REAL-TIME PERFORMANCE ESTIMATION AND OPTIMIZATION OF DIGITAL COMMUNICATION LINKS

by

Jason B. Scholz

Submitted April 1992
for the degree of

Doctor of Philosophy

University of Adelaide
Department of Electrical and Electronic Engineering

TABLE OF CONTENTS

Abstract

The subject of this thesis is performance estimation and optimization of digital communications links.

The demand for ever higher data throughput rates and lower error rates drives much communications research. As a result of this research some sophisticated error control strategies and higher capacity modulation schemes have evolved which now make it possible to design excellent digital communications links for which the channel statistics are stable. In contrast, variable quality channels pose a difficult design problem. The design method usually employed is to select the link configuration (modulation, error control, protocol, etc.) to suit the most probable channel statistics and restrict the data rate to give acceptable availability. Clearly, such a link will rarely be operating at its optimum performance. Vast improvements in performance would be possible, however, if performance measurement/estimation is rapid and the link configuration then tracks channel conditions.

In order to improve performance, it was firstly considered necessary to examine what constitutes "performance", how it is defined and determine which metrics of performance should be measured.

Error probability was considered to be the most important metric. However, the review showed that there was not a single technique which could be universally applied to any type of digital link to provide rapid, accurate measurements with minimal (and preferably zero) link overhead.

In response to this deficiency, a new method, which exploits the receiver decision variable more fully and uses a priori knowledge of the types of link condition to be expected was developed. This method coined EVEREST (Extremely Versatile Error Rate ESTimator) which can accurately estimate the decision variable probability density function and hence the probability of error, is the main subject of this thesis.

The theory of the new error measurement technique is developed and supported by an extensive set of simulations investigating the performance of the EVEREST for binary and quaternary signalling schemes with a wide range of channels including AWGN, Rayleigh fading, inter-symbol interference and carrier interference. Simulations are further supported by live High Frequency (HF) link measurements.

EVEREST uniquely offers high accuracy in a time-varying channel for a small number of samples. This characteristic is exploited in an innovative adaptive feedback communication system. It is shown that such a scheme can maximize data throughput whilst maintaining a ceiling on the error probability, thus allowing a guaranteed quality of service for the user or network.

The performance of the adaptive link strategy is simulated for a skywave HF link, using the Watterson channel model and parallel-tone modems. A significant improvement in terms of throughput, end-to-end delay and reliability over conventional coding schemes is demonstrated, supporting the utility of the EVEREST technique.

DECLARATION

I hereby declare that:

(a) This thesis contains no material which has been accepted for the award of any other degree or diploma in any University and that to the best of my knowledge and belief, the thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis; and

(b) I consent to the thesis being made available for photocopying and loan if accepted for the award of the degree.

Jason Beaufort Scholz

ACKNOWLEDGEMENT

I wish to thank my sponsor, the Defence Science and Technology Organisation (DSTO) for the opportunity to partake paid full-time study for this work.

I am indebted to my University supervisor, Dr. Bruce Davis for his guidance and inspiration. I am also indebted to my DSTO supervisor, Dr. Stephen Cook.

I wish to thank Mr. Timothy Giles for many hours of stimulating discussions and cooperative effort with the adaptive code rate feedback communication system, in particular, for developing and implementing the fading predictor. Thanks to Mr. Martin Gill the BCH coder and decoder software was quickly and efficiently implemented. Mr. Jeffrey Ball provided valuable technical support and assistance. Mr. Brian Andrews provided essential encouragement. Dr. Langford White gave inspiration and helpful guidance in early stages of the IEEE paper and Mr. John Wibrow for his enthusiasm and effort in the commercialisation of these ideas.

I also wish to thank my wife, Sabine for her patience and understanding during many long nights of study.

Chapter 1

PERFORMANCE MEASUREMENT OF DIGITAL COMMUNICATIONS SYSTEMS

Abstract

This chapter defines and establishes performance criteria for digital communication links. The criteria of accuracy is singled out as an important primary criterion. Emphasis is placed on surveyed *definitions* of performance *accuracy* which it will be seen are *all* (logically) based on digital *errors* and the ways that they occur.

Measurement of these digital errors is addressed in two sections covering error rate measures and error distribution measures. Error rate gets prime attention due to its wide-spread use and ease of measurement and interpretation. Error distribution measures are considered because they examine the *way* that errors occur and as such they have the ability to describe detail that may be very relevant to communications system performance which cannot be extracted from a simple error rate measure.

It is resolved that error rate is truly an estimate of error probability and that it is error probability that should be estimated.

In new, original work it is shown that the Bit Error Rate (BER), as defined by standards associations, is equal to the probability of bit error if and only if the sequence of data bits is ergodic (stationary and regular) over the period of measurement. Over a time-varying channel, these conditions may hold only over a short period of time. Consequently, attempts to estimate the error probability by, say, the error counting method are not necessarily valid. This endorses the necessity to estimate the error probability over a short period of time.

1.1 Introduction

In the conception, design, commissioning, and use of any communication system a key consideration is the ability to gauge the performance of the system while it is in operation.

It is a general characteristic of *analog* communications system that a gradual degradation in system performance occurs in relation to the level of signal distortion. This characteristic makes analog system performance more readily assessable than a *digital* communications system because performance in the latter case is, in general, a highly non-linear function of the signal distortion. For example, in a Gaussian noise environment and in particular where error correction is applied, a well designed digital system can offer very high performance, until a point where a small increase in signal distortion causes the performance to degrade rapidly. However, the nature of this non-linearity can vary and in contrast, a digital system operating in a severe environment (eg. a HF channel with flat fading) may offer an irreducible error probability, so that a large increase in the signal to noise ratio provides no performance increase.

1.2 Performance Criteria

The ultimate performance criterion for *any* communications system will be its fitness for purpose and ability to meet agreed user requirements. This user-based perspective is advocated by Richters and Dvorak (1988). Such criteria would imply that information transfer be achieved with sufficient *speed*, *accuracy* and *security* with minimum *cost*.

1.2.1 The Speed Criterion

The minimum acceptable *speed* with which information can be transferred between terminals depends on the nature of the information, the type of service used (voice, video or data) and physical limitations. Such a requirement is usually met by application of the appropriate form of communications system for the application. For example, a digital voice service would usually require that end-to-end delay be under 0.5 second

(and preferably far less) in order to maintain the satisfaction of its users. In order to achieve this, an appropriate form of communications channel and equipment set needs to be chosen.

For a *data* communications system, the performance criterion regarding the overall speed of information transfer has been addressed by the American National Standards Institute (ANSI) standard 3.44 (1974) and Moore (1971). The communications system is segmented into five "call phases" (each of which may involve *delay*) as follows:

(i) Connection establishment
(ii) Link establishment
(iii) Information transfer
(iv) Link termination
(v) Connection clearing

ANSI 3.44 defines three performance criteria related to *speed*, the Transfer Rate of Information Bits (TRIB) and the Transfer Overhead Time (TOT) which apply only to phase (iii), the information transfer phase of a call; and the Availabity, which applies to all five phases.

Four user-oriented performance criteria and definitions for data networks (related to *speed*) were proposed by McManamon et. al. (1975):

(i) Accessibility

A property of the system which enables a user to obtain permission, liberty, or to be able to enter a telecommunication system, or to communicate with a telecommunication system, in order to employ an available service, such as a transfer or exchange of information among users.

Measured by the probability of no *delay* at access and the average *waiting time* if there is delay.

(ii) Delay

The elapsed time experienced by an origination-destination user pair between the time of the start of information transfer by the origination user to the time at which the first bit of the transmission arrives for the destination user.

Measured as the absolute end-to-end *delay* time and the standard deviation of the delay time.

(iii) Efficiency

The degree to which the resources made available by a telecommunication system as a part of a service are utilised solely for transfer or exchange of information among users.

Measured as the percentage of transmission *rate* (bandwidth) used by the user. Lower transmission rate will incur a greater *delay* in receiving all of that information.

(iv) Reliability

The confidence that the telecommunications service, once access has been completed and the service initiated, will not be terminated by the network for a time duration at least as great as the service time interval.

Measured as the probability of a given *service time* interval.

McManamon's work formed a part basis for Federal Standard INT-FED-STD-001033 (1979). This standard specified digital communication performance parameters for U.S. Government agencies. Twenty six performance parameters were described and categorised under efficiency, accuracy and reliability. Efficiency was defined in terms of access time, bit and block transfer time, data transfer rate and disengagement time.

Throughput is a measure of transmission efficiency and is a key figure of merit in the evaluation of many digital communications systems (Burton and Sullivan, 1972):

$$T_e = \frac{N_a}{R\ T} \qquad (1.1)$$

Where:

N_a = The number of information bits accepted by the receiver.
R = Channel bit rate.
T = Measurement time.

Throughput is similarly defined in Lin and Costello (1983, pp. 461-462) as the ratio of the average number of information bits accepted by the receiver per unit of time to the total number of bits that could be transmitted per unit time.

Throughput is directly related to *speed* since if the throughput through a channel is limited, more *time* is required to send the information. For the case of Automatic Repeat Request (ARQ) systems the only uncertain variable affecting the throughput is the block error probability. In this situation the *speed* of the service is a direct function of link *accuracy*.

1.2.2 The Accuracy Criterion

The *accuracy* of received information may be defined as by McManamon et. al. (1975):

The freedom from mistake or error of a telecommunication system when providing a service such as a transfer or exchange of information among users.

The ANSI (1974) standard for *data* communications systems defines only one performance criteria related to *accuracy*. This is the Residual Error Rate (RER), which applies only to the information transfer phase. The definition of RER is:

$$RER = \frac{C_e + C_u + C_d}{C_t} \qquad (1.2)$$

Where:

C_e = Erroneous information characters accepted by the receiving terminal.

C_u = Information characters that were transmitted and assumed by the sending terminal to be accepted, but were not accepted by the receiving terminal (undelivered).

C_d = Information characters that were accepted in duplicate by the receiving terminal, but were not intended for duplication (duplicates).

C_t = Information characters contained in the source data.

McManamon et al (1975) described four accuracy performance criteria also related to digital data networks. The first of these was "Bit Error Ratio" which is defined exactly as RER, but it is more specific since it

was defined for bits only. The remaining three criteria are as follows:

(i) False Bit Ratio

The ratio of the number of correct and incorrect bits accepted by the user at incorrect destinations, to the total number of user originated information bits.

(ii) Duplicate Bit Ratio

The ratio of the number of correct and incorrect but duplicated bits accepted by the user at a correct destination, to the total number of user originated information bits.

(iii) Erasure Bit Ratio

The ratio of the number of bits which are not accepted by the user at a correct destination, to the total number of user originated information bits.

McManamon (1975) pointed out that error ratio measurements were an average measure and that more than an average would be necessary to reflect the fact that bit errors on telephone and data networks tend to cluster in time.

U.S. Federal Standard INT-FED-STD-001033 (1979) specifies the following primary digital communication accuracy parameters:

(i) Bit and block error probability
(ii) Bit and block misdelivery probability
(iii) Extra bit and block probability
(iv) Incorrect access probability

The incorrect access probability parameter is the only unique parameter in this standard and relates to the *misdirection* of information which may occur at network (or higher) levels.

Munday (1975) described error rate objectives for various services in a digital network. These objectives are detailed in table 1.1 for a hypothetical 6000 km reference path.

SERVICE	NOMINAL RATE	DESIGN OBJECTIVE	OPERATIONAL REQUIREMENT
Colour TV	70 Mbits/s	2.2 in 10^6	1 in 10^4
Switched monochrome visual visual telephone	8M bits/s	2.3 in 10^6	1 in 10^4
Monophonic sound programme	320 kbits/s	7 in 10^6	1 in 10^5
N-supergroup FDM assembly	70 Mbits/s	2.2 in 10^6	1 in 10^5
Public switched telephony	64 kbits/s	7.1 in 10^6	1 in 10^5
Common signalling channel	64 kbits/s	7 in 10^6	1 in 10^5
Alphabetic telegraphy	50 baud	7.1 in 10^6	1.3 in 10^6
Low-speed facsimile	4.8 kbits/s	7.1 in 10^6	1 in 10^6
High-speed facsimile	48 kbits/s	7 in 10^6	1 in 10^5
Data transmission	64 kbits/s	7.1 in 10^6	see table 2.2

TABLE 1.1. Summary of the design objectives for (Poisson) error rate for a 6000 km hypothetical reference path. Note add 0.5 in 10^6 to the design objective if a satellite link is used. (reproduced from Munday, 1975).

CATERGORY OF DATA	MAXIMUM PERMISSIBLE CHARACTER ERROR RATE
Non self-evident:	
irrecoverable, critical data	
eg. air traffic control or computer program data	1 in 10^9
general data	1 in 10^7
low-grade data	1 in 10^5
Superseded information:	
slowly superseded	1 in 10^6
rapidly superseded	1 in 10^5
Self-evident text:	
computer conversational mode	1 in 10^4
high-grade correspondence	1 in 10^4
general message	1 in 10^3

TABLE 1.2. Summary of ISO guidance on character error rates for data reproduced from Munday (1975).

The evolution towards an Integrated Services Digital Network (ISDN) for voice and data traffic has led to new quality specifications for digital transmission performance. The Comité Consultatif Internationale Télégraphique et Teléphonique (CCITT) recommendation G.821 (1988a) defines three quality parameters relating to the occurrence of errors on a 64kbps

link as given in Table 1.3.

TERM	DEFINITION	PERFORMANCE OBJECTIVE
Degraded minutes	One minute interval with a BER>10^{-6}	<10% degraded minutes over a time period T
Severely errored seconds	One second interval with a BER>10^{-3}	<0.2% severely errored seconds over time T
Errored seconds	One second interval with any error	<8% errored seconds over time period T

TABLE 1.3. Error performance specifications, from CCITT (1988a).

Table 1.3 also shows the performance objective over some period of time T. This period is not specified in the standard, but a period of the order of one month is suggested as a reference.

In assessing these parameters, periods of unavailability are excluded.

These parameters clearly recognise the burst nature of error occurrences and use different measurement time intervals (or block lengths) to account for the different ways that errors can occur.

1.2.3 The Security Criterion

Security was defined by McManamon et. al. (1975):

The degree to which a user's information, while in a telecommunications system, including the network and terminals, is guarded against exposure to terminals and persons, and interpretation by terminals and persons, not explicitly authorized by the user.

Security is a performance parameter which applies during the entire information transfer interval.

1.2.4 The Cost Criterion

The cost criterion relates to providing a particular level of service. In general, increased speed and security directly affect service cost.

The *cost* of communications between users depends on the cost and

sophistication of the communication equipment used. This includes the cost of performance monitoring equipment.

1.3 Performance Accuracy Metrics

1.3.1 Error Rate

Error rate is the most commonly used measure of performance in a digital communications system. It is an important measure since it generally gives a good indication of the quality of the received information.

If the communications system includes error correction coding, the error rate of the data both before and after decoding is important. The post-decoding error rate measure refers more directly to the quality of the *information* (as correctable errors and redundancy are removed). The post-decoding error rate is related to the pre-decoding error rate by the structure of the decoder and the statistical nature of error occurrences prior to decoding. Most importantly, pre-decoding error rate gives a greater indication of how close the link is to failure.

Error Rate (ER) is defined by Newcombe and Pasupathy (1982) as the ratio of the number of units received containing errors (N_E) to the total number of units received (n) during a specified measurement interval:

$$ER = \frac{N_E}{n} \tag{1.3}$$

A fundamental flaw of this metric (1.3) is the fact that the measurement interval *is not* specified in its calculation. The length of the measurement interval however, does affect the accuracy of the ER value. The meaningfulness of the metric as a performance measure outside the nominal measurement period comes into question if the statistics of error occurrences are bursty in nature, or varying during the measurement interval.

If this measurement is performed over a communications link whose statistics are fixed, then the error rate would approach the probability of error as the measurement time extended to infinity (Jeruchim, 1984).

The error rate may be defined at a multitude of levels in a communications system hierarchy. The Open Systems Interconnection (OSI) 7-layer model is described by Stallings (1988). This model describes the transfer of information in a communications system. In this model hierarchy, error rate is usually associated with lower layers, including the physical link, data link and network layers.

If the information bits are considered at the lowest level, the smallest unit for measure is bits and thus the Bit Error Rate (BER) is obtained. At higher levels, the bits are usually bunched into symbols at the modulation or lower link level and larger units such as characters, blocks and packets at higher system levels. The presence of one or more bit errors in a symbol would constitute a symbol error and thus a Symbol Error Rate (SER) may be defined. Similarly the Character Error Rate (CER), BLock Error Rate (BLER), and Packet Error Rate (PER) may be defined. All of these error rate measures may be defined as prior or post coding error rates, if error control coding is used at these levels.

1.3.1.1 Statistical Uncertainty in the Error Rate

The calculation of bit error rate from equation (1.3) requires the number of bits in error N_E to be calculated. The process of forming N_E is a Bernoulli trial described by equation (1.4).

$$N_E = X_1 + X_2 + \ldots + X_n \tag{1.4}$$

N_E is thus a random variable which is the sum of n binary random variables X_i, where $X_i = 0$ if received bit i was correct or $X_i = 1$ if bit i was in error.

Combining equations (1.3) and (1.4), the BER is thus calculated as:

$$BER = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{1.5}$$

The *time-average* mean of the process x(t) is:

$$\eta_x = \lim_{n\to\infty} \left[\frac{1}{n} \sum_{i=1}^{n} X_i \right] \qquad (1.6)$$

A natural (unbiased) estimate of the time-average mean is:

$$\hat{\eta}_x = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad (1.7)$$

Note that this estimate is equivalent to the BER in equation (1.5).

The *ensemble* mean of the random variable X_i is:

$$\eta_{x_1} = E[X_1] = \sum_{\text{all } X_1} x_1 \, p(x_1)$$

$$= 0.P(X_1=0) + 1.P(X_1=1)$$

$$= P_e \qquad (1.8)$$

Where E [.] denotes the expectation operator.

If the stochastic process x, is regular (the time-average mean does not depend on the particular realisation) and stationary (the ensemble mean is not a function of time) then the process is said to be *ergodic*. In this case, the time-averaged mean and the ensemble mean are equivalent. Thus, equations (1.6) and (1.8) may be equated, so the probability of bit error is:

$$P_e = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i \qquad (1.9)$$

Then, from equations (1.7) and (1.8), an unbiased estimate of the probability of bit error is:

$$\hat{P}_e = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$= \text{BER} \qquad (1.10)$$

Thus, for finite n, an estimate of the probability of bit error is the BER.

For finite n, the reliability of the estimate may be quantified in terms of confidence intervals. These curves are given in Portny (1966) and Jeruchim (1984) as a function of the number of errors detected. The curves are derived from formulae given by Burington and May (1958) which assume that the x_1 random variables are independent and identically distributed (iid), thus producing a binomial distribution for N_e. From these curves the rule that a sufficient number of samples is approximated by (1.11) becomes clear.

$$n = \frac{10}{P_e} \tag{1.11}$$

The relationship given in (1.11) coincides with a confidence interval of about 0.3 in log error rate about the estimate for a 95th percentile confidence. For example, a bit error rate estimate of 10^{-4} would imply a probability of error of between $10^{-3.7} = 2 \times 10^{-4}$ and $10^{-4.3} = 5 \times 10^{-5}$. Figure 1.1 is a reproduction of the confidence interval graph given in Jeruchim (1984).

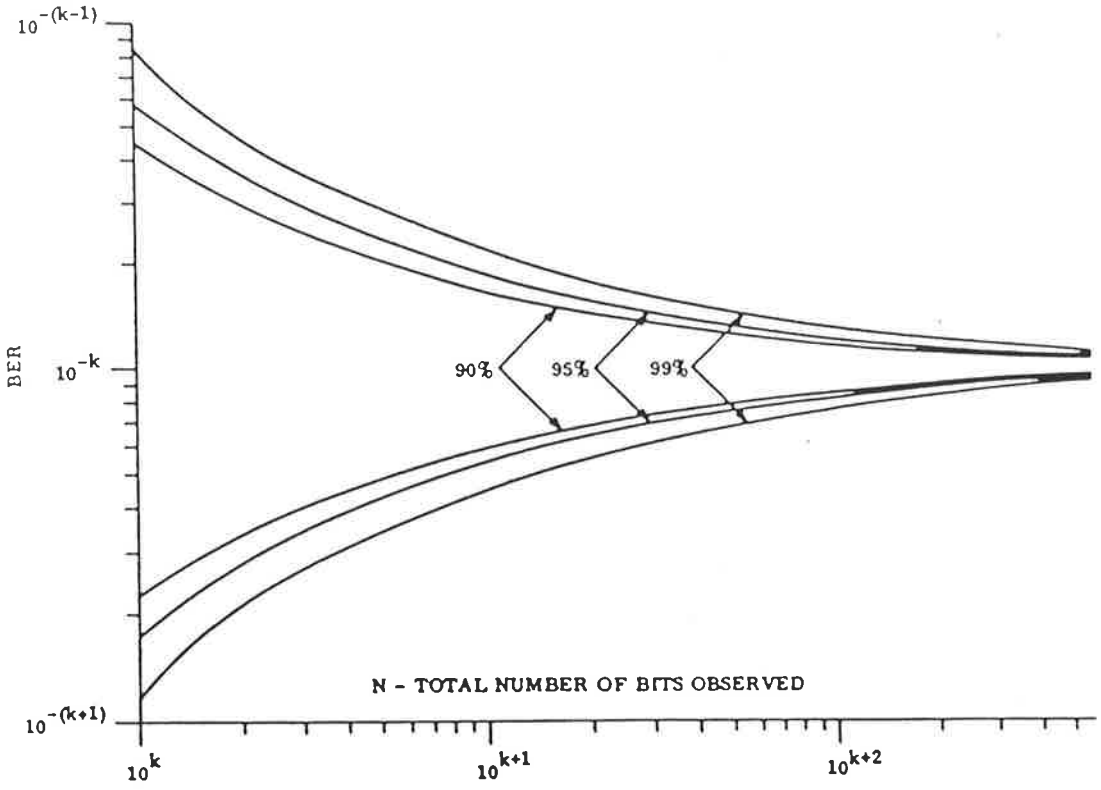


FIGURE 1.1. Confidence Bands on BER when observed value is 10^{-k} (reproduced from Jeruchim, 1984).

The formulation of the confidence intervals is made tractable by the assumption of independence between error events. Any practical communications link will introduce some degree of dependence. Dependence is introduced by memory, which may be present as a result of encoding, modulation, the channel, or other forms of filtering. Jeruchim (1984) indicates how such dependence may be taken into account in the computation of the confidence interval. Assuming a simple model whereby a single error gives rise to a burst of m consecutive errors, followed again by independence (until the occurrence of another error and so on...), the variance in the probability of error estimate is approximated by (for n>>m):

$$\sigma_{P_e}^2 \simeq \frac{1}{n} pq(1 + 2m) \qquad (1.12)$$

where, as usual, p = probability of error, and q = 1-p.

The confidence interval (which is proportional to σ) is therefore increased by a factor of $\sqrt{(1+2m)}$.

Davis (1988) also describes the effect of sample correlation when the error probability is estimated by counting errors.

1.3.1.2 Sufficient Measurement Time

The definition of Error Rate (ER) includes an implicit measurement interval which is not specified in equation (1.3). The main factor governing the required measurement time is the reliability with which the ER is to be calculated. Assuming the x_1 random variables in (1.4) are iid, then the BER estimate will be within a factor of 2 of the probability of error if (1.7) holds and hence:

$$T \geq \frac{10}{R\ P_e} \qquad (1.13)$$

Where: R = The net channel bit rate.

Measurement time is also a key issue in understanding what the error rate means. For example, Kim (1989) proposed that the performance of a land mobile satellite link be described by a metric called the "Bit by Bit Error Rate" (BBER). It was recognised that the error rate according to equation (1.3) is a measure which would require a long measurement time (to obtain accurate results) (1.11) and therefore be unsuited to describing the dynamic nature of the channel in this form of link. The BBER metric was really an attempt to define the error rate over a *very short* time scale, so that it would be more meaningful for a rapidly dynamic channel. Unfortunately, the BBER was only usable for simulation purposes since the instantaneous SNR and form of fading was known for each data bit generated, and hence the BER (actually probability of error) *per data bit* could be calculated. Results show that BBER as a metric can be used to predict the exact location of error bits under fast fading conditions.

1.3.2 Error Distribution Measurements

The error rate is undoubtedly the most important first order statistic of digital communications system performance. However, the error rate is an *average* measure and does not reflect *how* the errors occur.

Many communications channels experience effects which give rise to bursts or clusters of errors. For example, transmission signal fading is experienced over ionospheric, tropospheric and mobile radio channels, giving rise to bursts of errors. Natural and man-made interference can also cause burst errors. Error bursts have also been observed on digital transmission lines by Brillant (1978), Urien and Rault (1982), and Wataya (1989).

Error bursts imply some form of *dependence* between errors. This dependence may be viewed in terms of the statistical distribution of some definable metrics. Knowledge of these distributions can allow the communications link to be optimised, either during the design stage, as proposed by Rahman and Bulmer (1990); or during operation, by measuring the distribution and choosing the best form of equipment to yield optimum system performance. Typically, the distributions can be used to select a burst-error correcting code, or where random error correcting codes are

employed, a sufficient interleaving length as discussed by Fritchman (1967) and Tsai (1969).

1.3.2.1 Error Distribution Metrics

As outlined in the introduction the CCITT (1988a) defined three performance metrics which may be related to burst error measurements. These were defined in table 1.3 earlier. Becam et al (1987) describe error distribution measurements for HDB3 and 4B3T line codes.

It is possible to define a far wider variety of metrics, depending on the nature of the bursts. In order to physically measure and gather these distributions, it is necessary to employ a Monte Carlo type method, whereby a stream of data bits known to the receiver is sent and straight-forward bit-by-bit comparisons are made to form an error stream.

The following definitions of error distributions are from Tsai (1969):

Gap Distribution. A gap is a region of error-free bits between two errors; the length of the gap is the number of those error-free bits. The gap distribution is a plot of the cumulative relative frequency of the gap length versus length. The gap distribution gives some indication of the randomness of the link, or the dependency between consecutive errors.

Burst Distribution. If a channel is not random, there are more errors in certain parts of the data stream. A density Δ is defined:

$$\Delta = \frac{\text{number of errors}}{\text{total number of bits in a region}} \qquad (1.13)$$

A burst is defined to begin with an error and end with an error. If the successive inclusion of the next error keeps the density above a specified value, the burst is continued, otherwise, the burst is defined to end at that location. Also, a burst should not begin with an error belonging to a previous burst.
The burst distribution is then a plot of the cumulative relative frequency of bursts of length m versus the length m.

Burst Interval Distribution. A burst interval is the region between two bursts. The error density of a burst interval must be less than the prescribed density Δ chosen to define the burst. This distribution is a high-order statistic and gives some indication of the dependence between bursts.

Cluster Distribution. A cluster is a region of consecutive errors. The cluster distribution is a plot of the probability of m consecutive errors versus the number m.

Wataya (1989) reports on practical measured gap distributions for a metallic cable channel. In particular, conversion of error rate into CCITT terms (table 1.3) of errored seconds is discussed.

1.4 Chapter Summary

This chapter has examined the fundamental user-based criteria of performance for digital communication links. Attention was focussed on one key criterion, called *accuracy*. The emphasis on definitions from reviewed Standards showed that error-based measures predominate as the universal metrics of performance accuracy.

A more detailed examination of the most common error-based metric, *error rate*, showed the true relationship to error probability. The error rate measure is criticised from the aspect that there is no component of time in its dimension, unlike error probability which is defined as time tends to infinity.

A brief examination of other error-based metrics did not yield a single metric which would be universally applicable.

From these considerations, it is advocated that *error probability* is the truly fundamental parameter which must be estimated to determine *performance accuracy*.

Chapter 2

REVIEW OF ERROR MEASUREMENT AND ESTIMATION TECHNIQUES

Abstract

This review chapter aims to examine and compare methods of error monitoring with regard to the *desirable features* that *any* monitor should possess. These features were defined by Newcombe and Pasupathy (1982):

(i) SIMPLICITY
Simple implementation so that monitoring does not substantially increase the cost or complexity of the system.

(ii) SPEED
Fast response to changes in the ER to allow rapid detection of faults.

(iii) THROUGHPUT
Ability to estimate the ER with little or no interruption in the flow of useful information.

(iv) PRECISION and REPEATABILITY
Accuracy and reproducability to a level adequate for the use being made of the information provided by the error monitor.

(v) PERFORMANCE FOR REAL CHANNELS
Sensitivity to the parameters that affect the ER and robustness to those that do not.

2.1 Introduction

The last known reviews on digital communications system performance techniques were by Smith (1973), and Newcombe and Pasupathy (1982), which concentrated on error monitoring.

In order to embrace the history of error monitoring, the approaches taken by researchers have been assigned to one of several categories. The most used technique in practice is the *direct* method, which involves measuring the actual errors occurring. Alternatively, if the errors are not to be measured directly, they must be *estimated* from the receiver system. Two classes of estimating monitor systems may be identified, *parametric* and *non-parametric* monitors. A parametric monitor examines a parameter or parameters which are (hopefully) related to error occurrences and translates the parameter(s) values into an error metric. A section is devoted to each of these categories.

2.2 Direct Measurement Methods

2.2.1 Test Signals and Sequences

These methods of error rate measurement involve the transmission of digital and/or analogue test sequences or signals.

2.2.1.1 Digital Test Sequences - The Monte Carlo Method

The only true method of error rate *measurement* is the Monte Carlo method as examined by Jeruchim (1984) and involves sending a known sequence of bits to the receiver. The actual number of errors can be counted directly by accumulating a count of the number of units (bits, symbols, characters, blocks, etc..) received which contain errors. The error rate may then be formed as described in equation (2.1). Pseudo-random binary sequences are widely used for the transmission bit stream. Such systems commonly use the complementary self-synchronising error detection scheme patented by Westcott (1972) which allows for simple generator/receiver synchronism.

Given that a level of uncertainty in the BER estimate of at least a factor of two is acceptable, the measurement interval required (2.7) may be prohibitively long for some applications. For example, 33 minutes is

required to determine a 10^{-4} error probability on a 50 bits/sec link, or 83 minutes to determine a 10^{-9} error probability for a 2M bits/sec link.

The BER measurement range for 64kbits/s links is specified as 10^{-2} to 10^{-7} in CCITT requirement 0.152 (1988b). This requires a measurement time (2.7) of at least 1,560 seconds (26 minutes) to resolve a BER of 10^{-7}. In fact, a much longer measurement period is required considering the burst nature of error occurrences on these links, as examined by Brillant (1978), and Urien and Rault (1982).

In CCITT recommendation 0.151, the range of BER measurement is specified as 10^{-3} to 10^{-10} for bit rates of 2.048M bits/sec. This requires a measurement time (2.7) of 48,800 seconds (or 13.5 hours) to resolve a BER of 10^{-10}!

Besides the long measurement time, this method has the disadvantage that it cannot function in an on-line mode and information transmission must be stopped, thus reducing the throughput to zero. Such methods are, however, pertinent for off-line link tuning and diagnostic work and have been applied to very high rate (Gigabit/sec) data systems by O'Reilly and Rampaigul (1987).

2.2.1.2 Analogue Test Signals

The name of the method refers to analogue signals which are added to the transmitted signal to allow measurement of error rate.

A patent by Crawford et. al. (1983) described a method of inserting a variable amplitude pulse into a baseband digital stream. These pulses were used to build up an estimate of the noise margin probability density *at the decision threshold*. The variable pulses were inserted at positions in the data sequence known to the receiver and the pulse amplitudes were progressively increased. As the amplitudes increased, the receiver output state gradually shifts from one state to another. At some value of the variable pulse, the receiver starts detecting the pulse as the other state. Since there is noise in the system, the change happens over a range of amplitudes of the variable signal. If the probability of receiving that pulse incorrectly is plotted against the variable pulse

amplitude, some information about the amplitude probability density function and the variance can be deduced. The occurrences beyond the threshold are accumulated and a calculation of error probability made based on the assumption that the noise is Gaussian.

The major disadvantages of this method is that it requires adding overheads to the transmission and that it necessitates the use of linear output stages in the receiver.

2.2.2 Decoder Interrogation

Most modern digital communications systems employ some form of error control coding to protect digital data against the errors that occur during transmission through a communications channel. All forms of coding require adding redundancy to the information bits. Two basic types of error control strategy are employed. The first uses error *detection* coding, that is, a code which is optimum for the detection of errors in a block. Error control is achieved by request for retransmission of blocks found to contain errors. These schemes are referred to as Retransmission or Automatic Repeat Request (ARQ) schemes. The second utilises error *correction* coding or Forward Error Correction (FEC) which detects the location of errors and is therefore capable of correcting a number of errors in a sequence or block, without retransmission. Hybrid systems are also possible, combining both retransmission and FEC schemes as discussed by Lin and Costello (1983).

These codes besides fulfilling their primary function to control errors, may be exploited to give information about error occurrences, as described by Newcombe and Pasupathy (1982). It is necessary, however, to define at what point in the communications system that error occurrences are being examined. The probability of error *prior* to decoding may be estimated by obtaining the accumulated number of errors detected by the decoder over some period of time and dividing this number by the total number of units (bits, symbols etc..) received in this time, as described in a patent lodged by the Aust. Telecommunication Commission (1988). This method is similar to the Monte Carlo method described previously because errors are counted directly. However, this method has two problems:

(i) No decoder can detect *all* possible error patterns.

(ii) The required measurement time is long.

Any coding scheme is limited in the number of errors and the patterns of errors it can detect. If an error pattern corrupts the transmitted data stream in such a way that the received data then resembles a valid codeword, the decoder will interpret that data as error-free. Therefore by interrogating the decoder for the number of detected errors, the estimated pre-detection error probability will be *less* than it actually is.

By counting detected bit errors, the estimation of error probability is subject to counting statistics. It is therefore necessary to know what constitutes a sufficient sample size (or equivalently measurement time) in order to estimate the error probability with a given degree of accuracy. Equation (2.7) applies directly, so that as for the Monte Carlo method, a prohibitively long measurement time is required in many situations.

It should be emphasized that this method does have a distinct advantage over the Monte Carlo method in that error monitoring imposes no additional overhead on the link if error control is already employed.

2.2.2.1 Error Measurements Where Retransmission Schemes are Used

The most striking feature of retransmission schemes are their inherent reliability. This is a consequence of the fact that the fraction of *undetectable* error patterns is only $1/2^r$, where r is the number of parity bits, regardless of the length of the code (Burton and Sullivan, 1972). Since a code can be easily chosen to detect the vast majority of all error patterns, it does not matter very much *how* errors occur on the channel.

Retransmission schemes are usually designed to have an arbitrarily small post-decoder error rate and this quantity is not generally of interest.

2.2.2.2 Error Measurements Where Forward Error Correction Schemes are Used

In digital systems employing Forward Error Correction (FEC) coding, detection of pre-decoder errors may be achieved quite effectively by

examining the number of errors detected; as in a retransmission scheme.

The weaker error detection capability of FEC schemes results in non-negligible uncorrected or undetected post-decoder error rates. The probability of decoder error (decoder failure) is generally calculated only as an upper bound assuming random errors, as in Proakis (1989:p417,461). In practice, it may depend on the error pattern (particularly for convolutional codes).

Block Codes

When block codes are employed, the decoder output will *always* be one of the set of valid codewords, so the syndrome must be examined to determine errors.

A block code designed to detect random error patterns will generally be capable of detecting dependent (or burst) errors only if the burst contains less than number of detectable errors per block. Blahut (1984) describes Fire and Golay codes which are examples of block codes specifically applicable for the detection and correction of burst errors.

This method of error rate measurement is suffers the same limitations as the error counting (Monte Carlo) method since we are again faced with gathering a significant number of error occurrences.

Convolutional Codes

The difference between the raw (demodulated) bits and the decoded bits may be used to directly yield the error rate.

In a Viterbi decoder, the decoded bit stream is found by selecting the code sequence which is closest to the received sequence. The measure of closeness is carried in *branch metrics* which are summed continuously. The minimum (accumulated) branch metric thus represents the number of detected *errors* in the received sequence for a hard decision decoder. The BER may therefore be estimated by summing the minimum state metric over several constraint lengths of the code (so that a sufficient number of errors have been counted) and then dividing the sum by the number of bits decoded. This method equates with the Monte Carlo method of error counting and may thus require a long measurement period in many circumstances, as discussed

earlier.

2.2.3 Pattern Violation Detection for Line (Spectrum Shaping) Codes

Coding in line systems is very different from coding for error control. Error control codes are concerned primarily with the detection of errors whereas the ability of a line code to correct errors is useful, but is a secondary consideration to the following primary signal design requirements (Newcombe and Pasupathy, 1982):

(i) Efficient use of available bandwidth (shaped spectrum).
(ii) Ease of recovering a timing signal.
(iii) Small low frequency signal content (for AC coupled links).
(iv) System behaviour independent of data values.

Many codes do, however, have an inherent error detection capability as a result of fulfilling the above requirements, which facilitates methods of counting error occurrences.

2.2.3.1 Bipolar-Based Ternary Code Violation

For example consider a simple ternary code, each data bit may be mapped onto one of three symbols; Miller and Ahmed (1987). An Alternate Mark Inversion (AMI) or *bipolar code* for example, would map data state 0 to zero, while state 1 would be encoded to pulses with alternating polarity. Thus there is no DC energy in the signal if data bits are random, and appropriately chosen pulse shapes will yield good spectral response. A single error in transmission of the ternary stream will cause the introduction of an extra pulse, remove a pulse or reverse the polarity of a pulse. For AMI, this will result in two or more consecutive pulses with the same polarity, which is not a valid code sequence. Detection of these *code violations (or format violations)* for single errors is easily achieved in an AMI signal. However, if a second error occurs soon after another, the effects of the two errors may cancel one another and no pattern violation will be detected. Thus, the count of code violations does *not* necessarily correspond to the count of actual errors, since there are sequences of error patterns which cannot be detected.

The AMI code suffers if a long sequence of zero symbols occurs, since the energy in the received signal will be low and the recovery of bit timing will be difficult. To avoid this, more sophisticated codes substitute a string of the same length for the sequence of zero's. This string must incorporate transitions to boost the signal energy. High Density Bipolar (HDBn) codes place a single violation (pair of pulses of the same polarity) in the replacement sequence. The receiver identifies these violations and substitutes back the all zero sequence. To avoid introducing a low frequency component in the spectrum, the polarity of the introduced violations are alternated. For the case of the BZ6S (bipolar with 6 zero substitution) code, the substitution sequence is chosen to contain two violations of opposite polarity to eliminate low frequency components. The receiver identifies these violations and substitutes back the all zero sequence.

CCITT recommendation 0.161 (1988b) describes code violation monitoring equipment for AMI, HDB3, B6ZS and B8ZS codes required to operate at bit rates of 1.544Mbits/s to 8.448 Mbits/s. These monitors are specified to give only a fairly coarse indication of "violation rate" on a meter showing:

(i) No signal.(meter reading zero)
(ii) Violation rate$\geq 10^{-3}$.
(iii) Violation rate$\geq 10^{-4}$.
(iv) Violation rate$\geq 10^{-5}$.
(v) Violation rate$\geq 10^{-6}$.
(vi) Single code violations.
(vii) Valid signal.(meter reading full scale)

Code violations are defined in CCITT (1988b) for the AMI code as two consecutive marks of the same polarity and for the HDB3 code as two consecutive bipolar violations of the same polarity. For the B6ZS and B8ZS codes, a code violation is defined as two consecutive marks of the same polarity *excluding* violations caused by the all zero substitution code. Thus any error pattern resulting in an all zero substitution code will be undetectable and what is more, would cause the substitution of incorrect bits for that pattern.

Becam et. al. (1987) described error distribution measurements for HDB3 and 4B3T line codes concluding that 85% of single binary errors were detected using the HDB3 code, with at worst, less than 15% of errors detected if they occurred in bursts. No code error is detected in the case of consecutive binary errors. For the case of the 4B3T code with descrambler, code errors led to an underestimate of binary errors. Between 13% and 20% of isolated errors were detected. At worst less than 5% of burst errors were detected, which is made more serious as the combined effects of transcoder and descrambler create bursts even when the line errors are isolated.

2.2.3.3 Partial Response Code Violation

By introducing a controlled (deterministic) amount of Inter-Symbol Interference (ISI) into the formation of transmitted symbols, it is possible to produce a transmission signal with a more efficient frequency response (narrower bandwidth and less low frequency energy) than a bipolar code. The added ISI yields a deterministically correlated signal. Proakis (1989: p614) shows that this signal is optimally detected by maximum likelihood sequence estimation using the Viterbi algorithm. Many forms of pulse shape are possible but recovery of the data sequence becomes complex when the number of non-zero symbol states increases beyond two.

The error detection capability inherent in the correlated transmission signal was originally discussed by Lender (1964). This paper also described circuitry to perform error (violation) detection for a polybinary system.

Gunn and Lombardi (1969) described an algorithmic method of error detection for a partial-response system. The system used fifteen symbol states and examined the running sum of receiver detected states to see if a constraint bound was violated. The constraint was that the current sum of received symbol values should not lie outside an upper and lower bound, formed as a function of the state value from a time slot two slots previous. After an adjustment period (necessary due to the fact that the initial state is unknown), almost 100% of errors were detected at low error rates (below 10^{-3}). The ability to detect errors when they occur in

bursts was not discussed.

Lender (1968) described the error detection performance for a modified duo-binary signalling scheme. Due to the correlation span extending over three digits rather than two as in duo-binary, error detection was more effective. That is, both single and double errors were found to be detectable. Longer bursts would not necessarily be detectable.

In common with ternary codes, generally only single errors can be detected reliably using pattern violations for partial response systems.

2.3 Parametric Estimation Techniques

The performance of digital communications systems is a complex function of many signal parameters. The correlation between these parameters and probability of error is therefore often difficult to specify in quantitative terms. Regardless, certain signal parameters have been used to indicate communications system performance.

One advantage of parametric estimation techniques is that a performance metric is formed on the basis of received signal parameters *without* the need to send test signals or sequences. Thus they do not usually compromise information throughput and may be performed on-line.

2.3.1 Eye Pattern Monitor

The eye pattern is a display of the received waveform amplitude versus time overlaid for many symbol periods. This is most easily achieved by applying the signal to an oscilloscope and setting the timebase (repetition frequency) to be equal to one symbol period. If a storage-type oscilloscope is used, the trace from many symbol periods may be displayed simultaneously, as shown in Figure 2.1.

