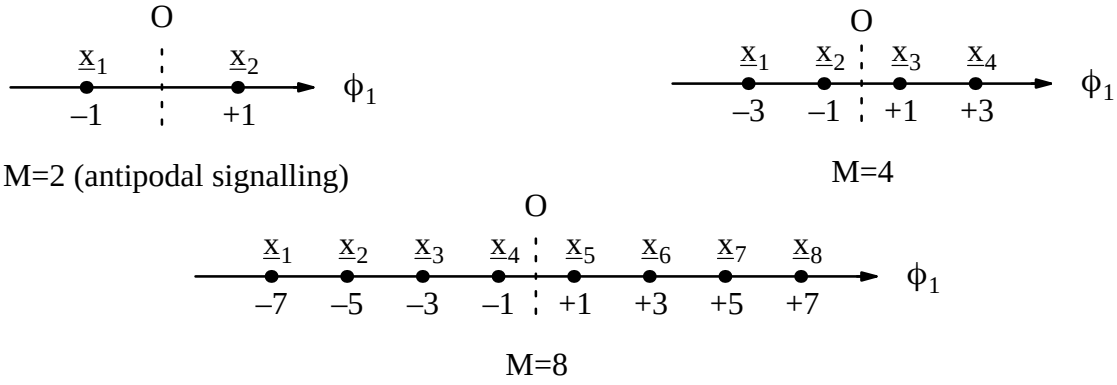


Figure 3.1 PAM constellations

3.1.2 Channel Capacity

The channel capacity of PAM is shown in Figure 3.2 . These results agree with [4] and [7, p. 275]. The error rate data is calculated in Appendix C1.

To interpret these results [4] consider the transmission of 2 bit/sym using M=4 PAM. At SNR=16.77 dB the symbol error rate is 10^{-5}. If the number of signals is doubled then error free transmission of 2 bit/sym is possible using M=8 PAM at SNR=12.6 dB (assuming unlimited coding and decoding effort). This reduction in SNR requirement represents a potential coding gain of 4.17 dB.

For any M, doubling of the number of signals yields nearly all the coding gain that is possible. Further expansion results in little additional coding gain.

The bandwidth efficiency of PAM is shown in Figure 3.3 . Here, the asymptotic value of bandwidth efficiency is double the channel capacity. This characteristic can be achieved by one of two methods [12, p. 277]. The first method involves transmitting PAM as single sideband (SSB) modulation. The second method is to split the information sequence into two parallel sequences and transmit them as two half–rate PAM signals on quadrature carriers.

3.2 Quadrature Amplitude Modulation

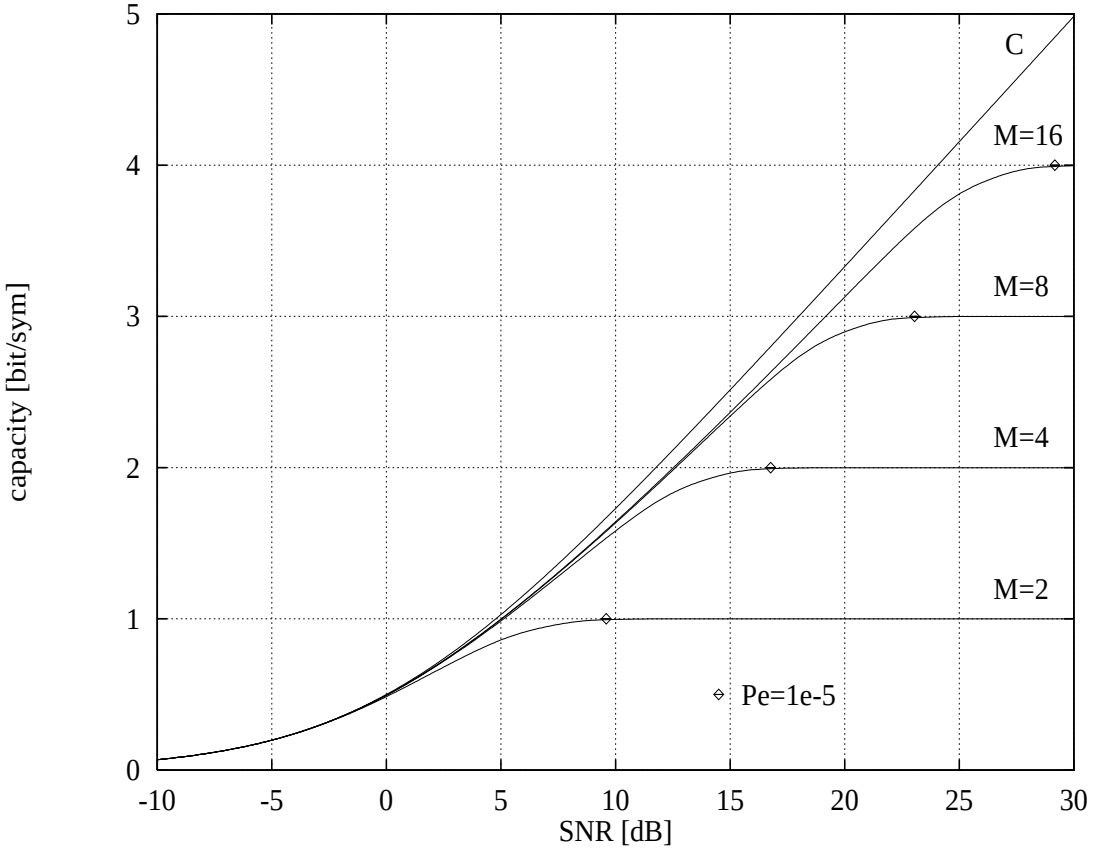


Figure 3.2 Channel capacity for PAM

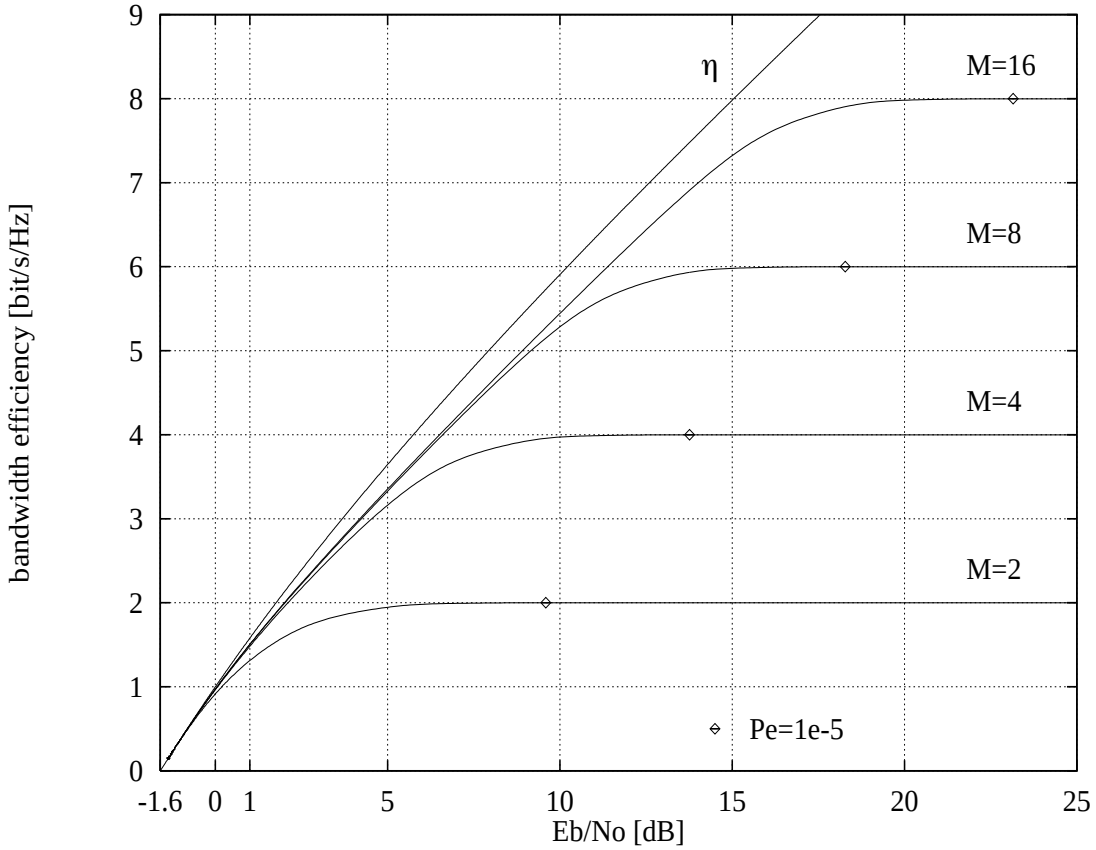


Figure 3.3 Bandwidth efficiency for PAM

3.2.1 Description

Quadrature amplitude modulation (QAM) is two PAM signals in quadrature

$$x_m(t) = \begin{cases} x_{m1}\sqrt{\frac{2}{T}}\cos\omega_0 t + x_{m2}\sqrt{\frac{2}{T}}\sin\omega_0 t\ , & 0 \le t \le T \\ 0 & , \text{ elsewhere} \end{cases}$$

$$, \quad m = 1, 2, \ldots, M\ . \tag{3.4}$$

Each message block is mapped onto a pair of coordinates on a rectangular grid. All QAM constellations used for capacity analysis are shown in Appendix D. A rectangular constellation is used for simplifying error bounding. Alternate QAM constellations will offer better error rate performance [12, pp. 278–285]. The signal vectors and orthonormal functions are given by

$$\begin{aligned} \underline{x}_m &= \left(x_{m1}, x_{m2}\right) \quad , \ m = 1, 2, \ldots, M\ , \\ \phi_1(t) &= \sqrt{\frac{2}{T}}\cos\omega_0 t\ , \ 0 \le t \le T \\ \phi_2(t) &= \sqrt{\frac{2}{T}}\sin\omega_0 t\ , \ 0 \le t \le T. \end{aligned} \tag{3.5}$$

3.2.2 Channel Capacity

The channel capacity of QAM is shown in Figure 3.4 . These results agree with [4] and [7, p. 277]. The error rate data is calculated in Appendix C2.

At a capacity of 3 bit/sym a potential coding gain of 7.57 dB is possible by doubling M=8 QAM to M=16 QAM. An additional coding gain of 0.3 dB is achieved using M=32 QAM. For any M, the doubling of the number of signals yields nearly all the coding gain that is possible.

The bandwidth efficiency of QAM is shown in Figure 3.5 . Here, the asymptotic value of bandwidth efficiency is equal to the capacity because the transmitted waveform consists of two modulated carriers transmitted in quadrature.

3.3 Phase Shift Keying

3.3.1 Description

Phase shift keying (PSK) is a scheme where the information is contained in the phase of the transmitted carrier. The set of waveforms is described by [18, pp. 340–341]

$$x_m(t) = \begin{cases} \sqrt{\frac{2E_s}{T}}\cos\left(\omega_0 t - \frac{2\pi m}{M}\right), & 0 \le t \le T \\ 0, & \text{elsewhere} \end{cases}$$

$$, \quad m = 1, 2, \ldots, M \; . \tag{3.6}$$

The PSK signal set is an equal energy set and the constellation is contained on a circle of radius $\sqrt{E_s}$. The orthonormal functions are the same as for QAM. A signal vector is described by the coefficients

$$x_{m1} = \sqrt{E_s}\cos\left(\frac{2\pi m}{M}\right) ,$$

$$x_{m2} = \sqrt{E_s}\sin\left(\frac{2\pi m}{M}\right) , m = 1, 2, \ldots, M \; . \tag{3.7}$$

The signal constellations for M=2, 4, and 8 are shown in Figure 3.6 .

3.3.2 Channel Capacity

The channel capacity of PSK is shown in Figure 3.7 . These results agree with [4] and [7, p. 277]. The error rate data is calculated in Appendix C3.

At a transmission of 2 bit/sym a coding gain of 7.1 dB is possible by doubling M=4 PSK to M=8 PSK. Further expansion results in little additional coding gain. For any M, the doubling of the number of signals yields nearly all the coding gain that is possible.

The coding gain obtained by doubling M for QAM is approximately 0.3 dB better than for PSK. The coding gain from MPSK to 2MQAM is larger than using 2MPSK.

The bandwidth efficiency of PSK is shown in Figure 3.8 . Here, the asymptotic value of bandwidth efficiency is equal to the capacity as for QAM.

