THE STUDY OF SPACE COMMUNICATIONS SPREAD SPECTRUM SYSTEMS

(PHASE IV)

REPORT NO. 87-1

INTERIM PROGRESS REPORT

PREPARED FOR

THE DEPARTMENT OF COMMUNICATIONS

UNDER DSS CONTRACT NO. 36001-6-3530/01-ST

P. H. WITTKE

P. J. McLANE

S. J. SIMMONS

M. G. WEARRING

Y. M. LAM

W. HOPKINS

DEPARTMENT OF ELECTRICAL ENGINEERING

QUEEN'S UNIVERSITY

KINGSTON, ONTARIO, CANADA.

Department of Electrical Engineering

Queen's University at Kingston

Kingston, Ontario, Canada

JUNE 1987

THE STUDY OF SPACE COMMUNICATIONS SPREAD SPECTRUM SYSTEMS

(Phase IV)

INTERIM PROGRESS REPORT

P.H. Wittke, P.J. McLane & S.J. Simmons

M.G. Wearring, Y.M. Lam. & W. Hopkins

Report No. 87-1

June 1987

Prepared for

The Department of Communications

Under DSS Contract No. 36001-6-3530/01-ST

Department of Electrical Engineering

Queen's University

Kingston, Ontario, Canada.

PREFACE

The work reported in this interim progress report is divided into four parts. The first two parts report the completion of work started under the previous contract. Part I presents the results of a study of the application of trellis codes to noncoherent frequency hopped spread spectrum systems. The second is the final results on the error performance of a high data rate, hopped spread spectrum transmission system that employs band efficient modulations that are coherent during each hop. Earlier reports provided results on the signalling spectrum, optimum receivers and system complexity.

In Part III, the first considerations of uplink synchronization of hopped spread spectrum systems are presented. A possible downlink synchronization procedure is described as well. Various aspects of synchronization are outlined, and options and strategies for system implementation are presented.

Part IV describes the research begun on the detection performance of intercept receivers for frequency-hopped, spread spectrum systems. The received signal is modelled as having an unknown but discrete frequency. Performance is determined for a coherent receiver, which represents an upper limit to attainable performance. Theory has been developed recently for the case of noncoherent interception. These results will be available for the final report.

I N D E X

Interim Progress Report

Part I

TRELLIS CODING FOR NONCOHERENT FREQUENCY-HOPPED COMMUNICATIONS SYSTEMS

by

M.G. Wearring,

P.H. Wittke.

ABSTRACT

In noncoherent frequency hopped spread spectrum communication systems, M-ary frequency shift keying (FSK) is the usual form of modulation. In this report, coded systems are studied which consist of convolutional codes with the codewords mapped onto multiple tone signal sets. The signals employ the same set of orthogonal tones as M-ary FSK, but several tones may be transmitted simultaneously. This allows an increase in the number of signals, to accommodate the redundant information introduced by the coding, without any bandwidth expansion. Some of Ungerboeck's rules for assigning signals to the trellis branches are employed.

The results of this study were obtained from simulation of transmission over an additive white Gaussian noise channel. Viterbi decoding was used, with both hard and soft decoding metrics. Performance was measured as the energy per bit to noise spectral density ratio (E_b/N_0) at a bit error rate (BER) of 10^{-4}. There are seven different coded systems which encompass information rates of 1, 2, and 3 bits per signalling interval, and employ sets of 2, 3, 4, or 8 orthogonal tones. There exists at least one coded system at each information rate which gives a performance improvement of 2 dB compared to the reference system of M-ary FSK. The maximum improvement achieved was 4.6 dB for the case of two tones. The best performance is obtained with soft decision decoding metrics. In some cases theoretical error bounds can be calculated, and they support the simulation results.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

1.1 Background

The purpose of this study is to investigate methods of improving the performance of a communications system by the use of coding. A block diagram of the coded system under consideration appears in Figure 1.1. The actual system is intended for digital satellite communications, and may be incorporated into a frequency hopped spread spectrum anti-jam application. The type of modulation considered is frequency shift keying (FSK) in which one of several different tones is sent in each signalling interval. Details of the modulation scheme and the receivers employed are given in Chapter two.

The basic principle of coding is that redundant information is added to the information bit stream so that transmission errors may be detected and/or corrected. Because of this redundant information, more signals must be sent over the channel to maintain the same information transfer rate as in the uncoded case. This can be accomplished by using the same set of signals at a faster signalling rate, or by increasing the number of signals used (signal set expansion). Both of these methods usually require a larger bandwidth than the original uncoded scheme for noncoherent FSK signalling. Alternatively, the information rate may be lowered to maintain the same bandwidth occupancy.

The goal of this study is to find a method of improving the system performance without sacrificing the data rate or increasing the bandwidth required. In frequency

