AD-A031 559 ILLINOIS UNIV AT URBANA-CHAMPAIGN COORDINATED SCIENCE LAB F/G 17/2
MODELS AND EFFICIENT RECEIVERS FOR COMMUNICATION CHANNELS WITH --ETC(U)
SEP 76 P K VARSHNEY DAAB07-72-C-0259
UNCLASSIFIED R-744 NL

1 OF 2
AD
A031559

FG.

ADA031559

REPORT R-744 SEPTEMBER 1976 UILU-ENG 76-2232

CSL COORDINATED SCIENCE LABORATORY

12

MODELS AND EFFICIENT RECEIVERS FOR COMMUNICATION CHANNELS WITH MEMORY

PRAMOD KUMAR VARSHNEY

UNIVERSITY OF ILLINOIS - URBANA, ILLINOIS

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) MODELS AND EFFICIENT RECEIVERS FOR COMMUNICATION CHANNELS WITH MEMORY,		5. TYPE OF REPORT & PERIOD COVERED (9) Technical Report,
		6. PERFORMING ORG. REPORT NUMBER R-744; UILU-ENG 76-2232
7. AUTHOR(s) (10) Pramod Kumar Varshney		8. CONTRACT OR GRANT NUMBER(s) (15) DAAB-07-72-C-0259, DAAG-29-76-G-0154
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory University of Illinois at Urbana-Champaign Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Joint Services Electronics Program		12. REPORT DATE (11) Sep 1976
		13. NUMBER OF PAGES 130
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) (14) R-744, UILU-ENG-76-2232		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. (12) 140p.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Digital Channels
Digital Communications
Channel Models
Receivers
Fading Channels

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In the present work, some aspects of digital communications over channels with memory are studied. The basic objective is to model channels with memory and to develop more efficient and reliable communication techniques over such channels. In particular, the receiver design problem for channels with memory is considered and then the relationship of receivers and channel models is examined. In the receiver design problem, fading is assumed to be the source of channel memory. First, the idea of receivers with memory is

over

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

097700

20. ABSTRACT (continued)

investigated and a suboptimal receiver with one-bit memory is derived which performs better than the optimal receiver without memory. A receiver with large memory is devised which consists of an estimator and a detector. The decision rule adapts to the channel conditions based on the information provided by the estimator. The performance of the receiver is examined by computing the average probability of error. The estimator is a limited-memory decision-feedback estimator, the estimation criterion being the MMSE. Finally, a methodology is described which can be used to develop channel models based on the physical processes involved. Two measures are defined which quantitatively characterize the correlation of errors and then the relationship of the receiver design problem to channel modeling is examined.

MODELS AND EFFICIENT RECEIVERS
FOR COMMUNICATION CHANNELS WITH MEMORY

BY

PRAMOD KUMAR VARSHNEY
B.S., University of Illinois, 1972
M.S., University of Illinois, 1974

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1976

Thesis Adviser: Professor A. H. Haddad

Urbana, Illinois

MODELS AND EFFICIENT RECEIVERS
FOR COMMUNICATION CHANNELS WITH MEMORY

Pramod Kumar Varshney, Ph.D.
Coordinated Science Laboratory and
Department of Electrical Engineering
University of Illinois at Urbana-Champaign, 1976

In the present work, some aspects of digital communications over channels with memory are studied. The basic objective is to model channels with memory and to develop more efficient and reliable communication techniques over such channels. In particular, the receiver design problem for channels with memory is considered and then the relationship of receivers and channel models is examined. In the receiver design problem, fading is assumed to be the source of channel memory. First, the idea of receivers with memory is investigated and a suboptimal receiver with one-bit memory is derived which performs better than the optimal receiver without memory. A receiver with large memory is devised which consists of an estimator and a detector. The decision rule adapts to the channel conditions based on the information provided by the estimator. The performance of the receiver is examined by computing the average probability of error. The estimator is a limited-memory decision-feedback estimator, the estimation criterion being the MMSE. Finally, a methodology is described which can be used to develop channel models based on the physical processes involved. Two measures are defined which quantitatively characterize the correlation of errors and then the relationship of the receiver design problem to channel modeling is examined.

ACKNOWLEDGMENTS

I would like to thank my advisor Professor A. H. Haddad for his encouragement, constructive suggestions and guidance throughout the course of this work. Thanks are extended to Professor M. B. Pursley for enlightening discussions on the subject. I am grateful to all of my teachers, colleagues and friends for a pleasant and fruitful stay at the University of Illinois. The assistance of Mrs. Hazel Corray, Mrs. Rose Schmidt and Mrs. Trudy Williams in the preparation of the manuscript is gratefully acknowledged.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. General

Error-free and reliable transfer of information is the basic goal of any communication system. Analog communication has been the traditional mode of communication but in the recent past digital communication has become increasingly popular. There are several reasons for this trend towards digital communication. One of the prime reasons for this gradual shift is the feasibility of relatively error-free transmission over long distances. Digital communication systems may also capitalize on the recent revolution in the integrated circuit technology and are, thus, cost effective. Some of the other features of digital communication systems are high speed transmission and error-control capabilities. In the present work, only digital communication systems are considered because of the current interest in such systems. A block diagram of a typical digital communication system is shown in Fig. 1.1.

Digital communication systems have conventionally been assumed to be memoryless. This assumption results in notational convenience and it also simplifies the analysis. In practice, however, most digital communication systems exhibit memory. By memory of a digital communication system, it is meant that the digital data at the receiving end is correlated. The major sources of the statistical dependence are the source, the channel encoder and the channel itself. Quite frequently the source is assumed to produce independent bits which is not valid in practice. Most of the practical sources generate statistically dependent data sequence to be transmitted. The source encoder, which has not been shown explicitly in Fig. 1.1, is considered

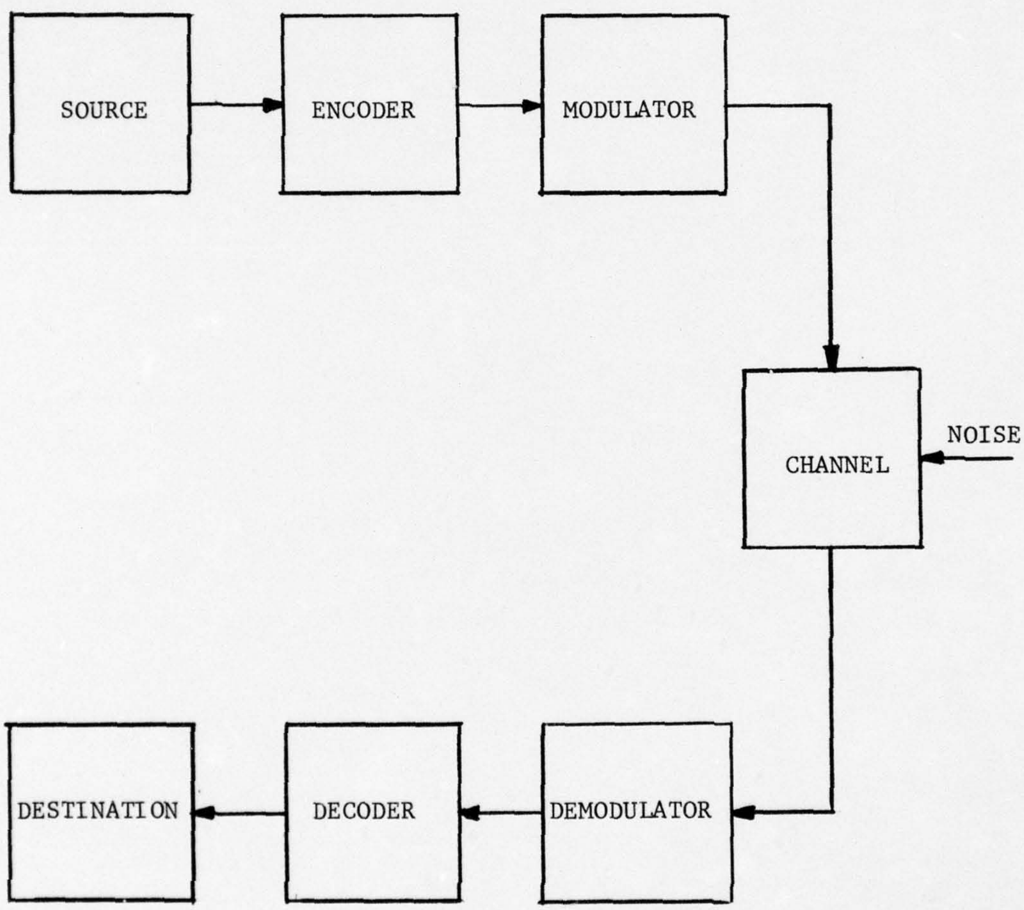

Figure 1.1. Block diagram of a digital communication system.

to be a part of the source and it also contributes to the statistical dependence of the data input to the channel encoder. Deterministic redundancy is introduced in the channel encoder for error-control. In both block coding and convolutional coding, parity check bits are added to every k information bits and a total of n bits are transmitted. The physical processes responsible for the channel memory are intersymbol interference, fading and correlated noise. The dependence introduced by the source and the channel encoder are beyond the scope of the work presented here and, therefore, these two have not received any further consideration. Thus, channel is considered to be the only source of memory in the received data sequence and communication over such channels is the subject of this thesis.

1.2. Channels with Memory

Channels with memory have been studied quite extensively in the literature. Development of reliable and more efficient communication schemes has been an area of intense research activity as evidenced by the survey article of Lucky [1]. A considerable amount of attention has been focussed on the problem of devising efficient detection schemes. The standard reception schemes assume the independence of the received data bits and a bit-by-bit detection without memory is performed, the design criterion being the minimization of the average probability of error. In most of the practical digital communication systems, however, the independence assumption is not valid due to the channel memory. The received data is correlated and it is expected that receivers with memory which exploit this dependence would yield better performance.

Intersymbol interference has traditionally been treated as the source of channel memory and the receiver design problem for such channels has been studied extensively [2-38]. Several classes of equalization techniques have been proposed. A brief summary of the effort in the general area of receiver design for intersymbol interference channels is presented here. The adaptive equalization systems were first formalized by Lucky [2,3]. An adaptive mean-square equalizer is a finite-length transversal filter whose coefficients are adjusted using decision-directed estimation loops so as to minimize mean-squared error (MSE) in the output samples. Proakis and Miller [4] and Gersho [5] have analyzed the structure with respect to convergence time and tap gain error. Several authors [6-10] have examined techniques for faster convergence. Structures which are different from the basic linear equalizer [1] have also been considered, such as in [11-17]. Nonlinear equalizers have also been considered. Several suboptimal nonlinear receivers have been explored in the literature [18-21]. Decision-feedback equalizer has been treated extensively. The output samples $\hat{a}_n$ are fed back through a transversal filter to subtract out the tails of these pulses from subsequent pulses. Austin [22] examined this "bootstrap" system and the optimization for minimum mean-squared error (MMSE) was performed by Monsen [23]. More recently Monsen [24] has compared the performance of a decision feedback equalizer and a linear equalizer. Price [25] considered the receiver in the zero-forcing mode of operation. Salz [26], Clark [27] and George, Bowen and Storey [28] have also treated different variations of the basic decision-feedback equalizer. Maximum-likelihood sequence estimation in the presence of intersymbol interference has been a popular area of investigation [29-38]. Chang and Hancock [29] devised an optimum decision procedure in which a decision was made on the basis of the complete message. Abend and Fritchman [30] modified the procedure and derived an optimum sequential compound detector under a fixed delay constraint. They considered optimal bit-by-bit detection based on the sequence received in the past. Gonsalves [31] has also considered a related problem. The Viterbi algorithm generated considerable interest and it found applications in receiver design for intersymbol interference channels [32-37], with the objective of finding the most likely sequence transmitted. The complexity increases for larger sequences. Recently Yao and Milstein [38] have considered a maximum likelihood bit detector in the presence of intersymbol interference.

The other major source of channel memory, fading, has received very little attention. The memory of the channel is characterized by the statistical dependence of the received data bits. Fading is assumed to be a slowly varying process (especially for high data rates) and its contribution during different signalling intervals is correlated. This correlation could be exploited to yield a receiver with better performance. The present work will concentrate of fading as the source of channel memory and will assume the absence of intersymbol interference. The goal here is, therefore, to derive receiver algorithms which exploit the continuity and correlation of the fading process to yield a better performance in the average probability of error sense. Memory will have to be incorporated in the receivers to attain the goal stated above. In the next section, the organization of this thesis is described and an outline is presented.

1.3. Thesis Outline

The major problem considered in the present work is the receiver design problem for channels with memory. The source of channel memory is assumed to be fading. Chapter two introduces the receiver design problem. The feasibility of receivers with memory is investigated and a receiver with one-bit memory is considered. The statistical correlation of the two adjacent data bits is utilized to achieve a receiver with better performance. The standard design criterion, i.e. the minimization of probability of error, is employed. The performance measure is the average probability of error. Two examples are considered and numerical results obtained indicate that an improvement in the performance is attained. This leads us to expect that receivers with larger memory would perform even better.

An alternative approach to receiver design problem is investigated and receivers with larger memory are considered next. The receiver is assumed to consist of an estimator and a detector. The estimator is responsible for the estimation of uncertain parameters present in the received signal. The estimator is followed by a detector which treats the estimate furnished by the estimator as the correct value of the uncertain parameter and adapts the decision rule to the existing channel conditions. In Chapters three and four the receiver design problem with large memory is discussed. In Chapter three, an estimate of the uncertain parameters is assumed to be available and the attention is directed towards the development of an optimum adaptive detector. The memory length is assumed to be large and asymptotic results are obtained. The optimum decision rule is determined and the performance is examined by considering the probability of error. An example is presented to illustrate the receiver. In Chapter four, the principles of estimator design are discussed. The estimation criterion which results in the optimum receiver is determined. A limited-memory estimator with decision-feedback is derived. The estimate is assumed to be a linear function of the observations with nonlinearity introduced through the coefficients due to decision-feedback. In some communication system applications, the received signal does not always contain the uncertain parameter to be estimated. For these applications, the estimator is to be modified to include this uncertainty about the presence of the parameter to be estimated. This situation arises in the on-off keying systems. This estimation algorithm also finds applications in control theory. While tracking the trajectory of a target, the observation mechanism may fail, for instance, due to misalignment of antennas. Examples are considered and the performance of the estimators is examined by computing the mean-squared error.

In Chapter five, another aspect of digital communications over channels with memory is considered. Modeling of digital channels is an important problem and has been discussed in this chapter. The concept of channel modeling based on the actual physical processes to characterize the input-output behavior of the channel is described. Relationship of receivers with memory and channel models is examined. Some measures to represent the channel memory quantitatively are defined and illustrated by an example. Finally, the work is summarized and the results are presented in the last chapter.

2. RECEIVER WITH ONE-BIT MEMORY

2.1. Introduction

As indicated in the first chapter, the objective of the present work is to design receivers with memory for channels which exhibit memory. The source of channel memory is assumed to be fading and the absence of intersymbol interference is assumed. Signal detection schemes with memory have been considered in the literature [39-42]. The receiver problem in such cases can be formulated as a hypothesis testing problem. A Bayesian approach results in a decision rule in which the likelihood ratio is compared to a threshold to decide as to which hypothesis is true. Cover and Hellman [39-41] have considered the hypothesis testing problem with finite memory. The data is reduced to an m-valued statistic and it is employed to make a decision. The hypotheses themselves are independent of each other. The difference between the design of receivers with memory and the hypothesis testing with memory is that the statistics corresponding to each signalling interval are statistically dependent in the receiver design problem whereas the experiments in the hypothesis testing problem are assumed to be independent. Baxa and Nolte [42] have considered the signal detection problem using Cover's approach. They discuss the problem under the constraint of finite soft memory. The memory is modeled as a finite-state machine and the memory is updated according to a suboptimal time dependent rule. Baxa and Nolte do not consider channels with memory and the statistical correlation of the received bits.

In the present work, receivers with memory are derived for digital communication systems where the received bits are correlated. Fading is assumed to be the disturbance process causing this statistical dependence. In this chapter, a receiver with one-bit memory is investigated so that the decision on a particular bit is based on the statistics and the decision of the previous bit. Channels generally exhibit a larger memory and, therefore, a receiver with a larger memory is expected to perform better but at the expense of implementation complexity. The discussion of such receivers is postponed for later chapters. The objective in this chapter is to investigate the idea of a receiver with memory for fading channels. The emphasis is on exploring the theoretical feasibility of the idea and not on deriving a relatively complex detection algorithm with larger memory. In the next section, the problem is defined and appropriate assumptions are stated. In the following two sections, an optimal solution and a decision-feedback suboptimal scheme are described. Finally, two examples are considered and numerical results are presented.

2.2. Problem Statement

The objective of this chapter is to design a receiver with one-bit memory for fading communication channels. It is assumed that a binary system is operating over a fading channel in the absence of intersymbol interference. The transmitted bits are assumed to be independent with zeros and ones equiprobable. The transmitted signal is $s_0(t)$ or $s_1(t)$ depending on whether a zero or a one is sent. The transmitted signal in any signalling interval $[0,T)$ is given as follows:

$$s_i(t) = \begin{cases} \sqrt{2}\, f_i(t) \cos(\omega_i t + \phi_i(t)) & 0 \leq t < T \\ 0 & \text{Otherwise} \end{cases} \tag{2.1}$$

with

$$\int_0^T f_i^2(t)dt = 1 \tag{2.2}$$

and where ω_i and ϕ_i correspond to frequency and phase modulations respectively. Orthogonal signalling is assumed, i.e.,

$$\int_0^T s_0(t)\ s_1(t)dt = 0 \tag{2.3}$$

The channel is assumed to be corrupted by additive white Gaussian noise $n(t)$ with power spectrum N_0. Fading is assumed to be slow so that the channel gain and phase are constant over the period of one signalling interval and thus may be represented by a sequence of dependent random variable pairs $\{v_k,\theta_k\}$. The sequence $\{v_k,\theta_k\}$ is defined by the expression of the received signal $r(t)$ during the k-th signalling interval $[(k-1)T,kT)$, namely, if H_k^i (the hypothesis that i is transmitted during the k-th interval) is true, then

$$\begin{aligned} r_k(t) &\triangleq r[t+(k-1)T] = v_k\sqrt{2}\ f_i(t)\ \cos(\omega_i t + \phi_i(t) + \theta_k) \\ &\quad + n[t+(k-1)T] \\ &= v_k\ s_i(t,\theta_k) + n_k(t) \quad , \quad 0 \leq t < T,\ i=0,1 \end{aligned} \tag{2.4}$$

Hence, v_k and θ_k denote the contributions of amplitude and phase fading during the k-th signalling period. The statistical dependence in the received data is introduced by the correlation of the sequence $\{v_k,\theta_k\}$. This statistical dependence as a function of the fading process is utilized to derive the detection scheme with memory. A correlation receiver is employed to process the incoming waveform, as shown in the block diagram of Fig. 2.1. The outputs of the integrators which serve as decision statistic are defined by

$$X_k^i = \int_0^T r_k(t)\,\sqrt{2}\,f_i(t)\cos(\omega_i t + \phi_i(t))dt \quad , \quad i=0,1 \tag{2.5a}$$

$$Y_k^i = \int_0^T r_k(t)\,\sqrt{2}\,f_i(t)\sin(\omega_i t + \phi_i(t))dt \quad , \quad i=0,1 \tag{2.5b}$$

The basic form of the desired recursive structure to be optimized is shown in Fig. 2.2. The sequences $\{X_k^0, Y_k^0, X_k^1, Y_k^1\}$ are functions of the sequence $\{v_k, \theta_k\}$, and this relationship is exploited to design the receiver with one-bit memory next.

2.3. An Optimal Solution

In order to obtain the optimal solution for the one-bit memory detection problem, it is treated as a four-hypothesis Bayesian decision problem. For notational convenience, only the coherent reception case (no phase uncertainty) is considered so that under the signal model (2.1) $Y_k^i = 0$, $i=0,1$. Let H_{ij} denote the hypothesis that H_{k-1}^i and H_k^j are true, i.e. during (k-1)-th signalling interval i is transmitted and j is sent during the k-th signalling interval. The four hypotheses are defined as follows:

$$H_{00}: \quad \begin{aligned} r_{k-1}(t) &= v_{k-1}s_0(t) + n_{k-1}(t) \\ r_k(t) &= v_k s_0(t) + n_k(t) \end{aligned} \tag{2.6a}$$

$$H_{01}: \quad \begin{aligned} r_{k-1}(t) &= v_{k-1}s_0(t) + n_{k-1}(t) \\ r_k(t) &= v_k s_1(t) + n_k(t) \end{aligned} \tag{2.6b}$$

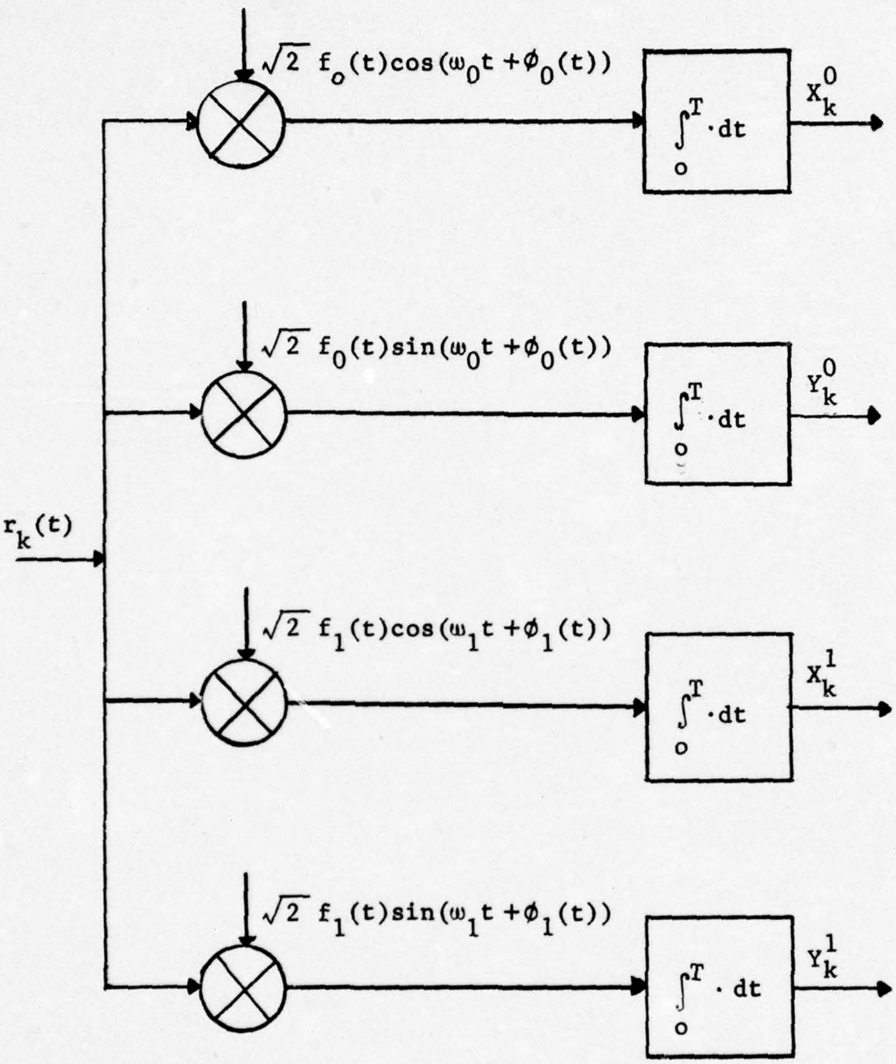

Figure 2.1. Block diagram of the demodulator.

