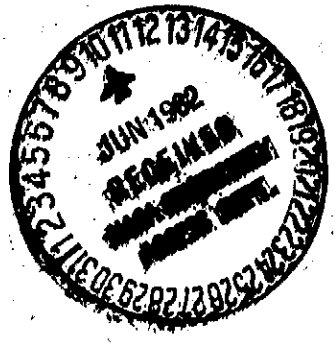
JPL PUBLICATION 81-73

Modulation/Demodulation Techniques for Satellite Communications

Part II: Advanced Techniques – The Linear Channel

Jim K. Omura
Marvin K. Simon



March 1, 1982

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL PUBLICATION 81-73

Modulation/Demodulation Techniques for Satellite Communications

Part II: Advanced Techniques – The Linear Channel

Jim K. Omura
Marvin K. Simon

March 1, 1982

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

FOREWORD

This report represents Part II of a series of reports to be published under the same title with the following subtitles:

Part I: Background
Part III: Advanced Techniques - The Nonlinear Channel
Part IV: Appendices

ABSTRACT

Part II of this report presents a theory for deducing and predicting the performance of transmitter/receivers for bandwidth efficient modulations suitable for use on the linear satellite channel. The underlying principle used throughout is the development of receiver structures based on the maximum-likelihood decision rule. Along with this overall theme is the desire to apply the performance prediction tools, e.g., channel cutoff rate and bit error probability transfer function bounds developed in great detail in Part IV to these modulation/demodulation techniques. Many of the numerical illustrations given in summary form in Part I are theoretically justified here in Part II.

CONTENTS

Figures

1.0 Introduction

Part I of this report presented a tutorial discussion of the various modulation and coding techniques commonly considered for radio and satellite communication links. Although not an exhaustive presentation, it did contain sufficient detail to serve as a background for the more detailed study of advanced modulation/demodulation techniques to be undertaken here in Part II and later in Part III.

In this part, we examine a general class of bandwidth efficient modulation techniques and their error probability performance over the linear channel. Part III examines these same modulation techniques for nonlinear satellite channels.

We begin by defining the mathematical model for the linear channel and accompanying generalized maximum-likelihood decision rules. Next a class of bandwidth efficient modulation techniques and the corresponding maximum likelihood demodulators are described. A new concept of simultaneous phase and data estimation is examined followed by error probability evaluation of these advanced bandwidth efficient modulations over the linear channel.

2.0 Linear Channel Model

The linear channel model we consider here is the classical additive white Gaussian noise (AWGN) channel sketched in Figure 1 with x(t) the transmitted signal and y(t) the signal at the input to the receiver. Then,

$$y(t) = x(t)+n(t) \tag{II.2.1}$$

where n(t) is the additive white Gaussian noise with autocorrelation

$$E\{n(t+t)n(t)\} = \frac{N_0}{2}\delta(t). \tag{II.2.2}$$

The AWGN channel is often used as a model for the temperature dependent noise in the front end of receiver systems that is due to the sum of many incremental voltages caused by electrons in thermal motion. Based on the Central Limit Theorem, one would expect the Gaussian assumption to be a good statistical

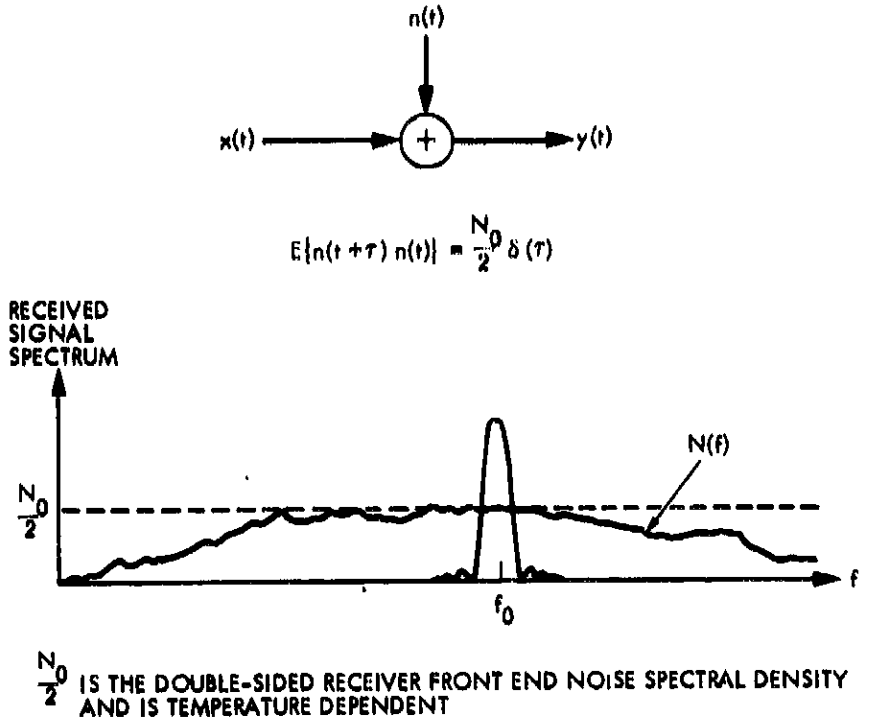


Figure 1. Additive White Gaussian Noise (AWGN) Channel

characterization of this additive noise. For the typical narrowband radio signal, the noise power distribution is generally much wider than the signal bandwidth as shown in Figure 1, in which case it is reasonable to assume that this noise energy distribution is flat over the signal bandwidth. Since we are only interested in the additive noise power distribution in the signal bandwidth, the above assumption is equivalent to assuming that the additive noise power distribution is constant for all frequencies. Although this approximation makes little difference in the signal band, it greatly simplifies the mathematical analysis for the linear channel. Hence, in what follows, we assume a constant noise power spectral density over all frequencies (double-sided) given by $N_0/2$.

2.1 Representation of AWGN

Let n(t) be an AWGN process with two-sided power spectral density given by $N_0/2$. Consider any set of orthonormal functions $\phi_1(t)$, $\phi_2(t)$, ... , $\phi_N(t)$

over the interval [0,T] and define the noise correlating components.

$$n_k = \int_0^T n(t)\phi_k(t)dt \quad ; \quad k=1,2,\ldots,N. \tag{11.2.3}$$

It is easy to verify that

$$E\{n_i n_j\} = \frac{N_0}{2}\delta_{ij} \tag{11.2.4}$$

where

$$\delta_{ij} = \begin{cases} 1 \; ; \; i=j \\ 0 \; ; \; i\neq j \end{cases} . \tag{11.2.5}$$

Thus, the noise components n_1, n_2, ... , n_N are independent identically distributed (i.i.d.) Gaussian random variables with zero mean and variance $N_0/2$. The random process

$$\hat{n}(t) = \sum_{k=1}^{N} n_k\phi_k(t) \tag{11.2.6}$$

is the projection of the AWGN process n(t) onto the space spanned by the orthonormal functions $\phi_1(t)$, $\phi_2(t)$, ... $\phi_N(t)$. The part of the AWGN process outside the subspace, namely,

$$\tilde{n}(t) = n(t)-\hat{n}(t) \tag{11.2.7}$$

satisfies the conditions

$$\int_0^T \tilde{n}(t)\phi_k(t)dt = 0 \; ; \; k=1,2,\ldots,N \tag{11.2.8}$$

and

$$E[\tilde{n}(t_1)\hat{n}(t_2)] = 0 \quad \text{for all } t_1, t_2 . \tag{II.2.9}$$

This means that in accordance with (II.2.7), n(t) can be separated into two independent random processes (see Figure 2) $\hat{n}(t)$ and $\tilde{n}(t)$ where $\hat{n}(t)$ can be obtained by direct correlation of the AWGN, n(t), with the orthonormal functions as in (II.2.6).

2.2 Representation of the Received Signal Plus Noise

Suppose that one of M signals is transmitted over the linear channel during the time interval [0,T]. Denote these signals as $s_1(t)$, $s_2(t)$, ... , $s_M(t)$ with energies

$$E_m = \int_0^T x_m^2(t)dt \; ; \; m=1,2,\ldots,M \tag{II.2.10}$$

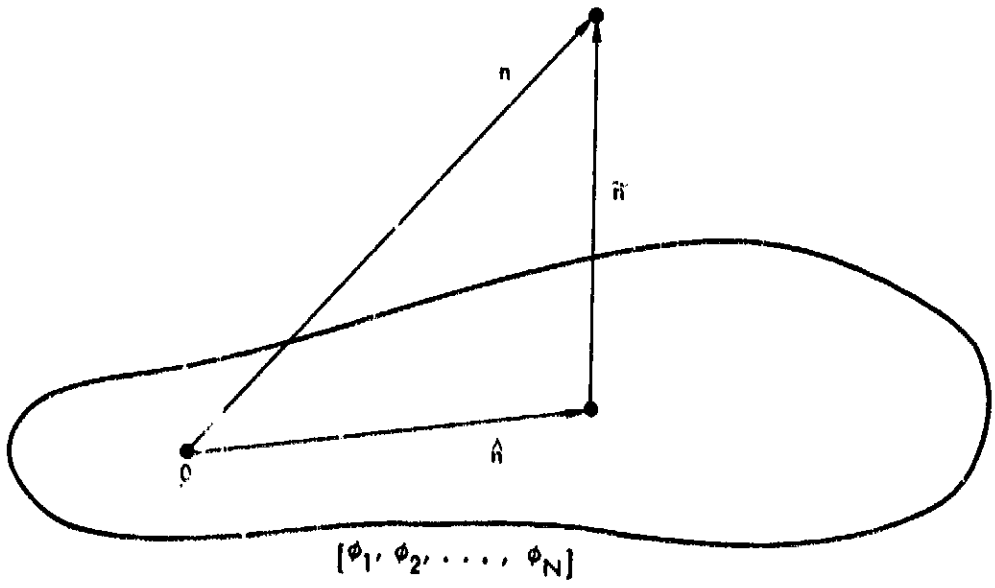


Figure 2. Representation of White Gaussian Noise

For the given signal set, we can always find an orthonormal basis $\phi_1(t)$, $\phi_2(t)$, ..., $\phi_N(t)$ with $N \leq M$ such that all signals are representable as a linear combination of this set. That is,

$$x_m(t) = \sum_{k=1}^{N} x_{mk}\phi_k(t) \quad ; \quad m=1,2,\ldots,M \tag{II.2.11}$$

where the signal coefficients are given by

$$x_{mk} = \int_0^T x_m(t)\phi_k(t)dt \quad ; \quad \begin{matrix} k=1,2,\ldots,N \\ m=1,2,\ldots,M. \end{matrix} \tag{II.2.12}$$

Arbitrarily, suppose that during the interval [0,T], the transmitted signal is $x_1(t)$. The signal out of the AWGN channel is then

$$\begin{aligned} y(t) &= x_1(t) + n(t) \\ &= \sum_{k=1}^{N} x_{1k}\phi_k(t) + \sum_{k=1}^{N} n_k\phi_k(t) + \tilde{n}(t) \\ &= \sum_{k=1}^{N} (x_{1k} + n_k)\phi_k(t) + \tilde{n}(t). \end{aligned} \tag{II.2.13}$$

At the receiver, we can find the basis components of the received signal, namely

$$\begin{aligned} y_k &= \int_0^T y(t)\phi_k(t)dt \\ &= x_{1k} + n_k \quad ; \quad k=1,2,\ldots,N \end{aligned} \tag{II.2.14}$$

from which we have the projection onto the signal space

$$\hat{y}(t) = \sum_{k=1}^{N} y_k \phi_k(t)$$

$$= \sum_{k=1}^{N} (x_{1k} + n_k)\phi_k(t) \qquad \text{(II.2.15)}$$

The only difference between the AWGN channel output, y(t), and its projection onto the signal space, $\hat{y}(t)$, is $\tilde{n}(t)$ of (II.2.7). Since this difference is independent of both the signal and noise $\hat{n}(t)$ in the signal space, there is no loss of optimality in restricting our attention to $\hat{y}(t)$ alone. Note that noise outside the signal space (or signal frequency band) is unimportant.

Since $\hat{y}(t)$ is equivalent to the vector $\underline{y} = (y_1, y_2, \ldots, y_N)$, we can take our channel output as the vector $\underline{y}$ and the set of possible channel inputs as

$$\underline{x}_m = (x_{m1}, x_{m2}, \ldots, x_{mN}) \quad ; \quad m=1,2,\ldots,N \qquad \text{(II.2.16)}$$

Then if $\underline{x}_1$ is transmitted, the channel output is

$$\underline{y} = \underline{x}_1 + \underline{n} \qquad \text{(II.2.17)}$$

where

$$\underline{n} = (n_1, n_2, \ldots, n_N) \qquad \text{(II.2.18)}$$

Based on the foregoing, the probability of $\underline{y}$ given that $\underline{x}_1$ is sent is given by

$$p_N(\underline{y}|\underline{x}_1) = \left(\frac{1}{\pi N_0}\right)^{\frac{N}{2}} \exp\left\{-\|\underline{y}-\underline{x}_1\|^2/N_0\right\} \qquad \text{(II.2.19)}$$

where

$$\|z\|^2 \triangleq \sum_{k=1}^{N} z_k^2 . \tag{II.2.20}$$

2.3 General Maximum-Likelihood Receivers

Let us now define $P_{E,m}$ as the probability that the receiver makes a decision error when x_m is transmitted. We implicitly assume that, given the channel output y, the receiver always decides that one of the M signals was transmitted. If all M possible signals are, a priori, equally likely to be transmitted, then the average error probability is

$$P_E = \frac{1}{M} \sum_{m=1}^{M} P_{E,m}. \tag{II.2.21}$$

For most cases of interest we wish to use the decision rule that minimizes P_E. This is the maximum-likelihood decision rule given by:

> Given y as the channel output, choose the signal x_m that has the largest conditional probability $p_N(y|x_m)$; $m = 1,2,\ldots,M$.

This is a general result for any vector channel characterized by conditional probabilities, $\{p_N(y|x_m)\}$. For the AWGN where $p_N(y,x_m)$ is given by (II.2.19) with x_1 replaced by x_m, the maximum-likelihood rule becomes:

> Given y as the channel output, choose the signal $x_{\hat{m}}$ that has the smallest distance $\|y-x_{\hat{m}}\|^2$ from y.

Note that in view of (11.2.22), we have for any m the relationships

$$E_m = \int_0^T s_m^2(t)\,dt$$

$$= \sum_{k=1}^{N} s_{mk}^2 = \|s_m\|^2 \qquad (11.2.22)$$

and

$$\int_0^T y(t)s_m(t)\,dt = \sum_{k=1}^{N} s_{mk} \int_0^T y(t)\varphi_k(t)\,dt$$

$$= \sum_{k=1}^{N} y_k s_{mk} = (y, s_m), \qquad (11.2.23)$$

Further noting that

$$\|y - s_m\|^2 = \|y\|^2 + \|s_m\|^2 - 2(y, s_m) \qquad (11.2.24)$$

then the maximum-likelihood decision rule also has the form:

Given $y(t)$ as the channel output, choose the signal $s_m(t)$ that has the largest biased correlation

$$(y, s_m) - \frac{1}{2}\|s_m\|^2 = \int_0^T y(t)s_m(t)\,dt - \frac{1}{2}\int_0^T s_m^2(t)\,dt \;;$$

$$m = 1, 2, \ldots, M \qquad (11.2.25)$$

Note that this form of the maximum-likelihood decision rule is coordinate free. That is, there is no need to pick an orthonormal basis $\phi_1(t), \phi_2(t), \ldots, \phi_N(t)$ and compute the channel output vector $\underline{y}$. Here we do direct correlation of the channel output signal $y(t)$ with each of the possible signals $x_1(t), x_2(t), \ldots, x_M(t)$. We shall use these two forms of the maximum-likelihood decision rule throughout the remainder of Part II. For the special cases where all signals have the same energy, such as with constant envelope signals, we can drop the $1/2 \|\underline{x}_m\|^2$ term since it is the same for all signals.

3.0 Modulation/Demodulation Techniques

In this section, we first briefly review some of the conventional modulation and demodulation techniques as discussed in Part I and then go on to examine more complex bandwidth-efficient modulations. The approach taken here will be to derive both maximum-likelihood and suboptimum demodulators and as such will provide a more detailed and broader understanding of the casual treatment of these same subjects given in Part I.

3.1 Coherent MPSK

Coherent MPSK signals have the form

$$x_m(t) = \sqrt{2S}\cos\left(\omega_0 t + \frac{2\pi m}{M}\right) \quad ; \quad \begin{matrix} m = 1,2,\ldots,M \\ 0 \le t \le T \end{matrix} \qquad \text{(II.3.1)}$$

Defining the orthonormal basis functions,*

$$\phi_c(t) = \sqrt{\frac{2}{T}}\cos\omega_0 t$$

$$\phi_s(t) = -\sqrt{\frac{2}{T}}\sin\omega_0 t \quad ; \quad 0 \le t \le T \qquad \text{(II.3.2)}$$

*We assume narrowband signals where these are orthogonal.

we see that all of the M signals can be represented as a sum of these basis functions as follows:

$$x_m(t) = \sqrt{E}\cos\left(\frac{2\pi m}{M}\right)\phi_c(t) + \sqrt{E}\sin\left(\frac{2\pi m}{M}\right)\phi_s(t) \qquad \text{(II.3.3)}$$

where

$$E = ST \qquad \text{(II.3.4)}$$

is the energy per MPSK pulse. The maximum-likelihood demodulator merely computes the received signal components

$$y_c = \int_0^T y(t)\phi_c(t)dt$$

$$y_s = \int_0^T y(t)\phi_s(t)dt \qquad \text{(II.3.5)}$$

and applies the maximum-likelihood decision rule:

Choose the signal $\underset{\sim}{x}_m$ that minimizes

$$\|\underset{\sim}{y}-\underset{\sim}{x}_m\|^2 = (y_c - x_{mc})^2 + (y_s - x_{ms})^2 \qquad \text{(II.3.6)}$$

where

$$x_{mc} = \sqrt{E}\cos\left(\frac{2\pi m}{M}\right) \qquad \text{(II.3.7a)}$$

and

$$x_{ms} = \sqrt{E}\sin\left(\frac{2\pi m}{M}\right) \qquad \text{(II.3.7b)}$$

for $m = 1,2,\ldots,M$. This demodulator structure is shown in Figure 3 for $M = 4$.

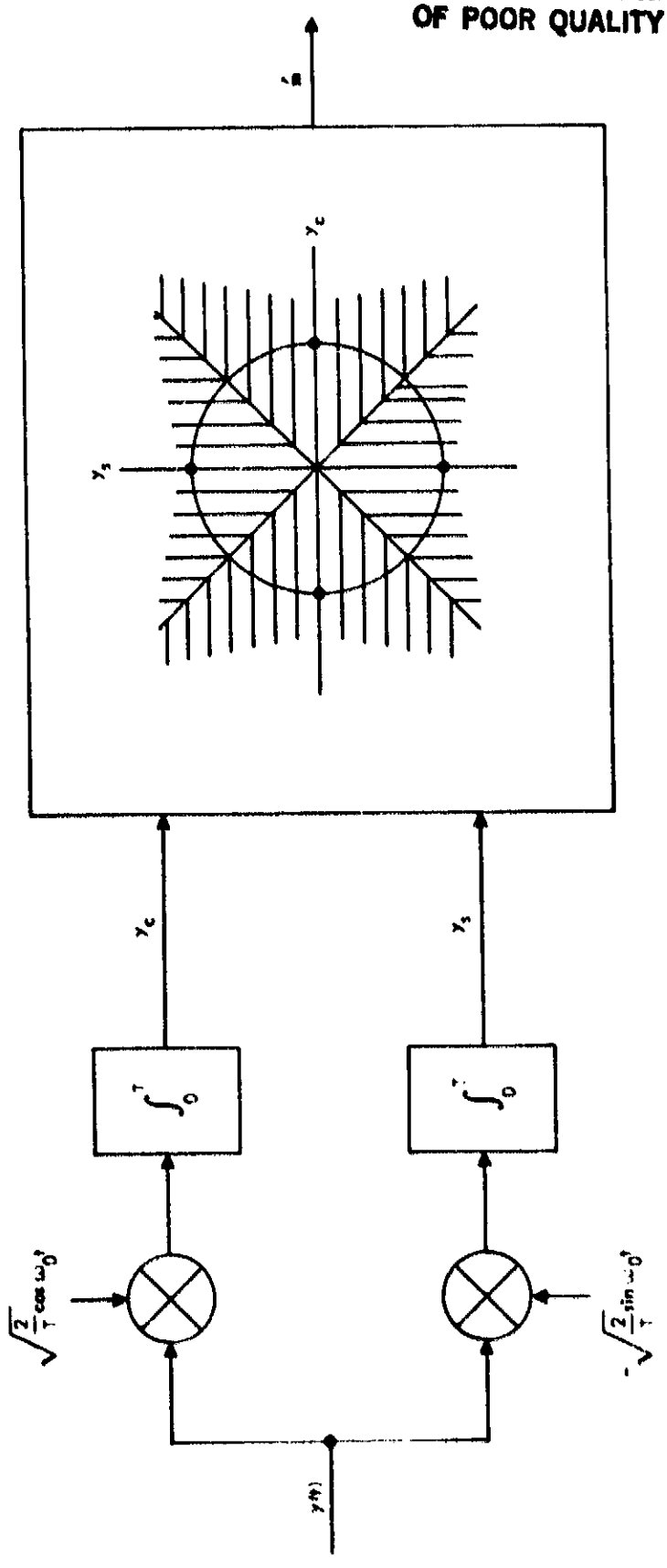


Figure 3. A Maximum-Likelihood Demodulator of QPSK

3.2 Coherent 16-QAM

For 16-QAM we have the signal representation

$$x_m(t) = a_m \phi_c(t) + b_m \phi_s(t) \qquad (II.3.8)$$

where $\phi_c(t)$ and $\phi_s(t)$ are the same basis functions as for MPSK. Here each amplitude,

$$a_m, b_m \in \{-3A, -A, A, 3A\} \qquad (II.3.9)$$

has one of four values resulting in the 16 possible signals. The maximum-likelihood demodulator computes y_c and y_s as in MPSK modulation and applies the decision rule:

Choose the signal $\underline{x}_{\hat{m}}$ that minimizes

$$||\underline{y} - \underline{x}_m||^2 = (y_c - x_{mc})^2 + (y_s - x_{ms})^2 \qquad (II.3.10)$$

where

$$x_{mc} = a_m \qquad (II.3.11a)$$

and

$$x_{ms} = b_m \qquad (II.3.11b)$$

This demodulator is sketched in Figure 4.

3.3 Noncoherent MFSK

Conventional MFSK signals have the form

$$x_m(t) = \sqrt{2S} \cos(\omega_m t + \theta); \; m = 1,2,\ldots,M \qquad (II.3.12)$$

$$0 \le t < T$$

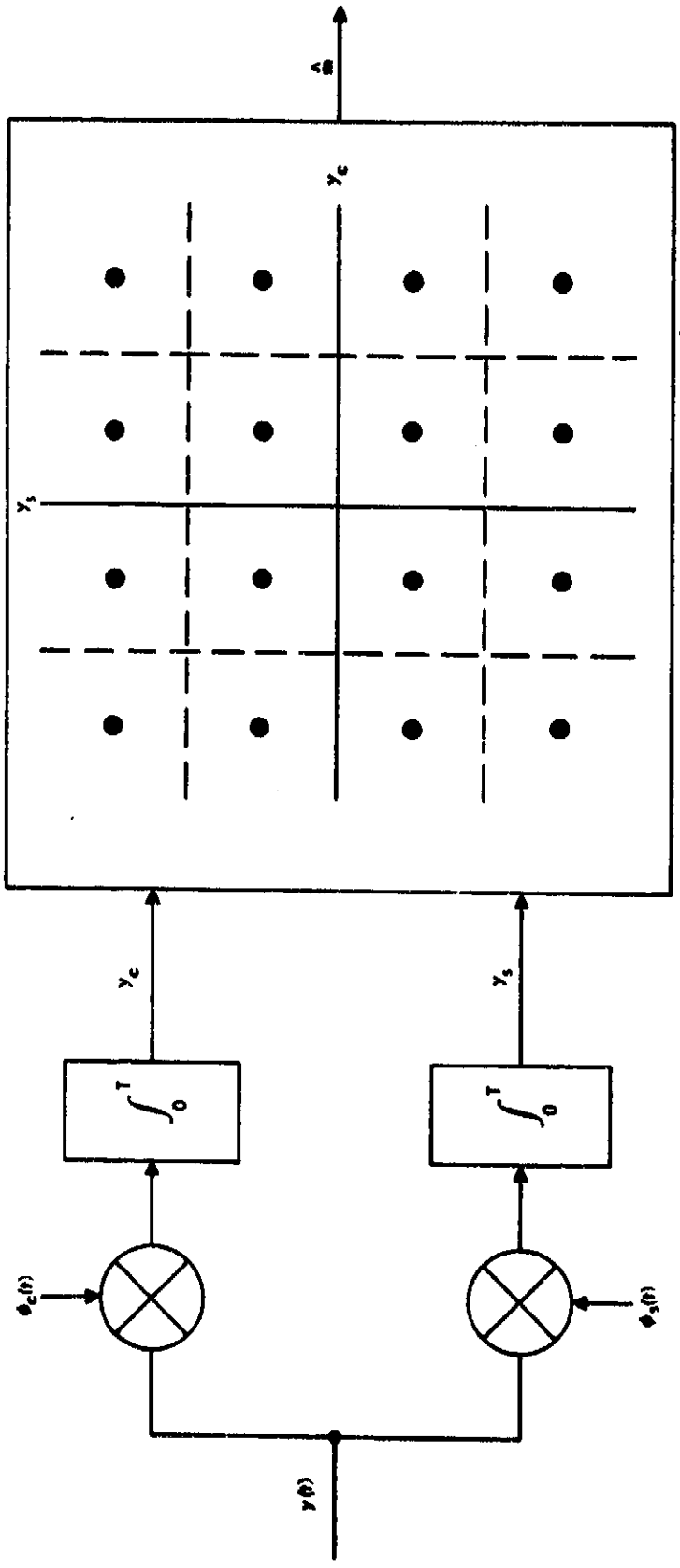


Figure 4. A Maximum-Likelihood Demodulator of 16-QAM

where θ is some arbitrary phase term and the frequency separations are multiples of $\Delta\omega = 2\pi/T$ so that

$$\int_0^T x_m(t)x_n(t)dt = E\delta_{mn};\ m,n = 1,2,\ldots,\ M \qquad \text{(II.3.13)}$$

where E as defined in (II.3.4) is the energy per MFSK pulse.