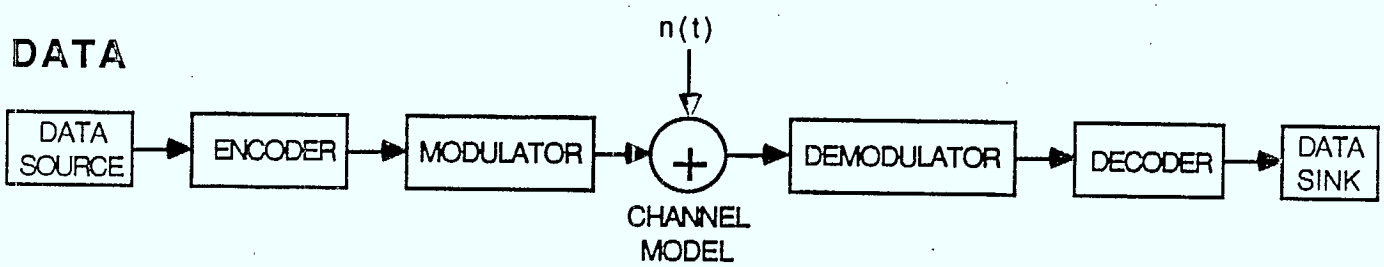


Figure 1.1 System Block Diagram

hopped systems, the processing gain is the ratio of the overall system bandwidth to the bandwidth occupied during each hop, and it indicates the ability of the system to reject jamming noise. Increasing the modulation bandwidth decreases the processing gain of the system, which is undesirable. The technique proposed in this study is to use convolutional codes with Viterbi decoding to improve error performance. The signal set is expanded by the use of multiple tone signals to accommodate the redundant bits, instead of the single tones which are used in conventional M-ary FSK. This does not increase the modulation bandwidth and thereby affect the processing gain. This study considers the performance of the encoder/decoder and modulator/receiver over an additive white Gaussian noise (AWGN) channel.

1.2 Literature Review

There have been several studies of the application of conventional coding to noncoherent anti-jam communication systems [1,2,3,4]. These papers explore the use of various codes to reduce the effect of a partial band jamming signal. Block codes, convolutional codes, and repetition codes (diversity) have all been investigated, as well as different combinations of concatenated codes. Reed-Solomon codes and convolutional codes both have good performance, especially when combined with diversity. Conventional M-ary FSK modulation was employed in all cases, and usually bandwidth expansion was allowed to accommodate the coding. Both hard and soft decision demodulation have been considered. Hard decision receivers select which signal is closest to the transmitted signal and relay that decision to the decoder. A metric, which is proportional to the logarithm of the probability that each signal was received, is computed by a soft decision receiver and used by the decoder. In a jamming environment, soft decision decoding is not desirable unless there is side information

about the presence of the jammer, because jammed signals seriously degrade the operation of the decoder. If side information is available, soft decision demodulation gives the best performance; otherwise hard decisions are superior [1]. This report does not consider the effects of jamming or frequency hopping on the communication system, but considers the performance in white noise. It is noteworthy that convolutional codes showed good performance under jamming conditions [3,4]. The Viterbi decoder is also readily adaptable to accept hard or soft demodulator decision variables.

There are two papers which consider the application of coding to noncoherent FSK signalling [5,6]. They both employ conventional M-ary modulation schemes and require bandwidth expansion or reduced data rate to accommodate the coding. The first study [5] considers continuous phase FSK, which has phase continuity between tones in subsequent signalling intervals, and also uses non-orthogonal tone spacings. The demodulation technique, although noncoherent, makes use of the phase continuity and employs an unconventional receiver structure. Practical frequency hopped systems, where the hopping occurs over large bandwidths, cannot maintain phase continuity between hops, and so these modulations are inappropriate. Keightley [6] studied the use of convolutional codes with binary and 4-ary FSK. Noncoherent demodulation with hard decisions was used for application to a frequency hopped spread spectrum system. The coding gain observed at a bit error rate (BER) of 10^{-5} was approximately 2 dB in the binary case, and negligible in the 4-ary case. The codes employed had rates of 1/2 and 1/3 with constraint lengths of 7 and 8 respectively. Thus the system transmission rates were reduced to 1/2 and 1/3.

Much work has been done recently on trellis coding [7,8,9,10] and a functioning system has been implemented in a modem. This technique provides large coding gains

without bandwidth expansion or reduction of data rate. All of the documented studies have been confined to coherent communication systems, in which the carrier may be both amplitude and/or phase modulated. The uncoded signal set is expanded by adding different levels of phase and amplitude modulation without requiring additional bandwidth. The resulting expanded signal constellation contains a symmetrical array of signal points. This signal set is partitioned or subdivided into subsets which have increasing distances between signal points in the subsets. The signal points are then assigned to the codewords of a convolutional code, according to a set of rules designed to provide maximum coding gain. This study employs some of the concepts of trellis coding applied to FSK signal sets. The signal constellations are expanded by using multiple tone signals so that additional bandwidth is not required. Signal set partitioning is also carried out, but not in the same manner as in coherent systems. The multiple tone signal constellations cannot be partitioned into subsets with increasing spacing between signal points. However, some of the rules from trellis coding are employed when mapping codewords onto signals. It is anticipated that the application of the principles of trellis coding will provide significant coding gains because of the success realized in coherent systems.

1.3 Report Summary

This report is divided into four chapters. An indication of the problem under consideration and the extent of research in this area has been given in the introductory chapter. Chapter two contains a detailed theoretical explanation of the various components in the system. First the modulation scheme and the structure of the different receivers is explained, followed by a description of the error correcting codes and the operation of the decoder. The second chapter concludes with an explanation of how the

coding and modulation schemes are combined in the overall system and presents some bounds on error performance. There are seven different cases which were investigated, and their characteristics are detailed in this chapter.

In general, the problem of theoretically evaluating the error performance of the noncoherent coded system does not appear tractable, and so computer simulations were carried out. Chapter three is a presentation of the results of the simulation programs. A brief description of the program structure is given first, followed by error performance curves for each of the seven cases considered. The results for all the cases are summarized in a table and then the performance of each case is discussed in detail.

Chapter four contains the general conclusions which may be derived from the simulation results. Suggestions for further research on this problem are given also.

CHAPTER TWO

THEORY

2.1 Modulation

2.1.1 Modulation Scheme

The reference modulation for this study is M-ary FSK. This scheme is commonly used in frequency hopped spread spectrum satellite systems for digital communication [1,6]. One of a set of M tones is sent in each signalling interval (T seconds). The number of tones (M) may be 2, 4, 8, or 16 corresponding to 1, 2, 3, or 4 bits of information per signal. The tones are orthogonally spaced in the frequency domain, which entails a minimum frequency difference of 1/2T hertz between adjacent tones.