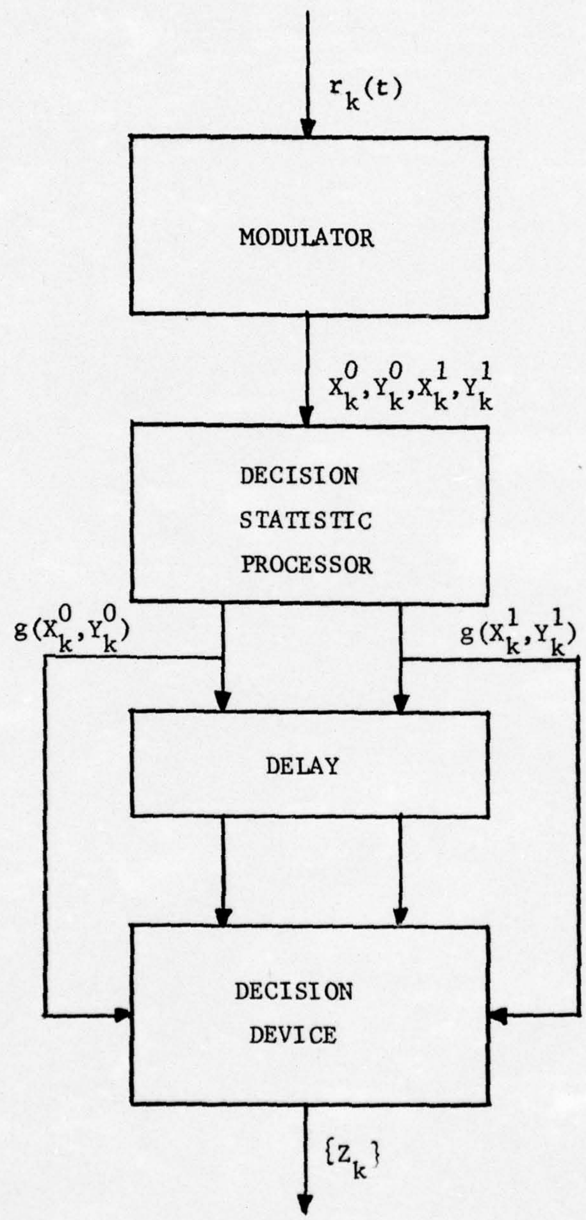

Figure 2.2. Optimum receiver structure.

$$H_{10}: \quad r_{k-1}(t) = v_{k-1}s_1(t) + n_{k-1}(t)$$
$$r_k(t) = v_k s_0(t) + n_k(t) \tag{2.6c}$$

$$H_{11}: \quad r_{k-1}(t) = v_{k-1}s_1(t) + n_{k-1}(t)$$
$$r_k(t) = v_k s_1(t) + n_k(t) \tag{2.6d}$$

Bayesian approach is used to obtain the optimal solution with the Hamming distance as the cost function. This is equivalent to minimizing the conditional probability of error given the previous bit statistics $P(e_k|X^0_{k-1},X^1_{k-1})$ where $\{X^0_k\}$ and $\{X^1_k\}$ are defined by (2.5). If Ω_0 and Ω_1 represent the optimum decision regions in the (X^0_k,X^1_k) plane for the hypotheses H^0_k and H^1_k, then the Bayes' risk may be expressed as

$$\mathcal{R} = \frac{1}{4}\Big[\iint_{\Omega_1} f_{00}(x,y|X^0_{k-1},X^1_{k-1})dxdy + \iint_{\Omega_1} f_{10}(x,y|X^0_{k-1},X^1_{k-1})dxdy$$
$$+ \iint_{\Omega_0} f_{01}(x,y|X^0_{k-1},X^1_{k-1})dxdy + \iint_{\Omega_0} f_{11}(x,y|X^0_{k-1},X^1_{k-1})dxdy\Big] \tag{2.7}$$

where $f_{ij}(x,y|X^0_{k-1},X^1_{k-1})$ denotes the conditional density of X^0_k,X^1_k given X^0_{k-1},X^1_{k-1} and hypothesis H_{ij}. The conditional densities are given by

$$f_{ij}(x,y|X^0_{k-1},X^1_{k-1}) = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h^0_k(x)h^0_{k-1}(u)h^1_k(y)h^1_{k-1}(v)f_{v_k v_{k-1}}(v_k,v_{k-1})dv_k dv_{k-1}}{\int_{-\infty}^{\infty} h^0_{k-1}(u)h^1_{k-1}(v)f_{v_{k-1}}(v_{k-1})dv_{k-1}} \tag{2.8}$$

where $f_{X^j_k|v_k,H^j_k}(x)$, denoted by $h^j_k(x)$, is the conditional density of X^j_k given the fading parameter v_k and the hypothesis H^j_k. It should be noted that the conditional independence of X^0_k, X^1_k, X^0_{k-1} and X^1_{k-1} has been utilized in obtaining (2.8). Furthermore, the conditional densities $h_k^j(x)$ are Gaussian and thus the resulting densities used in (2.7) may be easily derived if an appropriate model for the density of the fading sequence $\{v_k\}$ is assumed. The decision regions Ω_i, i=0,1, obtained by the minimization of the Bayes' risk given in (2.7) or equivalently obtained by the minimization of the conditional probability of error $P(e_k|X_{k-1}^0,X_{k-1}^1)$ are given by the following likelihood ratio test:

$$\Lambda = \frac{f_{01}(x,y|X_{k-1}^0,X_{k-1}^1) + f_{11}(x,y|X_{k-1}^0,X_{k-1}^1)}{f_{00}(x,y|X_{k-1}^0,X_{k-1}^1) + f_{10}(x,y|X_{k-1}^0,X_{k-1}^1)} \underset{H_k^0}{\overset{H_k^1}{\gtrless}} 1 \tag{2.9}$$

which is seen, of course, to depend on the previous bit statistics X_{k-1}^0 and X_{k-1}^1. Equation (2.9) determines the optimum decision regions in the two-dimensional space (X_k^0,X_k^1) as a function of variables X_{k-1}^0 and X_{k-1}^1. The resulting decision rule may also be expressed in terms of the likelihood ratios Λ_i and is given by

$$\Lambda_1 + \Lambda_3 \underset{H_k^0}{\overset{H_k^1}{\gtrless}} 1 + \Lambda_2 \tag{2.10}$$

where Λ_i are defined as follows

$$\Lambda_1 = f_{01}(x,y|X_{k-1}^0,X_{k-1}^1)/f_{00}(x,y|X_{k-1}^0,X_{k-1}^1) \tag{2.11a}$$

$$\Lambda_2 = f_{10}(x,y|X_{k-1}^0,X_{k-1}^1)/f_{00}(x,y|X_{k-1}^0,X_{k-1}^1) \tag{2.11b}$$

$$\Lambda_3 = f_{11}(x,y|X_{k-1}^0,X_{k-1}^1)/f_{00}(x,y|X_{k-1}^0,X_{k-1}^1) \tag{2.11c}$$

It is observed that the optimum detection scheme with one-bit memory is quite complex to implement because the decision rule depends upon the level of the previous output. However, the decision regions are different from the ones for the memoryless scheme which indicates that an improvement is feasible if a detection scheme with memory is employed. Therefore, in the next section a suboptimal detection algorithm with memory is considered which keeps the essential features of the optimal scheme in that it minimizes the bit-error probability and is simpler to implement.

2.4. Suboptimal Scheme

The suboptimal scheme considered in this section is a decision-feedback scheme for which it is assumed that the system error probability is sufficiently low so that the received digits can be assumed to be correct. This simplifies the receiver structure considerably. There is, however, a possibility of error propagation but it is not very serious. The problem of error propagation in decision-directed schemes has been considered previously by Davisson and Schwartz [43]. The suboptimal scheme may also be considered as a simplification of the optimal scheme in that the k-th bit decision depends only on the two-level quantized value of the previous signalling interval statistic. Namely, the k-th bit decision is assumed to be a function of the decision on the previous bit and not of the variables X^0_{k-1}, X^1_{k-1}. In the suboptimal scheme, two different likelihood ratios λ^0_k and λ^1_k are employed for the decision on the previous bit. If the output of the receiver is represented by the binary sequence $\{Z_k\}$, the expressions for the likelihood ratio $\lambda_k(Z_{k-1})$ can be written as

$$\lambda_k(Z_{k-1}) = \begin{cases} \lambda_k^0 & \text{if } Z_{k-1} = 0 \\ \lambda_k^1 & \text{otherwise} \end{cases} \tag{2.12}$$

where

$$\lambda_k^0 = \frac{f_{X_k^0 X_k^1 | Z_{k-1}, H_k^1}(x,y|Z_{k-1} = 0)}{f_{X_k^0 X_k^1 | Z_{k-1}, H_k^0}(x,y|Z_{k-1} = 0)} \tag{2.13a}$$

and

$$\lambda_k^1 = \frac{f_{X_k^0 X_k^1 | Z_{k-1}, H_k^1}(x,y|Z_{k-1} = 1)}{f_{X_k^0 X_k^1 | Z_{k-1}, H_k^0}(x,y|Z_{k-1} = 1)} \tag{2.13b}$$

The minimization of the conditional probability of error results in the decision rule

$$\lambda_k^i \underset{H_k^0}{\overset{H_k^1}{\gtrless}} 1 \quad \text{if } Z_{k-1} = i \tag{2.14}$$

A block diagram of the suboptimal receiver is shown in Fig. 2.3. The decision on the first bit for both the optimal and suboptimal scheme is made using the decision rule without memory. The likelihood ratio for the scheme without memory is given by

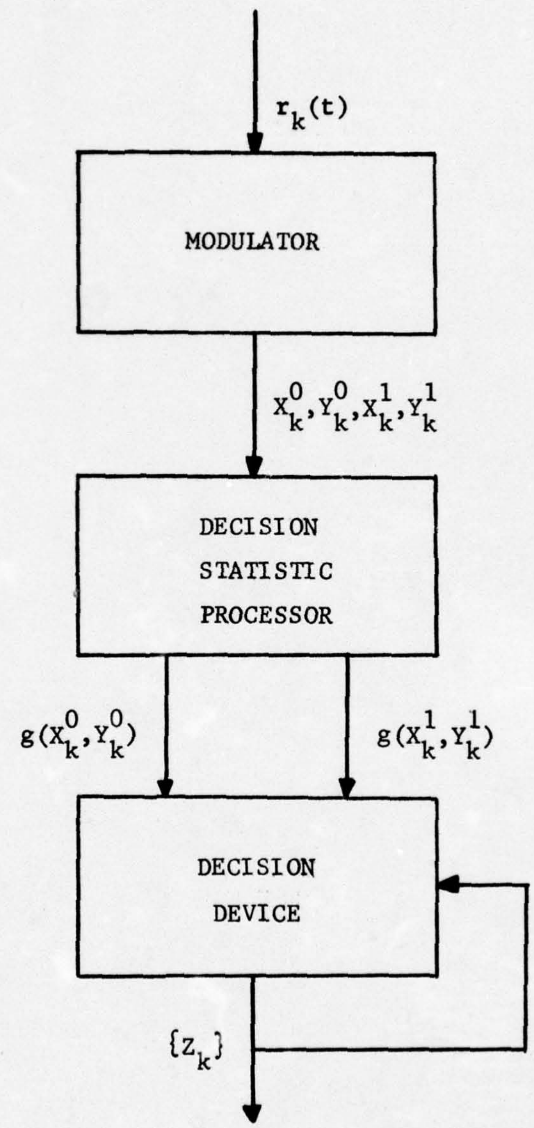

Figure 2.3. Suboptimal receiver structure.

$$\Lambda = \frac{f_{X_k^0 X_k^1 | H_k^1}(x,y)}{f_{X_k^0 X_k^1 | H_k^0}(x,y)} \tag{2.15}$$

and the decision rule is

$$\Lambda \underset{H_k^0}{\overset{H_k^1}{\gtrless}} 1 \tag{2.16}$$

Then, for all succeeding bits the decision rule is modified according to the scheme with memory. Two alternative schemes are considered here. First, the decision on each bit is made based on the decision rule which assumes that the decision rule used for the previous bit was memoryless, i.e., (2.16). Thus, the decision rule (2.14) is fixed and is a function of only the previous decision, i.e. Z_{k-1}. The second possible scheme is recursive in that the decision rule used for the detection of the previous bit is employed to obtain the decision rule for the k-th bit. It can be shown that the recursive algorithm attains a steady state if the sequence $\{v_k\}$ is stationary and the steady state decision rule may be employed for detection purposes and only its derivation is obtained recursively. In the next section, two examples are considered where further details are furnished about both the optimal and the suboptimal schemes. The performance of the detection schemes is also examined in terms of the probability of error.

2.5. Examples

In this section, two examples are considered to illustrate the detection algorithms with one-bit memory. Binary on-off keying systems are chosen as examples since the decision rules could be expressed as simple threshold tests, which are easy to visualize and represent. The two signals corresponding to the transmission of a zero and a one are given by

$$s_0(t) = 0$$
$$s_1(t) = s(t) = f(t)\cos(\omega_c t + \phi(t)) \tag{2.17}$$

where f(t) has been normalized as in (2.2). The first example considers the case of Rayleigh fading and the second treats Gaussian fading.

2.5.1. Rayleigh Fading Example

The communication system considered in this example is assumed to be operating over a channel where both the amplitude and the phase of the received signal vary. The envelope of the received signal is assumed to have Rayleigh density and the phase is assumed to have uniform density. This model is frequently used for ionospheric and tropospheric links. Using the signal model of (2.1) the received signal during the k-th signalling interval, represented by $r_k(t)$, may be written under the two hypotheses as

$$H_k^0:\ r_k(t) = n_k(t) \qquad , \quad 0 \le t < T \tag{2.18a}$$

$$H_k^1:\ r_k(t) = v_k f(t)\cos[\omega_c t + \phi(t) + \theta_k] + n_k(t), \quad 0 \le t < T \tag{2.18b}$$

The signal component could be expressed in terms of its quadrature components and $r_k(t)$ is given by

$$H_k^0:\ r_k(t) = n_k(t) \qquad , \quad 0 \le t < T \tag{2.19a}$$

$$H_k^1:\ r_k(t) = a_{1k} f(t)\cos[\omega_c t + \phi(t)] + a_{2k} f(t)\sin[\omega_c t + \phi(t)] + n_k(t)$$
$$\triangleq a_{1k} s_1(t) + a_{2k} s_2(t) + n_k(t), \quad 0 \le t < T \tag{2.19b}$$

where a_{1k} and a_{2k} are independent zero-mean Gaussian random variables with variance σ_a^2 (where $E[v_k^2] = 2\sigma_a^2$). Also, the two terms $s_1(t)$ and $s_2(t)$ are orthogonal. The sequences $\{a_{1k}\}$ and $\{a_{2k}\}$ are assumed to be correlated, otherwise a memoryless scheme results. The correlation coefficient of both sequences is denoted by r, namely

$$E\{a_{ik}a_{ik-1}\} = r\,\sigma_a^2 \quad , \quad i = 1,2 \tag{2.20}$$

The structure of the optimum receiver is shown in Fig. 2.4. The decision rule is of the form

$$\ell_k = (X_k)^2 + (Y_k)^2 \underset{H_k^0}{\overset{H_k^1}{\gtrless}} T_k \tag{2.21}$$

where X_k and Y_k correspond to X_k^1 and Y_k^1 defined in (2.5) and T_k is the optimum threshold. The objective is to devise a decision rule which utilizes the correlation of (X_k,Y_k) with the pair (X_{k-1},Y_{k-1}) to yield a better error performance. This is accomplished by computing the threshold T_k *at each step which minimizes* the probability of error in the k-th bit and the threshold T_k is a function of the signal-to-noise ratio, the correlation between (X_k,Y_k) and (X_{k-1},Y_{k-1}) and the statistic of the previous bit ℓ_{k-1}. The signal-to-noise ratio η and the correlation coefficient ρ between the pairs of random variables (X_k,X_{k-1}) and (Y_k,Y_{k-1}) are defined as:

$$\eta = \sigma_a^2(N_0)^{-1} \tag{2.22}$$

$$\rho = (\eta\, r)(\eta+1)^{-1} \tag{2.23}$$

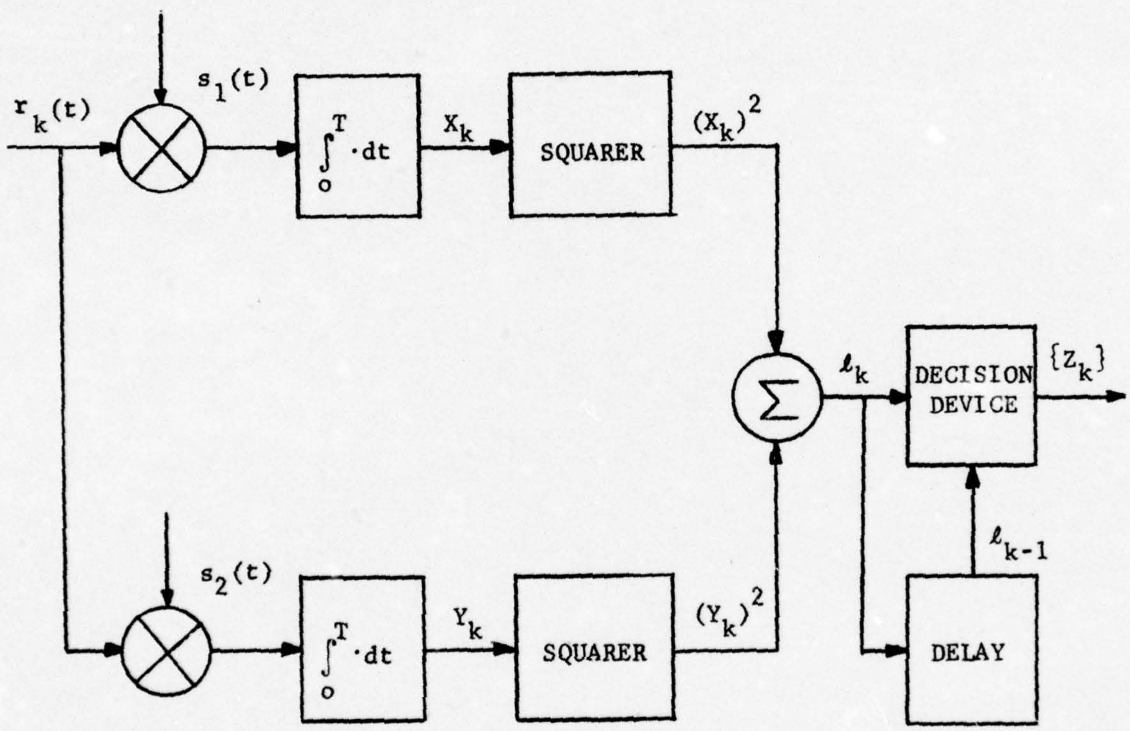

Figure 2.4. Optimum receiver structure for Rayleigh fading example.

where r has been defined in (2.20). It must be noted here that X_k and Y_k are independent of each other due to the independence of a_{1k} and a_{2k}. Hence the conditional joint densities of the pairs (X_k, X_{k-1}) and (Y_k, Y_{k-1}) are also Gaussian because of the previous assumptions. The optimal and the suboptimal schemes are now described. Only the results are presented here, the detailed derivations are given in the Appendix A.

2.5.1.1. Optimal Scheme

The optimal scheme involves the computation of the optimum decision threshold. The decision statistic ℓ_k is compared to the threshold T_k and a decision on the k-th bit is reached. The threshold for the detection of the first bit, T_1, is obtained by the minimization of the probability of error and is given by

$$T_1 = 2(1+\eta)\ell n(1+\eta)\eta^{-1} \quad . \tag{2.24}$$

Thresholds for the detection of succeeding bits are computed so as to minimize the conditional probability of error, i.e., T_k is given by

$$\partial P(e_k = 1 | X_{k-1}, Y_{k-1})/\partial T_k = 0 \tag{2.25}$$

For this example, the equation which determines T_k is:

$$\exp(-T_k/2) + \{2(1+\eta)(1-\rho^2)\}^{-1}\exp\{-(T_k+\rho^2\ell_{k-1})/2(1+\eta)(1-\rho^2)\}$$

$$I_0\{\rho\, T_k^{\frac{1}{2}} \ell_{k-1}^{\frac{1}{2}}/(1+\eta)(1-\rho^2)\} - \{2(1+\eta)\}^{-1}\exp\{-T_k/2(1+\eta)\} = 0 \tag{2.26}$$

where $I_0(\cdot)$ is a modified Bessel function of the first kind and $\ell_{k-1} = X_{k-1}^2 + Y_{k-1}^2$. Thus, T_k is a function of η, ρ and ℓ_{k-1} and is computed in a recursive fashion from (2.26). The performance of the optimal scheme could be evaluated by computing the probability of error which is given as

$$P(e_k) = \int_{w=0}^{\infty} \frac{1}{2} P(e_k = 1|\ell_{k-1}) f_{\ell_{k-1}}(w|H^0_{k-1})dw$$

$$+ \int_{w=0}^{\infty} \frac{1}{2} P(e_k = 1|\ell_{k-1}) f_{\ell_{k-1}}(w|H^1_{k-1})dw \qquad (2.27)$$

where

$$P(e_k = 1|\ell_{k-1}) = \frac{1}{2} \exp(-T_k/2) + \frac{1}{4}\{1-\exp(T_k/2(1+\eta))\}$$

$$+ \frac{1}{4}\left[1 - Q\{\rho\, \ell^{\frac{1}{2}}_{k-1} (1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}} \,;\, T_k^{\frac{1}{2}}(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\}\right]$$

$$f_{\ell_{k-1}}(w|H_0) = \frac{1}{2} \exp(-w/2)$$

$$f_{\ell_{k-1}}(w|H_1) = \frac{1}{2} (1+\eta)^{-1} \exp(-w/2(1+\eta))$$

$Q(\cdot,\cdot)$ is the Marcum's Q function [78,79] which, for completeness, is described in the Appendix B. The conditional density of ℓ_{k-1} is obtained using the fact that X_{k-1} and Y_{k-1} are independent Gaussians and $\ell_{k-1} = X^2_{k-1} + Y^2_{k-1}$ [44]. It is observed that the decision threshold obtained from (2.5) is different from T_1 and it is expected that the optimal scheme will perform better. However, the optimal scheme is quite complex to implement and, therefore, it is not explored any further and the suboptimal scheme is considered next.

2.5.1.2. Suboptimal Scheme

The suboptimal scheme is a decision-feedback scheme where two different decision thresholds are used depending upon the previous decision. Threshold R^0_k is used if the previous decision was a zero and R^1_k is employed otherwise. T_1, as determined in (2.24) is utilized for the detection of the first bit. Thresholds T_k which minimize the conditional probability of error are computed in a recursive fashion from the following equations

$$\partial P(e_k = 1 | Z_{k-1} = 0)/\partial R_k^0 = 0 \tag{2.28a}$$

$$\partial P(e_k = 1 | Z_{k-1} = 1)/\partial R_k^1 = 0 \tag{2.28b}$$

Thresholds R_k^0 and R_k^1 for the detection of the second bit are computed from (2.28), for k = 2. The resulting set of equations which yield the values of R_2^0 and R_2^1 are

$$\frac{1}{2}(1+\eta)^{-1}\exp\{-R_2^0/2(1+\eta)\}[2-\exp(-T_1/2) - Q\{\rho(R_2^0)^{\frac{1}{2}}(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\ ;$$
$$T_1^{\frac{1}{2}}(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}] - \frac{1}{2}\exp\{-R_2^0/2\}[2-\exp(-T_1/2)-\exp(-T_1/2(1+\eta))] = 0 \tag{2.29a}$$

$$\frac{1}{2}(1+\eta)^{-1}\exp\{-R_2^1/2(1+\eta)\}[\exp(-T_1/2)+Q\{\rho(R_2^1)^{\frac{1}{2}}(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}};\ T_1^{\frac{1}{2}}(1+\eta)^{-\frac{1}{2}}$$
$$(1-\rho^2)^{-\frac{1}{2}}\}] - \frac{1}{2}\exp(-R_2^1/2)[\exp(-T_1/2)+\exp(-T_1/2(1+\eta))] = 0 \tag{2.29b}$$

which is a set of nonlinear implicit equations which must be solved for R_2^i. The process may be continued recursively to obtain R_k^i as a function of R_{k-1}^i. It can be shown that a steady state for the thresholds is reached and the resulting values of the thresholds can be computed. Here only a special case is considered where for the detection of the previous bit it is assumed that the memoryless threshold T_1 is used. The value of the threshold T_k is, therefore, expressed as

$$T_k(Z_{k-1}) = \begin{cases} R_2^0, & \text{if } Z_{k-1} = 0 \\ R_2^1, & \text{if } Z_{k-1} = 1 \end{cases} \tag{2.30}$$

The performance of the suboptimal scheme can be evaluated by comparing the probability of error using the standard memoryless scheme and the suboptimal scheme described above. The probability of error using the standard scheme $P(e)_{nm}$ is given by

$$P(e)_{nm} = \frac{1}{2}\exp(-T_1/2) + \frac{1}{2}\left[1 - \exp(-T_1/2(1+\eta))\right] \tag{2.31}$$

and the probability of error in the second bit using the suboptimal scheme $P(e)_m$ is given by:

$$\begin{aligned} P(e)_m = \frac{1}{4}\Big[&2 + \exp(-R_2^0/2)\{2-\exp(-T_1/2)-\exp(-T_1/2(1+\eta))\} \\ &- \exp(-R_2^0/2(1+\eta))(2-\exp(-T_1/2))-\exp(-T_1/2)\exp(-R_2^1/2(1+\eta)) \\ &+ \exp(-R_2^1/2)\{\exp(-T_1/2)+\exp(-T_1/2(1+\eta))\} \\ &+ \int_{(R_2^0)^{\frac{1}{2}}}^{(R_2^1)^{\frac{1}{2}}} \frac{w}{(1+\eta)}\exp(-w^2/2(1+\eta))Q\{\rho w(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\,; \\ &T_1^{\frac{1}{2}}(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\}dw\Big] \end{aligned} \tag{2.32}$$

It can be observed that the probabilities of error using the scheme without memory and the one with memory are equal if $R_2^0 = R_2^1 = T_1$. Numerical results were obtained for the suboptimal scheme. The numerical values of thresholds and the probabilities of error obtained for different values of signal-to-noise ratio are summarized in Table 2.1. The correlation coefficient r for these computations was assumed to be one.

Table 2.1

η	T_1	R_0	R_1	$P(e)_{nm}$	$P(e)_m$
10	5.27537	4.87149	5.80982	.14236	.14146
10^2	9.32254	8.66881	10.17792	.02728	.02694
10^3	13.83133	13.10032	14.80521	.00394	.00389
10^4	18.42272	13.67107	19.44922	.00051	.00050

2.5.2. Gaussian Fading Case

In this example, it is assumed that the communication system is operating over an amplitude fading channel with coherent reception. The received signal during the k-th signalling interval, denoted by $r_k(t)$, may be represented by

$$H_k^0: \quad r_k(t) = n_k(t) \quad , \qquad 0 \leq t < T \tag{2.33a}$$

$$H_k^1: \quad r_k(t) = v_k f(t) \cos[\omega_c t + \phi(t)] + n_k(t), \qquad 0 \leq t < T \tag{2.33b}$$

where f(t) is normalized as in (2.2). The sequence of random variables $\{v_k\}$ is assumed to be a stationary Gaussian sequence with zero mean and variance σ^2. The correlation coefficient of v_{k-1} and v_k is denoted by r. ρ and η are as defined in (2.22) and (2.23). The basic structure of the proposed optimum receiver is shown in Fig. 2.5. The decision rule is of the form

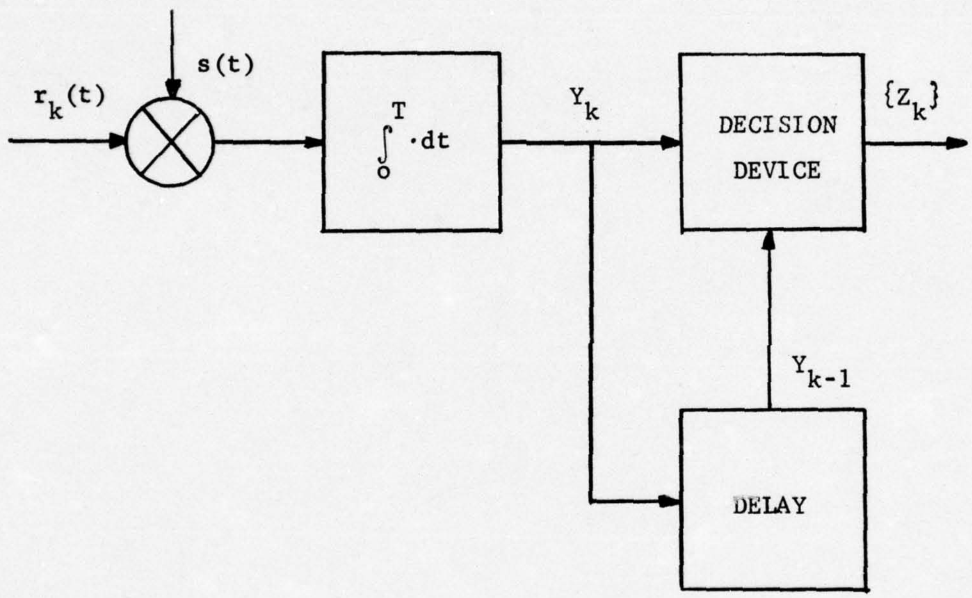

Figure 2.5. Optimum receiver structure for the Gaussian fading case.

$$Y_k \underset{H_k^0}{\overset{H_k^1}{\gtrless}} T_k \tag{2.34}$$

The objective is to compute the threshold T_k which minimizes the conditional probability of error. The threshold is a function of the signal-to-noise ratio and the correlation coefficient. The optimal and the suboptimal schemes are considered next. Again, the detailed derivations of the equations are referred to Appendix C.