Here the complete set of orthonormal basis functions is

$$\phi_{mc}(t) = \sqrt{\frac{2}{T}} \cos\ \omega_m t$$

$$\phi_{ms}(t) = -\sqrt{\frac{2}{T}} \sin\ \omega_m t;\ m = 1,2,\ldots,\ M \qquad \text{(II.3.14)}$$

where all of the MFSK signals can be represented as a sum of these basis functions. That is,

$$x_m(t) = \sqrt{E} \cos\ \theta\ \phi_{mc}(t) + \sqrt{E} \sin\ \theta\ \phi_{ms}(t),\ m = 1,2,\ldots,M \qquad \text{(II.3.15)}$$

The receiver computes components of the vector $\underline{y} = (\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_M)$ where

$$\underline{y}_m = (y_{mc}, y_{ms})$$

$$y_{mc} = \int_0^T y(t)\phi_{mc}(t)\ dt$$

$$y_{ms} = \int_0^T y(t)\phi_{ms}(t)\ dt;\ m = 1,2,\ldots,\ M. \qquad \text{(II.3.16)}$$

If θ is known at the receiver, then the demodulator chooses the signal that minimizes $\|\underline{y} - \underline{x}_m\|^2$ where

$$\underline{x}_m = (\underline{x}_{m1}, \underline{x}_{m2}, \cdots, \underline{x}_{mM})$$

$$\underline{x}_{mk} = (0, 0); \; m \neq k$$

$$\underline{x}_{mm} = (\sqrt{E} \cos \theta, \sqrt{E} \sin \theta); \; m = 1,2,\ldots, M \qquad (II.3.17)$$

When θ is unknown at the receiver and assumed to be a random variable uniformly distributed over $[0,2\pi]$, then we have the noncoherent demodulator. This is based on the conditional probabilities

$$p(\underline{y}|\underline{x}_m) = g(\underline{y}_1)g(\underline{y}_2) \cdots g(\underline{y}_m) \exp(-E/N_0) I_0(\sqrt{E}\|\underline{y}_m\|); \qquad (II.3.18)$$

$$m = 1,2,\ldots, M$$

where

$$g(\underline{y}_k) = \frac{1}{\pi N_0} \exp(-\|\underline{y}_k\|^2/N_0) \qquad (II.3.19)$$

The noncoherent maximum-likelihood rule is to choose the signal $\underline{x}_{\hat{m}}$ that has the largest value of $p(\underline{y}|\underline{x}_m)$ for the given received vector $\underline{y}$. Since $I_0(\cdot)$ is a monotone increasing function of its argument, we have the noncoherent maximum-likelihood decision rule:

Choose the signal $\underline{x}_{\hat{m}}$ where $\hat{m}$ has the largest energy output

$$\|\underline{y}_m\|^2 = y_{mc}^2 + y_{ms}^2; \; m = 1,2,\ldots, M \qquad (II.3.20)$$

The noncoherent MFSK demodulator merely records the energy about each of the M possible carrier frequencies. This is sketched in Figure 5 for M = 2. Since this demodulator does not require knowledge of the arbitrary signal phase θ, the demodulator is simplified since no carrier synchronization is necessary. There is, however, some loss in performance in not estimating or tracking the carrier phase.

3.4 DBPSK

There are several different suboptimum demodulators one can use for differentially coherent BPSK. Since the differential encoder that precedes the BPSK modulator introduces memory or dependency on past data symbols, we must now consider sequences of pulses formed by sequences of symbols.

Denote the data sequence as $\underline{u} = (\ldots, u_{-1}, u_0, u_1, \ldots)$ and define the BPSK basis function

$$\phi_c(t) = \sqrt{\frac{2}{T}} \cos \omega_0 t, \quad 0 \le t \le T \tag{II.3.21}$$

and its T-sec time shifts

$$\phi_n(t) \triangleq \phi_c(t - nT), \quad n = \ldots, -1, 0, 1, \ldots \tag{II.3.22}$$

as an orthonormal basis. The data "bits" $\{u_n\}$ are taken to be independent identically distributed (i.i.d.) random variables with

$$\Pr\{u_n = 1\} = \Pr\{u_n = -1\} = \frac{1}{2} \tag{II.3.23}$$

The data sequence $\underline{u}$ is converted into another binary sequence

$$\underline{x} = (\ldots, x_{-1}, x_0, x_1, \ldots) \tag{II.3.24}$$

by the differential encoder which outputs

$$x_n = x_{n-1}\, u_n \qquad \text{for all } n. \tag{II.3.25}$$

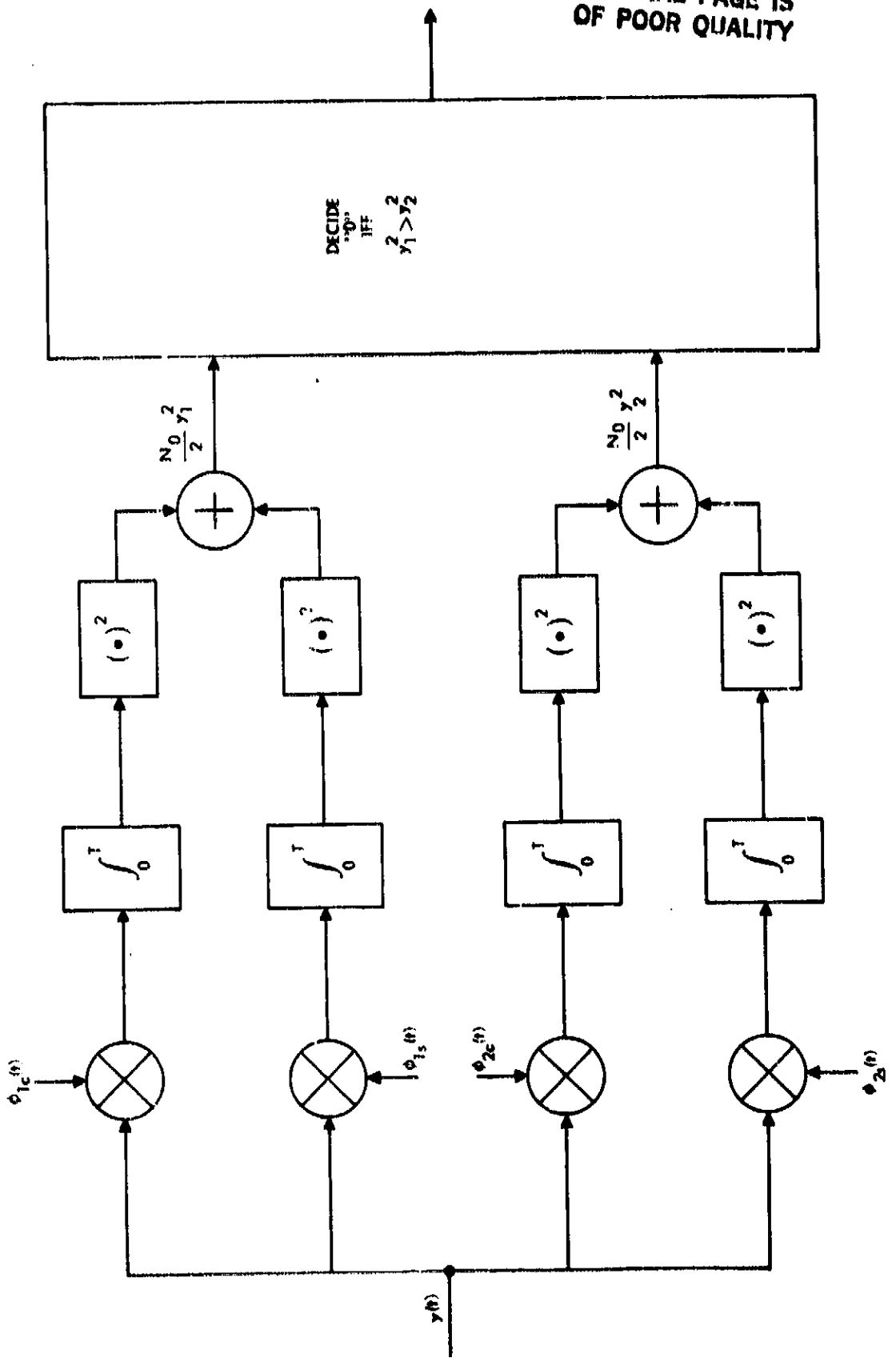


Figure 5. A Maximum-Likelihood Demodulator of Noncoherent FSK

Note that this results in a BPSK signal where the phase (0 or π) change is relative to the previously transmitted phase. The transmitted signal is

$$x(t;\underline{u}) = \sum_{n=-\infty}^{\infty} x_n \sqrt{E}\, \phi_n(t) \qquad (II.3.26)$$

where E is again the energy per BPSK pulse. We emphasize the dependence of the transmitted signal on the total past of the data sequence $\underline{u}$.

3.4.1 Differential Decoder #1

Assume the receiver does not estimate the signal phase but has some arbitrary phase reference 0. Then, it has an orthonormal basis set

$$\phi_c(t;0) = \sqrt{\frac{2}{T}} \cos(\omega_0 t - 0)$$

$$\phi_s(t;0) = -\sqrt{\frac{2}{T}} \sin(\omega_0 t - 0); \quad 0 \le t \le T$$

and

$$\phi_{n,c}(t;0) \triangleq \phi_c(t - nT;0)$$

$$\phi_{n,s}(t;0) \triangleq \phi_s(t - nT;0); \; n = \ldots, -1, 0, 1, \ldots \qquad (II.3.27)$$

The corresponding orthonormal basis set for the transmitted signal x(t) would be

$$\phi_n(t) = \cos 0\, \phi_{n,c}(t;0) + \sin 0\, \phi_{n,s}(t;0) \qquad (II.3.28)$$

which of course is independent of 0 and from (II.3.26) allows this signal to be written in the in-phase-quadrature form

$$x(t;\underline{u}) = \sum_{n=-\infty}^{\infty} x_n \sqrt{E} \cos 0\, \phi_{n,c}(t;0) + \sum_{n=-\infty}^{\infty} x_n \sqrt{E} \sin 0\, \phi_{n,s}(t;0) \qquad (II.3.29)$$

Thus, since the transmitted signal is representable in terms of the receiver's orthonormal basis, there is no loss of optimality in expressing the received signal $y(t)$ in terms of the sequence $\underline{y} = (y_{n,c}, y_{n,s})$ where

$$y_{n,c} \triangleq \int_{nT}^{(n+1)T} y(t)\phi_{n,c}(t;\theta)dt$$

$$= x_n \sqrt{E} \cos\theta + n_{n,c}$$

$$y_{n,s} \triangleq \int_{nT}^{(n+1)T} y(t)\phi_{n,s}(t;\theta)dt$$

$$= x_n \sqrt{E} \sin\theta + n_{n,s} \quad \text{for all } n \qquad (11.3.30)$$

where $n_{n,c}$ and $n_{n,s}$ are independent Gaussian random variables with zero mean and variance $N_0/2$.

If θ were known at the receiver then the maximum-likelihood demodulator would choose the data sequence $\underline{u}$ that results in the encoded sequence $\underline{x}$ which minimizes

$$\|\underline{y} - \underline{x}\|^2 = \sum_{n=-\infty}^{\infty} \left[(y_{n,c} - x_n \sqrt{E}\cos\theta)^2 + (y_{n,s} - x_n \sqrt{E}\sin\theta)^2\right] \qquad (11.3.31)$$

Since θ is indeed unknown at the receiver, as per our previous assumption, we must instead seek a suboptimum demodulator. In this regard, consider the terms

$$y_{n,c}\, y_{n-1,c} = (x_n \sqrt{E}\cos\theta + n_{n,c})\,(x_{n-1}\sqrt{E}\cos\theta + n_{n-1,c})$$

$$= x_n x_{n-1} E\cos^2\theta + n_{n,c}\, n_{n-1,c}$$

$$+ x_n n_{n-1,c}\sqrt{E}\cos\theta + x_{n-1} n_{n,c}\sqrt{E}\cos\theta$$

$$= u_n E\cos^2\theta + \hat{n}_{n,c} \qquad (11.3.32)$$

where we have made use of the relation (II.3.25) and

$$\hat{n}_{n,c} = n_{n,c}\, n_{n-1,c} + x_n\, n_{n-1,c}\, \sqrt{E}\cos\theta + x_{n-1}\, n_{n,c}\, \sqrt{E}\cos\theta \qquad \text{(II.3.33)}$$

Similarly,

$$y_{n,s}\, y_{n-1,s} = u_n\, E \sin^2\theta + \hat{n}_{n,s} \qquad \text{(II.3.34)}$$

where

$$\hat{n}_{n,s} = n_{n,s}\, n_{n-1,s} + x_n\, n_{n-1,s}\, \sqrt{E}\sin\theta + x_{n-1}\, n_{n,s}\, \sqrt{E}\sin\theta \qquad \text{(II.3.35)}$$

Thus, adding (II.3.32) and (II.3.34),

$$y_{n,c}\, y_{n-1,c} + y_{n,s}\, y_{n-1,s} = u_n\, E + \hat{n}_{n,c} + \hat{n}_{n,s} \qquad \text{(II.3.36)}$$

We see that with no knowledge of the reference phase θ, one can compute a test statistic that contains the signal $u_n E$ together with additive zero mean, signal dependent noise. A suboptimum demodulator based on this statistic is shown in Figure 6 where the decision rule

$$\hat{u}_n = \text{sgn}\,[y_{n,c}\, y_{n-1,c} + y_{n,s}\, y_{n-1,s}] \qquad \text{(II.3.37)}$$

has been applied to decode the original data sequence $\underline{u}$. A generalization of Figure 6 to apply to DMPSK is shown in Figure 7.

3.4.2 Differential Decoder #2

Suppose that during the time interval $-T \le t \le 0$, the transmitted encoded symbol was x_{-1}. Then, from (II.3.25) the next encoded symbol would be

$$x_0 = x_{-1} u_0 \qquad 0 \le t \le T \qquad \text{(II.3.38)}$$

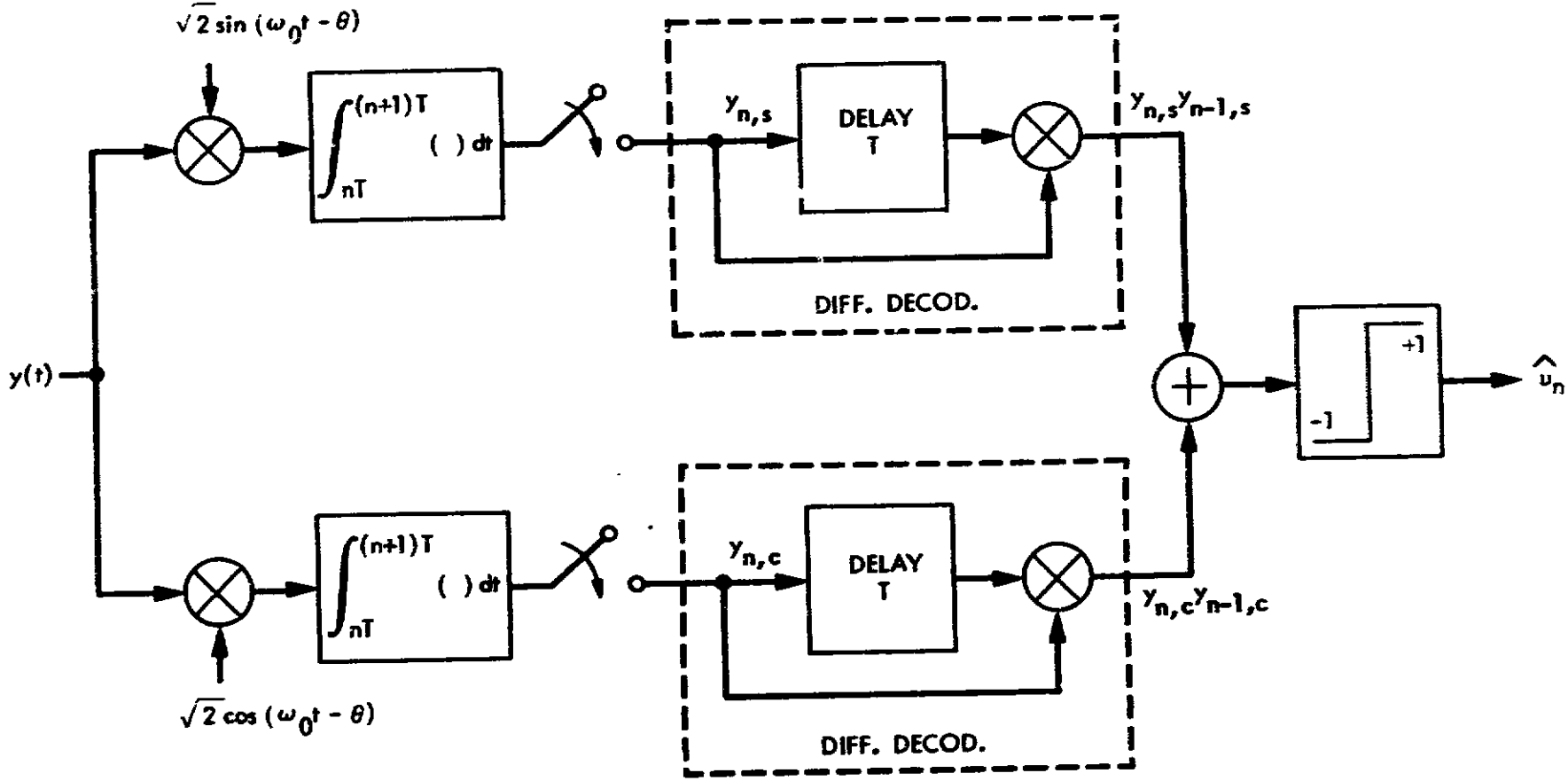


Figure 6. A DBPSK Receiver

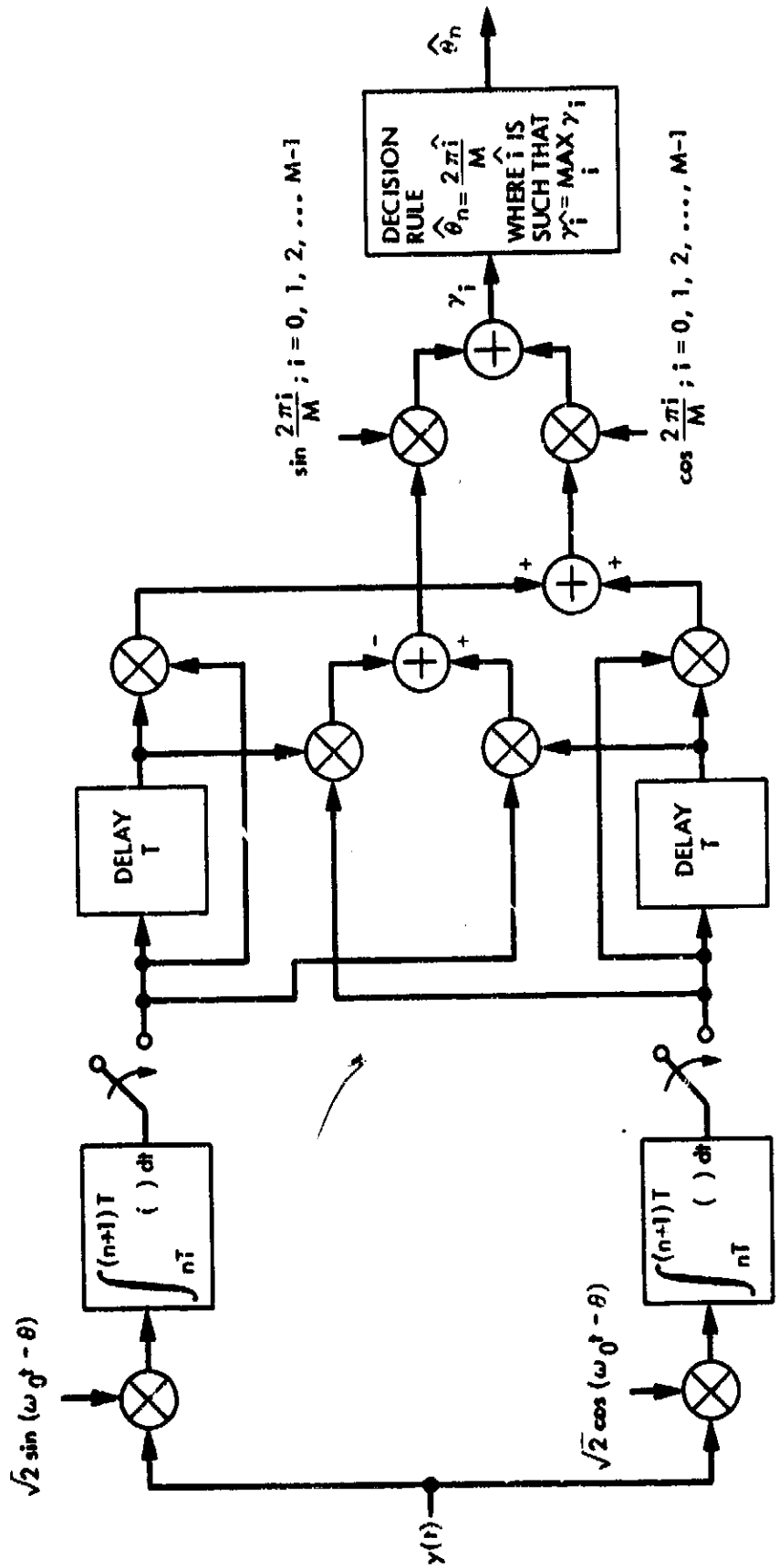


Figure 7. A DMPSK Receiver

where u_0 can be either 1 or -1. Assuming x_{-1} is fixed, the two possible pairs of encoded symbols (x_{-1}, x_0) are $(x_{-1}, 1)$ and $(x_{-1}, -1)$ and their inner product is given by

$$x_{-1}^2 - 1 = 1 - 1 = 0 \tag{11.3.39}$$

i.e., they are orthogonal to one another. Thus if a receiver is attempting to choose between $u_0 = 1$ and $u_0 = -1$ given only the observation of $y(t)$ during the interval $-T \leq t \leq T$ and without knowledge of the phase reference 0, it has the problem of noncoherently choosing between two orthogonal signals. This is essentially the same as the noncoherent BFSK receiver where the receiver must compare the energies in the two orthogonal signal coordinates. Thus, in general, another suboptimum demodulator computes for $u_n = 1$ the energy term

$$(y_{n,c} + y_{n-1,c})^2 + (y_{n,s} + y_{n-1,s})^2 \tag{11.3.40}$$

and for $u_n = -1$ the energy term

$$(-y_{n,c} + y_{n-1,c})^2 + (-y_{n,s} + y_{n-1,s})^2 \tag{11.3.41}$$

and decide $\hat{u}_n$ according to the larger of the two. A demodulator that implements this decision rule is illustrated in Figure 8.

3.4.3 Differential Decoder #3

Coherent demodulation such as for MPSK and 16-QAM discussed earlier requires some means of estimating the transmitted carrier phase. This is typically done with a Costas loop. Such carrier synchronization is not required for noncoherent MFSK and the above two suboptimum DBPSK demodulators. We now examine another DBPSK demodulator based on an approximation to coherent demodulation that results in open loop estimates of the carrier phase. With ever increasing speeds and complexity of digital processing, this approach may become increasingly important, particularly as applied to the newer bandwidth efficient modulation techniques. A more detailed discussion of these applications will follow later on.