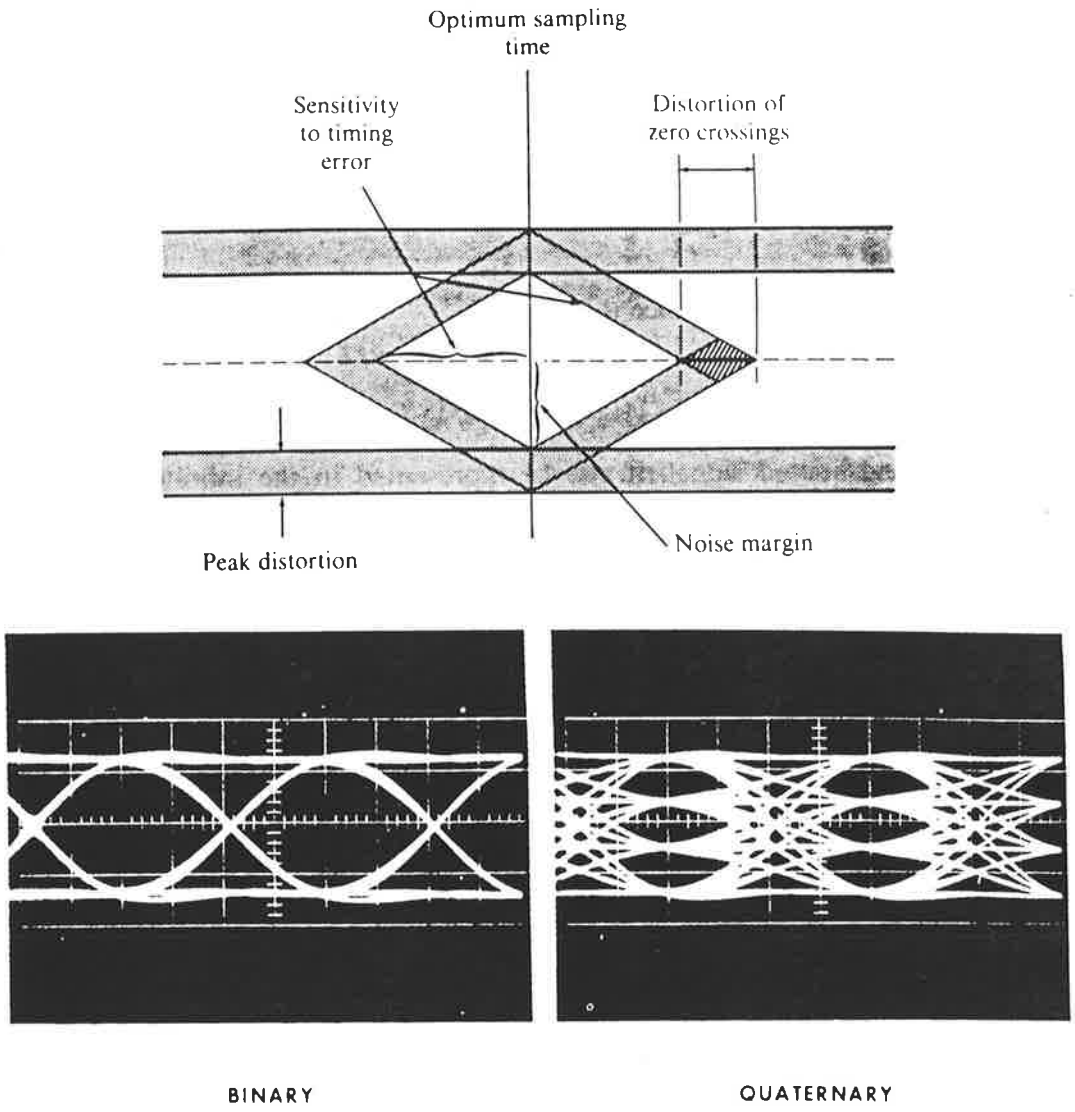


FIGURE 2.1 Eye pattern display as viewed on an oscilloscope display.

The basic shape of the eye gives some indication of the phase distortion and inter-symbol interference present in the signal. Increased noise and distortion in the signal makes the vertical dimension of the eye pattern close and changes in timing cause the horizontal dimension of the eye to close.

One possible means of estimating transmission quality proposed by Leon et. al. (1974) was to measure the maximum and minimum values of the eye opening. The noise margin is measured at the sampling time and is the inside height of the eye (see figure 2.1). This represents the amount of added noise required to cause a digital error. Therefore there is a relationship between noise margin and error rate. Large noise margin

values represent low error rates and small values represent high error rates. Eye opening measurement monitors have also been described by Leon et. al. (1974), and Argenzia and Campbell (1977).

To a trained human operator, the eye pattern contains a lot of useful information about the quality of the received signal. Automated eye pattern monitors can only observe a small amount of this information. An indication of the range and sophistication of possible measurements is given by Kawashima (1987).

2.3.2 Timing Jitter Monitor

"Jitter" is defined by the CCITT (1984) as short-term variations of the significant instants of a digital signal from their ideal positions in time. "Wander" is defined as longer-term variations.

According to Cock et. al. (1975), the bandwidth of jitter disturbances is usually proportional to the symbol rate. Low frequency jitter is caused by variations of propagation delays in transmission lines, in demultiplexer outputs (due to inaccuracy of frame boundaries) or from basic instabilities in system clocks. Jitter may also result from network synchronisation control.

A brief review of jitter measurement techniques is given by Cock (1987). Measurement of timing jitter involves the detection of timing displacements of the edges of digital signals and/or the zero-crossings of sinusoidal or rectangular clock signals. Timing jitter is equivalent to angle modulation and consequently the most common technique of measurement is to "demodulate" the digital signal by applying the jitter to a baseband calibrated phase demodulator. Such a system was described by Hoyer (1985) to meet CCITT recommendations. Hall and Lancaster (1975) described practical measurement techniques.

The effect of jitter on error performance has been examined experimentally by McManamon (1973) for an HF ionospheric channel, using Binary Frequency Shift Keying (B-FSK) at 75 bits/sec. McManamon's hypothesis was that the number of binary errors in a sample sequence of N bits is reasonably correlated with the number of data transitions near the expected centre of the bit intervals for $N \gg 1$. The experiment examined the distance of

transitions from the expected sampling point and the number of bit errors. By dividing each symbol period up into a number of sub-intervals, the cumulative distribution of offset transitions was measured. It was found that the number of bit errors counted (over some period of time) was usually between the number of counts accumulated in two of the sub-intervals. Thus the number of transition counts in sub-interval x and y, say, could provide an upper and lower bound to the number of true errors in that period of time.

The correlation coefficient (r) relating the number of counts in the upper bin (the upper bound) and the actual number of errors counted has been estimated using the formulae given by Walpole and Myers (1978) and using McManamon (1973:table 2.2) as r=0.995. The correlation coefficient relating the lower bin count and the number of errors was also estimated to be r=0.994. These are both high correlations, however several considerations should be made in considering these figures. Only 43 samples of error estimates were given, reflecting a total of 86 minutes of measurements. Of the total number of measurements, 33% of these were *below* the measurement resolution of the method (14 counted errors were zero) another four of the 43 total samples were *not* between the bounds hypothesised.

As stated by McManamon, the results apply only to the specific equipment and channel used in the experiment. There is no evidence to support any such correlation on any other type of link. It is clear that any relationship between timing jitter and error rate will generally be a complex function, particularly when combined with the effects of channel disturbances. In general, channel disturbances may cause significant errors totally unrelated to the timing jitter in the system.

Urien and Rault (1982) describe measured correlations between timing jitter and error rate on 2Mbit/s digital lines. The correlation between these parameters was visibly poor and the jitter measurement was used only to make a first estimate binary decision as to whether the link error rate was acceptable or not. One measurement problem was the fact that errors were recorded in bursts and the average error rate was low, ranging between 10^{-7} and 10^{-13}.

2.3.3 Line-code Spectrum Monitor

A desirable characteristic of line-coded communications systems is that the spectrum of the transmitted signal should contain little or no low frequency energy.

Pulse imbalance caused by errors will result in an increase in this low frequency energy. Hence, one way of detecting errors is to examine the amount of low frequency signal content, as proposed by Catchpole (1975). However, when used for in-traffic measurement, this method was found to be unsuitable because spectral fluctuations in normal traffic could not be discerned from errored data. In contrast, for out-of-traffic use, the technique was found to be far more effective for detecting error occurrences because a pseudo random test pattern (which does not contain any power at frequencies below its repetition frequency) was used. This is hardly a redeeming characteristic because if a known data sequence were sent, it would be more exact to actually count data errors.

2.3.4 Signal to Noise Ratio (SNR) Monitor

This is a commonly used performance metric. The popularity of this measure may be as a result of its effectiveness in indicating the performance of *analog* communications systems. This metric is pertinent in simple digital communications systems with an AWGN channel, where the received energy per bit is often the only required parameter to characterise the probability of error.

Nahdi and Gagliardi (1967) described a SNR estimator for a Gaussian noise corrupted signal. Gagliardi and Thomas (1968) described a maximum likelihood SNR estimator for a general signal waveform with additive Gaussian noise. Confidence bounds on the SNR estimate were formed for a Pulse Coded Modulation (PCM) system. A 90th percentile confidence interval yielded ± 0.5 dB using 30 words (1 word = 5 bits) for the SNR estimate. In terms of word error probability estimation, this gave a 90th percentile of 0.5 in log error rate at $P_e = 10^{-5}$ and worse at lower error rates. Naturally, more samples would better this performance.

Smith and Thomas (1989) described an adaptive meteor burst communications

modem which used signal to noise ratio estimation to rapidly evaluate link quality. The method used 40 samples of the decision variable to form an estimate which was accurate to within about 1dB.

The weakness with this method is that degradations to digital communications system performance may be totally unrelated to SNR, for example, because of non-Gaussian noise (eg. impulsive atmospheric, thermal, man-made), imperfections (eg. timing jitter, AGC), fading (eg. multipath) and other distortions.

2.3.5 Specialised Monitors

2.3.5.1 Error Estimator for FSK Over Fading Channel

Hingorani and Chesler (1968a) described a monitor for FSK transmissions over fading channels which estimates the probability of error from estimated statistics of the outputs of square-law detectors. The outputs from the square-law detectors in a differential binary FSK demodulator are the receiver decision variables. These variables are important because the largest of the pair is used to decide whether the output bit is a mark or space.

The technique estimates a number of statistics of the decision variable; β which is the average power ratio of the specular (deterministic) component of the received signal to the random component; r the ratio of the variances of active to inactive matched filters; ρ the normalised correlation coefficient between the outputs of the two matched filters.

The error probability estimate of the method was stated to be asymptotically unbiased. The variation in the estimate due to sampling a finite number of values is also examined and full details are given on how to calculate the required measurement time. Measurement time is a function of β, r, ρ and another parameter, α the ratio of the average power in the fading component of the inactive channel output to the total average power in the inactive channel output. The parameter α is important because independent measurements of the noise occur at the signalling rate whereas independent measurements of the signal occur at the fading rate which is usually much slower. An approximate expression for the measurement time required to estimate P_e, with variance σ^2, fade

rate B_f and parameter Q was:

$$T_m = \frac{Q}{B_f \sigma^2} \qquad (2.1)$$

As a measurement device, the measurement time needed to yield an estimation error within some desired limits over all possible conditions is of great interest. According to Maslin (1987:p78) a fade rate of around 1 second is typical for a fading HF channel and a normalised variance of 0.1 was considered a reasonable estimation variance according to Hingorani and Chesler (1968a); (ie. σ=0.316 giving a 95th percentile confidence bound for a Gaussian distribution of $\pm 2\sigma = \pm 0.6$ in log error rate). For the case of no diversity, the graphs yield a maximum value of Q=250. Thus a maximum measurement time of around 2500 seconds (42 minutes) would be needed to estimate the error probability to 10^{-4} for a 75 bits/s link (compare with 22 minutes by sending known data). This measurement time is prohibitively long, but might be made more acceptable if estimation accuracy is relaxed to σ=1, giving a ±1 order of magnitude 95th percentile error and about 4 minutes measurement time. This relaxed degree of accuracy could only be used as a coarse indicator of link quality.

2.3.5.2 Adaptive Channel Estimator

Jankauskas et. al. (1980) described planned digital monitoring techniques for a digital PCM/TDM network for the U.S. Defence Communications System. An "Adaptive Channel Estimator" (ACE) was described to estimate channel quality information. The ACE algorithm described by Jankauskas (1976), and Sunkenberg and Jankauskas (1976) assumed that the channel may be modelled as a tapped delay line (a channel with memory). The channel model used the transmitted data sequence as input, with each delay element output stage multiplied by a gain factor and all outputs summed together. Lastly, a DC level and Gaussian noise was added to form the received signal. An adaptive equaliser was devised based on a Mean Square Error (MSE) algorithm. Estimation of the bit error rate was achieved by using the tap gain values on the adaptive equaliser. The calculation used was not disclosed (proprietry information) but is believed to be as follows:

The probability of error performance of a linear MSE equaliser may be upper-bounded by a function of the noise variance and *the sum of tap gain values* as given by Proakis (1989:p584:eq(6.4.105)). Thus the error rate may be calculated. Jankauskas [38] states that a "modification to the truncated pulse train approximation" is used. Proakis gives seven references on methods of calculation approximations for symbol error probability when there are many taps in the equaliser. The performance of the ER estimation scheme was not given.

2.3.5.3 Pilot-Tone PSK (PT-PSK) Monitor

Hingorani and Chesler (1968b) described a performance monitoring method for a PT-PSK system, where the pilot tone was transmitted for a phase reference, over a fading radio channel. A relationship for the probability of error conditioned on the noise variances of the outputs of the information and pilot tone matched filters was derived. This relationship was formed on the basis of a binary (coherent) PSK signalling scheme in a flat Rician fading channel with AWGN and diversity. Making a number of simplifying assumptions (ie. high SNR's etc..), the average probability of error turns out to be approximated by:

$$\bar{P}_e \cong \frac{1}{2^L N L! g} \sum_{i=1}^{N} (y_i)^L \qquad (2.2)$$

Where:
- L = Order of diversity.
- N = Number of samples.
- y_i = Estimated sum of the square of noise components.
- g = Estimated average of the fading envelope.

The performance of the scheme for flat fading and additive Gaussian noise showed an estimation bias, which became vanishingly small at high SNR values. The bias was largest for higher orders of diversity and at worst ranged from 0.5 in log error rate for L=8 to about 0.1 in log error rate for L=1. The number of samples required to obtain a reasonable estimation variance was demonstrated to be small for the log-normal and Gaussian noise cases. Generally, for the same accuracy as direct error counting, fewer samples were required.

The performance of the technique for other forms of channels was not

detailed, nor was the situation where either AWGN *or* log-normal noise may occur. Also, the technique relied on the assumptions of coherent BPSK signalling with a pilot tone.

2.4 Non-Parametric Estimation Methods

2.4.1 Pseudo-Error Estimation Methods

The concept of a "pseudo error" seems to have been coined by Gooding (1968). A pseudo error is an artificial digital error. Pseudo errors may be formed in two basic ways, the first by adding noise in a parallel receiver path, the second, by introducing a decision threshold in the receiver that is lower than the actual decision threshold. The result of either technique is a stream of bits usually referred to as a pseudo error stream. In this stream there will be more "1's" or "pseudo errors" than in the actual data stream, and some correlation will exist between the number of pseudo errors and the number of real errors. Several authors including Brown (1976) and Feher (1977) have described the merits of pseudo errors as an end in themselves, but this treatment will concentrate on how pseudo errors relate to *real* errors.

2.4.1.1 The Additive Noise Method

An error rate monitor may be constructed by adding noise to the received signal and demodulating it in parallel with the unaltered received signal. This method has been described by Keelty and Feher (1978) and Takenaka et. al. (1980). The output of the parallel demodulator (processing the signal with added noise) is referred to as the pseudo-data stream. More errors will occur in the pseudo-data stream. By direct comparison of the pseudo-data stream and the actual data stream, a pseudo error rate is deduced. This pseudo error rate is defined in the same way as error rate (3.1). A lookup table may be used to obtain an estimate of the true error rate from the pseudo error rate.

The addition of noise to the signal degrades the SNR by a known amount so that an error rate can be measured quickly. The measured rate can then be used to estimate the actual error rate, if the calibration is known. This calibration curve will be unique to the channel type currently being observed. If the type of the channel disturbance is unknown, or may take

a number of different forms, a unique calibration (and hence error rate) cannot be chosen.

A monitor was described by Popovici (1987) for the case of an AWGN channel. Gaussian noise was added to the system, the pseudo error rate measured for a number of points and a Gaussian regression algorithm applied to estimate the true error rate. This method appeared to be accurate and robust to minor measurement error. The major drawbacks of the method are the computational complexity of the Gaussian regression algorithm, which requires a large number of iterative steps and the fact that the regression is applicable to a system degraded by AWGN *only*.

2.4.1.2 Lower Threshold Methods

The lower threshold methods require access to the receiver decision statistic (or decision variable). This is the signal in any demodulator/detector just prior to making the data decision. Thresholds are set at points *lower* than the actual data decision threshold level. This concept is illustrated in Figure 2.2, for the example of a QPSK modulation scheme. The decision thresholds are at $\phi=\{0,\pi/2,\pi,3\pi/2\}$ and the lower thresholds are illustrated. This method is applicable for low to very high rate signalling systems as demonstrated by Hogge (1977) who described electronic circuits to perform lower thresholding for a 90Mbits/s data rate with 8-PSK modulation.

The majority of research has been inspired by Gooding (1968), although the use of lower-threshold decision devices to estimate error rate may be traced back to Rush (1965). Rush described using two lower-threshold decision devices on a non-coherent FSK detector. By observing the ratio of the error rates between the two devices, information about the channel performance could be obtained leading to estimation of the error rate.

Linear Extrapolation Method

Gooding's work was influential on the Additive Noise Method and his Performance Monitor Unit (PMU) was the subject of a number of later papers, reviewed here. Figure 2.3 shows an example of the PMU. A number of thresholds are set which are below the actual decision threshold. A

counter is incremented each time the decision variable exceeds a given lower threshold. The count registered over the number of total units (bits, symbols, etc...) sent, is referred to as a pseudo error rate. From this range of pseudo error rate values, two (or more) pseudo error rate values are chosen and an extrapolation process applied to form an estimate of the actual error rate.

Gooding's hypothesis was that for a given type of modulation and probability distribution of the noise and fading processes, it is possible to define a threshold parameter K such that the logarithm of the pseudo error rate P_p is a linear function of K for a wide range of values of P_p.

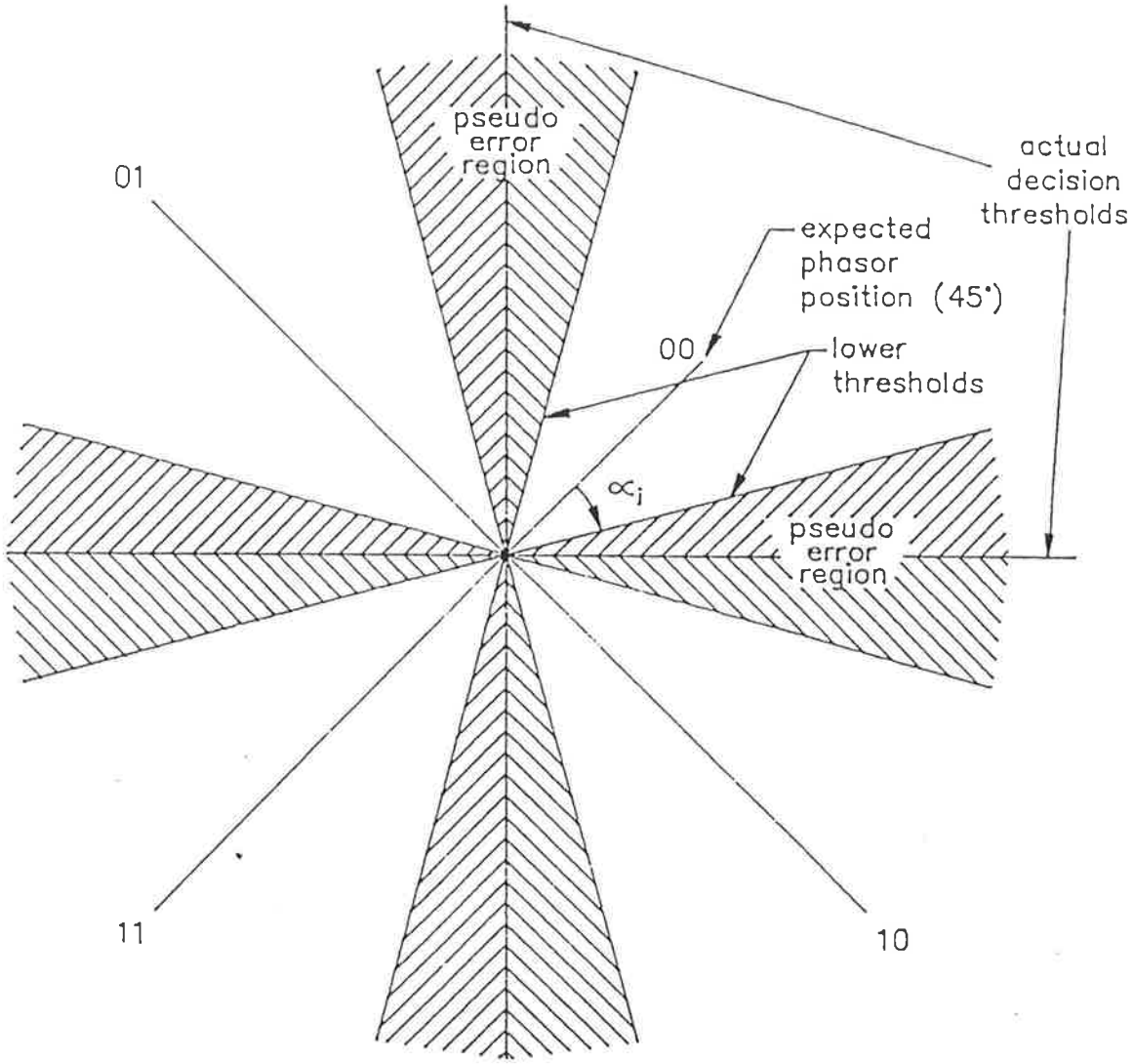


FIGURE 2.2. Example of lower thresholds for QPSK modulation.

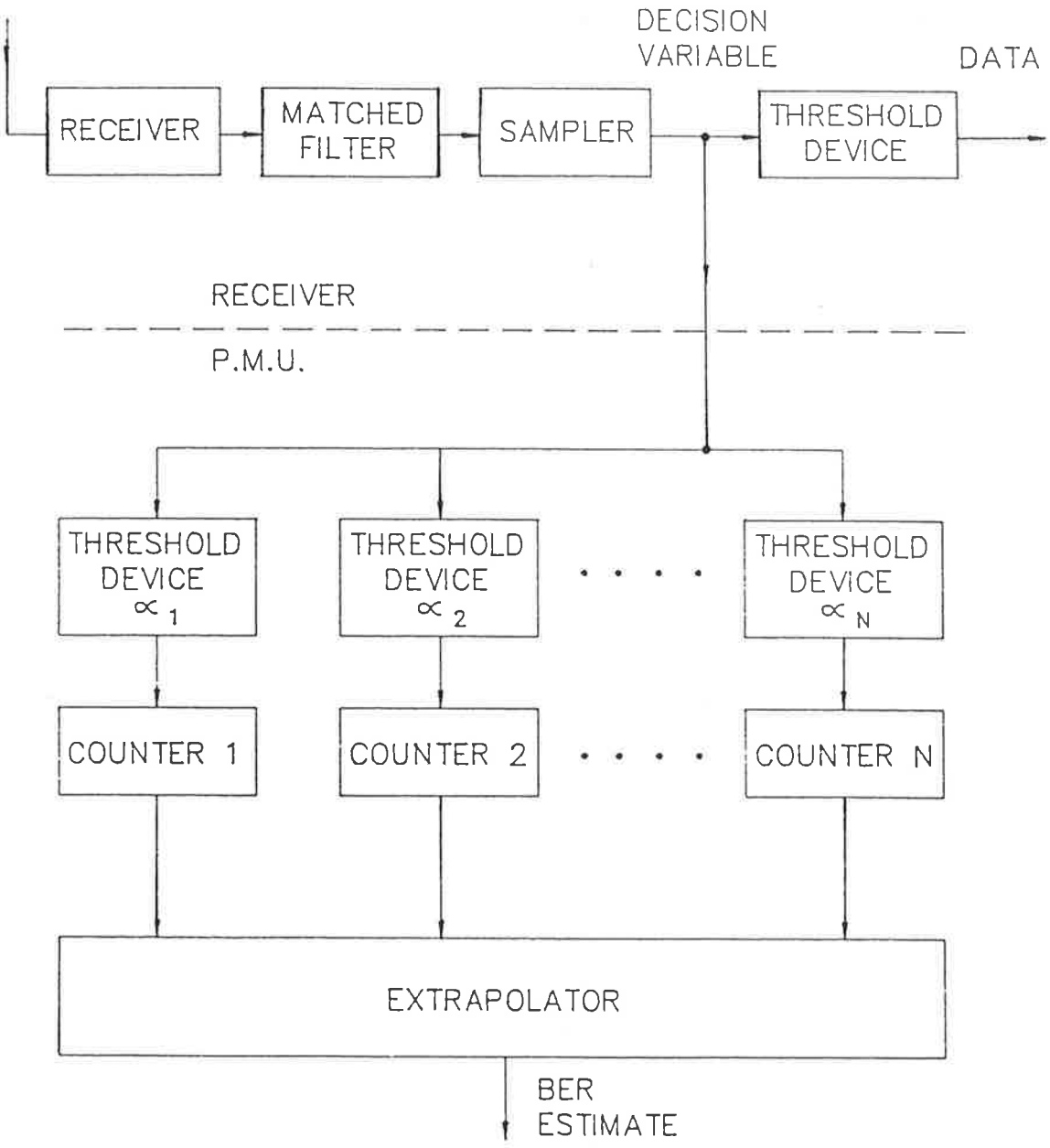


FIGURE 2.3. PMU shown attached to a receiver.

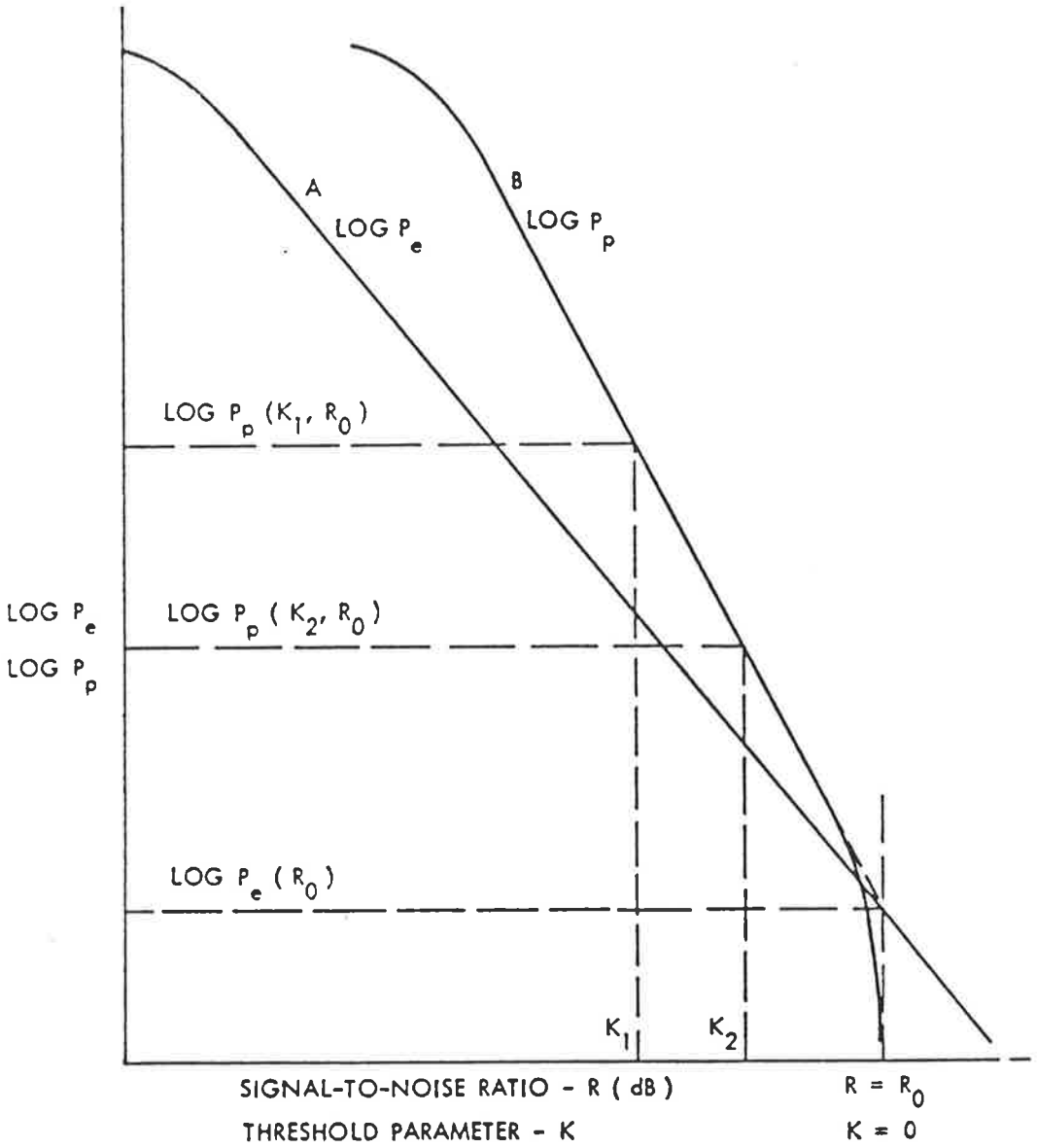


FIGURE 2.4. PMU extrapolation process.

This linear extrapolation in the log probability versus threshold parameter (K) axis is taken to the point K=0 (on the x ordinate), which is supposed to correspond to the logarithm of the *actual* ER on the y axis. This extrapolation process is depicted in Figure 2.4 and consists of solving the equation (2.3) as follows:

$$\log \hat{P}_e = \frac{K_2 \log P_p(K_1) - K_1 \log P_p(K_2)}{K_2 - K_1} \qquad (2.3)$$

Gooding further claimed that the method is *non-parametric*, that is, the linearity of the pseudo-error rate versus threshold characteristic appears to be only weakly dependent on the form of the probability distributions of the fading process and the additive noise. That claim was largely unsubstantiated, since the PMU performance was tested only for multipath Rayleigh fading with additive Gaussian and log-normal noise. This is supported, in the case of FSK transmission with Rician fading and AWGN, by McLarnon (1977) who claims that it may be shown that the linearity of the P_p versus K characteristic is almost independent of the specular to fading power ratio, the order of diversity and the SNR. This does not however, supply convincing evidence that the technique is not parametric and the results given also showed quite poor performance in some situations (Gooding:table I,p385). For example, estimated error rates for the fading with Gaussian noise case (M=8) showed that the mean of the error rate estimates was one order of magnitude worse than the actual error rate, and had a standard deviation of 0.203 in log error rate. Kreyzig (1983) states that if this distribution were Gaussian, 95% of estimates would lie within $\pm 2\sigma$ of the mean. In that case, 95th percentile confidence limits would be one half an order of magnitude either side of the estimated value. Similar results were obtained for the log-normal case (M=8), the estimate again being around one order of magnitude worse than the actual, but with a standard deviation of 0.59 (Gaussian 95th percentile of 1.2 orders of magnitude either side of the estimate).

Weinstein (1971) examined using an extrapolation monitor for estimating the error rate for a υ^{th}-order exponentially-based decision variable distribution, which would include the exponential distribution (υ=1) or

the Gaussian distribution (υ=2). For the case of the exponential distribution the log of the probability of exceeding a threshold x, is shown to be proportional to the threshold value x and hence the "linear" extrapolation described by Gooding (1968) is shown to be exact. He further demonstrates that for υ increasing beyond 1, a corresponding increase in bias (or offset) between the PMU estimate and the true error rate occurred, but would allow the device to be practically usable for the exponential and Gaussian cases. The extrapolating monitor was tested with a binary modulation scheme on a AWGN channel. Unfortunately, the results given by Weinstein (in Fig.5) did not examine the statistical variations in the estimate. It was also stated that the observation time required by the extrapolation monitor yielded a saving of 3.8 times for the exponential distribution and 18 times for the Gaussian distribution.

Hammond et. al. (1973) studied the relationship between P_p and P_e for the Gooding extrapolation monitor on PSK links. For the case of a AWGN channel with a single coherent interferer, it was found that if the Carrier to Interference Ratio (CIR) was allowed to vary, P_p was not a single valued function of P_e. This finding was echoed by Leon et. al. (1975). The impact of this result is if the link may adopt more than one form of channel statistic, the extrapolating PMU has no way of discerning which statistic is responsible for the disturbance and hence *which* P_p curve to apply to give P_e. It was indicated by Leon et. al. (1975) that use of a number of values of β in a composite indication (that is, use of more than two lower threshold values) may have reduced the confusion experienced by the monitor; but this approach was taken no further.

Newcombe and Pasupathy (1980) studied pseudo error monitor schemes for partial response systems where the monitoring unit was located on a parallel path to the main receiver with known applied amounts of degradation. The degradations included threshold modification, sampling offset, amplitude distortion and phase distortion. In each type of monitor, the Gooding linear extrapolation (3.3) was used to convert a pair of pseudo error rates into an actual error rate estimate. However, in place of K given in equation (3.3), the actual distortion measure was used (ie. K was replaced with the distortion amplitude, the phase distortion level, the sampling offset time or the lower threshold level for whichever type of monitor was to be used). Thus, the logarithm of P_p was assumed to

be a linear function of the distortion value, *regardless* of the type of distortion. Only one form of degradation was examined for each PMU, so that the technique did *not* have to cope with attempting to discern what form of degradation was *causing* errors to occur. Considering this, it was not surprising to find the results from Hammond et. al. (1973) and Leon et. al. (1975) reinforced and that for the case of sinusoidal interfering signals, serious estimation errors were noted; pointing to the realisation that probability of error depends on more than a pair of pseudo error values.

Leon and Kitahara (1973) examined the performance of a linear extrapolation monitor for a specific type of QPSK receiver in the presence of AWGN by simulation. In addition to the AWGN channel, other degradations were introduced, including a constant carrier phase error of 10 degrees, Gaussian random jitter with a standard deviation of 10 degrees and a constant carrier frequency offset of 200Hz at 70MHz IF frequency. The performance graphs given were evaluated for only *one* point at a given SNR and a line of best fit applied through these points. The points gathered in some of these figures showed a significant amount of scatter (particularly the last three figures shown) so as to make clear evaluation of PMU performance questionable. The lack of statistical gathering was possibly due to the prohibitively high expense of conducting the digital computer simulations (in 1973) which the authors emphasized. In addition, the authors stated that if there were more than one degradation variable, (operating conditions were variable) the PMU estimate of error probability was ambiguous ie. there was no way to know which degradation (or combination of degradations) was responsible for causing errors.

Davies and Sherif (1986) have examined the Gooding pseudo-error extrapolation method applied to an optical communications system. As stated by Cherin (1983) and Radcliffe (1991) these systems typically exhibit signal-dependent noise. As a result, for a fixed pair of lower thresholds, the technique could not accurately extrapolate over the wide range of mean signal levels commonly experienced. The conclusion from simulations was that the method gave a good estimate for low signal to noise ratios only.

Sunkenberg and Jankauskas (1976) examined the *mean* performance (estimation

bias) of a linear pseudo error rate extrapolator for binary signalling with a channel distorted by two equal amplitude discrete linear channel tap gains and AWGN. The method performed poorly over the range of varying amounts of distortion, but might be adequate if the SNR was always below 14dB. For SNR values above 14dB, the extrapolation performance degraded very rapidly. Sunkenberg (1978) examined the statistical *variance* of the error rate estimation. The BER estimator performance, was found to have a standard deviation of around 0.4 in log error rate (or a 95th percentile of 0.8 orders of magnitude) up to a SNR of 14dB.

A Further Simulation

Pseudo error rate values are random variables and are susceptible to variation due to counting statistics. Variation in these pseudo error rate values relate *directly* to variations in actual error rate estimates and thus a comprehensive study of the performance *must* include the effects of counting statistics. To demonstrate the PMU performance with regard to counting statistics, a simulation for coherent QPSK modulation in an AWGN channel was performed by Scholz et. al. (1990). Ten counters were used and the design was linearised for an error rate of 10^{-6} using the procedure described by Gooding (1968). Figure 2.5 shows the 95^{th} percentile of the difference in log of the actual error rate (or probability of error) and the log of the PMU estimated error rate for 1000, 4000 and 16000 samples.

It is clear from figure 2.5 that the performance of the PMU is poor. Part of the reason for this poor performance is due to the fact that the first counter value taken may be as low as five, (which was the value stated in Gooding's algorithm) but is not truly a sufficient number of counts. A large proportion of the error, however, is believed to be due to the basic inaccuracy of the extrapolation process, which uses only a small amount of the available information.

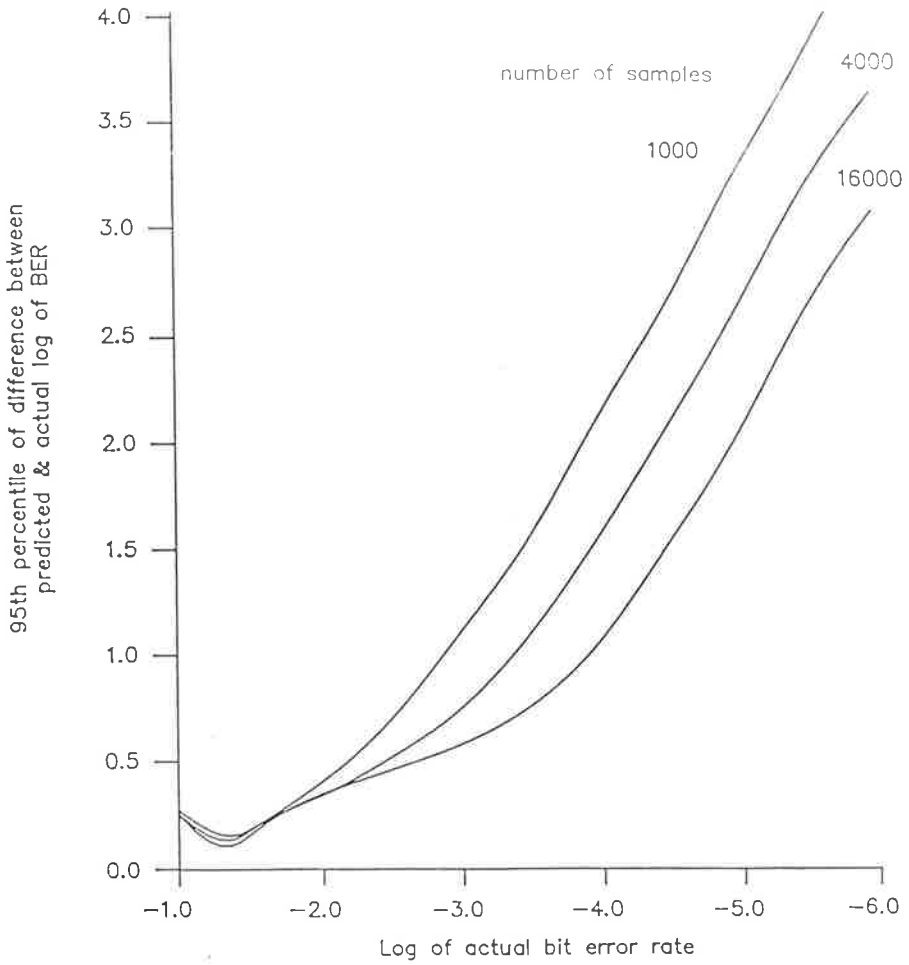


FIGURE 2.5 Simulation of the performance of Gooding's PMU technique for QPSK modulation in an AWGN channel.

Summary of Extrapolation Techniques

These methods are all deemed inadequate because they:

(i) Apply to only one type of noise/interference.

(ii) Are unbiased for exponentially-distributed noise only.

(iii) Require a large number of samples or are low in accuracy.

Other Techniques Based on Lower-Thresholds

Keelty and Feher (1978) plotted graphs showing the relationship between pseudo error rate and probability of error for several forms of degradation. The degradations included carrier tone interference with AWGN and non-Gaussian noise with AWGN. It was concluded that a pair of pseudo error rates may be related to the BER by a simple multiplication

factor. Unfortunately, the results presented were inconclusive since the graphs showed error rate versus SNR (for the AWGN) with only *one value of degradation parameter*. The interfering carrier tone case showed the curves for AWGN only and AWGN with a single Carrier to Interference Ratio (CIR) of 20dB. The relationship for any *other* CIR was not given. Likewise, for the non-Gaussian noise case graphs for AWGN only and AWGN with Non-Gaussian noise of 20dB were given. The relationship between P_p and P_e for any amount other than 20dB of non-Gaussian noise was not given. The hypothesis of error probability being a constant scaling factor multiplied by a pseudo error rate value was therefore untested for a range of varying amounts of degradation. In conclusion, the PMU was not shown to work with more than one type of noise simultaneously.

Kostic (1989) examined the relationship between pseudo error rate and actual error rate analytically for *M*-ary PSK system affected by Gaussian noise and co-channel interference. The equations and graphs derived were given as an aid to designing a monitor for a given set degradations. It was found that for Interference to Signal Ratio (ISR) values less than -20dB, for any threshold angle (γ), a value of error probability could be estimated directly from the pseudo error rate. It was also found that a unique estimate of the probability of error was *not* possible from a single P_p value due to the fact that there were two values of P_e that could be chosen (ISR $\geq$ -15dB, $\gamma \geq 15$ degrees and ISR $\geq$ -10dB, for any γ).

An interesting method was described by Brown and Wilson (1974). This technique sampled the detector eye pattern (actually the decision variable) and performed a threshold function to determine the percentage of decision vectors lying within preset boundaries. By determining the percentage of vectors which lie in the threshold region, estimates of the channel degradation and resulting error rate were obtained. The size of the region correlated to bit error rate was empirically determined, details of this test were not given. The performance of the technique was not detailed.

"Alpha flunk" is the term in a U.S. Army (1979) installation manual, used to describe the count of data bits that are below a set quality in a given number of total bits. The concept was used as a quality assessment for PSK satellite links. In a PSK modem (at least for Harris MD-921/G or

MD-1002/G modems), during any bit period, an integrator (in the demodulator's matched filter(s)) will be charging to a value proportional to the quality of the incoming bit. The level to which the integrator charged is quantised into a 2 bit binary word (that is, 0-25% = 00, 26-50% = 01, 51-75% = 10, 76-100% = 11). The most significant bit (MSB) of the two bits represents demodulated bits which are more likely to be in error. The MSB's are accumulated (one count each time the MSB = 0), since this represents the integrator charging below 51% of its desired steady state value. The ratio of the accumulated counts to the total number of bits is defined as the "alpha flunk". The "alpha flunk" was then plotted versus signal to noise ratio on the same graph with error rate performance of the modem and could be used to estimate the link error rate.

The "alpha flunk" idea may be viewed as a coarse, but simple method of pseudo error rate measurement, since it uses the receiver decision variable but quantises the value into only four levels (two bits). The method relies on the fact that the channel is AWGN only. If scintillation fading or jamming was present a different probability of error curve would need to be used to provide the calibration of "alpha flunk" reading into BER.

Sunkenberg and Jankauskas (1976) examined the performance of several pseudo error rate based schemes. The paper considered binary signalling with a channel distorted by two equal amplitude discrete linear channel tap gains and AWGN. The test examined and compared four methods of estimating the error probability by:

(i) Linear extrapolation of a pair of pseudo error rates.
(ii) Quadratic extrapolation of three pseudo error rates.
(iii) Using a constant k (amplification), where $P_e = kP_p$.
(iv) Using an adaptive channel estimation algorithm (called ACE).