The alternative modulation schemes, which allow coding without rate reduction, are divided into two types. The first expanded modulation technique employs the same set of orthogonal tones, but more than one tone can be transmitted in each interval (jT,(j+1)T). All possible combinations of the M tones are used, and the signal space is expanded to 2^M signals which transmit M bits per symbol period. This multiple tone signal constellation does not require any additional bandwidth, but there are other considerations. The transmitter requires increased power to send several tones simultaneously, if each tone has the same energy as a single tone in the reference system. The different modulation schemes will be compared with the same average signal energy, but the variation in energy between signals for this multiple tone case

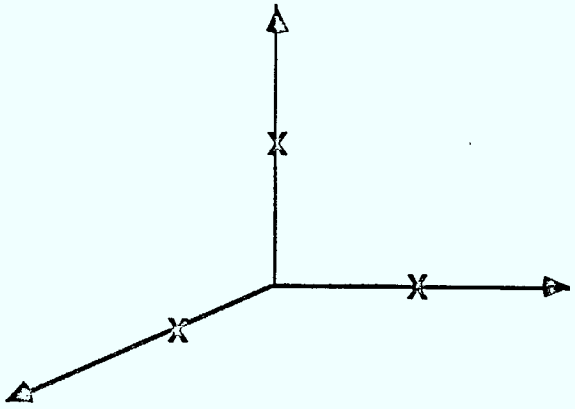
could require a transmitter with a larger peak power capability. This signal constellation also includes a zero signal (no tones sent), which could be undesirable in some systems.

The spacing of signals in the constellation has a significant effect on the system performance. In the reference M-ary FSK system, the single orthogonal tones are equidistant in the signal space. The multiple tone signals with constant energy tones form the vertices of a hypercube in M dimensional space, where M is the number of tones. This can best be visualized in three dimensions as shown in Figure 2.1. This first type of modulation is used in four of the seven cases to be considered. Case 1A employs two tones, Case 2A employs three tones, and both Cases 2B and 3A use four tone signal sets.

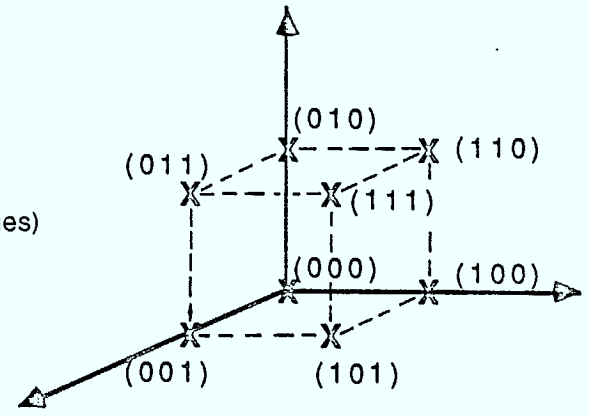
The second type of multiple tone signal set was devised to reduce the variation in signal amplitude among the various possible signals. The same set of orthogonal tones is again employed. Instead of permitting all possible tone combinations, a more limited subset is allowed to comprise the signal set. For the two and four tone cases, only single tones and pairs of tones are used as signals. The eight tone case employs single tones and sets of three tones. The pairs and triples of tones have the amplitude of each tone reduced so that the signal energy is equal to that of a single tone signal. The zero signal is excluded from these signal sets and therefore the signal energy is constant for this scheme. This makes demodulation simpler as will be explained in the discussion of receiver structure.

The constant signal energy multiple tone scheme is used in three cases which were simulated. Case 1B employs two tones and thus only three signals are available in each signal interval. These signals are $s_1(t) = \cos \omega_1 t$, $s_2(t) = \cos \omega_2 t$, and $s_3(t) = (\cos \omega_1 t + \cos \omega_2 t)/ \sqrt{2}$. The modulation is performed over two symbol intervals

(a) Single Tone (FSK)



(b) Multiple Tone (constant energy tones)



(c) Multiple Tone (constant signal energy)

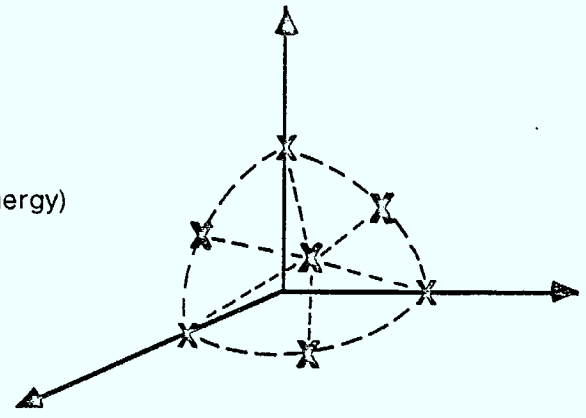


Figure 2.1 Signal Constellations

(2T) so that nine possible combinations, using three signals per period, are available. Eight of these points were chosen to represent three bits of data in each interval of 2T. For the case of four tones, there are four single tone signals and six different possible pairs of tones. Simulation Case 2C employs eight of these ten signals in the constellation. This allows three bits to be transmitted during each symbol interval (T seconds). The final case employing the second alternative modulation scheme is designated number 3B and uses eight orthogonal tones. There are eight signals containing a single tone and fifty-six possible three tone combinations. The combined signal constellation has sixty-four signal points which permits transmission of six binary digits with each signal.

The multiple tone modulations with constant signal energy have an irregular spacing between points in the signal constellation. The maximum spacing between signals is the same as the distance between orthogonal tones, and occurs between signals with no tones in common. When signals possess common tones, the spacing is reduced. The more tones in common between two signals, the smaller the distance between the two. The amplitudes of the tones in common between the signals also affect the spacing. This irregular spacing must be considered when mapping code words onto the signal points, and is discussed in the third section of this chapter.