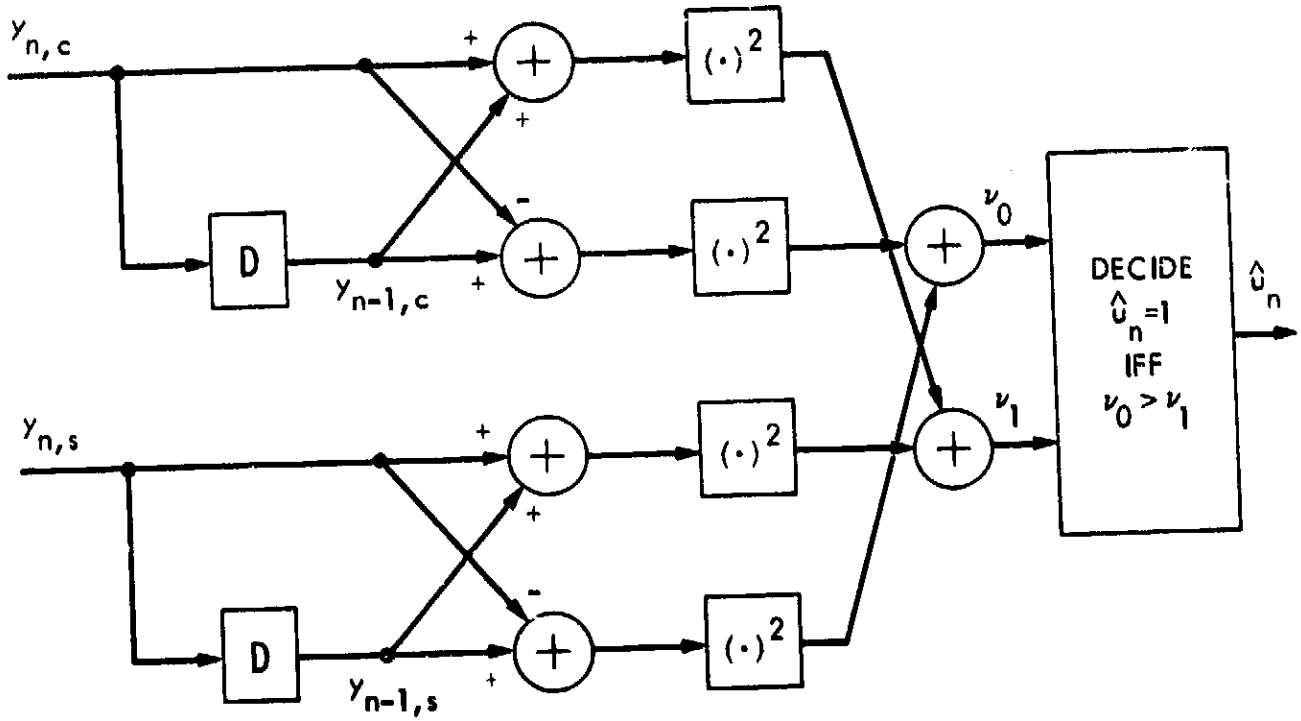


Figure 8. A DBPSK Receiver Based on Noncoherent Demodulation of Orthogonal Signals

Suppose that for DBPSK, we assume the unknown carrier phase θ is part of the data we wish to estimate. Let us quantize the phase space $0 \le \theta \le 2\pi$ into L values $\theta_1, \theta_2, \ldots, \theta_L$. To emphasize this formulation, we now rewrite the transmitted signal of (II.3.29) in the form

$$x(t;\underline{u},\theta_\ell) = \sum_{n=-\infty}^{\infty} x_n \sqrt{E}\left[\cos\theta_\ell\ \phi_{n,c}(t;\theta_\ell) + \sin\theta_\ell\ \phi_{n,s}(t;\theta_\ell)\right];$$

$$\ell = 1, 2, \ldots, L \qquad \text{(II.3.42)}$$

keeping in mind that the true carrier phase is indeed 0. As usual, the receiver computes $y_{n,c}$ and $y_{n,s}$ of (II.3.30) for all n. Without loss of optimality, all decisions can be based on these basis components.

Suppose for the moment we limit ourselves only to n = 1, 2, ..., N. Then we have 2^N possible sequences $\underline{x}$ and L possible quantized phase angles $\theta_1, \theta_2, \ldots, \theta_L$. Assuming there are then $M = L \times 2^N$ possible transmitted sequences $\underline{x}(\theta)$ with pairs of components

$$x_n \sqrt{E}\begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix};\ n = 1, 2, \ldots, N \qquad \text{(II.3.43)}$$

we have the maximum-likelihood decision rule*:

Choose $\hat{\underline{x}}$ and $\hat{\theta}$ that maximizes the inner product

$$(y, x(\theta)) = \sum_{n=1}^{N} y_{n,c}\ (x_n \sqrt{E}\cos\theta) + y_{n,s}\ (x_n \sqrt{E}\sin\theta)$$

$$= \sum_{n=1}^{N} x_n \sqrt{E}\left[y_{n,c}\cos\theta + y_{n,s}\sin\theta\right] \qquad \text{(II.3.44)}$$

over all 2^N sequences x and L phase angles θ.

*Here all signals are assumed to have the same energy.

Note that for a given fixed $\theta = \theta_\ell$ the best estimate for $\underline{x}$ is obtained from the bit by bit decision rule

$$\hat{x}_n = \text{sgn}\left[y_{n,c} \cos \theta_\ell + y_{n,s} \sin \theta_\ell\right] \qquad \text{(II.3.45)}$$

Because of this we can separate the estimates of the data $\underline{x}$ and phase θ into a two step demodulation structure. Thus we have the approximate maximum-likelihood sub-decoder shown in Figure 9. This sub-decoder must be duplicated L times, once for each quantized angle θ_ℓ. The output of the sub-decoder with the largest value of $[\underline{y},\hat{\underline{x}}(\theta_\ell)]$; $\ell = 1, 2, \ldots, L$ is then chosen as the desired output which consists now of both a decoded data sequence $\hat{\underline{u}}$ and a phase estimate $\hat{\theta}$.

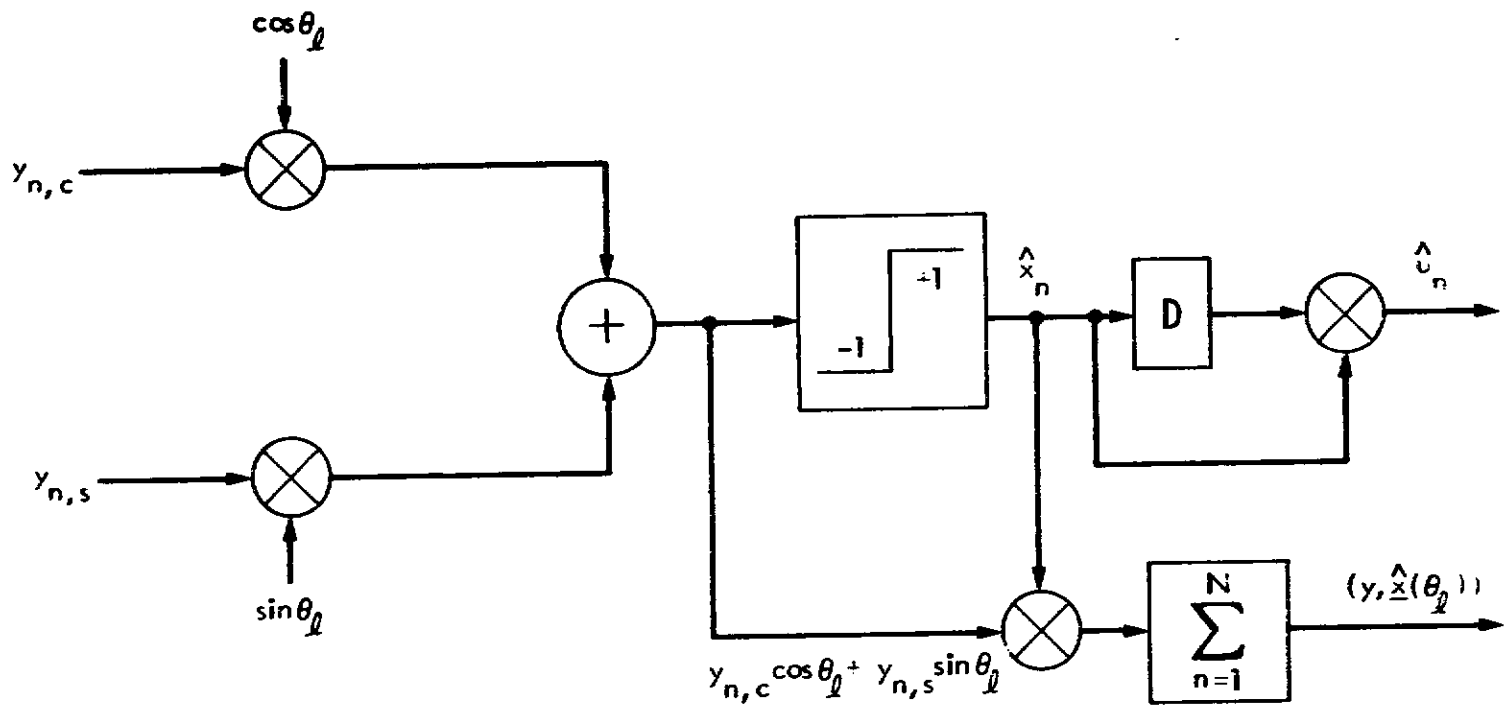
We have shown a form of the decoder that "blocks" the received sequence into N pulses. A continuous version of this decoder can be easily constructed. Our purpose in introducing this joint phase and data estimation demodulation structure here is to motivate looking at similar phase and data demodulation structures for more complex bandwidth efficient modulations where it might not be obvious how one would design the usual closed-loop type of carrier phase estimator. Also, to re-emphasize a previous statement, with new digital processing capabilities these phase and data open loop demodulators may be a cost effective alternative to having a separate closed loop phase estimator, especially in systems with burst mode operation such as TDMA or packet radios where fast acquisition is important.

3.5 Coherent BPSK with Intersymbol Interference

We now return our attention to coherent BPSK modulation to introduce use of the Viterbi algorithm [1] as a maximum-likelihood demodulator when intersymbol interference (ISI) is present at the receiver. For convenience in what follows, we shall use the nonoverlapping and therefore orthogonal set of pulses defined by (II.3.22) as the basis for representing the coherent BPSK signal in the absence of ISI. Thus, analogous to (II.3.26), we have

$$x(t;\underline{u}) = \sum_{n=-\infty}^{\infty} u_n \sqrt{E}\, \phi_n(t) \qquad \text{(II.3.46)}$$

where again the dependence on the i.i.d. data sequence $\underline{u}$ is to be emphasized.



$$\hat{\underline{x}}(\hat{\theta}) = \max_{\underline{x}_m(\theta_\ell)}{}^{-1}(y, \underline{x}_m(\theta_\ell)) = \max_{\hat{\underline{x}}(\theta_\ell)}{}^{-1}(y, \hat{\underline{x}}(\theta_\ell))$$

$$\hat{\theta} = \max_{\theta_\ell}{}^{-1}(y, \hat{\underline{x}}(\theta_\ell))$$

NOTE: THE NOTATION "$\hat{a} = \max_{a_i}{}^{-1} f(a_i)$" MEANS $\hat{a}$ IS THAT VALUE OF a_i THAT MAXIMIZES $f(a_i)$

Figure 9. One Element of a DBPSK Receiver Based on Simultaneous Data and Phase Estimation

RF filtering, channel multipath, and other channel propagation characteristics can often lead to intersymbol interference where the above BPSK signal arrives at the receiver in the form

$$x(t;\underline{u}) = \sum_{n=-\infty}^{\infty} u_n \, h(t-nT) \qquad \text{(II.3.47)}$$

where h(t) may span several symbol times. In Figure 10 we show an RF filter causing the input pulse $\phi_n(t) = \phi_c(t-nT)$ to appear at the filter output as h(t-nT).

Recall that the maximum-likelihood decision rule for selecting the data sequence is to choose $\underline{\hat{u}}$ that maximizes

$$\int_{-\infty}^{\infty} y(t)x(t;\underline{u})dt - 1/2 \int_{-\infty}^{\infty} x^2(t;\underline{u})dt \qquad \text{(II.3.48)}$$

Defining

$$y_n = \int_{-\infty}^{\infty} y(t) \, h(t-nT)dt \qquad \text{(II.3.49)}$$

and

$$h_{m-n} = \int_{-\infty}^{\infty} h(t-nT) \, h(t-mT)dt \quad \text{for all } m,n \qquad \text{(II.3.50)}$$

we have

$$\int_{-\infty}^{\infty} y(t)x(t;\underline{u})dt = \sum_{n=-\infty}^{\infty} u_n y_n \qquad \text{(II.3.51)}$$

$$x(t;\underline{u}) = \sum_{n=-\infty}^{\infty} u_n h(t-nT)$$

$$u_n \in \{-1, 1\}$$

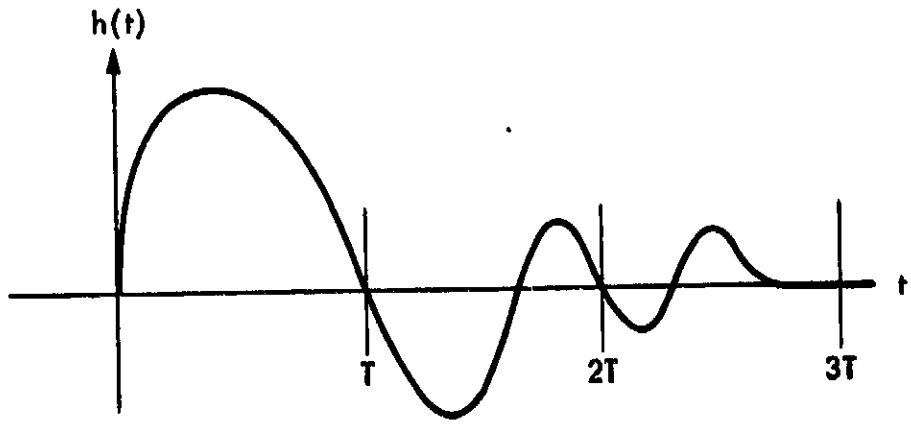


Figure 10. A Transmitter Model for BPSK with ISI

and

$$\int_{-\infty}^{\infty} x^2(t;\underline{u})dt = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} u_m u_n h_{m-n} \qquad (11.3.52)$$

Next, we assume that the intersymbol interference has finite span L. That is,

$$h_i = 0 \; ; \quad |i| \geq L \qquad (11.3.53)$$

Thus, using (11.3.53) in (11.3.52), we obtain*

$$\int_{-\infty}^{\infty} x^2(t;\underline{u})dt = \sum_{n=-\infty}^{\infty} u_n^2 h_0 + \sum_{m=-\infty}^{\infty} u_m \sum_{n=m-L}^{m-1} u_n h_{m-n} + \sum_{n=-\infty}^{\infty} u_n \sum_{m=n-L}^{n-1} u_m h_{m-n}$$

$$= \sum_{n=-\infty}^{\infty} h_0 + \sum_{m=-\infty}^{\infty} u_m \sum_{i=1}^{L-1} u_{m-i} h_i + \sum_{n=-\infty}^{\infty} u_n \sum_{i=1}^{L-1} u_{n-i} h_{-i}$$

$$= \sum_{n=-\infty}^{\infty} \left(h_0 + 2u_n \sum_{i=1}^{L-1} u_{n-i} h_i \right) \qquad (11.3.54)$$

and from (11.3.48) the maximum-likelihood decision rule becomes:

Choose the sequence $\hat{\underline{u}}$ that maximizes the sum

$$\sum_{n=-\infty}^{\infty} \left(u_n y_n - \frac{h_0}{2} - u_n \sum_{i=1}^{L-1} u_{n-i} h_i \right)$$

*Note that because of the symmetry in the definition of h_{m-n} given in (11.3.50), we have $h_i = h_{-i}$ for all i.

We now show that the maximum-likelihood decision rule can be implemented using the Viterbi algorithm and we refer to the resulting demodulator as a Viterbi demodulator. First define the signal sequence state s_n for the nth interval as the sequence of L-1 previous data bits, i.e.,

$$s_n = \left(u_{n-1}, u_{n-2}, \ldots, u_{n-(L-1)}\right) \quad \text{for all } n \tag{11.3.55}$$

Here we see that the state s_{n+1} for the n+1st interval can be formed from s_n and u_n by merely shifting the symbols in s_n one step to the right and placing u_n in the 1st position. This is represented as

$$s_{n+1} = g(s_n, u_n) \quad \text{for all } n \tag{11.3.56}$$

Thus, since the transmitted signal is given by (11.3.47), then for each interval indexed by n,

$$\begin{aligned} x_n &= \int_{-\infty}^{\infty} x(t;\underline{u})h(t-nT)dt \\ &= \sum_{m=-\infty}^{\infty} u_m \int_{-\infty}^{\infty} h(t-mT)h(t-nT)dt \\ &= \sum_{m=-\infty}^{\infty} u_m h_{n-m} \\ &= \sum_{i=0}^{L-1} u_{n-i} h_i \\ &= u_n h_0 + \sum_{i=1}^{L-1} u_{n-i} h_i \\ &= u_n h_0 + (h, s_n) \end{aligned} \tag{11.3.57}$$

where $\underline{h} \triangleq (h_1, h_2, \ldots, h_{L-1})$. Clearly, from (II.3.57), the signal components have the form

$$x_n = f(s_n, u_n) \tag{II.3.58}$$

and the signal process then has a finite state description.

The channel output as described by (II.3.49) can be expressed in terms of its signal and noise components, i.e.,

$$y_n = x_n + n_n \quad \text{for all } n \tag{II.3.59}$$

where

$$x_n = \int_{-\infty}^{\infty} x(t;\underline{u})h(t-nT)dt = \sum_{m=-\infty}^{\infty} u_m h_{m-n} \tag{II.3.60}$$

and

$$n_n = \int_{-\infty}^{\infty} n(t)h(t-nT)dt \tag{II.3.61}$$

Note that here the additive noise components $\{n_n\}$ are not independent. Finally defining the metric

$$m(y_n;u_n,s_n) = u_n y_n - \frac{h_0}{2} - u_n(\underline{h},s_n) \tag{II.3.62}$$

we see that the maximum-likelihood rule is to find the data sequence $\hat{\underline{u}}$ or equivalently the state sequence $\hat{\underline{s}}$ that maximizes the accumulated metric

$$\sum_{n=-\infty}^{\infty} m(y_n;u_n,s_n)$$

For this formulation of the finite state signal description with the additive metric, the Viterbi algorithm is known to be a solution. This is shown in detail in Appendix A. We show here a simple example where L = 2.

Example: $L = 2$

$$x_n = a_n h_0 + a_{n-1} h_1$$
$$= a_n h_0 + s_n h_1 \qquad (11.3.63)$$

The metric of (11.3.62) is

$$m(y_n; a_n, s_n) = a_n y_n - \frac{1}{2} h_0 - a_n s_n h_1$$
$$= a_n y_n - \frac{1}{2} h_0 - a_n a_{n-1} h_1 \qquad (11.3.64)$$

Since $s_n = a_{n-1}$ can have only two possible values we have a two state Viterbi demodulator which is shown in Figure 11 together with its corresponding trellis diagram. The branches of the trellis diagram are labelled with the values of the signal components (x_n) and the data bit corresponding to each state transition.

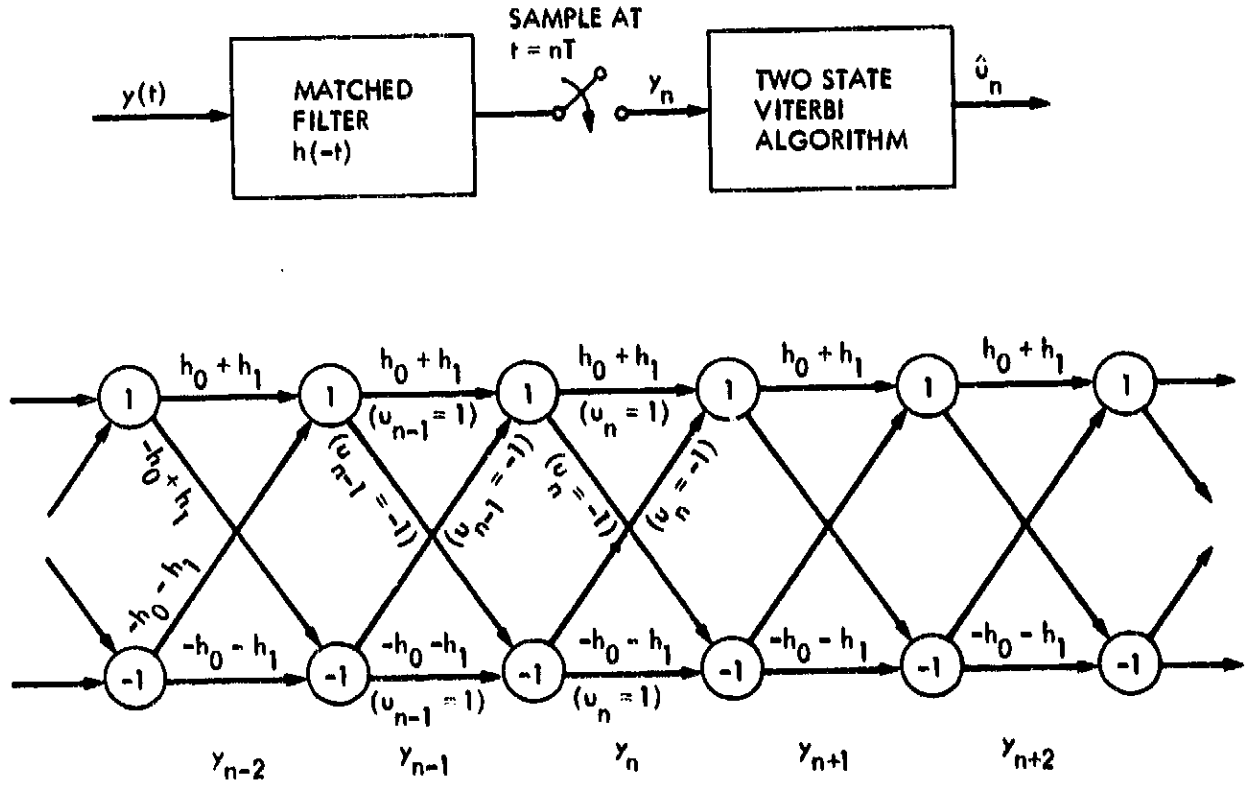


Figure 11. A Two State Viterbi Demodulator and its Trellis Diagram for BPSK with ISI

It is important to note that, in deriving the maximum-likelihood rule, we did not use any orthonormal basis but instead used the coordinate-free form of the rule. This resulted in taking as the channel outputs the matched filter output samples of (II.3.49). The noise components of these samples, as defined by (II.3.61), have the correlation function

$$E\{n_k n_j\} = \frac{N_0}{2} h_{k-j} \quad \text{for all } k,j \tag{II.3.65}$$

where we have made use of (II.3.50). Note that, despite the fact that the additive noise is correlated in the filter sample outputs, the Viterbi demodulator is still optimum in that it performs the maximum-likelihood decision rule.

4.0 Bandwidth Efficient Modulations

One approach to developing bandwidth efficient modulations is to take a conventional modulation and then use RF filtering to shape the signal spectrum. This is typically done to reduce out-of-band signal energy. On the other hand, as we have previously seen, this can also result in intersymbol interference which degrades system performance, particularly when a conventional demodulator designed for the original unfiltered modulation is employed. Using a maximum-likelihood Viterbi demodulator in this case can often dramatically improve the bit error rate performance.

In view of the above, it makes sense to consider the possibility of changing the modulation itself to shape the signal spectrum rather than use RF filtering on conventional modulations. In this section, we first examine some modulations that are used in practice today that attempt to accomplish this. Following this, new advanced bandwidth efficient modulations are defined and examined in detail.

4.1 Controlled ISI: Duobinary and Partial Response Signals

Duobinary signalling, first introduced by Lender [2], is an attempt to control the signal intersymbol interference at the modulator in some known way so as to minimize the unknown intersymbol interference caused by a telephone channel. The various generalizations of this modulation are referred to as partial response signals [3,4].

A duobinary modulator is shown in Figure 12. The i.i.d. data sequence $\{u_n\}$ is first differentially encoded to produce the i.i.d. sequence $\{z_n\}$ where

$$z_n = z_{n-1} u_n \quad \text{for all } n \tag{II.4.1}$$

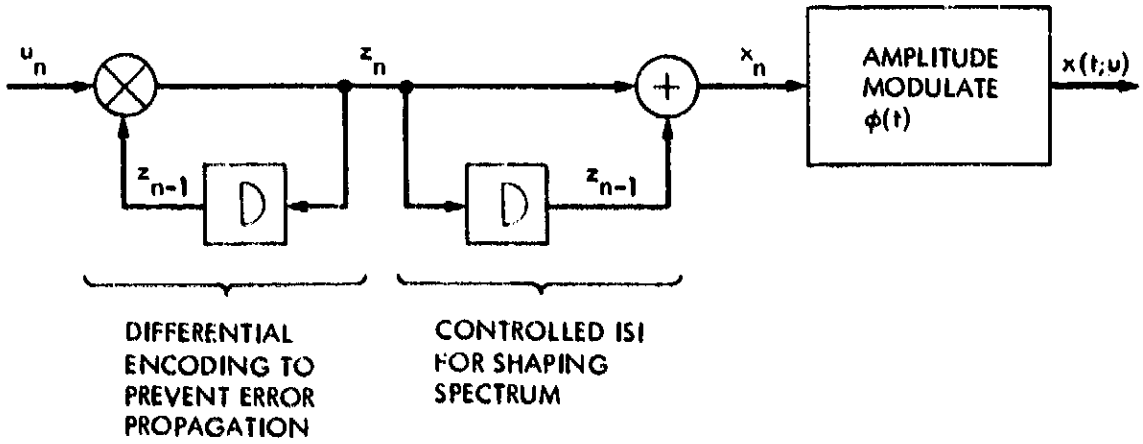
This encoding operation is necessary to prevent error propagation as will become evident in the demodulation structure which we examine next.