The ACE algorithm has been treated in the section with parametric estimation techniques. The linear extrapolator has been covered in the previous section. The performance of these techniques in terms of statistical variation in the error probability estimates was not stated. The results representing *mean* estimation accuracy. Since variation in pseudo error rate values relate *directly* to variations in actual error

rate estimates a comprehensive study of performance should have included the effects of counting statistics.

A quadratic extrapolation was also investigated by Sunkenberg and Jankauskas (1976), which attempted to fit measured points onto a quadratic curve. The method showed reasonable performance with mean values of the error rate estimate (bias) within 0.6 in log error rate terms over the range of degradations.

The last algorithm examined by Sunkenberg and Jankauskas (1976) was direct amplification (multiplication of the pseudo-error rate by a constant). The conclusion was that the algorithm was unsuitable for accurate error rate estimation, but might be used as a gross indicator of performance. Snyder and Hersey (1984) also described an application based on the direct amplification method for use on INTELSAT satellite Time-Division Multiple-Access (TDMA) links. Pseudo error rates were derived for coherent QPSK modulation in an AWGN channel with and without non-linearities, for several threshold levels. The effect of counting statistics on the accuracy of error probability estimates was examined for the AWGN channel. A ±21% tolerance on probability of error estimation was stated with 95% confidence using 10^6 symbols. In deriving these figures, *estimation bias* was ignored, which has been demonstrated by Weinstein (1971) to be a significant consideration in this type of monitor. Tesla (1988) and (1989) also described an amplification type monitor. This monitor was derived for a differential BFSK modulation scheme operating over a Rician fading channel. It was stated that for this particular situation, the monitor performed well.

Brown (1976) examined confidence limits for the uncertainty of counting pseudo error rate values in a binary AWGN channel. Translation of pseudo error rate value(s) to an actual error rate was not considered. The number of samples used to obtain a pseudo error rate value was examined to accept or reject the hypothesis with a given confidence, that some channel SNR was present. This same approach could have been used to accept the hypothesis whether a particular error probability was present. However, the approach is of limited value as it assumed Gaussian distribution of the decision variable. Also, the formation of confidence intervals for other distributions (particularly with dependent samples) is generally

intractable; a notable exception being the treatise by Davis (1988).

Milstein (1976a & b) described a method of bit error rate estimation based on Gumbels (1966) "Extreme Value Theory" (EVT). The method assumed the distribution of the decision variable(s) may be expressed in an "asymptotic form". An "exponential type" asymptotic form was assumed for the tail of the decision variable distributions.

Consider a binary signalling system with states ±1. Samples *below* the decision threshold were grouped together in the -1 class and placed into N groups of n samples per group. The largest sample of each of the N groups was used to find the sample mean and variance for the N maxima. Using the same procedure on minimum samples from groups of samples detected *above* the threshold, the sample mean and variance of the N minima were found. Thus, sample statistics of the *extreme* values only, were collected. These means and variances were substituted into four equations to (numerically) determine the error rate estimate.

The error probability estimates from this technique were stated to be biased estimates. The size of the bias and the estimation variance was not given. In addition, the calculation required appeared to be intensive since it could only be solved numerically (by iteration) and the number of required iterations was not disclosed.

Schwartz and Richman (1968) also described error probability estimation based on EVT. For the Gaussian distribution 10,000 samples allowed error probability estimation as low as 10^{-5} with a ± 50% error. The confidence level for this performance was not stated!

Guida et. al. (1988) examined the performance of EVT and Generalised EVT (GEVT) methods for estimation of low error probabilities for Gaussian and Weibull distributions. The GEVT method introduces another variable, υ (in the case of EVT υ=1). This variable allows for greater accuracy in the estimation process (ie. faster convergence) than for EVT. Comparisons showed that the EVT estimators were more or less equivalent to classical counting estimators whereas the GEVT estimators were consistently more efficient (particularly for the Weibull distribution with $F(x)=1-\exp(-x^3)$,$x \geq 0$.

2.4.2 Pilot-Tone Channel Estimator

This is a technique proposed by Betts et. al. (1970) for HF channels with a 3kHz bandwidth.

The concept relies on transmission of a diagnostic pilot tone and measurement of its phase. This procedure is carried out repeatedly and whenever the phase exceeds an arbitrarily set threshold (θ), the event is stored to give a measure of the "phase error" count associated with the pilot tone. The relationships between the "phase error rates" and data channel error rates for various conditions were calculated.

The practical performance of the technique on a live channel was given by Betts et. al. (1975). The modulation scheme used was wide-shift FSK (850 Hz) with data tones at 800 and 1650 Hz with the additional pilot-tone at 1200 Hz. The pilot-tone "phase error rate" in this situation was a function of the signal to noise ratio, the phase comparison interval τ and the threshold setting θ. Optimal values of τ=10ms and $\theta=\pi/8$ were chosen so that the "phase error rate" should directly translate into data error rate. This optimisation or "trade-off" between two vastly different error rate conditions (flat fading versus AWGN channel) was stated to be "extremely demanding". Surprisingly, the measured performance of the technique appeared to be reasonably good over a wide range of conditions. However, statistical correlation figures were not given and there appeared to be significant estimation bias in some situations (notably in fig. 13).

2.4.3 Importance Sampling

Importance sampling is a technique for generating error events more often so that the variance in the estimation of the probability of error is reduced. In the simulation of digital communications systems, as examined by Jeruchim (1984), and Davis (1988) this has generally been achieved by modifying the input sequence (the sequence of received decision variable values). The input sequence may be applied to a processor which in general may be non-linear and have memory. This processor produces "important" samples whose average is related to the error probability by some function. For the case of a Gaussian received sequence, this function is a simple scaling factor. Jeruchim showed that a useful

reduction in the variance of the error probability estimate with respect to the Monte Carlo method is achievable in the Gaussian case.

This method does not seem to have been applied to the real-time estimation of error probability, but only as a "speed up" for calculating error rate (when the error rate is very small) in digital simulations.

The advantages of this method, if it were used in a real time estimation system, is that it does not require the transmission of a known data sequence and that a faster result than the Monte Carlo method may be achieved. The main disadvantage relates to the assumption that the form of the channel distribution is known so that the function can be calculated.

2.4.4 Density Deviation Monitor (Link Failure Monitor)

Leon et. al. (1974) proposed a monitor to give a binary decision to indicate whether or not a fault had occurred on a digital communications link. The method required that a collection of amplitude windows be applied to sort samples of the decision variable into amplitude ranges, or class intervals. A count of the number of inputs in each amplitude range was collected. This collection of counts in each interval provided a time-varying histogram of the decision variable cumulative density function. When the link was operating normally, the monitor was put into a measurement mode which stored in memory, the histogram for normal operation.

The method involved deriving a Density Deviation Estimate (DDE), defined by (2.4) and comparing this value with a threshold value to decide whether or not a link failure had occurred.

$$DDE = \frac{1}{k} \sum_{i=1}^{k} | A_i - B_i | \qquad (2.4)$$

Where:

A_i = Number of averaged counts in the i^{th} class interval from current data.

B_i = Number of averaged counts stored from normal operation for

the i^{th} class interval.

k = Number of class intervals in the estimating histogram.

The test (2.4) was chosen for its simplicity and intuitive nature. The monitor was examined in terms of its probability of detection P_d (if a fault does exist) and probability of false alarm (ie. indicated a fault when none occurred). These two probabilities were derived as functions of the threshold value, the number of class intervals and the number of samples, to allow the device to be designed.

The performance of the device was *not* evaluated by Leon et. al. (1974) or elsewhere in the literature, yet it appeared (to the author) to be the most promising approach in evaluating link performance. The reason for this observation was due to the fact that the method examined *more* than a single pair of pseudo-error counts (or counts from two intervals). The method was actually an *estimator* of the cumulative density of the decision variable. Armed with this estimated distribution a comparison of the measured and the expected *distribution* was made to examine link condition.

2.5 Chapter Summary

The choice of a particular error monitoring technique for a specific application depends largely on the nature of that application. For specific uses where the type of link is fixed (ie. known class of channel type at any measurement time) parametric techniques may be suitable. Unfortunately, a large number of applications have requirements that can *in no way* be met by using the techniques reviewed. For example, to take accurate measurement of error rate whilst maintaining 100% throughput on a time-varying channel (eg. HF radio link, mobile satellite) is not possible. If the type of channel statistic at the time of measurement is unknown, only the direct methods (reduced throughput) can be used to give accurate results in all but a few special cases.

Table 2.1 summarizes the features of the various categories of error monitoring schemes reviewed, which may be applied to *any* digital communications link. Those methods which are specifically suited to a single application or class of applications have not been included. It is quite clear from the table that no single technique scores highly in all desirable aspects required of a generic error monitor.

TYPE	SPEED	PRECISION	THROUGH-PUT (%)	REPEAT-ABILITY	PERFORMANCE FOR REAL CHANNELS
DIRECT Monte Carlo Method	slow	excellent for *large* samples	0	subject to counting statistics	excellent
DIRECT Error Detect-ing Codes	very slow	good for *large* samples	50-100	worse than Monte Carlo	excellent but degrades for burst errors
PARAMETRIC	medium	poor to fair	100	poor	poor
PSEUDO-ERROR	fast	poor	100	poor	poor
IMPORTANCE SAMPLING	medium	fair	100	fair	fair

TABLE 2.1 Features of generic error monitoring techniques.

The parametric techniques by their nature, rely on accurate prior knowledge of the form of the channel statistics at the time of measurement to form a reliable estimate of error probability.

The pseudo-error techniques, although appearing at first to be promising have, after almost thirty years since their conception, proven to be difficult to design and inaccurate. As a result, they have had little impact as measurement techniques.

What is required is a *generic* form of error monitor which is fast, precise (for a modest number of samples), allows 100% throughput, is repeatable in its measurements and has very good performance on *real* channels. This aim has been achieved in a new link monitoring technique called the Extremely Versatile Error Rate ESTimator (EVEREST). As will be seen in the succeeding chapters, this technique will give not only provide error probability estimates, but also an indication of the *class* (or type) of channel which is currently under observation. It will be shown that this information can be used to form the basis of a novel performance optimization scheme.

Chapter 3

A NEW ERROR PROBABILITY ESTIMATION TECHNIQUE

Abstract

The review has shown the dearth of a *generic* form of error monitor which is fast, precise (for a modest number of samples), allows 100% throughput, is repeatable in its measurements and has very good performance on *real* channels. This aim has been achieved in a new link monitoring technique coined EVEREST (Extremely Versatile Error Rate ESTimator). As will be seen, this technique will estimate the *class* (or type) of link which is currently under observation in order to obtain the error probability estimate. In this chapter, the theory of the new technique is developed and analysed. The relative performance of different EVEREST structures are compared through simulation for binary signalling in an AWGN channel using a coherent detector.

3.1 Introduction

The probability of error is the most important measure of the performance accuracy for a digital communications system. The importance of this measure is witnessed by the quantity of research effort applied to the problem of its measurement as reviewed in chapter 2.

The new method of error probability estimation is coined EVEREST. The early concept (in particular, using the Weighted Least Squares test) is co-patented with Dr. Stephen Cook and Mr. Timothy Giles (Scholz et al, 1990b). Research solely attributable to the author involves all later material, including some theoretical analysis of the WLS test, Adaptive Bin Boundary and Maximum Likelihood tests. The versatility of the technique is in its applicability to all known forms of channel and modulation type, as will be demonstrated in subsequent chapters. It is the first truly *non-parametric* method of estimating probability of error. The reason for its non-parametric nature is that it examines the *decision variable(s)* in the demodulation/detection process of reception.

The review in Chapter 2 has made clear the fact that the characteristics of the decision variable(s) are solely responsible for the error probability of the link since these are the only variables used in the data decision process.

This research was initially inspired by Gooding (1968). Although Gooding's approach lacked precision and was dependent on channel statistics, the concept of extracting information from the received signal by using lower (more sensitive) thresholds remained attractive. It was recognised that Gooding's method was an attempt to identify the BER from points on the probability distribution of the receiver decision variable(s). However, information available about the *shape* of the distribution far below the decision threshold had been ignored. Additional information such as the modulation scheme employed and the class of channels expected which could also be used to assist in identifying the error probability, was also ignored.

Quite some way into this research (after the weighted least squares test was derived and tested on a live circuit) work by Leon et. al. (1974) was brought to my attention. Leon's research on the Density Deviation Monitor

initially ran very close parallels with the EVEREST technique and would have saved considerable effort had it been published in a journal or conference proceedings as was other work contained in that report. Leon realised the importance of the decision variable in evaluating link performance and that an estimate of the decision variable probability distribution could be made by employing lower thresholds. However, Leon was concerned *only* with identifying link failure and not link quality per se (see Chapter 2 discussion).

3.2 The EVEREST Concept

The technique measures the receiver decision statistic (or decision variable). The signal is usually the output of matched filters, located *just prior* to making data decisions. This signal is present in all digital communications receivers. The location of the decision variable(s) in a digital communications receiver is illustrated in figure 3.1.

The general form of the EVEREST is shown in figure 3.2. The device extracts features of the decision variable(s) and then performs a classification on these features (Van Trees, 1968). The classification is a matching process, which involves choosing a model whose describing statistical features *most closely resemble* the statistical features extracted from the decision variable. Each model has an error probability value (error rate) associated with it. Thus the error probability estimate for the link is taken as the error probability associated with the chosen model.

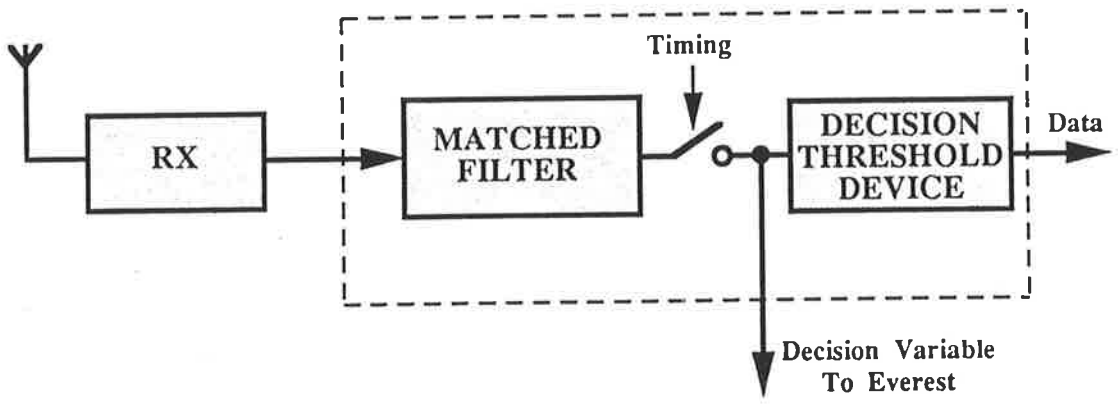


FIGURE 3.1. Digital receiver structure, showing the decision variable(s).

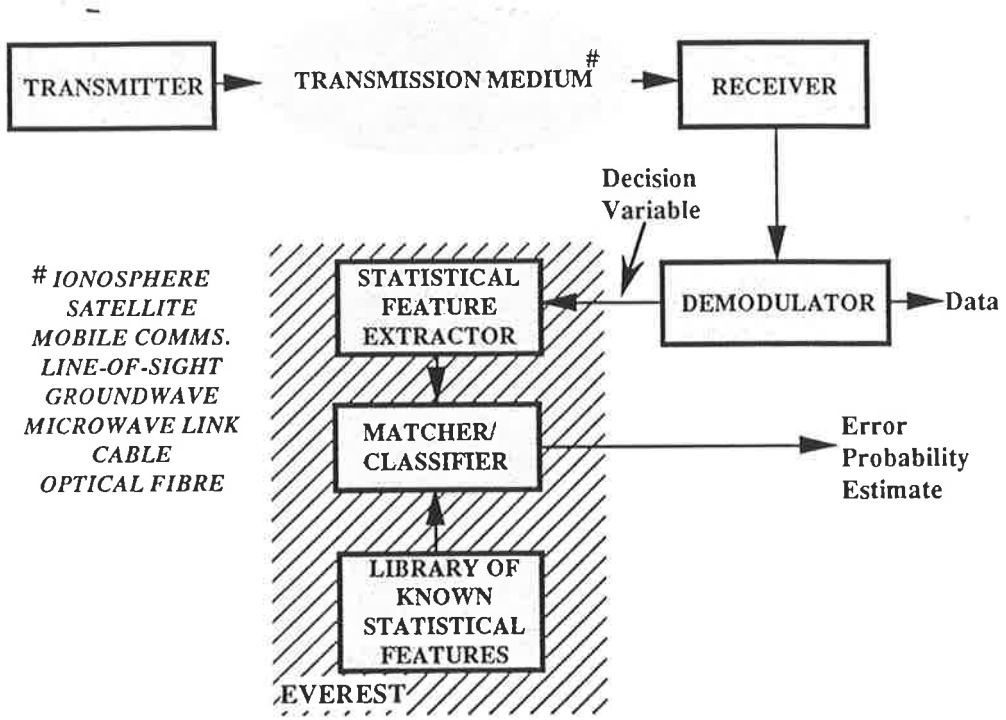


FIGURE 3.2. General form of the EVEREST.

It would be possible to estimate the error probability without using a priori "library" models by using some form of calculation (eg. extrapolation used in pseudo-error monitors). However, since the nature of disturbances is generally known for a given link, to assume unknown structure would ignore valuable available information and undoubtedly weaken the estimation accuracy. In all of the following studies, knowledge of the types of link model are incorporated in the library set, or as part of the test structure.

3.3 Sub-optimal Estimators

3.3.1 Weighted Least Squares Test

3.3.1.1 Derivation

Consider a feature extractor which consists of a set of lower thresholds covering the entire space of values that the decision variable can take. The first form of the EVEREST to be examined was based on an intuitive

idea inspired by Pearson's (1894) Chi-Squared Test for Goodness of Fit. The first part of the Chi-Squared test procedure involves forming a statistic $\bar{}(\chi^2)$ which is a measure of the "closeness" of two discrete distributions (histograms) as follows:

$$\chi^2 = \sum_{J=1}^{M} \frac{\left[u_j - e_j \right]^2}{e_j} \qquad (3.1)$$

Where: M = Number of bins (or regions).
u_j = Measured count in bin j.
e_j = Expected count in bin j.

The value of χ^2 approaches zero as the two distributions become more alike.

Instead of using this test as a hypothesis test for a given percentile confidence of distribution **u** resembling distribution **e**, it may be used by considering the fact that the statistic (χ^2 in equation (3.1)) is a weighted mean square error. The weights on the squared differences are $1/e_j$. The Weighted Least Squares test then, is to choose the error probability given by model r for which $\vartheta(r)$ is the minimum value of $\vartheta(k)$, k=1,2,...,L; where:

$$\vartheta(k) = \sum_{J=1}^{M} \frac{\left[u_j - e_{jk} \right]^2}{e_{jk}} \qquad (3.2)$$

Where: u_j = Measured count in bin (region) j.
e_{jk} = Model count for bin j, model k.
M = Number of histogram bins.
L = Number of library models.

The test given in (3.2) should not be used with integer values for the model counts e_{jk} because as can be seen from (3.2), as $e_{jk} \rightarrow 0$ the value of $\vartheta(k)$ becomes large and is coarse for $e_{jk} \leq 5$.

Note that the expected number of counts in bin j, for model k is given by:

$$e_{jk} = N P_{jk} \tag{3.3}$$

Where: N = Number of samples taken of the decision variable.

P_{jk} = Model probability for bin j, model k.

The test described in equation (3.2) may be simplified by substituting equation (3.3) and simplifying, to give:

$$\xi(r) = \min_{(\text{over } k)} \left\{ \sum_{j=1}^{M} \frac{u_j^2}{P_{jk}} \right\} \tag{3.4}$$

Equation (3.4) describes the preferred form of the WLS test.

The squared value of the count in each bin is weighted by the expected bin probability for model k. This weighting helps the test by giving greater significance to terms when the expected probability is small.

3.3.1.2 Performance of the WLS Statistic

Consider the expected value of the estimate $\xi(k)$:

$$E[\xi(k)] = \eta = \sum_{j=1}^{M} \frac{E\left[u_j^2\right]}{P_{jk}} \tag{3.5}$$

Now, u_j is binomially distributed, with expectation NP_j and variance $V[u_j] = NP_j(1-P_j)$, so:

$$E\left[u_j^2\right] = V[u_j] + \left(E[u_j] \right)^2$$

Giving:

$$\eta = \sum_{j=1}^{M} \frac{NP_j + N(N-1) P_j^2}{P_{jk}} \tag{3.6}$$

Let γ be some parameter describing the system. The library models have

bin probabilities $P_{jk}(\gamma)$ where γ spans all possible models. The actual observed system has bin probabilities $P_j = P_{jk}(\gamma_0)$.

To examine if the expected value, η is a minimum when $\gamma = \gamma_0$, let:

$$\eta' = \eta(P_{jk}) + \mu \sum_{j=1}^{M} P_{jk} \tag{3.7}$$

Using the method of Lagrange multipliers (Leithold, 1981). Where μ is the Lagrange multiplier parameter. Differentiating (3.7), where η is given by (3.6) gives:

$$\frac{\partial \eta'}{\partial P_{jk}} = -\frac{NP_j + N(N-1)P_j^2}{P_{jk}^2} + \mu = 0 \tag{3.8}$$

Rearranging:

$$P_{jk} = \frac{1}{\sqrt{\mu}} \sqrt{NP_j + N(N-1)P_j^2} \tag{3.9}$$

$$\sum_{j=1}^{M} P_{jk} = 1 = \frac{1}{\sqrt{\mu}} \sum_{j=1}^{M} \sqrt{NP_j + N(N-1)P_j^2}$$

$$\sqrt{\mu} = \sum_{j=1}^{M} \sqrt{NP_j + N(N-1)P_j^2}$$

Substitute back into (3.9):

$$P_{jk} = \frac{\sqrt{P_j + (N-1)P_j^2}}{\sum_{j=1}^{M} \sqrt{P_j + (N-1)P_j^2}} \tag{3.10}$$

In the limit as $N \to \infty$, (3.10) becomes:

$$P_{jk} = P_j$$

Therefore, for small N, we expect that the WLS test will yield a *biased* estimate, which is undesirable since the case $P_{jk} = P_j$ does not yield the *minimum* expectation for $\xi(k)$.

3.3.1.3 EVEREST Design Based on the WLS Test

A number of parameters must be determined when applying the WLS test. The performance for given channel(s) and modulation scheme depends on choice of the sample size N and the position (and number of) bins M. For a given N it would be straight forward to simply choose some number of equally spaced bins and examine the performance (through simulation), increasing the number of bins (if necessary) until satisfactory performance is attained. A more efficient approach, detailed in the appendix involves maximising the *sensitivity* of the WLS statistic for a given type of model. By maximising the sensitivity, the optimum placement of bin boundaries may be chosen for a given *single* model. If this model is chosen at a particular error probability, EVEREST performance may be optimised at (and around) this point.

To show the effect on performance that optimising boundary positions at a particular error rate has, the performance of the WLS test was simulated using M=3 bins and N=16000 samples. An AWGN channel was assumed with binary coherent signalling. The decision variable takes the form derived in chapter 4. Figure 3.3 shows the performance optimised for error rates of 10^{-2}, 10^{-4} and 10^{-6} using the procedure detailed in the appendix.

It is clear from figure 3.3 that the performance is indeed optimised at designed BER values of 10^{-2} and 10^{-4}. The curve for optimised design at 10^{-6} shows a shift towards the designed optimal point, but does not achieve a clear "minimum" at that point. In fact, the optimisation of bin boundary positions at a given error rate does *not* mean that performance is optimised over the full range of required operation. Design of an EVEREST using the WLS test optimised over the full range is a more difficult problem and has not been addressed here.

A second interesting example is shown in figure 3.4. This shows the performance using the WLS test optimised at BER values of 10^{-2} and $3\text{x}10^{-2}$. The performance difference between these two designs is large. The design

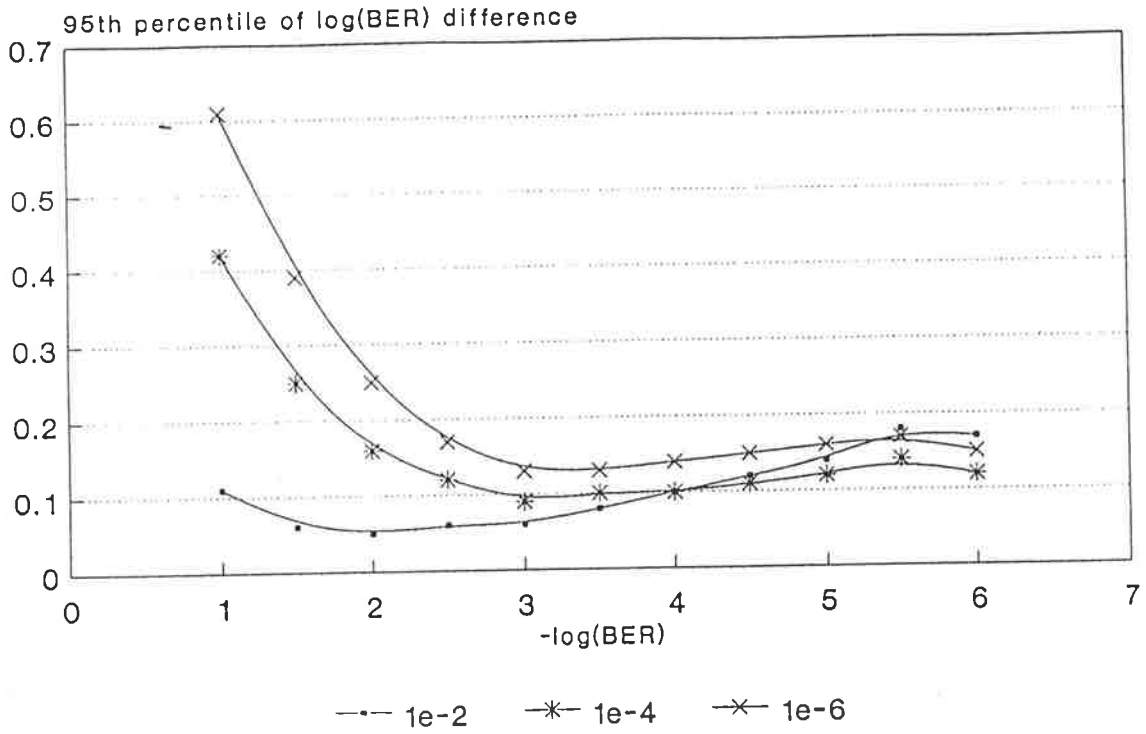


FIGURE 3.3. EVEREST performance using the WLS test, AWGN channel, binary antipodal signalling, coherent detection, with optimised bin positions for BER=10^{-2},10^{-4},10^{-6} using M=3 bins and N=16000 samples.

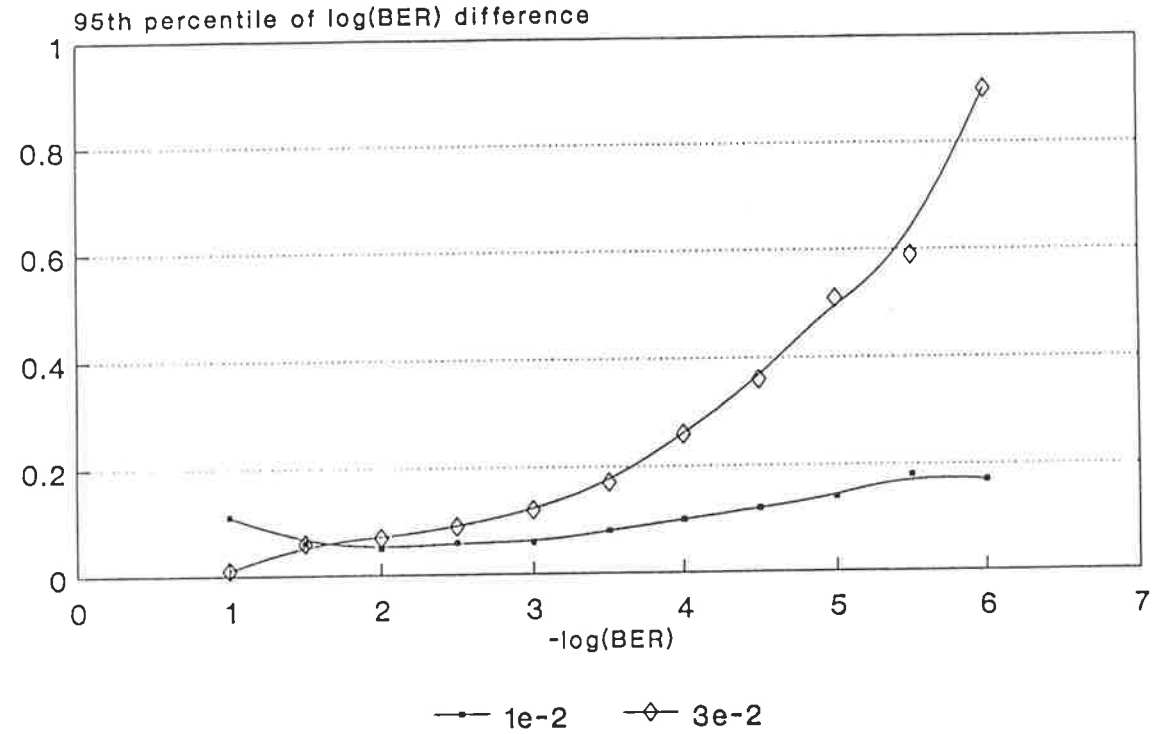


FIGURE 3.4. EVEREST performance using the WLS test, AWGN channel, binary antipodal signalling, coherent detection, with optimised bin positions for BER=3×10^{-2} and 10^{-2} using M=3 bins and N=16000 samples.

optimised at $3x10^{-2}$ achieves better performance at high error rates but suffers much worse at lower error rates. This is because the values recorded in the bins at high error rates are more uniformly spread between the bins. At low error rates counts are largely confined to one bin and thus all models between error rates of 10^{-5} and 10^{-6} appear to be indistinguishable.

3.3.2 Other Tests

Any of a multitude of tests are possible. Tests for "goodness of fit" are associated with the statistical (hypothesis) testing of measured samples to see if they *fit* a given model. Tallis (1983) described a variety of such tests.

Bradley (1960) described a special class of tests called "distribution-free" statistical tests which endeavor to estimate goodness of fit *without* the need for assumptions about the form of the expected distributions.

Other sub-optimal tests were not examined further after the maximum likelihood tests were devised (next sections).

3.4 Optimal Estimators

3.4.1 Low Complexity Estimator Based on Sample Independence

The derivation of the Chi-Squared test given in Fisher (1958) shows that its origins are in the multinomial distribution. Due to the fact that large factorials were difficult to calculate in forming the multinomial probability, simplifying assumptions that the counts in a bin are Gaussian distributed and independent were made. If these assumptions were true, the statistic χ^2 (from equation 3.1) would *truly* be Chi Square distributed (the sum of the squares of independent normal random variables). The optimal estimator derived here, in contrast, does not assume Gaussian distribution of bin counts and uses the multinomial distribution as the likelihood function.

3.4.1.1 Derivation of Optimal Estimator

Consider a stationary, ergodic sequence of random variables v_i (i=1,2,...,N) representing the decision variables. The binning process may be described as follows:

$$e_j(i) = \begin{cases} 1 & v_i \in \mathbb{R}_j \\ 0 & \text{otherwise} \end{cases}$$

Where $\mathbb{R}_j$ is the range of values of v_i which define region j and one symbol period, one count is recorded in only one of the regions $\mathbb{R}_j$. Then:

$$u_j = \sum_{i=1}^{N} e_j(i) \qquad (3.11)$$

and,

$$\sum_{j=1}^{M} u_j = N \qquad (3.12)$$

Disturbances occurring in a communications channel are generally random in nature. As a result, the observations v_i and the accumulated count u_j in region j is thus also a random variable.

The accumulated count u_j is thus the sum of N independent binary random variables. The process of forming u_j is a Bernoulli trial and for the special case of M=2, the probability of k counts in the first region (and hence N-k counts in the second region) is given by the binomial distribution as follows:

$$P(u_1=k, u_2=N-k) = \binom{N}{k} p^k q^{N-k} \qquad \{ q=1-p \} \qquad (3.13)$$

Where: p = Probability of a count occurring in region 1.

q = Probability of a count occurring in region 2.

In general, if there are M regions (or bins), then the probability of u_1 counts in bin 1, u_2 counts in bin 2, (etc.) is given by the *multinomial* distribution (Walpole and Myers, 1975) as follows:

$$P(u_1, u_2, \ldots, u_M \mid \text{model } k) = N! \prod_{j=1}^{M} \frac{\left(P_{jk}\right)^{u_j}}{u_j!} \qquad (3.14)$$

Where: N = Total number of decision variable samples (observations).

u_j = Accumulated count in bin j.

M = Number of bins (or regions).

P_{jk} = Probability of a count occurring in bin j for model k.

k = 1,2,...,L.

L = Number of library models.

Thus (3.14) may form the basis of a test since the probability $P(\mathbf{u}|\text{model k})$ is a measure of the closeness of the observed accumulated counts to a set of probabilities representing model k. If this probability is evaluated for all values of k (ie. against all models) the highest value obtained for some model k=r will correspond to the most likely model. The communications link error probability associated with model x will be taken as the estimated error probability for the link over that measurement interval.

The test based on equation (3.14) may be employed in a simplified form. A Maximum A posteriori Probability (MAP) estimator may be devised by considering the a posteriori probability from Bayes law:

$$P(M_k \mid u_1, u_2, \ldots u_M) = \frac{P(u_1, u_2, \ldots u_M \mid M_k)\ P(M_k)}{P(u_1, u_2, \ldots, u_M)} \qquad (3.15)$$

A MAP estimation requires that the aposterior probability be maximised over all possible models. Let:

$$\psi(r) = \max_{(\text{over } k)} \left\{ P(M_k \mid u_1, u_2, \ldots, u_M) \right\} \qquad (3.16)$$

Where $\psi(r)$ is thus the maximum likelihood, corresponding to model k=r.

Substituting (3.15),

$$\zeta(r) = \max_{(\text{over } k)} \left\{ P(u_1, u_2, \ldots, u_M | M_k) P(M_k) \right\} \qquad (3.17)$$

Noting that the $P(u_1,u_2,\ldots,u_M)$ term is irrelevant since it is not a function of k and has been dropped in (3.17). It is equivalent to maximise the logarithm of $P(\mathbf{u}|\text{model k})P(\text{model k})$, since log is a monotonically increasing function. So, substituting equation (3.14) into (3.17) and taking the logarithm gives:

$$\theta(r) = \max_{(\text{over } k)} \left\{ \log N! - \sum_{J=1}^{M} \log u_j ! + \sum_{J=1}^{M} \log \left[P_{Jk} \right]^{u_j} + \log P(M_k) \right\} \qquad (3.18)$$

The first two terms are independent of k and may thus be neglected, so (3.18) reduces to:

$$\xi(r) = \max_{(\text{over } k)} \left\{ \sum_{J=1}^{M} u_j \log \left[P_{Jk} \right] + \log P(M_k) \right\} \qquad (3.19)$$

or,

$$\xi(r) = \max_{(\text{over } k)} \left[\langle \mathbf{u}, \mathbf{v}_k \rangle + \lambda_k \right] \qquad (3.20)$$

Where $\langle \mathbf{u}, \mathbf{v}_k \rangle$ denotes the inner product of vectors $\mathbf{u}$ and $\mathbf{v}_k$.

Equation (3.20) describes a MAP scheme. The structure of the MAP processor is shown in figure 3.5. If equal weighting is given to priors this reduces to a Maximum Likelihood (ML) test, as given by (3.21) or (3.22).

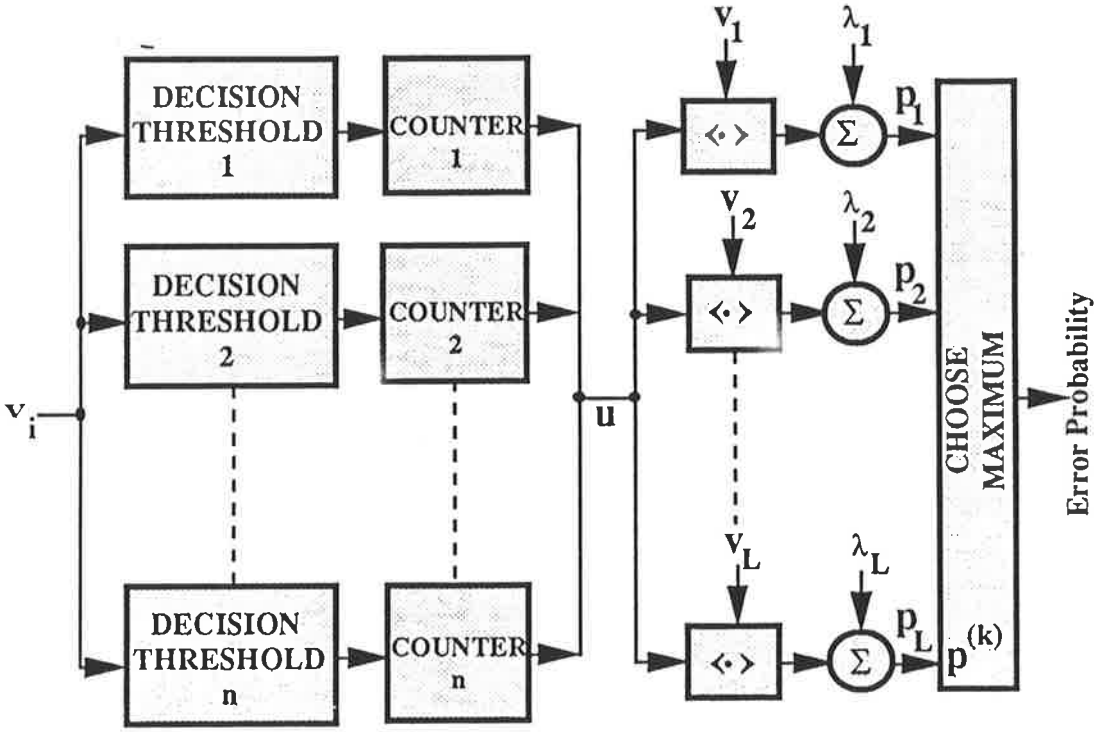


FIGURE 3.5. Maximum A Posteriori Probability (MAP) EVEREST structure.

$$\xi(r) = \max_{(\text{over } k)} \left[\sum_{j=1}^{M} u_j \log \left[P_{jk} \right] \right] \tag{3.21}$$

$$\xi(r) = \max_{(\text{over } k)} \left[< \mathbf{u}, \mathbf{v}_k > \right] \tag{3.22}$$

An implementation of a complete ML estimator of error probability thus consists of the lower-threshold feature extractor (which constructs a histogram) followed by a bank of matched filters (or correlators) as shown in figure 3.6. Note also that the ML test given in (3.21) is *simpler* than the WLS test (3.4). That is, the ML test requires *less* processing or calculation to evaluate.

The performance of the multinomial-based ML test for binary signalling, AWGN channel and coherent detection is shown in figure 3.11, in conjunction with the "continuous-form" ML test for comparison and is described in a later section.

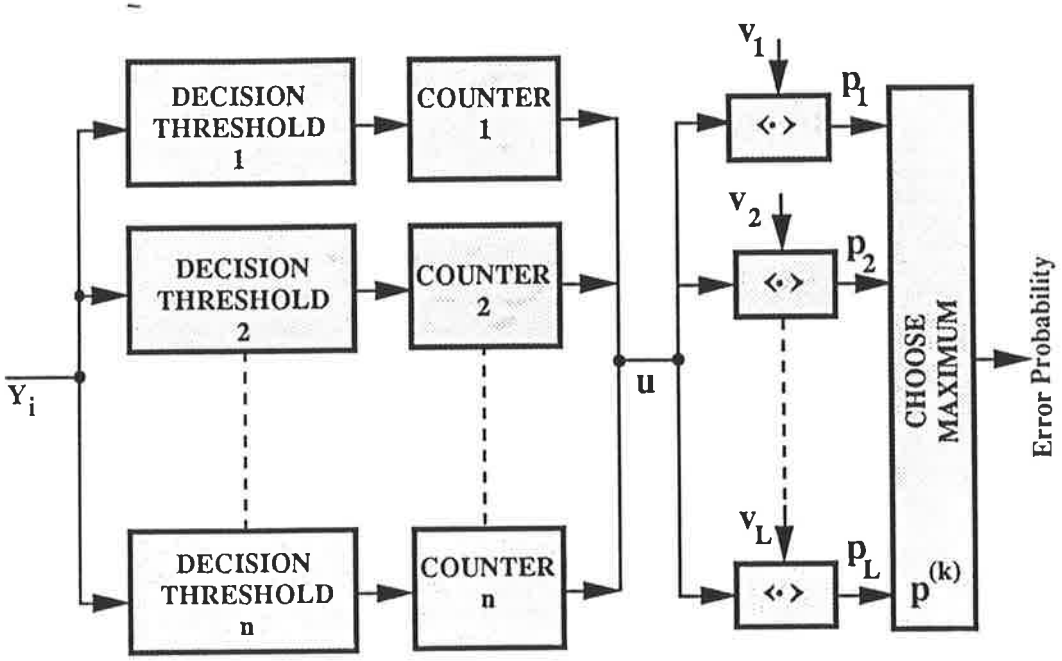


FIGURE 3.6. Maximum Likelihood (ML) EVEREST structure.

3.4.1.2 Behaviour of the Maximum Likelihood Statistic

The EVEREST firstly accumulates samples in each bin. This process is, in effect, an attempt to estimate the individual bin probabilities:

$$\hat{P}_j = \frac{1}{N} \sum_{i=1}^{N} e_j(i) \qquad \{ j=1,2,\ldots,M \} \tag{3.23}$$

Then, the estimate $\hat{P}_j$ is a random variable with mean:

$$E[\hat{P}_j] = P_j \tag{3.24}$$

and variance (Davis, 1988):

$$V[\hat{P}_j] = \frac{1}{N^2} \sum_{g=-N}^{N} (N - |g|)\, C_e(j,g) \tag{3.25}$$

$$= \frac{1}{N} P_j (1 - P_j) \quad \text{if the } e_j(i) \text{ are uncorrelated} \tag{3.26}$$

Where $C_e(j,g)$ is the auto-covariance defined by:

$$C_e(j,g) = E [(e_j(i) - P_j)(e_j(i+g) - P_j)] \tag{3.27}$$

$$= P_j (1 - P_j) \delta_{g0} \quad \text{if the } e_j(i) \text{ are uncorrelated}$$

$C_e(j,g)$ is independent of i because the sequence $e_j(i)$ is assumed stationary.