2.5.2.1. Optimal Scheme

Detection scheme without memory is used for the detection of the first bit. The threshold T_1 is computed so as to minimize the probability of error and is given by

$$T_1 = [(1+\eta)\ell n(1+\eta)]^{\frac{1}{2}} \eta^{-\frac{1}{2}} \tag{2.35}$$

Thresholds for the detection of succeeding bits, T_k, are given by

$$(1+\eta)^{-\frac{1}{2}}\exp(-T_k^2/2(1+\eta)) + \{(1+\eta)(1-\rho^2)\}^{-\frac{1}{2}}$$

$$\exp\{-(T_k-\rho y_{k-1})^2/2\,(1+\eta)(1-\rho^2)\} - 2\exp(-T_k^2/2) = 0 \tag{2.36}$$

Again it is observed that the threshold depends on the statistic of the previous bit and it is quite complex to implement in real time. Consequently, the suboptimal scheme is considered next.

2.5.2.2. Suboptimal Scheme

As discussed earlier, two different decision thresholds are employed for the detection of the k-th bit. The decision threshold T_k, which is a function of the previous decision, is given by:

$$T_k(Z_{k-1}) = \begin{cases} Q_k & \text{if } Z_{k-1} = 0 \\ R_k & \text{otherwise} \end{cases} \tag{2.37}$$

The equations which yield the values of Q_2 and R_2 are obtained by minimizing the conditional probability of error $P(e_2 = 1|Z_{k-1} = i)$ and are given by

$$a/(a+b) + [\mathrm{erf}\{(T_1 - \rho Q_2)\alpha\} + \mathrm{erf}\{(T_1 + \rho Q_2)\alpha\}]/2(a+b)$$
$$= (1+\eta)^{\frac{1}{2}} \exp\{-\eta Q_2^2/2(1+\eta)\} \tag{2.38}$$

and

$$c/(c+d) + [\mathrm{erfc}\{(T_1 - \rho R_2)\alpha\} + \mathrm{erfc}\{(T_1 + \rho R_2)\alpha\}]/2(c+d)$$
$$= (1+\eta)^{\frac{1}{2}} \exp\{-\eta R_2^2/2(1+\eta)\} \tag{2.39}$$

where

$$\mathrm{erf}(x) = (2\pi)^{-\frac{1}{2}} \int_0^x \exp\{-t^2/2\}dt$$

$$a = \mathrm{erf}(T_1)$$

$$b = \mathrm{erf}(T_1(1+\eta)^{-\frac{1}{2}})$$

$$c = 1. - a$$

$$d = 1. - b$$

$$\alpha = (1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}$$

The thresholds are computed in a recursive fashion. The equations which yield the thresholds Q_k and R_k are similar to (2.38) and (2.39) and are given by

$$g/(g+h) + [(\mathrm{erf}\{\alpha(Q_{k-1} - \rho Q_k)\} + \mathrm{erf}\{\alpha(Q_{k-1} + \rho Q_k)\})P(Q_{k-1})$$

$$+ (\mathrm{erf}\{\alpha(R_{k-1} - \rho Q_k)\} + \mathrm{erf}\{\alpha(R_{k-1} + \rho Q_k)\})P(R_{k-1})]/2(g+h)$$

$$= (1+\eta)^{\frac{1}{2}} \exp\{-\eta Q_k^2/2(1+\eta)\} \tag{2.40}$$

and

$$m/(m+n) + [(\mathrm{erfc}\{\alpha(Q_{k-1} - \rho R_k)\} + \mathrm{erfc}\{\alpha(Q_{k-1} + \rho R_k)\})P(Q_{k-1})$$

$$+ (\mathrm{erfc}\{\alpha(R_{k-1} - \rho R_k)\} + \mathrm{erfc}\{\alpha(R_{k-1} + \rho R_k)\})P(R_{k-1})]/2(m+n)$$

$$= (1+\eta)^{\frac{1}{2}} \exp\{-\eta R_k^2/2(1+\eta)\} \tag{2.41}$$

where

$$g = \mathrm{erf}(Q_{k-1})P(Q_{k-1}) + \mathrm{erf}(R_{k-1})P(R_{k-1})$$

$$h = \mathrm{erf}(Q_{k-1}(1+\eta)^{-\frac{1}{2}})P(Q_{k-1}) + \mathrm{erf}(R_{k-1}(1+\eta)^{-\frac{1}{2}})P(R_{k-1})$$

$$m = \mathrm{erfc}(Q_{k-1})P(Q_{k-1}) + \mathrm{erfc}(R_{k-1})P(R_{k-1})$$

$$n = \mathrm{erfc}(Q_{k-1}(1+\eta)^{-\frac{1}{2}})P(Q_{k-1}) + \mathrm{erfc}(R_{k-1}(1+\eta)^{-\frac{1}{2}})P(R_{k-1})$$

It can be shown that a steady-state is eventually reached and the steady-state solution may be obtained by setting $Q_{k-1} = Q_k = Q$ and $R_{k-1} = R_k = R$ in (2.40) and (2.41). The resulting equations cannot be solved analytically and, therefore, approximate solutions for Q and R are obtained by linearization. Let ΔQ and ΔR represent the deviations of Q and R from the no-memory threshold $T = T_1$, i.e.

$$Q = T + \Delta Q \tag{2.42a}$$

$$R = T + \Delta R \tag{2.42b}$$

The approximate solution for the deviations ΔQ and ΔR is obtained from the linearized equations as a set of simultaneous linear equations:

$$a_{11}\Delta Q + a_{12}\Delta R = b_1 \tag{2.43a}$$

$$a_{21}\Delta Q + a_{22}\Delta R = b_2 \tag{2.43b}$$

where

$$a_{11} = u(a+b)(\alpha_2\beta + \alpha_3/\beta) - \alpha_1(a+b)(2\pi)^{-\frac{1}{2}} + v\rho(c+d)(\alpha_3 - \alpha_2)$$
$$+ \eta T(a+b)(1+\eta)^{-1}$$

$$a_{12} = v(c+d)(\alpha_2 + \alpha_3) - \alpha_1(c+d)(2\pi)^{-\frac{1}{2}}$$

$$a_{21} = \alpha_1(a+b)(2\pi)^{-\frac{1}{2}} - v(a+b)(\alpha_2 + \alpha_3)$$

$$a_{22} = v\rho(a+b)(\alpha_2 - \alpha_3) + \alpha_1(c+d)(2\pi)^{-\frac{1}{2}} - u(c+d)(\alpha_2\beta + \alpha_3/\beta)$$
$$+ \eta T(a+b)(1+\eta)^{-1}$$

$$b_1 = b - \frac{1}{2}(\mathrm{erf}(T\beta(1+\eta)^{-\frac{1}{2}}) + \mathrm{erf}(T\beta^{-1}(1+\eta)^{-\frac{1}{2}}))$$

$$b_2 = d - \frac{1}{2}(\mathrm{erfc}(T\beta(1+\eta)^{-\frac{1}{2}}) + \mathrm{erfc}(T\beta^{-1}(1+\eta)^{-\frac{1}{2}}))$$

$$\beta = (1-\rho)^{\frac{1}{2}}(1+\rho)^{-\frac{1}{2}}, \quad u = \frac{1}{2}(2\pi(1+\eta))^{-\frac{1}{2}}, \quad v = u(1-\rho^2)^{-\frac{1}{2}},$$

$$\alpha_1 = \exp(-T^2/2), \quad \alpha_2 = \exp(-T^2\beta^2/2(1+\eta)), \quad \alpha_3 = \exp(-T^2/2\beta^2(1+\eta)).$$

In (2.43), $a_{11}, a_{22} > 0$, $a_{12} > 0$, $a_{21} < 0$, $b_1 < 0$ and $b_2 > 0$ so that $\Delta Q < 0$ and $\Delta R > 0$, unless $r = 0$, for which case the receiver reduces to the zero memory receiver.

The performance of the receiver is measured in terms of the probability of error. The bit-error rate using both schemes is given by:

$$P(e)_{nm} = b + c \tag{2.44}$$

$$\begin{aligned} P(e)_m = {} & (a+b)\mathrm{erfc}(Q) + (c+d)\mathrm{erfc}(R) + c\ \mathrm{erf}(R(1+\eta)^{-\frac{1}{2}}) \\ & + (a+\tfrac{1}{2})\mathrm{erf}(Q(1+\eta)^{-\frac{1}{2}}) \\ & + \frac{1}{2}\int_{u=Q}^{R} (2\pi(1+\eta))^{-\frac{1}{2}}\exp(-u^2/2)[\mathrm{erfc}\{\alpha(T-\rho u)\} \\ & + \mathrm{erfc}\{\alpha(T+\rho u)\}]du \end{aligned} \tag{2.45}$$

The relationship between the probabilities of error using the decision scheme with and without memory is stated in the following theorem.

Theorem 2.1: The probability of error using the optimal detector without memory, $P(e)_{nm}$, is greater than or equal to the probability of error $P(e)_m$ using the proposed suboptimal decision scheme with one-bit memory, i.e.,

$$P(e)_{nm} \geq P(e)_m$$

The equality is achieved if and only if the correlation coefficient r is zero.

An illustration of the theorem for the Gaussian fading case is shown in Appendix D. Numerical results were obtained for different values of the signal-to-noise ratio η. The correlation coefficient r was assumed to be one. The numerical values of the thresholds and the probabilities of error are presented in Table 2.2.

Table 2.2

η	T	Q	R	$P(e)_{nm}$	$P(e)_m$
10	1.6241	1.5307	1.7403	.2399	.2391
10^2	2.1590	2.0058	2.3152	.1004	.0994
10^3	2.6298	2.4576	2.7857	.0374	.0369
10^4	3.0350	2.8876	3.1802	.0133	.0131

2.6. Discussion

In this chapter, the idea of introducing memory into the receiver with applications to channels with memory was explored. In particular, receivers with one-bit memory for fading channels were considered. An optimal detection scheme and a suboptimal decision-feedback scheme were devised. Their performance was measured in terms of probability of error. Two examples were considered. Improvement in the performance was achieved indicating the theoretical feasibility of the idea. It is expected that introducing larger memory would result in more improvement. The emphasis in this chapter was to examine the feasibility of the idea and receivers with larger memory are discussed in the next chapter.

3. RECEIVER WITH LARGE MEMORY

3.1. Introduction

The results obtained in the last chapter indicated the feasibility of receivers with memory. A suboptimal decision-feedback receiver with one-bit memory was shown to perform better than a receiver without memory. The amount of improvement obtained leads us to the consideration of an alternative approach to the receiver design problem which is the subject of this chapter. The receiver is assumed to consist of an estimator and a detector. The estimator is employed to estimate the existing channel conditions and an estimate of the uncertain parameters is furnished to the detector. The detector implements an adaptive decision rule based on the information provided by the estimator about the channel conditions. The memory is incorporated into the estimator and the estimate is based on the observations from a finite past. The memory length of the receiver is denoted by M and this means that the estimator utilizes the signals received during the previous M signalling intervals to yield an estimate of the uncertain parameters. In this chapter, we concentrate on the adaptive detector and the discussion on the estimator design is postponed to the next chapter. It is assumed in this chapter that the estimator yields a known functional of the past M observations as estimates of the uncertain parameters. The estimation criterion and the selection of the estimation functional are discussed in chapter four.

As in the previous chapter, absence of intersymbol interference is assumed and fading is treated as the only source of channel memory.

A large memory length, M, is required to estimate the fading parameters of the channel due to the assumed slowly varying nature of the fading process relative to the signalling rate. It is expected that such a receiver with a large memory would perform significantly better than the zero-memory receiver. It was mentioned earlier that only the adaptive detector is discussed at length in this chapter. In section two, the problem is stated along with the necessary assumptions. The receiver structure is described in the third section. The performance evaluation is considered in the fourth section. Finally, the receiver operation is illustrated by considering a specific example and numerical results are obtained.

3.2. Problem Statement

The goal here is to develop a receiver with an M-bit memory for a binary communication system operating over a fading channel in the absence of intersymbol interference. The transmitted bits are again assumed to be independent of each other with zeroes and ones equiprobable. The received signal in any signalling period [0,T) under the two hypotheses is given by:

$$H_0: r(t) = s_0(t,\underline{\theta}) + n(t) \qquad 0 \leq t < T \qquad (3.1a)$$

$$H_1: r(t) = s_1(t,\underline{\theta}) + n(t) \qquad 0 \leq t < T \qquad (3.1b)$$

where $\underline{\theta}$ denotes the uncertain parameters due to fading i.e. both amplitude and phase fadings may be included in $\underline{\theta}$. Fading is again assumed to be slow so that is can be assumed to be constant during a signalling interval and the effect of fading may be represented as a sequence of dependent random variables. The signal $s_i(t,\underline{\theta})$ is conditionally deterministic, i.e., it is completely known if $\underline{\theta}$ is known. Since the absence of intersymbol interference has been assumed, the signal model is the same as that given in (2.1) and (2.2). The additive noise n(t) is again assumed to be white Gaussian with power spectrum N_0. The contributions of fading, i.e., amplitude and phase fading parameters are denoted, as in Chapter 2, by $\{v_k,\theta_k\}$. Hence the received signal during the k-th signalling interval $[(k-1)T,kT))$ under hypothesis H_k^i is then represented as

$$r_k(t) \triangleq r[t + (k-1)T]$$

$$= v_k\sqrt{2}\ f_i(t)\ \cos\{\omega_i t + \phi_i(t) + \theta_k\} + n[t + (k-1)T]$$

$$= v_k\ s_i(t,\theta_k) + n_k(t) \qquad 0 \le t < T,\ i = 0,1 \qquad (3.2)$$

where H_k^i has been defined in Chapter 2. Orthogonal signalling is assumed for simplicity. A correlation receiver as shown in Fig. 2.1 is used to demodulate the incoming signal. The decision statistic is obtained in terms of X_k^i and Y_k^i which are defined by (2.5), and are functions of the dependent random variable pairs (v_k, θ_k) corresponding to the fading process. The statistical properties of the sequence $\{v_k,\theta_k\}$ are employed to derive the receiver with memory. The relationship between (X_k^i, Y_k^i) and (v_k,θ_k) if H_k^i is true, for i = 0,1, is given by

$$X_k^j = \delta_{ij} v_k \cos\theta_k + n_k \equiv \delta_{ij} a_k + n_k \quad , \quad j = 0,1 \tag{3.3a}$$

$$Y_k^j = -\delta_{ij} v_k \sin\theta_k + n_k \equiv \delta_{ij} b_k + n_k \quad , \quad j = 0,1 \tag{3.3b}$$

The optimal solution to the receiver with an M-bit memory problem can be obtained by treating the problem as a 2^M - hypothesis decision problem. The structure of the possible optimum receiver is shown in Fig. 3.1. The number of hypotheses increase exponentially with M. Therefore, for a large M, the problem becomes too complex and one must resort to suboptimal receivers which are easier to implement. A suboptimal constrained receiver is, therefore, considered in the next section.

3.3. The Constrained Receiver

As indicated in the previous section, the complexity of the optimal solution leads naturally to the investigation of a constrained suboptimal receiver. The details of the receiver are described in this section.

3.3.1. System Structure

The adaptive receiver considered here consists of an estimator and a detector. A limited-memory filter computes an estimate of the fading parameters on the basis of the previous M observations and decisions. The estimate is furnished to the detector which treats the estimate as the correct value of the uncertain parameter in making the decision on the present bit. This idea has previously been explored by Price [45] and Kailath [46]. Jointly optimum combined estimation-detection schemes have recently been considered (e.g. [47-49]). The problem considered

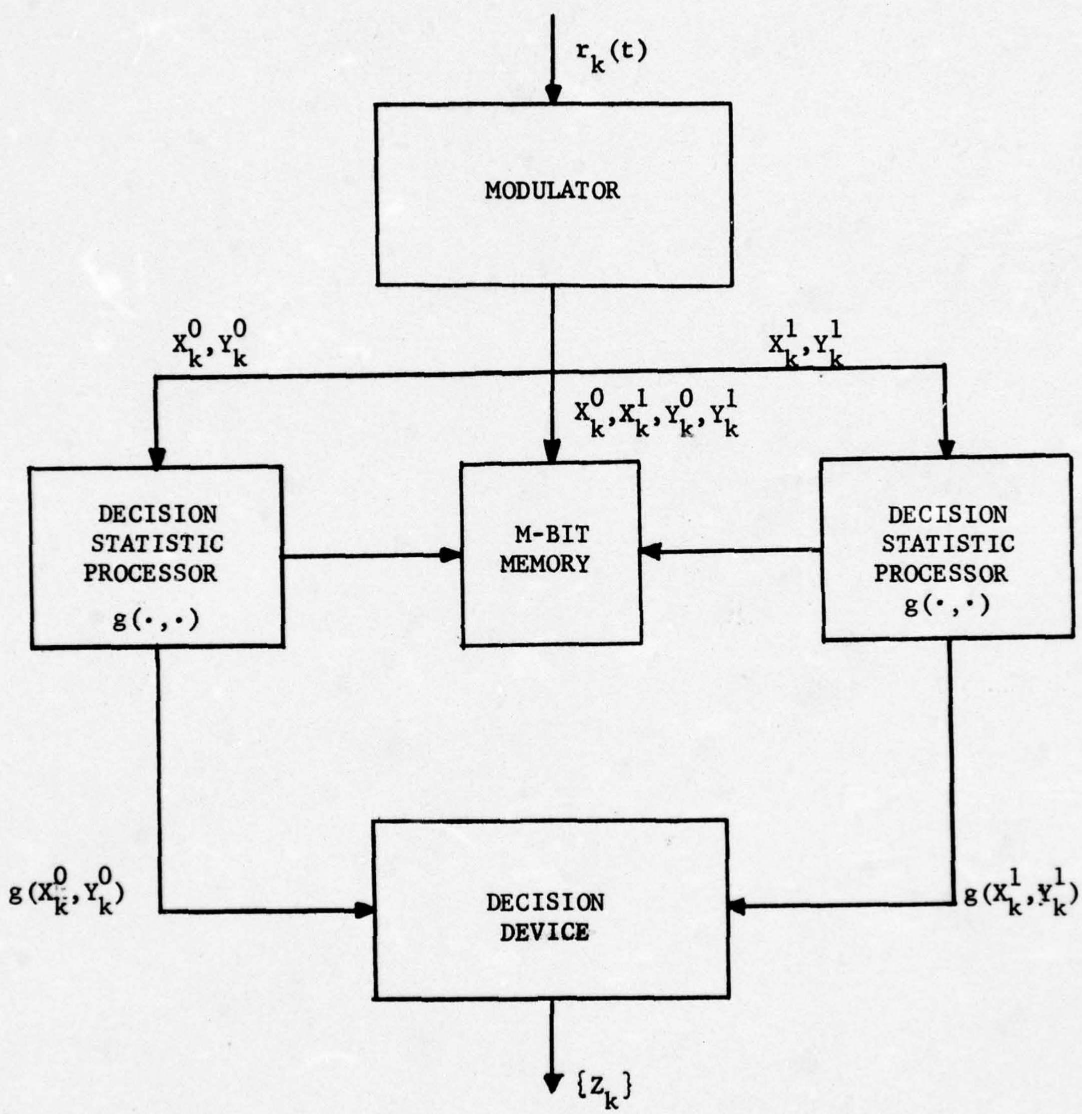

Figure 3.1. Optimum receiver with M-bit memory.

here is related but different because we are only interested in the detection problem and estimation is present only to aid in a better detection. Decision-feedback is employed to make the algorithm implementation simpler. It is assumed that the system error probability is low so that all the past decisions are considered to be correct. This is a standard assumption with the decision-feedback schemes. This assumption, however, causes error propagation and, thus, there is a possibility of runaway. The problem of runaway in decision-directed schemes has been considered by Davisson and Schwartz [43]. A block diagram of the proposed receiver is shown in Fig. 3.2. The detector adapts the decision rule to the existing channel conditions and which obviously depends upon the estimate furnished by the estimator and its statistics. An estimate of the uncertain parameters a_k and b_k is evaluated and furnished to the adaptive detector. In the next section, estimator design is briefly considered, the details being postponed to the next chapter.

3.3.2. Estimator Description

As discussed earlier, the estimator is responsible for the estimation of the uncertain parameters $\{a_k, b_k\}$. The estimator design is based on the classical estimation theory results. The class of linear estimators is selected for the estimator implementation. This implies that the estimates $\hat{a}_k$ and $\hat{b}_k$ of a_k and b_k are linear functionals of the past observations. One major difference, however, is that the memory of the filter is limited to the past M observations where M is the memory length of the receiver. Limiting the memory of the filter is essential

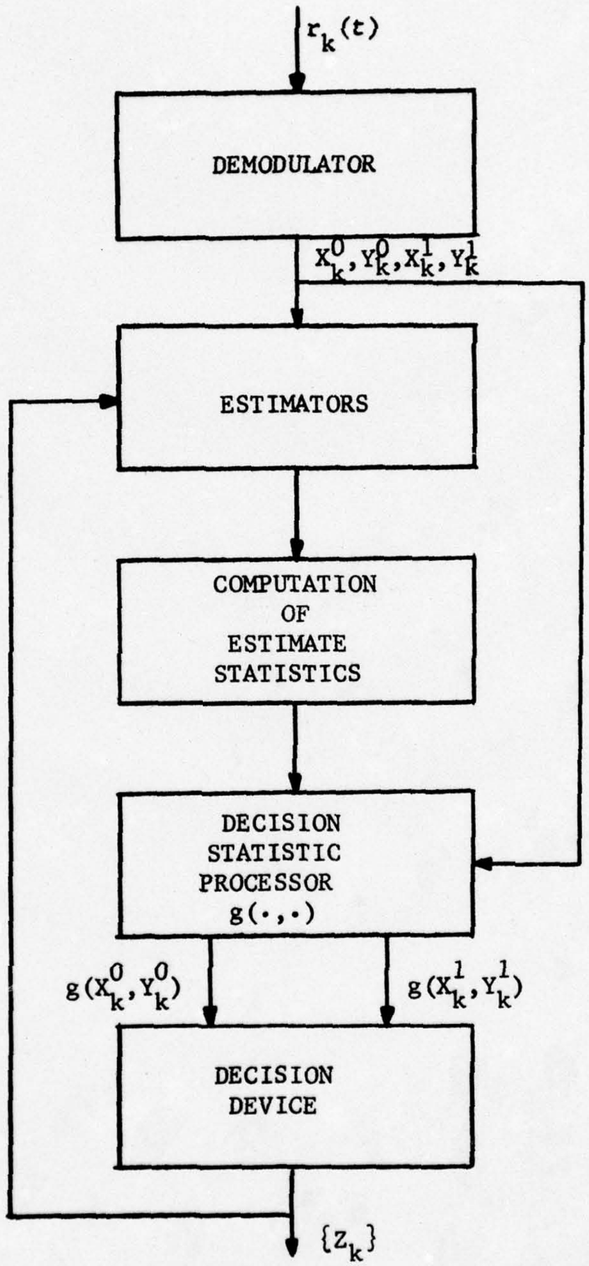

Figure 3.2. Block diagram of the suboptimal constrained receiver.

due to the nature of the fading process. An estimate based on all the past observations may result in a wrong estimate. The class of limited-memory filters have been studied by Jazwinski [50]. The presence of the detector is utilized and decision-feedback is used in the estimator which causes it to be nonlinear. The sequences $\{a_k\}$ and $\{b_k\}$ are assumed to be ergodic Markov sequences with means and autocorrelation functions

$$E\{a_k\} = m_a \tag{3.4a}$$

$$E\{b_k\} = m_b \tag{3.4b}$$

$$E\{a_k a_\ell\} = R_a(k-\ell) \tag{3.4c}$$

$$E\{b_k b_\ell\} = R_b(k-\ell) \tag{3.4d}$$

The limited-memory estimate $\hat{a}_k$ and $\hat{b}_k$ with decision-feedback are assumed to have the following structure.

$$\hat{a}_k = \sum_{i=1}^{M} [Z_{k-i} X^1_{k-i} + (1-Z_{k-i}) X^0_{k-i}]\alpha_i \tag{3.5}$$

and

$$\hat{b}_k = \sum_{i=1}^{M} [Z_{k-i} Y^1_{k-i} + (1-Z_{k-i}) Y^0_{k-i}]\gamma_i \tag{3.6}$$

where $\{Z_k\}$ is the binary output sequence of the detector. The constants $\{\alpha_i, \gamma_i\}$ are computed so as to minimize the mean squared error (MSE) and the detailed derivations of the estimator are presented in the next chapter. The estimate obtained here is a conditional estimate based on the past M decisions $\{Z_{k-1}, \ldots Z_{k-m}\}$ which is denoted by $\mathbf{Z}_k$. The estimate being conditioned on $\mathbf{Z}_k$ is an approximation to the estimate obtained by applying the classical estimation theory results. In some communication systems, the received signal does not always contain the uncertain parameters to be estimated. An example of this class of communication systems is the on-off keying system. The estimator design is modified to be able to use it with the on-off keying system. This point will be discussed in further detail in the next chapter.