2.1.2 Coherent Receiver

The coherent receiver is the optimum receiver for reception in the presence of additive white Gaussian noise (AWGN) when the phase of the transmitted signal is known. Then the local oscillator is perfectly matched to the incoming tones to eliminate any phase offset. This receiver cannot be used in the frequency hopped spread

spectrum system because phase continuity between subsequent signal tones and between frequency hops is not provided by practical transmitters. These phase jumps, and the phase offsets introduced during transmission, preclude the matching of the local oscillator to the transmitted signal. However, the coherent receiver provides the best possible performance that could be achieved, and will be included for purposes of comparison. The structure of the coherent receiver is simpler than the noncoherent receiver, and this allows a simulation to be performed more quickly and easily. Also, the performance of the coherent receiver can be readily analyzed, and there are theoretical results that provide verification of the simulation program. If coherent demodulation was actually used, a more efficient modulation scheme than FSK would likely be employed.

The usual optimum coherent receiver [11, p. 235; 12, p. 49] calculates the squared Euclidean distance between the received signal (r(t)), and each of the possible transmitted signals ($s_i(t)$). A decision is made in favour of the signal closest to the received signal, based on the minimum of the computed distances. The squared distance is given by

$$\begin{aligned} d_i^2 &= \int_0^T [r(t) - s_i(t)]^2 dt \\ &= \int_0^T r^2(t)dt - 2\int_0^T r(t)s_i(t)dt + \int_0^T s_i^2(t)dt \end{aligned} \tag{2.1}$$

The first term in (2.1) is constant, independent of the index i, and so may be neglected in the search for the nearest $s_i(t)$. The remaining two terms can be multiplied by -1/2 to form a decision variable which is now maximized.

$$L_i = \int_0^T r(t)s_i(t)dt - \frac{1}{2}\int_0^T s_i^2(t)dt$$
$$= q_i - \frac{1}{2}E_i \qquad (2.2)$$

The structure of this receiver is well-known [11, p. 235; 12, p. 49] and is shown for two cases in Figure 2.2. The receiver for conventional binary FSK is shown in part (a) of the diagram, where $s_1(t) = \cos \omega_1 t$ and $s_2(t) = \cos \omega_2 t$. The multiple tone scheme with different signal energies is demodulated by the system in part (b). In this case, the signals are $s_0(t) = 0$, $s_1(t) = \cos \omega_1 t$, $s_2(t) = \cos \omega_2 t$, and $s_3(t) = \cos \omega_1 t + \cos \omega_2 t$. For constant energy signals, E_i is constant for all i, and the decision variable becomes simply q_i . The decision variables for multiple tone signals are in general obtained by summing the variables for each tone in the signal, and scaling to account for the signal energy. This means that the multiple tone receiver requires only one correlator for each tone.

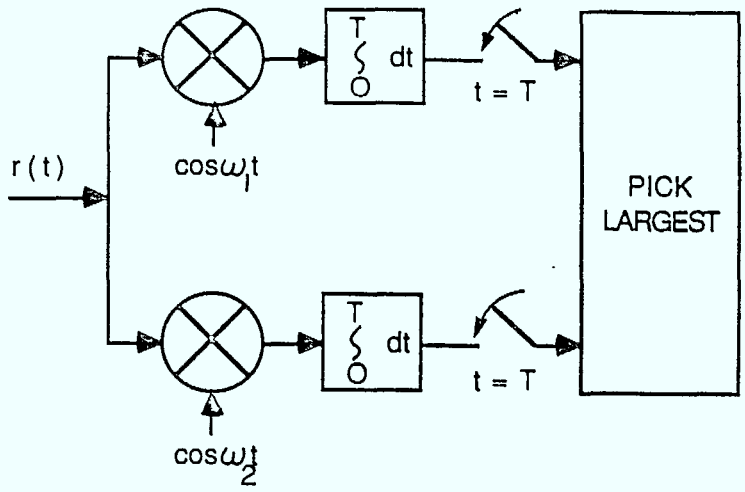
The theoretical error probabilities for uncoded signalling are easily obtained for the baseline system of M-ary FSK with coherent reception. The probability of symbol error for an orthogonal signal set over an AWGN channel is given by [11, p. 257; 13, p. 120]

$$P_S = 1 - \int_{-\infty}^{\infty} p_n(\alpha - \sqrt{E_S})d\alpha \left[\int_{-\infty}^{\alpha} p_n(\beta)d\beta\right]^{M-1} \qquad (2.3)$$

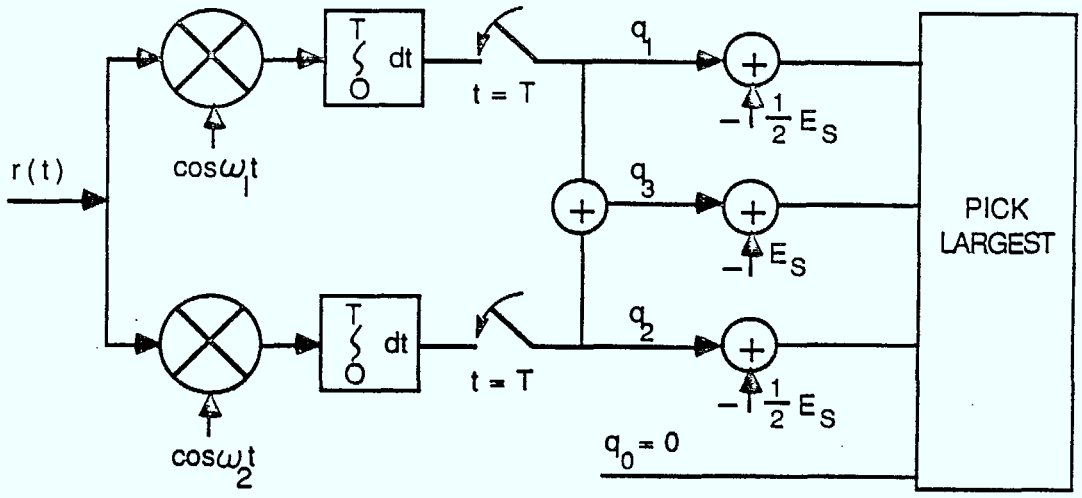
where M is the number of signals (ie. tones) and

$$p_n(\alpha) = \frac{1}{\sqrt{2\pi}} \exp \frac{-\alpha^2}{2}$$

is the Gaussian probability density function. The bit error probability is obtained from the symbol error probability as