The technique chooses the model r for which $\rho(k)=\rho(r)$ is a maximum:

$$\rho(k) = \sum_{j=1}^{M} \hat{P}_j \log P_{jk} \tag{3.28}$$

The mean of $\rho(k)$ is:

$$E[\rho(k)] = \sum_{j=1}^{M} E \left[\hat{P}_j \log P_{jk} \right]$$

$$= \sum_{j=1}^{M} P_j \log P_{jk} \tag{3.29}$$

Over all possible models, the one with $P_j = P_{jk}$ gives the maximum $E[\rho(k)]$ as desired (Proakis, 1989:p71). If $E[\rho(k)]=\eta$, the variance of $\rho(k)$ is:

$$V[\rho(k)] = E [(\rho(k) - \eta)^2]$$

$$= E \left[\left(\sum_{j=1}^{M} \hat{P}_j \log P_{jk} - \sum_{j=1}^{M} P_j \log P_{jk} \right)^2 \right]$$

$$= E \left[\left(\sum_{j=1}^{M} (\hat{P}_j - P_j) \log P_{jk} \right)^2 \right] \tag{3.30}$$

Equation 3.30 shows that for the case of perfect estimation where $\hat{P}_j = P_j$, the variance is zero, which can only happen in the limit as $N \to \infty$.

$$V[\rho(k)] = \frac{1}{N} \left\{ \sum_{j=1}^{M} P_j (\log P_{jk})^2 - \left[\sum_{j=1}^{M} P_j \log P_{jk} \right]^2 \right\} \qquad (3.31)$$

Equation 3.31 shows that the variance of the estimate is minimised by taking a larger number of samples. It is not clear from these expressions, however, as to how increasing the number of bins (M) would effect estimation accuracy. In practice it has been found (see Chapter 4) that increasing M increases estimation accuracy asymptoting to the performance of the continuous ML test, which is the subject of a later section in this chapter.

In conclusion, the ML test is unbiased and the accuracy of the test increases as N increases.

3.4.2 Self-Optimising Estimator

The motivation behind the "self-optimising" or "Adaptive Bin Boundary" estimator was to develop a scheme which would take greater advantage of the concentration of points on the decision variable probability density function. That is, the *shape* of the pdf may be such that significant detail may be lost by coarse binning of decision variable points. By *making* the distribution uniform we will be maximising the *entropy* or information content.

The entropy of a discrete probability distribution is defined by Shannon and Weaver (1949):

$$H = - \sum_{i=1}^{M} p_i \log_2 p_i \qquad (3.32)$$

The maxmimum entropy (ie. *information*) in a probability distribution function occurs when $p_1 = p_2 = p_3 = \ldots\ldots = p_M$. That is when $p_i = 1/M$, (Proakis,1989:p71) which gives:

$$H_{max} = \log_2 M \qquad (3.33)$$

3.4.2.1 Adaptive Bin-Boundary Algorithm

3.4.2.1.1 Feature Extraction

The decision variable is a random variable which may take any of a continuum of values at discrete time instants. These values are stored in an array v(j) for j=1,2,...,N. The elements of the array are *reordered* from lowest value to highest value (or vice-versa).

Assuming that a value has been chosen for the number of bins, M. The choice of M may be made from a trade-off between:

- Implementation complexity (decreases for smaller M)
- Estimation accuracy (increases for larger M)

The value of M should also be chosen to be an integer divisor of N. That is N divides M exactly.

Then, at positions in the array **v** corresponding to multiples of N/M, boundaries are constructed as shown in figure 3.7.

Thus there will be an integer number of random variable values in each bin equal to N/M. The accumulated distribution will then be uniform in height, and the information carried in the measured boundary positions d_1, d_2, etc.

The boudary measurements may be taken as the mean value between the two measured points jN/M and jN/M + 1 (j=1,2,...,M) as:

$$d_j = \frac{\left| v_{jN/M} \right| + \left| v_{jN/M+1} \right|}{2} \qquad (3.34)$$

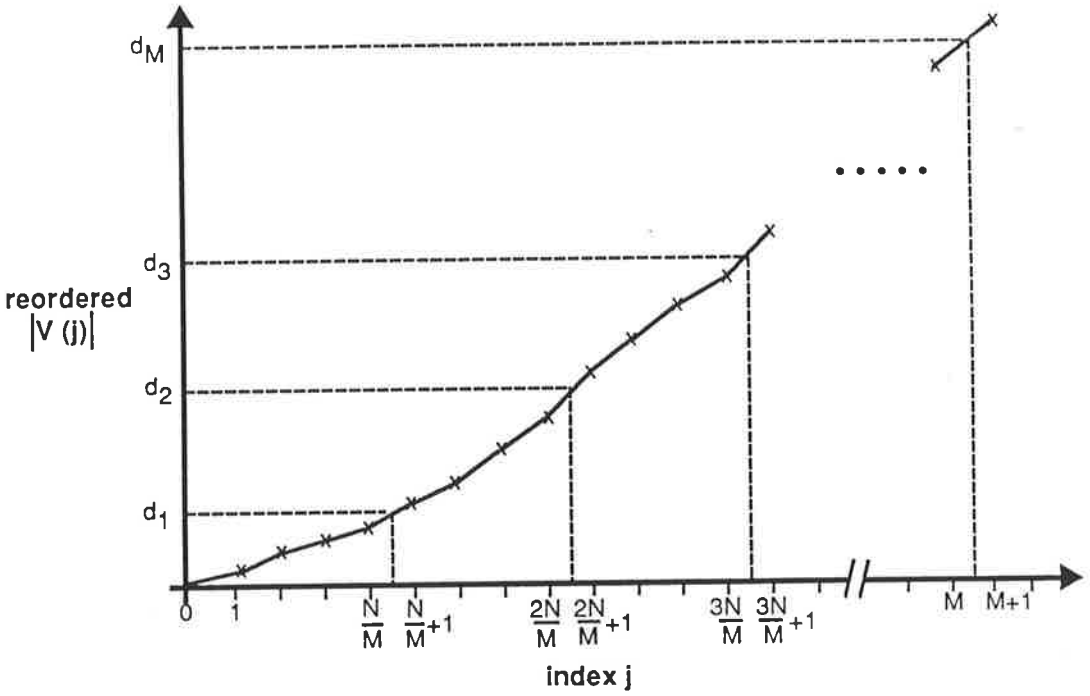


FIGURE 3.7. Example for determining bin boundary values with N = 4M.

3.4.2.1.2 Matcher / Classifier

The classification problem is then to "best" match the set of measured boundaries { d_1, d_2,...,d_M } to a collection of library models.

A maximum likelihood estimator may be found if a likelihood function can be derived, however, no analytical solution was found for this problem.

The *new* model bin probabilities are then given by:

$$P'_{jk} = \int_{d_{j-1}}^{d_j} p_k(v)\, dv \qquad (3.35)$$

Now, using the multinomial likelihood function, we choose the model k=r

for which equation (3.14) is maximised. That is:

$$\rho(r) = \max_{(\text{over } k)} \left(\prod_{j=1}^{M} \left[P'_{jk} \right]^{x_j} \right) \tag{3.36}$$

Taking the logarithm of (3.36) and noting that the observations x_j are all equal to N/M, then (3.36) reduces to:

$$\xi(r) = \max_{(\text{over } k)} \left(\sum_{j=1}^{M} \log P'_{jk} \right) \tag{3.37}$$

Substituting (3.35) into (3.37) for P'_{jk} :

$$= \max_{(\text{over } k)} \left(\sum_{j=1}^{M} \log \left[\int_{d_{j-1}}^{d_j} p_k(v)dv \right] \right) \tag{3.38}$$

Equation (3.38) is the form of the adaptive bin boundary test. This test requires the pdf to be integrated between the measured boundary points or the cumulative distribution at the measured points be known.

3.4.2.1.2 Performance of the Adaptive Bin Boundary Estimator

The feature extraction process requires sorting an array of length N which is equal to the number of samples to be used in the estimation. Clearly, performing a sort on many thousands of samples would be computationally intensive for a real-time application and hence this form of estimator would be best suited to very short-term (small sample) estimates. The additional computation required in this form of estimator however, does result in better estimation performance for a given number of bins than using fixed bin boundaries and the multinomial-based maximum likelihood test, as described in the previous section.

The performance gains are illustrated in the following example, for binary signalling over an AWGN channel. Coherent detection was assumed using a simple correlator/ matched filter. The performance of the Adaptive Bin Boundary (ABB) EVEREST is shown in figure 3.8 using M=4 bins, N=40 and

N=100 samples. The models were chosen to cover the range of error probabilities from 10^{-1} to 10^{-10}. The apparent levelling-off in performance at very low error rates is due to the fact that error probability models below 10^{-10} cannot be chosen, because they are not included in the model library, and thus the chance of choosing a wrong model is reduced. This type of effect will always occur in a practical EVEREST system.

Figures 3.9 and 3.10 show the comparative performance between the ABB EVEREST and the Fixed Bin Boundary (FBB) EVEREST for N=40 and N=100 samples, respectively. The performance of the fixed bin estimator degrades rapidly at low error probability values, since it can no longer distinguish between distributions, having all counts arriving in the first bin (closest to the mean). The ABB estimator, in contrast, continues to distinguish between distributions even at low error rates since it can resolve the fine detail on the distribution.

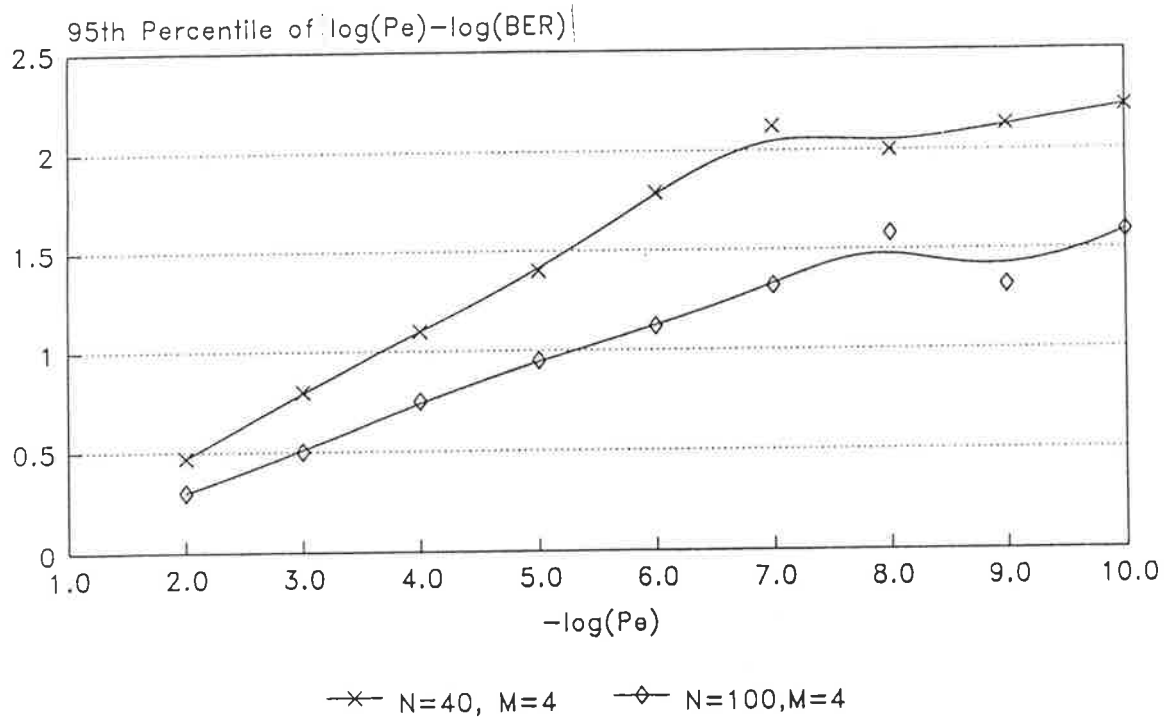


FIGURE 3.8. Adaptive Bin Boundary EVEREST performance for N=40 and 100 samples, M=4 bins.

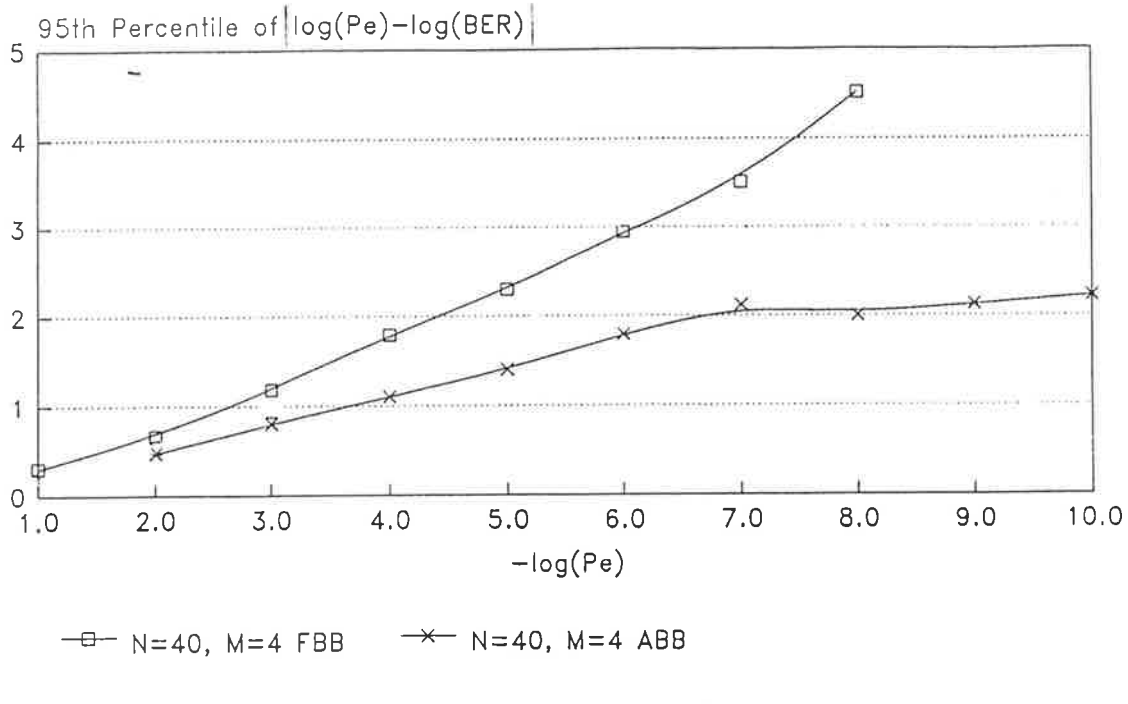


FIGURE 3.9. Comparison of ABB and FBB EVEREST for N=40 samples.

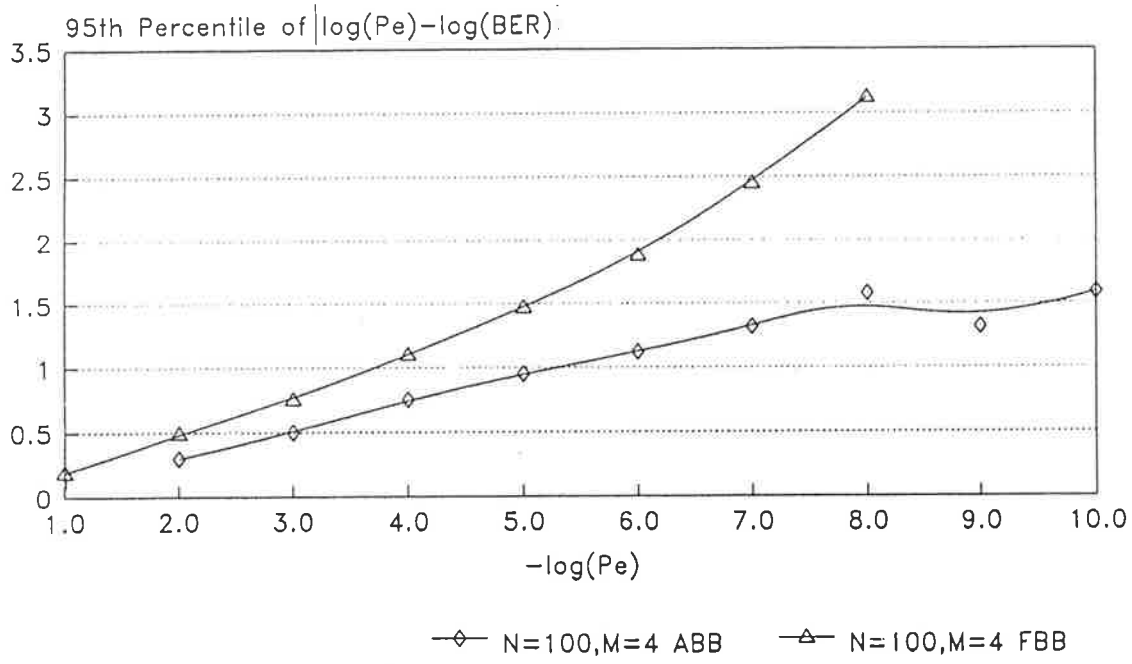


FIGURE 3.10. Comparison of ABB and FBB EVEREST for N=100 samples.

3.4.3 High Performance Estimator Based on Sample Independence

In order to gain the most information from the set of decision variable observations, it would be preferable not to perform *any* binning. Consider, then two possible interpretations of Bayes rule. Where P(.) represents the cumulative distribution, p(.) the probability density, u_i the decision variable samples and M_k, the model k:

$$P(M_k | u_1, u_2, \ldots, u_N) = \frac{P(u_1, u_2, \ldots, u_N | M_k)\, P(M_k)}{P(u_1, u_2, \ldots, u_N)} \tag{3.40}$$

or,

$$P(M_k | u_1, u_2, \ldots, u_N) = \frac{p(u_1, u_2, \ldots, u_N | M_k)\, P(M_k)}{p(u_1, u_2, \ldots, u_N)} \tag{3.41}$$

The first formulation (3.40) uses the cumulative distribution functions, the second (3.41) uses the probability *density* functions. In all previous formulations of the EVEREST problem, the probability of receiving the sample set u_i (i=1,2,...,N) given that the link model k was present, was considered as a distribution function (3.40). This meant that an estimate of the *distribution* $P(u_1,u_2,\ldots,u_N|M_k)$ was used to estimate the aposteriori distribution. Consider the *density* $p(u_1,u_2,\ldots,u_N|M_k)$ for the case of independent samples u_i:

$$p(u_1, u_2, \ldots, u_N | M_k) = p(u_1 | M_k)\, p(u_2 | M_k) \ldots p(u_N | M_k) \tag{3.42}$$

Substituting equation (3.42) into (3.41) gives:

$$P(M_k | u_1, u_2, \ldots, u_N) = \frac{P(M_k) \prod_{i=1}^{N} p(u_i | M_k)}{p(u_1, u_2, \ldots u_N)} \tag{3.43}$$

We wish to maximise the a posteriori probability (3.43) with respect to k.

Since $p(u_1,u_2,\ldots,u_N)$ is independent of k, it may be neglected. A further simplification results by taking the logarithm of (3.43) since in maximising a function, it is equivalent to consider maximising the logarithm of that function. This gives a MAP estimator:

$$\xi(k) = \sum_{i=1}^{N} \ln p(u_i \mid M_k) + \ln P(M_k) \qquad (3.44)$$

Thus, we choose a model k=r such that $\xi(r)$ is maximised over all k. If all priors have equal weighting, this reduces to an ML estimator:

$$\xi(k) = \sum_{i=1}^{N} \ln p(u_i \mid M_k) \qquad (3.45)$$

Again, we choose a model k=r such that $\xi(r)$ is maximised over all k. This test requires use of either the continuous form of the pdf (if known) and evaluating it at each measured point u_i, or by interpolating between pdf points stored in memory to determine the pdf at the measured point. The test, as would be expected, performs exactly as the multinomial-based test would with an infinite number of bins. This additional performance comes at the cost of greater computational complexity.

An example of the performance of this "continuous-form" ML test for binary signalling, AWGN and coherent detection is given in figure 3.11. This figure shows the 95th percentile of the log of the estimation error versus the log of the actual error probability. The simulation used N=1000 samples, with the continuous-form test (M=infinity), and the multinomial test with M=5 and M=9 regions and L=400 library models ranging from 10^{-1} to 10^{-5}. Further simulation details are given in the next chapter. Clearly, the continuous-form test performs best, although not significantly better than using the multinomial test with M=9. Note also that a "dip" in the curves at low error probability values is due to the fact that in the simulation, models were used only to error probabilities of 10^{-5} and classification error is therefore reduced.

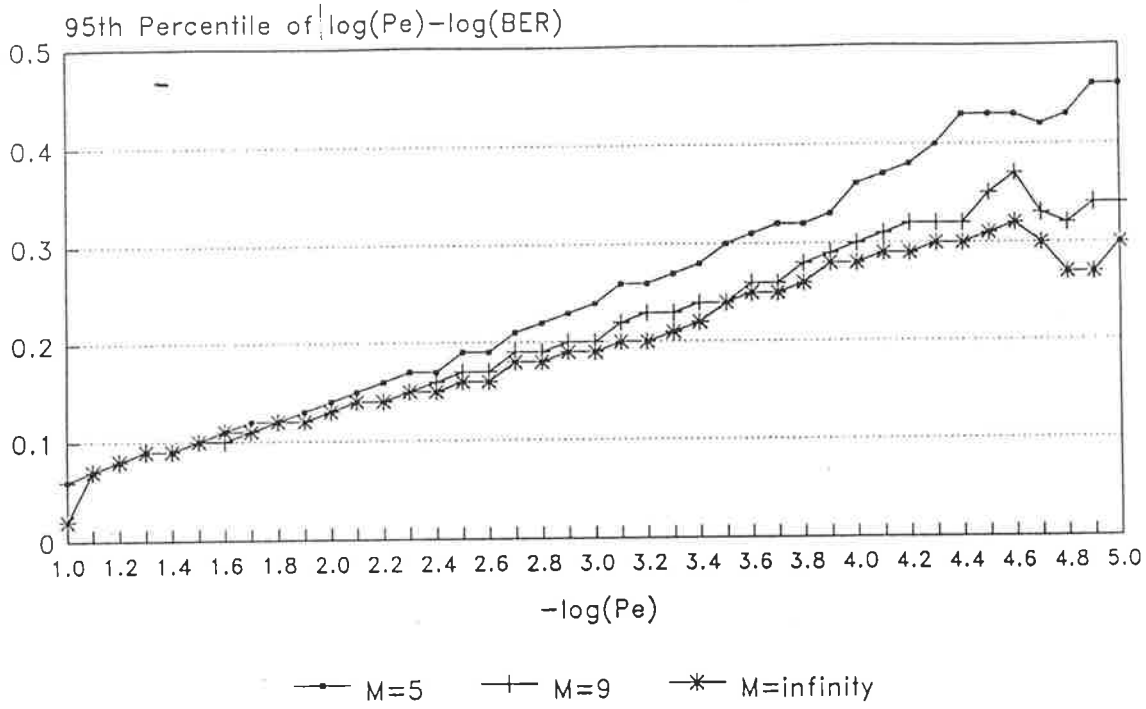


FIGURE 3.11. Continuous-form ML test performance (M=infinity), compared with multinomial-based ML test for M=5 and M=9 regions and N=1000 samples. AWGN channel, binary signalling, coherent detection (see chapter 4).

3.5 Correlation and EVEREST Performance

The derivation of optimal estimator structures in the previous sections were based on the assumption of sample independence. In a practical application some degree of correlation will generally exist between samples. This correlation may be a result of the modulation scheme (ie. a differential scheme will cause two samples to be correlated), the channel (ie. memory due to multipath etc.), receiver filtering, estimators etc.

3.5.1 The Effect of Unexpected Correlation on Performance

Consider the case where the estimator uses library models which were derived assuming no sample dependence, and the multinomial-based ML test is used, but the samples actually do contain correlation. The degree of performance degradation will clearly depend on the exact form of the correlation. The correlation model is a simple one inspired by Jeruchim (1984) and is represented by the Markov process shown in figure 3.12.

This model has been used by Jeruchim in examining the effect of correlation on error rate estimation using the direct method of transmitting a known data sequence. The model is for independent (random) errors and whenever one such error occurs, D further errors follow with probability equal to 1.0. This model is realistic for differential modulation schemes, where an error in one bit causes an error in the next bit. The model may be considered an extreme or "worst-case" form of correlation.

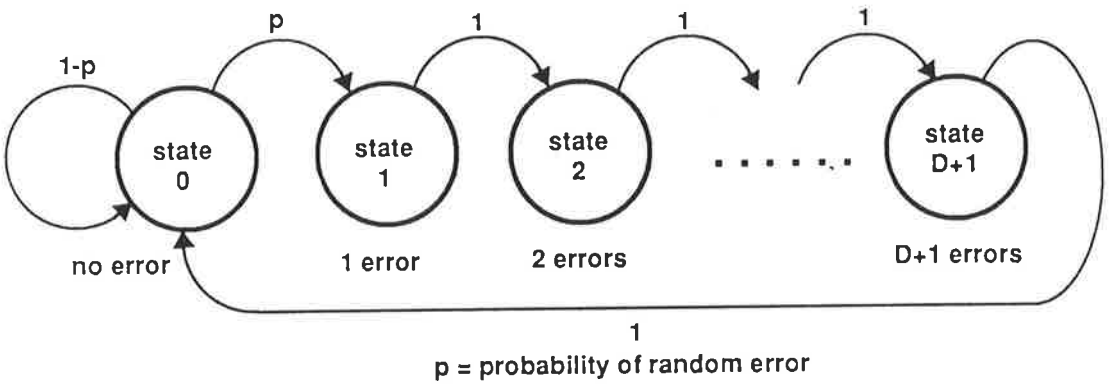


FIGURE 3.12. Decision variable correlation Markov model for random error followed by D dependent errors.

Figure 3.13 shows the simulated worst-case performance of the technique under these assumptions for varying correlation depths (D). The example is for binary coherent signalling in Additive White Gaussian Noise (AWGN) using N=1000 samples and the multinomial-based Maximum Likelihood test with M=9 regions. The case D=0 is the memoryless case, for which performance is optimum. Models spanning the range of error probabilities from 0.5 to 10^{-5} were used. Jeruchim showed that for the error counting method of error probability estimation, that the confidence interval was stretched by the factor $\sqrt{(1+2D)}$. For the ML estimator, figure 3.13 indicates that the increased correlation may be viewed as a reduction in the number of independent samples by a factor of (1+D), thus reducing the effective sample size to N/(1+D), and the estimation performance in accordance.

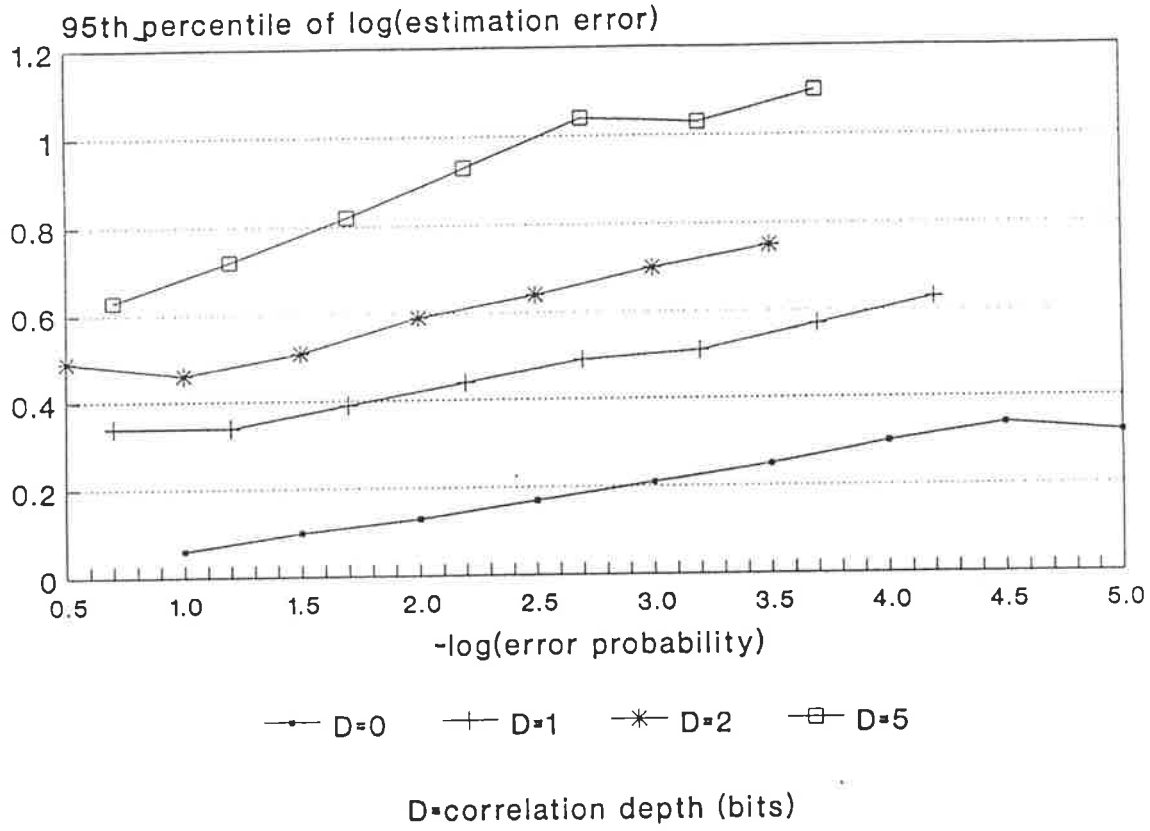


FIGURE 3.13. EVEREST performance when presented with unexpected correlation. N=1000 samples, M=9 bins. Correlation model from Jeruchim (1984).

3.5.2 Ideas for Performance Improvement

The first and most obvious way of improving estimation performance is to include models of the correlation in the estimator. If a fixed form of ISI is primarily responsible for the correlation, say for example a receive filter, the correlation could be accounted for in the building the models. Of course the ML tests given are not necessarily optimum, since they are derived on the assumption of independence. Nonetheless, a performance improvement over over ignoring the correlation would be expected. The major difficulty with this scheme would be if a variety of forms of correlation could occur, due say, to various channel conditions that in order to include all possible forms of correlation, a very large number of samples may be required.

To decrease the additional variance in the error probability estimate one method would be to take time-separated samples of the decision variable(s). Periodic sampling could effectively eliminate correlation between each sample in the set of measurements. This technique implies *not* using all possible samples. For the performance estimation of high speed links this compromise may be quite acceptable.

Communication systems required to operate over fading channels are often subject to bursts of errors, which are periods of high correlation. Data interleavers/ de-interleavers as described by Proakis (1989) are used to remove the correlation in the data, so that random error correction may be used. If a communication system employs soft-decision deinterleaving, the decision variable may be extracted *after* the deinterleaver, therefore removing the correlation effects.

3.6 Chapter Summary

This chapter has developed the theory for a new error probability estimation technique. The estimator structures based on the assumption of sample independence, are demonstrated to operate with a very high degree of accuracy and repeatability. The Weighted Least Squares test was found to possess some undesirable properties, the most significant of which are estimation bias and applicability to Gaussian distributed bin values (which is true only for a large number of samples in each bin). To overcome these limitations, the multinomial ML estimator was developed, which also resulted in a computationally simpler estimator structure.

An Adaptive Bin Boundary (ABB) structure, based on a maximum entropy principle, was proposed to improve performance for a fixed number of bins. Simulation results found the ABB structure to be marginally superior in performance to the multinomial-based Maximum Likelihood estimator structure in some situations. However, implementation complexity, the main driving inspiration for the proposed test, became its main drawback. As a result, the unbiased and efficient Maximum Likelihood estimators are the preferred structures.

The effect of introducing correlation on estimation performance is to reduce the effective number of samples. For the severe correlation model examined, performance was degraded significantly. In practice, demodulation of the signal would necessitate the use of a ML Sequence Estimator and the error probability estimator would be designed to integrate with this new detection strategy as appropriate.

Chapter 4

PERFORMANCE OF THE NEW ERROR PROBABILITY TECHNIQUE FOR VARIOUS MODULATIONS AND CHANNEL TYPES

Abstract

In the chapter 3, proposed error probability estimator structures have been examined on their relative merits using a simple channel and modulation type. In order to examine the practical performance of the EVEREST (Extremely Versatile Error Rate ESTimator), a range of different realistic channels and modulations have been examined.

This chapter describes the design methodology and performance of EVEREST, based on simulation, mostly using the multinomial-based test, due to its simplicity. Both simulated and live channel performances are presented.

The effects of altering the number of bins (regions) and the number of samples are examined. The sensitivity of the EVEREST to these design parameters is shown to be strongly dependent on the form of the decision variable distribution.

4.1 Introduction

The design of EVEREST for a particular application must be achieved by simulation, since no purely analytical means of predicting performance has been discovered. This is not a practical limitation since simulation for design can be easily implemented and is not computationally expensive.

Referring below, "Program 1" *calculates* EVEREST performance, for the multinomial-based test producing the log difference between the actual and estimated error probability for a desired percentile confidence level and has the following flow:

PROGRAM 1

1. *Generate library models* for the desired error probability estimation range and resolution.
2. For each desired point "i" ($1 \leq i \leq L$) on the percentile graph:
3. Use library model "i":
4. For iterations = 1 to 2000:
   { for a 95th percentile, 2000 iterations gives 5%=100 counts beyond the 95th percentile}
5. *Perturb* the pdf model "i" as if it had been subject to counting statistics, using the multinomial distribution.
6. Use the perturbed model as the measurement and find the best *matching* library model.
7. Calculate the log difference "d" between the error probability of the best matching model and the true model "i".
8. Record the distribution of matching error by incrementing an array value at index "d", test_store[d] = test_store[d] + 1.
9. Next iteration.
10. For j=0 to 2000:
    Add up values of test_store[j] until the accumulated sum is the desired percentile. (eg. 95th percentile over 2000 iterations is 1900).
11. Record point on performance graph of that value of j at which the accumulated total equalled the percentile.
12. Next "i" (next point on the graph).

The advantage of program 1 is that it is a high level calculation and does not require the time consuming bit by bit simulation required of a Monte Carlo performance simulation.

A Monte Carlo (bit-by-bit) type simulation was written and used for determining the EVEREST performance when the form of the decision variable distribution was not known or if the continuous-form Maximum Likelihood test was used.

4.2 Additive White Gaussian Noise Channels

4.2.1 Binary Modulations, Coherent Detection

The transmitted signal may be represented, as given by Proakis (1989):

$$s_{1,2}(t) = \mathrm{Re}\left\{ u_{1,2}(t)\ e^{j\omega_c t} \right\} \qquad (4.1)$$

Where $u_{1,2}(t)$ is a real-valued signal pulse representing the binary digit to be transmitted. At the receiver, the signal is down-converted and for the case of coherent detection, the detector is assumed to form the decision variable **U** from the received signal r(t), for each symbol i=1,2,...,N:

$$U_{1,2}(t=iT) = \mathrm{Re}\left\{ \int_{(i-1)T}^{iT} r(t)\ u_{1,2}^{*}(t)\ dt \right\} \qquad (4.2)$$

Where: T = Signalling period.

* Denotes conjugation.

For the case of an AWGN channel with coherent detection, the optimal detector is a bank of correlators (or matched filters) and a largest-of decision function as shown in figure 4.1. The matched filter is designed to minimise the probability of error for transmission over the channel, which is the same as maximizing the energy per bit.

In this case, the decision variable distribution is Gaussian distributed

and is the sum of two correlation components as follows:

$$U(t=iT) = \mathbb{R}_{SS}(iT) + \mathbb{R}_{SN}(iT) \qquad (4.3)$$

The autocorrelation component $\mathbb{R}_{SS}$ is a constant of value $-\mu$ or $+\mu$, depending on which symbol was sent. The signal / noise cross-correlation component is zero mean Gaussian so **U** is Gaussian with mean of either $-\mu$ or $+\mu$.

Most occurrences of the decision variable will be around the mean of the distribution and it would therefore be intuitive that optimum placement of bin boundaries would be near the mean. For simplicity, an even number of equally spaced bins was chosen between the decision threshold and twice the mean, so that most information is captured, with a single additional bin extending from twice the mean to infinity as shown in figure 4.2. In a practical receiver, the decision variable value will be limited by either voltage rails in an analog system or the number of bits in a digital system, in which case the additional bin would probably include values at and beyond this "saturation" point.

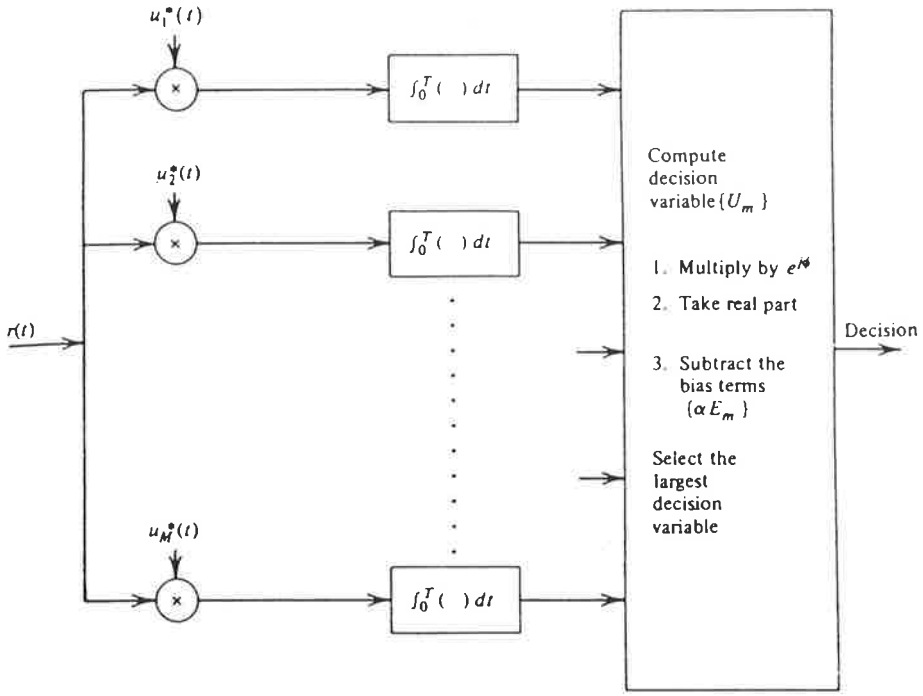


FIGURE 4.1. Optimal coherent detector for AWGN channel - Proakis (1989).

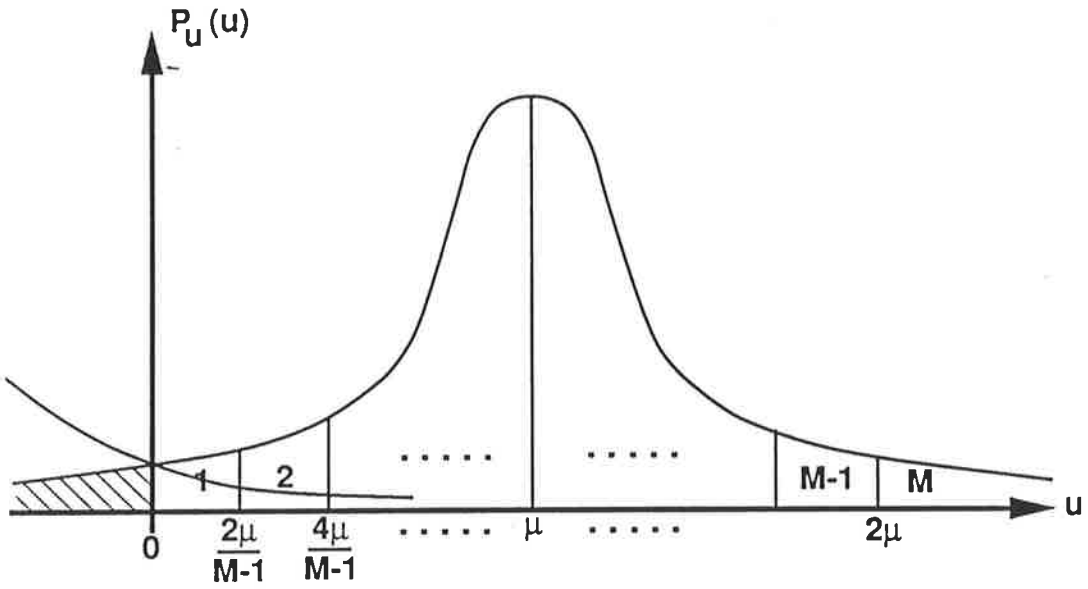


FIGURE 4.2 Decision variable bin segregation.

The cumulative probability of the decision variable is given by (4.4), which becomes the probability of error when a=0.

$$P(u \leq a) = 1 - Q\left\{(a-\mu)/\sigma\right\} \tag{4.4}$$

Where:

$$\mu/\sigma = \sqrt{R(1-\rho_r)} \tag{4.5}$$

R = Energy per bit ratio

ρ_r = Correlation coefficient , $-1 \leq \rho_r \leq 1$

and,

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(\frac{-s^2}{2}\right) ds \tag{4.6}$$

Note that each bin contains the overlapping distribution from the other decision. Each bin should include this overlap probability although it becomes miniscule for bins further from the decision threshold, and for

increasing SNR. Then, the bin probability for a region (a,b) is given by:

$$P_{ab} = P(u \leq b) - P(u \leq a) + P(u \leq -a) - P(u \leq -b) \qquad (4.7)$$

One may arbitrarily take μ=1 to evaluate P_{ab}, using (4.4), (4.5) and (4.7) for a given E_b / N_0 and thus determine the library (reference) models over the desired range of error probability values. Taking μ as a constant assumes perfect Automatic Gain Control (AGC) which maintains a constant average received power level. The effect of imperfect AGC characterisitcs has not been treated.

The number of library models required depends simply on the desired resolution. The following simulations used a large number of models only so that a continuous performance curve could be obtained. Each model used was separated by a value of 0.01 in log error probability terms, over a range extending from $P_e = 10^{-1}$ to 10^{-6}, thus giving 500 models. In a practical application, a model resolution of 0.2 in log error probability would probably be more than sufficient, therefore requiring only around *25 models*.

The following figures show the results of simulations of EVEREST performance. The vertical scale indicates the 95th percentile of the logarithm of the difference of the actual probability of error and the EVEREST estimated BER. The results are shown in terms of a confidence interval due to the uncertainty involved in sampling a random process. The horizontal scale indicates the actual BER (or probability of error). The simulation program was written in C language.

Figures 4.3 and 4.4 show the performance of the EVEREST for an AWGN channel using 1000, 4000 and 16000 samples, where the number of regions (bins) is 5 and 9, respectively. The library models were derived using equations (4.4) to (4.6).