3.3.3. Asymptotic Results

The estimates $\hat{a}_k$ and $\hat{b}_k$ as given by (3.5) and (3.6) are linear functions of the past M observations and the nonlinearity is introduced by the decision-feedback. The memory length M is a function of the fading process and its statistics, e.g., average fade duration, fading rate and distribution of fade duration. Since fading is assumed to be a slowly varying process relative to the signalling rate, M is assumed to be large and asymptotic results are obtained. It is noted that the observations are not independent of each other and, therefore, the standard central limit theorem is not applicable. Central limit theorem for dependent random variables [51-53] is used to obtain the asymptotic results. It has already been stated that the sequence $\{a_k, b_k\}$ is assumed to be an ergodic Markov sequence and, therefore, it satisfies the strong mixing condition. The central limit theorem for dependent random variables as stated in [51-53] is then used to conclude that the estimates $\hat{a}_k$ and $\hat{b}_k$ are asymptotically normal. Details of this central limit theorem and the strong mixing condition are furnished in the Appendix E. In communication theory context this theorem has been utilized by Kanefsky and Thomas [54] for nonparametric detection systems. The conditional means and the variances of the Gaussian random variables $\hat{a}_k$ and $\hat{b}_k$ are denoted by $\xi_a(k)$, $\xi_b(k)$ and $\zeta_a(k)$, $\zeta_b(k)$. These means are conditioned on the sequence of previous M decisions, i.e., on $\mathbf{Z}_k$. The conditional means are given by

$$\begin{aligned}\xi_a(k) &= E\{\hat{a}_k|\mathbf{Z}_k\} \\ &= E\{\sum_{i=1}^{M}[Z_{k-i}\,X^1_{k-i} + (1-Z_{k-i})X^0_{k-i}]\alpha_i\} \\ &= \sum_{i=1}^{M}\alpha_i m_a \qquad (3.7)\end{aligned}$$

and $\xi_b(k)$ is given by

$$\xi_b(k) = \sum_{i=1}^{M}\gamma_i m_b \qquad (3.8)$$

The conditional variances are given by

$$\begin{aligned}\zeta_a(k) &= E\{(\hat{a}_k-\xi_a)^2|\mathbf{Z}_k\} \\ &= E\{\hat{a}_k{}^2|\mathbf{Z}_k\} - \xi_a^2(k) \\ &= \sum_{i=1}^{M}\sum_{j=1}^{M}\alpha_i\alpha_j[Z_{k-i}Z_{k-j}R_a(i-j) + N_0\delta_{ij}] - \xi_a^2 \qquad (3.9)\end{aligned}$$

and similarly

$$\zeta_b(k) = \sum_{i=1}^{M}\sum_{j=1}^{M}\gamma_i\gamma_j[Z_{k-i}Z_{k-j}R_b(i-j) + N_0\delta_{ij}] - \xi_b^2 \qquad (3.10)$$

where δ_{ij} denotes the Kronecker delta function. It should again be emphasized that these expressions for the means and variances are approximate because they are conditional on $\mathbf{Z}_k$. The detector optimization is considered next.

3.3.4. Optimization of the Detector

The design criterion for the receiver is the minimization of the probability of error. Therefore, the detector is obtained so as to minimize the conditional probability of error. A Bayesian approach is used to obtain the optimum decision rule. The decision rule is based on the estimate of the uncertain parameters, its statistics and the decision sequence $\mathbf{Z}_k$. The decision rule is, therefore, adaptive. The variables X_k^i and Y_k^i under the two hypotheses are given by (3.3).

For minimization of the probability of error criterion, the cost is the usual Hamming distance. If Ω_k^0 and Ω_k^1 represent the optimum decision regions, the expression for the conditional Bayes risk function can be written as

$$\mathcal{R}_k = \frac{1}{2}\iint_{\Omega_k^0} f_{X_k^1 Y_k^1}(x,y\,|H_k^1,\mathbf{Z}_k)dxdy + \frac{1}{2}\iint_{\Omega_k^1} f_{X_k^0 Y_k^0}(x,y\,|H_k^0,\mathbf{Z}_k)dxdy \tag{3.11}$$

where $f(\cdot,\cdot|H_k^i, \mathbf{Z}_k)$ represents the conditional density under a given hypothesis H_k^i and the past decisions. The conditional density based on the estimates of the uncertain parameters can be used to obtain the density functions described in (3.11), i.e.,

$$f_{X_k^i Y_k^i}(x,y|H_k^i,\mathbf{Z}_k) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_k^i Y_k^i}(x,y|H_k^i,\hat{a}_k,\hat{b}_k)\; f(\hat{a}_k,\hat{b}_k|\mathbf{Z}_k)d\hat{a}_k d\hat{b}_k \tag{3.12}$$

where $f(\hat{a}_k,\hat{b}_k)$ is the joint density of $\hat{a}_k$ and $\hat{b}_k$ and it is Gaussian from the asymptotic normality of $\hat{a}_k$ and $\hat{b}_k$, i.e.,

$$f_{\hat{a}_k \hat{b}_k | \mathbf{Z}_k}(u,v|\mathbf{Z}_k) = (2\pi)^{-1} \zeta_a^{-\frac{1}{2}} \zeta_b^{-\frac{1}{2}} \exp\{-\frac{(u-\xi_a)^2}{2\zeta_a} - \frac{(v-\xi_b)^2}{2\zeta_b}\} \tag{3.13}$$

Since orthogonal signalling has been assumed in the presence of additive white Gaussian noise, the conditional density based on the estimates can be expressed as

$$f_{X_k^i Y_k^i | \hat{a}_k, \hat{b}_k, H_k^i}(x,y|\hat{a}_k,\hat{b}_k,H_k^i) = (2\pi\, N_0)^{-1} \exp\{-\frac{(x-\hat{a}_k)^2 + (y-\hat{b}_k)^2}{2N_0}\} \tag{3.14}$$

The optimum decision rule is obtained by minimizing the conditional risk function of (3.11). It can be expressed as the following likelihood test.

$$\Lambda_k = \frac{f_{X_k^1 Y_k^1}(x,y|H_k^1,\mathbf{Z}_k)}{f_{X_k^0 Y_k^0}(x,y|H_k^0,\mathbf{Z}_k)} \underset{H_k^0}{\overset{H_k^1}{\gtrless}} 1 \tag{3.15}$$

The decision rule is adaptive since the likelihood ratio Λ_k depends upon the estimates and their statistics. Equation (3.15) can be explicitly written as

$$\Lambda_k = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_k^1 Y_k^1 | \hat{a}_k, \hat{b}_k, H_k^1}(x,y|u,v,H_k^1)\, f_{\hat{a}_k \hat{b}_k}(u,v|\mathbf{Z}_k)\,du\,dv}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_k^0 Y_k^0 | \hat{a}_k, \hat{b}_k, H_k^0}(x,y|u,v,H_k^0)\, f_{\hat{a}_k \hat{b}_k}(u,v|\mathbf{Z}_k)\,du\,dv} \underset{H_k^0}{\overset{H_k^1}{\gtrless}} 1 \tag{3.16}$$

where the expressions for the densities are given in (3.13) and (3.14). The performance of the receiver is examined in the next section.

3.4. Performance Evaluation

The performance of the constrained receiver is investigated by evaluating the average probability of error. Stationarity is assumed and, therefore, any arbitrary bit k could be selected and its probability of error can be computed. The adaptive decision rule developed in the previous section, given by (3.16) is conditioned on the decision sequence Z_k. The decision regions determined by the decision rule are denoted by Ω_k^0 and Ω_k^1. These decision regions are also conditioned on the knowledge of the decision on the previous M transmitted digits. Therefore, the probability of error computed for any bit using the adaptive decision rule also depends upon Z_k. This conditional probability of error, in fact, depends upon the estimate furnished by the estimator and its statistics. The conditional probability of error is given by

$$P(e_k|Z_k) = \frac{1}{2}\iint_{\Omega_k^0} f_{X_k^1 Y_k^1}(x,y|H_k^1,Z_k)\ dxdy + \frac{1}{2}\iint_{\Omega_k^1} f_{X_k^0 Y_k^0}(x,y|H_k^0,Z_k)\ dxdy \tag{3.17}$$

This can be written as

$$P(e_k|Z_k) = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big[\iint_{\Omega_k^0} f_{X_k^1 Y_k^1}(x,y|\hat{a}_k,\hat{b}_k,H_k^1)\ dxdy + \iint_{\Omega_k^1} f_{X_k^0 Y_k^0}(x,y|\hat{a}_k,\hat{b}_k,H_k^0)dxdy\Big] f_{\hat{a}_k\hat{b}_k}(u,v|Z_k)dudv \tag{3.18}$$

The density expressions are given in (3.13) and (3.14). The means ξ_a and ξ_b and the variances ζ_a and ζ_b are functions of the decision sequence Z_k as evident from (3.7)-(3.10).

The average probability of error can be computed by first evaluating the conditional probability of error using (3.18) for each of the 2^M possible bit configurations in the previous M digits and then computing an average of all of these equally likely conditional probabilities of error. For a large M this exhaustive method may not be computationally attractive since the number of terms increase exponentially with M. Computation of probability of error in the presence of intersymbol interference and additive Gaussian noise is a similar problem and has received considerable attention recently. Exact calculation is quite tedious and, therefore, bounds on the probability of error have been computed. Lucky [2,3] considered the worst case operation of equalizers and obtained lower bounds on the probability of error. The second class of bounding techniques stem from the use of the Chernoff bound. The most notable bounds of this class have been obtained by Saltzberg [55] and Lugannani [56]. These bounds have been found to be quite loose and other modified bounds have also been considered. Approximations to the probability of error have also been computed by using the series expansions of the Gram-Charlier type. This series expansion technique has been applied to the intersymbol interference problem by Ho and Yeh [57] and Shimbo and Celebiler [58].

In the next section, an example is considered. The average probability of error is computed using the exhaustive method for small M. For more complicated communication systems, and for larger memory lengths, bounding techniques and approximations similar to the ones described above can be used to evaluate the probability of error in the presence of fading and additive Gaussian noise. It should be noted again that for communication systems where observations do not always contain the uncertain parameters to be estimated, the whole analysis needs to be modified. The example considered in the next section illustrates this aspect also.

3.5. Example

In this section a coherent on-off keying system is considered as an example to illustrate the receiver. This example also illustrates the case when all the observations do not contain the uncertain parameters to be estimated and the estimator described earlier is to be modified. The adaptive decision rule can be expressed in terms of an adaptive threshold which is a function of the existing channel conditions. This results in notational convenience and simplicity in presentation of the example. The received signal under the two hypotheses is given by

$$H_k^0: \ r_k(t) = n_k(t) \qquad 0 \leq t < T \qquad (3.19a)$$

$$H_k^1: \ r_k(t) = v_k\sqrt{2}\ f(t)\cos\{\omega t + \phi(t) + \theta_k\} + n_k(t), \quad 0 \leq t < T \qquad (3.19b)$$

In the coherent system, θ_k is assumed to be known and the decision statistic obtained from the demodulator is given by

$$H_k^0 : X_k = n_k \qquad (3.20a)$$

$$H_k^1 : X_k = v_k + n_k \qquad (3.20b)$$

where X_k is, in fact, X_k^1 and the superscript has been dropped for notational convenience. Also $X_k^0 = Y_k^0 = Y_k^1 = 0$ for coherent on-off keying system.

The estimator computes an estimate $\hat{v}_k$ of the uncertain parameter v_k which is furnished to the detector to be employed as the correct value in the decision making process. As indicated previously, the estimator design is modified for this communication system with uncertain observations. The mean and the correlation function of the sequence $\{v_k\}$ are represented by

$$E\{v_k\} = m \tag{3.21}$$

$$E\{v_k v_\ell\} = R(k-\ell) \tag{3.22}$$

The estimator is assumed to have the following structure

$$\hat{v}_k = \sum_{i=1}^{M} Z_{k-i}\alpha_i X_{k-i} \tag{3.23}$$

The estimate $\hat{v}_k$ depends on Z_k as discussed earlier. The set of M equations which yield the constants $\{\alpha_i, \beta_i\}$ are given by

$$\sum_{i=1}^{M} \alpha_i [Z_{k-i} Z_{k-j} R(i-j) + N_0 \delta_{ij}] = Z_{k-j} R(j), \qquad j = 1, \ldots M \tag{3.24}$$

The conditional mean ξ_k and the conditional variance ζ_k are obtained from

$$\xi_k = \sum_{i=1}^{M} Z_{k-i} \alpha_i m \tag{3.25}$$

$$\zeta_k = \sum_{i=1}^{M} \sum_{j=1}^{M} Z_{k-i} Z_{k-j} \alpha_i \alpha_j [R(i-j) + N_0 \delta_{ij}] - \xi_k^2 \tag{3.26}$$

The optimum adaptive decision rule is computed in terms of the threshold which minimizes the conditional probability of error. The expression for the conditional probability of error in this case can be written as

$$P(e_k|Z_k) = \frac{1}{2} \int_{-\infty}^{T_k} [2\pi(N_0 + \zeta_k)]^{-\frac{1}{2}} \exp\{-(x-\xi_k)^2/2(N_0+\zeta_k)\}\, dx$$

$$+ \frac{1}{2} \int_{T_k}^{\infty} [2\pi N_0]^{-\frac{1}{2}} \exp\{-x^2/2N_0\}\, dx \tag{3.27}$$

The threshold T_k is computed by minimizing the conditional probability of error and is computed from

$$\partial P(e_k|Z_k)/\partial T_k = 0 \tag{3.28}$$

which results in

$$T_k = N_0 \xi_k \zeta_k^{-1} [\{1 + \zeta_k N_0^{-1} \xi_k^{-2} (\xi_k^2 + (N_0 + \zeta_k) \ln(1 + \zeta_k N_0^{-1}))\}^{\frac{1}{2}} - 1] \tag{3.29}$$

As expected, the threshold T_k is a function of ξ_k, ζ_k and the signal to noise ratio $1/N_0$. The performance of the receiver is given by the following conditional probability of error

$$P(e_k|Z_k) = \frac{1}{2}[\mathrm{erf}\{(T_k - \xi_k)/(\zeta_k + N_0)^{\frac{1}{2}}\} + \mathrm{erfc}\{T_k/N_0^{\frac{1}{2}}\}] \tag{3.30}$$

where

$$\mathrm{erf}(x) = \int_{-\infty}^{x} (2\pi)^{-\frac{1}{2}} \exp\{-u^2/2\}\, du$$

The conditional probability of error is data dependent and the average probability of error could be computed by the exhaustive method discussed previously. The conditional probability is computed for each of the 2^M possible bit configurations and an average is computed which is the average probability of error.

A numerical example is considered. $\{v_k\}$ is assumed to be a Markov sequence with mean m and autocorrelation

$$R(k-\ell) = \rho^{|k-\ell|} + m^2 \tag{3.31}$$

so that its variance is normalized to unity. The numerical values chosen are $m = .5$, $\rho = .9$ and $M = 8$. The average probability of error as a function of signal to noise ratio (SNR) is computed and is presented in Table 3.1 and Fig. 3.3. The improvement in the performance is significantly better than the improvement obtained from the receiver with one-bit memory. The constrained receiver is quite robust as long as the received sequence of length M contains at least a single transmission of one which is the case for a meaningful communication system. If a sequence of M zeroes is transmitted, the performance of the constrained receiver is worse than the performance of a receiver without memory for lack of any observations containing the uncertain parameters.

Table 3.1

SNR	Average Probability of Error	
	Optimal receiver without memory	Constrained receiver with memory
1	0.4276	0.4255
10	0.2761	0.2753
10^2	0.2396	0.1866
10^3	0.2386	0.1536
10^4	0.2385	0.1420
10^5	0.2385	0.1380

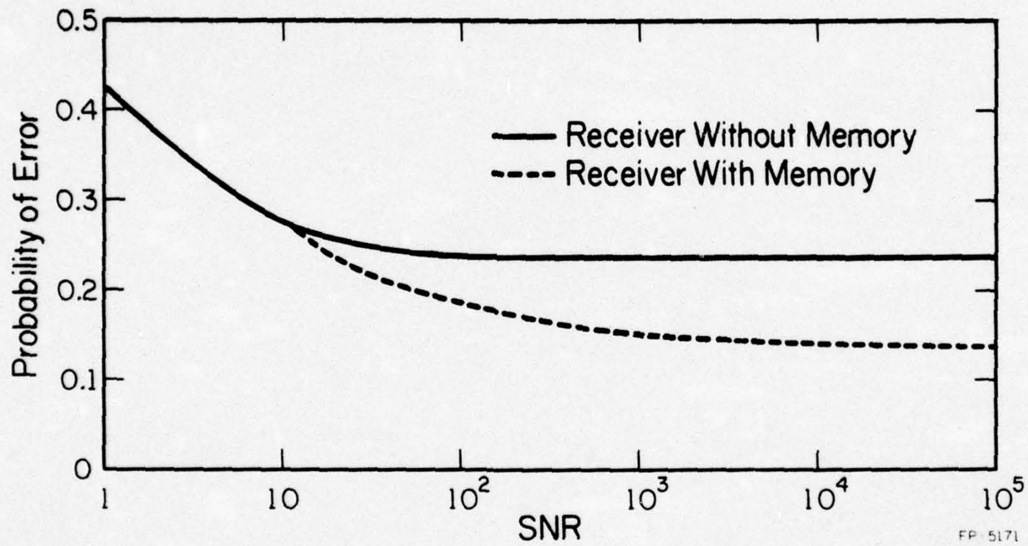

Figure 3.3 Performance of the constrained receiver.

3.6. Discussion

In this chapter an adaptive receiver with large memory is developed. The receiver consists of an estimator and a detector. The estimator furnishes an estimate of the uncertain fading parameters to the detector which treats the estimate as the correct value of the uncertain parameter. The optimum adaptive decision rule is obtained by minimizing the conditional Bayes' risk and the decision rule is based on the previous decisions. The memory length is assumed to be large and asymptotic results are obtained. The performance of the receiver is examined by computing the average probability of error. A methodology is described to compute the bounds and approximations to the probability of error for receivers with large memory lengths. A coherent on-off keying system is considered as an example and the performance of the receiver is examined. The constrained receiver performs significantly better than an optimum receiver without memory. In this chapter, the estimates were assumed to be available and the design of estimators is discussed in the next chapter.

4. ESTIMATION ALGORITHMS FOR RECEIVERS WITH MEMORY

4.1. Introduction

In the previous chapter a receiver with memory was discussed. The receiver consisted of an estimator and a detector. The estimator is responsible for the computation of the estimates of the channel uncertain parameters. The design of estimators is considered in this chapter. Classical estimation theory provides the basic tools for the estimator design. While a recursive estimator (e.g. Kalman filter) is simple to implement, it is based on all the available data (i.e. it has infinite memory) and requires the knowledge of the dynamics of the signal model. If the system dynamics is not known completely or if it is not known accurately, the estimate based on all the past information may not result in a convergent estimate. Jazwinski [50] proposed limited-memory filters to avoid such divergence. The memory length depends upon the time interval over which the model represents a satisfactory approximation to reality. Fading process, which is to be estimated is one such process whose dynamics and statistical parameters are not known accurately. It is, therefore, expected that an estimator based only on the recent past would be more suitable for the estimation of fading parameters. Physically also, the fading passes through periods of severe fading and an estimate based on the observations from the severe fading period but which did not occur in the recent past may result in a wrong estimate. The size of the filter memory required depends upon the parameters of the fading process, e.g., fading rate and average fade duration.

As noted earlier, the idea of limiting the filter memory has been treated in the literature. Jazwinski [50] has devised optimal limited-memory filters for systems where the system dynamics is not completely known. Limited-memory filters are also used in the receivers considered here, which also utilize the presence of the detector in a decision-feedback structure (see (3.5) and (3.6)). The objective of the combined estimator-detector structure is signal detection. Therefore, the estimator should be designed so as to minimize the probability of error in the detection process. In the next section the estimation criterion required to accomplish the basic goal of the receiver is considered. In section 4.3, a limited-memory estimator with decision-feedback using the appropriate estimation criterion is derived.

The estimator discussed in section 4.3 assumes the presence of the parameter to be estimated in all the observations. In many practical situations, however, the observations do not always contain the parameter to be estimated, i.e., the probability that the observations contain the parameter to be estimated is less than one. In digital communications, this problem occurs in the on-off keying systems. When a zero is transmitted, the received signal does not contain any information about the fading parameters. Under these circumstances, the estimator discussed above cannot be used directly and a modified version of the estimator is needed. In section 4.4, a limited-memory filter with decision-feedback and uncertain observations is derived. In section 4.5, several aspects of the resulting estimator are discussed.

4.2. Estimation Criterion

It has been discussed earlier that the goal of the combined estimator-detector structure is signal detection. In the previous chapter a limited-memory estimator with decision-feedback was assumed available. The estimate was assumed to be a linear function of the observations with nonlinearity due to the decision-feedback being introduced in the coefficients. In this section the problem of the desired estimation criterion to achieve optimum detection is considered.

The discussion in this section applies Esposito's results [59] on the subject to the problem under consideration here. For simplicity a binary coherent on-off keying system is assumed. The results obtained similarly apply to more general communication systems as well but at the expense of more complex analysis. For this simple case, the decision statistic X_k^1 in (3.3) reduces to

$$H_k^0: \quad X_k = n_k \tag{4.1a}$$

$$H_k^1: \quad X_k = a_k + n_k \tag{4.1b}$$

since $X_k^0 = 0$, $Y_k^i = 0$, $i = 0,1$ and where the superscript in X_k^1 has been dropped for notational convenience. The optimum decision rule for the minimization of the probability of error criterion involves the computation of a likelihood ratio and then it is compared to a threshold, namely

$$\Lambda_k = \frac{\int_{-\infty}^{\infty} f_1(x_k|a)\, f_{a_k}(a)\, da}{f_0(x_k)} \mathop{\gtrless}_{H_k^0}^{H_k^1} 1 \tag{4.2}$$

where $f_i(x_k|a)$ is the conditional density of X_k under the hypothesis H_k^i and given fading parameter a_k, $f_{a_k}(a)$ is the density function of the fading parameter a_k and where $f_0(\cdot)$ does not depend on a_k for obvious reasons.

In the following, a theorem relating the estimation criterion for the estimator and the optimum performance of the constrained receiver is presented and proved. It uses a basic theorem due to Esposito [59] which is presented first, for completeness.

Theorem 4.1 (Esposito): If the conditional density of v given the independent variable s, i.e., $f(v|s)$ is bounded and continuous for every s and for every value of v, then for each v there exists a value $\hat{s}(v)$ such that

$$\int_{-\infty}^{\infty} f(v|s)\, f_S(s)ds = f(v|\hat{s}) \tag{4.3}$$

Proof: Let $L(v) = \underset{s}{\mathrm{Min}}\, f(v|s)$ and $U(v) = \underset{s}{\mathrm{Max}}\, f(v|s)$ be the lower and upper bounds of $f(v|s)$ for a given v. Since $f_S(s)$ is the density function of S, we have

$$f_S(s) \geq 0$$

and

$$\int_{-\infty}^{\infty} f_S(s)\, ds = 1$$

It, therefore, follows that

$$L(v) \leq \int_{-\infty}^{\infty} f(v|s)\, f_S(s)\, ds \leq U(v) \tag{4.4}$$

By hypothesis $f(v|s)$ is, for each v, a continuous function of s. It follows that for each v there exists a value $\hat{s}(v)$ such that

$$\int_{-\infty}^{\infty} f(v|s)\ f_S(s)\ ds = f(v|\hat{s})$$

Q.E.D.

The main theorem of the section is now considered.

Theorem 4.2: The optimized constrained receiver under the minimization of probability of error criterion is obtained by using a minimum mean-squared error (MMSE) estimator.

Proof: First, an expression for the estimate which minimizes the probability of error is obtained using the result of Theorem 4.1. The estimate is found to have a structure similar to the Kalman filter which is a MMSE estimator. For the coherent on-off keying system, the optimum decision rule is given by (4.2). Since the noise is assumed to be white Gaussian, the conditional densities are given by

$$f_1(x_k|a_k) = (2\pi N_0)^{-\frac{1}{2}} \exp\{-(x_k - a_k)^2/2N_0\} \tag{4.5}$$

$$f_0(x_k) = (2\pi N_o)^{-\frac{1}{2}} \exp\{-x_k^2/2N_0\} \tag{4.6}$$

For simplicity $\{a_k\}$ is assumed to be a Markov sequence so that

$$a_{k+1} = ra_k + w_k \tag{4.7}$$

The model may be generalized to higher-order Markov cases as well. The MMSE estimate of a_k given the past observations is given by

$$\hat{a}_{k|k-1} = r\hat{a}_{k-1} \tag{4.8}$$

with variance

$$V_{k|k-1} = r^2 V_{k-1} + (1-r^2)\ \sigma_a^2 \tag{4.9}$$

Hence, the conditional density function of the uncertain parameter given the past observations is given by

$$f_{a_k|Z_k}(a) = (2\pi V_{k|k-1})^{-\frac{1}{2}} \exp\{-(a - \hat{a}_{k|k-1})^2/2V_{k|k-1}\} \tag{4.10}$$

The optimum likelihood ratio for minimizing the conditional error probability of (4.2) is given by

$$\Lambda_k = \int_{-\infty}^{\infty} \exp\{-[(x_k - a)^2/2N_0] + [x_k^2/2N_0]\}$$
$$(2\pi V_{k|k-1})^{-\frac{1}{2}} \exp\{-(a - \hat{a}_{k|k-1})^2/2V_{k|k-1}\}\, da \tag{4.11}$$

After algebraic manipulations, integration and simplification, (4.11) results in

$$\Lambda_k = \{N_0/(N_0 + V_{k|k-1})\}^{\frac{1}{2}} \exp\{[x_k^2 V_{k|k-1}^2 + 2x_k \hat{a}_{k|k-1} V_{k|k-1} N_0$$
$$-\hat{a}_{k|k-1}^2 N_0 V_{k|k-1}]/[2N_0 V_{k|k-1}(N_0 + V_{k|k-1})]\} \tag{4.12}$$

Theorem 4.1 is now used to find an estimate which yields the minimum probability of error. The conditional density is computed using the expression for conditional density on the right hand side of (4.3)

$$\Lambda_k = f_1(x_k/\hat{a}_k)/f_0(x_k)$$
$$= \exp\{[x_k^2/2N_0] - [(x_k - \hat{a}_k)^2/2N_0]\}$$
$$= \exp\{\frac{2\hat{a}_k x_k - \hat{a}_k^2}{2N_0}\} \tag{4.13}$$

The exponents in the expressions (4.12) and (4.13) are compared and the resulting equation is solved for $\hat{a}_k$. The estimate $\hat{a}_k$ is given by the form

$$\hat{a}_k = \hat{a}_{k|k-1} + K(k)[x_k - \hat{a}_{k|k-1}] \tag{4.14}$$

which is the optimum linear MMSE estimate and, therefore, the constrained receiver is optimized under the minimization of probability of error criterion by using a MMSE estimator.