(a) Single Tone Binary FSK



(b) Multiple Tone FSK (2 tones)

Figure 2.2 Coherent Receiver

$$P_B = \frac{\frac{1}{2}M}{M-1} P_S \tag{2.4}$$

The integral in equation (2.3) is tabulated in Golomb [13, p. 196] and these results were used for the baseline performance curves with coherent reception.

The theoretical performance may also be obtained for the case of multiple tone signalling with constant energy **tones.** With the ideal coherent receiver, the signal constellation may be considered as a rectangular signal set [11, p. 254] with error probabilities given by

$$P_S = 1 - (1-p)^M \tag{2.5}$$

$$P_B = p = Q\left(\frac{d}{\sqrt{2N_0}}\right) \tag{2.6}$$

where

M is the number of dimensions (ie. the number of tones),

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp -\frac{t^2}{2} dt$$

$$d = \sqrt{E_S}$$

E_S is the energy of a single tone, and

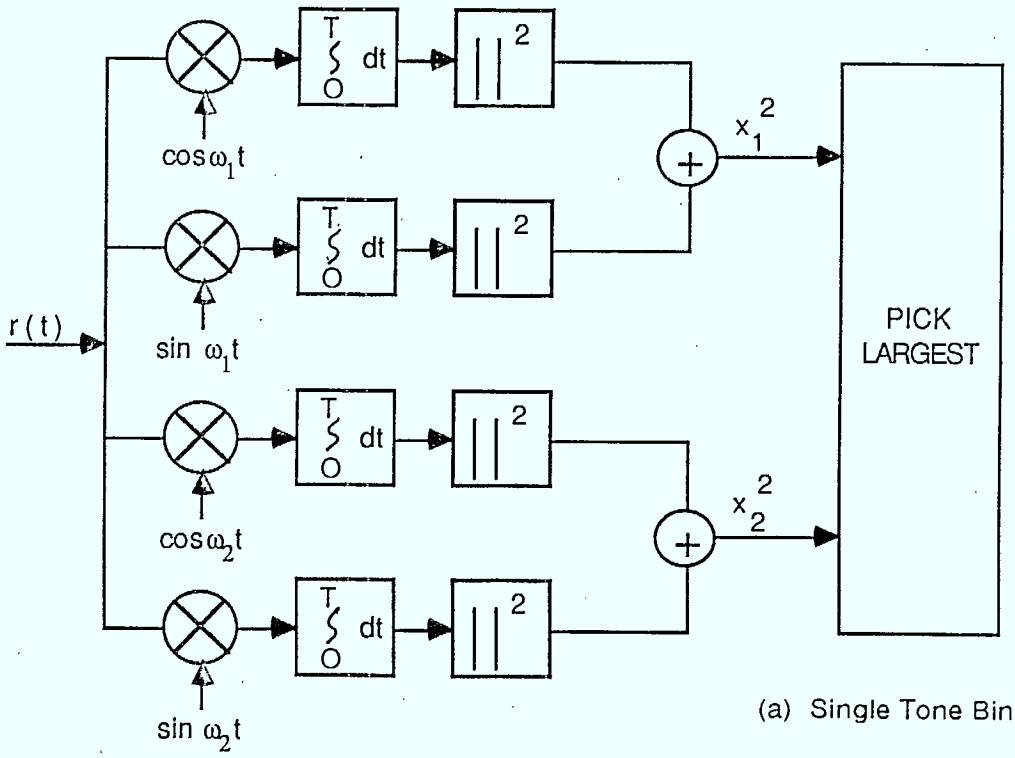
N_0 is the spectral density of the Gaussian noise in watts/hertz.

The multiple tone signalling schemes with constant signal energy do not form a rectangular signal set. Because of the irregular spacing of the signals, the probability of error varies with the signal that was sent, due to the different proximities of other signals. This does not allow for a simple expression for the error performance as in the previous cases, and the results are obtained strictly by simulation.