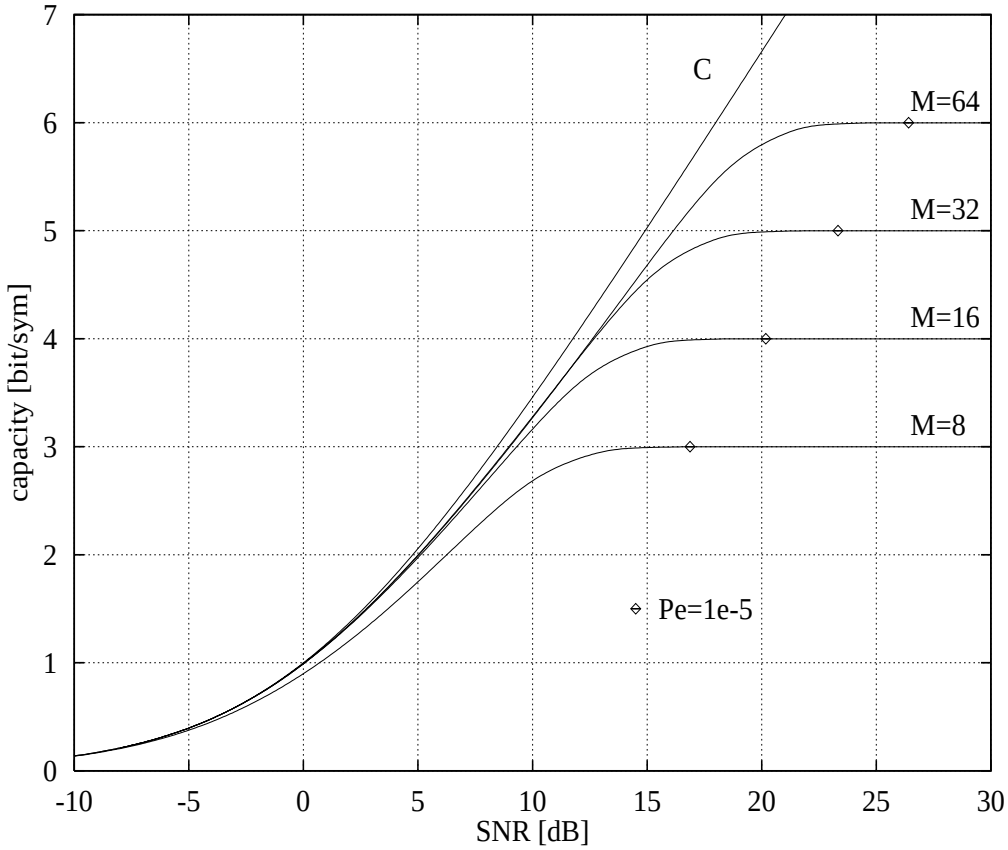


Figure 3.4 Channel capacity for QAM

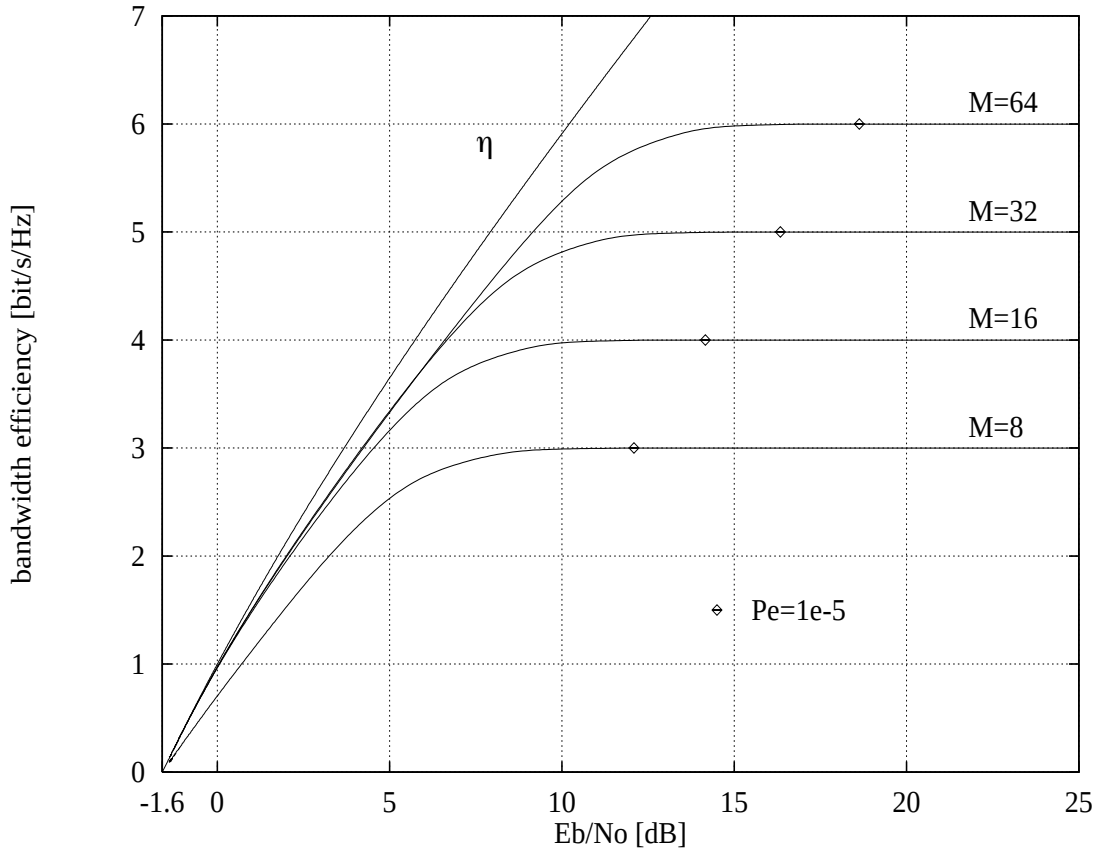


Figure 3.5 Bandwidth efficiency for QAM

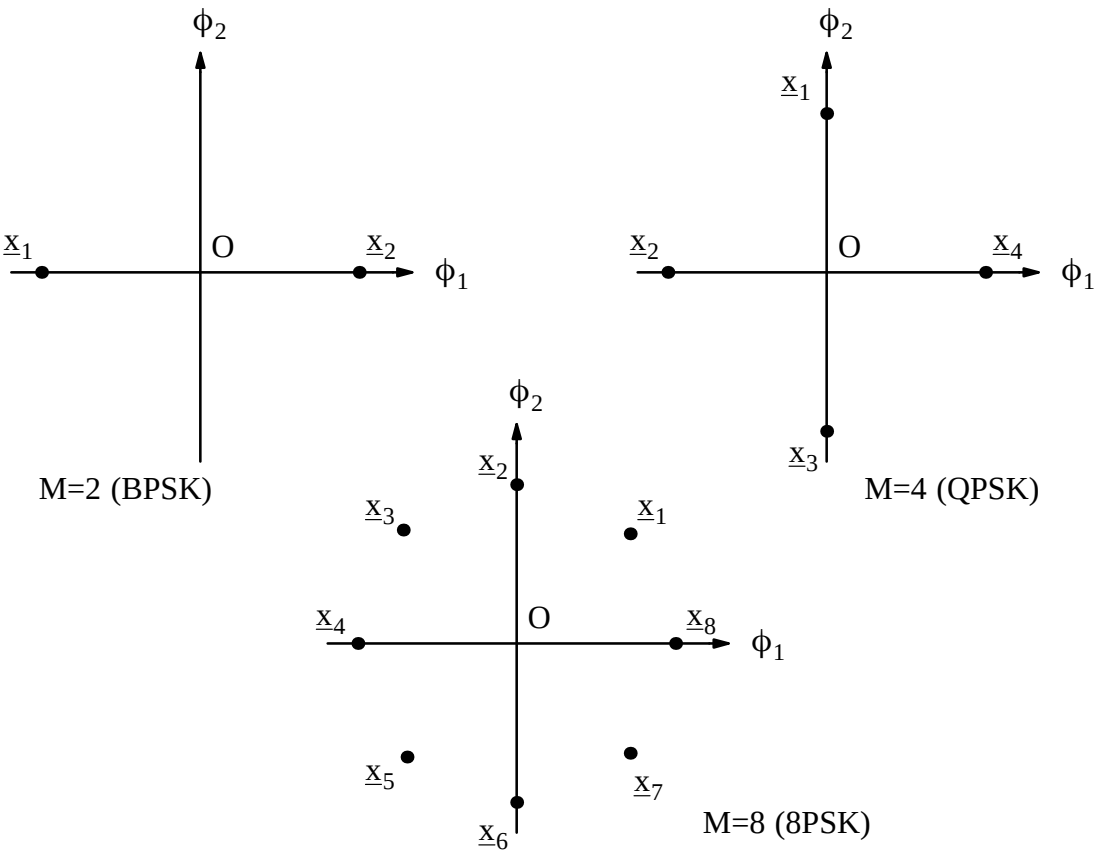


Figure 3.6 PSK constellations

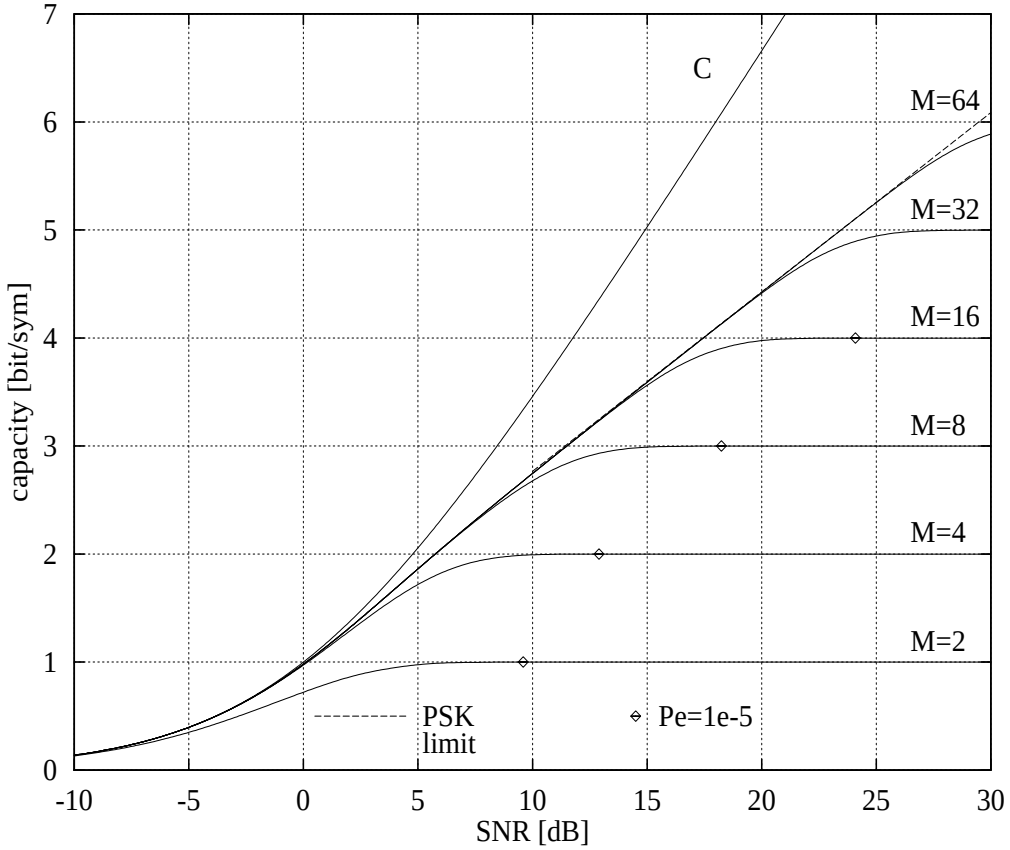


Figure 3.7 Channel capacity for PSK

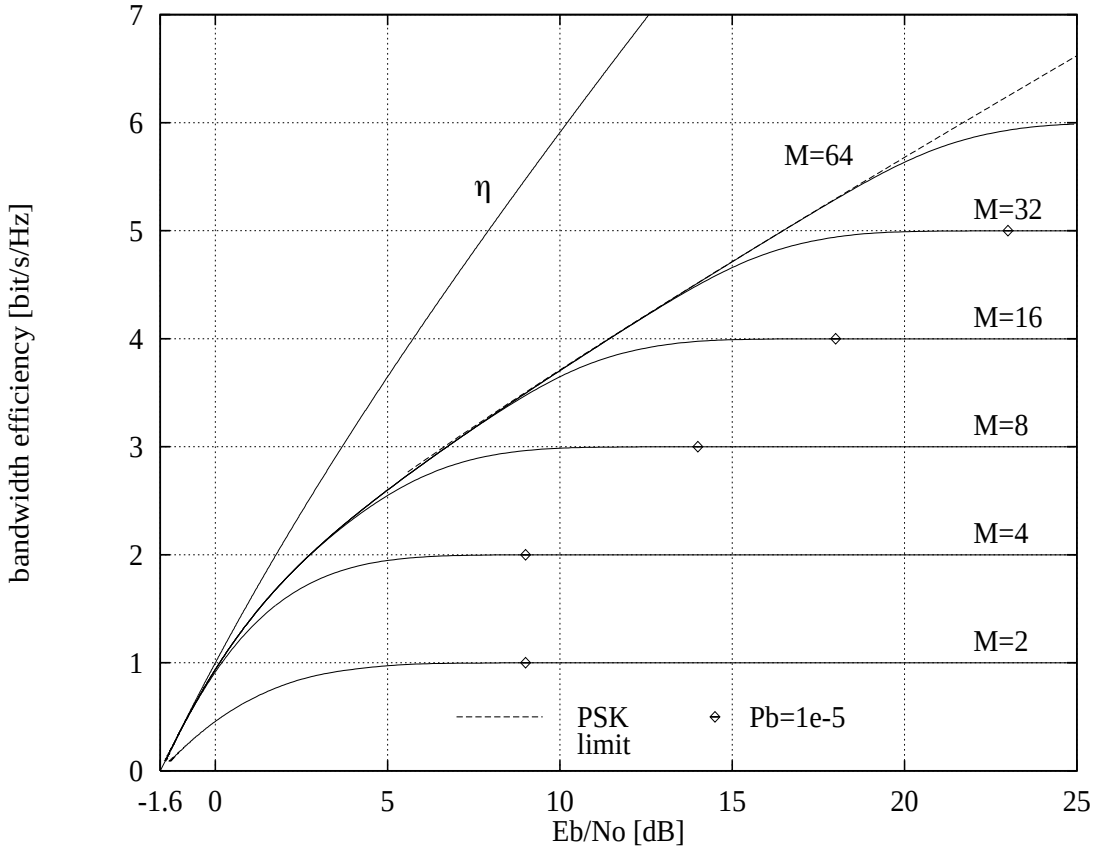


Figure 3.8 Bandwidth efficiency for PSK

The gap between the Shannon bound and capacity for MPSK becomes larger for increasing M. The PSK curves converge to the PSK limit [7, pp. 276–279]

$$C_{\text{PSK LIMIT}} = \log_2 \sqrt{\frac{4\pi}{e}\text{SNR}} \ \text{[bit/sym]}. \tag{3.8}$$

3.4 M–ary Orthogonal Signalling

3.4.1 Description

M–ary orthogonal signalling is a scheme where each signal is orthogonal to each other [13, p.238]. The signal constellation is interpreted as a collection of points in M–dimensional space with one point located on each coordinate axis. Each point is located at a distance of $\sqrt{E_s}$ from the origin.

The vector representation is described by

$$\begin{aligned} \underline{x}_1 &= \sqrt{E_s}(1, 0, \ldots, 0) = \sqrt{E_s}\phi_1 \ , \\ \underline{x}_2 &= \sqrt{E_s}(0, 1, \ldots, 0) = \sqrt{E_s}\phi_2 \ , \\ &\vdots \\ \underline{x}_M &= \sqrt{E_s}(0, 0, \ldots, 1) = \sqrt{E_s}\phi_M \ . \end{aligned} \tag{3.9}$$

Constellations for orthogonal signalling are shown in Figure 3.9 .

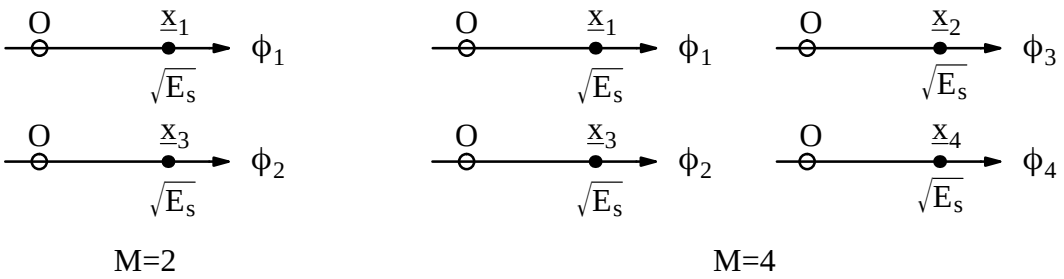
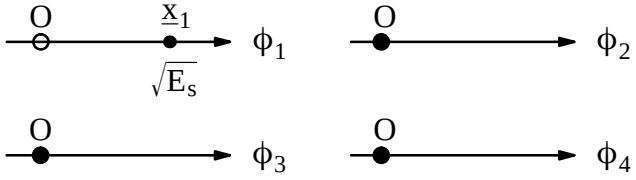


Figure 3.9 M–ary orthogonal signalling constellations

The signal constellations are interpreted in the following manner. The empty circles represent the origin. Since a transmitted signal is represented as a linear combination of $\{\phi_n\}$ then all orthonormal functions not associated with the transmitted $\underline{x}_m$ are weighted with zero. Figure 3.10 shows an example when $\underline{x}_1$ is transmitted for M=4 with the origin of the other $\{\phi_{n\neq 1}\}$ marked by a solid dot.



transmit $\underline{x}_1$

Figure 3.10 Example of signal transmission for M=4 orthogonal signalling

Orthogonal signal sets can be constructed in several ways. One method is M–ary frequency shift keying (MFSK) where the set of possible transmitted frequencies are spaced sufficiently far apart to be orthogonal [14, pp. 642–645]. Another method is to subdivide the interval, T, into M time slots and transmit one of M non–overlapping waveforms [14, pp. 290–292]. A carrier may be modulated by one of $M = 2^k$ binary sequences generated by a $\left(2^k, k\right)$ Haddamard code [19, p. 190].