The duobinary sequence $\{x_n\}$ is then obtained from $\{z_n\}$ by simply adding z_n to a one bit delayed version of itself, i.e.,

$$x_n = z_n + z_{n-1} \quad \text{for all } n \tag{II.4.2}$$

Note that $\{x_n\}$ is a correlated three-level (-2,0,2) sequence where

$$\begin{aligned} \Pr\{x_n = 2\} &= \Pr\{x_n = -2\} = 1/4 \\ \Pr\{x_n = 0\} &= 1/2 \end{aligned} \tag{II.4.3}$$



$$x(t;u) = \sum_{n=-\infty}^{\infty} x_n \sqrt{E}\, \phi(t-nT)$$

$$u_n \in \{-1,1\}$$

Figure 12. A Duobinary Modulator

Finally, the transmitted duobinary signal $x(t;\underline{u})$ is formed by amplitude modulating the pulse shape sequence $\{\phi(t-nT)\}$ with $\{x_n\}$ resulting in

$$x(t;\underline{u}) = \sum_{n=-\infty}^{\infty} x_n \sqrt{E}\, \phi\,(t-nT) \tag{II.4.4}$$

Here the amplitude modulated pulse $\phi(t)$ has the property

$$\int_{-\infty}^{\infty} \phi(t-nT)\phi(t-mT)dt = \delta_{mn} \tag{II.4.5}$$

That is, shifts of the pulse are orthogonal even though the pulse is no longer restricted to lie in a T-second interval. Typically $\phi(t)$ is chosen to have spectral characteristics matched to those of the channel. This pulse can also represent the combined controlled ISI pulse and the channel filter impulse response.

Transmitting $x(t;u)$ over the AWGN channel results in the channel output

$$y(t) = x(t;\underline{u}) + n(t) \tag{II.4.6}$$

where again it is convenient to take as our basis

$$\phi_n(t) = \phi(t-nT) \text{ for all } n \tag{II.4.7}$$

Without loss of generality, we again also use as channel outputs the components

$$\begin{aligned} y_k &= \int_{-\infty}^{\infty} y(t)\phi_n(t)dt \\ &= x_k\sqrt{E} + n_k \end{aligned} \tag{II.4.8}$$

where $\{n_k\}$ is an i.i.d. sequence of zero mean Gaussian random variables with variance $N_0/2$.

Note that, from (11.4.1) and (11.4.2), we have the signal term

$$\begin{aligned} x_n &= z_{n-1}u_n + z_{n-1} \\ &= z_{n-1}(1 + u_n) \end{aligned} \tag{11.4.9}$$

Defining the state at time n to be

$$s_n = z_{n-1}, \tag{11.4.10}$$

then, analogous to (11.3.56) and (11.3.58), we have the general forms

$$\begin{aligned} s_{n+1} &= z_n \\ &= z_{n-1}u_n \\ &= g(s_n, u_n) \end{aligned} \tag{11.4.11}$$

and

$$\begin{aligned} x_n &= s_n(1 + u_n) \\ &= f(s_n, u_n) \end{aligned} \tag{11.4.12}$$

Since the maximum-likelihood receiver maximizes the function

$$\left(y, x(u)\right) - \frac{1}{2}\,\|x(u)\|^2 = \sum_{n=-\infty}^{\infty}\left[y_n x_n\sqrt{E} - \frac{1}{2}x_n^2 E\right] \tag{11.4.13}$$

we have the metric

$$m(y_n; x_n) = y_n x_n\sqrt{E} - \frac{1}{2}x_n^2 E \text{ for all } n \tag{11.4.14}$$

or from (11.4.12),

$$m(y;u_n,s_n) = y_n f(s_n,u_n)\sqrt{E} - \frac{1}{2}f^2(s_n,u_n)E \qquad (11.4.15)$$

With this formulation it is evident that the maximum-likelihood demodulator can be realized by the Viterbi algorithm. This is shown in Figure 13 where the branches are labelled with the values of the signal terms $\{x_n\}$.

Note from Figure 13 that the two shifted alternating state sequences

$$\cdots\cdot 1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1 \quad \cdots\cdot$$

$$\cdots\cdot -1 \quad 1 \quad -1 \quad 1 \quad -1 \quad 1 \quad -1 \quad 1 \quad \cdots\cdot$$

both yield $x_n = 0$ for all n and thus are indistinguishable at the demodulator. Without the differential encoder the state sequence and the data sequence would be the same and thus we could have an error propagation condition if we selected the wrong shifted alternating sequence. With differential encoding both of the two shifted state alternating sequences have only one unique data sequence thus resolving any ambiguity. In Figure 13 we show the data bit below each branch corresponding to each state transition.

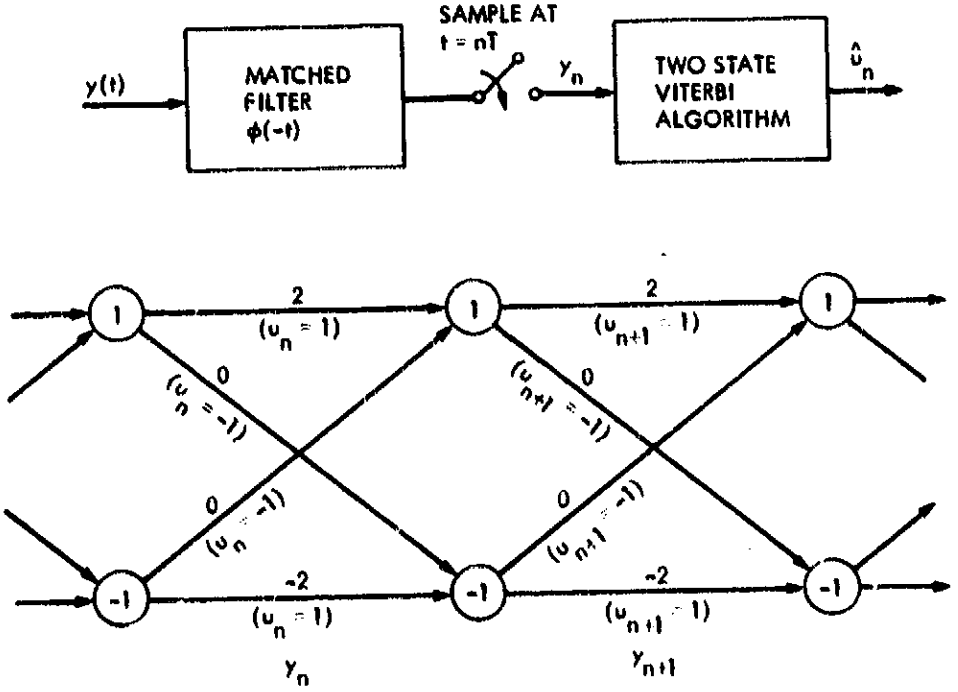


Figure 13. A Two State Viterbi Demodulator and its Trellis Diagram for Duobinary Modulation

Although the Viterbi demodulator is a maximum-likelihood demodulator, in practice, suboptimum demodulators are often used because of their simplicity. For example, note from (II.4.9) that

$$|x_n| = \begin{cases} 2; & u_n = 1 \\ 0; & u_n = -1 \end{cases} \qquad \text{(II.4.16)}$$

Hence a suboptimum decision rule based on a single channel output sample could be:

For each n, choose $\hat{u}_n = 1$ if and only if $|y_n| \geq \eta\sqrt{E}$.

The implementation of this decision rule as in Figure 14a is the conventional receiver used for demodulation of duobinary signals where η is a normalized threshold whose value satisfies $0 < \eta < 2$.

A suboptimum decision rule based on two successive channel output observations can be obtained by considering the sum $2y_n + y_{n+1}$. In particular, from (II.4.8) together with (II.4.1) and (II.4.2), we have

$$\begin{aligned} 2y_n + y_{n+1} &= (3z_n + 2z_{n-1} + z_{n+1})\sqrt{E} + 2n_n + n_{n+1} \\ &= z_{n-1}\underbrace{(3u_n + 2 + u_n u_{n+1})}_{w_n}\sqrt{E} + 2n_n + n_{n+1} \end{aligned} \qquad \text{(II.4.17)}$$

Thus, since

$$|w_n| = \begin{cases} 4 \text{ or } 6; & u_n = 1 \\ 0 \text{ or } 2; & u_n = -1 \end{cases}$$

where the pair-wise choice of values for $|w_n|$ occurs with equal probability, then an analogous suboptimum decision rule to that given above would be:

For each n, choose $\hat{u}_n = 1$ if and only if $|2y_n + y_{n+1}| \geq \eta\sqrt{E}$

where now η is chosen to satisfy $2 < \eta < 4$. This decision rule has the implementation illustrated in Figure 14b.

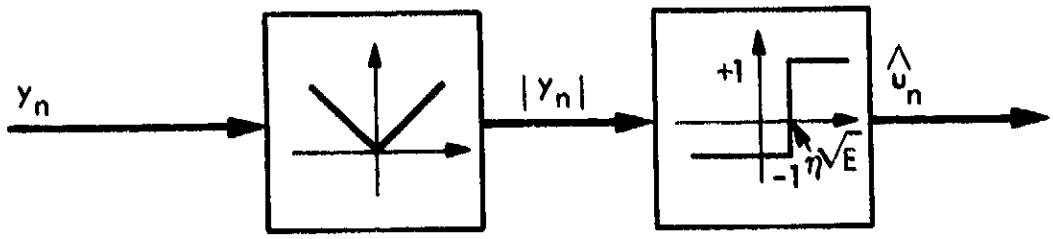


Figure 14a. A Suboptimum Receiver for Duobinary Signals Based Upon a Single Channel Output Observation

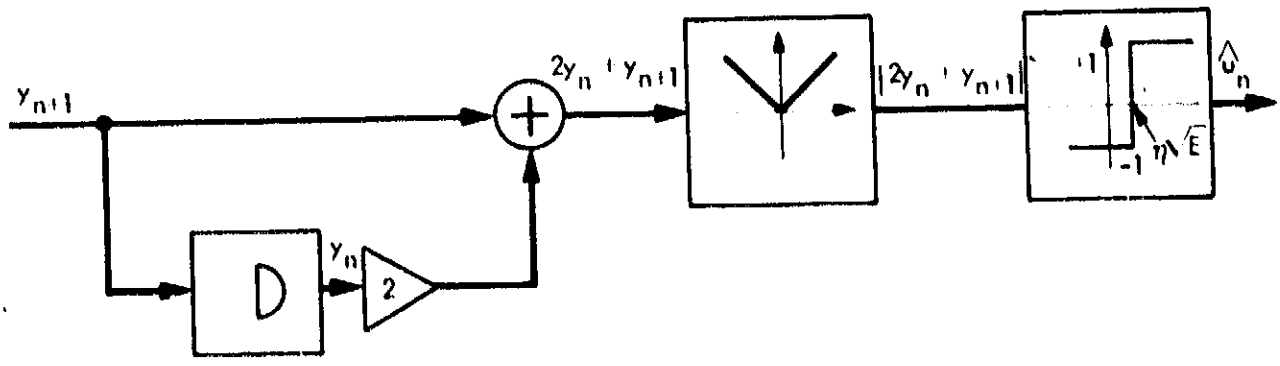


Figure 14b. A Suboptimum Receiver for Duobinary Signals Based Upon Two Successive Channel Output Observations

4.2 MSK

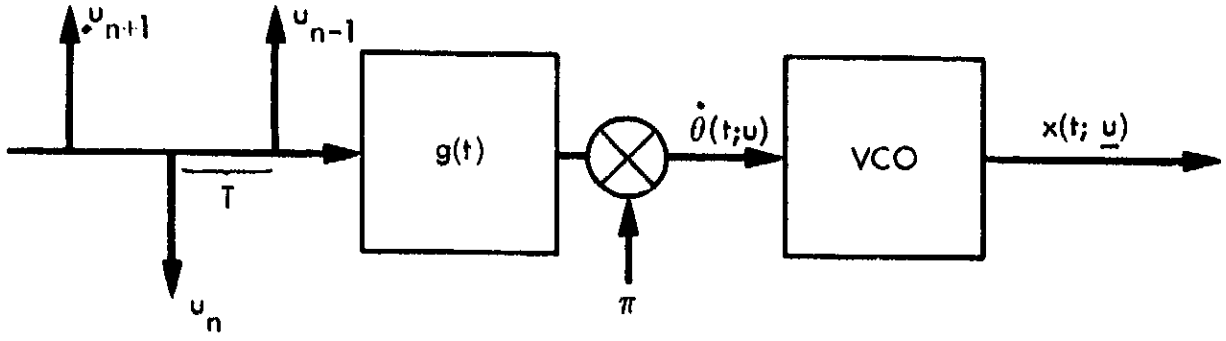
Minimum shift keying (MSK) is the simplest form of a new class of bandwidth efficient modulation techniques referred to as continuous phase frequency shift keying (CPFSK), or, more generally, continuous phase modulation (CPM). In this section, we discuss the transmission of MSK over the linear channel with particular emphasis on the realization of the maximum-likelihood demodulator by the Viterbi algorithm. In the next section, we extend these notions to the broader class of modulations mentioned above.

For an input data sequence $\underline{u}$, the MSK signal is given by (see Figure 15).

$$x(t;\underline{u}) = \sqrt{2S}\cos[\omega_0 t + \theta(t;\underline{u})] \tag{II.4.18}$$

where

$$\theta(t;\underline{u}) = \sum_{n=-\infty}^{\infty} u_n \int_{-\infty}^{t} g(\tau - nT)d\tau \tag{II.4.19}$$



$u_n \in \{-1, 1\}$

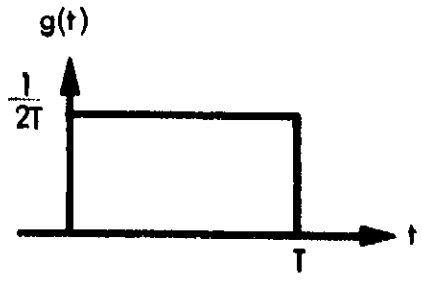


Figure 15. An MSK Modulator

and

$$g(t) = \begin{cases} \frac{1}{2T}; & 0 \leq t \leq T \\ 0; & \text{otherwise} \end{cases} \tag{II.4.20}$$

Note that g(t) is a rectangular pulse of T-second duration.

In view of (II.4.20),

$$\int_{-\infty}^{t} g(\tau - nT)d\tau = \int_{0}^{t-nT} g(\tau)d\tau = \begin{cases} 0; & t < nT \\ \frac{t-nT}{2T}; & nT \leq t \leq (n+1)T \\ \frac{1}{2}; & (n+1)T \leq t \end{cases} \tag{II.4.21}$$

Thus, for the interval $nT \leq t \leq (n+1)T$ we have

$$\theta(t;\underline{u}) = u_n \pi \left(\frac{t-nT}{2T}\right) + \frac{\pi}{2} \sum_{i=-\infty}^{n-1} u_i \tag{II.4.22}$$

We define the state s_n at the beginning of the nth time interval $nT \leq t \leq (n+1)T$ as

$$s_n = \frac{\pi}{2} \sum_{i=-\infty}^{n-1} u_i \quad \text{modulo } 2\pi \tag{II.4.23}$$

Note that s_n can have only four possible values belonging to the set

$$\mathcal{S} = \left\{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\right\} \tag{II.4.24}$$

We also can describe the state transition by

$$s_{n+1} = s_n + \frac{\pi}{2} u_n \quad \text{modulo } 2\pi \tag{II.4.25}$$

which is illustrated in Figure 16. In terms of the state s_n, the MSK phase term is then given by,

$$\theta(t;\underline{u}) = u_n \pi\left(\frac{t-nT}{2T}\right) + s_n;\ nT \le t \le (n+1)T \qquad (II.4.26)$$

Figure 17 is a sketch of the set of all possible phases, $\theta(t;\underline{u})$ for all input sequences $\underline{u}$. This portrayal of MSK clearly demonstrates its characterization as a continuous phase binary frequency shift keying.

Since $x(t;\underline{u})$ is a constant envelope signal all sequences have the same signal energy. Hence, the maximum-likelihood demodulator looks for the sequence $\hat{\underline{u}}$ that maximizes

$$\int_{-\infty}^{\infty} y(t)x(t;\underline{u})dt = \sum_{n=-\infty}^{\infty} \int_{nT}^{(n+1)T} y(t)x(t;\underline{u})dt \qquad (II.4.27)$$

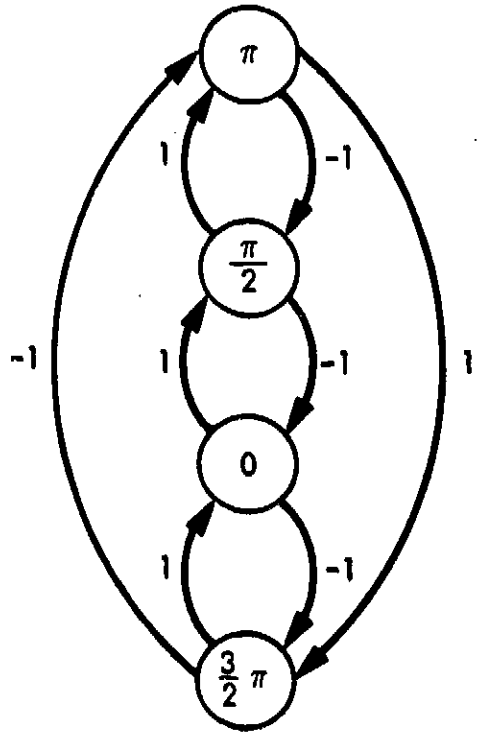


Figure 16. A State Transition Diagram for MSK

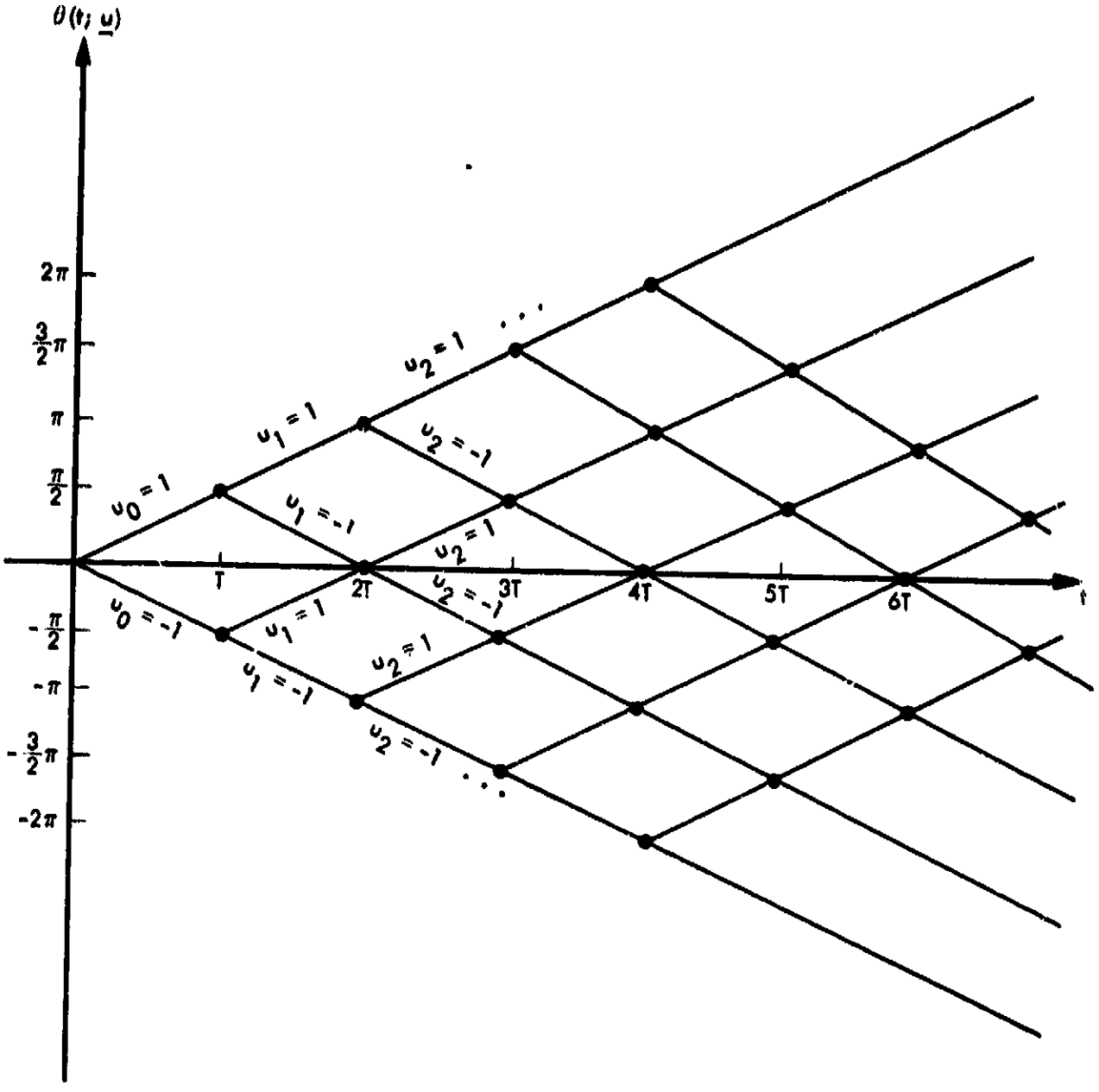


Figure 17. A Sketch of $\theta(t;\underline{u})$ for all Possible Input Sequences $\underline{u}$

Considering each sub-integral, we have

$$\int_{nT}^{(n+1)T} y(t)x(t;u)dt = \int_{nT}^{(n+1)T} y(t)\sqrt{2S}\cos\left[\omega_0 t + u_n\pi\left(\frac{t-nT}{2T}\right) + s_n\right]dt$$

$$= (\cos s_n)\int_{nT}^{(n+1)T} y(t)\sqrt{2S}\cos\left[\omega_0 t + u_n\pi\left(\frac{t-nT}{2T}\right)\right]dt$$

$$- (\sin s_n)\int_{nT}^{(n+1)T} y(t)\sqrt{2S}\sin\left[\omega_0 t + u_n\pi\left(\frac{t-nT}{2T}\right)\right]dt$$

(II.4.28)

Next, defining the integrals

$$y_{n,c}(u_n) = \int_{nT}^{(n+1)T} y(t)\sqrt{\frac{2}{T}}\cos\left[\omega_0 t + u_n\pi\left(\frac{t-nT}{2T}\right)\right]dt$$

(II.4.29)

$$y_{n,s}(u_n) = \int_{nT}^{(n+1)T} y(t)\sqrt{\frac{2}{T}}\sin\left[\omega_0 t + u_n\pi\left(\frac{t-nT}{2T}\right)\right]dt$$

then (II.4.28) becomes

$$\int_{nT}^{(n+1)T} y(t)x(t;u)dt = \sqrt{E}\, y_{n,c}(u_n)\cos s_n + \sqrt{E}\, y_{n,s}(u_n)\sin s_n$$

(II.4.30)

whereupon the total correlation of (II.4.27) simplifies to

$$\int_{-\infty}^{\infty} y(t)x(t;u)dt = \sqrt{E} \sum_{n=-\infty}^{\infty} \left[y_{n,c}(u_n) \cos s_n + y_{n,s}(u_n) \sin s_n \right] \qquad (II.4.31)$$

During the nth interval the receiver (Figure 18) computes the four-component vector

$$\underline{y}_n \triangleq \left(y_{n,c}(1),\ y_{n,c}(-1),\ y_{n,s}(1),\ y_{n,s}(-1) \right) \qquad (II.4.32)$$

and uses the metric

$$m(\underline{y}_n;\ u_n, s_n) = y_{n,c}(u_n) \cos s_n + y_{n,s}(u_n) \sin s_n \qquad (II.4.33)$$

in the four state Viterbi algorithm illustrated in Figure 19. Here, the branches are labeled with the metric values. Also note that by assuming we begin with a