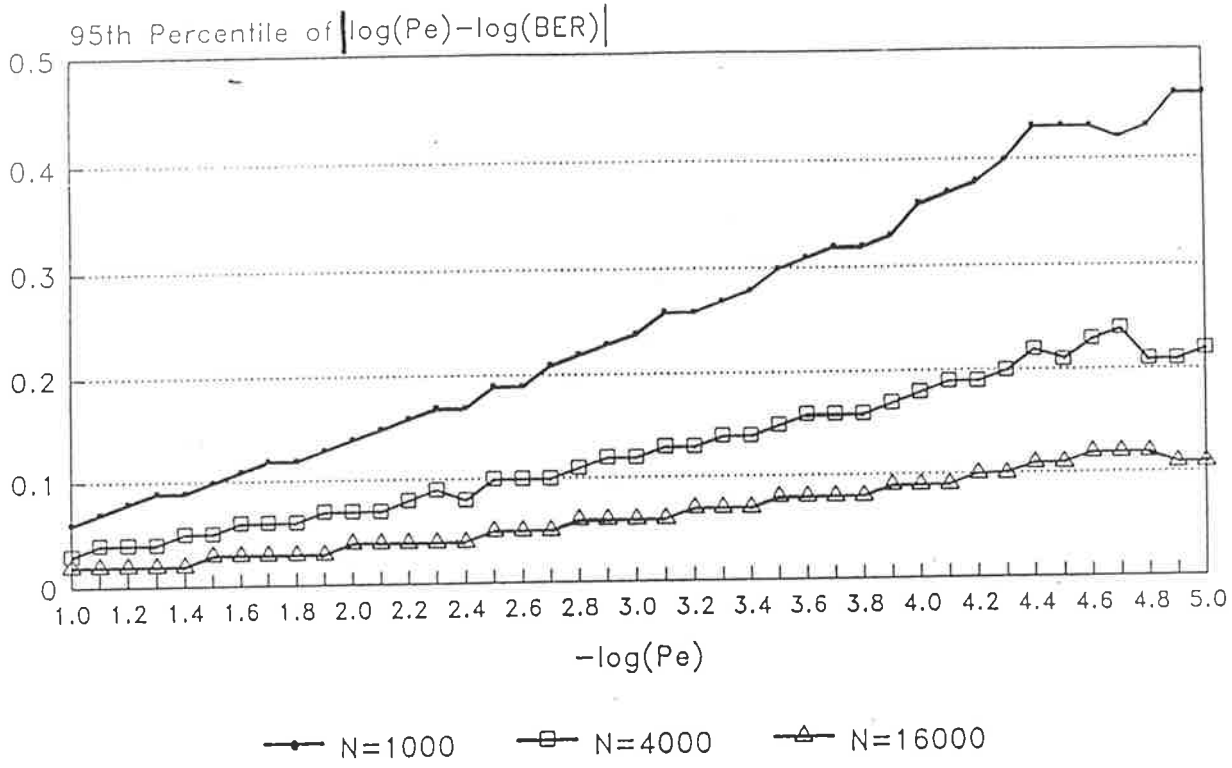


FIGURE 4.3 EVEREST performance with M=5 regions, AWGN channel model.

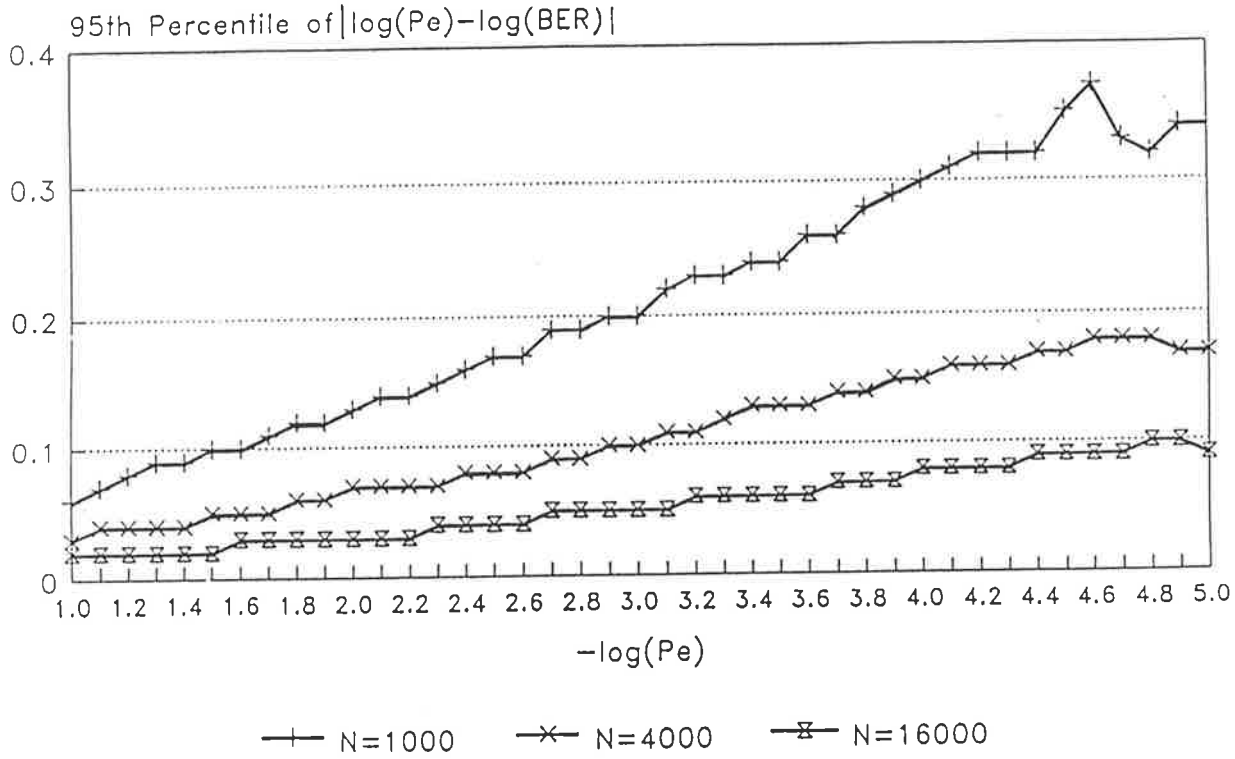


FIGURE 4.4 EVEREST performance with M=9 regions, AWGN channel model.

As an example for interpreting these graphs, in the case of figure 4.4 (M=9) for N=1000, at an error rate of 10^{-3}, the 95^{th} percentile is 0.2. Thus, 95% of all estimates at an error rate of 10^{-3} lie within 0.2 of an order of magnitude of 10^{-3}. That is, 95% of measurements will be between 0.63×10^{-3} and 1.58×10^{-3}.

These figures indicate that a high degree of accuracy is attainable with only a few bins, and *very* few samples compared with, say the Monte Carlo error counting method. It can be seen that a four-fold increase in the number of samples halves the log of the estimation error.

The performance of the EVEREST for a *very small* number of samples is shown in figure 4.5, with only 40 and 100 samples and M=9 bins. To appreciate this performance, compare it with the following method based on variance estimation. The variance of a Gaussian distributed random data sequence may be estimated by the following formula from Beyer (1988):

$$\hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 \qquad (4.8)$$

Where: n = Number of samples.

μ = Mean (expected value) of the sequence.

Since, for AWGN, the mean is considered constant, the variance is all that is required to determine the error probability. In order to draw a direct comparison, consider that a lookup table (or library) of variances exists, covering the range of expected error probabilities, analagous to the ones EVEREST uses. Then, the error probability of the model whose variance is closest to the variance estimated by equation (4.8) is chosen. The performance of this scheme has been simulated, using the same number and range of models used by the EVEREST and is also shown in figure 4.5. The performance of both schemes is identical at small error probabilities, but at high error probabilities, the EVEREST scheme is superior. The main reason for this is that the ML estimator is able to exploit the information in the distribution overlap, which is not taken into account using equation (4.8). The further disadvantage of the variance estimator is that if the data distribution were non-Gaussian, variance could not be translated into error probability.

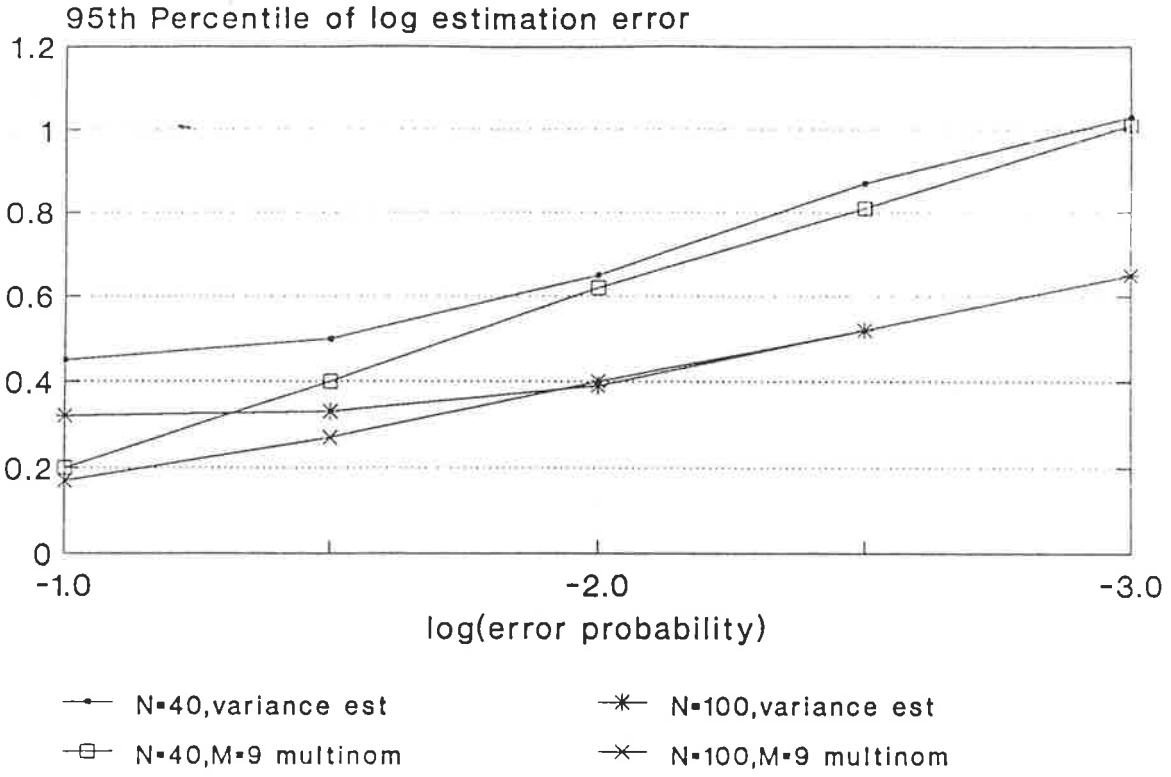


FIGURE 4.5 Performance of the EVEREST with M=9 regions, N=40 and 100 samples, and a variance estimation technique, AWGN channel model.

4.1.2 Multiphase Signalling, Coherent Detection

In multiphase signalling, the general form of the optimum demodulator for detecting one of a set of symbols, using Gray encoding in an AWGN channel is given by Proakis (1989,p.259). Optimum demodulation may be accomplished by phase detection of the received vector and selecting from the set $\{s_m\}$ that signal having a phase closest to ϕ.

Performing this, the probability density function for the phase is given by Proakis (1989,p.262):

$$p(\phi) = \frac{1}{2\pi} e^{-\gamma} \left(1 + \sqrt{4\pi\gamma}\, \cos\phi\, e^{\gamma \cos^2 \phi} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{2\gamma}\cos\phi} \exp(-y^2/2)\, dy \right) \qquad (4.9)$$

Where: γ = Signal to noise ratio per symbol.

Noting that the integral in (4.9) may be reduced:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{2\gamma}\cos\phi} \exp(-y^2/2)\, dy = 1 - Q\left(\sqrt{2\gamma}\cos\phi\right) \qquad (4.10)$$

The probability of receiving a phasor in the interval (a,b) is then:

$$P(a \le \phi \le b) = \frac{e^{-\gamma}}{2\pi} \left[b - a + \sqrt{4\pi\gamma} \int_a^b \left(1 - Q\left(\sqrt{2\pi}\cos\phi\right) \right) \cos\phi \right.$$

$$\left. \exp(\gamma\cos^2\phi)\, d\phi \right] \qquad (4.11)$$

The integral in equation (4.11) must be solved numerically.

Equation (4.11) is the required expression to evaluate library model probabilities, noting that in this case the density functions centred around each expected constellation point have tails that flow into each other. Figures 4.6a and 4.6b illustrate the binning process for the case of Quadrature PSK (QPSK) signalling with M=4 equal regions between $(-\pi/4, \pi/4)$.

The performance of the error estimation technique for multiphase signalling in AWGN is shown for the case of M=4 equal sized bins in figure 4.7. This figure shows results for N=1000 and N=16000 samples, with L=40 and L=400 library models equally spaced between error probabilities of 10^{-1} and 10^{-5}. The effect of a reduced set of models simply makes the estimation more coarse. The number of library models chosen for a practical application depends on the degree of coarseness acceptable.

At a greater extreme, the EVEREST performance at *very low* error probabilities is examined for QPSK in figure 4.8. In this situation, a practical value of M=8 was taken with equally spaced bins to twice the distribution mean and examined over a range of error probabilities extending from 10^{-1} to 10^{-11}.

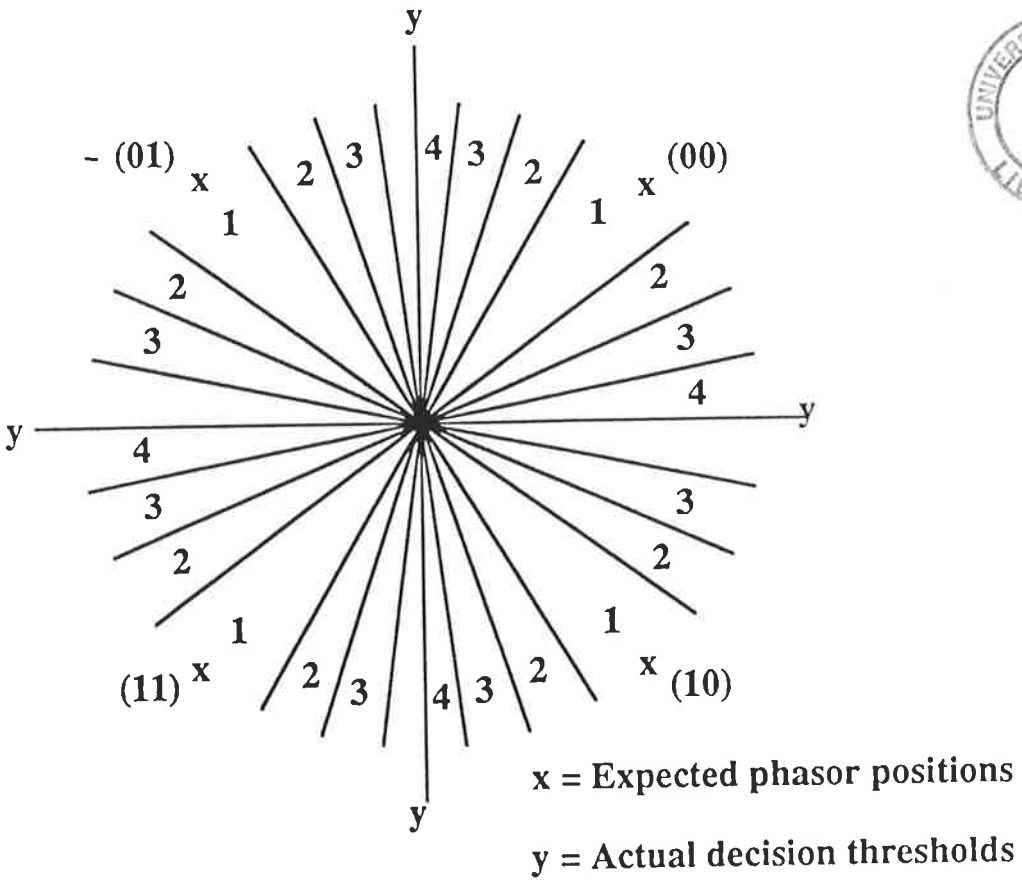


FIGURE 4.6a. Decision variable bin regions (1,2,3,4) for example QPSK signalling with M=4 bins.

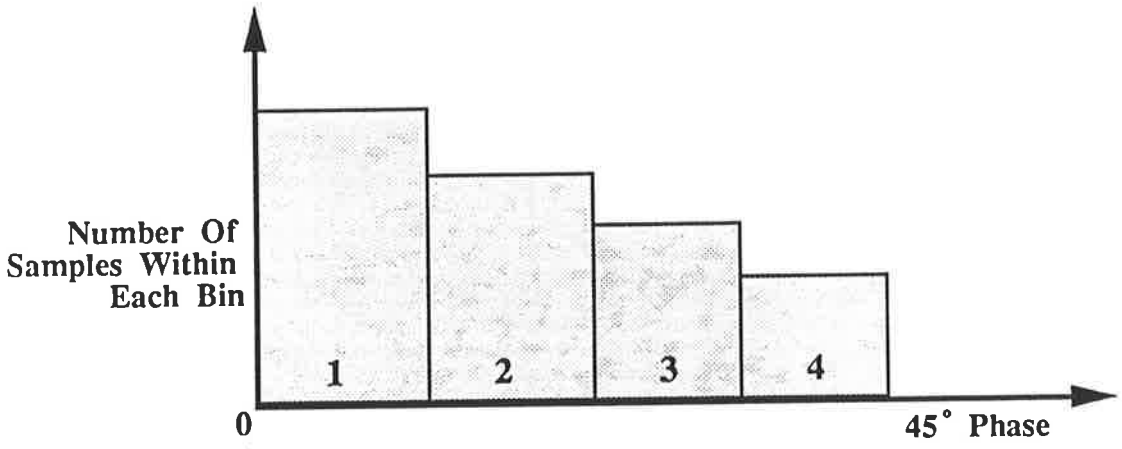


FIGURE 4.6b. Decision variable histogram resulting from the example M=4 bin regions shown in figure 4.6a.

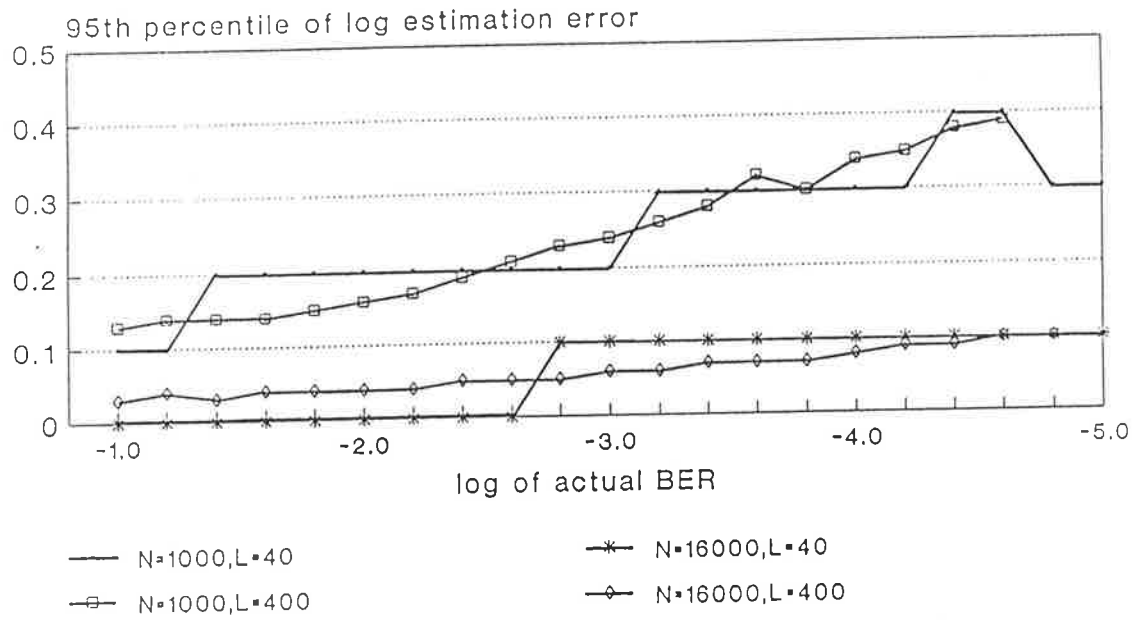


FIGURE 4.7 Estimator performance for Multiphase PSK signalling in an AWGN channel with M=4 equal sized bins.

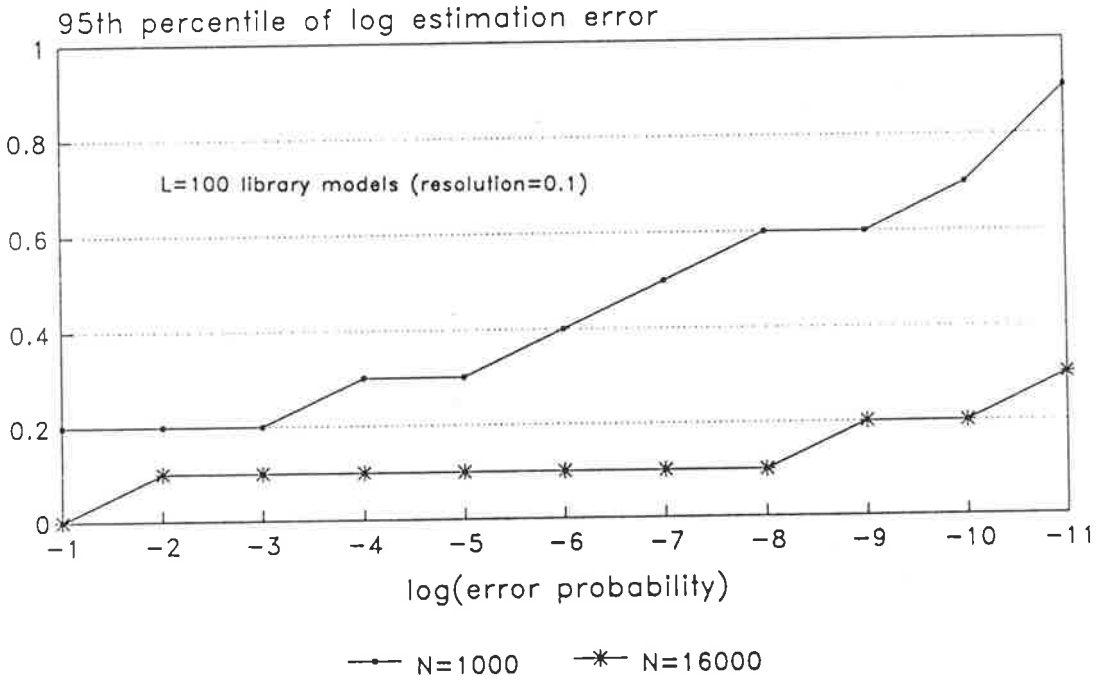


FIGURE 4.8 Performance of the EVEREST with M=8 regions, N=1000 and 16000 samples, AWGN channel model.

4.3 Fading Channels

4.3.1 Slow Flat Rayleigh Fading Channel - Binary Orthogonal Modulation with Square-Law Detection

This case assumes binary orthogonal modulation over a frequency non-selective (flat) fading Rayleigh channel with AWGN and L-order diversity. The matched filter detector firstly forms two random variables U_1 (containing the signal) and U_2, as given by Proakis (1989,p726):

$$U_1 = \sum_{k=1}^{L} \left| 2\varepsilon\, \alpha_k\, e^{-j\phi_k} + N_{k1} \right|^2$$

$$U_2 = \sum_{k=1}^{L} \left| N_{k2} \right|^2 \qquad (4.29)$$

According to Pierce (1958), U_1 and U_2 are independent, and each Chi-Squared distributed with 2L degrees of freedom as follows:

$$p(U_1) = \frac{u_1^{L-1}\, e^{-u_1/2\sigma_1^2}}{(2\sigma_1^2)^L\, (L-1)!}$$

$$p(U_2) = \frac{2\, u_2^{L-1}\, e^{-u_2/2\sigma_2^2}}{(2\sigma_2^2)^L\, (L-1)!} \qquad (4.30)$$

Where:

$\sigma_1^2 = 2\varepsilon N_0(1 + \bar{\gamma}_c)$

$\sigma_2^2 = 2\varepsilon N_0$

$\bar{\gamma}_c$ = Average SNR per diversity channel.

N_0 = Noise power spectral density.

ε = Signal Energy.

The decision variable is then $U = U_1 - U_2$, so the density of U is the convolution of $p(U_1)$ and $p(U_2)$. Consider the simplified case of L=1, performing the convolution integral gives the following pdf:

$$p_u(u) = \begin{cases} \dfrac{e^{u/2\sigma_1^2}}{2(\sigma_1^2 + \sigma_2^2)} & ; \quad u \le 0 \\ \dfrac{e^{-u/2\sigma_2^2}}{2(\sigma_1^2 + \sigma_2^2)} & ; \quad u > 0 \end{cases} \qquad (4.31)$$

For simulation, the continuous-form ML test was chosen.

Figure 4.9 shows the performance of the EVEREST for a Rayleigh Fading channel with AWGN. A "dip" in the curves at low probabilities of error occurs since the models in the simulation were chosen to extend only to $P_e = 10^{-5}$. This system requires more samples to achieve good performance than say, the Gaussian noise case, because the pdf's appear similar to one another over the range of average snr values.

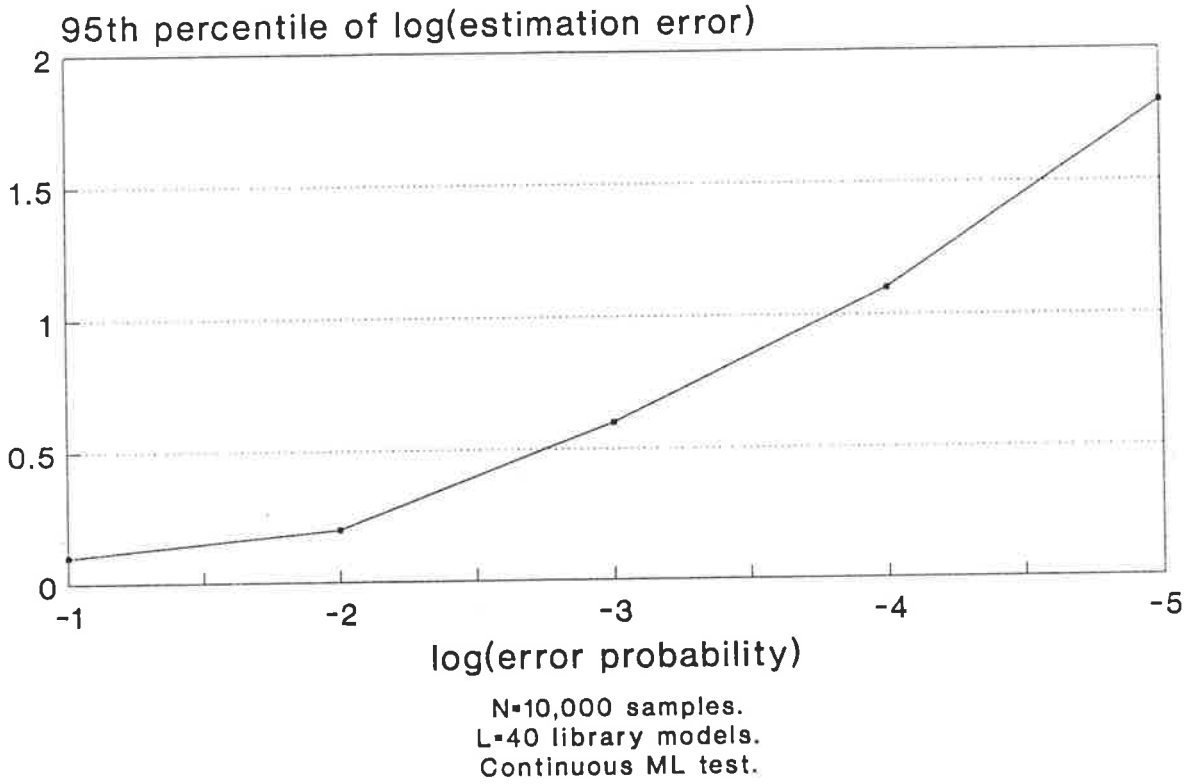


FIGURE 4.9 EVEREST performance with continuous-form ML test, N=10000 samples, Rayleigh Fading with AWGN channel model.

4.3.2 Mixed Rayleigh Fading and Additive Noise Channel, QPSK Modulation, Differentially-Coherent Detection

For this channel, two *types* of library model were considered. The models were AWGN and Rayleigh fading with AWGN. These models were deemed to be the two most important channel models for a particular case of a skywave HF link using a parallel tone DQPSK modem. A Magnavox MX-513B modem was available for testing and therefore the simulated and live-channel performance of the EVEREST could be compared. This modem by virtue of its parallel tone signal structure, eliminates multipath ISI, since each tone has a symbol period much greater than the largest possible path delay. As a result, the decision variables mostly experience Rayleigh fading, additive noise and interference. The simulated performance is therefore for a single DQPSK tone.

4.3.2.1 Simulated Performance

Determination of the EVEREST models for DQPSK modulation with flat Rayleigh fading and AWGN was achieved by hardware simulation.

Whilst models may be derived analytically, this method was felt more appropriate for a direct comparison of simulated performance and on-air measured performance since using the same set of models.

Models were derived by passing known data through Magnavox MX-513B modems and a Cossor ionospheric simulator and measuring the distribution of decision variable occurences in each bin, for a period of time sufficient to allow calculation of the data error probability.

For the simulation, 20 models for AWGN and 68 models for flat (non-frequency selective) Rayleigh fading with AWGN were used over a range of error probabilities from 10^{-1} to 10^{-4}. The simulation was conducted as described in section 4.1. The worst case performance for each actual error probability was used in the figure. The worst case performance is shown in figure 4.10 with N=3750 and M=4 bins. The roughness of the curve is due to two main effects, the non-uniform (and sometimes fairly wide)

spacing of library models over the measurement range and the fact that the worst-case performance (between AWGN and Rayleigh) is shown. In addition, as was later discovered, some inaccuracies in library models was present due to the fact that some were not measured for long enough (accounting for the peak at an error probability of 10^{-4}).

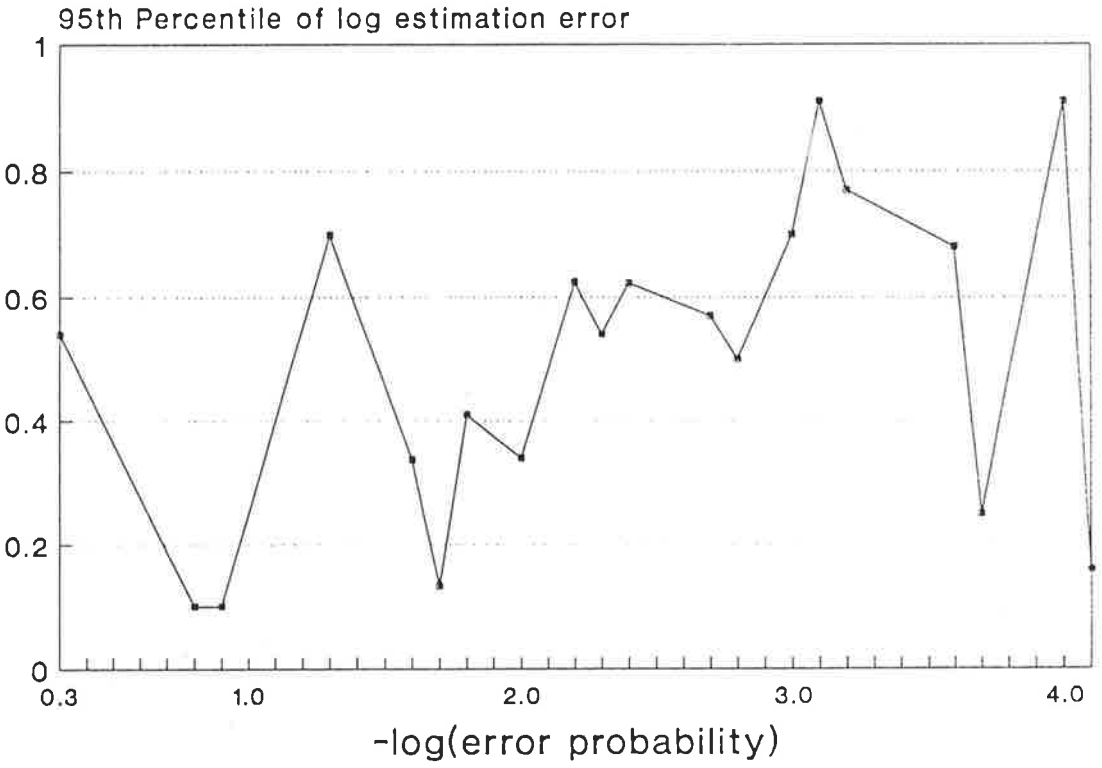


FIGURE 4.10 Simulated worst-case performance for DQPSK with both Rayleigh Fading and AWGN models for M=4 bins, N = 3750 samples using models from hardware simulator.

4.2.2.2 Live-Channel Performance

The EVEREST performance was tested using a Magavox MX-513B parallel tone (Kineplex) modem on a 3000 km skywave HF link between Darwin and Melbourne in 1989 and is reported in Scholz (1991a). This modem uses differential QPSK signalling at 75 symbols/sec on each of sixteen parallel tones in a

3kHz bandwidth. The Weighted Least Squares test was used in the trial with M=4 equal-sized bins, as the optimum tests had not at that time been developed. Four bins only were chosen so that the processing would be minimal.

The monitor was required to provide an estimate of the error probability at least once every minute for an intelligent frequency management system (Scholz, 1990). A measurement interval of 50 seconds was subsequently chosen, thereby allowing N = 75x50 = 3750 symbol samples per tone. It was decided that the error probability of each tone would be examined independently and the overall error probability would be estimated as the average of the sixteen estimated values. Over this measurement period, the two types of library model considered to be most important were AWGN and flat (non-frequency selective) Rayleigh fading with AWGN. Models for frequency selective fading were not necessary because the parallel-tone structure of the modem signal eliminates the effect of multipath. Frequency tracking was also employed in the modem so that models for ionospheric Doppler shift distortion were not required.

The actual error probability of the link was estimated using the Monte Carlo method by transmitting pseudo random data known to the receiver. Using the Monte Carlo method, the overall error rate could be measured from a total of 3750x16 = 60000 symbol samples. For QPSK this means 120000 bits were sent, thus allowing about a 95th percentile confidence in the actual value to be extended to an error rate as low as:

$$P_e \cong \frac{10}{120\ 000} = 8 \times 10^{-5} \cong 10^{-4}$$

Therefore, error probability estimates can only be verified in the range 0.5 to 10^{-4}.

The decision variables were extracted directly from the modem and applied via an analog to digital converter into an IBM AT-PC. The EVEREST was implemented entirely in the PC.

Figure 4.11 is a photograph of the PC display, showing the status of the link. The display was updated continually as new measurements were made.

The sixteen displayed histograms show the accumulated decision variable counts in each of four bins. Results for each QPSK tone are represented in the frequency band from 935Hz to 2585Hz. Below the bar graph display, the actual number of counts registered in each bin is shown. Below this, the "CHI VALUES" represent the lowest metric obtained as a result of matching each measured histogram with the "closest" model (ie. the model from the WLS test (equation 3.2) with the lowest chi value or metric). A value of zero for a "CHI VALUE" would indicate a perfect model match, larger values show a worsening correspondence. The bit error rate associated with the "closest" model is shown under this, and the average of the sixteen values is the "OVERALL BIT ERROR RATE" for the link.

Figure 4.12 shows the results which represent about 350 hours of measurements as reported in Scholz et. al. (1992). The library contained a total of L=300 models, made up of 250 flat Rayleigh fading models for fade rates of 0.2, 0.4, 0.6, 0.8, 1.0 Hz and 50 AWGN models covering the range of error probabilities from 0.5 to 10^{-6}.

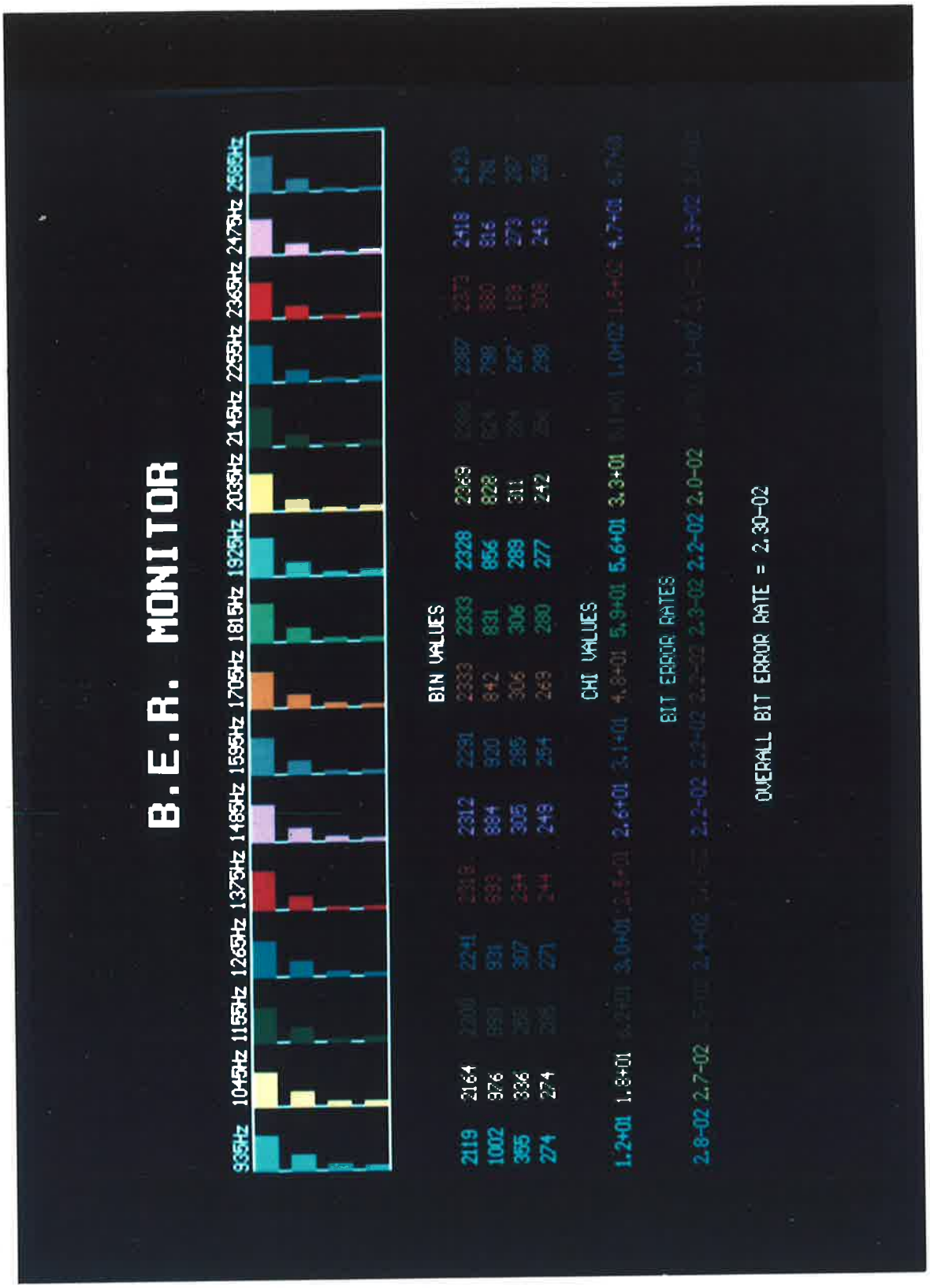


FIGURE 4.11. EVEREST display for live HF link, connected to Kineplex parallel-tone modem.

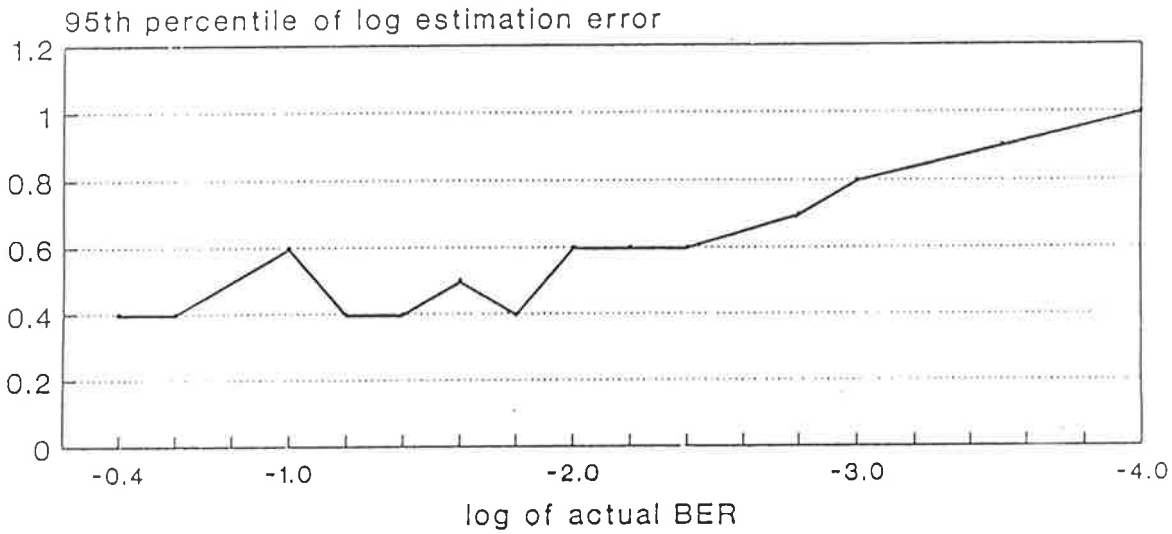


FIGURE 4.12. Live Performance of EVEREST for QPSK, 3000 km HF link, AWGN and Rayleigh fading models, N=3750, M=4, WLS test.

Measurements during the entire three week period were made with the error probability estimated using integer values in library model bins. The following figures 4.13 to 4.16 show the results segmented into four mutually exclusive sections based on the time of day. "Ionospheric sunrise" was defined as the period from 1 hr before to 1 hr after actual sunrise at the Melbourne site. "Daytime" was defined from the end of ionospheric sunrise to the start of "ionospheric sunset". Ionospheric sunset was defined as from 1 hr before actual sunset to 1 hr after. "Night time" was the remaining period until the beginning of the next ionospheric sunrise.

There is a good correlation between the estimated and observed BER for all four diagrams. The daytime data show a wider spread of plotted points than the night-time data. This result was not expected because the absence of models for interferers would be expected to cause a wider spread in the night-time results when the level of interference from distant sources would be higher. Although it is possible that the models were not as good a representation of link conditions at daytime compared with night-time, the most probable explanation is that co-channel

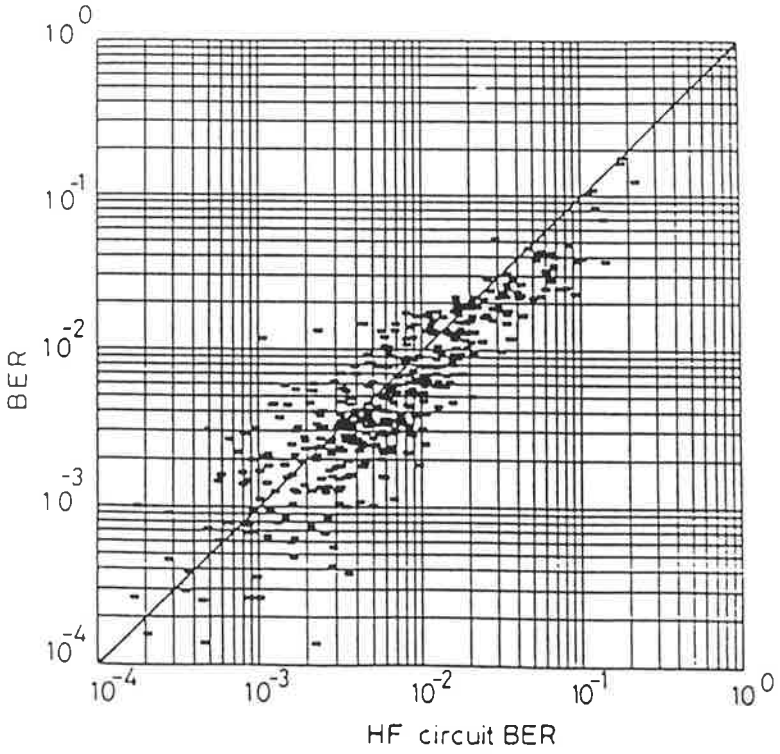


FIGURE 4.13. Performance of EVEREST during "ionospheric sunrise" periods.

occupancy due to local transmitters, which was observed to be common during the daytime, was responsible because no models were included for this. Time restrictions in trial preparation precluded the incorporation of such models.