3.4.2 Channel Capacity

The channel capacity of M–ary orthogonal signalling is shown in Figure 3.11 . Here, the bandwidth requirement for each signal set is different and so the Shannon bound is not plotted. The error rate data is calculated in Appendix C4. In this instance we have attempted N–dimensional integration for N=8.

The channel capacity curves in Figure 3.11 follow a different characteristic compared to $N \leq 2$ signal sets. Here, the capacity curves begin to asymptote at a lower SNR for increasing M.

The bandwidth efficiency of M–ary orthogonal signalling is shown in Figure 3.12 . Here, orthogonal signalling is characterised by none of the efficiency curves converging to the Shannon bound.

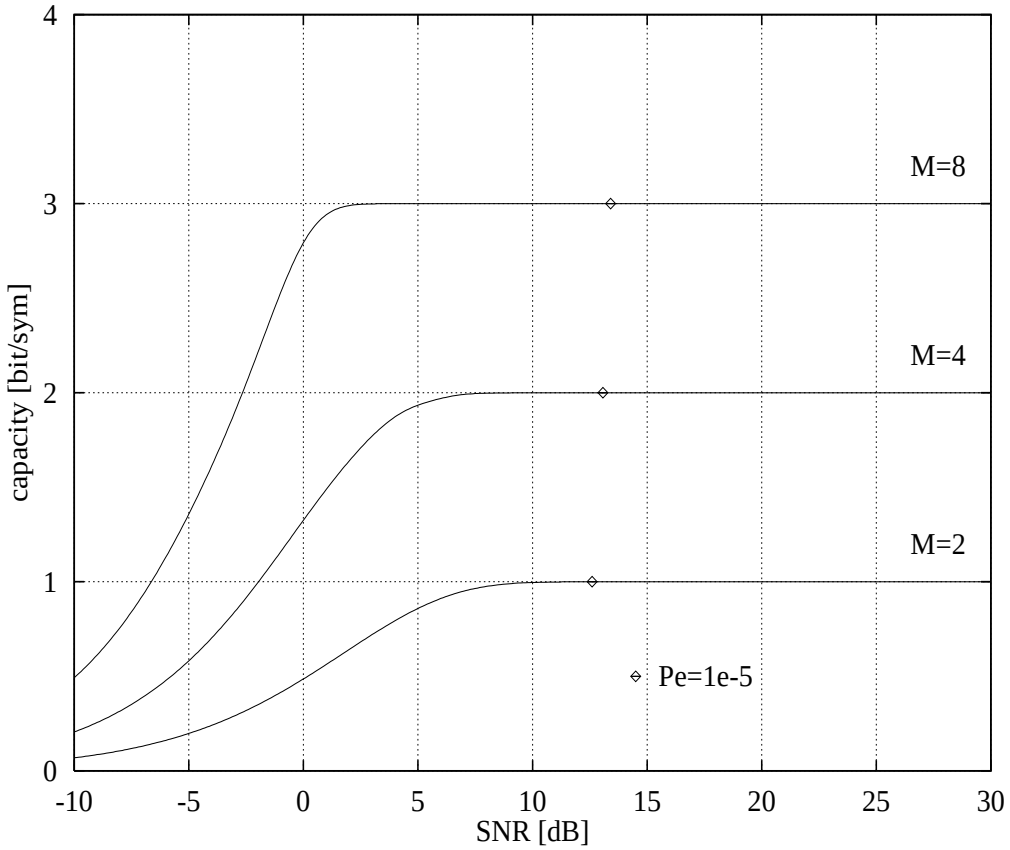


Figure 3.11 Channel capacity for M–ary orthogonal signalling

M–ary orthogonal signal sets are characterised by a non–zero centre of gravity [13, p. 245], $\underline{\bar{x}}$,

$$\underline{\bar{x}} = \frac{1}{M}\sum_{m=1}^{M} \underline{x}_m \ . \tag{3.10}$$

Figure 3.13 shows $\underline{\bar{x}}$ for M=2. Here, $|\underline{\bar{x}}|^2 = E_s/2$ which indicates that 3dB of the transmitted power is a constant and conveys no information. This result implies that η_{DCMC} for M=2 will be zero at E_b/N_0 = 1.4 dB. For any M, the energy of $\underline{\bar{x}}$ is

$$\begin{aligned} |\underline{\bar{x}}|^2 &= \frac{E_s}{M^2}\sum_{n=1}^{M}|\phi_n|^2 \\ &= \frac{E_s}{M} \ . \end{aligned} \tag{3.11}$$

The fraction of transmitted power conveying information is (M–1)/M. Table 3.1 lists the E_b/N_0 at 0 bit/s/Hz for orthogonal signalling. The bandwidth efficiency curves in Figure 3.12 are extended to 0 bit/s/Hz. The extension of the efficiency curves are indicated by dashed lines.

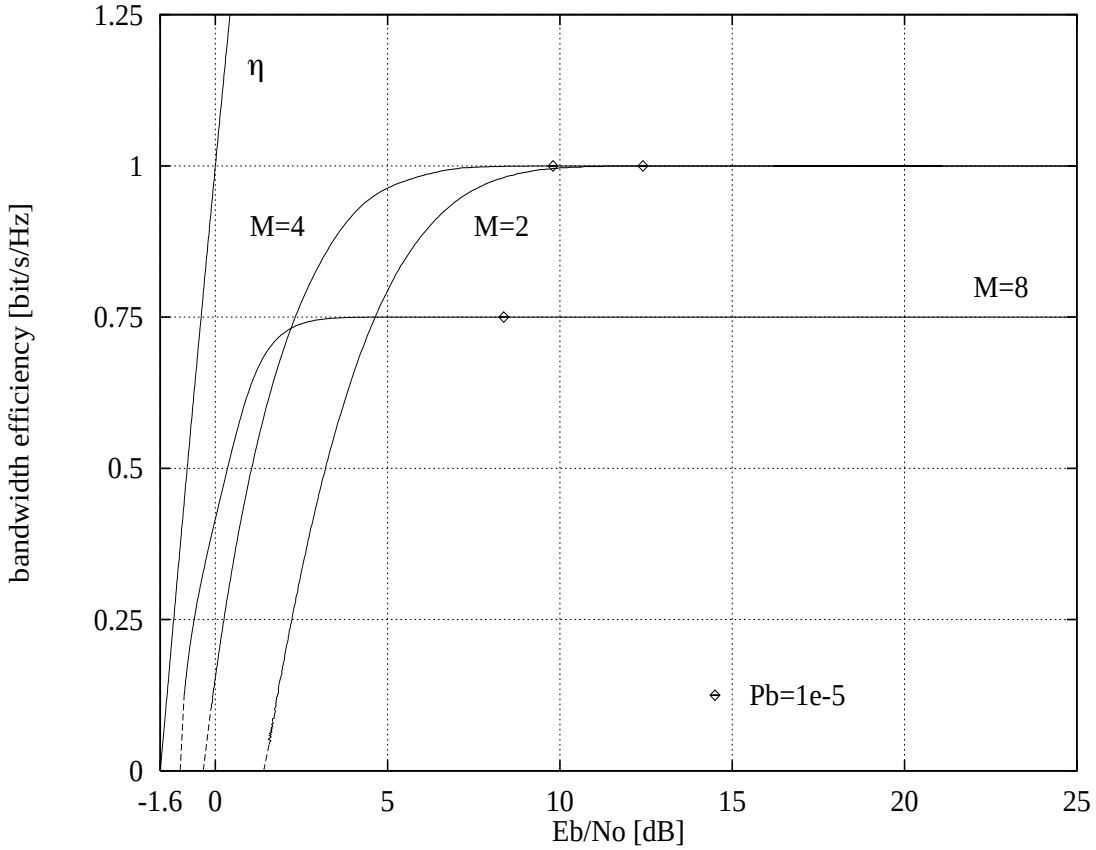


Figure 3.12 Bandwidth efficiency for M–ary orthogonal signalling

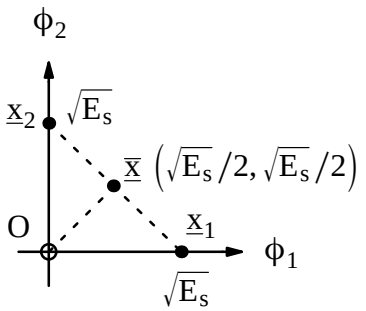


Figure 3.13 Centre of gravity for M=2 orthogonal signalling

Table 3.1 E_b/N_0 for M–ary orthogonal signalling at 0 bit/s/Hz

M	E_b/N_0 [dB]
2	+1.40
4	–0.35
8	–1.02

The slope of M=2 and M=4 efficiency curves are linear to 0 bit/sec/Hz. The slope for M=8 does not follow this characteristic. This result indicates that capacity calculation for M=8 is not accurate. The accuracy of computation is discussed in Section 4.3.5.

As M increases the curves converge closer toward the Shannon bound. This characteristic indicates that as $M \rightarrow \infty$ then η_{DCMC}=0 bit/s/Hz at E_b/N_0=–1.6 dB. This

result implies that as $M \rightarrow \infty$ the capacity for orthogonal signalling reaches the Shannon bound [20, pp. 226–229] but we will have zero bandwidth efficiency.

3.5 L–Orthogonal Signalling

3.5.1 Description

An L–orthogonal signal set is a hybrid form of orthogonal and PSK signalling. An L–orthogonal signal set is comprised of V independent LPSK subsets. The total number of waveforms is M=VL. All waveforms are of equal energy, E_s. The number of dimensions is N=2V. L–orthogonal signalling was first introduced by [21] and later studied by [19, pp. 235–240], [22], and [23]. The capacity data presented in this Section is new.

An L–orthogonal signal can be expressed in the following form [23]

$$x_m(t) = \begin{cases} \sqrt{E_s}\left[\phi_{v,1}(t)\cos\theta_l + \phi_{v,2}(t)\sin\theta_l\right], & 0 \le t \le T \\ 0, & \text{elsewhere} \end{cases}$$
$$, \; v = 1, 2, \ldots, V$$
$$, \; l = 1, 2, \ldots, L$$
$$, \; m = (v-1)L + l \qquad (3.12)$$

where $\theta_l = 2(l-1)\pi/L$. The signal constellation for V=2, L=8 is shown in Figure 3.14 . The signal constellation for L–orthogonal signalling is interpreted in the same manner as for M–ary orthogonal signalling.

L–orthogonal signalling allows for a trade–off between the power efficiency of orthogonal signal sets and bandwidth efficiency of PSK signal sets. In recent times there has been a renewed interest in signal sets which can be shown to be L–orthogonal. An easily implemented scheme called Differentially Coherent Block Coding (DCBC) [6] creates independent LPSK subsets using binary waveforms generated by a Haddamard

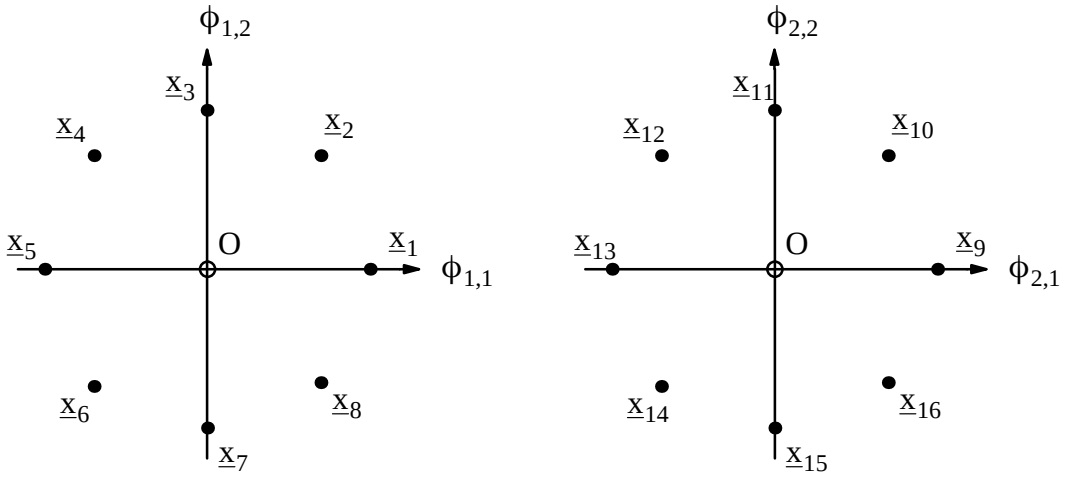


Figure 3.14 V=2, L=8 L–orthogonal signalling constellation

code. Another scheme, Frequency/Differential Phase Shift Keying (2FSK/MDPSK) [24], has two independent differential MPSK subsets created using binary FSK.

3.5.2 Channel Capacity

The channel capacity of V=2 L–orthogonal signalling is shown in Figure 3.15 . All error rate data is calculated in Appendix C5.

At a transmission rate of 3 bit/sym a coding gain of 11.7 dB is possible by doubling the number of phases from L=4 to L=8. Practically no additional coding gain results from increasing the number of phases beyond L=8. The V=2 L–orthogonal signalling set converges to a characteristic similar to the PSK limit.

The bandwidth efficiency of V=2 L–orthogonal signalling is shown in Figure 3.16 . Here, the bandwidth efficiency characteristic converges to the Shannon bound because $\underline{x} = 0$. V=2 L–orthogonal signalling in this instance appears to be more bandwidth efficient compared to M=4 orthogonal signalling for the same number of dimensions, N=4.