Q.E.D.

4.3. Estimator with Certain Observations

4.3.1. Estimator Design

The objective here is to develop limited-memory estimators with decision-feedback for the two sequences of uncertain parameters $\{a_k\}$ and $\{b_k\}$ which were defined in the previous chapter. The means and the correlation parameters of the two sequences are as defined in (3.4). The estimation criterion used is MMSE and this optimizes the constrained receiver as discussed in the previous section. This implies that the estimates $\hat{a}_k$ and $\hat{b}_k$ are computed so as to minimize $E\{(a_k - \hat{a}_k)^2|\mathbf{Z}_k\}$ and $E\{(b_k - \hat{b}_k)^2|\mathbf{Z}_k\}$. The estimate is assumed to be a linear function of the observations and nonlinearity is introduced in the coefficients by the decision sequence $\mathbf{Z}_k$. The structure of the estimators is assumed to be

$$\hat{a}_k = \sum_{i=1}^{M} \alpha_i[Z_{k-i}X^1_{k-i} + (1 - Z_{k-i})X^0_{k-i}] \tag{4.15}$$

$$\hat{b}_k = \sum_{i=1}^{M} \gamma_i[Z_{k-i}Y^1_{k-i} + (1 - Z_{k-i})Y^0_{k-i}] \tag{4.16}$$

The memory length of the estimator is M in that the estimates are based only on the past M observations and decisions. The observations X^j_{k-i}, Y^j_{k-i}, j = 0,1 are as defined in (3.3). The estimator design involves the evaluation of the coefficients $\{\alpha_i,\gamma_i\}$ which are computed so as to minimize the MSE. If the decisions are assumed to be error-free then the binary sequence $\{Z_k\}$ represents the sequence of transmitted digits which are assumed to be independent of each other for the digital communication system under consideration. In this case Z_{k-i} is just a binary variable and estimators given by (4.15) and (4.16) are still linear. All the observations X^i_k, Y^i_k, i = 0,1 contain the uncertain parameters to be estimated. Under these assumptions, the constants $\{\alpha_i,\gamma_i\}$ which yield the optimal estimators are computed from the following orthogonality conditions

$$E\{(a_k - \hat{a}_k)[Z_{k-i}X^1_{k-i} + (1-Z_{k-i})X^0_{k-i}]|\mathbf{Z}_k\} = 0 , \qquad i = 1..M \qquad (4.17)$$

$$E\{(b_k - \hat{b}_k)[Z_{k-i}Y^1_{k-i} + (1-Z_{k-i})Y^0_{k-i}]|\mathbf{Z}_k\} = 0 , \qquad i = 1..M \qquad (4.18)$$

For convenience only (4.17) is simplified and the derivations leading to an equation in α_j are shown. The equation in γ_j can be obtained in an identical fashion and only the final result is given. Equation (4.17) yields

$$E\{\hat{a}_k[Z_{k-i}X^1_{k-i} + (1-Z_{k-i})X^0_{k-i}]|\mathbf{Z}_k\}$$
$$= E\{a_k[Z_{k-i}X^1_{k-i} + (1-Z_{k-i})X^0_{k-i}]|\mathbf{Z}_k\},$$
$$i = 1,\ldots M \qquad (4.19)$$

or

$$E\{[\sum_{j=1}^{M}\alpha_j(Z_{k-j}X^1_{k-j} + (1-Z_{k-j})X^0_{k-j}]\,[Z_{k-i}X^1_{k-i} + (1-Z_{k-i})X^0_{k-i}]\,|Z_k\}$$

$$= R_a(i), \qquad i = 1,\ldots M \qquad (4.20)$$

or

$$\sum_{j=1}^{M} \alpha_j[R_a(i-j) + N_0\delta_{ij}] = R_a(i), \qquad i = 1,\ldots M \qquad (4.21)$$

Similarly, for γ_j,

$$\sum_{j=1}^{M} \gamma_j[R_b(i-j) + N_0\delta_{ij}] = R_b(i), \qquad i = 1,\ldots M \qquad (4.22)$$

The coefficients $\{\alpha_j,\gamma_j\}$ may be obtained by solving (4.21) and (4.22). An appropriate model for the sequences $\{a_k\}$ and $\{b_k\}$ is needed prior to the computation of $\{\alpha_j,\gamma_j\}$. An initial value of the estimate is selected for the initialization of the estimator. If $k < M$, the estimate $\hat{a}_k$ is based on the previous k observations. If $k \geq M$, M of the most recent observations are employed to compute the estimate. In high speed digital communication systems, the steady-state operation of the estimator is important and not the initialization period.

In the above analysis, the decisions are assumed to be error-free. In practice, however, the assumption is not valid. If a wrong decision is reached, then the corresponding observation does not contain the uncertain parameter to be estimated. The equations (4.21) and (4.22) can be extended to include the decision uncertainty but they result in suboptimal estimators since the orthogonality condition which yields optimal linear estimators does not result in optimal estimators for this nonlinear problem. The extended equations in α_j and γ_j are

$$\sum_{j=1}^{M} \alpha_j[(1-p)^2 R_a(i-j) + N_0\delta_{ij}] = (1-p)\, R_a(i), \quad i = 1,\ldots M \tag{4.23}$$

$$\sum_{j=1}^{M} \gamma_j[(1-p)^2 R_b(i-j) + N_0\delta_{ij}] = (1-p)\, R_b(i), \quad i = 1,\ldots M \tag{4.24}$$

where p is the average probability of error. The performance of the estimators is investigated in the next section.

4.3.2. Estimator Performance

The performance of the estimator is examined by computing the MSE, which for the estimation of a_k is given by

$$MSE(\hat{a}_k) = E\{(a_k - \hat{a}_k)^2\} \tag{4.25}$$

Since the estimate $\hat{a}_k$ is a linear function of the observations, the orthogonality condition implies that

$$E\{(a_k - \hat{a}_k)\hat{a}_k\} = 0 \tag{4.26}$$

and, therefore,

$$\begin{aligned} MSE(\hat{a}_k) &= E\{(a_k - \hat{a}_k)a_k\} \\ &= R_a(0) - \sum_{i=1}^{M} \alpha_i R_a(i) \end{aligned} \tag{4.27}$$

Similarly, the MSE for the estimate $\hat{b}_k$ is given by

$$MSE(\hat{b}_k) = R_b(0) - \sum_{i=1}^{M} \gamma_i R_b(i) \tag{4.28}$$

In (4.27) and (4.28) the estimator performance is computed based on the assumption of error-free decisions. In practice, however, the detector makes errors and the performance of the decision-feedback estimator is worse than is given in (4.27) and (4.28). If the detector makes $N(N \leq M)$ errors in the previous M decisions, the MSE could be expressed as

$$\begin{aligned} MSE(\hat{a}_k) &= R_a(0) - \sum_{i=1}^{M} \alpha_i R_a(i) + p \sum_{i_1=1}^{M} \alpha_{i_1} R_a(i_1) \\ &+ \frac{p^2}{2!} \sum_{i_1=1}^{M} \sum_{\substack{i_2=1 \\ i_2 \neq i_1}}^{M} [\alpha_{i_1} R_a(i_1) + \alpha_{i_2} R_a(i_2)] + \ldots \\ &+ \frac{p^n}{n!} \sum_{i_1=1}^{M} \cdots \sum_{\substack{i_N=1 \\ i_N \neq i_j}}^{M} [\alpha_{i_1} R_a(i_1) + \ldots + \alpha_{i_N} R_a(i_N)] \end{aligned} \tag{4.29}$$

where p is the average probability of error. A similar expression for $MSE(\hat{b}_k)$ can be obtained.

4.3.3. Example

It is assumed that $\{a_k\}$ and $\{b_k\}$ are Markov sequences with correlation coefficient ρ. The means are assumed to be zero and variances are assumed to be one. The correlations are then given by

$$R_a(k-\ell) = R_b(k-\ell) = \rho^{|k-\ell|} \tag{4.30}$$

It is assumed that the decisions are error-free so that the equations which yield the coefficients $\{\alpha_i\}$ and $\{\gamma_i\}$ are

$$\sum_{j=1}^{M} \alpha_j[\rho^{|j-i|} + N_0\delta_{ij}] = \rho^i , \qquad i = 1,\ldots M \qquad (4.31)$$

and

$$\sum_{j=1}^{M} \gamma_j[\rho^{|j-i|} + N_0\delta_{ij}] = \rho^i, \qquad i = 1,\ldots M \qquad (4.32)$$

In this example, MSE for the two estimators based on the assmuption of error-free reception are obtained. Equation (4.29) can be employed to obtain the MSE when the assumption of error-free reception is not valid.

$$\mathrm{MSE}(\hat{a}_k) = 1 - \sum_{i=1}^{M} \alpha_i\rho^i \qquad (4.33)$$

and

$$\mathrm{MSE}(\hat{b}_k) = 1 - \sum_{i=1}^{M} \gamma_i\rho^i \qquad (4.34)$$

The MSE was computed as a function of M, the memory length. The numerical results obtained with $\rho = .9$ and $N_0 = 1$ are presented in Table 4.1 and plotted in Fig. 4.1. It is obvious from the expressions (4.33) and (4.34) that the MSE decreases as M increases. The results obtained assume the steady-state operation of the estimator. As indicated earlier in high speed digital communication system applications, steady state operation of the estimator is of interest and not the initialization period.

4.4. Estimator with Uncertain Observations

4.4.1. General

In the digital communication systems where the received signal does not always contain the uncertain parameters to be estimated, the

$$\sum_{j=1}^{M} \alpha_j[\rho^{|j-i|} + N_0\delta_{ij}] = \rho^i , \qquad i = 1,\ldots M \tag{4.31}$$

and

$$\sum_{j=1}^{M} \gamma_j[\rho^{|j-i|} + N_0\delta_{ij}] = \rho^i, \qquad i = 1,\ldots M \tag{4.32}$$

In this example, MSE for the two estimators based on the assmuption of error-free reception are obtained. Equation (4.29) can be employed to obtain the MSE when the assumption of error-free reception is not valid.

$$MSE(\hat{a}_k) = 1 - \sum_{i=1}^{M} \alpha_i\rho^i \tag{4.33}$$

and

$$MSE(\hat{b}_k) = 1 - \sum_{i=1}^{M} \gamma_i\rho^i \tag{4.34}$$

The MSE was computed as a function of M, the memory length. The numerical results obtained with $\rho = .9$ and $N_0 = 1$ are presented in Table 4.1 and plotted in Fig. 4.1. It is obvious from the expressions (4.33) and (4.34) that the MSE decreases as M increases. The results obtained assume the steady-state operation of the estimator. As indicated earlier in high speed digital communication system applications, steady state operation of the estimator is of interest and not the initialization period.

4.4. Estimator with Uncertain Observations

4.4.1. General

In the digital communication systems where the received signal does not always contain the uncertain parameters to be estimated, the

Table 4.1

Memory length	MSE
1	.5950
2	.4922
3	.4572
4	.4441
5	.4391
6	.4371
7	.4369
8	.4361
9	.4360
10	.4359

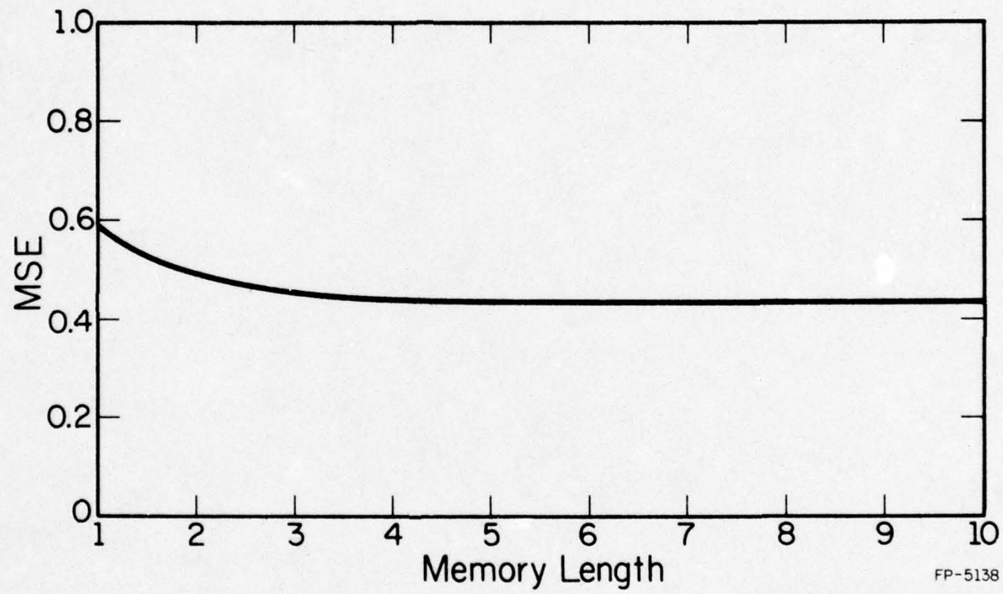

Figure 4.1 MSE of the estimator with certain observations versus memory length.

estimation algorithm described in the previous section needs to be modified. In binary on-off keying systems, the received signal does not contain the uncertain parameter to be estimated if a zero is transmitted. The estimation problem with uncertain observations has been treated in the literature recently [60-62]. Nahi [60] developed an optimal MMSE linear recursive estimator. His estimate was based on all the available data. The binary indicator random variables were assumed to be independent of each other. Jackson and Murthy [61] generalized the estimation algorithm to include limited dependence of the binary random variables characterizing the uncertainty. Sawaragi et al. [62] employed the Bayesian approach to obtain an approximate nonlinear estimator for sequential state estimation with interrupted observation mechanism.

In the present work, it is desired to derive an estimator with uncertain observations which is used in conjunction with a detector for digital data reception over channels with memory. The desired estimator is a modified version of Nahi's linear recursive estimator in that the filter memory is limited so as to avoid filter divergence. It is assumed that the transmitted digits are independent of each other. A finite-state Markov source could also be considered and an analysis similar to Jackson and Murthy [61] may be used for estimator design. The presence of the detector is utilized and a decision-directed scheme is derived in the next section.

4.4.2. Estimator Derivation

The objective is to derive MMSE limited-memory estimators with decision-feedback and uncertain observations for the sequences of channel uncertain parameters $\{a_k\}$ and $\{b_k\}$. The means and the correlations of the sequences are as defined previously in (3.4). The estimates $\hat{a}_k$ and $\hat{b}_k$ are computed so as to minimize the MSE, i.e., $E\{(a_k - \hat{a}_k)^2 | \mathbf{Z}_k\}$ and $E\{(b_k - \hat{b}_k)^2 | \mathbf{Z}_k\}$. The estimates are assumed to be linear functions of the observations and the nonlinearity is introduced in the coefficients through the decision sequence $\mathbf{Z}_k$. The structure of the estimators is assumed to be the same as (4.15) and (4.16). The estimator coefficients $\{\alpha_i, \gamma_i\}$ are evaluated so as to minimize the MSE. It is assumed that all the decisions are error-free and, therefore, $\{Z_k\}$ represents the sequence of transmitted digits. The set of estimator coefficients for the linear problem is given by the orthogonality conditions of (4.17) and (4.18). Proceeding in a similar manner as outlined in the previous section, the equations which yield the coefficients $\{\alpha_i, \gamma_i\}$ are obtained and are given by

$$\sum_{i=1}^{M} \{Z_{k-i} Z_{k-j} R_a(i-j) + N_0 \delta_{ij}\} \alpha_i = Z_{k-j} R_a(j), \qquad j = 1, \ldots M \tag{4.35}$$

$$\sum_{i=1}^{M} \{Z_{k-i} Z_{k-j} R_b(i-j) + N_0 \delta_{ij}\} \gamma_i = Z_{k-j} R_b(j), \qquad j = 1, \ldots M \tag{4.36}$$

Notice the difference between equations (4.21), (4.22), and (4.35), (4.36) due to uncertain observations. Equations (4.35) and (4.36) can be used to compute the estimator coefficients. These equations are, however, data dependent in that they depend upon the decision sequence $\mathbf{Z}_k$. An on-line computation of the estimator coefficients is necessary which is not attractive computationally. The equations are, therefore, averaged over all possible combinations of Z_k. The estimator coefficients are computed off-line from the resulting equations and stored for on-line applications. The equations which yield the estimator coefficients $\{\alpha_i, \gamma_i\}$ are given as

$$\{p_1 R_a(0) + N_0\}\,\alpha_j + \sum_{\substack{i=1\\ i\neq j}}^{M} p_1^2 R_a(i-j)\,\alpha_i = p_1 R_a(j), \qquad j = 1,\ldots M \qquad (4.37)$$

$$\{p_1 R_b(0) + N_0\}\,\gamma_j + \sum_{\substack{i=1\\ i\neq j}}^{M} p_1^2 R_b(i-j)\,\gamma_i = p_1 R_b(j), \qquad j = 1,\ldots M \qquad (4.38)$$

where p_1 denotes the probability that a one is transmitted. Under the assumption of error-free decisions, $p_1 = .5$.

In practice the assumption of error-free decisions is not valid. If a wrong decision is reached, then the information provided by the decision about the presence of the uncertain parameter in the observation is not correct. Equations (4.37) and (4.38) can be extended to include this decision uncertainty but they result in suboptimal estimators since the orthogonality condition which yields optimal linear estimators does not result in optimal estimators for this nonlinear problem. Equations (4.37) and (4.38) when extended become

$$\{p_1(1-p)R_a(0) + N_0\}\,\alpha_j + \sum_{\substack{i=1\\ i\neq j}}^{M} \{p_1^2(1-p)^2 R_a(i-j)\}\,\alpha_i = p_1(1-p)R_a(j),$$

$$j = 1,\ldots M \qquad (4.39)$$

$$\{p_1(1-p)R_b(0)+N_0\}\,\gamma_j + \sum_{\substack{i=1\\ i\neq j}}^{M} \{p_1^2(1-p)^2R_b(i-j)\}\,\gamma_i = p_1(1-p)R_b(j),$$

$$j = 1,\ldots M \qquad (4.40)$$

where p denotes the probability of error and p_1 is given by

$$p_1 = \frac{1}{2}\Big[\int\int_{\Omega_k^1} f_{X_k^0Y_k^0}(x,y\,|\,H_k^0,\mathbf{Z}_k)\ dxdy$$

$$+ \int\int_{\Omega_k^1} f_{X_k^1Y_k^1}(x,y\,|\,H_k^1,\mathbf{Z}_k)\ dxdy\Big] \qquad (4.41)$$

Thus, equations which can be used to derive the estimator coefficients have been developed and in the next section the performance of the estimators is considered.

4.4.3. Estimator Performance

The performance of the estimator is evaluated by computing the MSE. For the class of estimators under consideration here, (4.26) still holds and the MSE is obtained in an identical fashion as for the estimator with certain observations.

$$\mathrm{MSE}(\hat{a}_k) = E\{(a_k - \hat{a}_k)a_k\}$$

$$= R_a(0) - \sum_{i=1}^{M} \alpha_i Z_{k-i} R_a(i) \qquad (4.42)$$

Similarly,

$$\mathrm{MSE}(\hat{b}_k) = R_b(0) - \sum_{i=1}^{M} \gamma_i Z_{k-i} R_b(i) \qquad (4.43)$$

Equations (4.42) and (4.43) are based on the assumption of error-free decisions. An expression for the MSE when all the decisions are not error-free can be obtained in a similar manner as (4.29). The performance of the estimator is data dependent since it is based on the decision sequence $\mathbf{Z}_k$. Bounds on this conditional MSE can be computed and they are evaluated by computing the MSE under the best and worst operating conditions. The worst case occurs when all of the M previously received digits are zeroes and the MSE in this case is the upper bound (UB) on the estimator performance. From (4.42) and (4.43)

$$UB(\hat{a}_k) = R_a(0) \tag{4.44}$$

$$UB(\hat{b}_k) = R_b(0) \tag{4.45}$$

The lower bound (LB) is obtained by evaluating the MSE under the best operating conditions which occurs when all the previously received digits are ones. Again using (4.42) and (4.43)

$$LB(\hat{a}_k) = R_a(0) - \sum_{i=1}^{M} \alpha_i R_a(i) \tag{4.46}$$

$$LB(\hat{b}_k) = R_b(0) - \sum_{i=1}^{M} \gamma_i R_b(i) \tag{4.47}$$

The performance of the estimator as discussed above is conditioned on the decision sequence $\mathbf{Z}_k$. The MSE can be averaged over all the possible sequences and an average MSE can also be evaluated. The expressions for the average MSE are given by

$$\mathrm{MSE}_{ave}(\hat{a}_k) = R_a(0) - \sum_{i=1}^{M} p_1 \alpha_i R_a(i) \tag{4.48}$$

$$\mathrm{MSE}_{ave}(\hat{b}_k) = R_b(0) - \sum_{i=1}^{M} p_1 \gamma_i \; R_b(i) \tag{4.49}$$

A numerical example is considered next.

4.4.4. Example

For the numerical example, $\{a_k\}$ and $\{b_k\}$ are again assumed to be Markov sequences with correlation coefficient ρ. The means are assumed to be zero and variances are assumed to be one. The correlation functions are

$$R_a(k-\ell) = R_b(k-\ell) = \rho^{|k-\ell|} \tag{4.50}$$

The decisions are assumed to be error-free. The set of equations which yield the estimator coefficients $\{\alpha_i\}$ and $\{\gamma_i\}$ are

$$\sum_{i=1}^{M} \alpha_i [Z_{k-i} Z_{k-j} \; \rho^{|i-j|} + N_0 \delta_{ij}] = Z_{k-j} \; \rho^j, \quad j = 1,\ldots M \tag{4.51}$$

$$\sum_{i=1}^{M} \gamma_i [Z_{k-i} Z_{k-j} \; \rho^{|i-j|} + N_0 \delta_{ij}] = Z_{k-j} \; \rho^j, \quad j = 1,\ldots M \tag{4.52}$$

In the present work, a simulation is not performed and two specific examples are solved. Therefore, the actual equations described by (4.51) and (4.52) are employed for the evaluation of the estimator coefficients and not the equations given in (4.37) and (4.38). In practice, the estimator coefficients will be evaluated off-line and equations (4.37) and (4.38) will be employed. The conditional MSE for both the estimators with the assumption of error-free decisions is given by

$$\mathrm{MSE}(\hat{a}_k) = 1 - \sum_{i=1}^{M} \alpha_i Z_{k-i}\, \rho^i \tag{4.53}$$

$$\mathrm{MSE}(\hat{b}_k) = 1 - \sum_{i=1}^{M} \gamma_i Z_{k-i}\, \rho^i \tag{4.54}$$

The MSE for two decision sequences with different memory lengths are computed. Numerical results are obtained with $\rho = .9$ and $N_0 = 1$. The steady state operation of the estimator is assumed. The numerical results for the memory length four are presented in Table 4.2 and plotted in Fig. 4.2. The decision sequence $\mathbf{Z}_k$, i.e. $\{Z_{k-1}, Z_{k-2}, Z_{k-3}, Z_{k-4}\}$, was assumed to be $\{1001\}$. In the second example, the memory length is ten and the decision sequence $\mathbf{Z}_k$ is $\{110\ 1111\ 001\}$. The results are shown in Table 4.3 and Fig. 4.3.

4.5. Discussion

In this chapter, estimation algorithms for receivers with memory are discussed. The basic goal of the combined estimator-detector structure is signal detection. The estimation criterion which optimizes the receiver is derived. Limited-memory estimators with decision-feedback are discussed. The estimates are assumed to be linear functions of the observations with nonlinearity introduced through the coefficients which are functions of the past decisions. Two cases are considered. First, when all the observations contain the uncertain parameters to be estimated, and secondly when all the observations do not contain the channel uncertain parameters. Decisions are assumed to be error-free but results are extended to include the case

Table 4.2

Signalling Interval	Decision during the signalling interval	MSE
k	1	.5329
k + 1	1	.4922
k + 2	1	.4572
k + 3	0	.4441
k + 4	1	.5603
k + 5	1	.4901
k + 6	0	.4655
k + 7	0	.5887
k + 8	1	.6668
k + 9	1	.5329
k + 10	1	.4922
k + 11	0	.4572
k + 12	1	.4441
k + 13	1	.4441
k + 14	1	.4441
k + 15	1	.4441

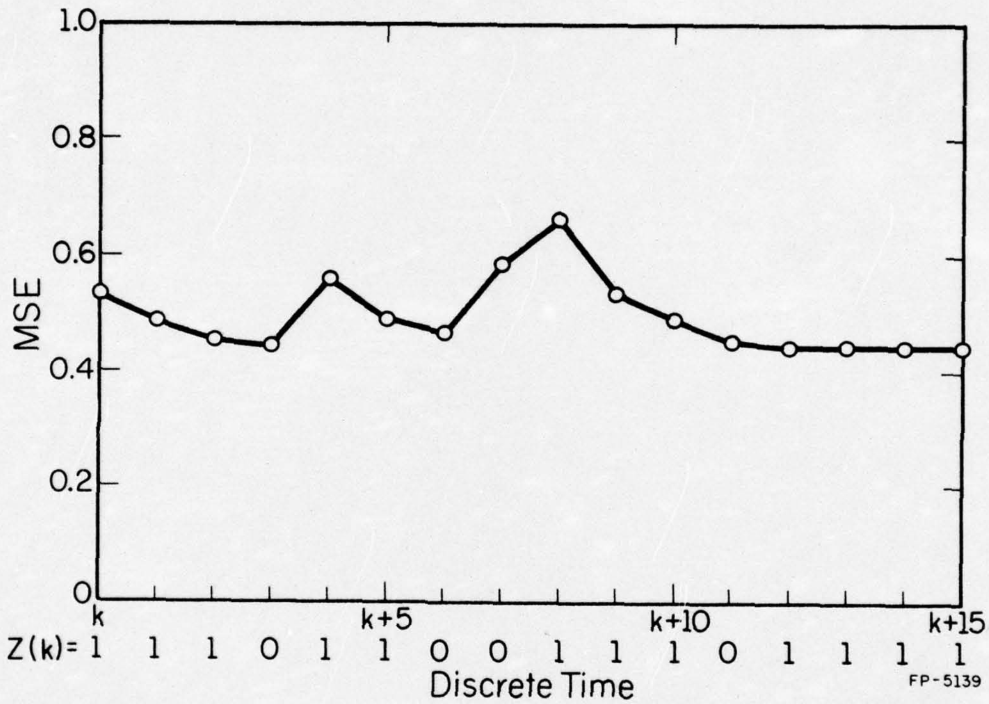

Figure 4.2 MSE of the estimator with uncertain observations with memory length four.