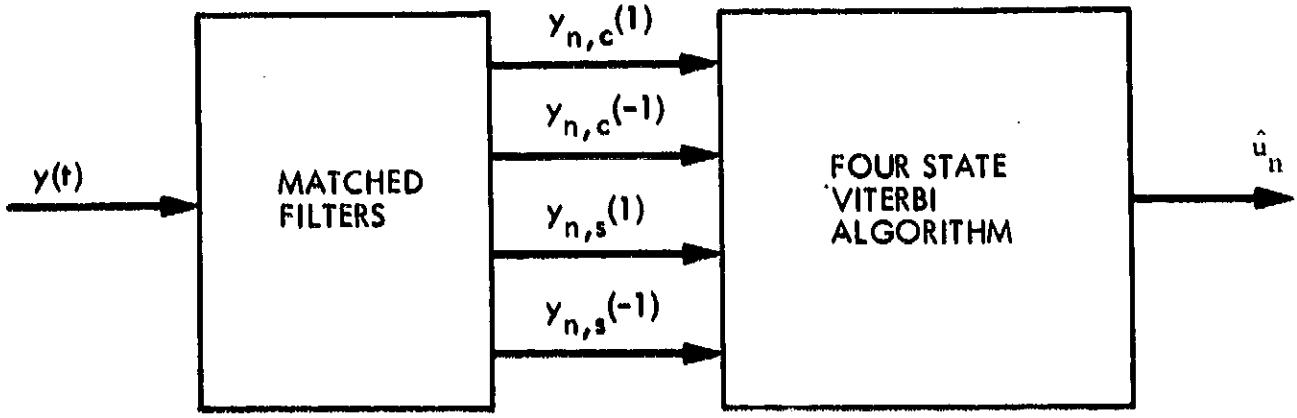


Figure 18. An MSK Receiver

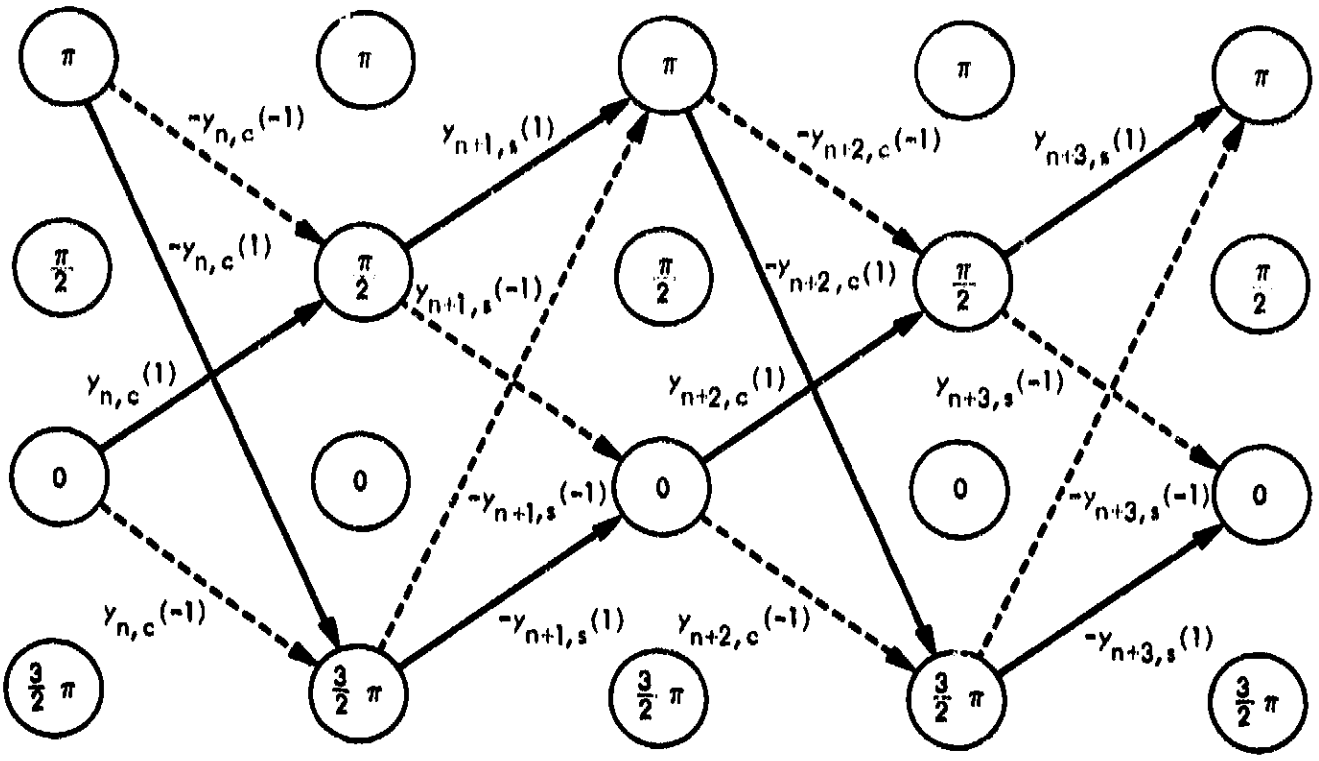


Figure 19. A Four State Trellis Diagram for Viterbi Demodulation of MSK

known initial state, this four state viterbi algorithm can be reduced to a simple two state algorithm shown in Figure 20. This reduction follows from the fact that out of four phase angles only two can be allowable states at the beginning of any time interval.

A slight modification of the MSK modulator due to Massey [5] allows us to reduce the Viterbi algorithm to an implementation with simple delay circuits and a threshold decision rule. In particular, suppose that the true data sequence is now denoted by $\{d_n\}$ where

$$\Pr(d_n = 1) = \Pr(d_n = -1) = \frac{1}{2} \qquad (11.4.34)$$

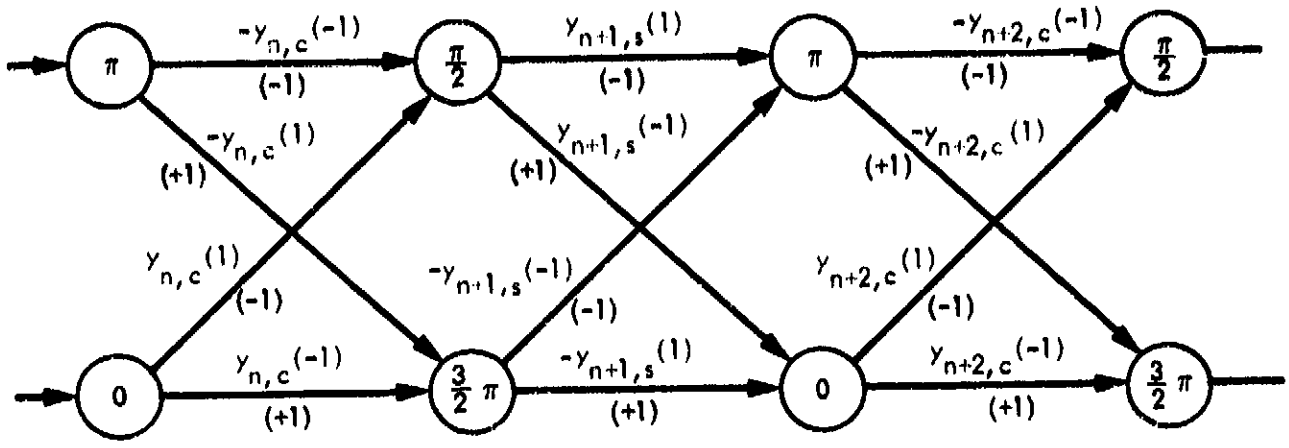


Figure 20. A Reduced (Two State) Trellis Diagram for Viterbi Demodulation of MSK

The input sequence $\{u_n\}$ to the MSK modulator of Figure 15 is obtained by a precoding (see Figure 21) of $\{d_n\}$ according to the relation

$$u_n = -d_n \left[\sin s_n + \cos s_n\right] \qquad \text{(II.4.35)}$$

where the state s_n is still defined by (II.4.23). We now examine the trellis diagram as before with the addition of a labelling below each branch to denote the value of the data bit d_n that corresponds to the particular state transition.

In the trellis diagram, we are primarily concerned with the probability of correctly choosing the correct data sequence $\{d_n\}$. Consider the nth data bit d_n and suppose that we have a "magic genie" who aids us in making a decision on this bit by telling us the states s_n and s_{n+2}. The four possible pairs of states that can be told to us by the genie are (see Figure 20):

$$\begin{aligned}
&\text{(a)} \quad s_n = \pi, \; s_{n+2} = \pi \\
&\text{(b)} \quad s_n = \pi, \; s_{n+2} = 0 \\
&\text{(c)} \quad s_n = 0, \; s_{n+2} = \pi \\
&\text{(d)} \quad s_n = 0, \; s_{n+2} = 0
\end{aligned} \qquad \text{(II.4.36)}$$

If the genie told us that the pair (a) were true, then we would have two possible paths from $s_n = \pi$ to $s_{n+2} = \pi$. One of these paths is given by

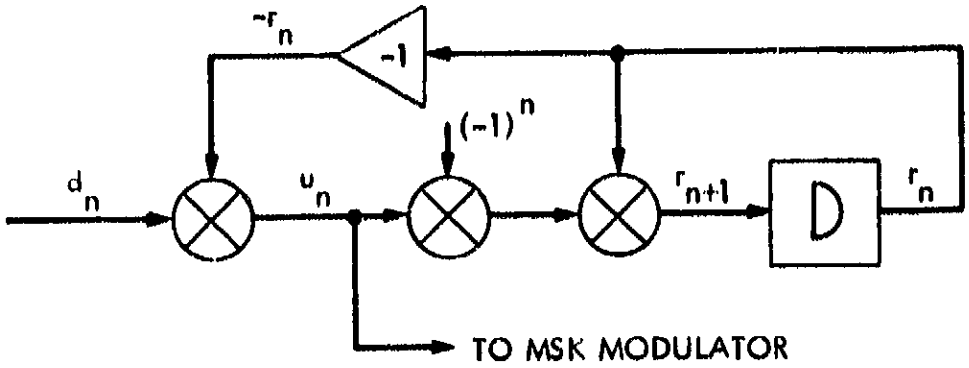


Figure 21. A Precoder for MSK

$$s_n = 0,\ s_{n+1} = \frac{\pi}{2},\ s_{n+2} = \pi \tag{11.4.37}$$

with accumulated metric [see (11.4.33)]

$$m(y_n;-1,\pi) + m(y_{n+1};1,\pi/2) = -y_{n,c}(-1) + y_{n+1,s}(1) \tag{11.4.38}$$

and data pair during the transitions

$$d_n = -1,\ d_{n+1} = -1 \tag{11.4.39}$$

The other path is given by

$$s_n = 0,\ s_{n+1} = \frac{3}{2}\pi,\ s_{n+2} = \pi \tag{11.4.40}$$

with accumulated metric

$$m(y_n;1,\pi) + m(y_{n+1};-1,3\pi/2) = -y_{n,c}(1) - y_{n+1,s}(-1) \tag{11.4.41}$$

and data pair during the transitions

$$d_n = 1,\ d_{n+1} = -1 \tag{11.4.42}$$

The best decision rule given the information provided by the genie is to pick the path with the larger accumulated metric. Hence, we would follow the rule:

$$\text{Choose } \hat{d}_n = 1 \text{ if and only if } -y_{n,c}(1) - y_{n+1,s}(-1) > -y_{n,c}(-1) + y_{n+1,s}(1)$$

If, on the other hand, the genie told us that the pair (b) were true, then we would have to compare the path given by

$$\begin{aligned} &s_n = \pi,\ s_{n+1} = \tfrac{\pi}{2},\ s_{n+2} = 0 \\ &-y_{n,c}(-1) + y_{n+1,s}(-1) \quad \text{(metric sum)} \\ &d_n = -1,\ d_{n+1} = 1 \end{aligned} \tag{II.4.43}$$

with the path given by

$$\begin{aligned} &s_n = \pi,\ s_{n+1} = \tfrac{3}{2}\pi,\ s_{n+2} = 0 \\ &-y_{n,c}(1) - y_{n+1,s}(1) \quad \text{(metric sum)} \\ &d_n = 1,\ d_{n+1} = 1 \end{aligned} \tag{II.4.44}$$

Here the decision rule is:

$$\text{Choose } \hat{d}_n = 1 \text{ if and only if } -y_{n,c}(1) - y_{n+1,s}(1) > -y_{n,c}(-1) + y_{n+1,s}(-1)$$

Note that this decision rule is the same as that when the genie told us pair (a) was true. In fact, examination of the decision rules for all four possible pairs of states s_n and s_{n+2} that the genie could tell us reveals them to be all identical. Thus, without the aid of the genie, the best decision rule is always:

$$\text{Choose } \hat{d}_n = 1 \text{ if and only if } y_{n,c}(-1) - y_{n+1,s}(-1) > y_{n,c}(1) + y_{n+1,s}(1)$$

The above rule assumes that we started with $s_0 = 0$ or $s_0 = \pi$ and that n is an even number. When n is an odd number, we have a similar rule, namely:

$$\text{Choose } \hat{d}_n = 1 \text{ if and only if } y_{n,s}(-1) + y_{n+1,c}(-1) > y_{n,s}(1) - y_{n+1,c}(1)$$

Let us define

$$p_n = \begin{cases} y_{n,c}(-1); & n \text{ even} \\ y_{n,s}(-1); & n \text{ odd} \end{cases}$$

$$q_n = \begin{cases} y_{n,c}(1); & n \text{ even} \\ y_{n,s}(1); & n \text{ odd} \end{cases} \qquad (11.4.46)$$

Then, the two rules combine to become:

Choose $\hat{d}_n = 1$ if and only if

$$p_n - p_{n+1} > q_n + q_{n+1}; \quad n \text{ even}$$

$$p_n + p_{n+1} > q_n - q_{n+1}; \quad n \text{ odd}$$

or, equivalently

$$p_n + (-1)^{n+1} p_{n+1} > q_n + (-1)^n q_{n+1}$$

This simplified demodulator is illustrated in Figure 22.

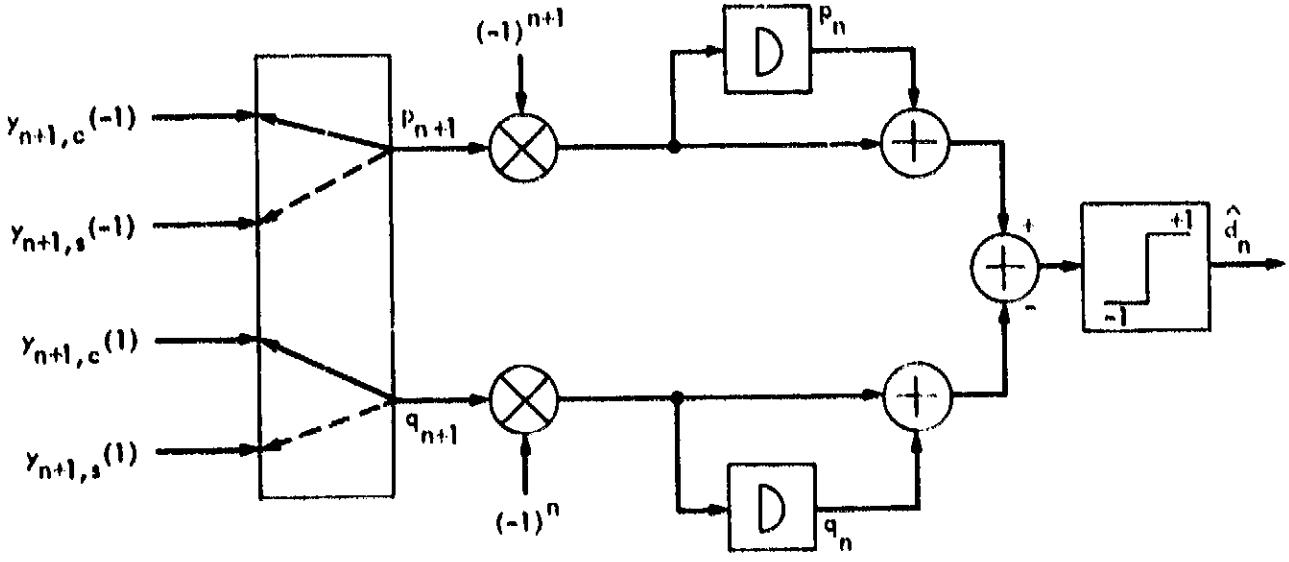


Figure 22. A Variation of Massey's MSK Receiver

4.3 General Continuous Phase Modulations

MSK is a special case of a general class of bandwidth efficient modulations referred to as continuous phase modulation (CPM). A general CPM signal can still be written in the form of (II.4.18) where now

$$\theta(t;\underline{u}) = 2\pi h \sum_{n=-\infty}^{\infty} u_n \int_{-\infty}^{t} g(\tau - nT)d\tau \qquad \text{(II.4.46)}$$

and the data sequence $\{u_n\}$ consists of i.i.d. symbols with probabilities

$$\Pr\{u_n = k\} = \frac{1}{M}; \quad k = \pm 1, \pm 2, \ldots, \pm \frac{M}{2} \qquad \text{(II.4.47)}$$

Typically M is taken to be a power of 2, e.g., 2, 4, 8, and we shall do so in our discussion. Also, in (II.4.46), h is the modulation index and g(t) is any pulse shaping function satisfying the normalization

$$\int_{-\infty}^{\infty} g(t)dt = \frac{1}{2} \qquad \text{(II.4.48)}$$

Thus, we see from our discussion in section 4.2 that MSK is the special case where h = 1/2 and g(t) is a rectangular pulse of T-second duration [see (II.4.20)].

The modulator for CPM is shown in Figure 15 with the multiplication by π replaced with a multiplication by $2\pi h$. The signal spectrum can be controlled by the choice of the pulse shaping function g(t), the modulation index h, and the alphabet size M. The bit error probability and data throughput measured in bits/second/Hz are also complex functions of these parameters. In Part I, we summarized some of the performance results obtained in this area. Here, we expand upon the details of this work emphasizing the structure of the maximum-likelihood demodulator.

From the practical point of view, we must limit the span of g(t). Assume that ν is the smallest integer such that

$$g(t) = 0 \quad \text{for } t > \nu T \qquad \text{(II.4.49)}$$

where we assume $g(t) = 0$ for $t < 0$. Then, we define

$$G(t) = \int_0^t g(\alpha)d\alpha \qquad (II.4.50)$$

where

$$G(t) = 0 \quad \text{for } t \leq 0 \qquad (II.4.51)$$

and, from (II.4.48),

$$G(t) = \frac{1}{2} \quad \text{for } t \geq \nu T \qquad (II.4.52)$$

The signal phase in the nth transmission interval $nT \leq t \leq (n+1)T$ is given by

$$\begin{aligned}\theta(t;\underline{u}) &= 2\pi h \sum_{k=-\infty}^{n} u_k\, G(t-kT) \\ &= 2\pi h \left[u_n G(t-nT) + u_{n-1} G(t-(n-1)T)\right. \\ &\quad \left. + \ldots + u_{n-(\nu-1)}\, G(t-[n-(\nu-1)]T)\right] \\ &\quad + \pi h \sum_{i=-\infty}^{n-\nu} u_i \qquad (II.4.53)\end{aligned}$$

For the same interval, also define the phase state

$$\gamma_n = \pi h \sum_{i=-\infty}^{n-\nu} u_i \qquad (II.4.54)$$

the pulse shaping function state vector

$$\underline{v}_n = (u_{n-1}, u_{n-2}, \ldots, u_{n-(\nu-1)}) \tag{11.4.55}$$

and the total state vector

$$s_n = (\underline{v}_n, \gamma_n) \tag{11.4.56}$$

To emphasize the dependence on γ_n, $\underline{v}_n$ we rewrite $\theta(t;\underline{u})$ as

$$\theta(t;\underline{u}) \triangleq \theta(t; u_n, \underline{v}_n, \gamma_n);\ nT \le t \le (n+1)T \tag{11.4.57}$$

As before the maximum-likelihood demodulator finds the sequence $\hat{\underline{u}}$ that maximizes (11.4.27). Considering each sub-integral, we obtain the generalization of (11.4.30), namely,

$$\int_{nT}^{(n+1)T} y(t)x(t;\underline{u})dt = \int_{nT}^{(n+1)T} y(t)\sqrt{2S}\cos\left[\omega_0 t + \theta(t; u_n, \underline{v}_n, \gamma_n)\right]dt$$

$$= \sqrt{E}\left[y_{n,c}(\underline{v}_n, u_n)\cos\gamma_n + y_{n,s}(\underline{v}_n, u_n)\sin\gamma_n\right] \tag{11.4.58}$$

where

$$y_{n,c}(\underline{v}_n, u_n) \triangleq \int_{nT}^{(n+1)T} y(t)\sqrt{\frac{2}{T}}\cos\left[\omega_0 t + u_n\, 2\pi h\, G(t-nT) + 2\pi h(G(t), \underline{v}_n)\right]dt$$

$$y_{n,s}(\underline{v}_n, u_n) \triangleq -\int_{nT}^{(n+1)T} y(t)\sqrt{\frac{2}{T}}\sin\left[\omega_0 t + u_n\, 2\pi h\, G(t-nT) + 2\pi h(G(t), \underline{v}_n)\right]dt$$

$$\underline{G}(t) = (G(t-(n-1)T), G(t-(n-2)T), \ldots, G(t-[n-(\nu-1)]T)) \tag{11.4.59}$$

Hence, during the nth interval the maximum-likelihood receiver computes M component pairs $y_{n,c}(\underline{v}_n, u_n)$, $y_{n,s}(\underline{v}_n, u_n)$ for all $\underline{v}_n$, u_n. The demodulator is realized by a Viterbi algorithm with state $s_n = (\underline{v}_n, \gamma_n)$ and branch metric

$$m(\underline{y}_n; u_n, \underline{v}_n, \gamma_n) = y_{n,c}(\underline{v}_n, u_n) \cos \gamma_n + y_{n,s}(\underline{v}_n, u_n) \sin \gamma_n \tag{II.4.60}$$

Note that unless h is chosen carefully the phase term

$$\gamma_n^* = 2\pi h \sum_{i=-\infty}^{n-\nu} u_i \text{ modulo } 2\pi \tag{II.4.61}$$

can take on an infinite number of possible values causing the state space to be infinite. From a practical viewpoint, however, we can always quantize the phase space and thus quantize γ_n^*. Another possibility is to carefully choose h to give a reasonably small set of possible phase values. Since the state is $s_n = (\underline{v}_n, \gamma_n)$, the size of the state space is $M^{\nu-1}|\Gamma|$ where $|\Gamma|$ is the number of distinct or quantized phase values.

One of the most commonly proposed pulse shaping functions, g(t), is the raised cosine of span ν. For M = 4 and $\nu = 3$ we have $M^\nu = 64$ component pairs $y_{n,c}(\underline{v}_n, u_n)$ and $y_{n,s}(\underline{v}_n, u_n)$ computed for each interval and a Viterbi algorithm with $16|\Gamma|$ states where $|\Gamma|$ is determined by the modulation index h. For $\nu = 2$ we have $M^\nu = 16$ and a Viterbi algorithm with $4|\Gamma|$ states. Clearly these new bandwidth efficient modulations require more complex signal processing for demodulation.

5.0 Simultaneous Phase/Data Demodulation

The coherent MSK and CPM modulations described in section 4.0 implicitly assumed an ideal carrier phase reference was available. At this point, it is not clear how conventional Costas phase tracking loops can be modified for application to some of these new bandwidth efficient modulations. Thus, in this section we describe a new open loop simultaneous phase/data estimation approach first discussed in section 3.4.3 in connection with differential BPSK. This approach is a natural extension of the Viterbi demodulator structure that, as we have already seen, is required for data demodulation alone.

Suppose that during the nth time interval $nT \le t \le (n+1)T$ the received signal contains an unknown phase perturbation denoted by ϕ_n. We assume this unknown phase sequence is a Markov chain described by

$$\phi_{n+1} = \phi_n + \Delta_n \quad \text{for all } n \tag{II.5.1}$$

where $\{\Delta_n\}$ is an i.i.d. sequence with

$$\Pr\{\Delta_n = k\Lambda\} = p_k \; ; \; k = 0, \pm 1, \pm 2, \ldots, \pm K \tag{II.5.2}$$

Our approach is to simultaneously estimate the phase sequence $\underline{\phi}$ and the data sequence $\underline{u}$. Since all phase sequences $\underline{\phi}$ are not equiprobable, we use the maximum a posteriori (MAP) rule:

Choose $\hat{\underline{u}}$ and $\hat{\underline{\phi}}$ corresponding to the largest biased correlation

$$\int_{-\infty}^{\infty} y(t)x(t;\underline{u},\underline{\phi})dt - \frac{N_0}{2} \ln p(\underline{\phi}) \tag{II.5.3}$$

where $p(\underline{\phi})$ is the probability of occurrence of the sequence $\underline{\phi}$. Here we have ignored the signal energy term since the signals of interest have constant envelope. The MAP rule is known to minimize the average probability of decision error where now we are also interested in choosing the phase sequence $\underline{\phi}$. The above Markov chain model is essentially a quantized approximation to a true phase process. With a sufficient number of quantization values determined by Λ and K, this can be as accurate as is necessary. Generally channel noise will be the dominant source of degradation beyond a certain level of quantization and finer quantization would not change the overall data bit error probability.

5.1 MSK Phase/Data Demodulation

The MSK signal with data sequence $\underline{u}$ and phase sequence $\underline{\phi}$ is given by

$$x(t,\underline{u},\underline{\phi}) = \sqrt{2S}\cos\left[\omega_0 t + \theta(t;\underline{u},\underline{\phi})\right] \tag{II.5.4}$$

where for the nth interval $nT \leq t \leq (n+1)T$,

$$\theta(t;\underline{u},\phi) = u_n \pi\left(\frac{t-nT}{2T}\right) + \frac{\pi}{2} \sum_{i=-\infty}^{n-1} u_i + \phi_n \qquad (II.5.5)$$

It is natural to now define the state at the beginning of nth interval as

$$s_n = \frac{\pi}{2} \sum_{i=-\infty}^{n-1} u_i + \phi_n \qquad (II.5.6)$$

where the phase space is now quantized to Q values (Q is a power of 2), namely,

$$\Phi = \{0, \Delta, 2\Delta, \ldots, (Q-1)\Delta\}; \Delta = \frac{2\pi}{Q} \qquad (II.5.7)$$

Note that the set Φ also characterizes the set of all possible values for the state s_n.