The channel capacity and bandwidth efficiency of V=4 L–orthogonal signalling are shown in Figures 3.17 and 3.18 , respectively. Here, we have attempted N=8 integration. The accuracy of computation is no longer reliable as the capacity data exceeds

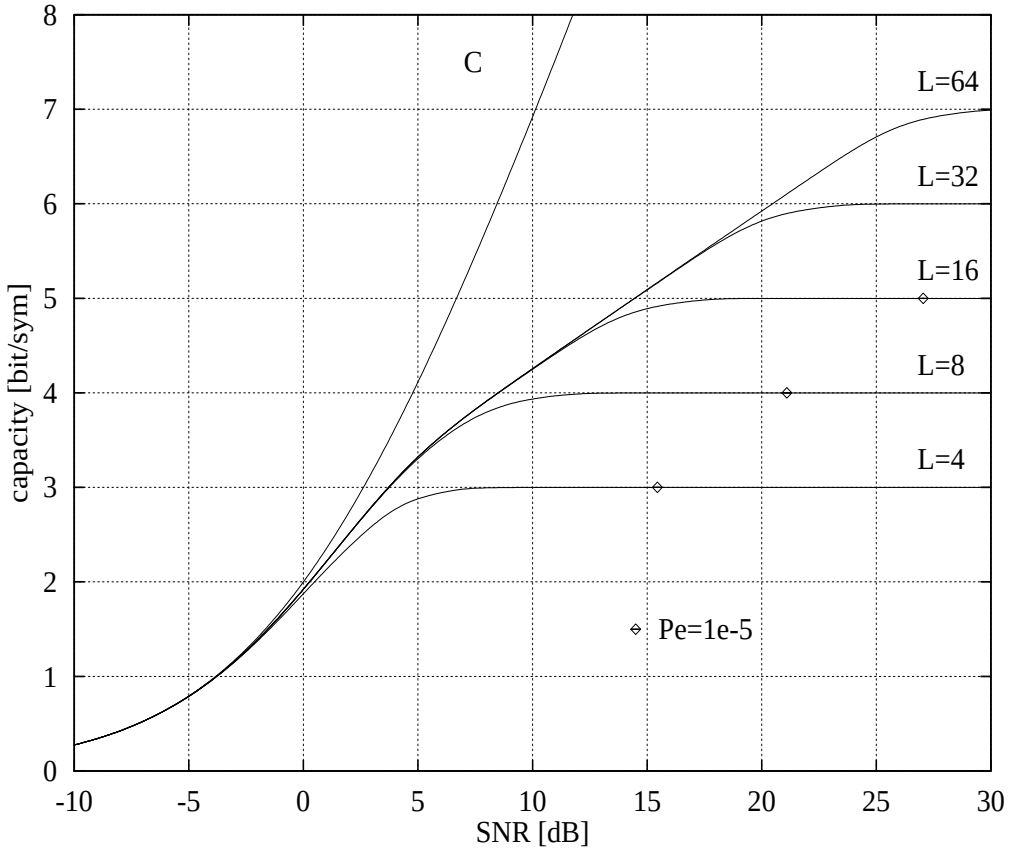


Figure 3.15 Channel capacity for V=2 L–orthogonal signalling

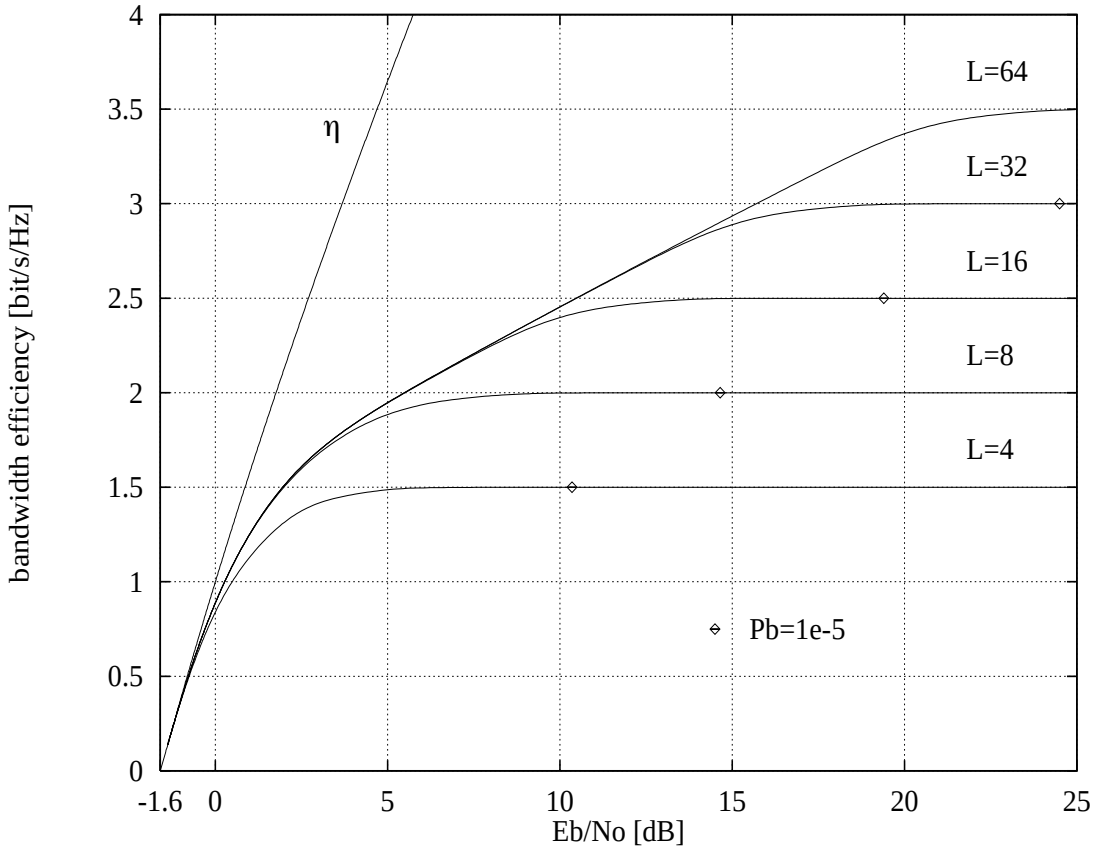


Figure 3.16 Bandwidth efficiency for V=2 L–orthogonal signalling

the Shannon bound at low SNR. The accuracy of computation is discussed in Section 4.3.5.

3.6 Conclusion

This Chapter has presented the channel capacity of well known M–ary signal sets and a hybrid scheme. The signal constellation for each signal set was presented first followed by plots of channel capacity versus SNR and bandwidth efficiency versus E_b/N_0.

The capacity of signal sets for $N \leq 2$ agree with earlier literature [4] and [7, pp. 272–279]. Computation of capacity for signal sets with $N \geq 4$ was attempted. We found that numerical accuracy was reliable only for N=4. Section 4.3.5 describes the cause for the decrease in numerical accuracy for N>4.

The capacity for PAM, QAM, and L–orthogonal signal sets converge toward the Shannon bound at low SNR. This characteristic was not the same for M–ary orthogonal signalling since $\underline{\bar{x}} \neq 0$. As M increases the capacity converges closer to the Shannon bound at 0 bit/s/Hz. The capacity for orthogonal signalling reaches the Shannon bound as $M \rightarrow \infty$ but the bandwidth efficiency becomes zero.

The potential coding gain obtained by signal set expansion was examined. In all cases the doubling of the number of signals yields most of the coding gain. This general

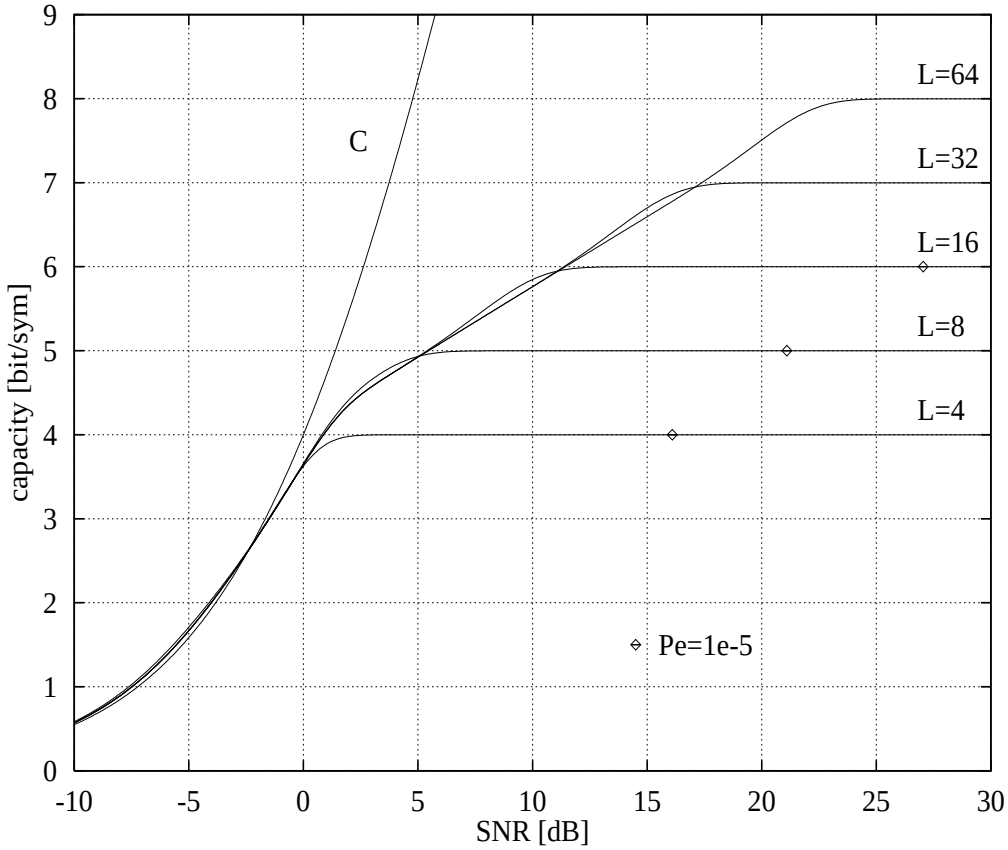


Figure 3.17 Channel capacity for V=4 L–orthogonal signalling

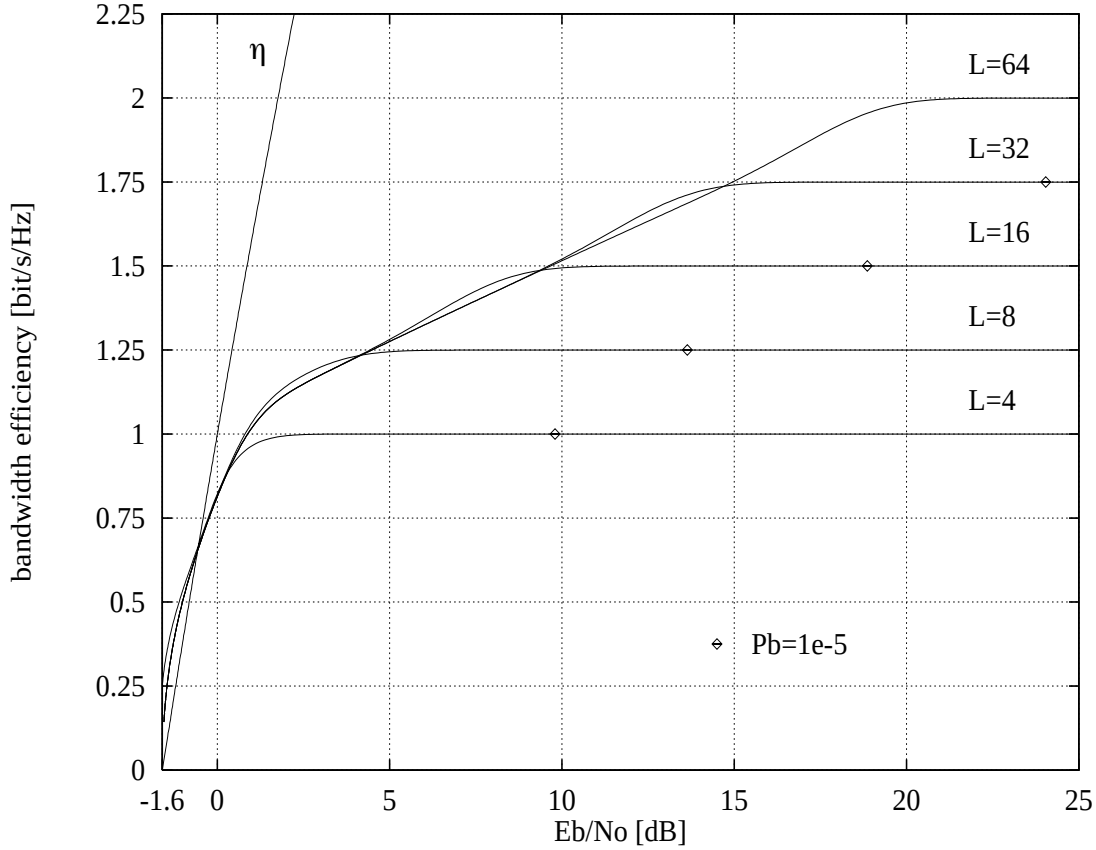


Figure 3.18 Bandwidth efficiency for V=4 L–orthogonal signalling

characteristic has been suggested in [4] and [5]. The highest coding gain for N=2 signal sets is obtained by doubling MPSK or MQAM to 2MQAM.

4 MULTIPLE INTEGRATION TECHNIQUES FOR CHANNEL CAPACITY CALCULATION

4.1 Introduction

The calculation of C_{DCMC} involves integration in N dimensions. Fortunately, the integral is identified as belonging to standard Gauss–Hermite quadrature. The technique for multiple integration is based on Cartesian products of Gauss–Hermite quadrature.

Initial attempts to simplify the multiple integral are given in Section 4.2. Section 4.3 discusses the multiple integration technique. Alternate integration methods requiring less computation are reviewed in Section 4.4.

4.2 Initial Strategies for Multiple Integration

The multiple integration problem involves the following term from (2.31)

$$\underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{\text{N–fold}} \exp\left(-|\underline{t}|^2\right) \log_2\left[\sum_{i=1}^{M} \exp\left(-2\underline{t}\cdot\underline{d}_i - |\underline{d}_i|^2\right)\right] d\underline{t} \tag{4.1}$$

where $\underline{t} = (t_1, t_2, \ldots, t_N)$. Changing of variables to N dimensional spherical coordinates [25, pp. 9–15] does not simplify the integral. The presence of a logarithm of a sum in (4.1) makes any reduction in dimensionality very difficult. The integral could be expressed as a sum of integrals of dimensionality $< N$ by writing a Taylor series expansion of the logarithm term [26, pp. 26–27]. However, the expansion is difficult to obtain because high order derivatives of the logarithm term are required and further complicated by the sum of exponential terms.

There is little alternative but to perform N–dimensional integration. Let the logarithm term of (4.1) be denoted

$$f(\underline{t}) = \log_2\left[\sum_{i=1}^{M} \exp\left(-2\underline{t}\cdot\underline{d}_i - |\underline{d}_i|^2\right)\right] . \tag{4.2}$$

We now proceed to evaluate the multiple integral

$$\underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{N-\text{fold}} \exp\left(-|\underline{t}|^2\right) f(\underline{t})\, d\underline{t} . \tag{4.3}$$

4.3 Approximation to Multiple Integration

4.3.1 Introduction

The integral in (4.3) can be approximated using products of the Gauss–Hermite quadrature. Section 4.3.2 presents the Gauss–Hermite quadrature for one dimension. The N–dimensional quadrature is presented in Section 4.3.3. The number of integration points are given in Section 4.3.4. The accuracy of this integration technique is discussed in Section 4.3.5.

4.3.2 Gauss–Hermite Quadrature

Gauss–Hermite quadrature is an approximation to integration over the region, $[-\infty, +\infty]$, with weight function, $\exp\left(-t^2\right)$,

$$\int_{-\infty}^{\infty} \exp\left(-t^2\right) f(t)\,dt \simeq \sum_{p=1}^{P} w_p f\left(t_p\right) \tag{4.4}$$

for some continuous smooth function, f(t). Tables of weights, $\{w_p\}$, and points, $\{t_p\}$, for different numbers of integration points, P, can be found in [27, p. 294], [28, pp. 343–346], and [29, pp. 359–366]. A quadrature formula is also referred to as a quadrature *rule*.

The accuracy of a quadrature formula can be described by stating the *degree of precision*, *d* [29, p. 1], [9, p. 3]. A quadrature rule has degree *d* if it is exact for f(t) being any polynomial of degree $\leq d$ and there is at least one polynomial of degree $d + 1$ for which it is not exact. Gauss–Hermite quadrature has the highest degree of precision [28, p. 100] with

$$d = 2P–1 \ . \tag{4.5}$$

4.3.3 Cartesian Product Formula

Quadrature formulas for N–dimensional integration can be constructed using products of rules for $<$ N dimensions and are classified as *product formulas* [9, p. 23]. We will construct our product formula by using N products of Gauss–Hermite quadrature. The resulting product formula is classified a *Cartesian product formula*.

Integration over N dimensions is written using products of (4.4) [9, p. 28]

$$\underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{\mathrm{N-fold}} \exp\left(-|\underline{t}|^2\right) f(\underline{t})\, d\underline{t} \simeq \sum_{p_1=1}^{P} w_{p_1} \sum_{p_2=1}^{P} w_{p_2} \cdots \sum_{p_N=1}^{P} w_{p_N} f\left(t_{p_1}, t_{p_2}, \ldots, t_{p_N}\right) . \quad (4.6)$$

The degree of precision is $d = 2P-1$ [9, pp. 26–27]. The total number of integration points is P^N.