The poorer performance during ionospheric sunset conditions was observed to be caused by interference from FSK, skywave HF radar and voice.

Improved overall performance would be expected by changing to a maximum likelihood test and incorporating models for impulsive noise and interferers.

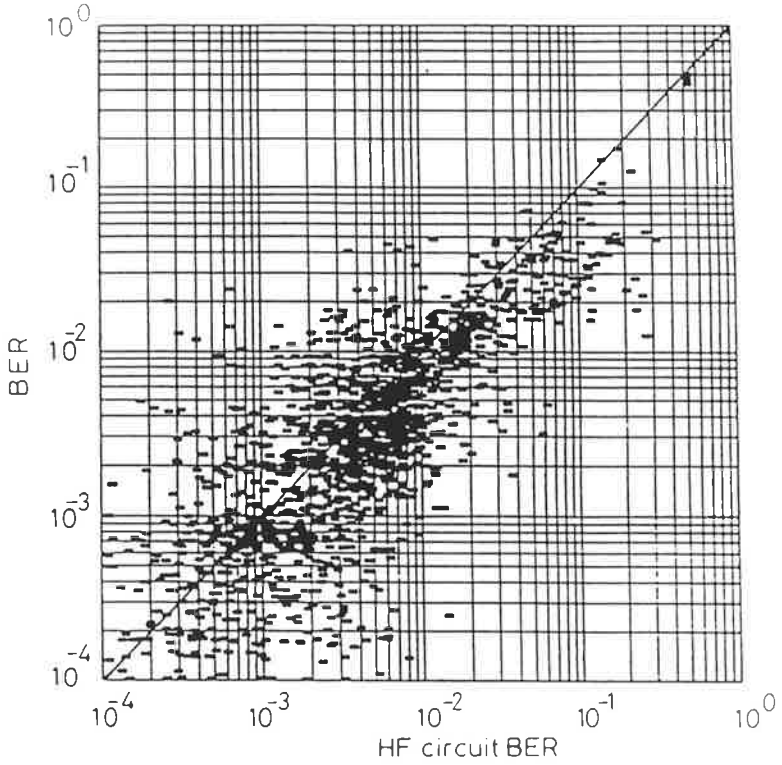


FIGURE 4.14. Performance of EVEREST during "day-time" periods.

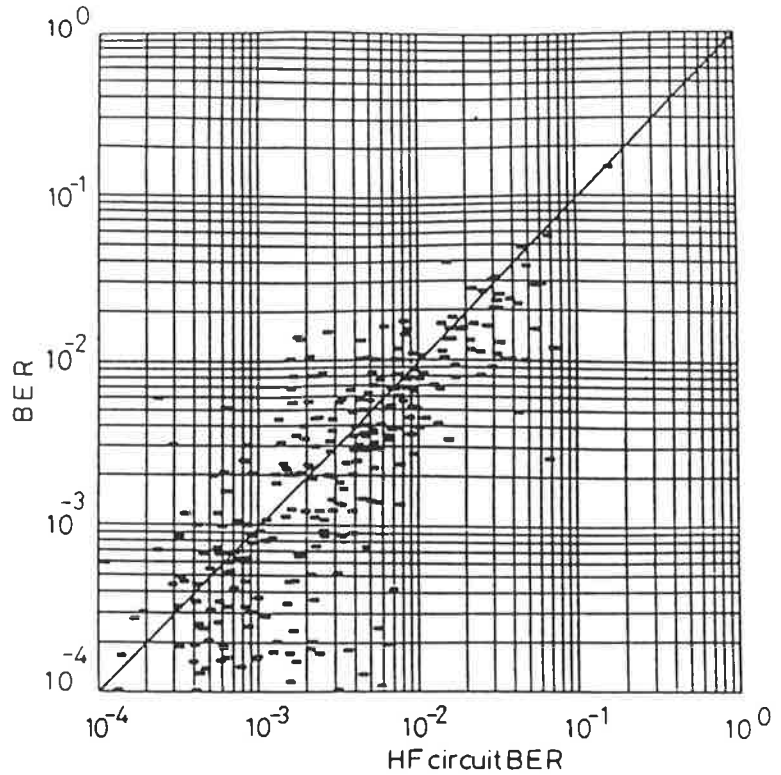


FIGURE 4.15. Performance of EVEREST during "ionospheric sunset" periods.

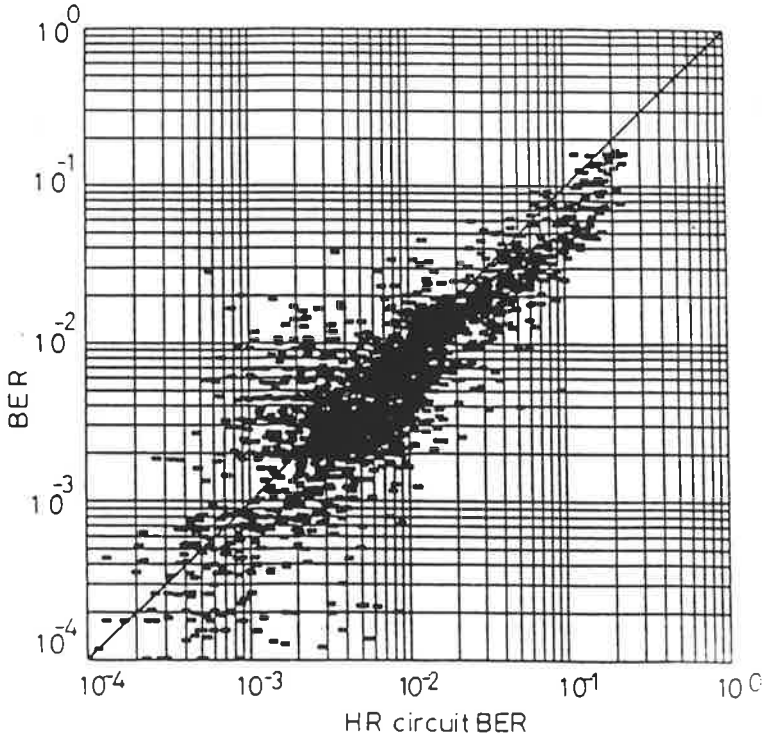


FIGURE 4.16. Performance of EVEREST during "night-time" periods.

4.4 Interfering Signals

4.4.1 Carrier (or Slow Antipodal Modulated) Interference with AWGN, Binary Signalling, Coherent Detection

In this situation the channel is assumed to be AWGN with an additive interference. The interfering signal x(t) is assumed to be a sinusoid with constant (but arbitrary) amplitude A_x, phase ϕ_x and frequency ω_x, over the period of measurement or may be a slow (with respect to the signalling rate) antipodal modulated signal (eg. morse). In this channel, the decision variable at time t=iT (i=1,2,...,N), as earlier, is given by:

$$U_{1,2}(t=iT) = \mu_{1,2} + R_{SN}(iT) + R_{SX}(iT) \qquad (4.31)$$

The decision variable is thus the sum of a constant ($\mu_{1,2} = R_{SS}(iT)$) and two independent random variables R_{SN} and R_{SX}. The probability density function of the decision variable is thus the convolution of the densities of the two independent random variables, as shown by Papoulis (1984):

$$p_u(u) = \int_{-\infty}^{\infty} p_N(u-x-\mu)\ p_x(x)\ dx \qquad (4.32)$$

The density of $\mathbb{R}_{SN}$ is zero-mean Gaussian with variance σ^2 as follows:

$$p_N(n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(- \frac{n^2}{2\sigma^2} \right) \qquad (4.33)$$

With the assumptions stated, the envelope of the random variable $\mathbb{R}_{SX}$ follows a sinusoidal curve and $\mathbb{R}_{SX}$ has a "sinusoidal" probability distribution (Papoulis,1984:p100) given by:

$$p_x(x) = \frac{1}{\pi\sqrt{k^2 - x^2}} \qquad (4.34)$$

Where k is a function of ω_x and A_x, which is a constant representing the maximum interference level which is the peak of a sine wave. It should be noted that (4.34) will also apply if the x(t) is a slow binary antipodal modulated signal (ie. with a signalling period « T) and a sufficient number of samples are taken. Then,

$$p_u(u) = \int_{-k}^{k} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(- \frac{(u-x-\mu)^2}{2\sigma^2} \right) \left(\pi \sqrt{k^2 - x^2} \right)^{-1} dx \qquad (4.35)$$

No analytical solution to this integral could be found. However, the cumulative distribution needed to calculate library models, which is P(u≤a;k), can be expressed as:

$$P(u \le a;\ k) = \int_{-\infty}^{a} p_u(u)\ du \qquad (4.36)$$

Substituting equation (4.35) into (4.36) and then swapping the order of integration yields:

$$P(u \le a;\ k) = \int_{-k}^{k} \left(1 - Q\left[\frac{a-x-\mu}{\sigma} \right] \right) \left(\pi \sqrt{k^2 - x^2} \right)^{-1} dx \qquad (4.37)$$

Let $x = k \sin\theta$, then $dx = k \cos\theta\ d\theta$. Substituting this into (4.37) and simplifying:

$$P(u \leq a;\ k) = 1 - \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} Q\left\{(a - k\sin\Theta - \mu)/\sigma\right\}\, d\Theta \qquad (4.38)$$

Therefore using (4.38) and (4.7), library models may be determined for a given signal to noise ratio (given by (4.5)) and Interference to Noise Ratio (INR) (defined as k/σ).

An even number of equally spaced bins was chosen between the decision threshold and twice the mean, with a single additional bin extending from twice the mean to infinity as shown in figure 4.2. Figure 4.17 shows the EVEREST performance with sinusoidal or slow antipodal interference and AWGN. The library models were derived using equations (4.38) and (4.7). Equation (4.38) was evaluated numerically. The performance shown is the worst-case performance since the EVEREST must choose from a two-dimensional set of models, specified by the SNR and the INR and hence an entire family of curves may be drawn for each INR value over the range of BER values.

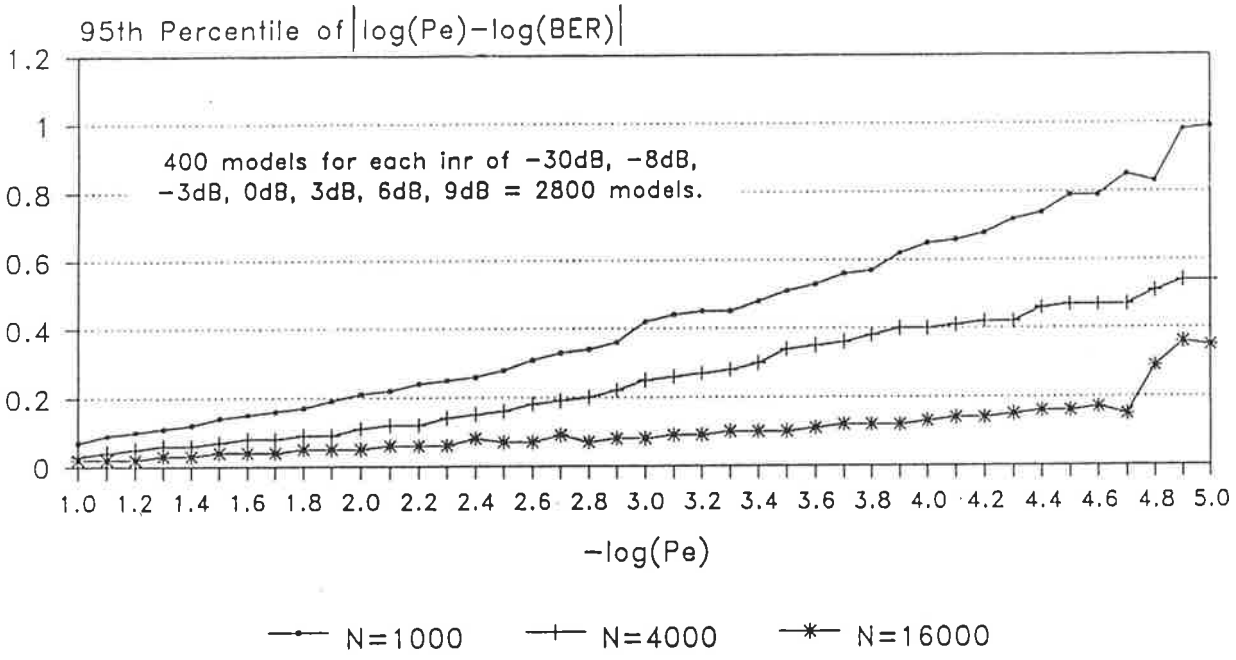


FIGURE 4.17 Worst-case EVEREST performance with M=9 regions, N=1000, 4000 and 16000 samples, sinusoidal (or slow antipodal) interference with AWGN channel model.

4.4 Chapter Summary

This chapter has detailed simulation results and given some example approaches to the design of an error probability estimator for a number of modulation and channel types.

The performance of the technique is strongly dependent on the ability to discern between library models. Thus, in order to design a monitor for arbitrary channel and modulation types, for a given sample size and implementation complexity, the use of simulation is recommended. The form of a general simulation program to achieve this has been given.

Simulation results have been supported by live trial results on a HF link using parallel-tone modems, demonstrating the ability to estimate error probability to an accuracy previously unattainable with such a small number of samples, regardless of the method used.

The EVEREST performance described has been for cases where no additional supporting measurements were taken. Clearly, performance improvements are possible if supplementary measurements are made of relevant parameters, which may then be used to direct the EVEREST classifier towards more probable model types instead of "blindly" searching the entire model space. For example, with the case of AWGN and Rayleigh fading models, if measurements of the received signal strength indicate a steady signal it would direct the EVEREST to ignore Rayleigh fading models.

Chapter 5

LINK PERFORMANCE OPTIMISATION BASED ON THE EVEREST ESTIMATION TECHNIQUE

Abstract

This chapter describes a unique feedback link control scheme that is made possible by the application of the EVEREST concept. The optimisation criteria is maximisation of data throughput for a guaranteed quality of service (maximum specified post-decoder error probability). This is achieved using a feedback control strategy, which adapts error control code rate.

The new scheme has been designed for a general fading channel model of the Watterson type and using short block codes. The performance of the strategy is compared with conventional fixed-code schemes and other forms of ARQ, demonstrating a significant throughput increase at very low signal to noise ratios.

The scheme may be employed as an applique to any parallel-tone modem.

5.1 Introduction

The introduction of fixed forms of error control coding has made a significant impact on the error performance of digital communication systems. However, a single code alone cannot provide optimum performance when faced with dynamic changes in channel quality and user requirements. Automatic Repeat reQuest (ARQ) schemes maintain acceptable maximum error rates over various channels but suffer very poor throughput performance at low signal-to-noise ratios.

Those digital communication links operating over relatively benign channels (eg. fibre-optic cable) have generally been well optimised and give reliable, virtually error-free performance. In contrast, "difficult" digital radio links such as mobile, High Frequency (HF) skywave, meteor burst and Super High Frequency (SHF) troposcatter often display non-random errors and dynamic error rates. This dynamic nature may be exploited in order to increase throughput.

Many theoretical studies have been performed on the potential use of feedback control for time-varying communication links. Most of these schemes have suffered from over-simplifications in the modelling of channel conditions and modems so that further research would be required before these schemes could be implemented. A wide range of techniques have been proposed when faced with flat fading as in Hayes (1968), Cavers (1972), Skinner and Cavers (1973), Feldman and Li (1988), Katakol and Maskara (1989). In contrast, a literature search yielded no papers on feedback communication systems (except for ARQ) which accounted for multipath conditions. In multipath environments, fading is frequency selective, so that these simple models are inadequate for both mobile radio and HF radio applications. Also, the presence of non-Gaussian noise and interferers has not been accounted for in these models.

Two categories of Adaptive Link Control (ALC) strategies may be identified. The first of these involves changing the receiver structure only. No change to the transmitted information is made. Such "self-adaptive" methods have been common to HF serial-tone modem technology in which the pre-filter or equaliser is updated by sending blocks of known data between normal data traffic for training purposes as

in Jayasinghe (1989). Another class of "self-adaptive" schemes involves changing the detector structure dynamically, to cope best with current channel noise and interference conditions such as those described by Hancock (1963) and Groginsky (1966).

Self-adaptive schemes will not be considered further, since it is believed that the greatest gains may be achieved by the second category of ALC strategies, the "feedback" system.

Feedback schemes involve changing the transmitted information or signal to suit prevailing channel conditions and user needs. Change in transmission parameters must be instructed by the receiver, whose structure is also altered to suit. Feedback schemes include:

- Re-transmission Schemes. Goodman and Farrell (1975)
  - ARQ
  - Hybrid-ARQ
  - Adaptive Hybrid-ARQ
- Transmit power change.
- Data rate change.
  - Bandwidth
  - Constellation packing
- Interleaving length change.
- Selective choice of channel diversity.
  - Frequency
  - Space (antennas)
- Change in the class of modulation scheme.
- Change in error control code rate.

The use of feedback systems is limited to applications where two-way communication is used, or a dedicated feedback link is available. The benefits of such systems appear not to have been so widely recognised. The benefits to the communications *user* as identified by Cook, Scholz and Giles (1990) include:

- *Reduced Waiting Time* - for systems which optimise their throughput, the transmission time is consequently minimised.

- *Optimising Link Performance for the Particular Service Type* - Throughput may be optimised whilst maintaining a ceiling on the maximum permissible error rate for the current service type.

Another possible benefit for the military user is *Lowering the Probability of Signal Interception* (LPI), which may be achieved say, by dynamically minimising the transmitted power whilst maintaining a required maximum error rate.

5.2 Feedback ALC Schemes Based on the EVEREST Estimate

The general feedback link control philosophy is:

(i) Estimate link environment pdf using EVEREST technique.

(ii) Calculate link error performance for various changes in transmission parameter(s).

(iii) Decide on the appropriate parameter change.

(iv) Feedback the decision.

In order to *track* a dynamically changing channel, the time required to estimate the nature of the pdf, devise the best adaptation strategy and engineer the change must be around twenty times faster than the rate of change of the channel (Cavers,1972). The high speed of pdf estimation provided by the EVEREST makes this possible. It is believed that link pdf estimation has not been applied to the dynamic control of link performance before and is the subject of a provisional patent by Scholz and Cook (1991).

5.3 Simulation of an Adaptive Rate Strategy for a Fading HF Channel

In a HF channel, fading is the principal error mechanism and must therefore be tracked. Noise (and possibly interference) will also be present and causing errors.

On a HF link, the distance between sites may be a few kilometres ranging to several thousand kilometers, causing propagation delays to vary

accordingly. As far as a feedback communication system is concerned, path delay is a most significant parameter affecting performance as recognised by Cavers (1972). A propagation delay of around 10 ms would be expected on a 3000 km link and may be considered as a worst-case value. The transmission path delay is not the only significant delay in such a system, delays due to transmit and receive filters are also significant. For a voice-band HF link, a value of 5 ms worst-case would be reasonable. These facts strengthen the need for fade tracking and prediction.

The system examined will have a choice of R error correcting codes to use. Any one of these codes may be used as error protection for the duration of one block. The code to be used on any given block will be determined from a prediction of the expected channel state at that time. For convenience, BCH block codes were used since a large range of code rates are available even for short codes. Software for a BCH coder and decoder was available, written by Martin Gill. A thorough investigation of the best FEC codes suited to this application was not done since it was intended only to demonstrate the principle. There is no reason why Rate-Compatable Punctured Convolutional (RCPC) codes such as those described by Hagenauer (1988) could not be used. The primary disadvantage of using a convolutional code in this system is that the decoder delay can become severe during periods of very poor channel quality. In this system, each block code will produce a codeword N bits in length, but each will have a different number of information bits, k. As each code produces codewords of the same length the data rate for the modulator is constant. The choice of codes should be such that a range of code rates are available to cover all possible channel conditions.

The block lengths are required to be considerably shorter than the "period" at which the channel conditions change, so that these changes can be *tracked*.

Figure 5.1 shows the concept.

In order to choose the most appropriate code for a future block of data, it is necessary to predict the expected number of errors in that block. Error prediction requires some knowledge of the error mechanisms experienced by the link.

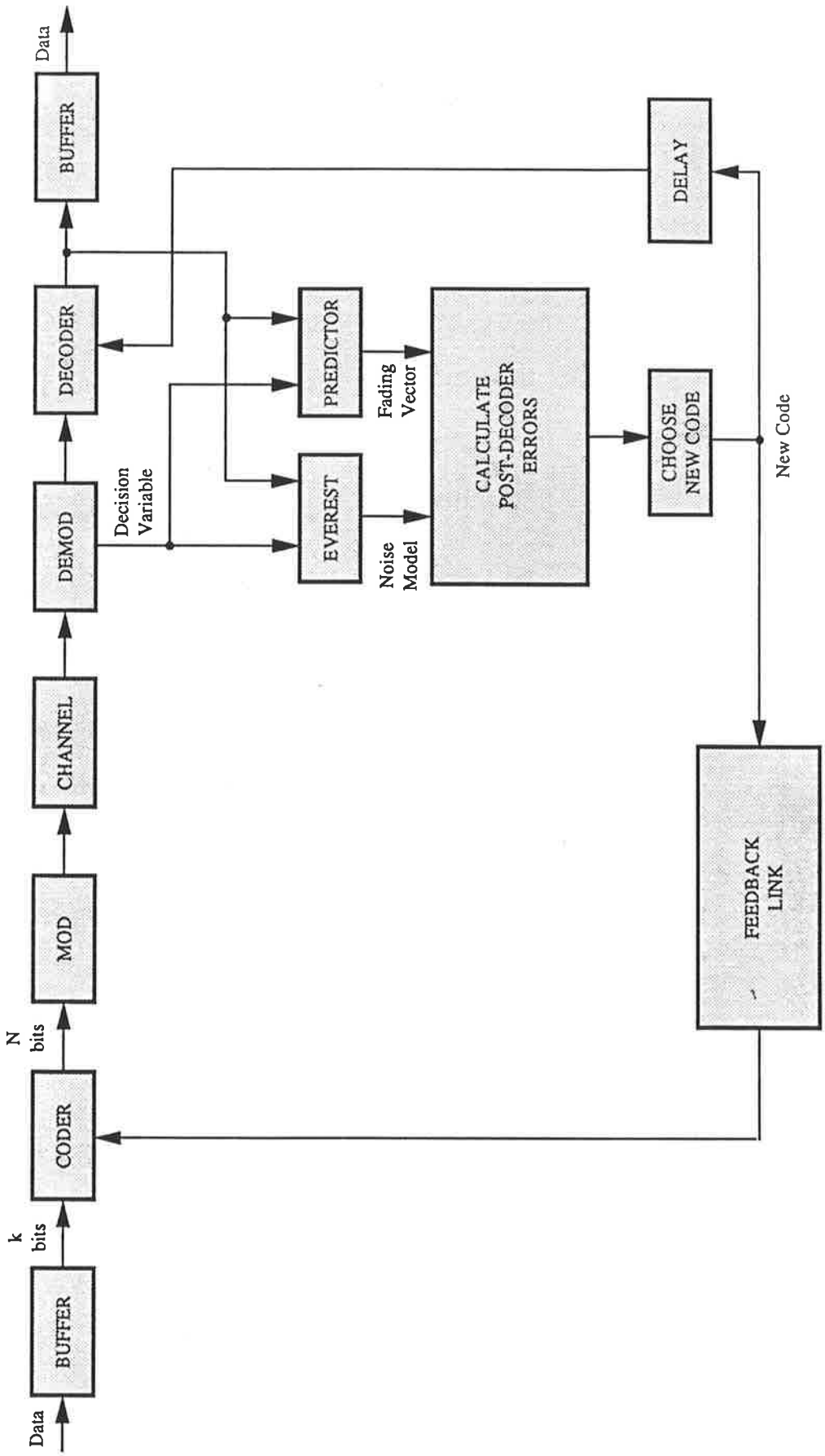


FIGURE 5.1. Feedback communications system with adaptive FEC codes.

Errors at link level may occur due to a variety of mechanisms. However, these mechanisms may be classed into one of the three general categories:

- Random Noise Processes (Generally unpredictable)
  - AWGN
  - Atmospheric
  - Galactic
  - Ignition

- Interfering Signals (generally man-made in origin)
  - Digital modulations
  - Voice

- Distortions
  - Fading
  - Multipath

The statistics of errors due to random noise processes would generally be slow varying in nature and therefore be effectively estimated (without the need for prediction) using the EVEREST technique as demonstrated in previous chapters.

To track fading in the channel, the length of the block should be much smaller than the shortest fading period. Over HF links, the fade rate may vary considerably over a range of tens of milliseconds to tens of minutes according to Maslin (1987). According to the Commité Consultatif Internationale Radio: CCIR (1982), a fading period of around 1 s worst case may be expected. Shorter fading periods up to 10ms are possible due to spread F conditions in the ionosphere, however, these effects occur relatively infrequently.

It is desirable to choose block lengths less than 1/10 th of the worst fading period to allow reasonable fade tracking, however, increasing size allows the application of more powerful FEC codes (Lin and Costello, 1983).

The HF modems to be used with the link operate at a transmission rate of 2400 bits/s. The Magnavox modem frames data into 32 bit words, so a multiple of 32; say 128 bits (T=53.3ms) and 64 bits (T=26.6ms) block

lengths were examined as an appropriate compromise between error control capability and block-length-induced delay.

Figure 5.2 shows the events occuring in the system for blocks of length T milliseconds. This diagram assumes that data is always available in a buffer such that when the new code rate has been determined, k information bits can be directly taken from the buffer, the code rate index added and the entire block coded for transmission with negligible delay, δ.

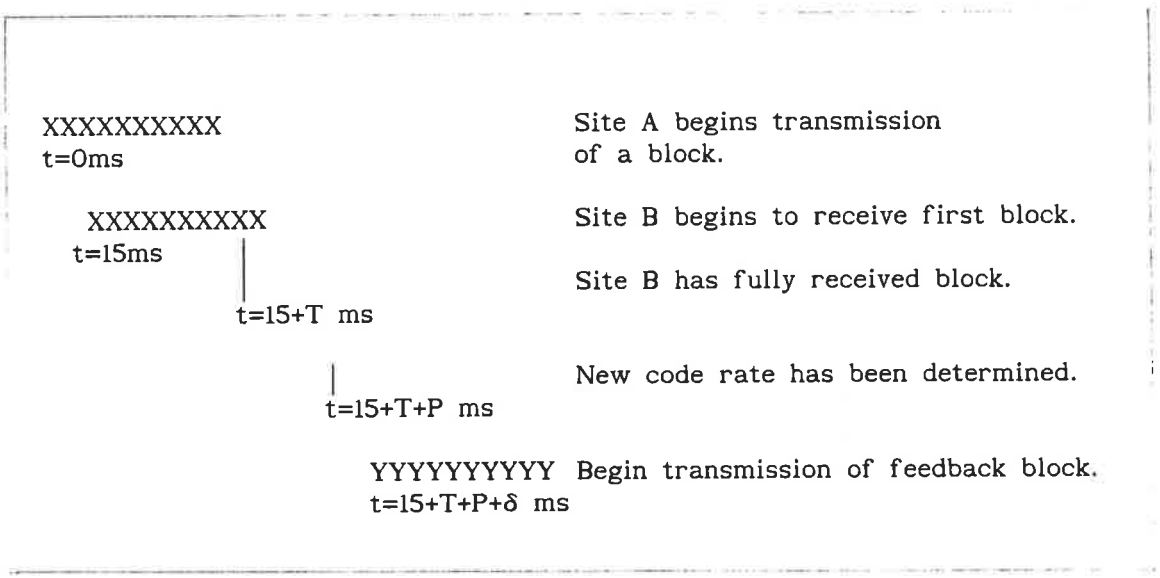


FIGURE 5.2. Sequence of events in feedback communications system for signalling with blocks of T (ms), and processing delay of P (ms).

Thus, the predictor calculation performed using the entire block beginning at time t=0ms needs to be for the block which is beginning transmission from site A at time t_p milliseconds later:

$$t_p \cong 15 + T + P + \delta + (15 + T + \delta)$$

$$= 2T + P + 2\delta + 30 \qquad (5.1)$$

Take P≅6ms and δ=0 ms. Then, for 128 bit blocks (T=53.3ms) $t_p \cong 143$ ms; which implies predicting 3 blocks ahead. For 64 bit blocks (T=26.6ms), giving $t_p \cong 90$ ms which implies predicting 4 blocks ahead.

5.3.1 EVEREST Design and Simulated Performance

Consider the general case of an M symbol transmitted alphabet. The transmitted signal may be expressed in the form:

$$s(t) = u(t)\ \exp(j\omega_c t) \tag{5.2}$$

Where: $u(t)$ = Complex valued signal pulse representing a symbol.

In most modems designed to operate over fading channels, differential data modulation is employed, so that the phase imparted by the channel on the signal is simple to remove. In which case, the data is recovered by calculating the product of two successive vector matched filter outputs. This decision variable is then thresholded to extract the symbols. The matched filter produces:

$$\mathbf{V}(T) = \int_{T-1}^{T} r(t)\ u^{*}(t)\ dt \tag{5.3}$$

Where: $r(t)$ = Received baseband signal.

T = Symbol period.

$*$ Denotes conjugation.

The decision variable is:

$$\mathbf{U}(T) = \mathbf{V}(T)\ \mathbf{V}^{*}(T-1) \tag{5.4}$$

In general, fading, noise and interference will corrupt the transmitted signal. The channel model used for the fading is by Watterson et. al. (1970). This model assumes that the tap gain functions have Gaussian spectra. However, as this assumption has been disputed by Vogler and Hoffmeyer (1990) and also by Serrat-Fernandez et. al. (1985), and since it would impact on the predictor design, the tap gain spectra are taken to be arbitrary. The received signal may thus be described as follows:

$$r(t) = \left(\sum_{w=1}^{H} \alpha_w(t)\exp(-j\ \phi_w(t)) \right) u(t) + n(t) + z(t) \tag{5.5}$$

Where: $\alpha_w(t)$ = Rayleigh fading amplitude for path w.

$\phi_w(t)$ = Phase for path w.

n(t) = Noise process.

z(t) = Direct (non-fading) interference.

H = Number of multipath components.

Substituting (5.5) into (5.3) gives:

$$V(T) = \int_{T-1}^{T} n(t)\, u^{*}(t)\, dt + \int_{T-1}^{T} z(t)\, u^{*}(t)\, dt +$$

$$\sum_{w=1}^{H} \int_{T-1}^{T} \alpha_w(t) \exp(-j\, \phi_w(t))\, u^{*}(t)\, u(t)\, dt \qquad (5.6)$$

Over a symbol interval, it is assumed that the Rayleigh fading amplitude $\alpha_w(t)$ and the phase $\phi_w(t)$ will be constant. This will be true if the symbol period is short compared with the rate of fading. The matched filter output becomes:

$$V(T) = \mathbb{R}_{SN}(T) + \mathbb{R}_{SZ}(T) + \sum_{w=1}^{H} \alpha_w(T) \exp(-j\, \phi_w(T))\ \mathbb{R}_{SS}(T) \qquad (5.7)$$

$$\{ T = 1,2,\ldots,N \}$$

Where: $\mathbb{R}_{SN}$ Denotes the cross-correlation between the transmitted signal and the noise.

$\mathbb{R}_{SZ}$ Denotes the cross-correlation between the transmitted signal signal and interfering signals.

$\mathbb{R}_{SS}$ Denotes the correlation between the transmitted signal and the expected signal.

The $\mathbb{R}_{SN}$ and $\mathbb{R}_{SZ}$ components of the decision variable will be independent random variables. The $\mathbb{R}_{SS}$ component will take one of M values. For the case of a QPSK signalling scheme, M=4 and $\mathbb{R}_{SS}$ = $1\angle 0$, $1\angle 90$, $1\angle 180$, $1\angle 270$. However, the term;

$$F(T) = \sum_{w=1}^{H} \alpha_w(T) \exp(-j\ \phi_w(T))$$

$$= \alpha(T)\ \angle\phi(T) \qquad (5.8)$$

assumed constant for a single symbol period, will vary from one symbol period to the next.

In this system, the last component of equation (5.7):

$$G(T) = F(T)\ \mathbb{R}_{SS}(T) \qquad (5.9)$$

is subtracted from the matched filter output. This is achieved by firstly taking the data from after the decoder, re-encoding it and re-modulating it to derive an estimate of the data vector $\mathbb{R}_{SS}(T)$. An estimate of the fading vector $F(T)$ is derived by low-pass filtering the matched filter output $\mathbf{V}$, to remove the noise. The product of the data vector estimate and fading vector estimate is then subtracted from equation (5.7).

The remaining terms from equation (5.7) are then:

$$\mathbf{V}'(T) = \mathbb{R}_{SN}(T) + \mathbb{R}_{SZ}(T) \qquad (5.10)$$

The EVEREST uses this "filtered" decision variable to estimate the combined noise and interference pdf. For the initial simulation, the interference term was ignored. Thus:

$$\mathbf{V}'(T) = \mathbb{R}_{SN}(T) = \rho(T) + j\ \upsilon(T) \qquad (5.11)$$

The noise vector terms $\mathbb{R}_{SN}(T)$, were assumed independent and identically distributed as zero-mean Gaussian. Thus, the distribution of $\mathbf{V}'$ is Gaussian with variance σ^2 in each orthogonal axis. To simplify calculation, a random variable transformation was applied reducing the distribution to one dimension. The Rayleigh transformation was applied as follows (Proakis, 1989:p29):

$$\mathbb{J}(T) = \sqrt{\rho(T)^2 + \upsilon(T)^2} \qquad (5.12)$$

The distribution of $\mathbb{J}$ is therefore Rayleigh, as follows:

$$p(x) = \frac{x}{\sigma^2} \exp\left(- \frac{x^2}{2\sigma^2} \right) \qquad x \geq 0 \qquad (5.13)$$

It was decided to use the continuous-form maximum likelihood test for EVEREST. The EVEREST chooses model s=r for which $\rho(r)$ is a maximum, repeating equation (3.45):

$$\rho(s) = \sum_{i=1}^{N} \ln p(x_i \mid M_s) \qquad (5.14)$$

Substituting (5.13) into (5.14):

$$\rho(s) = \sum_{i=1}^{N} \ln \left(\frac{x_i}{\sigma_s^2} \exp\left[- \frac{x_i^2}{2\sigma_s^2} \right] \right) \qquad (5.15)$$

Simplifying, gives the EVEREST test, to choose model s=r for which $\rho(r)$ is the maximum value of $\rho(s)$, over all possible values of s.

$$\rho(s) = - \sum_{i=1}^{N} \left[2 \ln \sigma_s + \frac{x_i^2}{2\sigma_s^2} \right] \qquad (5.16)$$

In this case, the noise models are completely specified by the Gaussian variance σ_s^2 ,(s=1,2,...,L).

On an HF channel the range of instantaneous energy per bit to noise ratio may extend to at least 30dB. The range chosen for simulation was 0dB to 30dB. The minimum necessary model resolution that should be used is difficult to ascertain, so an excess of models (L=300) were used, giving a high resolution of 0.1 dB.

5.3.2 Calculation of Post-Decoder Error Probability and Choosing the Optimum Code

Before calculating the post-decoder error probability, the predicted pre-decoder error distribution was calculated. This was achieved by

combining the EVEREST-estimated pdf model and the predicted fading vector then calculating the area of the pdf tail beyond the decision thresholds. In the parallel-tone modem, there are sixteen tones and this calculation is done for each tone. This is shown in figure 5.3.

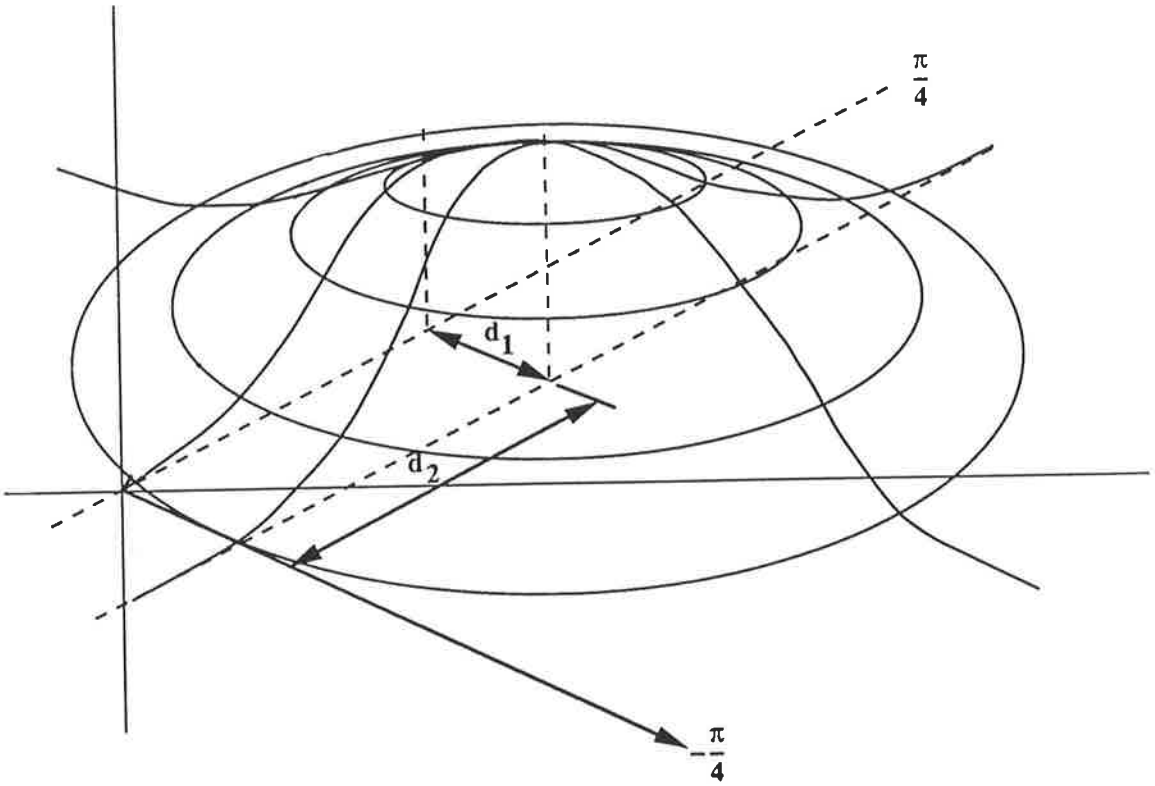


FIGURE 5.3. Calculation of the future probability of error for each QPSK tone.

Due to the fact that each tone represents two bits, each tone could contain zero, one or two errors. The calculation of pre-decoder *bit* error probability is thus, with aid of figure 5.3:

$$\begin{aligned} P(\text{ zero errors }) &= Q_1 Q_2 \\ P(\text{ one error }) &= P_1 Q_2 + P_2 Q_1 \\ P(\text{ two errors }) &= P_1 P_2 \end{aligned} \qquad (5.17)$$

Where: $P_1 = Q(d_1)$; $P_2 = Q(d_2)$

$Q_1 = 1 - P_1$; $Q_2 = 1 - P_2$

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp(-z^2/2)\, dz$$

This calculation is valid due to the independence of the real and imaginary components. If we further assume that the noise components on each of the 16 tones are independent, the composite pre-decoder probability of getting N errors per block may be calculated from the convolution of the 16 distributions each calculated as in equation set (5.17) for each predicted fading vector.

The transformation from the pre-decoder random variable to the post-decoder random variable is, as expected, a linear relationship, when the number of pre-decoder errors exceeds the number correctable as depicted in figure 5.4 and verified by simulation for all N=127 BCH codes used in later simulations.

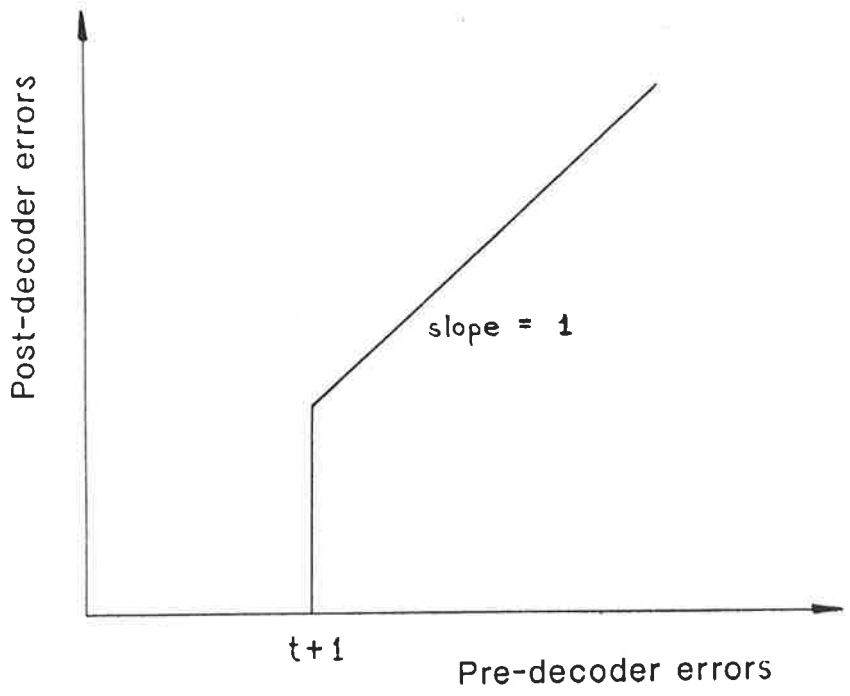


FIGURE 5.4. BCH decoder "transfer function".

The transformation results in the post-decoder error probability for j errors; $P_a(k)$, as a function of the pre-decoder error probability P_b, as

follows:

$$P_a(0) = \sum_{k=0}^{t_l} P_b(k)$$

$$P_a(k) = 0 \quad ; \quad 0 < k \leq t_i$$

$$P_a(k) = P_b(k) \quad ; \quad k \geq t_i + 1 \qquad (5.18)$$

Where: t_i = The number of errors correctable by code i.

The expected number of post-decoder errors is then, in terms of the pre-decoder error distribution:

$$\overline{P} = \sum_{k=t_l+1}^{N} k\, P_b(k) \qquad (5.19)$$

The post-decoder bit error probability is then:

$$P_e = \frac{1}{N}\, \overline{P} \qquad (5.20)$$

The algorithm is then to choose the highest rate code that results in an acceptable post-decoder error probability.

If having tested the most powerful code, P_e still exceeds the maximum allowable error probability, data transmission is halted and a known test sequence is sent, until the channel recovers.

Random errors will be caused by additive Gaussian noise and atmospheric noise. The statistical nature of these noise processes would not be expected to change appreciably over a period of minutes (CCIR, 1986). Interference of a wide variety may be expected with voice, morse and carrier tones being most common. These interferers may come from distant sources and be undergoing fading making them difficult to predict accurately.