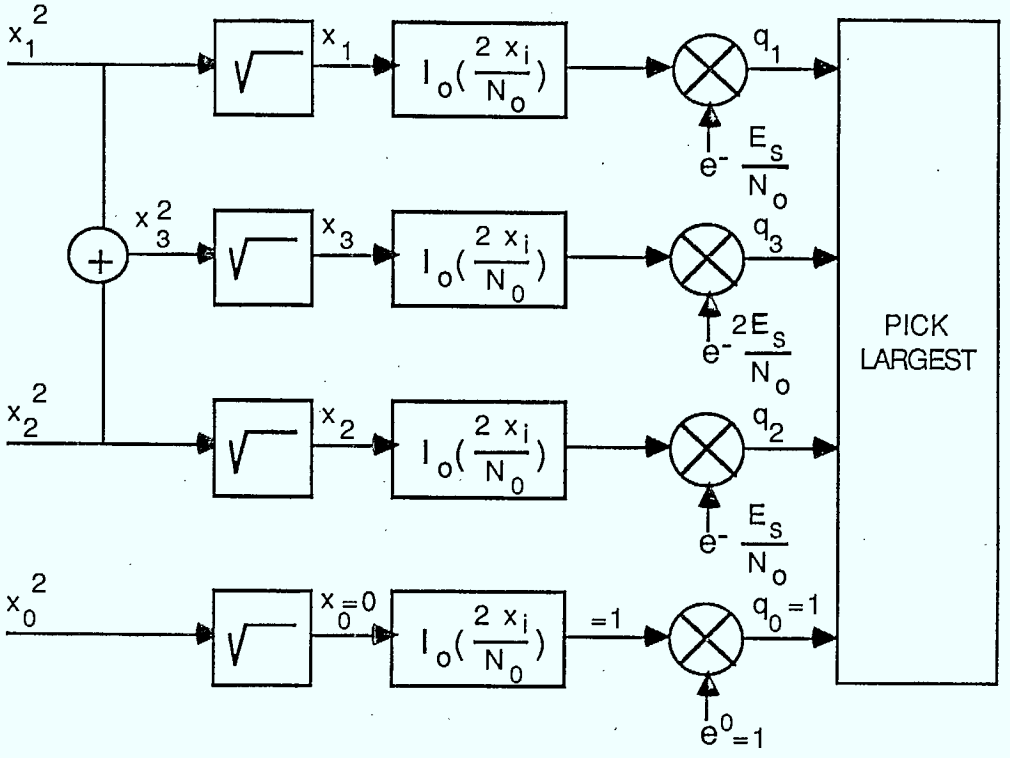
2.1.3 Noncoherent Receiver

The introduction of a random phase angle (θ) into the received signal, increases the complexity required for the optimum receiver. The structure of this noncoherent receiver is common [11, p. 519; 12, p. 104], and is shown in Figure 2.3. Two receivers appear in the diagram. Case (a) demodulates conventional binary FSK, where $s_1(t) = \cos \omega_1 t$ and $s_2(t) = \cos \omega_2 t$. In part (b), multiple tone signals with all possible combinations of two tones, and variable signal energy are demodulated. The signals for this receiver are $s_0(t) = 0$, $s_1(t) = \cos \omega_1 t$, $s_2(t) = \cos \omega_2 t$, and $s_3(t) = \cos \omega_1 t + \cos \omega_2 t$. There are two correlators for each signal, one in phase and one in quadrature, whose outputs are combined to negate the effect of the random phase. Multiple tone signals are assumed to have the same phase angle for all tones, and the probability distribution of the random phase is considered to be uniformly distributed between 0 and 2π radians.

This receiver calculates the probability of the received signal, given that the ith signal was sent, for all members i of the signal set. This probability is known as the likelihood function, and its logarithm is used as the decision variable. The signal with the maximum probability is chosen as the transmitted signal. The development of the theory for this receiver is rather lengthy, and can be found in the references [11, p. 511; 12, p.103], and so just the decision variable itself will be given here as



(a) Single Tone Binary FSK



(b) Multiple Tone FSK (2 tones)

Figure 2.3 Noncoherent Receiver

$$L_i = \ln\left[I_0\left(\frac{2X_i}{N_0}\right) \exp\left(-\frac{E_i}{N_0}\right)\right] \tag{2.7}$$

where

$$I_0(x) = \frac{1}{2\pi}\int_0^{2\pi} \exp[x\cos(\theta + \alpha)]d\theta$$

is the zero order modified Bessel function of the first kind,

$$E_i = \int_0^T s_i^2(t)dt$$

is the energy of the ith signal,

N_0 is the spectral density of the Gaussian noise in watts/hertz, and

$$X_i^2 = \left[\int_0^T r(t)s_i(t)dt\right]^2 + \left[\int_0^T r(t)\hat{s}_i(t)dt\right]^2 \tag{2.8}$$

is the sum of the squares of the correlator outputs for the in-phase and quadrature components of the correlation with the ith signal.

$s_i(t)$ is the in phase component of the signal since the signals are made up of cosine tones, and

$\hat{s}_i(t)$ is the quadrature component of the signal, which consists of sine tones at the same frequencies.

As was the case for the coherent receiver, the decision variables for multiple tone signals can be obtained from the correlator outputs for the individual tones. This means that only one pair of correlators is required for each tone in the system. The signal $s_i(t)$ giving rise to the maximum L_i is chosen as the signal sent.

For the baseline system and other constant signal energy cases, the decision variables can be further simplified. The exponential function in (2.7) will be a constant because E_i is the same for each index i, and so it can be ignored. To maximize the modified Bessel function, it is only necessary to maximize its argument because the Bessel function is a monotone increasing function. Therefore the optimum decision will be made by maximizing X_i or X_i^2. This receiver simplification results in the structure shown in Figure 2.3 (a), and is also known as square-law combining. The multiple tone modulation system with unequal signal energies cannot employ the simplified decision variables obtained from the correlator outputs. The modified Bessel function and the exponential function must be calculated in order to make a maximum likelihood decision.

The theoretical symbol error probability for noncoherent M-ary FSK is given by [11, p. 577]

$$P_S = \sum_{k=1}^{M-1} \frac{(-1)^{k+1}}{k+1} \binom{M-1}{k} \exp\left(-\frac{k}{k+1}\frac{E_S}{N_0}\right) \tag{2.9}$$

where

$$\binom{A}{B} = \frac{A!}{B!(A-B)!}$$

and the bit error probability is

$$P_B = \frac{\frac{1}{2}M}{M-1} P_S \tag{2.10}$$

The theoretical performance of the multiple tone noncoherent receiver is difficult to obtain. The presence of the Bessel function and the exponential function in the decision variables of the variable signal energy case makes the analysis difficult. The constant signal energy multiple tone modulation has irregular signal spacings which impedes the evaluation of theoretical error probabilities.