The multiple integration formula (4.6) has been implemented with double precision ANSI C programs. The integration formula is a nested sum of identical quadrature rules and a single C function is written to perform integration for any N and P. The C function is a recursive function [30, pp. 86–88] and is described in Appendix E. The product Gauss–Hermite rule was implemented on a modern workstation and run times were small for N=1,2, and 4 but N=8 requires a very long run time.

4.3.4 Number of Integration Points

The number of integration points for N=1 to N=8 are listed in Table 4.1 . The degree of precision for each case is $d \geq 9$. The value of P was chosen to satisfy a high degree of precision. The number of points were also limited to less than one million to minimise computing time.

The value of P should be at least 5 to ensure reliable integration [31]. The degree of N=4 integration could have been increased using P=10.

Table 4.1 Number of integration points

N	P	P^N	degree d
1	10	10	19
2	10	100	19
4	5	625	9
8	5	390,625	9

4.3.5 Decimal Place Accuracy and Error Estimates

It is a well known and accepted fact that repeated Gaussian quadrature is the most accurate technique for approximating multidimensional integration [31]. This Section discusses the decimal place accuracy of this technique and reviews methods of bounding the error in computation.

The decimal place accuracy was investigated by approximating the following integral with a known solution

$$\underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{N-\text{fold}} \exp\left(-|\underline{x}|^2\right)\left[1 + \sum_{i=1}^{N} \frac{x_i^2}{\sqrt{\pi}}\right] d\underline{x} = \left(\sqrt{\pi}\right)^N \left[1 + \frac{N}{2\sqrt{\pi}}\right]. \tag{4.7}$$

Table F.1 lists the decimal place accuracy for different N and P. The worst case accuracy is seven significant decimals. The minimum order of accuracy is said to be three significant decimal places for practical calculations [32] and so this technique should be acceptable.

The decimal place accuracy for numerical integration when $f(\underline{t})$ is given by (4.2) cannot be stated since the required degree for quadrature is not known. Multiple integration was performed for a large number of points to ensure acceptable accuracy [29, pp. 38–39].

Several methods for predicting the error of the current technique were investigated. The common method of obtaining error estimates for a quadrature rule requires computing derivatives of the function, $f(\underline{t})$. Expressions for error estimates in terms of derivatives are given in [29, pp. 72–76] but it is very difficult to generate high order derivatives of (4.2). Another method for obtaining error estimates without evaluating derivatives was found in [33] but was too difficult to attempt.

The results of Chapter 3 showed that capacity calculation was accurate for $N \leq 4$. We would have expected accurate calculations for N=8 because we used the same high degree of precision as for N=4. The errors occurring with N=8 results may be due to the increased number of summations performed when calculating C_{DCMC}.

The total number of summations in calculating C_{DCMC} can be obtained by observing (2.31) and (4.6). The total number of summations is M^2P^N. Tables G.1 to G.4 list the number of summations required to calculate C_{DCMC}.

Accurate results for $N \leq 4$ signal sets were obtained when the number of summations did not exceed 10.24×10^6. N=8 signal sets involve 25×10^6 and higher. The round–off error within the computing hardware becomes significant for such lengthy calculations.

4.4 Alternate Rules

Alternate rules can offer a large reduction in the total number of integration points. For $N \leq 4$, product Gaussian rules are superior to other rules in terms of the degree *d* and an economical number of integration points. This Section reviews a search for economical rules for $N \leq 8$.

Before discussing alternate rules it is helpful to give some background into the development of quadrature formulae. With this knowledge it is easier to describe relative merits and to classify different types of formulae [9, pp. 12–14], [34].

The two families of well known one dimensional formulae are the Newton–Cotes and Gaussian quadrature [35, pp. 245–261]. The Newton–Cotes rules include the trapezoidal, Simpson's 1/3, and Simpson's 3/8 rules for a set of evenly spaced integration points. The set of P weights are determined using the value of the function at the corresponding integration point. The degree of precision is $d = \text{P}-1$.

A Gaussian rule solves for 2P unknowns, the set of weights and integration points. A set of non–linear equations are required to be solved. This set of equations is difficult to solve but the resulting precision is $d = 2\text{P}-1$. For certain regions and weight functions the theory of orthogonal polynomials are used to obtain the set of weights and integration points [29, pp. 1–13]. We use Hermite polynomials in our case.

It is extremely difficult to apply the Gaussian method for developing rules for N dimensions. The resulting set of non–linear equations in N variables is very difficult to solve and the theory of orthogonal polynomials in N variables is complicated [9, p. 7], [36]. For small N a product rule is easily constructed and economical. It is recommended that product rules be used whenever possible.

The integration region and weight function of (4.3) are classified as *fully symmetric* (FS). A loose definition for FS is that both the region and weight function are symmetric about the origin. A quadrature rule is classified FS if it has weights and points symmetric about the origin. A quadrature formula is classified *good* (G) if the integration points lie inside the integration region and all the weights are positive. A combination of positive and negative weights may affect convergence. Another classification is a *minimal point rule* (M). An M rule is the only rule of the same degree which exists for the smallest number of integration points.

The product Gauss–Hermite rules are FSG and in some cases FSGM for (4.3). From Section 4.3.4 the product Gauss–Hermite rules require an exponential increase in the number of points as a function of dimensionality. The work of [36–38] concentrated on developing economical FS formulas for $\text{N} \leq 20$. The work in [34] constructs FS and FSM formulas for N=2 and N=3 using a computer based systematic approach. Additional economical FS formulae for $\text{N} \leq 10$ are constructed in [39]. Table 4.2 compares the number of points for N=8 quadrature rules of degree 9 for (4.3).

Table 4.2 Degree 9 rules for N=8

Quadrature Rule	Number of Points	Classification
Product Gauss–Hermite	390,625	FSG
Products of N=2 rule [40]	160,000	FSG
[38]	2,561	FS
[39]	2,497	FS

Table 4.2 shows a large reduction in the number of integration points for N=8 integration. Tables of quadrature data for the method by [38] have been obtained but little

work has been carried out. The alternative rules offer faster run times for N=8 computations. The decimal place accuracy would need to be investigated before using these alternate rules.

A product rule for N=16 with $d \geq 9$ is not practical to construct. The product Gauss–Hermite or products of alternate rules require over a million integration points. Some methods for integration for a large number of dimensions can be found in [26, pp. 415–417] and [32] but decimal place accuracy becomes difficult to obtain.

4.5 Conclusion

A numerical integration technique for calculating C_{DCMC} has been presented. By using products of Gauss–Hermite quadrature we obtain a high degree of precision for $N \leq 4$. This technique is implemented as a recursive C function for any N and P.

The decimal place accuracy of repeated Gauss–Hermite quadrature for $N \leq 8$ was investigated. Results for a test function indicate accuracy of at least seven significant decimal places. Several methods exist for predicting the accuracy of repeated quadrature. However, these methods could not be used to predict decimal place accuracy when computing C_{DCMC}.

The decimal place accuracy for C_{DCMC} was found to be reliable for $N \leq 4$ from Chapter 3. We investigated the total number of summations performed when computing C_{DCMC}. The number of summations is a function of both M and N. For N=8 and $M \geq 8$ the number of summations is too large and computer round–off error becomes significant.

A search for economical quadrature rules for $N \leq 8$ was carried out. A large reduction in the number of integration points is possible for the same degree of precision. However, the decimal place accuracy of these alternate rules has yet to be determined. The calculation of C_{DCMC} for N=8 may become more accurate if the number of summations can be reduced using these alternate rules. Numerical integration with high precision for N>8 using a product rule was not found to be practical.

5 CONCLUSION

This thesis has presented a technique for calculating channel capacity of coherently detected M–ary digital modulation over an AWGN channel. A general channel capacity formula has been developed for N–dimensional signal sets. The capacity formula requires the signal constellation and noise variance for computation.

The channel capacity formula involves integration in N dimensions. A numerical integration technique based on repeated Gaussian quadrature allows accurate capacity calculation for $N \leq 4$.

Chapter 1 introduced the general uncoded M–ary waveform communications system. We were interested in finding capacity for signal sets with large dimensionality, N. A literature search yielded only two publications for $N \leq 2$ [4, 7]. An additional publication outlined capacity calculation for large N but for incoherently detected modulation [5]. The numerical integration technique used in these publications is Monte Carlo simulation. This thesis presents an alternative integration method suited to the type of integral encountered in the capacity formula.

Chapter 2 presented the channel capacity for certain bandlimited input AWGN channels. The capacity for the DMC was reviewed first. We then described two variations of the DMC to model the waveform channel from Chapter 1. These two channels are the continuous channel and DCMC. These two channels were then extended to N–dimensional vector channels. By restricting our analysis to finite–energy bandlimited waveforms of equal symbol period, T, we can represent these waveforms as N–dimensional vectors.

We developed channel capacity formulas for the continuous channel and DCMC. The Shannon bound is obtained using the continuous channel and the capacity for M–ary signal sets is obtained using the DCMC. The conversion of capacity to bandwidth efficiency was also given. A bound on C_{DCMC} which does not require N–dimensional integration should be investigated to enable capacity analysis of signal sets with large N.

Chapter 3 presented channel capacity and bandwidth efficiency for PAM, QAM, PSK, M–ary orthogonal signalling, and L–orthogonal signalling. For each signal set the signal constellation was described first followed by plots of capacity and bandwidth efficiency. The capacity data for L–orthogonal signalling is new.

The capacity for $N \leq 2$ signal sets (PAM, QAM, and PSK) agrees with earlier literature [4], [7, pp. 272–279]. The coding gain obtained by signal set expansion was also examined. In all cases the doubling of the number of signals yields most of the coding gain as suggested by earlier literature [4, 5].

We attempted capacity calculation for $N \geq 4$ signal sets. Accurate results could only be obtained for N=4.

The bandwidth efficiency for M–ary orthogonal signalling did not converge with the Shannon bound at low E_b/N_0. This characteristic was due to orthogonal signal sets having a non–zero geometric mean which conveys no information. As $M \rightarrow \infty$ the capacity of orthogonal signalling reaches the Shannon bound but we have zero bandwidth efficiency.

Chapter 4 presented the technique for integration in N–dimensions. The numerical integration is based on repeated Gauss–Hermite quadrature. The integration is implemented as a recursive C function.

The decimal place accuracy for repeated quadrature was found to be seven significant decimal places by testing with a function with a known solution. However, the number of summations performed when computing C_{DCMC} for N=8 becomes too large and computer round–off error occurs.

A search for economical rules has been carried out and a large reduction in the number of points is possible for the same degree of precision. These alternate rules would need to be tested for decimal place accuracy before being used. These alternate rules may enable accurate capacity calculation for $N \geq 8$.

REFERENCES

[1] C. E. Shanon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.* , vol. 27, pp. 379–423, July 1948.

[2] C. E. Shanon, "A Mathematical Theory of Communication (Concluded from July 1948 issue)," *Bell Syst. Tech. J.* , vol. 27, pp. 623–656, Oct. 1948.

[3] C. E. Shannon, "Communication in the Presence of Noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.

[4] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals," *IEEE Trans. Inform. Theory*, vol. IT–28, pp. 56–67, Jan. 1982.

[5] G. Garrabrant and J. Ehrenberg, "Trellis Coded Modulation Applied to Noncoherent Detection of Orthogonal Signals," *MILCOM '89*, vol. 3, pp. 774–778, Oct. 1989.

[6] S. A. Rhodes, "Differentially Coherent FEC Block Coding," *COMSAT Tech. Rev.*, vol. 17, pp. 283–310, Fall 1987.

[7] R. E. Blahut, *Principles and Practice of Information Theory.* Reading, MA, Addison–Wesley, 1987.

[8] M. Campanella and G. Mamola, "On the Channel Capacity for Constant Envelope Signals with Effective Bandwidth Constraint," *IEEE Trans. Commun.*, vol. 38, pp. 1164–1172, Aug. 1990.

[9] A. H. Stroud, *Approximate Calculation of Multiple Integrals*. Englewood Cliffs, NJ, Prentice–Hall, 1971.

[10] R. G. Gallager, *Information Theory and Reliable Communication*. New York, Wiley, 1968.

[11] R. M. Fano, *Transmission of Information a Statistical Theory of Communication*. Cambridge, MA, M.I.T. Press, 1961.

[12] J. G. Proakis, *Digital Communications Second Edition.* New York, McGraw–Hill, 1989.

[13] M. Kanefsky, *Communication Techniques for Digital and Analog Signals*. New York, Wiley, 1987.

[14] J. M. Wozencraft and I. M. Jacobs, *Principles of Communications Engineering*. New York, Wiley, 1965.

[15] A. D. Wyner, "The Capacity of the Band–Limited Gaussian Channel," *Bell Syst. Tech. J.* , vol. 45, pp. 359–395, March 1966.

[16] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. Tokyo, McGraw–Hill Kogashuka, 1979.

[17] C. H. Edwards and D. E. Penny, *Calculus and Analytic Geometry*. Englewood Cliffs, NJ, Prentice–Hall, 1982.

[18] E. Arthurs and H. Dym, "On the Optimum Detection of Digital Signals in the Presence of White Gaussian Noise – A Geometric Interpretation and a Study of Three Basic Data Transmission Systems," *IRE Trans. Commun. Syst.*, vol. CS–8, pp. 336–372, Dec. 1962.

[19] W. C. Lindsey and M. K. Simon, *Telecommunications System Engineering*. Englewood Cliffs, NJ, Prentice–Hall, 1973.

[20] A. J. Viterbi, *Principles of Coherent Communications*. New York, McGraw–Hill, 1966.

[21] I. S. Reed and R. A. Scholtz, "N–Orthogonal Phase Modulated Codes," *IEEE Trans. on Inform. Theory*, vol. IT–12, pp. 388–395, July 1966.

[22] A. J. Viterbi and J. J. Stiffler, "Performance of N–Orthogonal Codes," *IEEE Trans. on Inform. Theory*, vol. IT–13, pp. 521–522, July 1967.

[23] W. C. Lindsey and M. K. Simon, "L–Orthogonal Signal Transmission and Detection," *IEEE Trans. Commun.*, vol. COM–20, pp. 953–960, Oct. 1972.

[24] L. Wei and I. Korn, "Frequency Differential Phase Shift Keying in the Satellite Mobile Channel with a Bandlimited Filter," *SICON/ICIE '93*, pp. 221–225, Sept. 1993.

[25] K. S. Miller, *Multidimensional Gaussian Distributions*. New York, Wiley, 1964.

[26] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration Second Edition*. New York, Academic Press, 1984.

[27] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. New York, Dover, 1972.

[28] V. I. Krylov, *Approximate Calculation of Integrals*. New York, MacMillan, 1962.

[29] A. H. Stroud and D. Secrest, *Gaussian Quadrature Formulas*. Englewood Cliffs, NJ, Prentice–Hall, 1966.

[30] B. W. Kernighan and D. M. Ritchie, *The C Programming Language Second Edition*. Englewood Cliffs, NJ, Prentice–Hall, 1988.

[31] I. H. Sloan, private communications, Nov. 1993.

[32] T. Tsuda, "Numerical Integration of Functions of Very Many Variables," *Numerische Mathematik*, vol. 20, pp. 377–391, 1973.