Table 4.3

Signalling interval	Decision during the signalling interval	MSE
k	0	.4514
k + 1	0	.5559
k + 2	1	.6403
k + 3	1	.5062
k + 4	1	.4623
k + 5	0	.4462
k + 6	1	.4400
k + 7	1	.4376
k + 8	1	.4366
k + 9	1	.4362
k + 10	0	.5435
k + 11	0	.6302
k + 12	1	.5031
k + 13	1	.4612
k + 14	0	.5637
k + 15	1	.4821
k + 16	0	.5807
k + 17	0	.6608
k + 18	1	.5126
k + 19	1	.4649

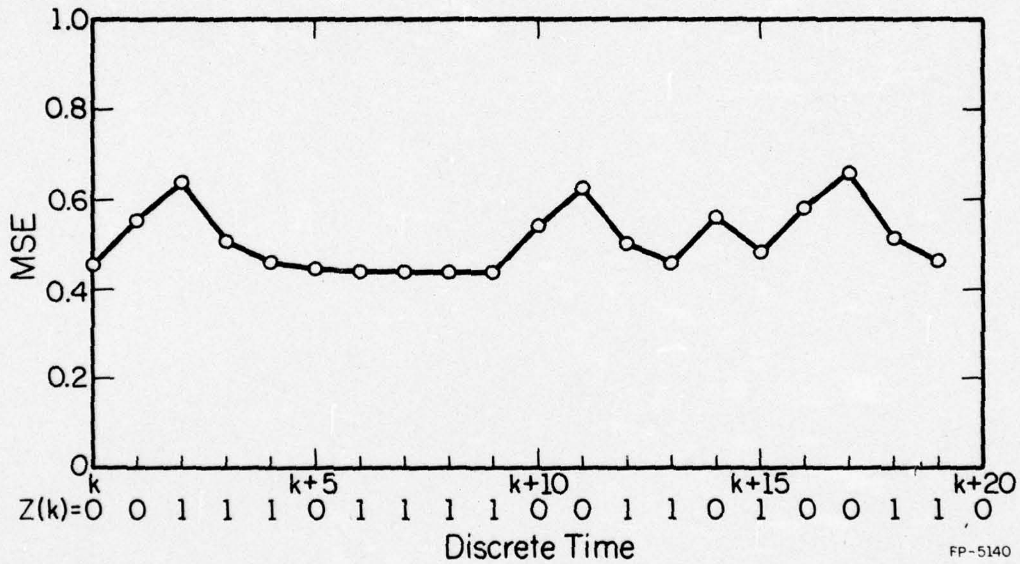

Figure 4.3 MSE of the estimator with uncertain observations with memory length ten.

when all the decisions are not assumed to be correct. The performance of the estimators is computed in terms of the MSE.

In practice, the detector is expected to make errors and the estimation problem is nonlinear. However, the average probability of error for digital communication systems is expected to be low and the probability of N errors ($N \geq 2$) in M decisions for a reasonable size M can be assumed to be negligible. Therefore, an estimator for this problem which is obtained by using the orthogonality condition is not optimal but is expected to be nearly optimal due to the almost error-free decision sequence.

The estimation algorithms developed in this chapter may also be used in control theory. One practical application could be tracking of a target trajectory where the target return is processed at discrete intervals and uncertain parameters are estimated. If observation mechanism breaks down, for instance, due to misalignment of antennas etc. the observations do not contain the parameters to be estimated and the estimation algorithm of section 4.4 could be employed in such cases. Thus, the estimation algorithms developed may find applications in a variety of practical problems.

5. MODELING OF DIGITAL CHANNELS

5.1. Introduction

Modeling of digital channels is an important problem in the theory of digital communications receiving considerable attention in the literature. Analytical models of digital channels find extensive use in the theoretical prediction of error-rate and other channel error statistics. These predicted values of error statistics assist in the comparative analysis of different modems and also in the analysis and performance evaluation of various error-control techniques. It is desirable for a communication system designer to be able to predict the preformance before selecting the system, i.e., the modem and the coding techniques. Digital channel models also provide an insight into the clustering of errors and this may help in the development of more efficient and reliable communication techniques. Thus, the channel modeling problem is of current interest with potential applications in the design of data communication networks where it is essential that the data links be characterized accurately.

There have been two basic approaches to the channel modeling problem. The first approach has been pursued by Bello and his colleagues at SIGNATRON Inc. They consider the actual physical processes present in the transmission media which are responsible for the channel behavior observed in practice. Basic results from the electromagnetic propagation theory have been used and the relationship between the random disturbance processes and the actual antenna parameters of the communication links have been derived. Bello [63] developed a mathematical model for the troposcatter channel. Scattering and other related phenomena are used to derive the channel model. The transfer function is assumed to describe the channel and the statistics are completely determined by the time-frequency correlation function. The model has been employed to predict the error rates and for the comparison of the performances of different modems [64-65]. The model has also been used to predict the performance of various codes and coding techniques [66] by computing the probability of m errors in a block of length n. This error statistics provides a tool for the performance evaluation of block codes but does not give an insight into the actual stochastic behavior of the channel error sequence.

The second class of channel models try to characterize the input-output behavior of the channel. The actual physical processes present in the channel are not taken into account and accurate stochastic models which represent the stochastic behavior of the channel error sequence are developed. Three major types of input-output models have been proposed. The simplest of them all are the renewal models [67-69]. They assume that the error sequence is a discrete renewal process and the occurrence of an error depends only on the time elapsed since the last error occurrence. The second group consists of models which consider the bit-by-bit behavior of the error sequence. Gilbert [70] proposed a two-state Markov model which was later generalized by Fritchman [71] who discussed a partitioned Markov chain model. Haddad et.al. [72] considered the statistical behavior of gaps and developed the third class of models which are the Markov gap models. These models were later extended to include some additional memory which characterized the short-term error behavior more accurately [73,74]. The input-output models have been employed successfully to predict the performance of error-control techniques. Clustering of errors and the concept of multigaps has been explored by Adoul [75].

It has been discussed above that the model developed by Bello considers the actual physical processes involved whereas the input-output models try to model the stochastic behavior of the error sequence. It is expected that a model which incorporates both features would represent a channel more exactly. It is, therefore, desirable to consider a unified treatment of the channel modeling problem in that a description of the actual processes is utilized to model the input-output behavior. In section 5.2, attention is focussed on some aspects of a unified treatment of the channel modeling problem. The main objective of this chapter, however, is to examine the relationship between receivers with memory and error clustering. This fits into the general framework of the channel modeling problem since the channel model depends upon the actual communication system in use and, therefore, it depends upon the receiver. To examine the relationship of receivers and error clustering it is necessary to define some measures which quantitatively characterize this relationship. These measures would describe the statistical behavior of errors in the channel error sequence. In section 5.3, the measures are defined and numerical results are obtained in the last section which represent the relationship of receivers and channel models quantitatively for special cases.

AD-A031 559 ILLINOIS UNIV AT URBANA-CHAMPAIGN COORDINATED SCIENCE LAB F/G 17/2
MODELS AND EFFICIENT RECEIVERS FOR COMMUNICATION CHANNELS WITH --ETC(U)
SEP 76 P K VARSHNEY DAAB07-72-C-0259
UNCLASSIFIED R-744 NL

2 OF 2
AD
A031559

END
DATE
FILMED
12-76

5.2. Channel Modeling Considerations

It has been indicated previously that a unified treatment of the channel modeling problem will be briefly presented in this section. Actual physical processes are taken into account while modeling the input-output behavior of the channel. To accomplish this, it is essential to consider the actual communication system. The communication system description would help in incorporating the knowledge about the actual physical processes involved into the channel model. It is assumed that the communication system described in Chapter three is under operation. The memory in the channel is introduced by the pair of dependent random variables $\{v_k, \theta_k\}$. The statistics and other description of these random variables is employed in deriving a channel model.

First the basic framework and some terminology is introduced. The channel is assumed to be as shown in Fig. 5.1. It consists of an information source which generates the input sequence $\{x_i\}$. The channel corrupts the transmitted information with noise and produces an output sequence $\{y_i\}$. It is assumed that the noise sequence, which is denoted by $\{n_i\}$, is independent of the input sequence $\{x_i\}$. The input and output symbols are the elements of the Galois Field GF(q) with operation $\oplus$, which is addition modulo q. The output sequence is generated by the summation of $\{x_i\}$ and $\{n_i\}$ over GF(q). The channel error sequence $\{e_i\}$ is defined as a mapping ϕ from the noise sequence $\{n_i\}$ onto the set $\{0,1\}$ so that the value of e_i is given by

$$e_i = \phi(n_i) = \begin{cases} 0 & \text{if } n_i = 0 \\ 1 & \text{otherwise} \end{cases} \tag{5.1}$$

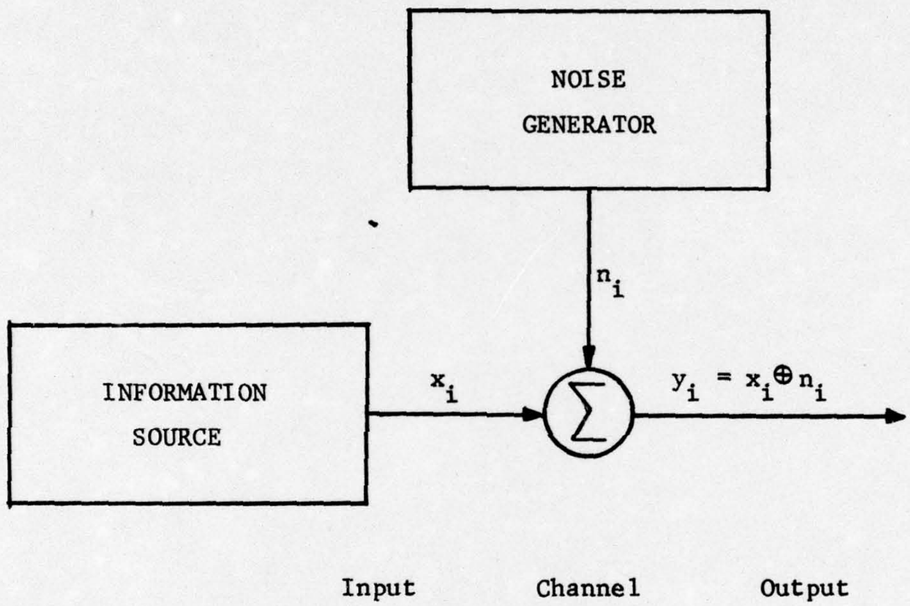

Figure 5.1. Digital communication channel.

Thus, $\{e_i\}$ is a discrete random process taking on values zero and one. In the binary sequence $\{e_i\}$, an occurrence of one indicates the presence of an error. It is this discrete error sequence $\{e_i\}$ whose stochastic behavior is to be described.

Digital channel models could be generated which utilize the properties of the channel in representing the input-output behavior. The generation of model requires the computation of probabilities of various error clusters or configurations in the sequence $\{e_i\}$. The error sequence is assumed to be stationary and ergodic so that the computation of the statistics of any particular sequence does, indeed, yield the channel model. An adequate description of a memoryless channel is the average probability of error. The models for channels with memory are obtained by evaluating the probabilities of error clusters. For example, the Gilbert's model could be generated by computing the conditional probabilities $P\{(e_k = 0)|(e_{k-1} = 0)\}$, $P\{(e_k = 0)|(e_{k-1} = 1)\}$, $P\{(e_k = 1)|(e_{k-1} = 0)\}$ and $P\{(e_k = 1)|(e_{k-1} = 1)\}$ for any pair of adjacent error bits. The expression for one of these four probabilities is evaluated for illustration purposes.

$$P\{(e_k = 1)|(e_{k-1} = 1)\} = P\{(e_k = 1)\cap(e_{k-1} = 1)\}/P(e_{k-1} = 1) \qquad (5.2)$$

The joint probability may be obtained from the following

$$P((e_k=1)\cap(e_{k-1}=1)) = \iiiint\limits_{\Omega_{00}} f_{X^1_{k-1}Y^1_{k-1}X^1_kY^1_k}(u,v,x,y)\,du\,dv\,dx\,dy$$

$$+ \iiiint\limits_{\Omega_{01}} f_{X^0_{k-1}Y^0_{k-1}X^1_kY^1_k}(u,v,x,y)\,du\,dv\,dx\,dy$$

$$+ \iiiint_{\Omega_{10}} f_{X^0_{k-1}Y^0_{k-1}X^1_kY^1_k}(u,v,x,y)\,dudvdxdy$$

$$+ \iiiint_{\Omega_{11}} f_{X^0_{k-1}Y^0_{k-1}X^0_kY^0_k}(u,v,x,y)\,dudvdxdy \tag{5.3}$$

where Ω_{ij} represents the optimum decision regions obtained using the appropriate detection scheme when the combination of the hypotheses $H^i_{k-1}H^j_k$ is true. If all the four conditional probabilities are evaluated it results in the parameters required to characterize the Gilbert model. It does not necessarily prove the validity of the model. Other more complicated models like the partitioned Markov chain models and the Markov gap models which are based on the knowledge of the physical processes can be derived in a similar manner but at the expense of a considerable amount of computational complexity. A receiver with or without memory could be employed and the channel model obtained will correspond to that particular communication system.

In the present work, the objective is not to generate complex channel models and, therefore, the subject is not pursued any further. In this section, the approach to be used for the modeling of a communication system has been indicated. The goal in this chapter is to examine the relationship between the receivers with memory and error clustering. In the next section, the measures which are used to achieve this goal are defined and discussed. These measures try to characterize the statistical dependence of errors and are, thus, essential to the understanding of channels with memory and their modeling.

5.3. Measures of Channel Memory

In this section, an attempt is made to define some measures of channel memory in terms of the input-output model. It is expected that these measures will have applications in modeling of channels and also communication system design. These measures will also describe the clustering properties of the error sequence of a communication system. In other words, these measures provide information about the occurrence of error events based on the past errors.

The channel error sequence $\{e_i\}$ is assumed to be stationary with

$$P\{e_i=1\} = p \tag{5.4}$$

where p is the average error rate. Two measures are considered. The first measure is obtained from statistical considerations and the other derived from the information theoretic point of view. In statistics, the correlation of two random variables is determined by means of the correlation coefficient. A similar measure is defined here using the same concept to describe the dependence of error events $\{e_i\}$.

Definition: The (n+1)-th order generalized correlation coefficient $\alpha_{n+1}(m_1,m_2\ldots m_n)$ of the sequence of random variables $\{e_i\}$ is defined as

$$\alpha_{n+1}(m_1,m_2\ldots m_n) = \frac{E\{(e_k-p)(e_{k-m_1}-p)\ldots(e_{k-m_n}-p)\}}{\sigma^{n+1}} \tag{5.5}$$

where $\sigma^2 = p(1-p)$ is the variance of e_i. It is observed that $\alpha_1 = 0$ and the values of α_i, $i > 1$, can be computed from (5.5).

The second measure is derived using the information theoretic approach. The dependence of error events is considered and a measure in terms of the mutual information [76] is described.

<u>Definition</u>: The (n+1)-th order error clustering coefficient $\beta_{n+1}(m_1,\ldots m_n)$ is defined as

$$\beta_{n+1}(m_1,m_2,\ldots m_n) = \ell n \frac{P[(e_k=1)\cap(e_{k-m_1}=1)\cap\ldots\cap(e_{k-m_n}=1)]}{P(e_k=1)P(e_{k-m_1}=1)\ \ldots\ P(e_{k-m_n}=1)} \tag{5.6}$$

This coefficient provides a quantitative measure of the dependence of various error events. If $\beta_{n+1}(m_1,\ldots m_n) = 0$, the error events are independent. A negative value of $\beta_{n+1}(m_1,\ldots m_n)$ indicates negative correlation and a positive value implies positive correlation. If the event $[(e_{k-m_j}=1)\cap\ldots\cap(e_{k-m_n}=1)]$ is denoted by E_j and the event $(e_{k-m_j}=1)$ by Ψ_j, then $\beta_{n+1}(m_1,\ldots m_n)$ can be expressed in terms of the mutual informations

$$\beta_{n+1}(m_1,\ldots m_n) = \ell n \frac{P[\Psi_0|E_1]}{P(\Psi_0)} + \ell n \frac{P[\Psi_1|E_2]}{P(\Psi_1)} + \ldots$$

$$+ \ell n \frac{P[\Psi_{m-2}|E_{m-1}]}{P(\Psi_{m-2})} + \ell n \frac{P[\Psi_{m-1}]}{P(\Psi_{m-1})}$$

$$= I(\Psi_0,E_1) + I(\Psi_1,E_2) + \ldots + I(\Psi_{m-2},E_{m-1})$$

$$= \sum_{j=0}^{n-1} I(\Psi_j,E_{j+1}) \tag{5.7}$$

where $I(\Psi_j,E_{j+1})$ represents the mutual information of the two events Ψ_j and E_{j+1}. Thus, $\alpha_{n+1}(m_1,\ldots m_n)$ and $\beta_{n+1}(m_1,\ldots m_n)$ define two measures which quantitatively specify the dependence of various error events. These measures can be employed to examine the relationship between receivers and channel models. It is expected that a receiver which exploits the correlation of the received data would alter the correlation properties of the error sequence and thereby changing the channel model. The measures can be computed for the actual communication system in use if an appropriate model for the uncertain parameters of the channel is assumed. In the next section, the computational procedure is illustrated by means of an example.

5.4. Receivers and Error Clustering

In this section, the effect of the receivers on error clustering is examined. The measures defined earlier provide an insight into the clustering of errors. The measures describing the correlation properties of the error sequence can be computed for different communication systems. In particular, in this section the coefficients $\alpha_{n+1}(m_1,\ldots m_n)$ and $\beta_{n+1}(m_1,\ldots m_n)$ are computed for the communication systems using the receivers both with and without memory. As noted these coefficients provide quantitative information about the occurrence of errors conditioned on past errors. A higher value of $\alpha_{n+1}(m_1,\ldots m_n)$ and $\beta_{n+1}(m_1,\ldots m_n)$ provides more information about the occurrence of errors. Thus, the errors are more predictable for such communication systems and it is expected that this may help in the development of more efficient and reliable communication techniques.

The computation procedure is illustrated by considering an example. The coherent on-off keying system example considered in chapter three is pursued and the correlation and error clustering coefficients are computed.

In particular, $\alpha_2(1)$ and $\beta_2(1)$, the second order coefficients are evaluated. Higher order coefficients $\alpha_{n+1}(m_1,\ldots m_n)$ and $\beta_{n+1}(m_1,\ldots m_n)$ can be computed in a similar fashion. As special cases of (5.5) and (5.6),

$$\alpha_2(1) = \frac{E\{(e_k-p)(e_{k-1}-p)\}}{p(1-p)}$$

$$= \frac{P\{(e_k=1)\cap(e_{k-1}=1)\}-p^2}{p(1-p)} \tag{5.8}$$

and

$$\beta_2(1) = \ell n\, \frac{P\{(e_k=1)\cap(e_{k-1}=1)\}}{P(e_k=1)\; P(e_{k-1}=1)}$$

$$= \ell n\, \frac{P\{(e_k=1)\cap(e_{k-1}=1)\}}{p^2} \tag{5.9}$$

In order to be able to compute α_2 and β_2 it is necessary to evaluate the joint probability $P\{(e_k=1)\cap(e_{k-1}=1)\}$. For the example under consideration, the joint probability can be expressed as

$$P\{(e_k=1)\cap(e_{k-1}=1)\} = \mathrm{erfc}(T_k/N_0^{\frac{1}{2}})\{\mathrm{erfc}(T_k/N_0^{\frac{1}{2}}) + 2\ \mathrm{erf}[(T_k-\xi_k)/(N_0+\zeta_k)^{\frac{1}{2}}]\}$$

$$+ \int_{-\infty}^{T_k}\int_{-\infty}^{T_k} [2\pi(N_0+\zeta)(1-\rho^2)^{\frac{1}{2}}]^{-1} \exp\{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)(N_0+\zeta)}\}dxdy \tag{5.10}$$

The coefficients α_2 and β_2 are computed as a function of signal to noise ratio and the results are obtained for both the receivers i.e. the optimum receiver without memory and the constrained receiver. These results are presented in Tables 5.1 and 5.2 and in Fig. 5.2 and Fig. 5.3. It is

Table 5.1

SNR	α_2 Optimal receiver without memory	α_2 Constrained receiver with memory
1	.3020	.2901
10	.5050	.5066
10^2	.5832	.6640
10^3	.5878	.7320
10^4	.5883	.7573
10^5	.5884	.7662

Table 5.2

SNR	β_2	
	Optimal receiver without memory	Constrained receiver with memory
1	.3418	.3328
10	.8488	.8529
10^2	1.0545	1.3692
10^3	1.0633	1.6279
10^4	1.0642	1.7314
10^5	1.0643	1.7693

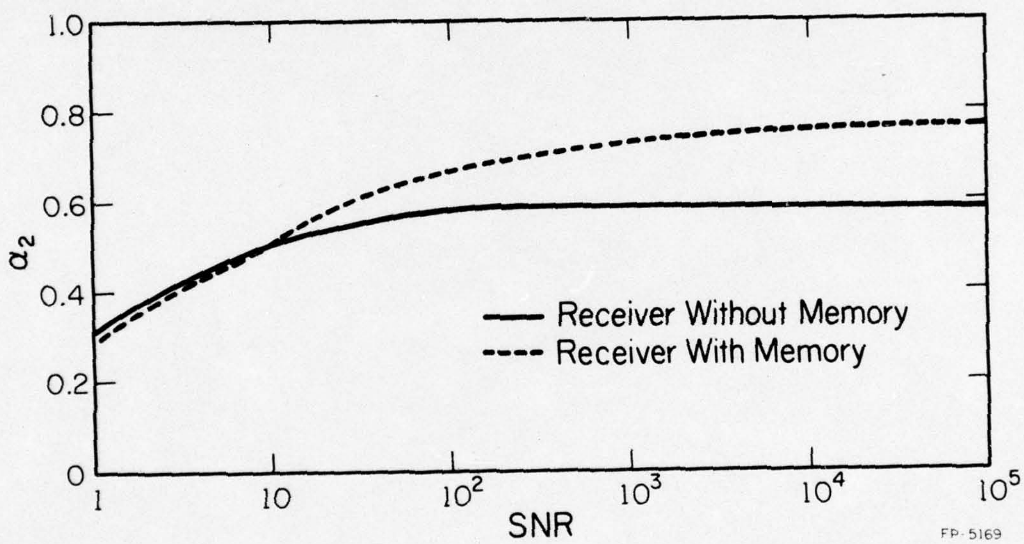

Figure 5.2 α_2 for the receivers with and without memory.

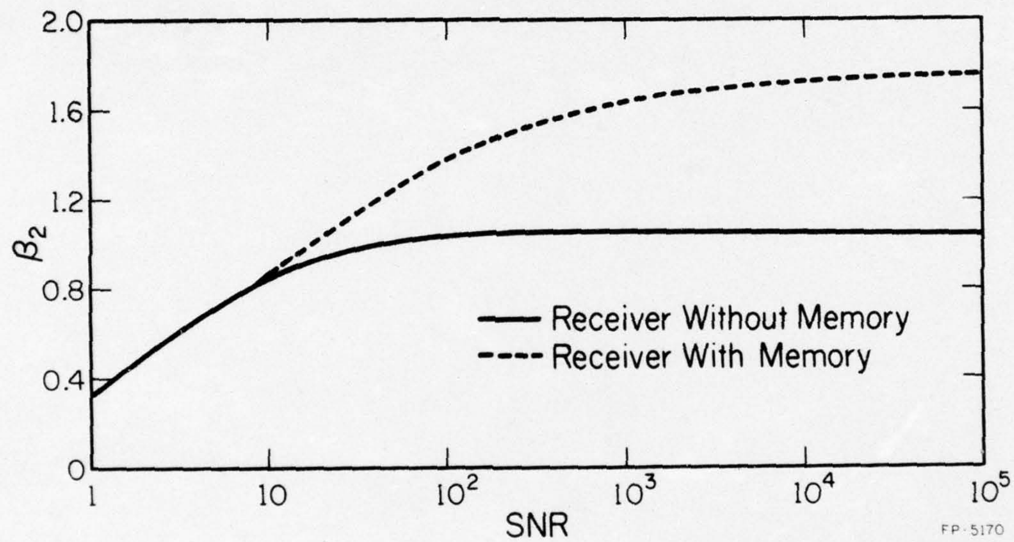

Figure 5.3 β_2 for the receivers with and without memory.

observed that the values of α_2 and β_2 for communication system using the receiver with memory are greater than the ones obtained when using the receiver without memory. This indicates that more information about the occurrence of errors is available and errors are more predictable when using the receiver with memory. This is the reason that the performance of receivers with memory is better and also the fact that the errors are more predictable can be used to devise more efficient communication techniques. The discrepancy in the behavior of α_2, β_2 for low SNR can be attributed to the approximations made in the design of the receiver with memory. The approximate design implied small error probability which certainly is not valid for the low SNR case.

5.5. Discussion

In this chapter, modeling of digital channels is considered. In the past two different approaches to the channel modeling problem were proposed. In the present work, a methodology for a unified treatment of the channel modeling problem is described. The objective is to utilize the knowledge of the actual physical processes in characterizing the input-output behavior of the channel more accurately. The emphasis is only on describing the methodology and not on generating complex channel models. An attempt is made to characterize the channel memory quantitatively. Two measures are defined which quantitatively describe the clustering of errors in the channel error sequence. Relationship of receivers with memory and channel models is examined. Receivers with memory which utilize the correlation of the received data are expected to alter the clustering properties of the channel error sequence.

communication systems using a receiver with memory and without memory are compared by computing the correlation measures. The errors generated by a system using a receiver with memory are more predictable and occurrence of an error provides more information about the occurrence of further errors. The example considered illustrated these points.

6. SUMMARY AND CONCLUSIONS

In this thesis some aspects of digital communications over channels with memory have been studied. The basic objective was to model channels with memory and to develop more efficient and reliable communication techniques over such channels. In particular, the receiver design problem for channels with memory was considered and also the relationship of receivers and channel models is examined. In the receiver design problem, the source of channel memory was assumed to be fading. Receivers with memory were derived for applications with channels with memory. In Chapter two, a receiver with one-bit memory was considered. The optimal receiver was derived but it was too complex to implement and, therefore, a suboptimal decision-feedback receiver with one-bit memory was considered. The suboptimal receiver was shown to perform better than the optimal receiver without memory. The emphasis in the second chapter was on the investigation of the theoretical feasibility of receivers with memory and not on actually deriving optimum but complex receivers.