2.2 Coding

2.2.1 Description of Codes

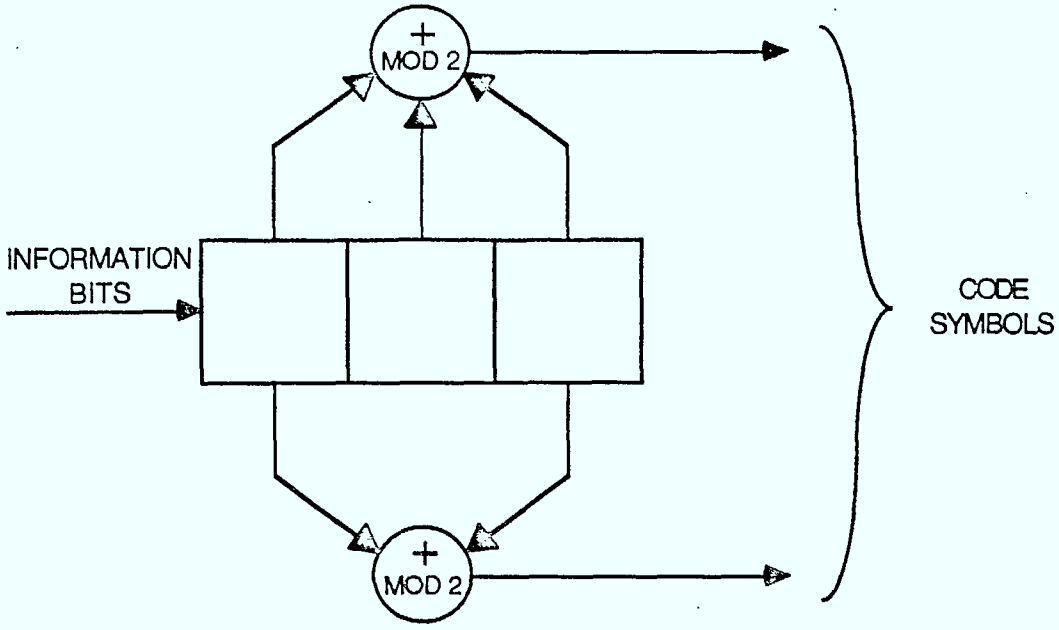
Convolutional encoding with Viterbi decoding is one of the more widely used methods of forward error correction. This is due to the ease of implementation and the relatively large coding gains obtainable from simple codes. As previously mentioned, this coding technique was chosen for consideration in this report. There are several different ways to describe convolutional codes, which will be briefly summarized below. More information on convolutional codes may be obtained from the references [12, p. 227; 14, p.227].

The first way to visualize a convolutional encoder is as a binary shift register with taps connected to modulo two adders. The information bit stream is shifted into the register in groups of b bits, and there are n modulo two adders which produce n output bits for each codeword. The rate of a code is given by the ratio b/n, which is the ratio

of input to output bits. The number of b-tuples in the shift register is denoted by k, and so the encoder retains b(k - 1) bits of the previous input data which define the state of the encoder. The value $\nu = b(k - 1)$ is defined as the constraint length of the code, and is the logarithm to the base two of the number of states. The length of the shift register (bk) is sometimes considered as the constraint length, but the previous definition (ν) will be used here as it is more useful for comparing codes of different rates. As the b-tuples of input data are shifted into the register, the state of the encoder changes and the output data is determined by the tap connections from the register to the adders. These connections are usually specified by generator polynomials or a generator matrix. A simple code with rate 1/2 and constraint length 2 is used as an example throughout this section. A diagram showing the shift register tap connections and the generator matrix appears in Figure 2.4.

Another way to describe a convolutional code is by means of a trellis diagram. The states of the code are assigned to nodes in the trellis and branches between states indicate a particular input and corresponding output symbol. It is obvious that a given input data sequence can be mapped onto a certain path through the trellis. There are 2^b branches which emerge from each node, corresponding to each of the possible inputs. The branches remerge at the next stage in the trellis in groups of 2^b at each state. The trellis for the example code is shown in Figure 2.5.

The final representation of a convolutional code is by a state diagram. The states of the code are again assigned to nodes in the diagram, and directed paths between states correspond to particular input and output symbols. The state diagram can be used to obtain the generating function of a code, which allows the weight profile to be determined. The state diagram for the model code appears in Figure 2.6.