For simulation, the effect of interferers was ignored, however, the effect is easily taken into account since the EVEREST can be used to estimate the decision variable pdf, with interference present, by including these

models, in the library set.

5.3.3 Results

Figures 5.5 and 5.6 show the results of simulation of the performance of the adaptive code rate strategy for a Rayleigh fading channel with AWGN. This figure shows the throughput performance attained for a guaranteed maximum post-decoder error probability of 10^{-4}. The simulation is for a single DQPSK tone signalling at 75 symbols/s. Average flat-fade rates of 0.1Hz and 0.5Hz were used, corresponding to CCIR "good" and "moderate" respectively, as per recommendation 520-1 CCIR (1982). These results were obtained using either the N=127 or the N=63 BCH codes given in table 5.1:

N	Feedback code Index	Code Rate (k/N)	Number of bits Correctable (t)
127	0	1	0 (no code)
127	1	120/127	1
127	2	99/127	4
127	3	71/127	9
127	4	43/127	14
127	5	29/127	21
127	6	8/127	31
127	7	0	transmit no data
63	0	1	0 (no code)
63	1	57/63	1
63	2	45/63	3
63	3	30/63	6
63	4	18/63	10
63	5	10/63	13
63	6	7/63	15
63	7	0	transmit no data

TABLE 5.1. Specifications of the binary BCH codes used for simulations with 127 and 63 bit blocks (code words).

Three binary digits (out of N) are required in each feedback packet to signal the new code to the transmitter. This reduction in throughput has been accounted for in the presentation of results. Also shown with the adaptive code rate scheme throughput in figures 5.5 and 5.6, is the throughput performance of fixed BCH codes over the same channel. The enormous gain achieved by using feedback control is clear.

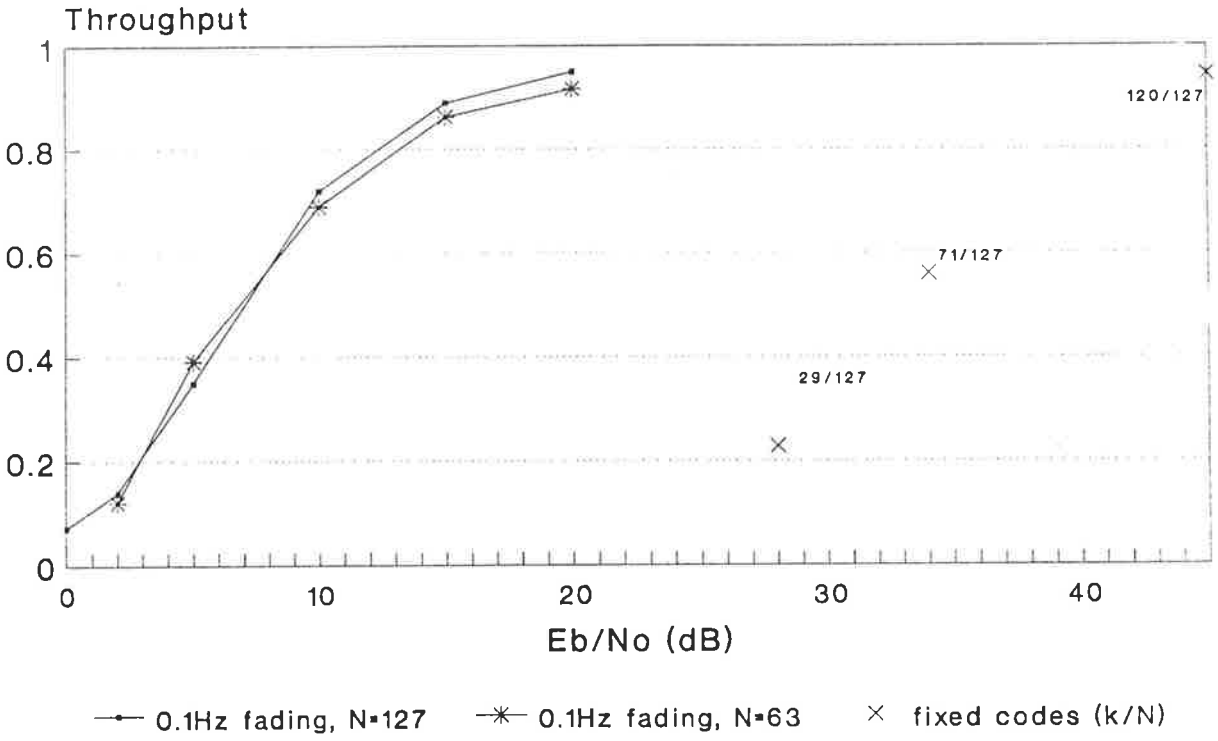


FIGURE 5.5. Throughput performance versus average energy per bit to noise, for a post-decoder error probability of 10^{-4}, 0.1 Hz average fade rate.

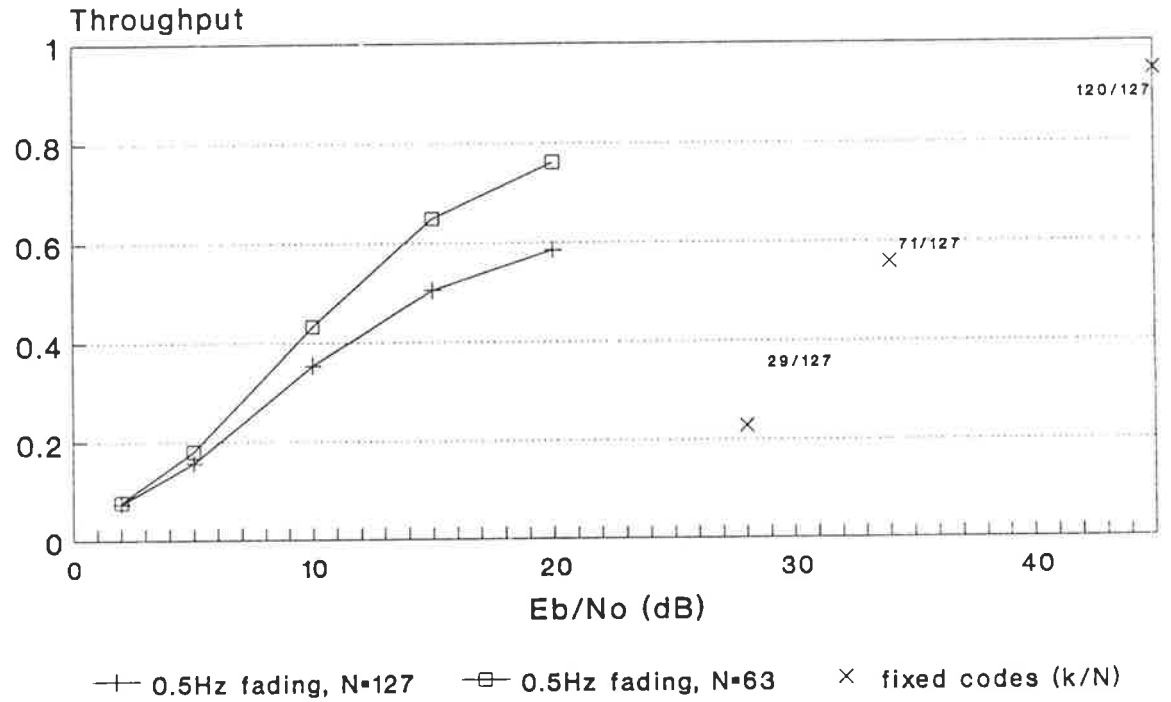


FIGURE 5.6. Throughput performance versus average energy per bit to noise, for a post-decoder error probability of 10^{-4}, 0.5 Hz average fade rate.

Figure 5.7 shows the actual post-decoder error probability obtained using the adaptive code rate strategy. At the higher fade rate of 0.5 Hz, the target error probability of 10^{-4} is not quite met. For the 0.1 Hz fade rate, this error probability is well met. In fact, for the 0.1Hz case, *no errors* were recorded over the entire range of E_b/N_0 from 2dB (22,000 transmitted bits) up to 20dB (137,000 bits).

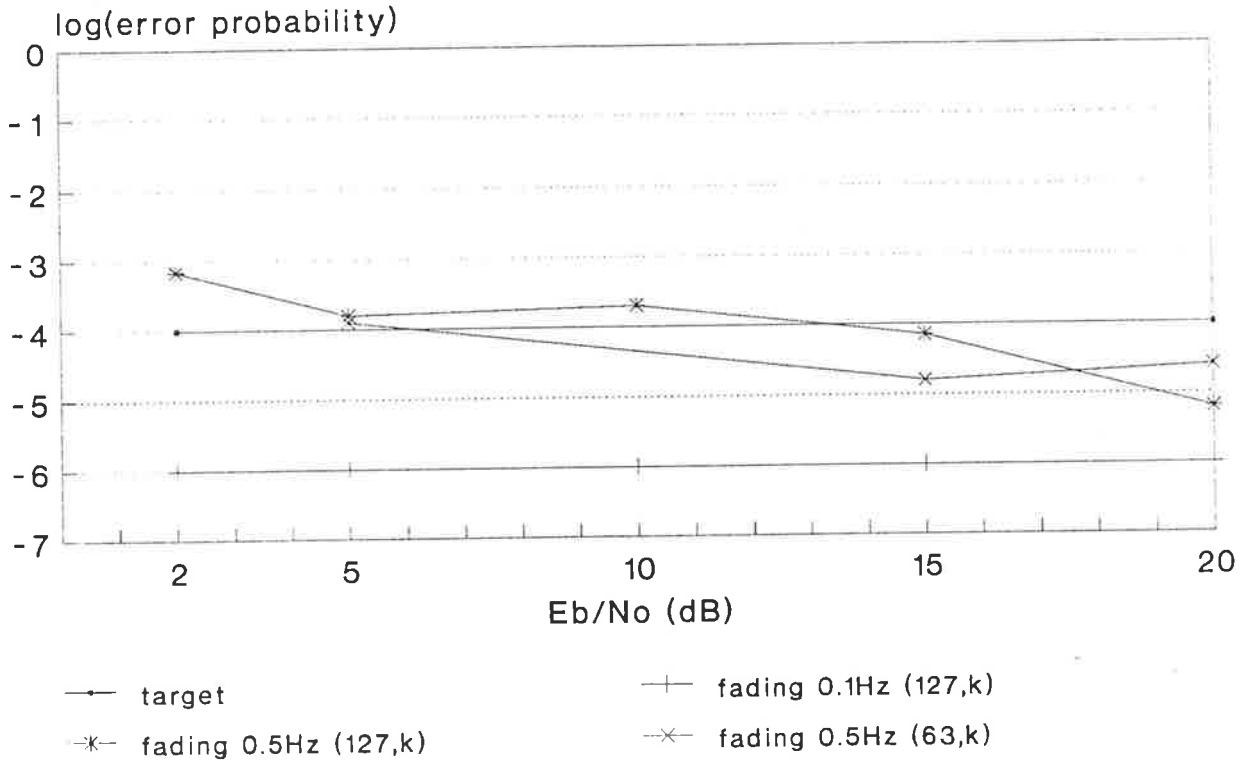


FIGURE 5.7. Post-decoder error performance of Adaptive Code Rate strategy versus average Energy per bit to noise, for a targeted post-decoder error probability of 10^{-4}.

Figure 5.8 shows a further result which gives the *performance limits* of the technique for N=63 using 6 BCH codes as well as on and off. CCIR "moderate" and "poor" channels were used as per recommendation 520-1 (1982), with average fade rates of 0.5Hz and 1.0Hz respectively. The bounds shown are for perfect prediction. Clearly, the bound is nearly reached with the 0.5Hz case but is far off for the 1.0Hz case.

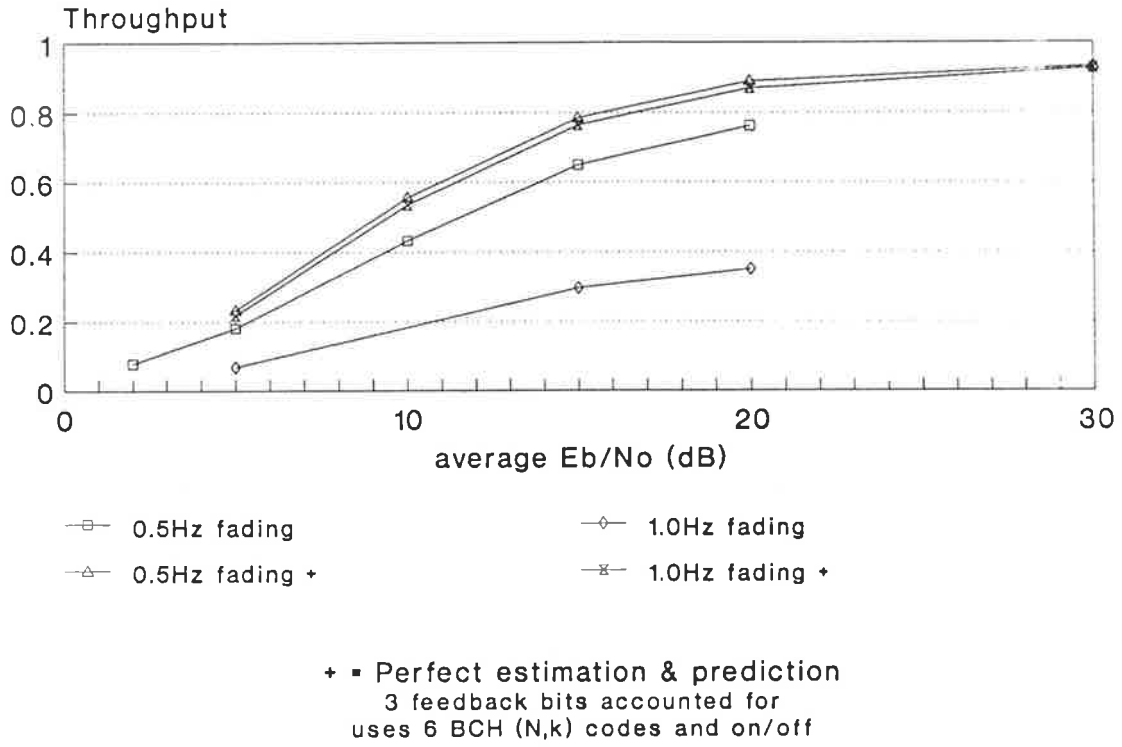


FIGURE 5.8. Performance of Adaptive Code Rate Strategy for CCIR "moderate" and "poor" channels, N=63, post-decoder error probability $\leq 10^{-4}$.

The prediction accuracy may be improved by use of a more sophisticated predictor. The simple linear predictor used was designed and developed by Timothy Giles and further work on this will continue. Since performance is strongly tied to link delays, the use of an engineering channel which could provide instantaneous feedback of code rate information (without the need to wait for link control data to be put into blocks) would significantly improve performance.

Figure 5.9 shows the total usage of each code type over the range 5 to 30dB. The situation is for perfect prediction and to maintain a post-decoder error probability of 10^{-4}. For a large proportion of the time no code is used. However, the true performance gains of the technique are apparent by the proportion of the time that the six codes are used, which occurs particularly for average signal to noise ratios of 5 to 12dB.

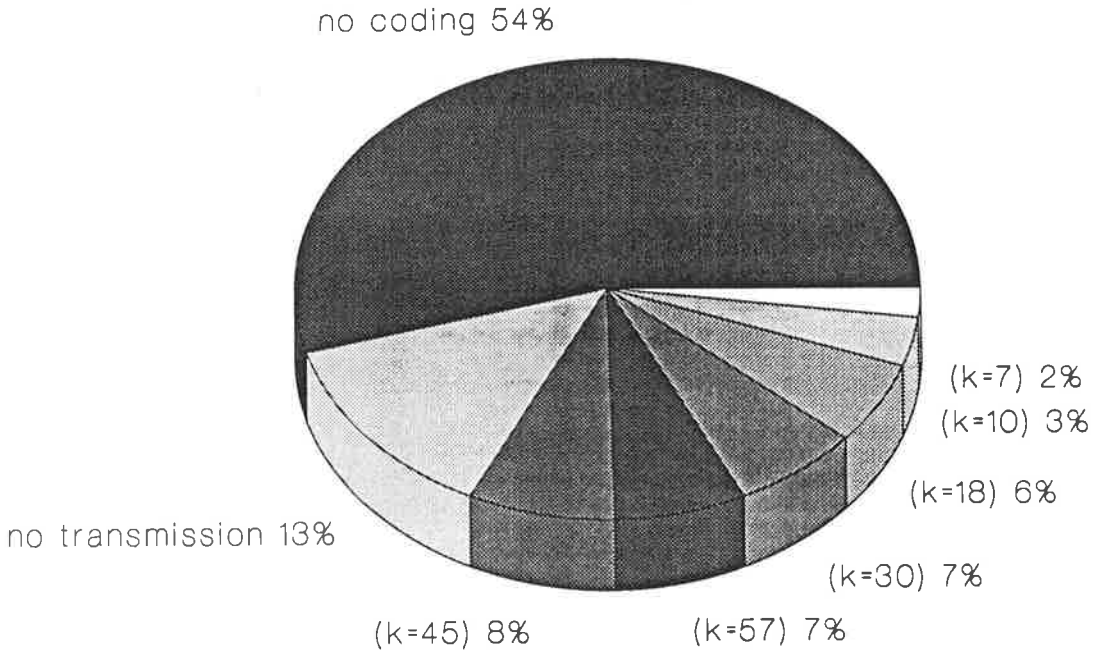


FIGURE 5.9. Adaptive code rate system usage of BCH(63,k) codes over a 5 to 30 dB range. Post-decoder error probability below 10^{-4}, with perfect prediction.

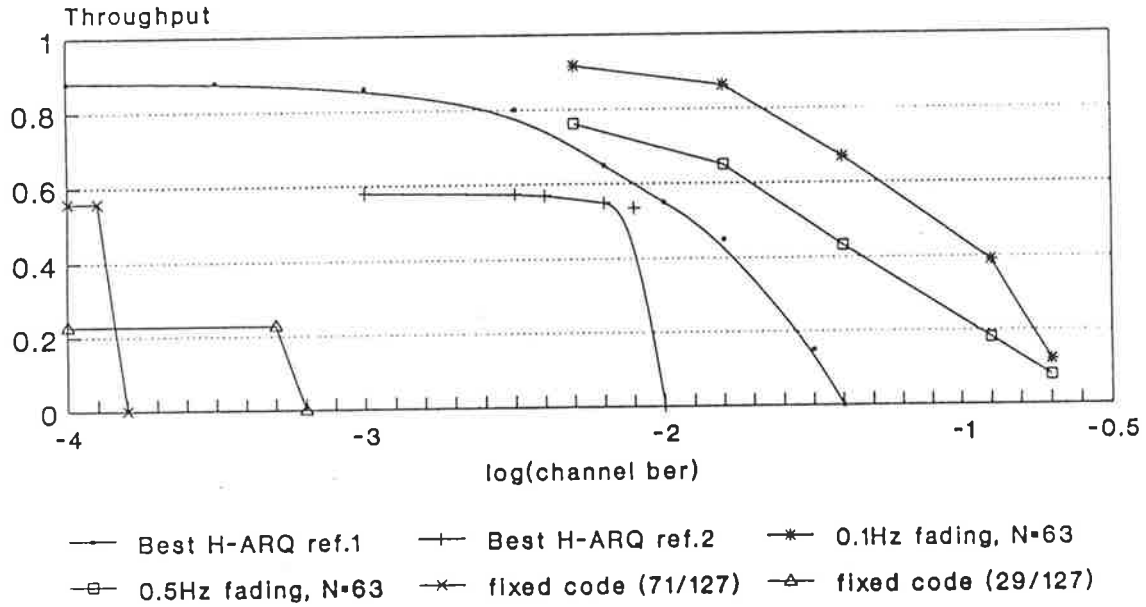
In order to allow direct comparison with two hybrid-ARQ schemes, the average energy to bit ratio was translated to channel BER, for the differential QPSK detection scheme used with Rayleigh fading. Figure 5.10 shows the comparative results with two H-ARQ schemes proposed by Wu et al (1990) and by Dalmau-Royo and Serrat-Fernandez (1991).

Wu et al examined a type-II hybrid ARQ scheme. The scheme used an error detection only code for normal transmission and a half-rate invertible error correction code when a block with errors was detected. Stop-and-wait type ARQ was used and *delays in transmission were ignored.* The link was modelled by a partitioned Markov chain model.

Dalmau-Royo and Serrat-Fernandez proposed three H-ARQ schemes. The

best-performing of these is shown in figure 5.10. The scheme used a BCH (31,11) code but transmitted only 16 of the 31 bits (11 segment and 5 parity bits). The decoder could detect up to 2 errors in each 16 bit block and would retransmit the remaining 15 parity bits (with 1 stuffing bit) if an error was detected. Again, a partitioned Markov link model was used. No details on the nature of the ARQ or assumptions with time delays were disclosed.

Figure 5.10 shows that the adaptive code rate system, which accounts for all time delays, out-performs the other techniques, particularly at high error rates.



Post-Decoder bit error probability ≤ 10^{-4}

ref 1: Wu, G. et al (1990) "A New Hybrid ARQ Scheme for Burst Error Environment" *ICCS '90 Conf. Communication Systems: Towards Global Integration, vol. 2, p.20.6.1-5. Singapore.*

ref 2: Dalmau-Royo, J. and Serrat-Fernandez, J. (1991) "Performance Analysis of a Hybrid ARQ System in Half-Duplex Transmission at 2400 bps" *Fifth Int. Conf. on HF Radio Systems and Techniques, p.248-52, July, U.K.*

FIGURE 5.10. Comparative performance of Adaptive Code Rate Strategy, two fixed FEC codes and two H-ARQ schemes for Rayleigh fading channel, with post-decoder error probability≤10^{-4}.

5.4 Chapter Summary

This chapter has described a method of control of link performance which uniquely optimises link throughput for a user-specified error probability. The technique achieves a guaranteed quality of service with extremely high throughput for a HF link.

Logical extensions to the study include examination of the effects of:

Performance in non-Gaussian noise (eg. impulsive) and interference.
The number of codes used.
Feedback errors on throughput and error performance.
Using a second ARQ "outer loop" to more thoroughly control errors.

Extension of this work to serial-tone modems and to the mobile channel may also be possible.

Chapter 6

Conclusions and Suggestions for Further Work

6.1 Conclusions

The decision variable is that point in any digital receiver just prior to where data decisions are made. Examination of the decision variable uniquely yields information about digital link performance, specifically, the probability of error. This information is amplified by the use of prior knowledge to provide error probability estimation performance hitherto unachieved. Estimation accuracy is now sufficient to facilitate dynamic link feedback control with very high performance.

In this thesis, digital communications link performance was comprehensively reviewed with the conclusion that there was no accurate on-line method for estimating error probability in existence. A concept for error probability estimation was developed from a raw form at the commencement of this work with a "Chi-Squared" test structure for the estimator, through to a deep understanding. This development began with the analysis of the chi-squared test and gauging its deficiencies. From this, numerous improvements were proposed culminating in the maximum likelihood tests with the support of extensive simulations examining a wide variety of channel types and demodulation schemes.

A link feedback control strategy was successfully developed to exploit the power of the estimation technique in such a way as to optimise throughput for a specified error performance.

6.2 Suggestions for Further Work

Extensions of the Density Deviation Monitor (Leon et. al.:1974) to the *prediction* of link failure has not been pursued and could be worthwhile.

Rapid estimation of post-decoder error probability would be of great use to network and higher level layers in a communication system. This was

performed for the case of a linear block code in the adaptive code rate system. Extension to convolutional codes would allow those systems using these codes to be characterised.

The combinations of different modulations and channel types for which estimator performance may be examined are innumerable. Two areas that could be examined further are modulations with memory (eg. CPFSK and MSK) and Maximum Likelihood Sequence Estimators (MLSE's).

Performance in the presence of other channel disturbances including impulsive noise could also be studied.

The examined performance has also been for cases where no additional supporting measurements are taken. Clearly, performance improvements are possible if supplementary measurements are made of relevant parameters (eg. fade rate, number of multipaths etc.), and used to direct the test towards more probable model types instead of searching the entire model space.

The adaptive code rate system is being taken further by the author and Mr. Timothy Giles to try better predictors, and include impulsive noise and interference. This system will be implemented and trialled as part of the DSTO research programme in 1992.

The application of the EVEREST concept to optimising the performance of other digital links (eg. with serial-tone modems) may now be examined, since performance may now be rapidly and accurately determined.

Advances in Digital Signal Processing (DSP) technology enables the implementation of detection schemes in software. This makes access to the decision variable straightforward and facilitates the realisation of the EVEREST.

APPENDIX

APPENDIX 1

Derivation of Optimised Bin Positions for the WLS Test

(B.R. Davis 15/8/90)

Consider the expected value η of the WLS statistic given in equation (3.6). This equation may be expanded:

$$\eta = N \sum_{j=1}^{M} \frac{P_j}{\lambda_j} + N(N-1) \sum_{j=1}^{M} \frac{P_j^2}{\lambda_j} \qquad (A.1)$$

Where: $\lambda_j = P_{jk}$ (k notation is dropped for convenience)

Let, $$\eta = N A + N(N-1) B \qquad (A.2)$$

For large N, we are interested in maximising the change in B for a change in γ, some parameter which characterises the system (as before).

$$B = \sum_{j=1}^{M} \frac{P_j}{\lambda_j} \qquad (A.3)$$

Put, $\lambda_j = \lambda_j + d\lambda_j$

$$B + \delta B = \sum_{j=1}^{M} \frac{P_j^2}{\lambda_j + d\lambda_j} \qquad (A.4)$$

$$= \sum_{j=1}^{M} \frac{P_j^2}{\lambda_j} \frac{1}{1 + \frac{d\lambda_j}{\lambda_j}} \qquad (A.5)$$

Applying a series expansion:

$$= \sum_{j=1}^{M} \frac{P_j^2}{\lambda_j} \left(1 - \frac{d\lambda_j}{\lambda_j} + \left(\frac{d\lambda_j}{\lambda_j} \right)^2 - \ldots \right) \qquad (A.6)$$

If $\lambda_j = P_j$, then B = 1 and (A.6) simplifies to:

$$1 + \delta B = \sum_{j=1}^{M} \lambda_j - 0 + \sum_{j=1}^{M} \frac{d\lambda_j^2}{\lambda_j} \qquad (A.7)$$

$$\delta B = \sum_{j=1}^{M} \frac{d\lambda_j^2}{\lambda_j} = (d\gamma)^2 \sum_{j=1}^{M} \frac{1}{\lambda_j} \left(\frac{d\lambda_j}{d\gamma} \right)^2 \qquad (A.8)$$

Hence we need to maximise:

$$f = \sum_{j=1}^{M} \frac{1}{\lambda_j} \left(\frac{d\lambda_j}{d\gamma} \right)^2 \qquad (A.9)$$

Subject to the constraints:

$$\sum_{j=1}^{M} \lambda_j = 1 \quad \text{or} \quad \sum_{j=1}^{M} \frac{d\lambda_j}{d\gamma} = 0 \qquad (A.10)$$

Using the method of Lagrange multipliers, let:

$$f' = \sum_{j=1}^{M} \left(\frac{1}{\lambda_j} \left(\frac{d\lambda_j}{d\gamma} \right)^2 + \mu \frac{d\lambda_j}{d\gamma} \right) \qquad (A.11)$$

Now, $$\lambda_j = \int_{x_{j-1}}^{x_j} p(x;\gamma)\, dx \qquad (A.12)$$

Where: x_j and x_{j-1} represent the bin boundary points desired.

$$g_j = \frac{d\lambda_j}{d\gamma} = \int_{x_{j-1}}^{x_j} q(x;\gamma)\, dx \qquad (A.13)$$

Where: $$q(x;\gamma) = \frac{\partial}{\partial\gamma} p(x;\gamma) \qquad (A.14)$$

We can now form:

$$\frac{\partial f'}{\partial x_j} = \frac{1}{\lambda_j} 2 g_j \; q(x_j;\gamma) - \frac{g_j^2}{\lambda_j^2} p(x_j;\gamma)$$

$$- \frac{1}{\lambda_{j+1}} 2 g_{j+1} q(x_j;\gamma) + \frac{g_{j+1}^2}{\lambda_{j+1}^2} p(x_j;\gamma)$$

$$= 0 \qquad (A.15)$$

Rearranging:

$$2q(x_j;\gamma) \left[\frac{g_j}{\lambda_j} - \frac{g_{j+1}}{\lambda_{j+1}} \right] = p(x_j;\gamma) \left[\frac{g_j^2}{\lambda_j^2} - \frac{g_{j+1}^2}{\lambda_{j+1}^2} \right] \qquad (A.16)$$

So:

$$\frac{q(x_j;\gamma)}{p(x_j;\gamma)} = \frac{1}{2} \left(\frac{g_j}{\lambda_j} + \frac{g_{j+1}}{\lambda_{j+1}} \right) \qquad (A.17)$$

For the case of a Gaussian distribution for the decision variable:

$$p(x;\gamma) = \frac{1}{\gamma\sqrt{2\pi}} \exp\left(-x^2/2\gamma^2 \right) \qquad (A.18)$$

So, from (A.14):

$$q(x;\gamma) = \frac{1}{\sqrt{2\pi}} \left[-\gamma^{-2} e^{-x^2/2\gamma^2} + \gamma^{-1} e^{-x^2/2\gamma^2} \left(\frac{x^2}{2\gamma^2} \right) \right] \qquad (A.19)$$

Figure A.1 shows the general shape of $p(x;\gamma)$ and $q(x;\gamma)$ with an example of bin boundary positions for M = 3 bins.

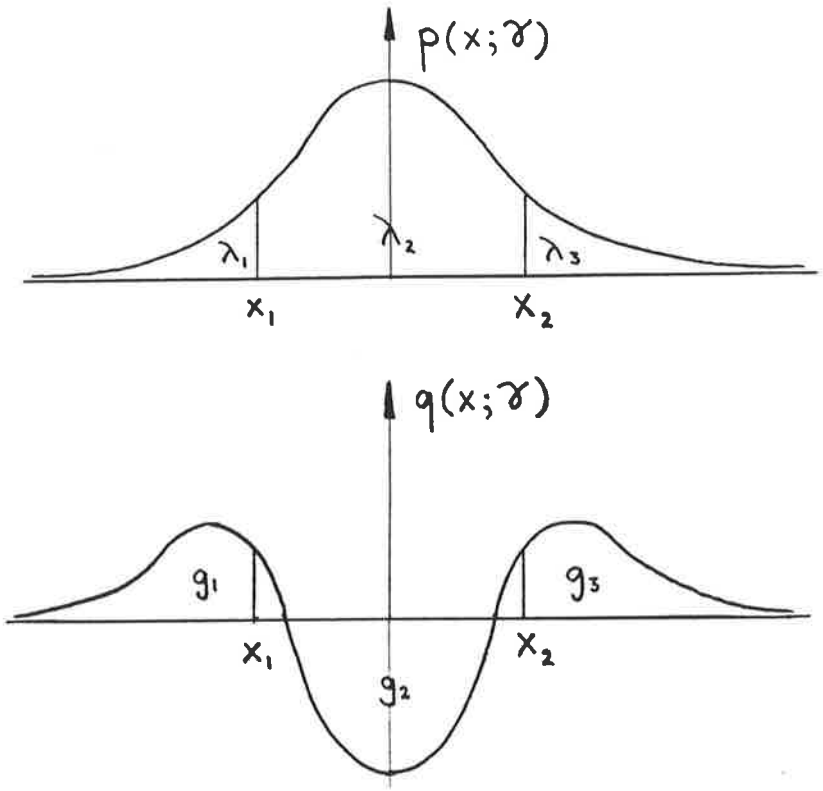


Figure A.1. General form of $p(x;\gamma)$ and $q(x;\gamma)$, with example boundaries for M=3 bins.

The bin boundary positions are x_1 and x_2, which may be determined as follows:

Substitute (A.18) and (A.19) into (A.17):

$$\frac{q(\,x\,;\gamma)}{p(\,x\,;\gamma)} \;=\; -\gamma^{-1} + x^2\,\gamma^{-3} \qquad \text{(A.20)}$$

Then, from (A.13):

$$g \;=\; \int_a^b \left(-\gamma^{-2} + x^2\,\gamma^{-3}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-x^2/2\gamma^2\right) dx \qquad \text{(A.21)}$$

put $t = x/\gamma$

$$g = \gamma^{-1} \int_{a/\gamma}^{b/\gamma} (t^2 - 1) \frac{1}{\sqrt{2\pi}} \exp\left(-t^2/2\right) dt \qquad (A.22)$$

$$= \left[- \frac{t}{\gamma \sqrt{2\pi}} \exp\left(-t^2/2\right) \right]_{a/\gamma}^{b/\gamma}$$

$$g = \frac{1}{\gamma^2 \sqrt{2\pi}} \left(a \exp(-a^2/2\gamma^2) - b \exp(-b^2/2\gamma^2) \right) \qquad (A.23)$$

Put $a = x_2$, $b = \infty$

$$g_3 = \frac{1}{\gamma^2 \sqrt{2\pi}} x_2 \exp\left(-x_2^2/2\gamma^2\right)$$

$$\lambda_3 = Q(x_2/\gamma)$$

Since $x_1 = x_2$ ...

$$g_2 = -2 g_3$$

$$\lambda_2 = 1 - 2\lambda_3$$

Equating (A.17) and (A.20) gives:

$$-\gamma^{-1} + \gamma^{-3} x_2^2 = \frac{1}{2\gamma^2 \sqrt{2\pi}} x_2 \exp\left(-x_2^2/2\gamma^2\right) \left[\frac{1}{Q(x_2/\gamma)} - \frac{2}{1-2Q(x_2/\gamma)} \right]$$

let $r = x_2/\gamma$

$$r^2 - 1 = \frac{1}{2\sqrt{2\pi}} r \exp(-r^2/2) \left[\frac{1}{Q(r)} - \frac{2}{1 - 2Q(r)} \right] \qquad (A.24)$$

Solving the transcendental equation (A.24) for r gives:

$$r = 1.482072$$

So,

$$x_2 = -x_1 = 1.482072\gamma \qquad (A.25)$$

APPENDIX 2

Published Papers

A Scheme for High Performance Real-Time BER Measurement

J.B. Scholz, S.C. Cook, T.C. Giles

Communications Division
Electronics Research Laboratory
Defence Science and Technology Organisation
Australia

ABSTRACT

Bit Error Rate (BER) is a fundamental measure of performance in digital communication systems, since it is a measure of the integrity of received information.

A statistically-based method of BER estimation has been devised which does not necessitate the transmission of test patterns or the interruption of data for measurement purposes. It is demonstrated that the method allows accurate estimation of BER several orders of magnitude faster than the Monte Carlo method. Design of an optimal BER estimator based on the Maximum-Likelihood principle is proposed.

The performance of the technique is presented from simulations for the case of QPSK modulation in AWGN and for a live Rayleigh fading HF link. The ability of the method to accurately predict BER using relatively few samples means that it is particularly suitable for high-speed BER estimation.

Accepted for publication in IEEE Transactions on Communications

I. INTRODUCTION

The development of the BER monitoring technique we have called EVEREST (Extremely Versatile Error Rate ESTimator) was instigated by the need to measure the performance of low-rate terrestrial digital radio links subject to natural and deliberate degradation. Such links are characterized by Rayleigh fading, intersymbol interference, additive and impulsive noise as well as frequency selective fading from multipath components. Noise distributions are generally non-Gaussian and channel statistics are liable to alter drastically in just a few minutes. It is therefore necessary to perform the measurement quickly.

Regardless of channel characteristics, the error rate provides the ultimate measure of performance for a digital communication system. The importance of this measure has been recognised [1] and quite some research effort has been applied to the problems of its measurement. Of the numerous methods proposed [1,2] few are applicable to rapid measurement, and fewer still appear to be applicable to non-Gaussian (eg. fading) channel statistics.

II. THE EVEREST APPROACH

The EVEREST method requires access to the receiver decision statistic (decision variable) or if this is not available, the demodulation process must be emulated to derive it. The decision variable is that signal in a digital demodulator/ detector, just prior to making the data (symbol) decisions. In a demodulator, the decision variable may be represented as Y_i, i=1,2,...,N representing samples taken at time intervals corresponding

to the symbol interval.

We wish to test the hypothesis that a set of decision variable observations correspond to a given prior model A_k' (k=1,2,...,L).

$$H_k : \{ Y_l \} \in A_k \quad (1)$$

We wish to test this hypothesis against each model (k=1,2..,L) and choose the best matching a priori model (in a maximum likelihood sense) fitting these collective observations. The chosen model A_k will have a BER value associated with it, so that the actual estimated BER will be taken as the BER of the chosen model. Due to the fact that the decision variable is the sole signal which constitutes the received data (and perturbations), we claim that its distribution directly relates to error occurrences and the error rate.

Each value of Y_l is recorded by accumulating a count of the number of times Y_l falls within a region R_j. There are M such regions spanning the entire range of Y_l. Thus, at the end of N observations, a histogram has been built, representing the distribution of the decision variable. Figure 1 shows an example of decision regions for a QPSK demodulator and figure 2 illustrates a histogram of counts for the regions shown in figure 1 for AWGN and $E_b/N_0 \cong 3$ dB.

In one symbol period, one count is recorded in only one of the regions R_j (j=1,..M). If the observations Y_l are random variables, then the accumulated histogram counts are also random variables.

Figure 3 shows a block diagram of the system. The statistical test which

selects the library model and hence the BER, is the subject of the following.

If the observations Y_1 are random, stationary and independent then so are the counts u_j for region j. Then, a likelihood function for the set of u_j states may be formed from the multinomial distribution, which is the extension of the binomial distribution from two to M regions [3]:

$$P(u_1, u_2, \ldots, u_M | A_k) = N! \prod_{j=1}^{M} \frac{P_{jk}^{u_j}}{u_j!} \qquad (2)$$

where: P_{jk} = probability of a count in region R_j for model k.

We wish to choose the model A_k for which the a posteriori probability $P(A_k | u_1, u_2, \ldots, u_M)$ is maximised. By Bayes rule this is equivalent to the maximization of ρ_k in (3) since $P(u_1, u_2, \ldots, u_M)$ is not a function of k.

$$\rho_k = P(u_1, u_2, \ldots, u_M | A_M) P(A_k) \qquad (3)$$

Maximizing (3) is the same as maximising its logarithm. Taking the logarithm and substituting equation (2):

$$\ln \rho_k = \ln N! - \sum_{j=1}^{M} \ln u_j! + \sum_{j=1}^{M} u_j \ln (P_{jk}) + \ln P(A_k) \qquad (4)$$

The first two terms of (4) are independent of k, and may be neglected, so the maximization may be reduced to maximizing $\ln \theta_k$ as shown in (5):

$$\theta_k = \sum_{j=1}^{M} u_j \upsilon_{jk} + \lambda_k \qquad (5)$$

where: $\upsilon_{jk} = \ln P_{jk}$

$\lambda_k = \ln P(A_k)$

So, we choose the BER model $A_{\hat{k}}$ where:

$$\theta(\hat{k}) = \max_k \left[\sum_{j=1}^{M} u_j \upsilon_{jk} + \lambda_k \right] \qquad (6)$$

The BER estimate is then the known BER of the library model $A_{\hat{k}}$.

This method of BER estimation has proven to be far superior to estimating the BER from the measured data [4,5]. Equation (6) describes a Maximum A Posteriori (MAP) scheme. If we decide to give equal weighting to priors, this reduces to a Maximum Likelihood (ML) BER estimator (which may be implemented as a bank of correlators or matched filters).

The library should contain an adequate set of models which cover each type of condition to be expected, with a resolution compatible with the required precision. Each model will contain a histogram of the decision variable and the corresponding error rate for that situation. The library models may be compiled from theory, simulation and/or from live measurements on links. The latter method would be ideal, in situations where it is possible to accurately measure the true error rate (say, by sending a known data sequence and counting actual errors offline) and the decision variable distribution over the measurement period. Actual channel measurements would have the advantages of allowing the effects of receiver calibration subtleties and non-linearities to be incorporated into the models. If the channel is time-varying, with a period comparable to the measurement interval, additional models accounting for the

different possible initial channel conditions would need to be incorporated.

The performance of the ML EVEREST has been simulated using an arbitrary choice of equal sized regions for a coherent QPSK modulation scheme in an AWGN channel environment. A small number of equally spaced regions has proven to be effective in practice. Figure 4 shows the performance using four equal sized regions (M=4) and L=400 and L=40 library models covering a range of BER values from 10^{-1} to 10^{-5} in equal steps for 1000 and 16000 symbol samples.

Figure 4 shows the performance using L=400 models in order to obtain a smooth curve. The practical performance using L=40 models demonstrates that the accuracy is not significantly compromised by lowering the precision and a coarser, nonetheless accurate estimation results.

The ML implementation of the previous section used the multinomial distribution to describe the histogram forming process. This assumes that the region probabilities P_j are constant over the period of accumulating N samples. If the channel distribution is non-stationary (eg. Rayleigh fading) over the sampling interval, samples may be correlated. The effect of correlated samples is to increase the variance of sample counts u_j in a region. This is effectively the same as reducing the number of independent samples [6]. This will increase the variance in θ_k and thus the effective sample size of the EVEREST will be reduced.

The EVEREST technique relies on characterising the error rate from an accumulated distribution of the decision variable(s). If an ambiguity occurs whereby similar distributions correspond to a different error rate,

a larger number of samples (N) and/or regions (M) would need to be used to make the models as distinguishable from one another as possible. Increasing M would, of course, imply increased estimator complexity.

The EVEREST has been implemented and trialled over a skywave HF radio link between Darwin and Melbourne, Australia. The results represent about 350 hours of measurements in August 1989. The decision variables were extracted directly from a parallel tone Kineplex® (2400 bps) QPSK MODEM and applied via an A to D converter into an IBM AT-PC. The EVEREST was implemented entirely in the PC, using M=4 regions and N=3750 samples for each of the sixteen parallel tones. A pseudo-random sequence was transmitted over the same MODEM in order to allow the error rate to be measured. The EVEREST measurement time was 50 seconds, in this time a total of $16*3750=6*10^4$ symbols were sent, allowing the error rate to be calculated with reasonable accuracy to as low as 10^{-4}. The EVEREST error rate was taken as the average of the sixteen error rates estimated for each tone. Figure 5 shows the results. The library contained a total of L=300 models, made up of 250 flat Rayleigh fading models for fade rates of 0.2, 0.4, 0.6, 0.8 and 1.0 Hz and 50 AWGN models covering the range of error rates from 0.5 to 10^{-6}. The accuracy of error rate estimates as low as 10^{-6} could not be confirmed without extending the error counting measurement time to around 80 minutes, so the results to 10^{-4} are shown.

III. CONCLUSIONS

A method of BER estimation has been proposed which is suitable for applications requiring rapid measurement and where it is undesirable or infeasible to send a known data stream and count actual bit errors. A maximum likelihood implementation of the method has been simulated for the case of QPSK modulation, but is equally suited to any modulation scheme.

The simulated performance of the method shows a BER estimation accuracy of within half an order of magnitude for 95% of measurements for BER values as low as 10^{-5} using only 1000 symbol samples in an AWGN environment. A sample size of 16000 symbols reduces the estimation error to about 0.1 in log error rate.

Live link results on a HF link using Rayleigh and AWGN channel models showed that 95% of all measurements were within one order of magnitude to as low as BER=10^{-4}, using only 3750 samples.