A receiver with large memory was considered in the third chapter. The receiver was assumed to consist of an estimator and a detector. The estimator provided the information about the existing fading conditions to the detector which adapted the decision rule accordingly. The memory length was assumed to be large and asymptotic results were obtained. The design criterion for the receiver was the minimization of the probability of error. The performance of the receiver was measured in terms of the average probability of error. Numerical results were obtained for a specific communication system and the performance of the constrained receiver with memory was compared to that of the optimal receiver without memory. The results indicated a significant improvement in the receiver performance when memory was introduced. In Chapter four, the estimator design was described. A limited-memory decision-feedback estimator was developed with applications both in communication and control theory. The performance of the minimum mean-squared error (MMSE) estimator was examined and numerical results were obtained. The results for estimator with certain observations indicated that the MSE attains its minimum for a memory length of four. The numerical results for the MSE with uncertain observations were also presented.

In Chapter five, the modeling of digital channels with memory was discussed. A unified treatment of the channel modeling problem was considered. Actual physical processes were employed to describe the input-output behavior of the channel. A methodology to be used for the development of such models was described and actual channel models were not derived. The role of actual communication systems in the clustering of errors in the channel error sequence was discussed. Two measures were defined which quantitatively characterize the correlation of errors. These measures are expected to have applications in the development of more efficient communication techniques. Finally, the effect of receivers on the correlation of errors was considered. The coefficients α_m and β_m are computed for the cases when the optimum receiver without memory was used and when the constrained receiver was employed. Numerical results obtained indicate that if the receiver with memory is employed in the communication system, the occurrence of an error provides more information about future errors. Thus, the occurrence of errors could be predicted more accurately.

Finally, the two major research problems which should follow the work presented here are indicated. The receiver design problem should be pursued in the presence of both fading and intersymbol interference. The unified treatment of the channel modeling problem should be considered and channel models, which take into account the actual physical processes involved while describing the input-output behavior of the channel should be developed.

APPENDIX A

Derivation of Equations from Section 2.5.1

The objective here in this appendix is to present derivations of some of the expressions obtained in Section 2.5.1. The threshold for the detection of the first bit, T_1, is obtained by minimizing the probability of error in the first bit, $P(e_1 = 1)$. Note that this is also the probability of error of the zero-memory receiver.

$$P(e_1 = 1) = \frac{1}{2} P\{\ell_1 > T_1 | H_1^0\} + \frac{1}{2} \{\ell_1 < T_1 | H_1^1\} \quad \text{(A.1)}$$

where ℓ_1 is the decision statistic defined in (2.21). If the optimal decision regions are denoted by Ω_0 and Ω_1, the above can be written as

$$\begin{aligned} P(e_1 = 1) &= \frac{1}{2} \iint_{\Omega_1} f_{X_1,Y_1|H_1^0}(x,y)dxdy \\ &+ \frac{1}{2} \iint_{\Omega_0} f_{X_1,Y_1|H_1^1}(x,y)dxdy \\ &= \frac{1}{2} \iint_{\Omega_1} (2\pi)^{-1} \exp\{-(x^2+y^2)/2\}dxdy \\ &+ \frac{1}{2} \iint_{\Omega_0} (2\pi(1+\eta))^{-1} \exp\{-(x^2+y^2)/2(1+\eta)\}dxdy \quad \text{(A.2)} \end{aligned}$$

which when changed into the polar coordinate system results in

$$P(e_1 = 1) = \frac{1}{2}\left[\int_{r=(T_1)^{\frac{1}{2}}}^{\infty} r\exp(-r^2/2)dr + \int_{r=0}^{(T_1)^{\frac{1}{2}}} r(1+\eta)^{-1}\exp(-r^2/2(1+\eta))dr\right]$$

$$= \frac{1}{2}\exp(-T_1/2) + \frac{1}{2}\{1-\exp(-T_1/2(1+\eta))\} \quad (A.3)$$

The value of the threshold T_1 which minimizes the right hand side of (A.3) is given by

$$\partial P(e_1 = 1)/\partial T_1 = -\frac{1}{4}\exp(-T_1/2) + \frac{1}{4}(1+\eta)^{-1}\exp(-T_1/2(1+\eta)) = 0 \quad (A.4)$$

This results in (2.24), i.e.,

$$T_1 = 2\eta^{-1}(1+\eta)\ell n(1+\eta) \quad (A.5)$$

The thresholds T_k for the optimal scheme are determined by (2.26) and are obtained so as to minimize the following conditional probability of error:

$$\begin{aligned} P(e_k = 1|X_{k-1},Y_{k-1}) &= \frac{1}{4}[P(\ell_k > T_k|X_{k-1},Y_{k-1},H_{00}) \\ &+ P(\ell_k > T_k|X_{k-1},Y_{k-1},H_{10}) + P(\ell_k < T_k|X_{k-1},Y_{k-1},H_{01}) \\ &+ P(\ell_k < T_k|X_{k-1},Y_{k-1},H_{11})] \\ &= \frac{1}{4}\left[\iint_{\Omega_1} f_{X_kY_k|X_{k-1},Y_{k-1},H_{00}}(x,y|u,v)dxdy + \iint_{\Omega_1} f_{X_kY_k|X_{k-1},Y_{k-1},H_{10}}(x,y|u,v)dxdy\right. \\ &+ \iint_{\Omega_0} f_{X_kY_k|X_{k-1},Y_{k-1},H_{01}}(x,y|u,v)dxdy \\ &\left.+ \iint_{\Omega_0} f_{X_kY_k|X_{k-1},Y_{k-1},H_{11}}(x,y|u,v)dxdy\right] \end{aligned}$$

$$= \frac{1}{4}\Big[\iint_{\Omega_1} (2\pi)^{-1}\exp(-(x^2+y^2)/2)dxdy + \iint_{\Omega_1} (2\pi)^{-1}\exp(-(x^2+y^2)/2)dxdy$$

$$+ \iint_{\Omega_0} (2\pi(1+\eta))^{-1}\exp(-(x^2+y^2)/2(1+\eta))dxdy$$

$$+ \iint_{\Omega_0} (2\pi(1+\eta)(1-\rho^2))^{-1}\exp\{-[(x-\rho u)^2+(y-\rho v)^2]/2(1+\eta)(1-\rho^2)\}dxdy\Big] \tag{A.6}$$

Denote the four integrals in (A.6) by I_j, $j=1,2,3,4$, in the order they are listed. The integrals may be written in polar coordinates as

$$I_1 = I_2 = \int_{r=(T_k)^{\frac{1}{2}}}^{\infty} r\exp(-r^2/2)dr = \exp(-T_k/2) \tag{A.7}$$

$$I_3 = \int_{r=0}^{(T_k)^{\frac{1}{2}}} r(1+\eta)^{-1}\exp\{-r^2/2(1+\eta)\}dr$$

$$= 1 - \exp\{-T_k/2(1+\eta)\} \tag{A.8}$$

$$I_4 = \int_{\phi=0}^{2\pi}\int_{r=0}^{(T_k)^{\frac{1}{2}}} r\{2\pi(1+\eta)(1-\rho^2)\}^{-1}\exp\{-[(r\cos\phi-\rho u)^2+(r\sin\phi-\rho v)^2]$$

$$/2(1+\eta)(1-\rho^2)\}drd\phi$$

$$= \int_{r=0}^{(T_k)^{\frac{1}{2}}} r\{(1+\eta)(1-\rho^2)\}^{-1}\exp\{-(r^2+\rho^2\ell_{k-1})/2(1+\eta)(1-\rho^2)\}$$

$$I_0\{\rho r(\ell_{k-1})^{\frac{1}{2}}/(1+\eta)(1-\rho^2)\}dr$$

$$= 1 - Q\{\rho(\ell_{k-1})^{\frac{1}{2}}[(1+\eta)(1-\rho)]^{-\frac{1}{2}}, T_k^{\frac{1}{2}}[(1+\eta)(1-\rho^2)]^{-\frac{1}{2}}\} \tag{A.9}$$

We substitute from (A.7) - (A.9) into (A.6) and set the derivative with respect to T_k equal to zero to obtain (2.26).

For the suboptimal scheme, (2.29) is employed to compute the threshold values R_2^0 and R_2^1. These expressions for R_2^0 and R_2^1 are obtained so as to minimize the conditional probabilities of error $P(e_2 = 1|Z_1 = 0)$ and $P(e_2 = 1|Z_1 = 1)$. Here the computation of R_2^0 is illustrated by minimizing $P(e_2 = 1|Z_1 = 0)$. R_2^1 can be obtained in an identical fashion.

$$P(e_2 = 1|Z_1 = 0) = \frac{1}{2} P(Z_2 = 1|Z_1 = 0, H_2^0) + \frac{1}{2} P(Z_2 = 0|Z_1 = 0, H_2^1)$$

$$= \frac{1}{4} [P(Z_1 = 0)]^{-1} [P(Z_2 = 1, Z_1 = 0|H_{00}) + P(Z_2 = 1, Z_1 = 0|H_{10})$$

$$+ P(Z_2 = 0, Z_1 = 0|H_{01}) + P(Z_2 = 0, Z_1 = 0\ |H_{01})] \quad \text{(A.10)}$$

$\frac{1}{4} [P(Z_1 = 0)]^{-1}$ is a constant as far as the computation of R_2^0 is concerned so it is denoted by C_0 and also R_2^0 is denoted by T_2 for convenience.

$$P(e_2 = 1|Z_1 = 0) = C_0[P\{\ell_2 > T_2, \ell_1 < T_1|H_{00}\} + P\{\ell_2 > T_2, \ell_1 < T_1|H_{10}\}$$

$$+ P\{\ell_2 < T_2, \ell_1 < T_1|H_{01}\} + P\{\ell_2 < T_2, \ell_1 < T_1|H_{11}\}]$$

$$= C_0[\{1-\exp(-T_1/2)\}\exp(-T_2/2) + \{1-\exp(-T_1/2(1+\eta))\}$$

$$\exp(-T_2/2) + \{1-\exp(-T_1/2)\}\{1-\exp(-T_2/2(1+\eta))\}$$

$$+ \int_{\ell_2=0}^{(T_2)^{\frac{1}{2}}} \ell_2(1+\eta)^{-1}\exp\{-\ell_2^2/2(1+\eta)\}(1-Q\{\rho\ell_2(1+\eta)^{-1}(1-\rho^2)^{-1},$$

$$T_1^{\frac{1}{2}}(1+\eta)^{-1}\ (1-\rho^2)^{-1})]d\ell_2 \quad \text{(A.11)}$$

The derivative of (A.11) is set to zero to obtain (2.29a). Equation (2.29b) can be derived in a similar manner.

APPENDIX B

Some Important Functions

In communication theory, some functions occur quite frequently. A brief summary of the functions used in this thesis is presented in this appendix. The first function is the error function which is defined as

$$\mathrm{erf}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{u^2}{2}\}du \tag{B.1}$$

and the complement of the error function is given by

$$\begin{aligned} \mathrm{erfc}(x) &= 1 - \mathrm{erf}(x) \\ &= \int_{x}^{\infty} \frac{}{\sqrt{2\pi}} \exp\{-\frac{u^2}{2}\}du \end{aligned} \tag{B.2}$$

The function has been tabulated in most of the standard books on mathematical tables. For digital computer applications a rational approximation to the error function can be used [77].

$$\mathrm{erf}(x) = 1 - (a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5) \exp(-x^2) + \epsilon(x) \tag{B.3}$$

where

$$t = (1 + px)^{-1}$$

$$p = .3275911$$

$$a_1 = .254829592, \quad a_2 = -.284496736$$

$$a_3 = 1.421413741, \quad a_4 = -1.453152027$$

$$a_5 = 1.061405429$$

and the error $\epsilon(x)$ satisfies the following bound

$$|\epsilon(x)| \leq 1.5 \times 10^{-7}$$

The other major function of importance is the Marcum's Q function [78] which is defined as

$$Q(\alpha,\beta) = \int_{\beta}^{\infty} Z \exp\left(-\frac{Z^2+\alpha^2}{2}\right) I_0(\alpha Z)dZ \tag{B.4}$$

where $I_0(\cdot)$ is a modified Bessel function of the first kind. The Q function has been tabulated by Marcum [79]. The integral cannot be evaluated analytically. An asymptotic approximation to the Q-function is given by

$$Q(\alpha,\beta) \simeq \frac{1}{2}\left\{\mathrm{erfc}\left(\frac{\beta-\alpha}{\sqrt{2}}\right) + \frac{\exp[-(\beta-\alpha)^2/2]}{\sqrt{2\pi\alpha\beta}}\right\} \tag{B.5}$$

if $\beta \gg 1$, $\alpha \gg 1$ and $\beta \gg \beta-\alpha \geq 0$. DiDonato and Jarnagin [80] have described an efficient numerical technique for the computation of the Q-function. Their results are presented next. They define two functions V(K,c) and P(R,D) as

$$V(K,c) = \frac{1}{c}\int_{o}^{K} \exp\left(-\frac{Br^2}{2}\right) I_0\left(\frac{Ar^2}{2}\right) r\,dr \tag{B.6}$$

and

$$P(R,D) = \exp(-D^2/2)\int_{o}^{R} \exp(-R^2/2)\; I_0(r\,D)r\,dr \tag{B.7}$$

These two functions are related by

$$P(R,D) = \frac{1}{2}\left[1-\exp\left(-\frac{R^2+D^2}{2}\right) I_0(R\ D) \pm V\left(|R-D|, \frac{|R-D|}{R+D}\right)\right]$$

$$+ \text{ if } R > D$$

$$- \text{ if } R < D \qquad \text{(B.8)}$$

Recursive schemes are utilized to obtain the value of P. If $2RD \leq M$ where M is a positive constant, the following set of relations are employed

$$T_{2n} = \left(\frac{2n-1}{2n}\right)\left(\frac{2RD}{R^2+D^2}\right)^2 T_{2n-2} - \frac{|R^2-D^2|}{R^2+D^2}\left(1 + \frac{4n}{R^2+D^2}\right)\bar{S}_{2n} \qquad n \geq 1 \qquad \text{(B.9)}$$

and

$$\bar{S}_{2n} = \left(\frac{RD}{2n}\right)^2 \bar{S}_{2n-2} \qquad n \geq 1 \qquad \text{(B.10)}$$

The initial terms are given by

$$S_0 = \exp(-(R^2+D^2)/2) \qquad \text{(B.11)}$$

and

$$T_0 = \frac{|R^2-D^2|}{R^2+D^2}(1-\bar{S}_0) \qquad \text{(B.12)}$$

The recursive scheme is continued until

$$T_{2n} < \varepsilon \text{ and } \bar{S}_{2n} < \varepsilon \qquad (\varepsilon > 0)$$

P(R,D) is given by

$$P(R,D) \simeq \frac{1}{2}\left[1 - \sum_{n=0}^{N'} \bar{S}_{2n} \pm \sum_{n=0}^{N'} T_{2n}\right] \quad \begin{array}{l} (+) \text{ if } R > D \\ (-) \text{ if } R < D \end{array} \tag{B.13}$$

The above value is given correctly to at least $(|\log_{10}\epsilon|-1)$ decimal digits.

APPENDIX C

Derivation of Equations from Section 2.5.2

In this appendix, some of the equations in the Section 2.5.2 are derived. The threshold for the detection of the first bit, T_1, is computed by minimizing the probability of error in the first bit.

$$P(e_1=1) = \frac{1}{2} P(Y_1 > T_1|H_1^0) + \frac{1}{2} P(Y_1 < T_1|H_1^1)$$

$$= \frac{1}{2} \int_{T_1}^{\infty} (2\pi)^{-\frac{1}{2}} \exp(-y^2/2)dy + \frac{1}{2} \int_{-\infty}^{T_1} (2\pi(1+\eta))^{-\frac{1}{2}} \exp(-y^2/2(1+\eta))dy \tag{C.1}$$

When the derivative of the right hand side of (C.1) is set to zero, it results in (2.35), i.e.,

$$-\frac{1}{2} (2\pi)^{-\frac{1}{2}} \exp(-T_1^2/2) + \frac{1}{2} (2\pi(1+\eta))^{-\frac{1}{2}} \exp(-T_1^2/2(1+\eta)) = 0 \tag{C.2}$$

and this implies

$$T_1 = \{(1+\eta)\ell n(1+\eta)\}^{\frac{1}{2}} \eta^{-\frac{1}{2}} \tag{C.3}$$

The thresholds T_k for the optimum decision rule are obtained by minimizing the conditional probability of error

$$P(e_k=1|Y_{k-1}=v) = \frac{1}{4} [P(Y_k > T_k|Y_{k-1},H_{00}) + P(Y_k > T_k|Y_{k-1},H_{10})$$

$$+ P(Y_k < T_k|Y_{k-1},H_{01}) + P(Y_k < T_k|Y_{k-1},H_{11})]$$

$$= \frac{1}{4} [2 \cdot \int_{T_k}^{\infty} (2\pi)^{-\frac{1}{2}} \exp(-y^2/2)dy + \int_{-\infty}^{T_k} (2\pi(1+\eta))^{-\frac{1}{2}} \exp(-y^2/2(1+\eta))dy$$

$$+ \int_{-\infty}^{T_k} (2\pi(1+\eta)(1-\rho^2))^{-\frac{1}{2}} \exp(-(y-\rho v)^2/2(1+\eta)(1-\rho^2))dy] \tag{C.4}$$

The derivative of (C.4) is set to zero and it yields (2.36). To obtain (2.38) and (2.39), the expressions for the conditional probability of error $P(e_2=1|Z_1=0)$ and $P(e_2=1|Z_1=1)$ are computed and the derivatives are set to zero. Here (2.38) is obtained and (2.39) can be obtained similarly.

$$
\begin{aligned}
P(e_2=1|Z_1=0) &= \frac{1}{2} P(Z_2=1|Z_1=0,H_2^0) + \frac{1}{2} P(Z_2=0|Z_1=0,H_2^1) \\
&= \frac{1}{4P(Z_1=0)} [P(Z_2=1,Z_1=0|H_{00}) + P(Z_2=1,Z_1=0|H_{10}) \\
&\quad + P(Z_2=0,Z_1=0|H_{01}) + P(Z_2=0,Z_1=0|H_{11})] \\
&= C_0[4\,\mathrm{erf}(T_1)\mathrm{erfc}(Q_2) + 4\,\mathrm{erf}(T_1(1+\eta)^{-\frac{1}{2}})\mathrm{erfc}(Q_2) \\
&\quad + 4\,\mathrm{erf}(Q_2/(1+\eta)^{\frac{1}{2}})\mathrm{erf}(T_1) + 2(2\pi(1+\eta))^{-\frac{1}{2}} \int_{\beta=0}^{Q} \exp\{-\beta^2/2(1+\eta)\} \\
&\quad \{\mathrm{erf}((T-\rho\beta)(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}) + \mathrm{erf}((T+\rho\beta)(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}})\}\,d\beta
\end{aligned}
\tag{C.5}
$$

Setting the derivative of (C.5) to zero results in (2.38). The equations which yield the thresholds Q_k and R_k are similar to (2.38) and (2.39) and are obtained in a similar fashion as outlined above. The steady state solutions Q and R are obtained by setting $Q_{k+1} = Q_k = Q$ and $R_{k+1} = R_k = R$. Approximate solutions are obtained by linearizing the equations. It is assumed that the deviations ΔQ and ΔR are small. The nonlinear functions are expanded in Taylor series and only the linear terms are kept, e.g.,

$$
\mathrm{erf}(\alpha(T-\rho\beta)) \simeq \mathrm{erf}(\alpha T) - \frac{\rho\alpha\beta}{\sqrt{2\pi}} \exp(-\alpha^2 T^2/2) \tag{C.6}
$$

and

$$\exp\{-\eta(2(1+\eta))^{-1}T^2(1+\frac{\Delta Q}{T})^2\}$$

$$\simeq \exp\{-\eta(2(1+\eta))^{-1}T^2(1+\frac{2\Delta Q}{T})\}$$

$$\simeq \{1-\eta T\Delta Q(1+\eta)^{-1}\}\exp\{-\eta T^2(2(1+\eta))^{-1}\} \qquad (C.7)$$

Approximate values of terms which are similar to the terms shown in (C.6) and (C.7) are substituted in the exact set of equations for Q and R, and it results in the set of simultaneous linear equations (2.42) which are solved for the threshold values.

APPENDIX D

Illustration of Theorem 2.1

In this appendix, theorem 2.1 is verified for the Gaussian fading example. For convenience in illustration, the probability of error in the second bit using both the schemes is computed and compared. The expressions for the probability of error for any arbitrary bit will be similar but a little more complicated. When $r = 0$, both the thresholds Q and R are equal to the no memory threshold T and $P(e)_{nm}$ is given by

$$P(e)_{nm} = (a+b)\mathrm{erfc}(T) + (c+d)\mathrm{erfc}(T) + c\ \mathrm{erf}(T(1+\eta)^{-\frac{1}{2}}) + (a+\tfrac{1}{2})\mathrm{erf}(T(1+\eta)^{-\frac{1}{2}})$$

$$= \mathrm{erfc}(T) + \mathrm{erf}(T(1+\eta)^{-\frac{1}{2}}) = P(e)_{nm} \tag{D.1}$$

Thus, the equality is attained for $r = 0$. Now, it is shown that even for small r, $P(e)_m < P(e)_{nm}$. It will then intuitively follow that the inequality holds for large r.

$$P(e_2)_m = P\{(e_2=1)\cap(Z_1=0)\} + P\{(e_2=1)\cap(Z_1=1)\} \tag{D.2}$$

where

$$P\{(e_2=1)\cap(Z_1=0)\} = \mathrm{erf}(T)\mathrm{erfc}(Q) + \mathrm{erf}(T(1+\eta)^{-\frac{1}{2}})\mathrm{erfc}(Q) + \mathrm{erf}(Q(1+\eta)^{-\frac{1}{2}})\mathrm{erf}(T) + \frac{1}{2}(2\pi(1+\eta))^{-\frac{1}{2}}\int_{\beta=0}^{Q}\exp(-\beta^2/2(1+\eta))$$

$$\{\mathrm{erf}((T-\rho\beta)(1+\eta)^{-\frac{1}{2}}(1-\rho^2))^{-\frac{1}{2}} + \mathrm{erf}((T+\rho\beta)(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}})\} \tag{D.3}$$

and

$$P\{(e_2=1)\cap(Z_1=1)\} = \mathrm{erfc}(R)\mathrm{erfc}(T) + \mathrm{erfc}(R)\mathrm{erfc}(T(1+\eta)^{-\frac{1}{2}})$$
$$+ \mathrm{erfc}(T)\mathrm{erf}(R(1+\eta)^{-\frac{1}{2}}) + \frac{1}{2}(2\pi(1+\eta))^{-\frac{1}{2}}$$
$$\int_{\beta=-R}^{R} \exp\{-\beta^2(2(1+\eta))^{-1}\}\mathrm{erfc}\{(T-\rho\beta)(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\}d\beta \tag{D.4}$$

For small r and, therefore, for small ρ, linear approximations for the error functions are employed in (D.4) and (D.5).

$$\mathrm{erf}\{(T-\rho\beta)(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\} = \mathrm{erf}\{T(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\}$$
$$- \rho\beta(2\pi(1+\eta)(1-\rho^2))^{-\frac{1}{2}}\exp\{-T^2/2(1+\eta)(1-\rho^2)\} \tag{D.6}$$

$$\mathrm{erf}\{(T+\rho\beta)(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\} = \mathrm{erf}\{T(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\}$$
$$+ \rho\beta(2\pi(1+\eta)(1-\rho^2))^{-\frac{1}{2}}\exp\{-T^2/2(1+\eta)(1-\rho^2)\} \tag{D.7}$$

$$\mathrm{erfc}\{(T-\rho\beta)(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\} = \mathrm{erfc}\{T(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}}\}$$
$$+ \rho\beta(2\pi(1+\eta)(1-\rho^2))^{-\frac{1}{2}}\exp\{-T^2/2(1+\eta)(1-\rho^2)\} \tag{D.8}$$

For small ρ, therefore (D.3) and (D.4) reduce to

$$P\{(e_2=1)\cap(Z_1=0)\} = \{\mathrm{erf}(T) + \mathrm{erf}(T(1+\eta)^{-\frac{1}{2}})\mathrm{erfc}(Q)$$
$$+ \{\mathrm{erf}(T) + \mathrm{erf}(T(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}})\}\mathrm{erf}(Q(1+\eta)^{-\frac{1}{2}}) \tag{D.9}$$

$$P\{(e_2=1)\cap(Z_1=1)\} = \{\mathrm{erfc}(T) + \mathrm{erfc}(T(1+\eta)^{-\frac{1}{2}})\}\mathrm{erfc}(R)$$
$$+ \{\mathrm{erfc}(T) + \mathrm{erfc}(T(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}})\}\mathrm{erf}\{R(1+\eta)^{-\frac{1}{2}}\} \tag{D.10}$$

Therefore,

$$P(e_2)_m = \beta_1 \operatorname{erfc}(Q) + \beta_2 \operatorname{erf}(Q(1+\eta)^{-\frac{1}{2}})$$

$$+ (1-\beta_1)\operatorname{erfc}(R) + (1-\beta_2)\operatorname{erf}(R(1+\eta)^{-\frac{1}{2}}) \qquad (D.11)$$

where

$$\beta_1 = \operatorname{erf}(T) + \operatorname{erf}(T(1+\eta)^{-\frac{1}{2}})$$

$$\beta_2 = \operatorname{erf}(T) + \operatorname{erf}(T(1+\eta)^{-\frac{1}{2}}(1-\rho^2)^{-\frac{1}{2}})$$

Utilizing the fact that

$$Q = T + \Delta Q$$

$$R = T + \Delta R$$

and that ΔQ and ΔR are small, the following approximations are obtained.

$$\operatorname{erfc}(Q) = \operatorname{erfc}(T) - \Delta Q(2\pi)^{-\frac{1}{2}}\exp(-T^2/2) \qquad (D.12)$$

$$\operatorname{erfc}(R) = \operatorname{erfc}(T) - \Delta R(2\pi)^{-\frac{1}{2}}\exp(-T^2/2) \qquad (D.13)$$

$$\operatorname{erf}(Q(1+\eta)^{-\frac{1}{2}}) = \operatorname{erf}(T(1+\eta)^{-\frac{1}{2}}) + \Delta Q(2\pi(1+\eta))^{-\frac{1}{2}}\exp(-T^2/2(1+\eta)) \qquad (D.14)$$

$$\operatorname{erf}(R(1+\eta)^{-\frac{1}{2}}) = \operatorname{erf}(T(1+\eta)^{-\frac{1}{2}}) + \Delta R(2\pi(1+\eta))^{-\frac{1}{2}}\exp(-T^2/2(1+\eta)) \qquad (D.15)$$

Substituting these approximate values into (D.11), the result is

$$P(e_2)_m = \operatorname{erfc}(T) - (2\pi)^{-\frac{1}{2}}\{\beta_1 \Delta Q + (1-\beta_1)\Delta R\}\exp(-T^2/2)$$

$$+ \operatorname{erf}(T\ (1+\eta)^{-\frac{1}{2}}) + (2\pi(1+\eta))^{-\frac{1}{2}}\{\beta_2 \Delta Q + (1-\beta_2)\Delta R\}$$

$$\exp(-T^2/2(1+\eta))$$

$$= P(e_2)_{nm} + (2\pi)^{-\frac{1}{2}}\exp(-T^2/2)\{(\beta_2-\beta_1)\Delta Q - (\beta_2-\beta_1)\Delta R\} \qquad (D.16)$$

Since

$$P(e_2)_{nm} = \mathrm{erfc}(T) + \mathrm{erf}(T(1+\eta)^{-\frac{1}{2}})$$

and

$$(2\pi)^{-\frac{1}{2}} \exp(-T^2/2) = (2\pi(1+\eta))^{-\frac{1}{2}} \exp(-T^2/2(1+\eta))$$

we have

$$P(e_2)_{nm} - P(e_2)_m = (2\pi)^{-\frac{1}{2}} \exp(-T^2/2)(\beta_2-\beta_1)(\Delta R-\Delta Q) \qquad \text{(D.17)}$$

It was observed earlier that $\beta_2 > \beta_1$ and also $\Delta R > 0$, $\Delta Q < 0$ so that the right hand side of (D.17) is positive and thus

$$P(e_2)_{nm} > P(e_2)_m \quad .$$

APPENDIX E

Central Limit Theorem for Dependent Random Variables

The central limit theorem for indpendent random variables has been treated extensively in the literature. It has been shown that the central limit theorem holds for dependent random variables also under certain conditions [51-53]. In this appendix,a brief exposition to the theorem and the conditions under which it holds is presented. The material here follows the discussion on the subject in [52] very closely but is presented here for completeness.