[33] F. Stenger, "Error Bounds for the Evaluation of Integrals by Repeated Gauss–Type Formulae," *Numerische Mathematik*, vol. 9, pp. 200–213, 1966.

[34] F. Mantel and P. Rabinowitz, "The Application of Integer Programming to the Computation of Fully Symmetric Integration Formulas in Two and Three Dimensions," *SIAM J. Numer. Anal.*, vol. 14, no. 3, pp. 391–424, June 1977.

[35] C. F. Gerald and P. O. Wheatly, *Applied Numerical Analysis*. Reading, MA, Addison–Wesley, 1984.

[36] J. McNamee and F. Stenger, "Construction of Fully Symmetric Numerical Integration Formulas," *Numerische Mathematik*, vol. 10, pp. 327–344, 1967.

[37] F. Stenger, "Numerical Integration in *n* Dimensions," Master's Thesis, Dept. of Math., Uni. of Alberta, Edmonton, 1963.

[38] F. Stenger, "Tabulation of Certain Fully Symmetric Numerical Integration Formulas of Degree 7, 9, and 11," *Mathematics of Computation*, vol. 25, no. 116, p. 953 and Microfiche, Oct. 1971.

[39] P. Keast, "Some Fully Symmetric Quadrature Formulae for Product Spaces," *J. Inst. Maths. Applics.*, vol. 23, pp. 251–264, 1979.

[40] P. Rabinowitz and N. Richter, "Perfectly Symmetric Two–Dimensional Integration Formulas with Minimal Numbers of Points," *Mathematics of Computation*, vol. 23, pp. 765–779, 1969.

[41] B. P. Lathi, *Modern Digital and Analog Communication Systems*. New York, Holt, Rinehart and Winston, 1983.

APPENDIX A: SHANNON BOUND FOR BANDLIMITED AWGN CHANNEL

The following derivation of the Shannon bound is a confirmation of the results of [12, pp. 134–135] using formulas from [10, p. 32]. All logarithms are to base 2.

The channel capacity for a continuous input continuous output memoryless channel, C_{CCMC} , is written

$$C_{CCMC} = \max_{p(\underline{x})} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{\text{2N–fold}} p(\underline{y}|\underline{x}) p(\underline{x}) \log\left[\frac{p(\underline{y}|\underline{x})}{p(\underline{y})}\right] d\underline{x} d\underline{y} \text{ [bit/sym]}. \tag{A1}$$

The Shannon bound, C, is obtained for an AWGN channel with a white noise input. Assume that the input is a white Gaussian noise source [10, p. 32] with zero mean and variance, σ^2, described by

$$p(\underline{x}) = \prod_{n=1}^{N} p(x_n) \tag{A2}$$

where

$$p(x_n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-x_n^2}{2\sigma^2}\right). \tag{A3}$$

Assume AWGN with zero mean and variance $N_0/2$. The channel statistic is represented by

$$p(\underline{y}|\underline{x}) = \prod_{n=1}^{N} p(y_n|x_n) \tag{A4}$$

where

$$p(y_n|x_n) = \frac{1}{\sqrt{\pi N_0}} \exp\left(\frac{-(y_n - x_n)^2}{N_0}\right). \tag{A5}$$

The channel output random variable, $\underline{y}$, is the sum of two Gaussian random variables [10, pp. 509–510]. The mean is zero and variance $N_0/2 + \sigma^2$ hence

$$p(\underline{y}) = \prod_{n=1}^{N} p(y_n) \tag{A6}$$

where

$$p(y_n) = \frac{1}{\sqrt{\pi(N_0 + 2\sigma^2)}} \exp\left(\frac{-y_n^2}{N_0 + 2\sigma^2}\right). \tag{A7}$$

We write C

$$C = \underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{2N\text{–fold}} p(\underline{x}, \underline{y}) \log\left[p(\underline{y}|\underline{x})\right] d\underline{x} d\underline{y} - \underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{2N-\text{fold}} p(\underline{x}, \underline{y}) \log\left[p(\underline{y})\right] d\underline{x} d\underline{y}$$

$$= I_1 - I_2. \tag{A8}$$

We now solve I_1.

$$I_1 = \underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{2N-\text{fold}} p(\underline{x}, \underline{y}) \log\left[p(\underline{y}|\underline{x})\right] d\underline{x} d\underline{y}$$

$$= \underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{2N-\text{fold}} p(\underline{x}, \underline{y}) \log\left[\prod_{n=1}^{N} p(y_n|x_n)\right] d\underline{x} d\underline{y}$$

$$= \underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{2N-\text{fold}} p(\underline{x}, \underline{y}) \left[\sum_{n=1}^{N} \log\left[p(y_n|x_n)\right]\right] d\underline{x} d\underline{y}$$

$$= \sum_{n=1}^{N} \underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{2N\text{–fold}} p(\underline{x}, \underline{y}) \log\left[p(y_n|x_n)\right] d\underline{x} d\underline{y} \tag{A9}$$

This is a sum of N 2N–fold integrals:

$$I_1 = \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1, y_1) \log\left[p(y_1|x_1)\right] dx_1 dy_1 \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_2, y_2) dx_2 dy_2 \right.$$

$$\left. \cdots \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_N, y_N) dx_N dy_N \right\}$$

$$+ \cdots +$$

$$\left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_N, y_N) \log\left[p(y_N|x_N)\right] dx_N dy_N \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1, y_1) dx_1 dy_1 \right.$$

$$\left. \cdots \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_{N-1}, y_{N-1}) dx_{N-1} dy_{N-1} \right\}$$

$$= N \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y)\log\left[p(y|x)\right]dxdy$$

$$= N \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y)\log\left\{\frac{1}{\sqrt{\pi N_0}}\exp\left[\frac{-(y-x)^2}{N_0}\right]\right\}dxdy$$

$$= N \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y)\left[\log\left\{\frac{1}{\sqrt{\pi N_0}}\right\} - \frac{(y-x)^2 \log(e)}{N_0}\right]dxdy$$

$$= N \log\left\{\frac{1}{\sqrt{\pi N_0}}\right\} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y)dxdy$$

$$-\frac{N \log(e)}{N_0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y)(y-x)^2 dxdy$$

$$= N \log\left\{\frac{1}{\sqrt{\pi N_0}}\right\} - \frac{N \log(e)}{N_0} \int_{-\infty}^{\infty} p(x) \int_{-\infty}^{\infty} p(y|x)(y-x)^2 dydx \ .$$

(A10)

The term, $\int_{-\infty}^{\infty} p(y|x)(y-x)^2 dy$, is the variance of the pdf, $p(y|x)$, which is $N_0/2$ [10, p. 32].

$$I_1 = N \log\left\{\frac{1}{\sqrt{\pi N_0}}\right\} - \frac{N \log(e)}{N_0} \int_{-\infty}^{\infty} \frac{N_0}{2} p(x)dx$$

$$= -\frac{N}{2}\log(\pi N_0) - \frac{N \log(e)}{2}$$

$$= -\frac{N}{2}\log(e\pi N_0) \ . \qquad \text{(A11)}$$

Similarly:

$$I_2 = \underbrace{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty}}_{N-\text{fold}} p(\underline{y})\log\left[p(\underline{y})\right]d\underline{y}$$

$$= N \int_{-\infty}^{\infty} p(y)\log[p(y)]dy$$

$$= N \int_{-\infty}^{\infty} p(y)\log\left[\frac{1}{\sqrt{\pi\left(2\sigma^2 + N_0\right)}}\exp\left[\frac{-y^2}{\left(2\sigma^2 + N_0\right)}\right]\right]dy$$

$$= N \int_{-\infty}^{\infty} p(y)\log\left[\frac{1}{\sqrt{\pi\left(2\sigma^2 + N_0\right)}}\right]dy + N \int_{-\infty}^{\infty} p(y)\log\left[\exp\left[\frac{-y^2}{\left(2\sigma^2 + N_0\right)}\right]\right]dy$$

$$= N \log\left[\frac{1}{\sqrt{\pi\left(2\sigma^2 + N_0\right)}}\right]\int_{-\infty}^{\infty} p(y)dy + \frac{N\log(e)}{\left(2\sigma^2 + N_0\right)}\int_{-\infty}^{\infty} p(y)-y^2dy$$

$$= N \log\left[\frac{1}{\sqrt{\pi\left(2\sigma^2 + N_0\right)}}\right] - \frac{N\log(e)}{\left(2\sigma^2 + N_0\right)}\int_{-\infty}^{\infty} p(y)y^2dy$$

The term, $\int_{-\infty}^{\infty} y^2 p(y)dy$, is the variance of p(y).

$$I_2 = -\frac{N}{2}\log\left[\pi\left(2\sigma^2 + N_0\right)\right] - \frac{N\log(e)}{\left(2\sigma^2 + N_0\right)}\frac{\left(2\sigma^2 + N_0\right)}{2}$$

$$= -\frac{N}{2}\log\left[\pi\left(2\sigma^2 + N_0\right)\right] - \frac{N\log(e)}{2}$$

$$= -\frac{N}{2}\log\left[e\pi\left(2\sigma^2 + N_0\right)\right] .$$

(A12)

Thus channel capacity can be written as

$$C = I_1 - I_2$$

$$= -\frac{N}{2}\log\left(e\pi N_0\right) + \frac{N}{2}\log\left[e\pi\left(2\sigma^2 + N_0\right)\right]$$

$$= \frac{N}{2}\log\left(1 + \frac{2\sigma^2}{N_0}\right) . \qquad \text{(A13)}$$

We define the average transmitter power, P, as

$$P = \frac{1}{T}\sum_{n=1}^{N} E\left\{x_n^2\right\}$$

$$= \frac{N\sigma^2}{T} . \qquad \text{(A14)}$$

From Section 2.3.4

$$N = 2TW \qquad \text{(A15)}$$

where W is the channel bandwidth and T is the signalling period.

Substituting (A15) into (A14)

$$2\sigma^2 = \frac{P}{W} \ . \tag{A16}$$

Substituting (A15) and (A16) into (A13)

$$C = WT \log\left(1 + \frac{P}{WN_0}\right) \ . \tag{A17}$$

We define the signal to noise power ratio (SNR) as

$$SNR = \frac{P}{WN_0} \ . \tag{A18}$$

Thus, the Shannon bound is

$$C = WT \log_2(1 + SNR) \ [\text{bit/sym}] \ . \tag{A19}$$

APPENDIX B: CAPACITY FOR DCMC

The derivation of the capacity for a discrete input continuous output memoryless channel, C_{DCMC}, is presented. This derivation is based on the work of [4], [5], [7, pp. 272–276], [8], [10, p. 32], and [14, pp. 293–323]. All logarithms are to base 2.

The channel capacity for the DCMC is written

$$C_{DCMC} = \max_{p(\underline{x}_1)\ldots p(\underline{x}_M)} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{\text{N–fold}} p(\underline{y}|\underline{x}_m) p(\underline{x}_m) \log\left[\frac{p(\underline{y}|\underline{x}_m)}{p(\underline{y})}\right] d\underline{y} \text{ [bit/sym]}. \quad \text{(B1)}$$

Using the identity for $p(\underline{y})$

$$p(\underline{y}) = \sum_{i=1}^{M} p(\underline{y}|\underline{x}_i) p(\underline{x}_i) \ , \quad \text{(B2)}$$

then

$$\begin{aligned} C_{DCMC} &= \max_{p(\underline{x}_1)\ldots p(\underline{x}_M)} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{\text{N–fold}} p(\underline{y}|\underline{x}_m) p(\underline{x}_m) \log\left[\frac{p(\underline{y}|\underline{x}_m)}{\sum_{i=1}^{M} p(\underline{x}_i) p(\underline{y}|\underline{x}_i)}\right] d\underline{y} \\ &= \max_{p(\underline{x}_1)\ldots p(\underline{x}_M)} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{\text{N–fold}} p(\underline{y}|\underline{x}_m) p(\underline{x}_m) \log\left[p(\underline{y}|\underline{x}_m)\right] d\underline{y} \\ &\quad - \max_{p(\underline{x}_1)\ldots p(\underline{x}_M)} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{\text{N–fold}} p(\underline{y}|\underline{x}_m) p(\underline{x}_m) \log\left[\sum_{i=1}^{M} p(\underline{x}_i) p(\underline{y}|\underline{x}_i)\right] d\underline{y} \\ &= I_1 - I_2 \ . \end{aligned} \quad \text{(B3)}$$

We assume capacity is obtained for equally likely inputs

$$p(\underline{x}_m) = \frac{1}{M} \ ; \ m = 1, \ldots, M. \quad \text{(B4)}$$

The waveform channel is AWGN with zero mean and variance $N_0/2$. The channel statistic is represented by

$$\begin{aligned} p(\underline{y}|\underline{x}_m) &= \prod_{n=1}^{N} p(y_n|x_{mn}) \\ &= \prod_{n=1}^{N} \frac{1}{\sqrt{\pi N_0}} \exp\left[\frac{-(y_n - x_{mn})^2}{N_0}\right] \end{aligned}$$

$$= \frac{1}{\left(\sqrt{\pi N_0}\right)^N} \exp\left[\sum_{n=1}^{N} \frac{-(y_n - x_{mn})^2}{N_0}\right]$$

$$= \frac{1}{\left(\sqrt{\pi N_0}\right)^N} \exp\left[-\frac{|\underline{y} - \underline{x}_m|^2}{N_0}\right] \tag{B5}$$

We now evaluate I_1.

$$I_1 = \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} p(\underline{y}|\underline{x}_m) p(\underline{x}_m) \log\left[p(\underline{y}|\underline{x}_m)\right] d\underline{y}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} \frac{1}{\left(\sqrt{\pi N_0}\right)^N} \exp\left[-\frac{|\underline{y} - \underline{x}_m|^2}{N_0}\right] \log\left[\frac{1}{\left(\sqrt{\pi N_0}\right)^N} \exp\left[-\frac{|\underline{y} - \underline{x}_m|^2}{N_0}\right]\right] d\underline{y} \tag{B6}$$

Let

$$\frac{y_n - x_{mn}}{\sqrt{N_0}} = t_n \tag{B7}$$

then

$$\frac{|\underline{y} - \underline{x}_m|^2}{N_0} = \frac{1}{N_0} \sum_{n=1}^{N} (y_n - x_{mn})^2$$

$$= \sum_{n=1}^{N} \left[\frac{(y_n - x_{mn})}{\sqrt{N_0}}\right]^2$$

$$= \sum_{n=1}^{N} (t_n)^2$$

$$= |\underline{t}|^2. \tag{B8}$$

Also

$$\frac{dy_n}{\sqrt{N_0}} = dt_n$$

$$d\underline{y} = \left(\sqrt{N_0}\right)^N d\underline{t}. \tag{B9}$$

The limits of integration become

$$\left\{ \underline{y} \right\} \to +\infty \Rightarrow \{ \underline{t} \} \to +\infty$$

$$\left\{ \underline{y} \right\} \to -\infty \Rightarrow \{ \underline{t} \} \to -\infty. \tag{B10}$$

Now I_1 becomes

$$I_1 = \frac{1}{M}\sum_{m=1}^{M}\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{N-\text{fold}}\frac{\left(\sqrt{N_0}\right)^N}{\left(\sqrt{\pi N_0}\right)^N}\exp\left(-|\underline{t}|^2\right)\log\left[\frac{1}{\left(\sqrt{\pi N_0}\right)^N}\exp\left(-|\underline{t}|^2\right)\right]d\underline{t}$$

$$= \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{N-\text{fold}}\frac{1}{\left(\sqrt{\pi}\right)^N}\exp\left(-|\underline{t}|^2\right)\log\left[\frac{1}{\left(\sqrt{\pi N_0}\right)^N}\right]d\underline{t}$$

$$+ \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{N-\text{fold}}\frac{1}{\left(\sqrt{\pi}\right)^N}\exp\left(-|\underline{t}|^2\right)\log\left[\exp\left(-|\underline{t}|^2\right)\right]d\underline{t}\ . \qquad \text{(B11)}$$