The method outlined achieves a previously unattainable standard of speed and accuracy in BER measurement.

REFERENCES

[1] E.A. Newcombe, S. Pasupathy, "Error Rate Monitoring for Digital Communications". *Proceedings of the IEEE*, vol. 70, No. 8, pp 805-828, Aug. 1982.

[2] J.B. Scholz, "Error Performance Monitoring of Digital Communication Systems". *Aust. Telecommun. Research Journal*, vol. 25, No. 2, 1991.

[3] .R.E. Walpole and R.H. Myers, "Probability and Statistics for Engineers and Scientists". *Second Ed., MacMillan Publ.*, pp 89-90, 1978.

[4] D.J. Gooding, "Performance Monitor Techniques for Digital Receivers Based on Extrapolation of Error Rate". *IEEE Trans. on Commun. Tech.*, vol. COM-16, No. 3, pp. 380-387, Jun. 1968.

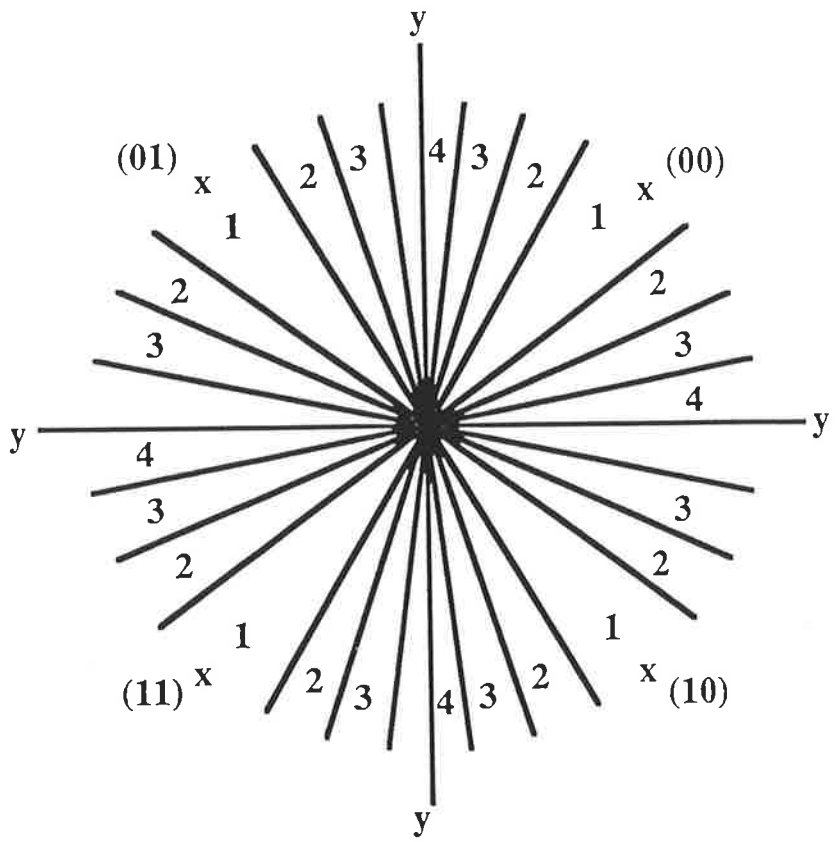
[5] S.B. Weinstein, "Estimation of Small Probabilities by Linearization of the Tail of a Probability Distribution Function". *IEEE Trans. on Commun. Tech.*, vol. COM-19, No. 6, pp. 1149-1155, Dec. 1971.

[6] M.C. Jeruchim, "Techniques for Estimating the Bit Error Rate in the Simulation of Digital Communication Systems". *IEEE Journ. on Selected Areas in Comms.*, vol. SAC-2, No. 1, pp. 153-170, Jan. 1984.

FIGURE LIST

A SCHEME FOR HIGH PERFORMANCE REAL-TIME BER MEASUREMENT

J.B. Scholz, S.C. Cook and T.C. Giles



x = Expected phasor positions

y = Actual decision thresholds

Figure1

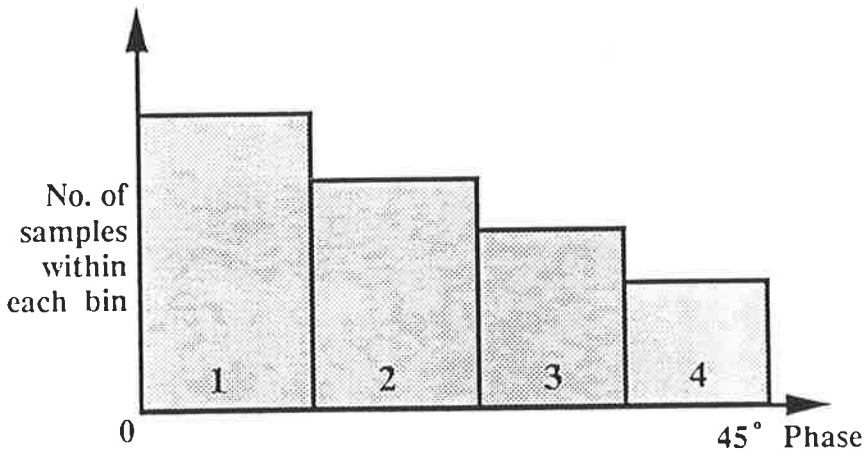


Figure 2

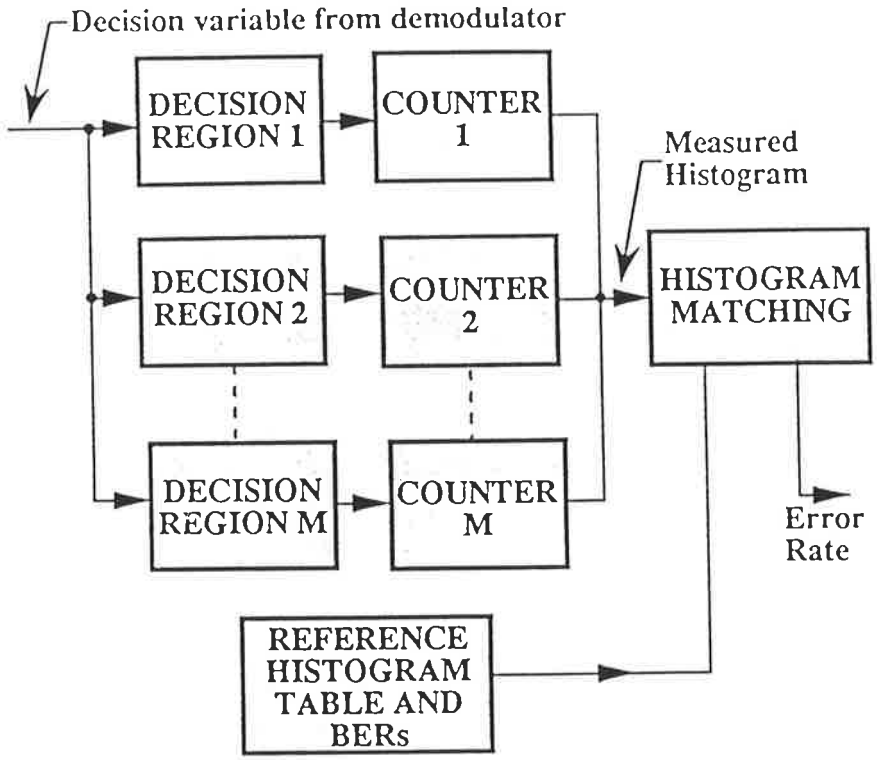

Decision variable from demodulator
DECISION REGION 1
COUNTER 1
Measured Histogram
DECISION REGION 2
COUNTER 2
HISTOGRAM MATCHING
DECISION REGION M
COUNTER M
Error Rate
REFERENCE HISTOGRAM TABLE AND BERs

Figure 3

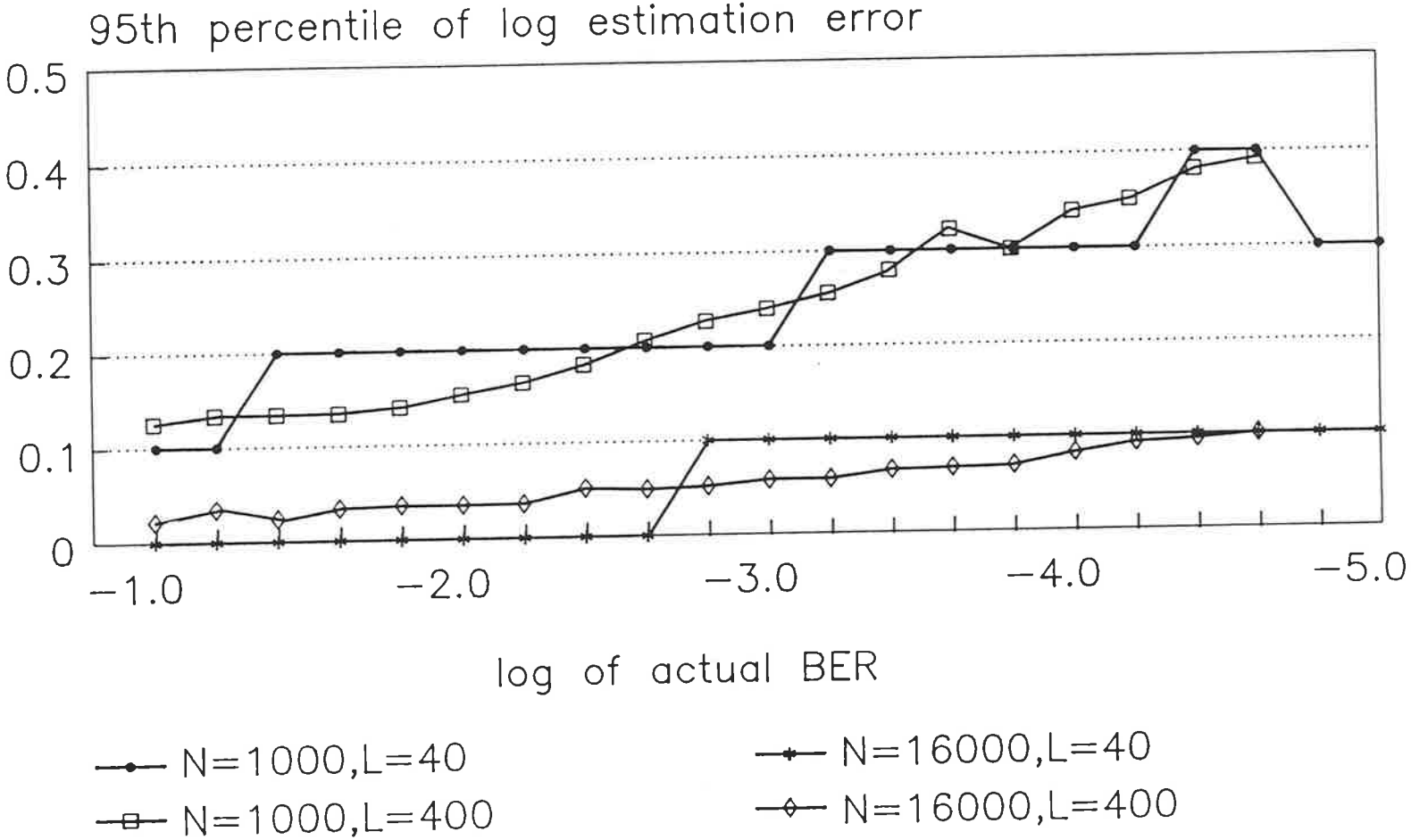

95th percentile of log estimation error
0.5
0.4
0.3
0.2
0.1
0
−1.0
−2.0
−3.0
−4.0
−5.0
log of actual BER
N=1000,L=40
N=1000,L=400
N=16000,L=40
N=16000,L=400
M=4 regions

Figure 4

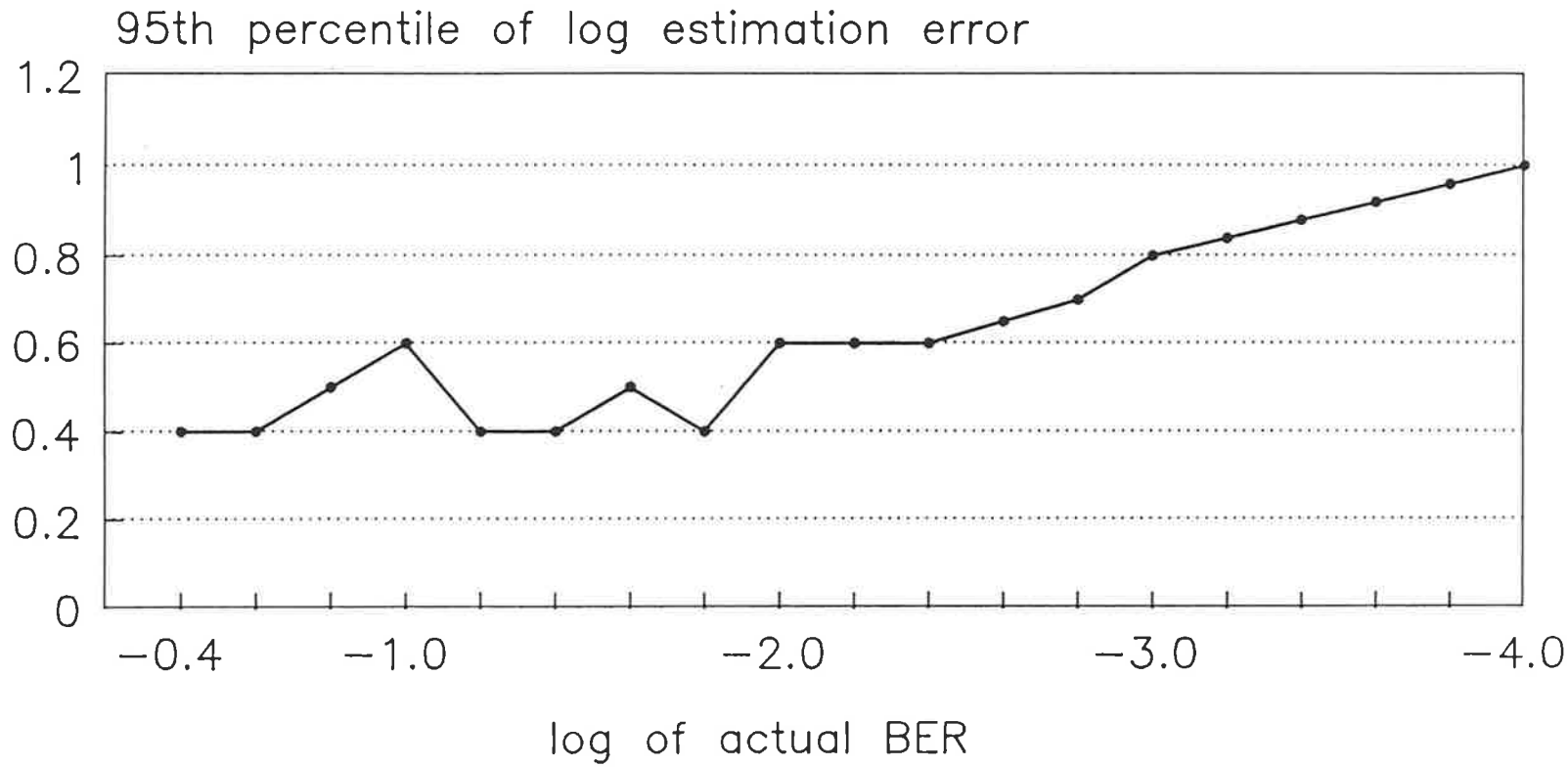

95th percentile of log estimation error
1.2
1
0.8
0.6
0.4
0.2
0
−0.4
−1.0
−2.0
−3.0
−4.0
log of actual BER
N=3750
M=4 regions, L=300 models

Figure 5

J.B. Scholz (1991) Trial results of novel bit error rate monitor scheme over skywave HF link.
Electronics Letters, v. 27 (19), pp. 1730–1732, 12 September 1991

J.B. Scholz (1991). Error Performance Monitoring of Digital Communications Systems.
Australian Telecommunication Research, v. 25 (2), pp. 1–25, 1991

J.B. Scholz. T.C. Giles, A Revolutionary High Performance Passive Real-Time Link Evaluator for Fast Adaptive Communications
The Institution of Engineers Australia Communications Conference, Melbourne 16-18 October 199O

J.B. Scholz. (1992) A digital communications link error probability estimation scheme and the effect of correlation on its performance
MILCOM '92 Conference

J.B. Scholz. (1990) Dynamically adaptive digital communication systems for improving throughput and error rates
The Institution of Engineers Australia Communications Conference, Melbourne 16 - 18 October 1990

Bibliography

American National Standards Institute (ANSI) (1974). "Determination of the Performance of Data Communication Systems" *ANSI Standard x3.44-1974.*

Argenzia, A.E. and Campbell, J.J. (1977). "Digital Automated Technical Control" *in Nat. Telecomm. Conf. Rec., pp. 43.1.1-43.1.7.*

Australian Telecommunications Commission (1988). "Digital Error Logger" *Australian Patent Spec., AU-A-21882/88, filed 5th Sept.*

Becam, D., Codet, A., Ruhault, F. (1987) "Error Detection in Operating Digital Systems: Influence of Error Distribution" *IEE Proc., Vol. 134, Pt. F, No. 5, Aug.*

Betts, J.A., Ellington, R.P. and Jones, D.R.L. (1970). "CW Sounding and its Use for Control of HF (3-30MHz) Adaptive Systems for Data Transmission" *Proc. IEE, Vol. 117, No. 12, pp. 108-111.*

Betts, J.A., Broom, R.S., Cook, S.J. and Clark, J.G. (1975). "Use of Pilot Tones for Real-Time Channel Estimation of HF Data Circuits" *Proc. IEE, Vol. 122, No. 9, Sept., pp. 887-896.*

Beyer, W.H. (1988). "Handbook of Mathematical Sciences" 6th Ed. CRC Press.

Blahut, R.E. (1984). "Theory and Practice of Error Control Codes" *Addison- Wesley Publ.*

Bradley, J. (1960). "Distribution Free Statistical Tests" *Behavioral Sc. Lab., Aerospace Medical Div., Wright Air Dev. Center, WADD Tech. Rep. 60-661, Aug.*

Brillant, M.B. (1978) "Observations of Errors and Error Rates on T1 Digital Repeated Lines" *Bell Syst. Tech. Journ., Vol. 53, pp. 711-746.*

Brown, R.O. and Wilson, G.G. (1974). "Automatic Data Quality Monitoring" *IEEE Canadian Commun. & Power Conf., pp. 83-84.*

Brown, R.O. (1976). "Performance of Pseudo Error Estimates" *IEEE Canadian Comms. Conf., pp. 109-112.*

Burington and May (1958) "Handbook of Probability and Statistics with Tables" *Sadunsky, Ohio: Handbook Publ. Inc., p. 186 ff.*

Burton H.H. and Sullivan D. (1972). "Errors and Error Control" *IEE Proc. Vol. 60, No. 11, Nov., pp. 1293-1301.*

Carlson, A. B. (1975). "Communication Systems" *Mc Graw-Hill, Second Ed.*

Catchpole, R.J. (1975). "Fault Location by Error Detection" *Instit. Elec. Eng. Conf. Telecomm. Trans., pp. 99-102.*

Cavers, J.K. (1972). "Variable-Rate Transmission for Rayleigh Fading Channels" *IEEE Trans. on Comms., Vol. COM-20, No. 1, Feb., pp. 15-22.*

CCIR (1982). Recommendation 520-1. "Use of High Frequency Ionospheric Channel Simulators" *Annex 1.*

CCIR (1986). Report 332-3. "Characteristics and Applications of Atmospheric Radio Noise Data".

CCITT (1984). Recommendation G702, *Red Book, Definition no. 2013.*

CCITT (1988a) "Digital Networks, Digital Sections and Digital Line Systems" *CCITT Recommendation G.821 (Blue Book, III.5), Nov.*

CCITT (1988b). "Specifications for Measuring Equipment" *CCITT Recommendation O.151, O.152, O.161, (Blue Book, IV.4), Nov.*

Cherin, A.H. (1983). "An Introduction to Optical Fibres" *McGraw-Hill Publ.*

Cock, C.C., Edwards, A.K. and Jessop, A. (1975). "Timing Jitter In Digital Line Systems" *IEE Conf. on Telecomm. Trans., 9-11 Sept., pp. 33-36.*

Cock, C.C. (1987). "Assessment of Timing Jitter in Digital Telecommunications Transmission Systems" *IEE Proc.,Vol. 134, Pt. F, No. 5, Aug., pp. 464-473.*

Cook, S.C., Scholz, J.B and Giles, T.C. (1990). "Dynamically Adaptive Digital Communication Systems for Improving Throughput and Error Rates", *Proc. Communications '90 Conf., IE Aust., Melbourne, Oct., pp. 192-196.* (see Appendix 2).

Crawford, T.M., Reynolds, A.S., Young, I.R. (1983). "Noise Margin Measurement and Error Probability Prediction" *U.S. Patent, 4 384 354, May 17.*

Dalmau-Royo, J. and Serrat-Fernandez, J. (1991). "Performance Analysis of a Hybrid ARQ System in Half-Duplex Transmission at 2400bps" *Fifth Int. Conf. on HF Radio Systems and Techniques, U.K., July, pp. 248-252.*

Davies, P.A. and Sherif, M. (1986). "Use of Pseudo-Error Monitoring Schemes in Optical Communications" *IEE Proc., Vol. 133, Pt. J, No. 6, Dec, pp. 371-376.*

Davis, B. R. (1988) "The Effect of Correlation on Probability of Error Estimates in Monte Carlo Simulations" *Journ. of Electrical & Electronic Eng., Aust., Vol. 8, No. 4, Dec., pp. 222-229.*

Feher, K. (1977). "Pseudoerror On-Line Monitoring: Concept Design and Evaluation" *Canadian Elec. Eng. Journ., Vol. 2, No. 2, pp. 33-36.*

Feldman, P.M. and Li, V.O.K. (1988). "Adaptive Coding for Discrete-Time Markovian Channels" *Proc. of the Third IFIP Int. Conf. on Data Comm. Systems and Their Performance., Rio De Janeiro, Brazil, 22-25 June, pp. 139-153.*

Fisher, R.A. (1958). "Statistical Methods for Research Workers" *13th Ed., New York, Hafner.*

Fritchman, B.D. (1967) "A Binary Channel Characterization Using Partitioned Markov Chains" *IEEE Trans. on Inform. Theory, vol. IT-13, no. 2, April.*

Gagliardi, R.M. and Thomas, M.C. (1968). "PCM Data Reliability Monitoring Through Estimation of Signal-to-Noise Ratio" *IEEE Trans. Com. Tech., Vol. COM-16, No. 3, June.*

Gooding, D.J. (1968). "Performance Monitor Techniques for Digital Receivers Based on Extrapolation of Error Rate" *IEEE Trans. on Comms., Vol. COM-16, No. 3, June, pp. 380-387.*

Groginsky, H.L. (1966). "Adaptive Detection of Statistical Signals in Noise" *IEEE Trans. on Info. Theory, Vol. IT-12, No. 3, July, pp. 337-348.*

Guida, M., Domenico, I. and Longo, M. (1988). "Comparative Performance Analysis of Some Extrapolative Estimators of Probability Tails" *IEEE Journ. on Sel. Areas in Comms., Vol. 6, No. 1, Jan., pp. 76-84.*

Gumbel, E.J. (1966). "Statistics of Extremes" *Columbia Uni. Press, New York, 3rd Ed., 1966.*

Gunn, J.F. and Lombardi, J.A. (1969). "Error Detection for Partial-Response Systems" *IEEE Trans. on Comm. Tech., Vol. COM-17, No. 6, Dec.*

Hagenauer, J. (1988). "Rate-Compatible Punctured Convolutional Codes (RCPC Codes) and Their Applications" *IEEE Trans. on Commun., Vol. 36, No. 4, April, pp. 389-400.*

Hall, R.D. and Lancaster, H.A. (1975). "Prediction and Measurement of Timing Jitter on Digital Line Links" *IEE Conf. on Telecomm. Transmission, 9-11 Sept., pp. 41-44.*

Hammond, J.L., Vena, P.A., Sears, W.E. (1973). "Extrapolation Monitors for Digital PSK Links" *Rec. of the 1973 Nat. Telecom. Conf., Vol. 1, pp. 4C-1/8, Nov.*

Hancock, J.C. (1963). "Optimum Performance of Self-Adaptive Systems Operating Through A Rayleigh Fading Medium" *IEEE Trans. on Commun. Sys., Vol. CS-11, Dec., pp. 443-453.*

Hayes, J.F. (1968). "Adaptive Feedback Communications" *IEEE Trans. on Comm. Tech., Vol. COM-16, No. 1, Feb., pp. 29-34.*

Hingorani, G.D. and Chesler, D.A. (1968a) "A Performance Monitoring Scheme for FSK Transmission Over Fading Channels" *IEEE Trans. Comm. Tech., Vol. COM-16, No. 3, June, pp. 430-435.*

Hingorani, G.D. and Chesler, D.A. (1968b) "A Performance Monitoring Technique for Arbitrary Noise Statistics" *IEEE Trans. Comm. Tech., Vol. COM-16, No. 3, June, pp. 430-435.*

Goodman, R.M.F. and Farrell, P.G. (1975). "Data Transmission with Variable-Redundancy Coding Over a High Frequency Channel", *Proc. IEE, Vol. 122, No. 2.*

Hogge, C.R. (1977). "Performance Monitoring of a Digital Radio by Pseudo-Error Detection" *Proc. of the Nat. Telecomm. Conf., pp. 43:3-1/3.*

Hoyer, W. (1985). "Jitter Measurements at Bitrates Up to 145Mbits/s" *IEE Conf. Publ. 256. (MTTS '85), Nov., pp. 166-170.*

Izzo, L., Panico, L. and Paura, L. (1982). "Character Error Probabilities of M-ary Non-Coherent Systems Due to Additive Combinations of Gaussian and Impulsive Noise" *Alta Frequenza, Vol. LI-N.4, July-Aug., pp. 184-191.*

Jankauskas, L. (1976). "Adaptive Estimation of Discrete Nonlinear Channels for Performance Assessment" *Presented at IEEE Canadian Conf. on Comms. and Power, Oct.*

Jankauskas, L.E. Landesberg, M.M. Spector, D. and Meyer, C. (1980). "Technical Control of the Digital Transmission Facilities of the Defense Communication System" *IEEE Trans Comm., vol. COM-28, no. 9, Sept.*

Jayasinghe, S. (1989). "Techniques of Detection, Estimation and Coding for Fading Channels" Ph.D. Dissertation, Loughborough Uni. of Technol., U.K.

Jeruchim, M.C. (1984). "Techniques for Estimating the Bit Error Rate in the Simulation of Digital Communications Systems" *IEEE Journ. Sel. Areas in Comms., vol. SAC-2, no. 1, Jan.*

Katakol, B.S. and Maskara, S.L. (1987). "Performance of Modified-Fano Algorithm in Adaptive-Variable-Rate Transmission System Over Fading Channels" *Int. Symp. on Electronic Devices Circuits and Syst., Dec. 16-18, Kharagpur.*

Kawashima, M. (1987). "Measurement Method for Eye Opening Estimation on Digital Signals" *IEE Proc., Vol. 134, Pt. F, No. 5, Aug.*

Keelty, J.M., and Feher, K. (1978). "On-Line Pseudo Error Monitors for Digital Transmission Systems" *IEEE Trans. on Comms., Vol. COM-26, No. 8, Aug., pp. 1275-1282.*

Kim, J. (1989) "Bit By Bit Error Rate Performance of Land Mobile Satellite System (LMSS) Under Fading Channel Using QPSK/MSK" *IEEE Conf. MILCOM, pp. 715-719.*

Kostic, I.M. (1989). "Pseudo Error Rate of a PSK System with Hardware Imperfections, Noise and Cochannel Interference" *IEE Proc., Vol. 136, Pt. I, No. 5, Oct., pp. 333-338.*

Kreyzig, E. (1983). "Advanced Engineering Mathematics" *5th Ed., John Wiley & Sons Publ., p. 929.*

Leithold, L. (1981). "The Calculus - With Analytic Geometry" *4th Ed., Harper & Row Publ., N.Y., pp. 1046.*

Leon, B.J. and Kitahara, R.T. (1973). "Monitor Simulation of the NEC TRP QPSK Receiver" *Nat. Telecommun. Cof. Rec., pp 4D-1/8.*

Leon, B.J., Hammond, J.L., Bode, R.B. and Sears, W.E. (1974). "Performance Monitors for Digital Communications Systems, Part II" *Rome Air Development Center, Griffiss Air Force Base, New York, Rep. RDC-TR-74-318, Dec.*

Leon, B.J., Hammond, J.L., Vena, P.A., Sears, W.E. and Kitahara, R.T. (1975). "A Bit Error Monitor for Digital PSK Links" *IEEE Trans. on Comms., Vol. COM-23, No. 5, May, pp. 518-525.*

Lender, A. (1964). "Correlative Digital Communication Techniques" *IEEE Trans. Comm. Tech., Vol. COM-12, pp. 128-135, Dec. 1964.*

Lender, A. (1968). "Correlative Data Transmission with Coherent Recovery Using Absolute Reference" *IEEE Trans. Comm. Tech., Vol. COM-16, No. 1, Feb., pp. 108-115.*

Lin, S. and Costello, D.J. (1983) "Error Control Coding: Fundamentals and Applications" *Prentice-Hall Publ., New Jersey.*

Maslin, N. (1987). "HF Communications - A Systems Approach" *Pitman Publ.*

McLarnon, B.D. (1977). "A Survey of Channel Evaluation Techniques for the Integrated Remote Communications System" *CRC Research Labs., Tech. Memo. No. CSR&D-94-77, Nov., pp. 33.*

McManamon, P. (1973). "Results of an Experiment in Channel Quality Monitoring" *IEEE Trans. on Comms., vol. COM-21, no. 9, Sept., pp. 1052-1054.*

McManamon, P.M., Rosich, R.K., Payne, J.A. and Miles M.J. (1975). "Performance Criteria for Digital Data Networks" *Report COM-75-10779, U.S. Dept. of Commerce, Office of Telecommunications, Jan.*

Miller, M.J. and Ahmed, S.V. (1987). "Digital Transmission Systems and Networks" *Vol. 1, Computer Science Press Inc.*

Milstein, L.B. (1976a). "On the Application of Extreme-Value Theory to Bit Error Rate Monitoring" *Int. Commun. Conf. Rec., pp. 44.7-44.11.*

Milstein, L.B. (1976b). "Performance Monitoring of Digital Communication Systems Using Extreme-Value Theory" *IEEE Trans. Comms., Vol. COM-29, No. 9, Sept., pp. 1032-1036.*

Moore, R.T. (1971). "Determination of Performance of Digital Data Communication Systems" *Int. Commun. Conf. Rec., pp. 12.1-12.3.*

Munday, S. (1975) "Error Rate Objectives for a Multi-Service Digital Network" *IEE Conf. on Telecommun. Transmission, 9-11 Sept., pp. 15-21.*

Nahdi, N.E. and Gagliardi, R.M. (1967). "Use of Limiters for Estimating Signal to Noise Ratio" *IEEE Trans. Info. Theory, Vol. IT-13, pp. 127-129.*

Newcombe, A.E. and Pasupathy, S. (1980). "Error Rate Monitoring in a Partial Response System" *IEEE Trans. on Comms., Vol. COM-28, No. 7, July, pp. 1052-1061.*

Newcombe, E.A. and Pasupathy, S. (1982). "Error Rate Monitoring For Digital Communications" *IEEE Proc., vol.70, no. 8, Aug.*

O'Reilly and Rampaigul (1987). "Circuit Structures for High-Digit- Rate Bit Error Ratio Measurements" *IEE Proc., vol. 134, Pt. F, no. 5, Aug.*

Papoulis, A. (1984). "Probability, Random Variables, and Stochastic Processes" *Mc Graw-Hill Publ., Second Ed.*

Pearson, K. (1894). "Contributions to the Mathematical Theory of Evolution" *PTRS, 185, p. 71.*

Pierce, J.N. (1958). "Theoretical Diversity Improvement in Frequency-Shift Keying" *Proc. IRE, Vol. 46, May, pp. 903-910.*

Popovici, A.C. (1987). "Fast Measurement of Bit Error Rate in Digital Links" *IEE Proc., Vol. 134, Pt. F, No. 5, Aug., pp. 439-447.*

Portny, S.E. (1966) "Large Sample Confidence Limits for Binary Error Probabilities" *IEEE Proc., vol. 54, pp. 1993, Dec.*

Proakis, J.G. (1989). "Digital Communications" *McGraw-Hill Publ., Second Ed.*

Radcliffe, J. (1991). "Fibre Optic Link Performance in the Presence of Reflection-Induced Intensity Noise" *Proc. of the SPIE, Vol. 1366, pp. 361-367.*

Rahman, M. and Bulmer, M. (1990) "Error Models For Land Mobile Satellite Channels" *International Mobile Satellite Comms. Systems Conf., Adelaide, Aust., Aug.*

Richters, J.S. and Dvorak, C.A. (1988). "A Framework for Defining the Quality of Communications Services" *IEEE Communications Magazine, Vol. 26, Oct., pp. 17-23.*

Rubtsov, V.D. (1977). "Statistical Characteristics of Signal and Atmospheric Noise Mixture in the HF Band" *translated in Radio Eng. and Electronic Physics, Vol. 22, No. 4, April, pp. 125 - 128.*

Rush, H.F. (1965). "Signal and Channel Parameter Determination Using Decision Threshold Devices" *First IEEE Ann. Comms. Convention, June, pp. 745-748.*

Scholz, J.B. (1990). "A Real-Time Knowledge-Based System for Frequency Management in Communications" *in Lecture Notes in Artificial Intelligence, No. 406, Springer-Verlag Publ.*

Scholz, J.B., Giles, T.C. (1990). "A Revolutionary High Performance Passive Real-Time Link Evaluator for Fast Adaptive Communications" *IE Aust. Communications Conf., Oct., pp. 186-191.* (see Appendix 2).

Scholz, J.B, Cook, S.C. and Giles, T.C. (1990). "Error Rate Monitor" *PCT/AU90/00581, Int. Patent Application, publ. under the Patent Cooperation Treaty, Dec.*

Scholz, J.B. (1991). "Trial Results of a Novel Error Rate Monitor Scheme Over a Skywave HF Link" *IEE Electronics Letters, Vol. 27, No. 19, Sept., pp. 1730-32.* (see Appendix 2).

Scholz, J.B. (1991). "Error Performance Monitoring of Digital Communication Systems" *Aust. Telecomm. Research (ATR) Journal. Vol. 25, No. 2, Nov.* (see Appendix 2).

Scholz, J.B. and Cook, S.C. (1991). "Digital Communication Link Performance Optimisation by Transmission Adaptation" *Provisional Patent Application, submitted Dec.*

Scholz, J.B. (1992). "The Effect of Correlation on the Accuracy of a Digital Communication Link Error Probability Estimation Scheme" *Accepted for MILCOM '92 Conference.* (see Appendix 2).

Scholz, J.B and Giles, T.C. (1992). "Performance Optimisation of Fading Digital Radio Links Through High-Speed Feedback Control" *Submitted to the Int. Symp. on Signals, Systems and Electronics, (ISSSE), Sept., Paris.*

Scholz, J.B. (1992). "Design of a High-Speed, On-Line Error Probability Estimator for Binary Signalling Schemes" *Submitted to Int. Symp. on Signals, Systems and Electronics, (ISSSE), Paris.*

Scholz, J.B., Cook, S.C., Giles, T.C. (1992). "A Scheme for High-Performance Real-Time BER Measurement" *Accepted for publication in IEEE Trans. on Comms.* (see Appendix 2).

Scholz, J.B. (1992). "Performance of a Real-Time Error Probability Estimator for Binary Modulations in the Presence of Noise, Fading and Intereference" *Submitted to IEE Proceedings Part I in 1991. Currently under review.*

Schwartz, M. and Richman, S.H. (1968). "Extremal Statistics in Computer Simulation of Digital Communication Systems" *Proc. Spring Joint Computer Conf., pp. 483-489.*

Serrat-Fernandez, J., Delgado-Penin, J.A., Munday, E., and Farrell, P.G. (1985). "Measurement and Verification of an HF Channel Model" *Third Int. Conf. on HF Communication Systems and Techniques, IEE, London, U.K., pp. 52-56.*

Shannon, C.E. and Weaver, W. (1949). "The Mathematical Theory of Communication" *Uni. of Illinois Press.*

Skinner, T.A. and Cavers, J.K. (1973). "Selective Diversity for Rayleigh Fading Channels with a Feedback Link" *IEEE Trans. on Comms., Vol. COM-21, No. 2, Feb., pp. 117-126.*

Smith, D.K. and Thomas, G.D. (1989). "Maximizing Throughput Under Changing Channel Conditions" *Signal, June, pp. 173-178.*

Smith, D.R. (1973). "Performance Assessment of Digital Transmission Systems" *Rec. of 1973 Nat. Telecomms. Conf., Nov., pp. 4F-1/5.*

Snyder, J.S. and Hersey, W.J. (1984). "Pseudo-bit-error-rate Measurement for 120-Mbits/s TDMA" *COMSAT Technical Review, Vol. 14, No. 2, pp. 285-311.*

Stallings, W. (1988). "Handbook of Computer Communications Standards, Volume 1" *MacMillan Publ., First Ed..*

Stremler, F.G. (1982). "Introduction to Communication Systems" *Second Ed., Addison-Wesley Publ.*

Sunkenberg, H.A. and Jankauskas, L.E. (1976). "Performance Assessment of High- Speed Digital Transmission Systems" *National Telecomm. Conf., pp. 51.5/1-5.*

Sunkenberg, H.A. (1978). "Pseudo Error Rate Measurements and Statistics" *IEEE Canadian Comms. & Power Conf., pp. 54-58.*

Takenaka, S., Katoh, T., Kurihara, H., Fukui, M. and Nakamura, H. (1980). "Bit Error Rate Monitor for Four Phase PSK System" *IEEE International Conf. on Comms., June, pp2.5/1-6.*

Tallis, G.M. (1983). "Goodness of Fit" *Encyclopaedia of Statistical Sciences, Volumme 3, John Wiley & Sons Publ., pp. 451-461.*

Tesla, D.D. (1988). "A Method of Real-Time Channel Evaluation in Adaptive HF Systems" *4th Conf. HF Systems and Techniques., London, pp. 38-42.*

Tesla, D.D. (1989). "A Method of Quick Channel Ranking in Adaptive HF Systems" *Mediterranean Electrotechnical Conf. (MELECON '89), Lisbon, Portugal, 11-13 April, pp. 374-377.*

Tsai, S. (1969) "Markov Characterization of the HF Channel" *IEEE Trans. on Comm. Tech., vol. COM-17, no. 1, Feb.*

Urien, M. and Rault, M. (1982) "Errors and Jitter Performances on a Digital Network" *Proc. of International Conf. on Comms., Vol. 2, pp. 2D4.1-5.*

U.S. Army (1979). "Standard Engineering Installation Package - Wideband Secure Voice Network Extension Project", *U.S. Army Communications Command, Fort Huachuca, AZ, SEIP 035, 18th May, pp. 7/8-23.*

U.S. Federal Standard (1979) INT-FED-STD-001033, *Aug 29.*

Van Trees, H.L. (1968). "Detection, Estimation, and Modulation Theory - Part I" *John Wiley & Sons Publ., NY.*

Viterbi, A.J. (1967). "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm" *IEEE Trans. Inform. Theory, Vol. IT-13, April, pp. 260-269.*

Vogler, L. and Hoffmeyer, J. (1990). "A New Approach to HF Modelling and Simulation - Part II: Stochastic Model" *NTIA Report, 90-255.*

Walpole, R.E. and Myers, R.H. (1975). "Probability and Statistics for Engineers and Scientists" *Second Ed., Macmillan Publ., p. 305.*

Wataya, H. (1989) "A Study of Bit Error Model on Digital Subscriber Transmission" *Electronics & Comms. in Japan, Part 1, Vol. 72, No. 9.*

Watterson, C.C., Juroshek, J.R., Bensema, W.D. (1970). "Experimental Confirmation of an HF Channel Model" *IEEE Trans. on Com. Tech., Vol. COM-18, No. 6, Dec., pp. 792-803.*

Weinstein, S.B. (1971). "Estimation of Small Probabilities by Linearization of the Tail of a Probability Distribution Function" *IEEE Trans. on Comm. Tech., Vol. COM-19, No. 6, Dec., pp. 1149-1155.*

Westcott, R.J. (1972). "Testing Digital Data Transmission Systems" *British Patent Spec. 1 281 390.*

Wolf, J.K., Michelson, A.M. and Levesque, A.H. (1982). "On the Probability of Undetected Error for Linear Block Codes" *IEEE Trans. on Comms., Vol. COM-30, No. 2, Feb., pp. 317-324.*

Wu, G., Mukumoto, K., Fukuda, A., Bao, Z. (1990). "A New Hybrid ARQ Scheme for Burst Error Environment" *ICCS '90 Conf. Communication Systems: Towards Global Integration, Vol. 2, pp. 20.6.1-5, Singapore.*

Glossary of Terms and Abbreviations

ABB	Adaptive Bin Boundary
ACE	Adaptive Channel Estimator
AGC	Automatic Gain Control
ALC	Automatic Link Control
AMI	Alternate Mark Inversion
ARQ	Automatic Repeat reQuest
AWGN	Additive White Gaussian Noise
BBER	Bit by Bit Error Rate
BER	Bit Error Rate
BFSK	Binary Frequency Shift Keying
CCIR	Commité Consultatif Internationale Radio
CCITT	Commité Consultatif Internationale Télégraphique et Teléphonique
CIR	Carrier to Interference Ratio
CPFSK	Continuous-Phase Frequency Shift Keying
DC	Direct Current
DDE	Density Deviation Estimate
DQPSK	Differential Quadrature Phase Shift Keying
DSP	Digital Signal Processing
DSTO	Defence Science and Technology Organisation
ER	Error Rate
EVEREST	Extremely Versatile Error Rate ESTimator
EVT	Extreme Value Theory
FBB	Fixed Bin Boundary
FEC	Forward Error Correction
FSK	Frequency Shift Keying
GEVT	Generalised Extreme Value Theory
H-ARQ	Hybrid Automatic Repeat reQuest
HDB	High Density Bipolar
HF	High Frequency
IF	Intermediate Frequency
INR	Interference to Noise Ratio
ISR	Interference to Signal Ratio
LPI	Low Probability of Intercept
MAP	Maximum A posteriori Probability
ML	Maximum Likelihood
MLSE	Maximum Likelihood Sequence Estimator
MSB	Most Significant Bit
MSE	Mean Square Error
MSK	Minimum Shift Keying
PC	Personal Computer
PCM	Pulse Coded Modulation
PDF (pdf)	Probability Density Function
PMU	Performance Monitor Unit
PSK	Phase Shift Keying
PT-PSK	Pilot Tone Phase Shift Keying
QPSK	Quadrature Phase Shift Keying
RCPC	Rate Compatible Punctured Convolutional
SHF	Super High Frequency
SNR	Signal to Noise Ratio
TDM	Time Division Multiplex
TDMA	Time Division Multiple Access
US	United States
WLS	Weighted Least Squares