The sequence of dependent random variables is assumed to be Markov. Let Ω be a space of points x representing the possible observations at any given fixed time. The possible events for which a probability is well defined are the elements of a σ-field $\mathcal{A}$ of subsets of Ω. The transition probability function of the Markov process is denoted by P(x,A). It is assumed to be $\mathcal{A}$-measurable and is defined as a function of x for each event A in $\mathcal{A}$ and a probability measure on the σ-field $\mathcal{A}$ for each x in Ω. A probability measure P_μ can be defined to describe the relative likelihood of observing the different possible trajectories $\omega = (x_0, x_1, \dots)$ of the random system being studied through time. The observation on the system at time n is given by the n-th coordinate function or random variable $X_n(\omega) = x_n$ and the random process is written as $\{X_n\} = \{X_n(\omega);\ n=0,1,\ \dots\}$. This random process $\{X_n\}$ has been assumed to be Markov above. The σ-field generated by the sets of the form $\overset{\infty}{\underset{t=0}{X}} A_t$ where $A_t \in \mathcal{A}$ is denoted by $\mathcal{A}_\infty$. $\Omega_\infty = \overset{\infty}{\underset{t=0}{X}} \Omega_t$, $\Omega_t = \Omega$. A shift transformation τ corresponding to a forward time shift for

$\omega = (x_0, \ldots x_n)$ is defined by

$$(\tau\omega)_n = x_{n+1} \tag{E.1}$$

If for each event $C \in \mathcal{A}_\infty$

$$P_\mu(\tau C) = P_\mu(C) \tag{E.2}$$

then the Markov process is called stationary. The σ-field $\mathcal{A}_\infty$ was constructed so that it is exactly the σ-field generated by the random variables $\{X_n\}$. The σ-field generated by a finite number of random variavles X_k, $m \leq k \leq n$, is denoted by $\mathcal{A}_m^n$.

Suppose $Y_n(\omega)$, $n = 0,1,\ldots$ is a sequence of real-valued random variables on the probability space of the Markov process $\{X_n\}$. The series $\{Y_n\}$ is called time-consistent if $Y_n(\omega)$ is measurable with respect to $\mathcal{A}_n^n$, $n = 0,1,\ldots$ and stationary if $Y_n(\tau\omega) = Y_{n+1}(\omega)$, $n = 0,1,\ldots$ where τ is the shift transformation defined above.

The conditions to be imposed on the stationary Markov sequence $\{X_n\}$ are considered next. The following proposition from ergodic theory is stated without proof. The proof is given in [81].

Proposition: Let μ be a probability measure on the σ-field $\mathcal{A}$ on Ω. If ϕ is a measure-preserving mapping of $\mathcal{A}$ onto itself, then ϕ is ergodic iff

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mu(A \cap \phi^{-j}B) = \mu(A)\mu(B) \tag{E.3}$$

for each pair of sets $A, B \in \mathcal{A}$.

In the case under consideration here, P is the probability measure and the measure-preserving shift transformation τ is ergodic. Let us define

$$a(n) = \sup_{\substack{B \in \mathcal{B}_0 \\ F \in \mathcal{F}_0}} \left| \frac{1}{n} \sum_{k=1}^{n} P(B \cap \tau^k F) - P(B)P(F) \right| \tag{E.4}$$

where $\mathcal{B}_0 = \mathcal{B}\{X_j, j \leq 0\}$ is the σ-field generated by X_j, $j \leq 0$ and $\mathcal{F}_0 = \mathcal{B}\{X_j, j \geq 0\}$ is the σ-field generated by X_j, $j \geq 0$. The stationary Markov sequence $\{X_n\}$ is called uniformly ergodic if $a(n) \to 0$ as $n \to \infty$.

A property of the Markov sequence $\{X_n\}$ is now discussed. This property is stronger than what is actually needed for the work in this thesis. Define

$$d(n) = \sup_{\substack{B \in \mathcal{B}_0 \\ F \in \mathcal{F}_n}} |P(B \cap F) - P(B)P(F)| \tag{E.5}$$

The Markov process is called strongly mixing if $d(n) \to 0$ as $n \to \infty$. This term strongly mixing as defined here is different and somewhat weaker [52] than the definition given in the standard literature on ergodic theory, e.g., Billingsley [81] and Arnold and Avez [82].

Let $\{Y_k^{(n)}, k=0,1 \ldots\}$ be a time consistent series. Such a series is called uniformly asymptotically negligible if for each $\epsilon > 0$

$$\lim_{n \to \infty} \sup_k P[|Y_k^{(n)}(\omega)| > \epsilon] = 0 \tag{E.6}$$

Define
$$Y^{(n)}(s,t) = \sum_{k=s}^{t} Y_k^{(n)} \quad . \tag{E.7}$$

The stationary Markov process $\{X_k\}$ is said to have <u>central structure</u> if for any uniformly asymptotically negligible stationary sequence $\{Y_k^{'(n)}\}$ of time consistent series, any partial sums $Y^{(n)}(s_n,t_n)$, $-\infty < s_n < t_n < \infty$, are well approximated in distribution by a limiting distribution. The following theorem is stated without proof.

<u>Theorem</u>: Let $\{X_n\}$ be a stationary Markov process. The process has central structure iff it is uniformly ergodic.

Finally, the central limit theorem is stated.

<u>Theorem</u>: Assume that $\{X_k\}$ is a stationary Markov process. Assume that $E\{X_k\} = 0$ and the following conditions are satisfied

(i) $E\{|Y^{(n)}(s_n,t_n)|^2\} \sim h(t_n-s_n)$

as $(t_n-s_n) \to \infty$, where $h(\ell) \to \infty$ as $\ell \to \infty$

(ii) $E\{|Y^{(n)}(s_n,t_n)|^{2+\delta}\} = o(h(t_n-s_n)^{1+\delta/2})$

as $(t_n-s_n) \to \infty$ for some $\delta > 0$.

(iii) $\{X_k\}$ is uniformly ergodic.

Then $Y^{(n)}(S_n,t_n)$ is asymptotically normally distributed.

The condition (iii) can be replaced by the strong mixing condition in the statement of the above theorem. As noted earlier strong mixing is a more stringent condition than the uniform ergodicity.

The stationary Markov process $\{X_k\}$ is said to have <u>central structure</u> if for any uniformly asymptotically negligible stationary sequence $\{Y_k^{(n)}\}$ of time consistent series, any partial sums $Y^{(n)}(s_n,t_n)$, $-\infty < s_n < t_n < \infty$, are well approximated in distribution by a limiting distribution. The following theorem is stated without proof.

<u>Theorem</u>: Let $\{X_n\}$ be a stationary Markov process. The process has central structure iff it is uniformly ergodic.

Finally, the central limit theorem is stated.

<u>Theorem</u>: Assume that $\{X_k\}$ is a stationary Markov process. Assume that $E\{X_k\} = 0$ and the following conditions are satisfied

(i) $E\{|Y^{(n)}(s_n,t_n)|^2\} \sim h(t_n-s_n)$

as $(t_n-s_n) \to \infty$, where $h(\ell) \to \infty$ as $\ell \to \infty$

(ii) $E\{|Y^{(n)}(s_n,t_n)|^{2+\delta}\} = o(h(t_n-s_n)^{1+\delta/2})$

as $(t_n-s_n) \to \infty$ for some $\delta > 0$.

(iii) $\{X_k\}$ is uniformly ergodic.

Then $Y^{(n)}(S_n,t_n)$ is asymptotically normally distributed.

The condition (iii) can be replaced by the strong mixing condition in the statement of the above theorem. As noted earlier strong mixing is a more stringent condition than the uniform ergodicity.

REFERENCES

1. R. W. Lucky, "A Survey of the Communication Theory Literature: 1968-1973," IEEE Trans. Inform. Th., Vol. IT-19, No. 6, pp. 725-739, Nov. 1973.

2. R. W. Lucky, "Automatic Equalization for Digital Communication," Bell Syst. Tech. J., Vol. 44, No. 4, pp. 547-588, April 1965.

3. R. W. Lucky, "Techniques for Adaptive Equalization of Digitial Communication Systems," Bell Syst. Tech. J., Vol. 45, No. 2, pp. 255-286, Feb. 1966.

4. J. G. Proakis and J. H. Miller, "An Adaptive Receiver for Digital Signalling Through Channels with Intersymbol Interference," IEEE Trans. Inform. Th., Vol. IT-15, No. 4, pp. 484-497, July 1969.

5. A. Gersho, "Adaptive Equalization of Highly Dispersive Channels for Data Transmission," Bell Syst. Tech. J., Vol. 48, No. 1, pp. 55-70, Jan. 1969.

6. T. J. Schonfeld and M. Schwartz, "A Rapidly Converging First-Order Training Algorithm for an Adaptive Equalizer," IEEE Trans. Inform. Th., Vol. IT-17, pp. 431-439, July 1971.

7. T. J. Schonfeld and M. Schwartz, "Rapidly Converging Second-Order Tracking Algorithms for Adaptive Equalization," IEEE Trans. Inform. Th., Vol. IT-17, pp. 572-579, Sept. 1971.

8. S. H. Richman and M. Schwartz, "Dynamic Programming Training Period for an MSE Adaptive Equalizer," IEEE Trans. Commun. Tech., Vol. COM-20, pp. 857-864, Oct. 1972.

9. T. Walzman and M. Schwartz, "Automatic Equalization Using the Discrete Frequency Domain," IEEE Trans. Inform. Th., Vol. IT-19, pp. 59-68, Jan. 1973.

10. K. H. Mueller and D. A. Spaulding, "Cyclic Equalization - A New, Rapidly Converging Equalization Technique for Synchronous Data Communication," Bell. Syst. Tech. J., Vol. 54, pp. 369-406, Feb. 1975.

11. R. D. Gitlin, E. Y. Ho, and J. E. Mazo, "Passband Equalization of Differentially Phase-Modulated Data Signals," Bell Syst. Tech. J., Vol. 52, No. 2, pp. 219-238, Feb. 1973.

12. R. Ramachandran and W. Steenart, "Optimal Equalization of Discrete Signals Passed Through a Random Channel," IEEE Trans. Comm. Tech., Vol. COM-20, pp. 884-890, Oct. 1972.

13. Y. S. Cho, "Optimal Equalization of Wideband Coaxial Cable Channels Using 'Bump' Equalizers," Bell Syst. Tech. J., Vol. 51, No. 6, pp. 1327-1345, July-Aug. 1972.

14. H. Rudin, "A Continuously Adaptive Equalizer for General-Purpose Communication Channels," Bell Syst. Tech. J., Vol. 48, No. 6, pp. 1865-1884, July-Aug. 1969.

15. R. E. Lawrence and H. Kaufman, "The Kalman Filter for the Equalization of a Digital Communication Channel," IEEE Trans. Commun. Tech., Vol. COM-19, pp. 1137-1141, Dec. 1971.

16. R. W. Chang, "Joint Optimization of Automatic Equalization and Carrier Acquisition for Digital Communications," Bell Syst. Tech. J., Vol. 49, No. 6, pp. 1069-1104, July-Aug. 1970.

17. M. J. DiToro, "Communication in Time-Frequency Spread Media Using Adaptive Equalization," Proc. IEEE, Vol. 56, pp. 1653-1679, Oct. 1968.

18. N. J. Bershad and P. A. Vena, "Eliminating Intersymbol Interference - A State-Space Approach," IEEE Trans. Inform. Th., Vol. IT-18, pp. 275-281, March 1972.

19. D. P. Taylor, "The Estimate Feedback Equalizer: A Suboptimum Nonlinear Receiver," IEEE Trans. Commun. Tech., Vol. COM-21, pp. 979-990, Sept. 1973.

20. G. Ungerboeck, "Nonlinear Equalization of Binary Signals in Gaussian Noise," IEEE Trans. Commun. Tech., Vol. COM-19, pp. 1128-1137, Dec. 1971.

21. S. U. H. Qureshi and E. E. Newhall, "An Adaptive Receiver for Data Transmission over Time-Dispersive Channels," IEEE Trans. Inform. Th., Vol. IT-19, pp. 448-457, July 1973.

22. M. E. Austin, "Decision-Feedback Equalization for Digital Communication Over Dispersive Channels," M.I.T. Lincoln Lab., Lexington, Mass. Tech. Rep. 437, Aug. 1967.

23. P. Monsen, "Feedback Equalization for Fading Dispersive Channels," IEEE Trans. Inform. Th., Vol. IT-17, pp. 56-65, Jan. 1971.

24. P. Monsen, "Adaptive Equalization of the Slow Fading Channel," IEEE Trans. Commun., Vol. COM-22, pp. 1064-1075, Aug. 1974.

25. R. Price, "Nonlinearly Feedback - Equalized PAM versus Capacity for Noisy Filter Channels," in Conf. Rec. Int. Conf. Comm., Sec. 22, pp. 12-16, Philadelphia, Pa. 1972.

26. J. Salz, "Optimum Mean-Square Decision Feedback Equalization," Bell Syst. Tech. J., Vol. 52, pp. 1341-1373, Oct. 1973.

27. A. P. Clark, "Adaptive Detection with Intersymbol Interference Cancellation for Distorted Digital Signals," IEEE Trans. Commun. Tech., Vol. COM-20, pp. 350-361, June 1972.

28. D. A. George, R. R. Bowen, and J. R. Storey, "An Adaptive Decision-Feedback Equalizer," IEEE Trans. Commun. Tech., Vol. COM-19, pp. 281-293, June 1971.

29. R. W. Chang and J. C. Hancock, "On Receiver Structures for Channels Having Memory," IEEE Trans. Inform. Th., Vol. IT-12, pp. 463-468, Oct. 1966.

30. K. Abend and B. D. Fritchman, "Statistical Detection for Communication Channels with Intersymbol Interference," Proc. IEEE, Vol. 58, pp. 779-785, May 1970.

31. R. A. Gonsalves, "Maximum-Likelihood Receiver for Digital Data Transmission," IEEE Trans. Commun. Tech., Vol. COM-16, pp. 392-398, June 1968.

32. H. Kobayashi, "Correlative Level Coding and Maximum-Likelihood Decoding," IEEE Trans. Inform. Th., Vol. IT-17, pp. 586-594, Sept. 1971.

33. J. K. Omura, "Optimal Receiver Design for Convolutional Codes and Channels with Memory via Control Theoretical Concepts," Inform. Sci., Vol. 3, pp. 243-266, July 1971.

34. G. D. Forney, Jr., "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference, Vol. IT-18, pp. 363-378, May 1972.

35. F. R. Magee, Jr. and J. G. Proakis, "Adaptive Maximum-Likelihood Sequence Estimation for Digital Signalling in the Presence of Intersymbol Interference," IEEE Trans. Inform. Th., Vol. IT-19, pp. 120-124, Jan. 1973.

36. M. J. Ferguson, "Optimal Reception for Binary Partial Response Channels," Bell Syst. Tech. J., Vol. 51, No. 2, pp. 493-505, Feb. 1972.

37. G. Ungerboeck, "Adaptive Maximum-Likelihood Receiver for Carrier-Modulated Data-Transmission Systems," IEEE Trans. Commun., Vol. COM-22, pp. 624-636, May 1974.

38. K. Yao and L. B. Milstein, "On ML Bit Detection of Binary Signals with Intersymbol Interference in Gaussian Noise," IEEE Trans. on Commun., Vol. COM-23, pp. 971-976, Sept. 1975.

39. T. M. Cover, "Hypothesis Testing with Finite Statistics," Ann. Math. Statist., Vol. 40, pp. 828-835, 1969.

40. M. E. Hellman and T. M. Cover, "Learning with Finite Memory," Ann. Math. Statist., Vol. 41, pp. 765-782, 1970.

41. M. E. Hellman and T. M. Cover, "The Two-Armed-Bandit Problem with Time-Invariant Finite Memory," IEEE Trans. Inform. Th., Vol. IT-16, pp. 185-195, March 1970.

42. E. G. Baxa and L. W. Nolte, "Adaptive Signal Detection with Finite Memory," IEEE Trans. Syst., Man, Cybern., Vol. SMC-2, pp. 42-49, Jan. 1972.

43. L. D. Davisson and S. C. Schwartz,"Analysis of a Decision-Directed Receiver with Unknown Priors," IEEE Trans. Inform. Th., Vol. IT-16, No. 3, pp. 270-276, May 1970.

44. A. Papoulis, Probability, Random Variables and Stochastic Processes, p. 194-195, McGraw Hill, New York, 1965.

45. R. Price, "Optimum Detection of Random Signals in Noise with Applications to Scatter-Multipath Communication, Part I," IRE Trans. Inform. Th., Vol. IT-2, pp. 125-135, April 1956.

46. T. Kailath, "Correlation Detection of Signals Perturbed by a Random Channel," IRE Trans. Inform. Th., Vol. IT-6, pp. 361-366, June 1960.

47. D. Middleton and R. Esposito, "Simultaneous Optimum Detection and Estimation of Signals in Noise," IEEE Trans. Inform. Th., Vol. IT-14, pp. 434-444, May 1968.

48. D. W. Kelsey and A. H. Haddad, "Detection and Prediction of a Stochastic Process Having Multiple Hypotheses," Information Sciences, Vol. 6, pp. 301-311, 1973.

49. D. G. Lainiotis and S. K. Park, "On Joint Detection Estimation and System Identification: Discrete Data Case," Int. J. of Control, Vol. 17, No. 3, pp. 609-633, 1973.

50. A. H. Jazwinski, "Limited Memory Optimal Filtering," IEEE Trans. Autom. Control, Vol. AC-13, No. 5, pp. 558-563, Oct. 1968.

51. M. Rosenblatt, "A Central Limit Theorem and a Strong Mixing Condition," Proc. Nat. Acad. Sciences, Vol. 42, pp. 43-47, 1956.

52. M. Rosenblatt, Markov Processes, Structure and Asymptotic Behavior, Springer-Verlag, 1971.

53. I. A. Ibragimov, "Some Limit Theorems for Stationary Processes," Theory Prob. Appl. (USSR), Vol. 7, pp. 349-383, 1962.

54. M. Kanefsky and J. B. Thomas, "On Adaptive Nonparametric Detection Systems Using Dependent Samples," IEEE Trans. Inform. Th., Vol. IT-11, No. 4, pp. 521-526.

55. B. R. Saltzberg, "Intersymbol Interference Error Bounds with Application to Ideal Bandlimited Signaling," IEEE Trans. Inform. Th., Vol. IT-14, pp. 563-568, July 1968.

56. R. Lugannani, "Intersymbol Interference and Probability of Error in Digital Systems," IEEE Trans. Inform. Th., Vol. IT-15, pp. 682-688, Nov. 1969.

57. E. Y. Ho and Y. S. Yeh, "A New Approach for Evaluating the Error Probability in the Presence of Intersymbol Interference and Additive Gaussian Noise," Bell Syst. Tech. J., Vol. 49, No. 9, pp. 2249-2265, Nov. 1970.

58. O. Shimbo and M. I. Celebiler, "The Probability of Error Due to Intersymbol Interference and Gaussian Noise in Digital Communication Systems," IEEE Trans. Commun. Tech., Vol. COM-19, pp. 113-120, April 1971.

59. R. Esposito, "A Class of Estimators for Optimum Adaptive Detection," Information and Control, Vol. 10, pp. 137-148, 1967.

60. N. E. Nahi, "Optimal Recursive Estimation with Uncertain Observation," IEEE Trans. Inform. Th., Vol. IT-15, No. 4, pp. 457-462, July 1969.

61. R. N. Jackson and D. N. P. Murthy, "Optimal Linear Estimation with Uncertain Observations," IEEE Trans. Inform. Th., Vol. IT-22, No. 3, pp. 376-378, May 1976.

62. Y. Sawaragi, T. Katayama, and S. Fujishige, "Sequential State Estimation with Interrupted Observation," Information and Control, Vol. 21, pp. 56-71, 1972.

63. P. A. Bello, "A Troposcatter Channel Model," IEEE Trans. Commun. Tech., Vol. COM-17, pp. 130-137, April 1969.

64. P. A. Bello and T. H. Crystal, "A Class of Efficient High-Speed Digital Modems for Troposcatter Links," IEEE Trans. Commun. Tech., Vol. COM-17, pp. 162-183, April 1969.

65. P. A. Bello and L. Ehrman, "Error Rates in Diversity FDM-FM Digital Troposcatter Transmission," IEEE Trans. Commun, Tech., Vol. COM-17, pp. 183-191, April 1969.

66. D. Arnstein, "Correlated Error Statistics on Troposcatter Channels," IEEE Trans. Commun. Tech., Vol. COM-19, pp. 225-228, April 1971.

67. J. M. Berger and B. Mandelbrot, "A New Model for Error Clustering in Telephone Circuits," IBM J. Res. Dev., Vol. 7, pp. 224-236, July 1963.

68. E. O. Elliott, "A Model of the Switched Telephone Network for Data Communications," Bell Syst. Tech. J., Vol. 44, pp. 89-109, Jan. 1965.

69. N. Muntner and J. K. Wolf, "Predicted Performances of Error-Control Techniques over Real Channels," IEEE Trans. Inform. Th., Vol. IT-14, pp. 640-650, Sept. 1968.

70. E. N. Gilbert, "Capacity of a Burst-Noise Channel," Bell Syst. Tech. J., Vol. 39, pp. 1253-1266, Sept. 1960.

71. B. D. Fritchman, "A Binary Channel Characterization Using Partitioned Markov Chains," IEEE Trans. Inform. Th., Vol. IT-13, pp. 221-236, April 1967.

72. A. H. Haddad, S. Tsai, B. Goldberg, and G. C. Ranierei, "Markov Gap Models for Real Communication Channels," IEEE Trans. Commun., Vol. COM-23, pp. 1189-1197, Nov. 1975.

73. P. K. Varshney and A. H. Haddad, "A Markov Gap Model with Memory for Digital Channels," Proc. Eleventh IEEE Int. Comm. Conf., San Francisco, June 1975.

74. P. K. Varshney, "Markov Gap Model with Memory for Digital Channels," CSL Report R-665, Univ. of Illinois, Oct. 1974.

75. J. P. A. Adoul, "Error Intervals and Cluster Density in Channel Modeling," IEEE Trans. Inform. Th., Vol. IT-20, pp. 125-129, Jan. 1974.

76. R. G. Gallager, Information Theory and Reliable Communication, p. 16, John Wiley, New York, 1968.

77. M. Abramovitz and I. A. Stegun, Handbook of Mathematical Functions, Dover Publications, New York, 1965.

78. J. Marcum, "A Statistical Theory of Target Detection by Pulsed Radar," IEEE Trans. Inform. Th., Vol. IT-6, No. 2, pp. 59-267, April 1960.

79. J. Marcum, "Table of Q-Functions," Rand Corporation Rpt. RM-339, Jan. 1, 1950.

80. A. R. DiDonato and M. P. Jarnagin, "A Method for Computing the Circular Coverage Function," Math. Comp., Vol. 16, pp. 347-355, July 1962.

81. P. Billingsley, Ergodic Theory and Information, John Wiley, New York, 1965.

82. V. I. Arnold and A. Avez, Ergodic Problems of Classical Mechanics, New York: Benjamin, 1968.

VITA

Pramod Kumar Varshney was born on July 1, 1952 at Allahabad, India. After completing his Bachelor of Science degree in Physical Sciences from the University of Allahabad in 1970, he came to the University of Illinois. He received his Bachelor of Science degree in Electrical Engineering and Computer Science with highest honors from the University of Illinois in 1972. He was a Bronze Tablet Senior and was also a James Scholar. He received his Master of Science in Electrical Engineering in 1974. He was a University of Illinois fellow in 1972 and has since then held teaching and research assistantships in the Department of Electrical Engineering and Coordinated Science Laboratory.

Mr. Varshney is a member of the Tau Beta Pi and IEEE.