From (II.5.6) the state transition equation is given by

$$s_{n+1} = \frac{\pi}{2} \sum_{i=-\infty}^{n} u_i + \phi_{n+1}$$

$$= s_n + \frac{\pi}{2} u_n + \delta_n \qquad (II.5.8)$$

Thus from each of the Q states in Φ, there are 2 possible values of u_n and 2K+1 possible values of δ_n resulting in 2(2K+1) possible transitions (assuming none overlap) from each of these states. As before we compute for each interval the vector

$$y_n = (y_{n,c}(1), y_{n,c}(-1), y_{n,s}(1), y_{n,s}(-1)) \qquad (II.5.9)$$

and use the Q state Viterbi algorithm with branch metric given by

$$m(\underline{y}_n;u_n,\delta_n,s_n) = y_{n,c}(u_n)\cos s_n + y_{n,s}(u_n)\sin s_n$$

$$- \frac{N_0}{2\sqrt{E}} \ln p(\delta_n) \qquad (II.5.10)$$

Although somewhat more complex, the basic demodulator for simultaneous phase and data estimation has the same realization as a Viterbi algorithm. For MSK applications, Jackson [6] has shown that for various ranges of signal-to-noise ratio of interest, Q = 32 is adequate. Also in most cases where the phase variations from symbol to symbol are small, K = 1 suffices. Thus, we have a 32 state Viterbi demodulator with 6 transitions from each state.

5.2 CPM Phase/Data Demodulation

For the general CPM signal with data sequence $\underline{u}$ and phase sequence $\underline{\phi}$, we still have a received signal in the form of (II.5.4) where now

$$\theta(t;\underline{u},\underline{\phi}) = 2\pi h\left[u_n G(t-nT) + u_{n-1}G(t-(n-1)T) + \ldots + u_{n-(\nu-1)}G(t-[n-(\nu-1)]T)\right]$$

$$+ \pi h \sum_{i=-\infty}^{n-\nu} u_i + \phi_n \; ; \; nT \le t \le (n+1)T \qquad (II.5.11)$$

Analogous to (II.5.6) it is natural here to redefine the phase state of (II.4.54) as

$$\gamma_n = \pi h \sum_{i=-\infty}^{n-\nu} u_i + \phi_n \qquad (II.5.12)$$

but maintain the pulse shaping state vector $\underline{v}_n$ and vector function $\underline{G}(t)$ as before [see (II.4.55) and (II.4.59) respectively].

For the nth interval the demodulator again computes the M^ν pairs $y_{n,c}(\underline{v}_n,u_n)$ and $y_{n,s}(\underline{v}_n,u_n)$ and uses this in a Viterbi algorithm with state at time n given by (II.4.56). Here the pulse shaping state vector satisfies

$$\underline{v}_{n+1} = f(\underline{v}_n,u_n) \qquad (II.5.13)$$

where $\underline{y}_{n+1}$ is simply a shifted version of $\underline{y}_n$ with u_n in the first position while the phase state equation is

$$\gamma_{n+1} = \gamma_n + \pi u_{n-\nu+1} + \delta_n \tag{11.5.14}$$

If we again quantize the phase space into the Q values of (11.5.7), then, analogous to MSK, we have Q phase states and $M^{\nu-1}$ pulse shaping states for a total of $QM^{\nu-1}$ states for the Viterbi algorithm. The branch metric for this Viterbi demodulator is given by

$$m(\underline{y}_n;u_n,\delta_n,\underline{y}_n,\gamma_n) = y_{n,c}(\underline{y}_n,u_n)\cos\gamma_n + y_{n,s}(\underline{y}_n,u_n)\sin\gamma_n$$

$$- \frac{N_0}{2\sqrt{E}} \ln p(\delta_n) \tag{11.5.15}$$

Note that the pulse shaping state vector which satisfies (11.5.13) and the phase state which satisfies (11.5.14) are loosely coupled through the data where the last component of $\underline{y}_n$ becomes the next input term to generate γ_{n+1}. At this time it is not clear if a reduced state Viterbi demodulator can be devised that has two separate trellis diagrams for these two state processes rather than a single very large state diagram for state $s_n = (\underline{y}_n,\gamma_n)$.

6.0 Bit Error Probability Bounds

Except for some special cases, exact error probability evaluations are difficult to obtain for the new bandwidth efficient modulations which introduce memory into the modulation process. In this section, we describe two basic means of approximately evaluating these error probabilities. The first, which is particularly suited to coded communication systems, expresses the result in terms of the cutoff rate of the coding channel consisting of the modulator, radio channel, and demodulator. The second approach makes use of transfer function bounds and is most suited to the evaluation of error probability in uncoded communication systems. A detailed treatment of this latter approach is given in Appendix A of Part IV of this report. Thus, here in Part II we shall merely summarize its salient features and go on from there to apply it to the evaluation of bit error probability in uncoded systems employing bandwidth efficient modulations of the CPM type.

6.1 Cutoff Rate

The cutoff rate R_0 is defined for the coding channel created by the modulator, radio channel, and demodulator. It represents the practically achievable code rate in bits/channel use. For any specific encoder preceding this channel and any specific decoder succeeding it, a bound can be derived on the coded bit error probability P_b in a form which is only a function of R_0. Since the function that relates P_b to R_0 is unique for each encoder/decoder combination, and R_0 is independent of the code used, then by expressing things in this form, we are able to decouple the coding from the rest of the communication system. As such, we can compare the performance of various communication systems by first evaluating them in terms of their cutoff rate parameter. The additional performance improvement obtained through coding can then be evaluated separately.

In Part I, we discussed this coding parameter and gave several examples of its application. Here, we derive the basic formula for the cutoff rate of the bandwidth efficient modulations discussed in Section 5. As always, we assume a white Gaussian noise channel.

Consider two data sequences of length L, namely,

$$\underline{u} = (u_0, u_1, \ldots, u_{L-1}) \tag{II.6.1}$$

$$\underline{\hat{u}} = (\hat{u}_0, \hat{u}_1, \ldots, \hat{u}_{L-1})$$

and their corresponding modulated waveforms $x(t;\underline{u})$ and $x(t;\underline{\hat{u}})$ of duration $0 \leq t \leq LT$. Next let $\phi_1(t), \phi_2(t), \ldots, \phi_N(t)$ be an orthonormal basis for the set of all such signals of LT second duration. Then the signal components are

$$x_n = \int_0^{LT} x(t;\underline{u})\phi_n(t)dt$$

$$\hat{x}_n = \int_0^{LT} x(t;\underline{\hat{u}})\phi_n(t)dt; \quad n = 1, 2, \ldots, N \tag{II.6.2}$$

Without loss of generality when $x(t;\underline{u})$ is transmitted we can compute at the receiver the components

$$y_k = \int_0^{LT} y(t)\phi_k(t)dt$$

$$= x_k + n_k;\ k = 1, 2, \ldots, N \qquad (II.6.3)$$

where $\{n_k\}$ is as before an i.i.d. sequence of Gaussian random variables with zero mean and variance $N_0/2$.

Letting $\underline{x} \triangleq (x_1, x_2, \ldots, x_N)$, $\hat{\underline{x}} \triangleq (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_N)$, and $\underline{y} \triangleq (\underline{y}_1, y_2, \ldots, y_N)$, we have the conditional probabilities

$$p_N(\underline{y}|\underline{x}) = (\frac{1}{\pi N_0})^{N/2} \exp(-\frac{1}{N_0} \| \underline{y}-\underline{x} \|^2)$$

and

$$p_N(\underline{y}|\hat{\underline{x}}) = (\frac{1}{\pi N_0})^{N/2} \exp(-\frac{1}{N_0} \| \underline{y}-\hat{\underline{x}} \|^2) \qquad (II.6.4)$$

If $\underline{x}$ and $\hat{\underline{x}}$ are the only two possible channel inputs then the maximum-likelihood decision regions are

$$\Lambda = \{\underline{y}: p_N(\underline{y}|\underline{x}) > p_N(\underline{y}|\hat{\underline{x}})\}$$

and

$$\hat{\Lambda} = \{\underline{y}: p_N(\underline{y}|\underline{x}) \leq p_N(\underline{y}|\hat{\underline{x}})\} \qquad (II.6.5)$$

Assuming $\underline{x}$ is the transmitted sequence, the pair-wise probability of error, denoted $\Pr\{\underline{u} \to \hat{\underline{u}}\}$, is given by

$$\Pr\{\underline{u} \to \hat{\underline{u}}\} = \int_{\hat{\Lambda}} p_N(\underline{y}|\underline{x})\,d\underline{y} \tag{II.6.6}$$

Next note from (II.6.5) that

$$1 \le \left[\frac{p_N(\underline{y}|\hat{\underline{x}})}{p_N(\underline{y}|\underline{x})}\right]^{1/2} \quad \text{for } \underline{y} \in \hat{\Lambda} \tag{II.6.7}$$

Hence, we have the Bhattacharyya bound

$$\Pr\{\underline{u} \to \hat{\underline{u}}\} \le \int_{\hat{\Lambda}} p_N(\underline{y}|\underline{x}) \left[\frac{p_N(\underline{y}|\hat{\underline{x}})}{p_N(\underline{y}|\underline{x})}\right]^{1/2} d\underline{y}$$

$$= \int_{\hat{\Lambda}} \sqrt{p_N(\underline{y}|\underline{x})\; p_N(\underline{y}|\hat{\underline{x}})}\; d\underline{y}$$

$$\le \int \sqrt{p_N(\underline{y}|\underline{x})\; p_N(\underline{y}|\hat{\underline{x}})}\; d\underline{y} \tag{II.6.8}$$

where the last inequality follows from the fact that we integrate over all sequences $\underline{y}$ and the integrand is non-negative. Noting from (II.6.4) that

$$\sqrt{p_N(\underline{y}|\underline{x})\; p_N(\underline{y}|\hat{\underline{x}})} = \left(\frac{1}{\pi N_0}\right)^{N/2} \exp\left(-\frac{1}{N_0}\left\| \underline{y} - \frac{\underline{x} + \hat{\underline{x}}}{2}\right\|^2\right) \exp\left(-\frac{1}{4N_0}\left\| \underline{x} - \hat{\underline{x}}\right\|^2\right) \tag{II.6.9}$$

then substituting (II.6.9) into (II.6.8), and noting that the integral of the first factor is unity, we get the final general form of the Bhattacharyya bound for the additive white Gaussian noise channel, namely,

$$\Pr\{\underline{u} \to \hat{\underline{u}}\} \le \exp\left(-\frac{1}{4N_0}\left\| x - \hat{x}\right\|^2\right) \tag{II.6.10}$$

Note that this bound is symmetric in $\underline{x}$ and $\hat{\underline{x}}$ and thus it also applies to the case where $\hat{\underline{u}}$ is the data sequence and $\hat{\underline{x}}$ is the transmitted sequence. Furthermore, regardless of the choice of the orthonormal basis we have

$$\|\underline{x} - \hat{\underline{x}}\|^2 = \int_0^{LT} [x(t;\underline{u}) - x(t;\hat{\underline{u}})]^2 dt \qquad \text{(II.6.11)}$$

Thus, (II.6.10) alternately becomes

$$\Pr\{\underline{u} \to \hat{\underline{u}}\} \leq \exp\left(-\int_0^{LT} [x(t;\underline{u}) - x(t;\hat{\underline{u}})]^2 dt\right) \qquad \text{(II.6.12)}$$

Averaging (II.6.12) over all independent equiprobable data sequences $\underline{u}$ and $\hat{\underline{u}}$ defines the function $r_0(L)$, namely,

$$E\left\{\Pr\{\underline{u} \to \hat{\underline{u}}\}\right\} \leq E\left\{\exp\left(-\int_0^{LT} [x(t,\underline{u}) - x(t,\hat{\underline{u}})]^2 dt\right)\right\} \triangleq 2^{-Lr_0(L)} \qquad \text{(II.6.13)}$$

If the channel is discrete memoryless, then $r_0(L)$ will be independent of L and equal to the cutoff rate r_0. For a discrete memory channel such as that created by the CPM modulation/demodulation techniques, $r_0(L)$ will, in general, be a function of L and hence cutoff rate is defined as the limit of $r_0(L)$ as L approaches infinity, and as such is independent of L. Thus, from (II.6.13),

$$\boxed{r_0 = -\lim_{L\to\infty} \frac{1}{L} \log_2 E\left\{\exp\left[-\frac{1}{4N_0}\int_0^{LT} [x(t;\underline{u}) - x(t;\hat{\underline{u}})]^2 dt\right]\right\}} \qquad \text{(II.6.14)}$$

Up to this point, we have not made use of the specific form of the CPM waveform as given by (11.4.18) together with (11.4.46). When this is done (11.6.11) and (11.6.12) respectively specialize to

$$\|\underline{x} - \underline{\hat{x}}\|^2 = \int_0^{LT} [x(t;\underline{u}) - x(t;\underline{\hat{u}})]^2 dt$$

$$= \sum_{k=0}^{L-1} \int_{kT}^{(k+1)T} [x(t;\underline{u}) - x(t;\underline{\hat{u}})]^2 dt$$

$$= \sum_{k=0}^{L-1} \left[2ST - 2 \int_{kT}^{(k+1)T} x(t;\underline{u}) x(t;\underline{\hat{u}}) dt \right]$$

$$= \sum_{k=0}^{L-1} \left\{ 2ST - 2S \int_{kT}^{(k+1)T} \cos[\theta(t;u_k,s_k) - \theta(t;\hat{u}_k,\hat{s}_k)] dt \right\} \tag{11.6.15}$$

and

$$\Pr\{\underline{u} \to \underline{\hat{u}}\} \leq \prod_{k=0}^{L-1} \exp\left\{ - \frac{ST - S \int_{kT}^{(k+1)T} \cos[\theta(t;u_k,s_k) - \theta(t;\hat{u}_k,\hat{s}_k)] dt}{2N_0} \right\} \tag{11.6.16}$$

where, analogous to (11.4.57), $\theta(t;u_k,s_k)$ represents the modulation angle in the interval $kT \leq t \leq (k+1)T$. [Note that, for simplicity of notation, we use the contracted form for the state at time k as given in (11.4.56).]

6.1.1 Coherent MSK

For coherent MSK, we have [see (II.4.22) and (II.4.23)]

$$\theta(t;u_k,s_k) = u_k\pi\left(\frac{t-kT}{2T}\right) + s_k;\ kT \le t \le (k+1)T \tag{II.6.17}$$

and thus

$$\theta(t;u_k,s_k) - \theta(t;\hat{u}_k\hat{s}_k) = (u_k-\hat{u}_k)\pi\left(\frac{t-kT}{2T}\right) + (s_k-\hat{s}_k) \tag{II.6.18}$$

where

$$s_k-\hat{s}_k = \frac{\pi}{2}\sum_{i=-\infty}^{k-1}(u_i-\hat{u}_i) \tag{II.6.19}$$

Furthermore, since we are only interested in $\cos[\theta(t;u_k,s_k) - \theta(t;\hat{u}_k,\hat{s}_k)]$, we can take all these differences modulo 2π.

Note that for this case we can define the error

$$\epsilon_k \triangleq u_k-\hat{u}_k \tag{II.6.20}$$

which can assume values $(-2,0,2)$ with probability

$$q(\epsilon_k) = \begin{cases} \frac{1}{4}; & \epsilon_k = -2,2 \\ \frac{1}{2}; & \epsilon_k = 0 \end{cases} \tag{II.6.21}$$

and the error state

$$\Delta_k \triangleq s_k-\hat{s}_k \tag{II.6.22}$$

which can take on values in $(0,)$. Using these definitions, we can simplify the integral in (II.6.15) and (II.6.16) as follows:

$$\int_{kT}^{(k+1)T} \cos[\theta(t;u_k,\theta_k) - \theta(t;\hat{u}_k,\hat{\theta}_k)]dt$$

$$= \int_{kT}^{(k+1)T} \cos\left[\epsilon_k \pi\left(\frac{t-kT}{2T}\right) + \Lambda_k\right]dt$$

$$= \cos\Lambda_k \int_{kT}^{(k+1)T} \cos\left[\epsilon_k \pi\left(\frac{t-kT}{2T}\right)\right]dt$$

$$- \sin\Lambda_k \int_{kT}^{(k+1)T} \sin\left[\epsilon_k \pi\left(\frac{t-kT}{2T}\right)\right]dt$$

$$= \cos\Lambda_k \int_0^T \cos\left(\frac{\epsilon_k \pi}{2T}t\right)dt - \sin\Lambda_k \int_0^T \sin\left(\frac{\epsilon_k \pi}{2T}t\right)dt$$

$$= \begin{cases} T\cos\Lambda_k; & \epsilon_k = 0 \\ 0; & \epsilon_k = -2,2 \end{cases} \qquad (II.6.23)$$

Next we define

$$f(\epsilon_k;\Lambda_k) \triangleq \exp\left\{-\frac{ST - S\int_{kT}^{(k+1)T} \cos\left[\epsilon_k \pi\left(\frac{t-kT}{2T}\right) + \Lambda_k\right]dt}{2N_0}\right\}$$

$$= \begin{cases} \exp -\left(\frac{ST}{2N_0}\right); & \epsilon_k = -2,2 \\ \exp\left[-\frac{ST}{2N_0}(1-\cos\Lambda_k)\right]; & \epsilon_k = 0 \end{cases} \qquad (II.6.24)$$

which results in the Bhattacharyya bound [see (11.6.16)]

$$\Pr(u \neq \hat{u}) \leq \prod_{k=0}^{L-1} f(\epsilon_k;\Lambda_k)$$

$$= \exp\left\{-\frac{ST}{2N_0}\left(L - \sum_{\epsilon_k=0} \cos\Lambda_k\right)\right\} \qquad (11.6.25)$$

To evaluate the cutoff rate r_0 we must consider the average*

$$E\left\{\prod_{k=0}^{L-1} f(\epsilon_k;\Lambda_k)\,|\,\Lambda_0\right\}$$

where $E\{\cdot\}$ is the expectation over the sequence of i.i.d. errors $\epsilon_0, \epsilon_1, \ldots, \epsilon_{L-1}$ and we have chosen $\Lambda_0 = 0$ as the initial error state. In terms of the probability $q(\epsilon_k)$ of (11.6.21) we have the conditional expectation

$$E\left\{\prod_{k=0}^{L-1} f(\epsilon_k;\Lambda_k)\,|\,\Lambda_0 = 0\right\} = \sum_{\epsilon_0}\sum_{\epsilon_1}\cdots\sum_{\epsilon_{L-1}}\prod_{k=0}^{L-1} q(\epsilon_k) f(\epsilon_k;\Lambda_k) \qquad (11.6.26)$$

where the summations on $\epsilon_0, \epsilon_1, \ldots, \epsilon_{L-1}$ each range over $(-2, 0, 2)$.

Next note from Figure 23 that the error state diagram is simply a two state diagram where both $\epsilon_k = -2$ and $\epsilon_k = 2$ result in a change of state and for $\epsilon_k = 0$ the state remains the same. Since, from (11.6.21), $q(-2) = q(2) = 1/4$, the probability of a change of state is $1/2$ as is the probability of no state change. Defining the matrix,

*Since r_0 is defined as a limit as L goes to infinity, the value of r_0 is independent of any initial assumption. Thus it is sufficient to consider the statistical average of above conditioned on any of the inital error states.

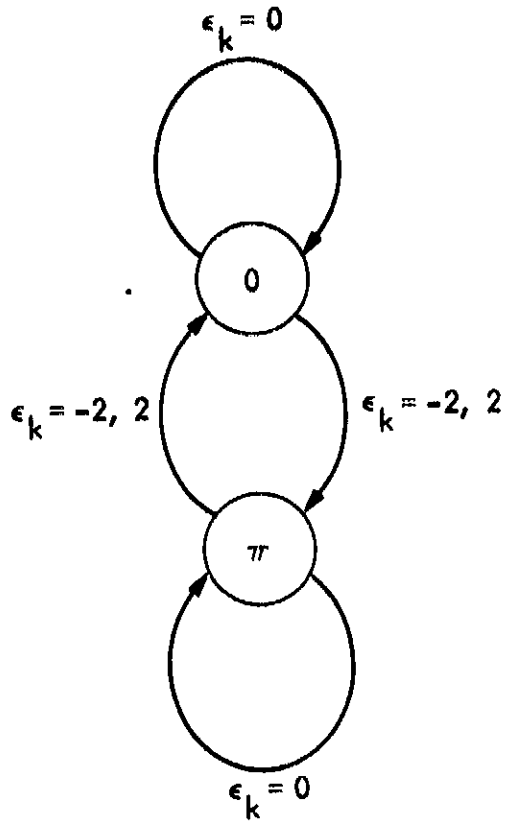


Figure 23. MSK Error State Diagram

$$\underline{\mathcal{M}} = \begin{bmatrix} \frac{1}{2}f(0;0) & \frac{1}{2}f(0;\pi) \\ \\ \frac{1}{2}f(2;0) & \frac{1}{2}f(2;\pi) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{2}\exp\left(-\frac{ST}{N_0}\right) \\ \\ \frac{1}{2}\exp\left(-\frac{ST}{2N_0}\right) & \frac{1}{2}\exp\left(-\frac{ST}{2N_0}\right) \end{bmatrix} \qquad \text{(II.6.27)}$$

we have the partial sum

$$\sum_{\epsilon_{L-1}} q(\epsilon_{L-1}) f(\epsilon_{L-1};\Lambda_{L-1}) = \begin{cases} \frac{1}{2}f(0;0) + \frac{1}{2}f(2;0) \; ; \; \Lambda_{L-1} = 0 \\ \\ \frac{1}{2}f(0;\pi) + \frac{1}{2}f(2;\pi) \; ; \; \Lambda_{L-1} = \pi \end{cases} \tag{II.6.28}$$

where the term for $\Lambda_{L-1} = 0$ corresponds to the first component of the vector

$$[1 \; 1]\,\underline{\mathscr{A}}$$

and the term for $\Lambda_{L-1} = \pi$ corresponds to the second component. Next note that

$$\sum_{\epsilon_{L-2}} \sum_{\epsilon_{L-1}} q(\epsilon_{L-2})q(\epsilon_{L-1})f(\epsilon_{L-2};\Lambda_{L-2})f(\epsilon_{L-1};\Lambda_{L-1})$$

$$= \begin{cases} \frac{1}{2}f(0;0)[\frac{1}{2}f(0;0) + \frac{1}{2}f(2;0)] \\ + \frac{1}{2}f(2;0)[\frac{1}{2}f(0;\pi) + \frac{1}{2}f(2;\pi)]; \; \Lambda_{L-2} = 0 \\ \frac{1}{2}f(0;\pi)[\frac{1}{2}f(0;0) + \frac{1}{2}f(2;0)] \\ + \frac{1}{2}f(2;\pi)[\frac{1}{2}f(0;\pi) + \frac{1}{2}f(2;\pi)]; \; \Lambda_{L-2} = \pi \end{cases} \tag{II.6.29}$$

Here the term for $\Lambda_{L-2} = 0$ is the first component of

$$[1 \; 1]\,\underline{\mathscr{A}}^2$$

while its second component is the above sum for $\Lambda_{L-2} = \pi$. Finally then for the L-fold sum required in (II.6.26), we obtain the desired result

$$E\left\{\prod_{k=0}^{L-1} f(\epsilon_k;\Lambda_k)\,|\,\Lambda_0 = 0\right\} = [1\ 1]\,\underline{\mathcal{A}}^L\begin{bmatrix}1\\0\end{bmatrix} \qquad \text{(II.6.30)}$$

Note that if we had assumed $\Lambda_0 = \pi$ as our initial state, then, analogous to (II.6.30), we would obtain

$$E\left\{\prod_{k=0}^{L-1} f(\epsilon_k;\Lambda_k)\,|\,\Lambda_0 = \pi\right\} = [1\ 1]\,\underline{\mathcal{A}}^L\begin{bmatrix}0\\1\end{bmatrix} \qquad \text{(II.6.31)}$$

The matrix $\underline{\mathcal{A}}$ is non-negative and irreducible. The Perron-Frobenius theorem [7] states that such a matrix has a real maximum eigenvalue λ and an associated positive left eigenvector. Defining $\alpha > 0$ to be the ratio of the largest component of the left eigenvector divided by its smallest component we have the inequalities [8; p. 338]

$$\frac{1}{\alpha}\lambda^L \le [1\ 1]\,\underline{\mathcal{A}}^L\begin{bmatrix}1\\0\end{bmatrix} \le \alpha\lambda^L \qquad \text{(II.6.32)}$$

Thus from (II.6.14) and (II.6.30) the cutoff rate is

$$\boxed{\begin{aligned} r_0 &= -\lim_{L\to\infty}\frac{1}{L}\log_2\left\{[1\ 1]\,\underline{\mathcal{A}}^L\begin{bmatrix}1\\0\end{bmatrix}\right\} \\ &= -\log_2\lambda \end{aligned}} \qquad \text{(II.6.33)}$$

Note that in arriving at the final equality in (II.6.33), we have made use of the fact that

$$\lim_{L\to\infty}\frac{1}{L}\log_2\frac{1}{\alpha} = \lim_{L\to\infty}\frac{1}{L}\log_2\alpha = 0$$

For the matrix $\underline{\mathcal{A}}$ of (II.6.27), we have the maximum eigenvalue

$$\lambda = \frac{1 + Z + \sqrt{1 - 2Z + Z^2 + 4Z^3}}{4} \qquad \text{(II.6.34)}$$

where

$$\gamma \triangleq \exp\left(-\frac{ST}{2N_0}\right) \tag{II.6.35}$$

Figure 24 is a plot of r_0 of (II.6.33) together with (II.6.34) and (II.6.35) versus ST/N_0 in dB.