Using the identities [10, p. 510]:

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{\pi}}\exp\left(-t^2\right)dt = 1\ ,$$

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{\pi}}t^2\exp\left(-t^2\right)dt = \frac{1}{2} \qquad \text{(B12)}$$

$$I_1 = \log\left[\frac{1}{\left(\sqrt{\pi N_0}\right)^N}\right]\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{N-\text{fold}}\frac{1}{\left(\sqrt{\pi}\right)^N}\exp\left(-|\underline{t}|^2\right)d\underline{t}$$

$$- \log(e)\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{N-\text{fold}}\frac{|\underline{t}|^2}{\left(\sqrt{\pi}\right)^N}\exp\left(-|\underline{t}|^2\right)d\underline{t}$$

$$= \log\left[\frac{1}{\left(\sqrt{\pi N_0}\right)^N}\right]\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{N-\text{fold}}\frac{1}{\left(\sqrt{\pi}\right)^N}\left[\exp\left(-t_1{}^2\right)\cdots\exp\left(-t_N^2\right)\right]d\underline{t}$$

$$- \log(e)\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{N-\text{fold}}\frac{1}{\left(\sqrt{\pi}\right)^N}\left(t_1{}^2+\ldots+t_N{}^2\right)\left[\exp\left(-t_1{}^2\right)\cdots\exp\left(-t_N{}^2\right)\right]d\underline{t}$$

$$= \log\left[\frac{1}{\left(\sqrt{\pi N_0}\right)^N}\right]\left[\int_{-\infty}^{\infty}\frac{1}{\sqrt{\pi}}\exp\left(-t^2\right)dt\right]^N$$

$$- \log(e).N.\left[\int_{-\infty}^{\infty} \frac{t_1^2}{\sqrt{\pi}} \exp\left(-t_1^2\right) dt_1\right].\left[\int_{-\infty}^{\infty} \frac{1}{\left(\sqrt{\pi}\right)} \left[\exp\left(-t_2^2\right)\right] dt_2\right]^{N-1}$$

$$= -\frac{N}{2}\log\left(\pi N_0\right) - \frac{N}{2}\log(e)$$

$$= -\frac{N}{2}\log\left(\pi e N_0\right) \qquad \text{(B13)}$$

We now evaluate I_2.

$$I_2 = \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} p\left(\underline{y}|\underline{x}_m\right) p(\underline{x}_m) \log\left[\sum_{i=1}^{M} p(\underline{x}_i) p\left(\underline{y}|\underline{x}_i\right)\right] d\underline{y}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} p\left(\underline{y}|\underline{x}_m\right) \log\left[\frac{1}{M} \sum_{i=1}^{M} p\left(\underline{y}|\underline{x}_i\right)\right] d\underline{y}$$

$$= \frac{1}{M} \log\left(\frac{1}{M}\right) \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} p\left(\underline{y}|\underline{x}_m\right) d\underline{y}$$

$$+ \frac{1}{M} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} p\left(\underline{y}|\underline{x}_m\right) \log\left[\sum_{i=1}^{M} p\left(\underline{y}|\underline{x}_i\right)\right] d\underline{y}$$

$$= -\log(M) + \frac{1}{M} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} p\left(\underline{y}|\underline{x}_m\right) \log\left[\sum_{i=1}^{M} p\left(\underline{y}|\underline{x}_i\right)\right] d\underline{y}$$

$$= -\log(M) + I_2' \qquad \text{(B14)}$$

We now solve for I_2'. We have

$$p\left(\underline{y}|\underline{x}_m\right) = \frac{1}{\left(\sqrt{\pi N_0}\right)^N} \exp\left[-\frac{\left|\underline{y} - \underline{x}_m\right|^2}{N_0}\right]$$

$$p\left(\underline{y}|\underline{x}_i\right) = \frac{1}{\left(\sqrt{\pi N_0}\right)^N} \exp\left[-\frac{\left|\underline{y} - \underline{x}_i\right|^2}{N_0}\right]. \qquad \text{(B15)}$$

Let

$$\frac{\left|\underline{y} - \underline{x}_m\right|^2}{N_0} = |\underline{t}|^2 \ , \ d\underline{y} = \left(\sqrt{N_0}\right)^N d\underline{t}$$

$$\left\{ \underline{y} \right\} \to + \infty \Rightarrow \{ \underline{t} \} \to + \infty$$

$$\left\{ \underline{y} \right\} \to - \infty \Rightarrow \{ \underline{t} \} \to - \infty . \tag{B16}$$

Expanding and making substitutions

$$\begin{aligned} \frac{y_n - x_{in}}{\sqrt{N_0}} &= \frac{y_n - x_{mn} + x_{mn} - x_{in}}{\sqrt{N_0}} \\ &= \frac{y_n - x_{mn}}{\sqrt{N_0}} + \frac{x_{mn} - x_{in}}{\sqrt{N_0}} \\ &= t_n + \frac{x_{mn} - x_{in}}{\sqrt{N_0}} \ . \end{aligned} \tag{B17}$$

Then

$$\begin{aligned} \frac{\left(y_n - x_{in}\right)^2}{N_0} &= \left(t_n + \frac{x_{mn} - x_{in}}{\sqrt{N_0}} \right)^2 \\ &= {t_n}^2 + \frac{2t_n\left(x_{mn} - x_{in}\right)}{\sqrt{N_0}} + \frac{\left(x_{mn} - x_{in}\right)^2}{N_0} \end{aligned} \tag{B18}$$

and

$$\begin{aligned} \frac{\left|\underline{y} - \underline{x}_i\right|^2}{N_0} &= \sum_{n=1}^{N} \frac{\left(y_n - x_{in}\right)^2}{N_0} \\ &= \sum_{n=1}^{N} {t_n}^2 + \sum_{n=1}^{N} \frac{2t_n\left(x_{mn} - x_{in}\right)}{\sqrt{N_0}} + \sum_{n=1}^{N} \frac{\left(x_{mn} - x_{in}\right)^2}{N_0} \\ &= |\underline{t}|^2 + \frac{2\underline{t} \cdot \left(\underline{x}_m - \underline{x}_i\right)}{\sqrt{N_0}} + \frac{\left|\underline{x}_m - \underline{x}_i\right|^2}{N_0} \ . \end{aligned} \tag{B19}$$

Let

$$\underline{d}_{mi} = \frac{\underline{x}_m - \underline{x}_i}{\sqrt{N_0}} \tag{B20}$$

then

$$\underline{d}_{mi} = \left(\frac{\underline{x}_{m1} - \underline{x}_{i1}}{\sqrt{N_0}}, \ldots, \frac{\underline{x}_{mN} - \underline{x}_{iN}}{\sqrt{N_0}} \right)$$

$$\begin{aligned} \left|\underline{d}_{mi}\right|^2 &= \sum_{n=1}^{N} {d_{min}}^2 \\ &= \sum_{n=1}^{N} \left(\frac{\underline{x}_{mn} - \underline{x}_{in}}{\sqrt{N_0}} \right)^2 \end{aligned}$$

$$= \frac{|\underline{x}_m - \underline{x}_i|^2}{N_0} \tag{B21}$$

thus

$$\frac{|\underline{y} - \underline{x}_i|^2}{N_0} = |\underline{t}|^2 + 2\underline{t} \cdot \underline{d}_{mi} + |\underline{d}_{mi}|^2 \ . \tag{B22}$$

$$I'_2 = \frac{1}{M} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} p(\underline{y}|\underline{x}_m) \log\left[\sum_{i=1}^{M} p(\underline{y}|\underline{x}_i)\right] d\underline{y}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} \frac{1}{\left(\sqrt{\pi N_0}\right)^N} \exp\left[-\frac{|\underline{y} - \underline{x}_m|^2}{N_0}\right] \log\left[\sum_{i=1}^{M} \frac{1}{\left(\sqrt{\pi N_0}\right)^N} \exp\left[-\frac{|\underline{y} - \underline{x}_i|^2}{N_0}\right]\right] d\underline{y}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} \frac{\left(\sqrt{N_0}\right)^N}{\left(\sqrt{\pi N_0}\right)^N} \exp\left(-|\underline{t}|^2\right).$$

$$\log\left[\frac{1}{\left(\sqrt{\pi N_0}\right)^N} \sum_{i=1}^{M} \exp\left[-\left(|\underline{t}|^2 + 2\underline{t} \cdot \underline{d}_{mi} + |\underline{d}_{mi}|^2\right)\right]\right] d\underline{t}$$

$$= \frac{1}{M} \sum_{m=1}^{M} \log\left[\frac{1}{\left(\sqrt{\pi N_0}\right)^N}\right] \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} \frac{1}{\left(\sqrt{\pi}\right)^N} \exp\left(-|\underline{t}|^2\right) d\underline{t}$$

$$+ \frac{1}{M} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} \frac{1}{\left(\sqrt{\pi}\right)^N} \exp\left(-|\underline{t}|^2\right) \log\left[\exp\left(-|\underline{t}|^2\right)\right] d\underline{t}$$

$$+ \frac{1}{M} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} \frac{1}{\left(\sqrt{\pi}\right)^N} \exp\left(-|\underline{t}|^2\right) \log\left[\sum_{i=1}^{M} \exp\left(-2\underline{t} \cdot \underline{d}_{mi} - |\underline{d}_{mi}|^2\right)\right] d\underline{t} \ . \tag{B23}$$

Using the derivation of (B13)

$$I'_2 = -\frac{N}{2} \log(\pi e N_0)$$

$$+ \frac{1}{M} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} \frac{1}{\left(\sqrt{\pi}\right)^N} \exp\left(-|\underline{t}|^2\right) \log\left[\sum_{i=1}^{M} \exp\left(-2\underline{t} \cdot \underline{d}_{mi} - |\underline{d}_{mi}|^2\right)\right] d\underline{t} \ . \tag{B24}$$

Now

$$I_2 = -\log(M) - \frac{N}{2}\log(\pi e N_0)$$

$$+ \frac{1}{M\left(\sqrt{\pi}\right)^N} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} \exp\left(-|\underline{t}|^2\right) \log\left[\sum_{i=1}^{M} \exp\left(-2\underline{t} \cdot \underline{d}_{mi} - |\underline{d}_{mi}|^2\right)\right] d\underline{t} \ . \tag{B25}$$

Thus channel capacity can be written as

$$C_{DCMC} = I_1 - I_2$$

$$= \log(M)$$

$$- \frac{1}{M\left(\sqrt{\pi}\right)^N} \sum_{m=1}^{M} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{N-\text{fold}} \exp\left(-|\underline{t}|^2\right) \log_2\left[\sum_{i=1}^{M} \exp\left(-2\underline{t} \cdot \underline{d}_{mi} - |\underline{d}_{mi}|^2\right)\right] d\underline{t} \quad [\text{bit/sym}] \tag{B26}$$

where $\underline{d}_{mi} = \left(\underline{x}_m - \underline{x}_i\right)/\sqrt{N_0}$.

APPENDIX C: ERROR RATE FOR M–ARY SIGNAL SETS

This appendix tabulates error rates for the M–ary signal sets studied in Chapter 3. The values of Q(x) are obtained from [41, pp. 373–380]. We use the following relationships

$$\begin{aligned} \text{erfc}(x) &= 2Q\left(x\sqrt{2}\right) , \\ k &= \log_2 M , \\ \text{SNR} &= k\frac{E_b}{N_0} . \end{aligned} \tag{C1}$$

Tables for both probability of symbol error, P_e, and bit error, P_b, are given where possible.

C1: PAM

The probability of symbol error is given by [12, pp. 276–277]

$$P_e = \frac{M-1}{M}\text{erfc}\left(\sqrt{\frac{3}{M^2-1}\text{SNR}}\right) . \tag{C2}$$

Table C.1 lists the required values of SNR and E_b/N_0 for $P_e = 10^{-5}$.

Table C.1 PAM – SNR and E_b/N_0 for $P_e = 10^{-5}$

M	SNR [dB]	E_b/N_0 [dB]
2	9.59	9.59
4	16.77	13.76
8	23.05	18.28
16	29.17	23.15

C2: QAM

The probability of symbol error is upper bounded by [12, p. 283]

$$P_e \leq 2\text{erfc}\left(\sqrt{\frac{3}{2(M-1)}\text{SNR}}\right) . \tag{C3}$$

(C3) holds for QAM constellations contained on a rectangular grid. Table C.2 lists the required values of SNR and E_b/N_0 for $P_e = 10^{-5}$.

Table C.2 QAM – SNR and E_b/N_0 for $P_e = 10^{-5}$

M	SNR [dB]	E_b/N_0 [dB]
8	16.87	12.09
16	20.18	14.16
32	23.33	16.34
64	26.41	18.63

C3: PSK

For M=2 [12, p. 622]

$$P_e = P_b = Q\left(\sqrt{2\frac{E_b}{N_0}}\right) . \tag{C4}$$

For M=4 [12, p. 623]

$$P_e = \text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)\left[1 - \frac{1}{4}\text{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)\right] . \tag{C5}$$

The probability of symbol error is upper bounded by [12, p. 265]

$$P_e \leq \text{erfc}\left(\sqrt{SNR}\sin\frac{\pi}{M}\right) , \; M > 4 . \tag{C6}$$

Table C.3 lists the required value of SNR for $P_e = 10^{-5}$.

Table C.3 PSK – SNR for $P_e = 10^{-5}$

M	SNR [dB]
2	9.59
4	12.90
8	18.24
16	24.09
32	30.07
64	36.08

The probability of bit error for a Gray mapped constellation is approximated by [12, p. 265]

$$P_b \simeq \frac{P_e}{k} . \tag{C7}$$

Table C.4 lists the required value of E_b/N_0 for $P_b = 10^{-5}$ using equations (C4) to (C7).

Table C.4 PSK – E_b/N_0 for $Pb = 10^{-5}$

M	E_b/N_0 [dB]
2	9.59
4	9.59
8	12.97
16	17.44
32	22.34
64	27.46

C4: M–ary Orthogonal Signalling

The probability of symbol error is obtained by the union bound [12, p. 251]

$$P_e \leq \frac{M-1}{2}\operatorname{erfc}\left(\sqrt{\frac{SNR}{2}}\right). \tag{C8}$$

Table C.5 lists the required value of SNR for $P_e = 10^{-5}$.

Table C.5 M–ary orthogonal signalling – SNR for $P_e = 10^{-5}$

M	SNR [dB]
2	12.60
4	13.07
8	13.41

Using the relationship between P_b and P_e [12, p. 250]

$$P_b = \frac{2^{M-1}}{2^M-1}P_e \tag{C9}$$

Table C.6 lists the required value of E_b/N_0 for $P_b = 10^{-5}$.

Table C.6 M–ary orthogonal signalling – E_b/N_0 for $P_b = 10^{-5}$

M	E_b/N_0 [dB]
2	12.41
4	9.80
8	8.37

C5: L–Orthogonal Signalling

The probability of symbol error, P_e, is upper bounded by [22]

$$P_e = P_e(V) + P_e(L)[1-P_e(V)] \ . \tag{C10}$$

$P_e(V)$ is the probability of symbol error for incoherent detection of orthogonal envelope.