6.1.2 MSK Phase/Data Demodulation

For the simultaneous phase and data demodulator discussed in section 5.1, we still have $\theta(t;u_k,s_k)$ as in (II.6.17). However, the state s_k is now given by (II.5.6) and includes the unknown phase ϕ_k which satisfies

$$\phi_{k+1} = \phi_k + \delta_k \tag{II.6.36}$$

Assuming again that the phase space is quantized into the Q values of (II.5.7), and the error and error states are still defined as in (II.6.20) and (II.6.22)

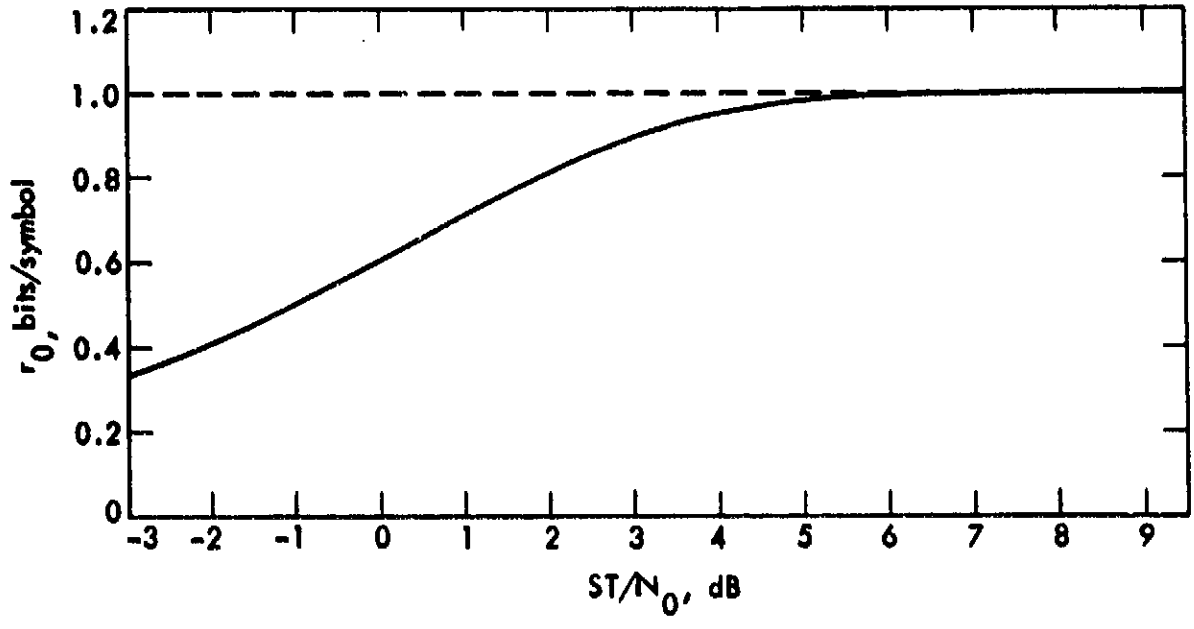


Figure 24. Computational Cutoff Rate versus Signal-to-Noise Ratio for MSK Modulation

respectively, then analogous to (II.6.26) we have*

$$E\left\{\sum_{k=0}^{L-1} f(\epsilon_k;\Delta_k)\,|\,\Delta_0 = 0\right\}$$

$$=\sum_{\epsilon_0}\sum_{\epsilon_1}\cdots\sum_{\epsilon_{L-1}}\sum_{\Delta\delta_0}\sum_{\Delta\delta_1}\cdots\sum_{\Delta\delta_{L-1}}\prod_{k=0}^{L-1} p(\Delta\delta_k)q(\epsilon_k)f(\epsilon_k;\Delta_k) \tag{II.6.37}$$

since the expectation must now be taken over both the data errors $\{\epsilon_k;k=0,1,\ldots L-1\}$ and the phase perturbation errors $\{\Delta\delta_k;k=0,1,\ldots,L-1\}$ defined by

$$\Delta\delta_k \triangleq \delta_k - \hat{\delta}_k \tag{II.6.38}$$

with probabilities $\{p(\Delta\delta_k)\}$. To illustrate the independence of $(\epsilon_k;\Delta_k)$ in (II.6.37) on the phase perturbation errors $\Delta\delta_k$, we write the error state Δ_k as

$$\Delta_k = \frac{\pi}{2}\sum_{i=-\infty}^{k-1}(u_i-\hat{u}_i) + \phi_k - \hat{\phi}_k \tag{II.6.39}$$

Then the error state transition equation becomes

$$\Delta_{k+1} = \frac{\pi}{2}\sum_{i=-\infty}^{k-1}(u_i-\hat{u}_i) + \frac{\pi}{2}(u_k-\hat{u}_k) + \phi_{k+1} - \hat{\phi}_{k+1}$$

$$= \frac{\pi}{2}\sum_{i=-\infty}^{k-1}(u_i-\hat{u}_i) + \frac{\pi}{2}\epsilon_k + \phi_k - \hat{\phi}_k + (\phi_{k+1}-\phi_k) - (\hat{\phi}_{k+1}-\hat{\phi}_k)$$

$$= \Delta_k + \frac{\pi}{2}\epsilon_k + \delta_k - \hat{\delta}_k$$

*Again insofar as computation of r_0 is concerned, it is sufficient to consider only the statistical average conditioned on one of the initial states.

$$= \Lambda_k + \frac{\pi}{2}\epsilon_k + \Delta\delta_k \qquad \text{(II.6.40)}$$

Note that while $\{\epsilon_k; k=0,1,\ldots,L-1\}$ still ranges over the values (-2, 0, 2) in their respective summations of (II.6.37), $\{\Delta\delta_k; k=0,1,\ldots,L-1\}$ ranges over the Q phase values of (II.5.7) in their respective summations. Thus, the number of states in the error state transition diagram for MSK phase/data demodulation is Q. (Note that since Q is assumed to be a power of 2, the changes in error state produced by $\epsilon_k=0$ and $\Delta\delta_k=\pi$, $\epsilon_k=2$ and $\Delta\delta_k=0$, or $\epsilon_k=-2$ and $\Delta\delta_k=0$ are all identical.)

Now, as before, we define the matrix

$$\underline{\mathcal{A}} = \left|\sigma_{ij}\right| \qquad \text{(II.6.41)}$$

where

$$\sigma_{ij} = \begin{cases} p(\Delta\delta)q(\epsilon)f(\epsilon;j\Lambda); & \text{if the state } i\Lambda \text{ can be reached from the state } j\Lambda \text{ for some pair } \epsilon \text{ and } \Delta\delta \\ 0; \text{ otherwise} \end{cases} \qquad \text{(II.6.42)}$$

Then, following steps analogous to those leading up to (II.6.30), it can be similarly shown that

$$E\left\{\prod_{k=0}^{L-1} f(\epsilon_k;\Lambda_k)\,\middle|\,\Lambda_0 = 0\right\} = [1\ 1\ \ldots\ 1]\,\underline{\mathcal{A}}^L \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{(II.6.43)}$$

Furthermore, using inequalities on the matrix multiplication in (II.6.43) identical to those in (II.6.32), then the cutoff rate is once again given by

$$r_0 = -\log_2 \lambda \quad \text{bits/symbol} \qquad \text{(II.6.44)}$$

where λ is the maximum eigenvalue of $\underline{\mathcal{A}}$ as defined in (II.6.41) together with (II.6.42).

6.1.3 CPM Phase/Data Demodulation

The most general CPM case including simultaneous phase and data demodulation follows the same pattern as the MSK results. Here we have for the interval $kT \le t < (k+1)T$ the phase term [see (II.5.11)]

$$\theta(t;\underline{u}_k,\underline{s}_k) = 2\pi h \underline{u}_k \underline{G}(t-kT) + 2\pi h\,(\underline{G}(t),\underline{v}_k) + \pi h \sum_{l=-\infty}^{k-\nu} u_l + \phi_k \qquad \text{(II.6.45)}$$

where $\underline{v}_k$ and $\underline{G}(t)$ are defined by (II.4.55) and (II.4.59) respectively. Furthermore the phase state γ_k is still as in (II.5.12) and the total state is once again given by

$$s_k = (\underline{v}_k, \gamma_k) \qquad \text{(II.6.46)}$$

As in the previous parts of Section 6, we can define the error ϵ_k by (II.6.20) and the error state Δ_k by (II.6.22) which in view of (II.6.46) becomes

$$\Delta_k = (\underline{v}_k - \hat{\underline{v}}_k,\ \gamma_k - \hat{\gamma}_k) \qquad \text{(II.6.47)}$$

To evaluate the computational cutoff rate r_0 for CPM phase/data demodulation, we need to evaluate a conditional expectation analogous to (II.6.37) where $f(\epsilon_k;\Delta_k)$ is still given by (II.6.24) but now

$$\begin{aligned}\theta(t;u_k,s_k) - \theta(t;\hat{u}_k,\hat{s}_k) &= 2\pi h\epsilon_k \underline{G}(t-kT) + 2\pi h(\underline{G}(t),\underline{v}_k-\hat{\underline{v}}_k) \\ &\quad + \pi h \sum_{l=-\infty}^{k-\nu} \epsilon_l + (\phi_k - \hat{\phi}_k) \\ &= 2\pi h\epsilon_k \underline{G}(t-kT) + 2\pi h(\underline{G}(t),\underline{v}_k-\hat{\underline{v}}_k) \\ &\quad + \gamma_k - \hat{\gamma}_k;\ kT \le t \le (k+1)T \qquad \text{(II.6.48)}\end{aligned}$$

which is a function of ε_k and Δ_k through (II.6.47). Thus, we have

$$E\left[\prod_{k=0}^{L-1} f(\varepsilon_k;\Delta_k)\,|\,\Delta_0 = \beta_0\right]$$

$$= \sum_{\varepsilon_0}\sum_{\varepsilon_1}\cdots\sum_{\varepsilon_{L-1}}\sum_{\Delta\delta_0}\sum_{\Delta\delta_1}\cdots\sum_{\Delta\delta_{L-1}}\prod_{k=0}^{L-1} p(\Delta\delta_k)q(\varepsilon_k)f(\varepsilon_k;\Delta_k)$$

$$= [1\ 1\ \ldots\ 1]\,\underline{\mathcal{A}}^{L}\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix} \qquad \text{(II.6.49)}$$

where the components of the matrix $\underline{\mathcal{A}}$ are given by

$$a_{ij} = \begin{cases} p(\Delta\delta)q(\varepsilon)f(\varepsilon;\beta_j); & \text{if error state } \beta_i \text{ can be reached} \\ & \text{from error state } \beta_j \text{ for some} \\ & \text{pair } \varepsilon \text{ and } \Delta\delta \\ 0; \text{ otherwise} & \end{cases}$$

(II.6.50)

and the error states are denoted by $\beta_0, \beta_1, \beta_2, \ldots, \beta_{N_\Delta-1}$. Here N_Δ is the number of error states, thus defining the dimension of $\underline{\mathcal{A}}$. Finally, the cutoff rate is again given by (II.6.44) where λ is the maximum eigenvalue of $\underline{\mathcal{A}}$ defined in (II.6.50).

6.1.4 Evaluation of r_0

The numerical evaluation of the cutoff rate, r_0, requires, in general, computing the maximum eigenvalue of a non-negative irreducible matrix $\underline{\mathcal{A}}$. We know from the bounds in (II.6.32) that as L increases, the quantity $\underline{x}\,\underline{\mathcal{A}}^{L}\,\underline{y}^{T}$ where $\underline{x} = [1\ 1\ \ldots\ 1]$ and $\underline{y} = [1\ 0\ \ldots\ 0]$, approaches a constant times λ^L or, equivalently

$$\lim_{L\to\infty}\frac{\underline{x}\,\underline{\mathcal{A}}^{L+1}\,\underline{y}^{T}}{\underline{x}\,\underline{\mathcal{A}}^{L}\,\underline{y}^{T}} = \lambda \qquad \text{(II.6.51)}$$

Thus, a good algorithm for computing λ requires successive evaluations of $\underline{x}\underline{\mathcal{A}}^L\underline{y}^T$ for several values of L until the above ratio converges. Generally since $\underline{\mathcal{A}}$ is a sparse matrix with few nonzero elements per row or column, an efficient algorithm for computing the vector sequence

$$\underline{y}_n^T = \underline{\mathcal{A}}\,\underline{y}_{n-1}^T \tag{II.6.52}$$

where

$$\underline{y}_0^T = \underline{y}^T \tag{II.6.53}$$

can be developed even for a large matrix $\underline{\mathcal{A}}$ [9].

In order to compare r_0 for different CPM schemes, it is important to normalize with respect to some common RF bandwidth. In Part I, we discussed the suitability of using the 99% power and FCC definitions of authorized bandwidth B = 2W for commercial communication within the United States. To promote a feeling of continuity between Parts I and II, we shall use these definitions in the examples which follow.

For given values of S, B, and N_0 a convenient parameter to use in our comparison is $S/N_0W = 2S/N_0B$. Thus, using the value of time-bandwidth product $\alpha \triangleq WT$ for the modulation technique under consideration, we have

$$\frac{S}{N_0W} = \frac{1}{WT}\left(\frac{ST}{N_0}\right) = \alpha\left(\frac{ST}{N_0}\right) \tag{II.6.54}$$

Also, it is convenient to normalize r_0 by $2\alpha = BT$ and define the computational cutoff rate

$$R_0 \triangleq \frac{r_0}{2\alpha} \quad \text{bits/sec/Hz} \tag{II.6.55}$$

The significance of R_0 is that it represents the maximum practical throughput of the modulation scheme. Thus, if from our previous results we determine r_0 versus ST/N_0 for a given modulation scheme, then using the corresponding value of α in

(II.6.54) and (II.6.55) we have an equivalent relation between R_0 and S/N_0W. Also note that the maximum value of R_0, namely [12, 13],

$$R_0^* = \left(\frac{\log_2 e}{2}\right)\left[1 + \frac{S}{2N_0W} - \sqrt{1 + \left(\frac{S}{2N_0W}\right)^2}\right]$$

$$+ 1/2 \log_2\left[1/2\left(1 + \sqrt{1 + \left(\frac{S}{2N_0W}\right)^2}\right)\right] \qquad \text{(II.6.56)}$$

is also given as a function of only S/N_0W.

Figure 25 is an illustration of R_0 versus S/N_0W for various CPM techniques all having a rectangular pulse shaping function (see section 4.3) of duration νT, i.e.,

$$g(t) = \begin{cases} \frac{1}{2\nu T}; & 0 \leq t \leq \nu T \\ 0; & \text{otherwise} \end{cases} \qquad \text{(II.6.57)}$$

but different values of h, ν, and alphabet $|U|$ for the data sequence $\{u_n\}$*. Also shown in this figure is the capacity C which is given by

$$C = 1/2 \log_2\left(1 + \frac{S}{N_0W}\right) \qquad \text{(II.6.58)}$$

For MSK (h=1/2, ν=1, and $|U|=\{\pm 1\}$), the 99% power requirement is satisfied when [6] $2\alpha = 1.17$. Thus, from (II.6.33), we have

$$R_0 = -\frac{1}{2\alpha} \ell n\lambda = -.8547 \ell n\lambda \qquad \text{(II.6.59)}$$

where λ is given by (II.6.34) with Z [see (II.6.35)] now expressed in terms of S/N_0W by

$$Z = \exp\left[-\frac{1}{2\alpha}\left(\frac{S}{N_0W}\right)\right] = \exp\left[-.8547\left(\frac{S}{N_0W}\right)\right] \qquad \text{(II.6.60)}$$

*In these results, we also allow $|U|$ to include the level zero in addition to the usual levels $\pm 1, \pm 2, \ldots, \pm M/2$.

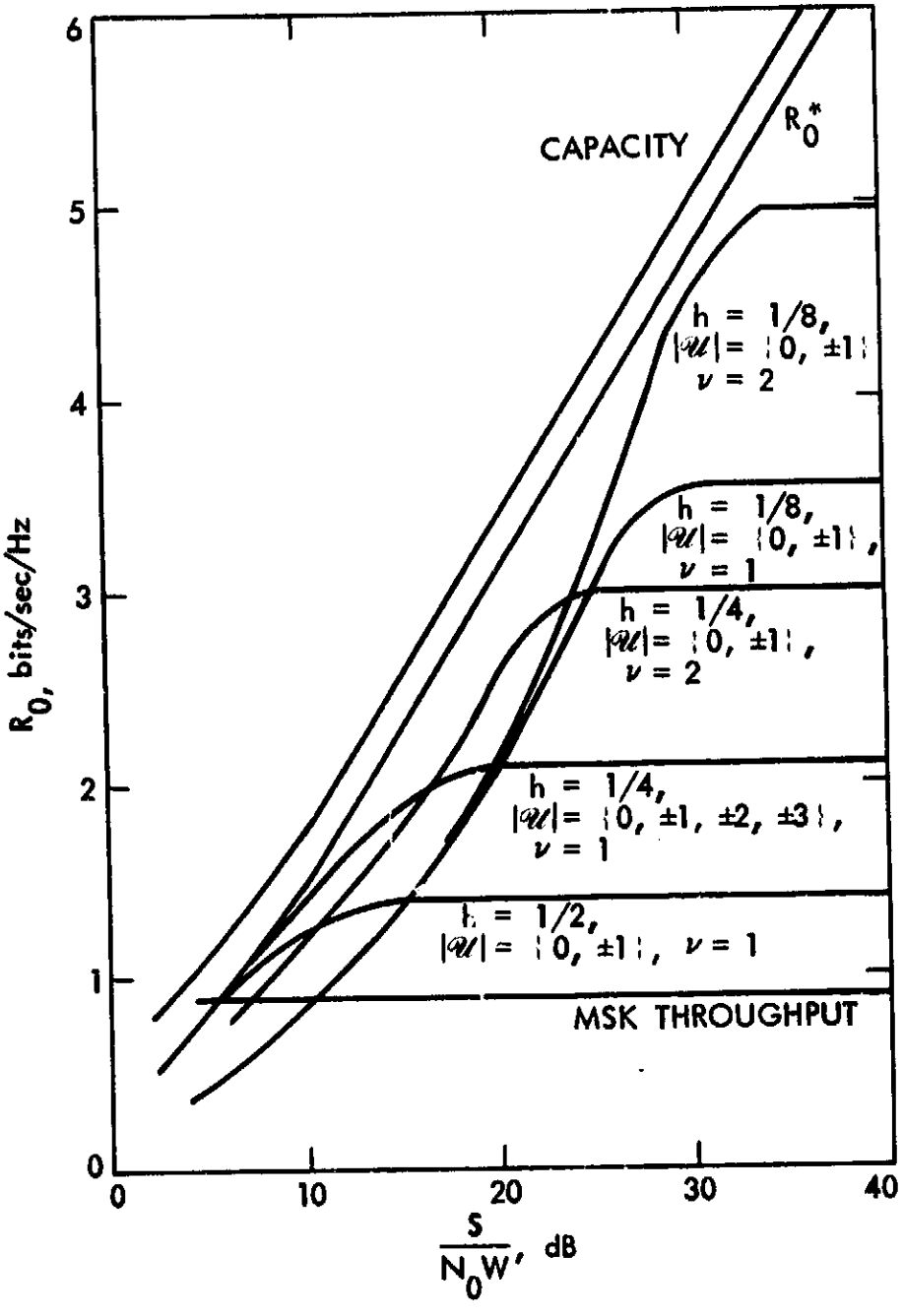


Figure 25. R_0 versus Signal-to-Noise Ratio for Coherent, Maximum-Likelihood Demodulation, 99 percent Bandwidth, Rectangular Filter (Reprinted from [6])

Since from Figure 24, we see that r_0 approximately attains its maximum value for ST/N_0>6 dB, then for S/N_0W >3.7 dB we would have R_0 = .8547. For the other CPM techniques, we observe that it is possible over certain ranges of S/N_0W values to attain close to the maximum practical bandwidth efficiencies as measured by R_0* even though these signals are restricted to be constant envelope.

Figures 26, 27 and 28 are plots of R_0 vs S/N_0W for the more restrictive FCC definition of bandwidth and several different pulse shaping functions. In particular, Fig. 26 corresponds to the rectangular pulse shape of (II.6.57), whereas Fig. 27 and 28 correspond to the raised cosine shape

$$g(t) = \begin{cases} \frac{1}{2\nu T}\left[1 - \cos\frac{2\pi t}{\nu T}\right]; & 0 \le t \le T \\ 0; \text{ otherwise} \end{cases} \qquad \text{(II.6.61)}$$

and the triangular shape

$$g(t) = \begin{cases} \frac{2}{\nu^2 T^2} t; & 0 \le t \le \nu T/2 \\ \frac{2}{\nu^2 T^2}(\nu T - t); & \nu T/2 \le t \le \nu T \\ 0; \text{ otherwise} \end{cases} \qquad \text{(II.6.62)}$$

respectively. In these figures the curves are labelled XZZY where

Y = ν, pulse memory in symbols
[duration of g(t)]

ZZ = { RE, rectangular pulse
RC, raised cosine pulse
TR, triangular pulse

X = { 1 h = 1/2, |U| = {0,+1}
2 h = 1/4, |U| = {0,+1,+2,+3}
3 h = 1/8, |U| = {0,+1,+2,+3,+4,+5} } (II.6.63)

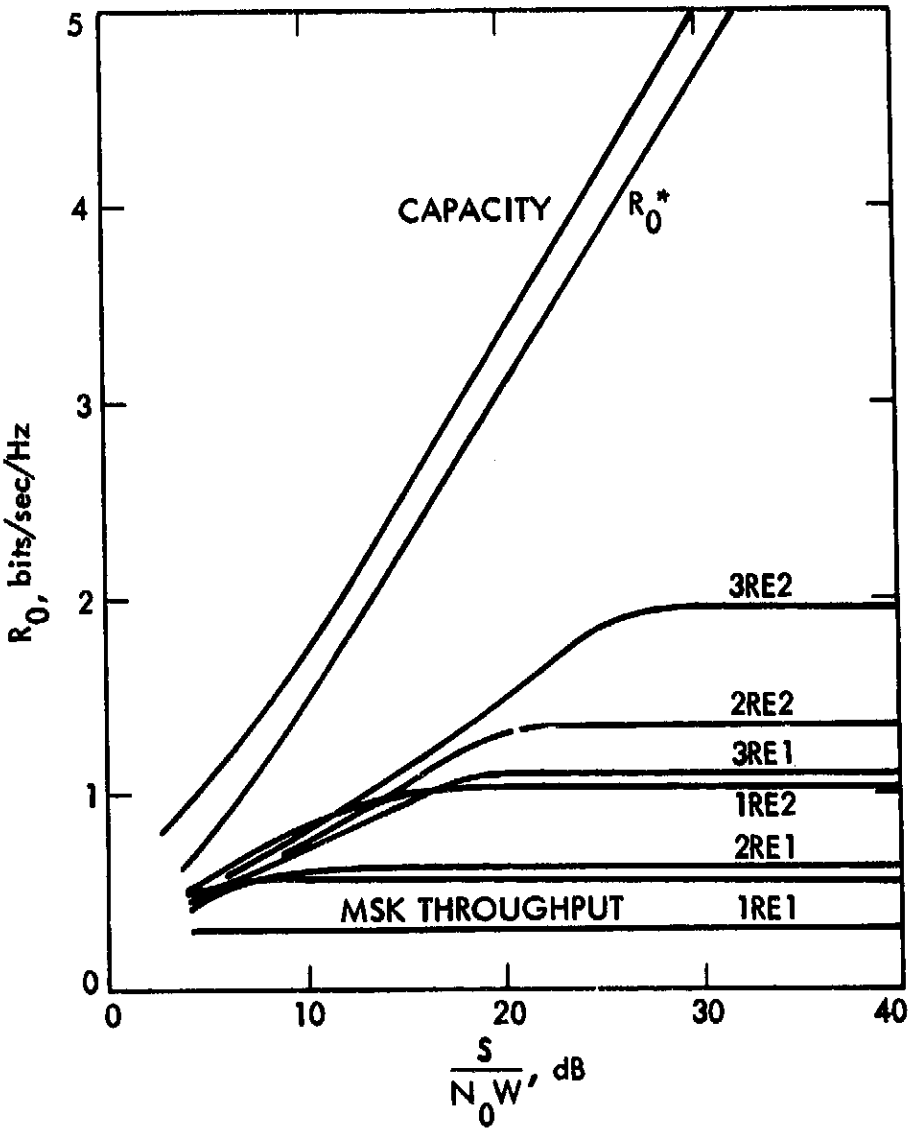


Figure 26. R_0 versus Signal-to-Noise Ratio for Coherent, Maximum-Likelihood Demodulation, FCC Bandwidth Rectangular Filter (Reprinted From [6])

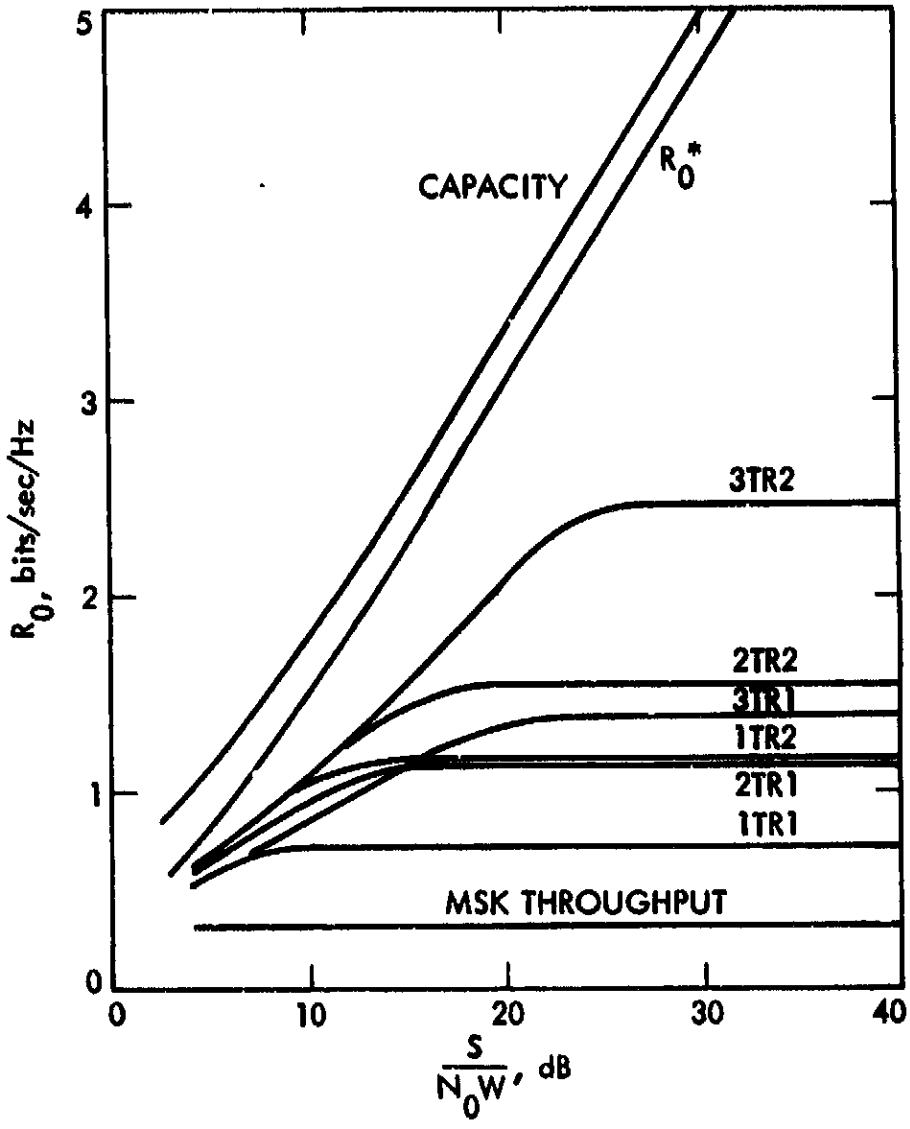


Figure 37. R_0 versus Signal-to-Noise Ratio for Coherent, Maximum-Likelihood Demodulation, FCC Bandwidth, Triangular Filter (Reprinted From [6])

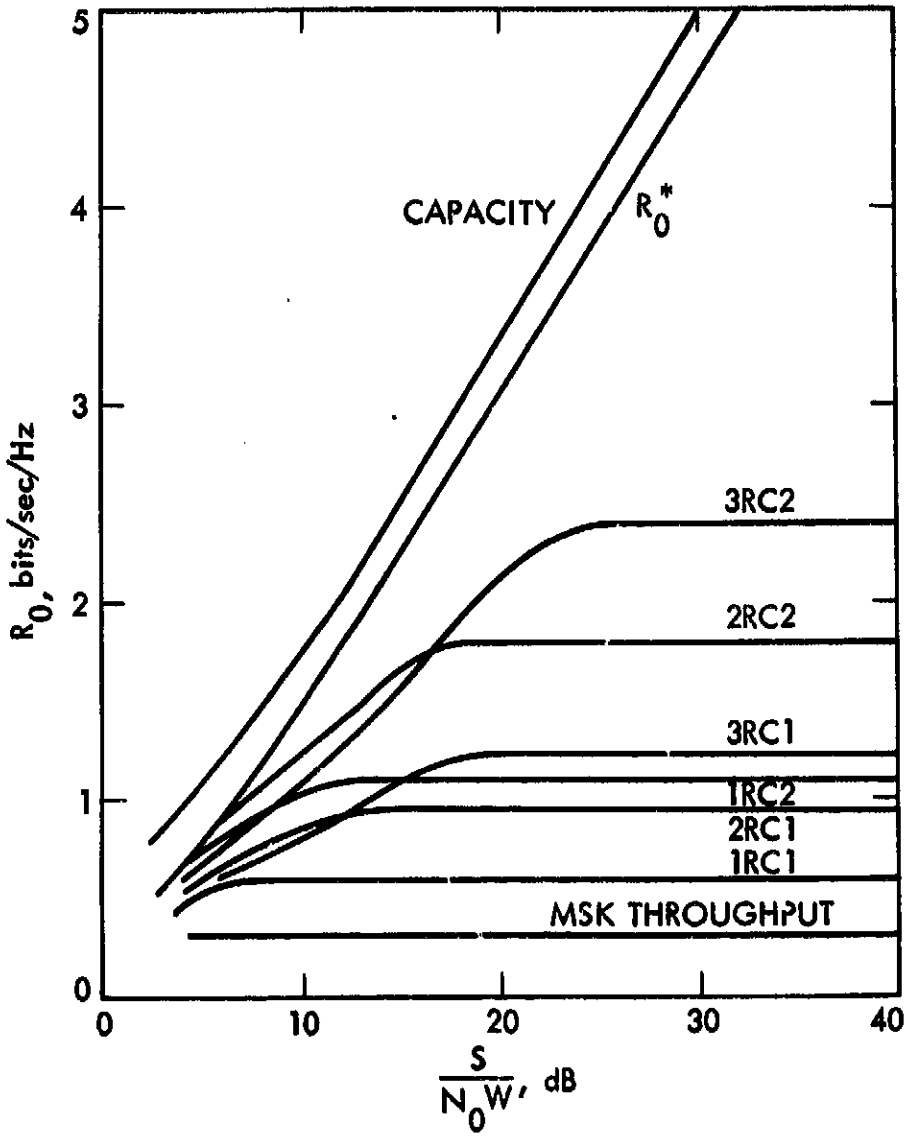


Figure 28. R_0 versus Signal-to-Noise Ratio for Coherent, Maximum-Likelihood Demodulation, FCC Bandwidth, Raised Cosine Filter (Reprinted From [6])

6.2 Transfer Function Bounds

Transfer function bounds were originally introduced by Viterbi [10] for evaluating the bit error probabilities of specific convolutional codes used in symmetric memoryless channels. Later these were applied to finding error probability bounds for amplitude-modulated signals used in intersymbol interference channels [11]. In Appendix A of Part IV we have generalized this approach so as to apply to almost all situations involving the Viterbi algorithm. We briefly summarize its salient features here and present, as an example, its application to MSK.

The generalized transfer function bound applies to discrete time systems where the signal is described by the general form

$$x_k = f(s_k, u_k)$$

$$s_{k+1} = g(s_k, u_k) \tag{11.6.64}$$

where the state s_k at time belongs to a finite set S of size $|S|$. Here $\{u_k\}$ is some i.i.d. sequence, usually a data sequence. The channel is characterized by

$$y_k = h(x_k, n_k) \tag{11.6.65}$$

where $\{n_k\}$ is an i.i.d. sequence that is independent of the data sequence $\{u_k\}$. Since we can assume any branch metric of the general form $m(y_k; s_k, u_k)$, then, given the channel output sequence $\{y_k\}$, the best estimate of the sequence $\{u_k\}$ that minimizes the total sum

$$\sum_{k=-\infty}^{\infty} m(y_k; s_k, u_k)$$

can be found using the Viterbi algorithm. Note that this is not necessarily a maximum-likelihood or maximum a-posteriori metric.

To evaluate the performance of the Viterbi algorithm we first define a distortion measure

$$d((\hat{s}_k,\hat{u}_k),(s_k,u_k)) \text{ for all } k$$

where $\{s_k\}$, $\{u_k\}$ are the actual state and data sequences and $\{\hat{s}_k\}$, $\{\hat{u}_k\}$ are the corresponding sequences found by the Viterbi algorithm. The transfer function bound is a bound on the average distortion

$$E\{d((\hat{s}_k,\hat{u}_k),(s_k,u_k))\}$$

where the expectation is over all sequences $\{u_k\}$ and channel noise sequences $\{n_k\}$. This measure of distortion is general and has as a special case the usual error distortion measure

$$d((\hat{s}_k,\hat{u}_k),(s_k,u_k)) = \begin{cases} 1; & \hat{u}_k \neq u_k \\ 0; & \hat{u}_k = u_k \end{cases} \qquad \text{(II.6.66)}$$

If the data sequence is binary, then the expected value of this distortion $\bar{d}$ is the average bit error probability.

The average distortion is given by the general expression

$$\bar{d} = E\{d((\hat{s}_k,\hat{u}_k),(s_k,u_k))\}$$

$$= \sum_{\underline{s}} q(\underline{s}) E\{d((\hat{s}_k,\hat{u}_k),(s_k,u_k))|\underline{s}\} \qquad \text{(II.6.67)}$$

where $\underline{s}$ is the actual transmitted state sequence and $q(\underline{s})$ is its probability. To evaluate the expectation required in (II.6.67) we examine the probability that the Viterbi algorithm selects the data sequence $\hat{\underline{u}}$ with corresponding state sequence $\hat{\underline{s}}$ given that $\underline{u}$ and $\underline{s}$ are the actual transmitted sequences. This probability is upper (union) bounded by pair-wise error probabilities of the form

$$Pr(\underline{u},\hat{\underline{u}}) = Pr\left\{\sum_{k=-\infty}^{\infty} m(y_k;\hat{s}_k,\hat{u}_k) \geq \sum_{k=-\infty}^{\infty} m(y_k;s_k,u_k)|\underline{u}\right\} \qquad \text{(II.6.68)}$$

which can be further bounded using the Chernoff bound

$$\Pr(u \neq \hat{u}) \leq E\left\{\exp\left[\lambda \sum_{k=-\infty}^{\infty} (m(y_k;\hat{s}_k,\hat{u}_k) - m(y_k;s_k,u_k))\right] \middle| \underline{u}\right\}$$

$$= \prod_{k=-\infty}^{\infty} D_\lambda((\hat{s}_k,\hat{u}_k),(s_k,u_k)) \qquad \text{(II.6.69)}$$

where

$$D_\lambda((\hat{s}_k,\hat{u}_k),(s_k,u_k)) = E\left\{\exp\left[\lambda \sum_{k=-\infty}^{\infty} (m(y_k;\hat{s}_k,\hat{u}_k) - m(y_k;s_k,u_k))\right]\right\} \qquad \text{(II.6.70)}$$

is a conditional expectation given the actual state s_n and data u_n at the beginning of the nth interval. The parameter λ is non-negative and can be chosen to minimize the Chernoff bound.

The parameters $\{D_\lambda((\hat{s}_n,\hat{u}_n),(s_n,u_n))\}$ play a key role in the evaluation of the upper bound to the average distortion. To evaluate these terms, we define super states S_k consisting of pairs of states $(s_k,\hat{s}_k)$ and super inputs U_k consisting of input pairs $(u_k,\hat{u}_k)$. Next we consider all possible transitions from super states to super states for all possible super inputs. The average distortion then has the bound

$$\bar{d} \leq 1/2 \left.\frac{dT(z)}{dz}\right|_{z=1} \qquad \text{(II.6.71)}$$

where

$$T(z) = \sum_{j=1}^{\infty} \sum_{S(0,j)} p(S_0) \prod_{k=0}^{j-1} z^{d(S_k,U_k)} q(U_k) D_\lambda(S_k,U_k) \qquad \text{(II.6.72)}$$

In (11.6.72), S_k is the kth super state, U_k is the kth super input, $d(S_k,U_k)$ is the distortion associated with the kth time interval, $p(\cdot)$ and $q(\cdot)$ are probabilities associated with super states and super inputs, and $S(0,j)$ is the set of all nonzero super state sequences.

The bound in (11.6.71) can be reformulated in matrix form where the key matrix A has elements a_{ij} representing the super state transitions, i.e.,

$$a_{ij} = \begin{cases} \sum_{U} q(U) D_{\lambda}(\delta_i, U); & \text{If } U \text{ exists such that super state } \delta_j \text{ can be reached from super state } \delta_i. \\ 0; & \text{otherwise} \end{cases} \qquad (11.6.73)$$

The details of this formulation are discussed in Part IV, Appendix A.

For the special case of CPM, and the error distortion in (11.6.66), the evaluation of the transfer function bound on the bit error probability is simplified because $D_{\lambda}((\hat{s}_k,\hat{u}_k),(s_k,u_k))$ depends only on the difference Δ_k states and difference ϵ_k between inputs as defined in (11.6.22) and (11.6.20) respectively. Thus, the key matrix is a <u>difference</u> state transition matrix analogous to the <u>super</u> state transition matrix discussed above only having much smaller dimension. In the following section, we illustrate this with the MSK example.

6.2.1 Coherent MSK

Consider now the case of MSK modulation with an ideal phase reference. For this case, the exact bit error probability is given by*

$$P_b = 2Q\left(\sqrt{\frac{2E_b}{N_0}}\right) - Q^2\left(\sqrt{\frac{2E_b}{N_0}}\right) \qquad (11.6.74)$$

*Note that if the data is first precoded (using a differential decoding operation), then the bit error probability performance of MSK will become the same as that of coherent BPSK namely $P_b = Q(\sqrt{2E_b/N_0})$.

where

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-y^2/2)dy \qquad (11.6.75)$$

is the Gaussian probability integral. We now wish to compare this exact result with the transfer function bound on bit error probability.

Recalling from Section 4.2 that the Viterbi demodulation algorithm for MSK uses the maximum-likelihood metric of (11.4.33), then the Chernoff bound of (11.6.69) becomes the Bhattacharyya bound and is minimized for $\lambda = 1/2$. Thus, comparing (11.6.69) with (11.6.25), we immediately observe that

$$D_{1/2}((\hat{s}_k,\hat{u}_k),(s_k,u_k)) \triangleq D_{1/2}(\Lambda_k;\epsilon_k) = f(\epsilon_k;\Lambda_k) \qquad (11.6.76)$$

where $f(\epsilon_k;\Lambda_k)$ is defined in (11.6.24). Furthermore the difference state transition diagram is shown in Figure 29a which is identical to Figure 23 except that the tranches have now been additionally labelled with the appropriate difference state transition matrix elements a_{ij}. These matrix elements are given by*

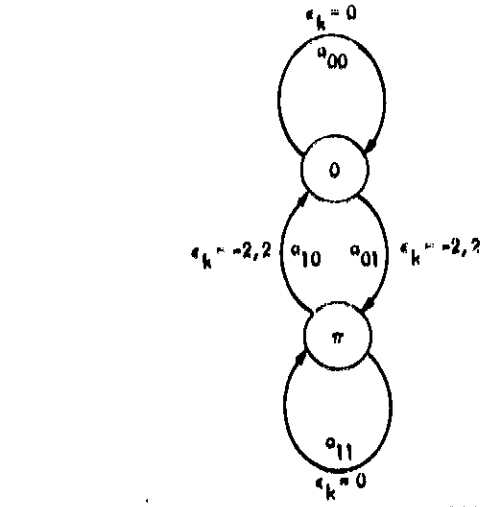
$$a_{00} = 1/2 \; ; \quad a_{11} = 1/2 \exp\left(-\frac{ST}{N_0}\right)$$

$$a_{01} = a_{10} = z \exp\left(-\frac{ST}{2N_0}\right) \qquad (11.6.77)$$

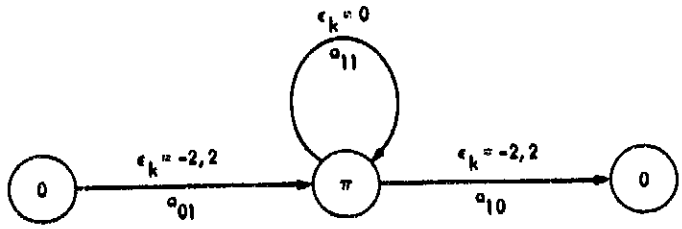
*Normally the kth transition has the form

$$1/2\, z^{d(\epsilon_k)} D_{1/2}(\Lambda_k;\epsilon_k).$$

However, transitions from 0 to π and π to 0 correspond to $\epsilon_k = -2$ or $\epsilon_k = 2$. Thus the 1/2 factor is added twice in a_{01} and a_{10}.



(a) DIFFERENCE STATE DIAGRAM



(b) MODIFIED DIFFERENCE STATE DIAGRAM

Figure 29. MSK Error State Diagrams

Next the modified difference state transition diagram is illustrated in Figure 29b where by inspection the transfer function is simply

$$T(z) = \frac{a_{01}\, a_{10}}{1 - a_{11}}$$

$$= \frac{z^2 \exp(-\frac{ST}{N_0})}{1 - 1/2 \exp(-\frac{ST}{N_0})} \qquad (11.6.78)$$

Thus, equating the average distortion of (11.6.71) with the average bit error probability, we obtain the bound

$$P_b \leq 1/2 \left.\frac{dT(z)}{dz}\right|_{z=1}$$

$$= \frac{\exp\left(-\frac{E_b}{N_0}\right)}{1 - 1/2 \exp\left(-\frac{E_b}{2N_0}\right)}$$

(II.6.79)

This result can be extended to the simultaneous phase/data demodulation receiver for MSK modulation. The primary change is that the signal state is now enlarged to the number of quantization levels Q for the phase space $[0,2\pi]$. Figure 30 shows the exact bit error probability for MSK as computed from (II.6.74) along with the transfer function bound for ideal (perfect phase reference) coherent MSK as computed from (II.6.79). Also shown are curves for simultaneous phase data/demodulation with Q=32 and Q=16. For the latter two cases, the unknown phase is assumed constant with a uniform distribution over $[0,2\pi]$. Note that Q=32 results in a very small degradation as compared to the ideal coherent case.

When a convolutional code is used together with any of these modulations with memory, the overall encoder and modulator together can be described as a finite state signal process and the maximum-likelihood receiver would again be realized with a Viterbi algorithm. This algorithm would combine the two operations of demodulation and decoding. The generalized transfer function bound approach can also be applied in these convolutionally coded CPM systems.

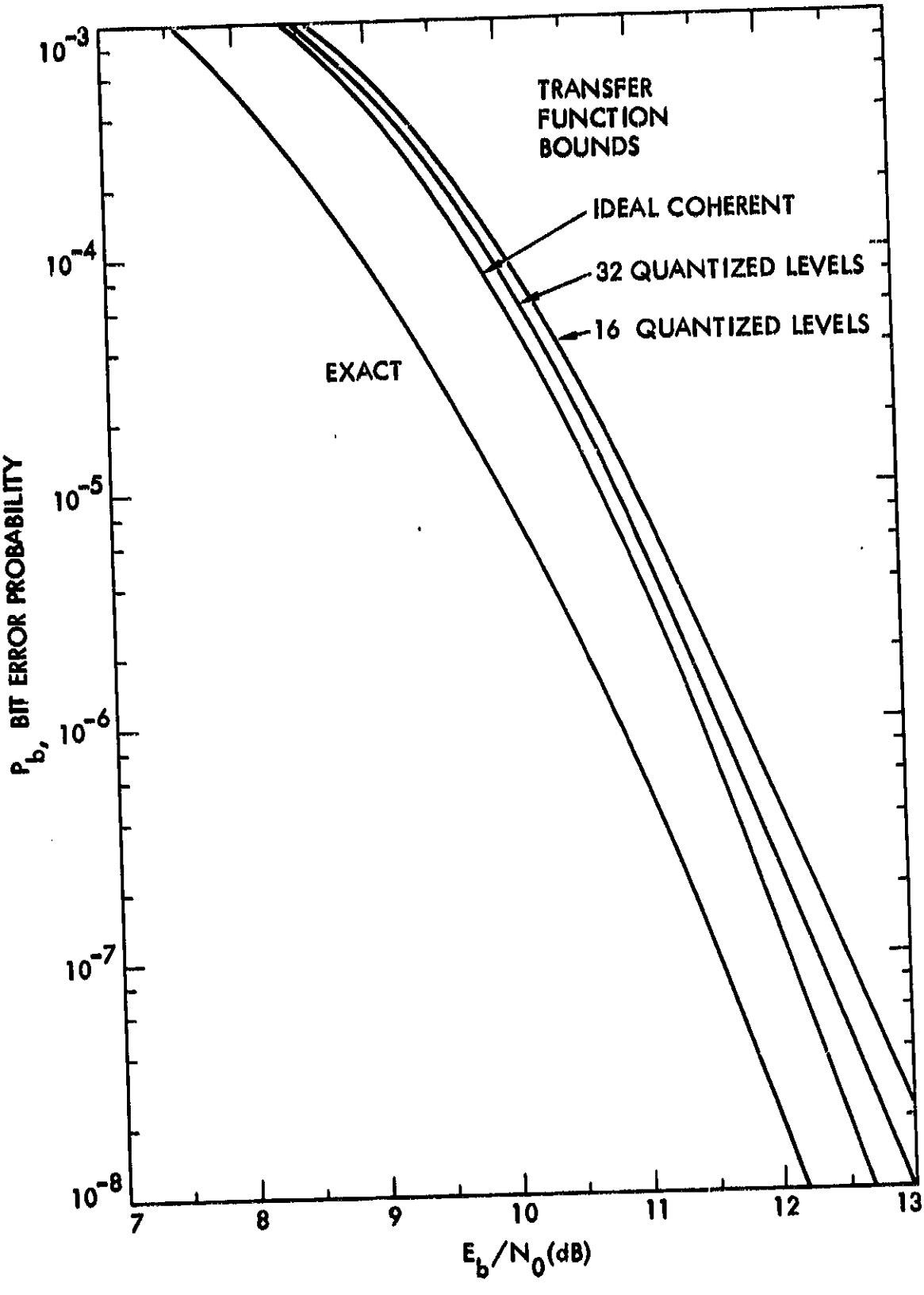


Figure 30. Bit Error Probability Performance of MSK

C-2

REFERENCES

1. Viterbi, A.J., "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm", IEEE Transactions on Information Theory, vol. IT-13, no. 2, April 1967, pp. 260-269.

2. Lender, A., "The Duobinary Technique for High-Speed Data Transmission", 1963 IEEE Winter General Meeting.

3. Kretzmer, E.R., "Generalization of a Technique for Binary Data Communication", IEEE Transactions on Communication Technology, vol. COM-14, no. 2, February 1966, pp. 67-68.

4. Kabal, P., and Pasupathy, S., "Partial Response Signalling", IEEE Transactions on Communications, vol. COM-23, no. 9, September 1975, pp. 921-935.

5. Massey, J.L., "A Generalized Formulation of Minimum Shift Keying Modulation", ICC '81 Conference Record, Seattle, Washington, June 8-12, 1980, pp. 26.5.1-26.5.4.

6. Jackson, D.E., "Bandwidth Efficient Modulation and Coding", Ph.D. Dissertation, UCLA, 1980.

7. Gantmacher, F.R., Applications of the Theory of Matrices, Interscience, New York, N.Y., 1959.

8. Viterbi, A.J., and Omura, J.K., Principles of Digital Communication and Coding, McGraw-Hill Book Company, New York, N.Y., 1979.

9. Divsalar, D., "Performance of Mismatched Receivers on Bandlimited Channels", Ph.D. Thesis, University of California, Los Angeles, 1978, Appendix B.

10. Viterbi, A.J., "Convolutional Codes and their Performance in Communication Systems", IEEE Transactions on Communication Technology, vol. COM-19, no. 5 (Part II), pp. 751-772.

11. Forney, G.D., "The Viterbi Algorithm", Proceedings of the IEEE, vol. 61, pp. 268-278.

12. Shannon, C.E., "Probability of Error for Optimal Codes in a Gaussian Channel", B.S.T.J., vol. 38, no. 3, May 1959, pp. 611-656.

13. Wozencraft, J.M., and Jacobs, I.M., Principles of Communications Engineering, John Wiley and Sons, Inc., New York, N.Y., 1968.