$P_e(L)$ is the probability of symbol error for incoherent detection of phase given correct

detection of orthogonal envelope. $P_e(V)$ and $P_e(L)$ are given by [18]

$$P_e(V) = \frac{\exp\left(\frac{-SNR}{2}\right)}{V} \sum_{v=2}^{V} \binom{V}{v} (-1)^v \exp\left\{\frac{-SNR(2-v)}{2v}\right\} ,$$

$$P_e(L) = 2Q\left(\sqrt{2SNR}\sin\frac{\pi}{L\sqrt{2}}\right) , \; L \geq 4 . \tag{C11}$$

We can upper bound bit error, P_b, by using equations (C7) and (C9) in (C10)

$$P_b = P_b(V) + P_b(L)\left[1 - P_b(V)\right] . \tag{C12}$$

Tables C.7 to C.10 list error probabilities for V=2 and V=4 L–orthogonal signalling. These results are calculated using double precision ANSI C programs. Overflow errors occurred for V=4, L=64 signalling.

Table C.7 V=2 L–orthogonal signalling – SNR for $P_e = 10^{-5}$

L	SNR [dB]
4	15.45
8	21.10
16	27.05
32	33.05
64	39.05

Table C.8 V=2 L–orthogonal signalling – E_b/N_0 for $P_b = 10^{-5}$

L	E_b/N_0 [dB]
4	10.35
8	14.65
16	19.40
32	24.50
64	29.75

Table C.9 V=4 L–orthogonal signalling – SNR for $P_e = 10^{-5}$

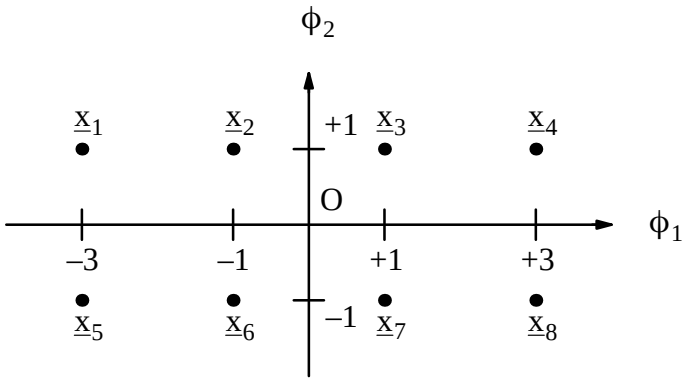
L	SNR [dB]
4	16.1
8	21.10
16	27.05
32	33.05

Table C.10 V=4 L–orthogonal signalling – E_b/N_0 for $P_b = 10^{-5}$

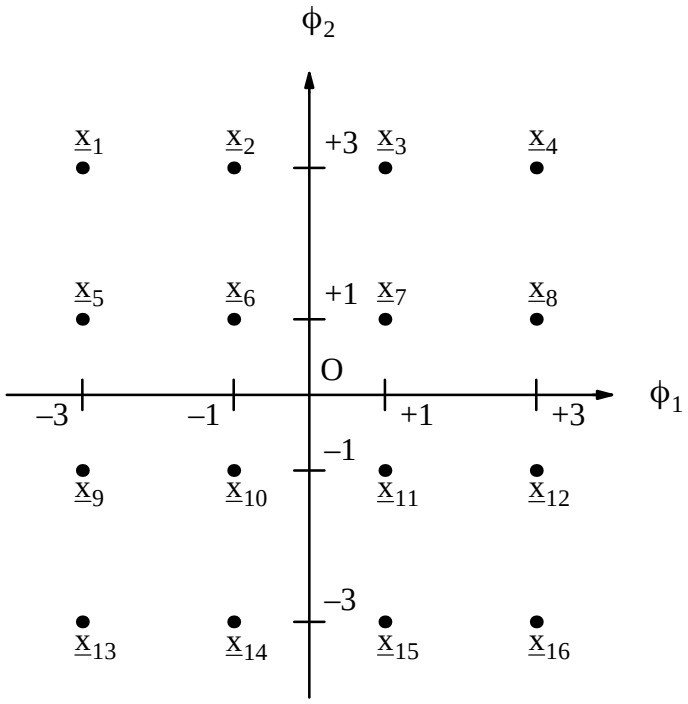
L	E_b/N_0 [dB]
4	9.8
8	13.64
16	18.86
32	24.04

APPENDIX D: QAM CONSTELLATIONS

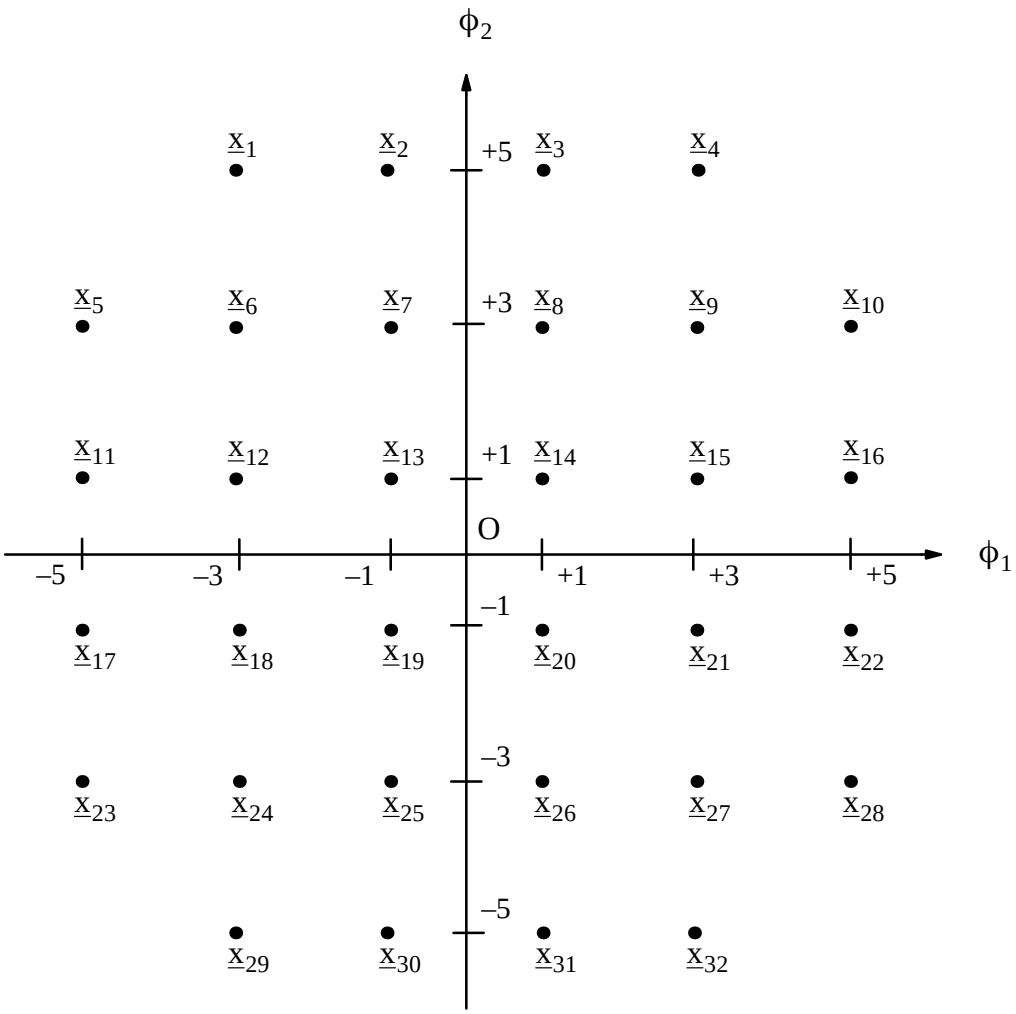
This Appendix contains QAM constellations used for channel capacity calculations.



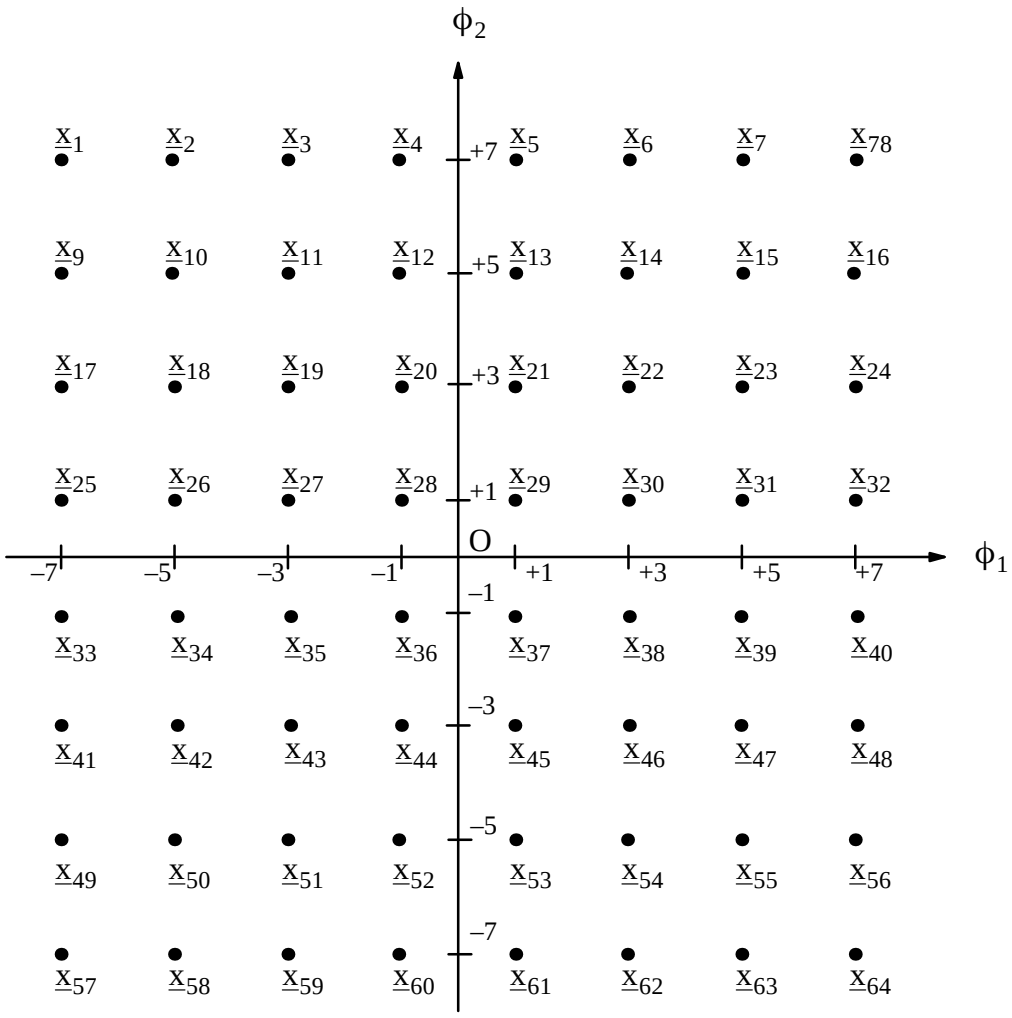
M=8



M=16


ϕ_2
$\underline{x}_1$
$\underline{x}_2$
$\underline{x}_3$
$\underline{x}_4$
$\underline{x}_5$
$\underline{x}_6$
$\underline{x}_7$
$\underline{x}_8$
$\underline{x}_9$
$\underline{x}_{10}$
$\underline{x}_{11}$
$\underline{x}_{12}$
$\underline{x}_{13}$
$\underline{x}_{14}$
$\underline{x}_{15}$
$\underline{x}_{16}$
$\underline{x}_{17}$
$\underline{x}_{18}$
$\underline{x}_{19}$
$\underline{x}_{20}$
$\underline{x}_{21}$
$\underline{x}_{22}$
$\underline{x}_{23}$
$\underline{x}_{24}$
$\underline{x}_{25}$
$\underline{x}_{26}$
$\underline{x}_{27}$
$\underline{x}_{28}$
$\underline{x}_{29}$
$\underline{x}_{30}$
$\underline{x}_{31}$
$\underline{x}_{32}$
+5
+3
+1
–1
–3
–5
O
–5
–3
–1
+1
+3
+5
ϕ_1

M=32


ϕ2
ϕ1
O
x1 x2 x3 x4 +7 x5 x6 x7 x78
x9 x10 x11 x12 +5 x13 x14 x15 x16
x17 x18 x19 x20 +3 x21 x22 x23 x24
x25 x26 x27 x28 +1 x29 x30 x31 x32
–7 –5 –3 –1 +1 +3 +5 +7
–1 x33 x34 x35 x36 x37 x38 x39 x40
–3 x41 x42 x43 x44 x45 x46 x47 x48
–5 x49 x50 x51 x52 x53 x54 x55 x56
–7 x57 x58 x59 x60 x61 x62 x63 x64

M=64

APPENDIX E: C FUNCTION FOR REPEATED QUADRATURE

The function is named **NGAUSS.c**. The nested sum (4.6) is rewritten in the form as evaluated by **NGAUSS.c**

$$\text{sumi} = \sum_{\substack{i=0 \\ \text{level}=N-1}}^{P-1} w[i] \sum_{\substack{i=0 \\ \text{level}=N-2}}^{P-1} w[i] \; \ldots \sum_{\substack{i=0 \\ \text{level}=0}}^{P-1} w[i] \; f(t) \; . \qquad \text{(E.1)}$$

The quadrature points and weights are stored in arrays **x** and **w** each of size P elements. The function, **f**, is written as a separate C function. The variable, **t**, is a P element array with each element initialised according to the corresponding value for **level**. The format of **NGAUSS.c** is:

```
double  NGAUSS(int level)
{
  int i;
  double sumi=0.0;

  if(level!=0)
    {
    for(i=0;i<P;i++)
      {
      t[level]=x[i];
      sumi+=w[i]*NGAUSS(level–1);
      }
    }
    else /*innermost level*/
    {
    for(i=0;i<P;i++)
      {
      t[0]=x[i];
      sumi+=w[i]*f(t);
      }
    }
    return(sumi);
}
```

APPENDIX F: DECIMAL PLACE ACCURACY OF NUMERICAL INTEGRATION

This Appendix tabulates the decimal place accuracy of repeated Gauss–Hermite quadrature for a function with known solution (4.7).

Table F.1 Decimal place accuracy

N	P	Decimal Place Accuracy
1	10	13
2	10	12
2	5	8
2	2	13
4	10	12
4	5	8
4	2	13
8	5	7
8	2	11

APPENDIX G: NUMBER OF SUMMATIONS FOR CAPACITY CALCULATION

This Appendix tabulates the number of summations performed for capacity calculations. The total number of summations is M^2P^N.

Table G.1 PAM – N=1, P=10

M	M^2P^N
2	40
4	160
8	640
16	2560

Table G.2 QAM, PSK – N=2, P=10

M	M^2P^N [10^3]
2	0.4
4	1.6
8	6.4
16	25.6
32	102.4
64	409.6

Table G.3 M=4 orthogonal, V=2 L–orthogonal – N=4, P=5

M	M^2P^N [10^6]
4	0.010
8	0.040
16	0.160
32	0.640
64	2.560
128	10.240

Table G.4 M=8 orthogonal, V=4 L–orthogonal – N=8, P=5

M	M^2P^N [10^9]
8	0.025
16	0.100
32	0.400
64	1.600
128	6.400
256	25.600