$G = [\ 11\ \ 01\ \ 11\ ]$

$\nu = 2$

$k = 3$

Figure 2.4 Convolutional Encoder

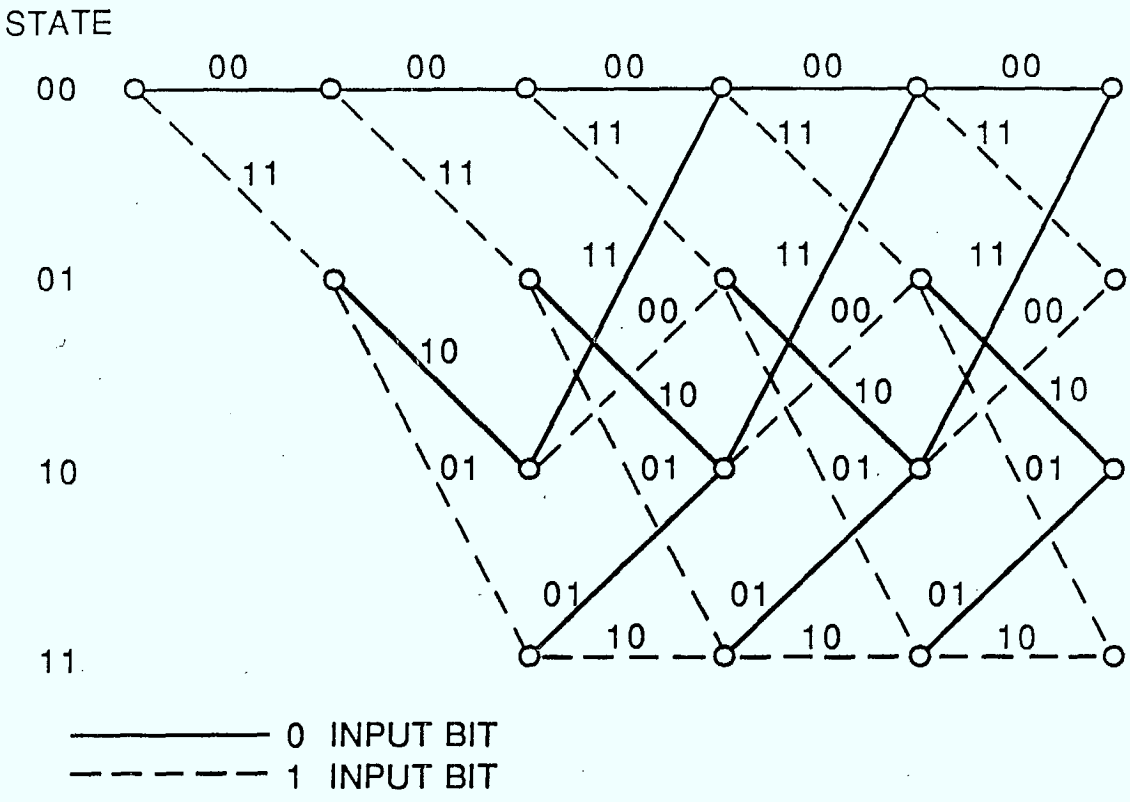


Figure 2.5 Trellis Diagram

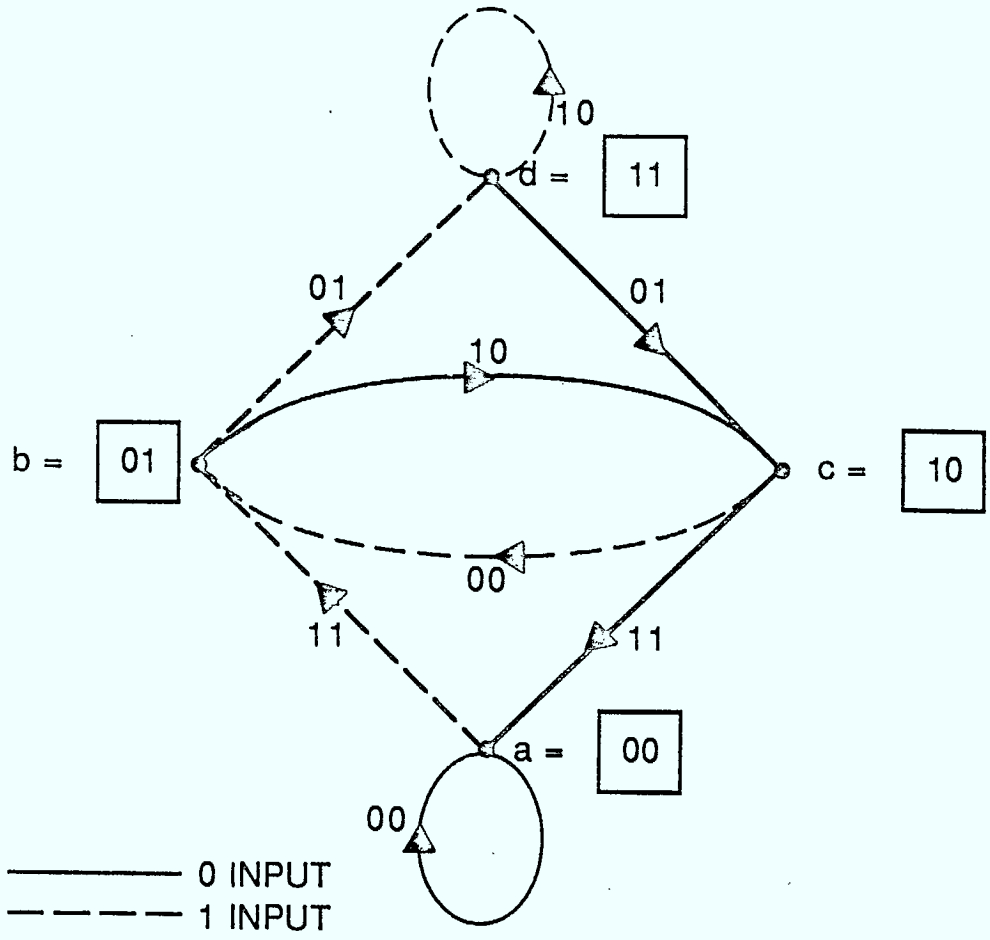


Figure 2.6 State Diagram

The weight profile is significant because it directly affects the error performance of the code. All the codes under consideration are linear, and therefore the all zeroes path may be considered as the correct path for the purpose of analysis. The weight or distance profile is obtained from the set of paths which diverge from the correct path and then remerge, corresponding to possible error events. The distance measure, for hard decision decoding, is the Hamming distance between the error path and the correct path. This is the number of output bits in which the two paths differ. These paths may be observed in the trellis diagram, and the weight profile obtained by adding up the Hamming weight of the output symbols along all paths which diverge from and then remerge with the all zeroes path.

A more tractable description of the weight profile is the code generating function. If the state diagram is redrawn with the zero state split, all the paths originating in one half of the zero state, passing through the other states, and entering the other half of the zero state will represent error paths. The branches of the state diagram are labelled

$$D^a L I^b \tag{2.11}$$

where

D represents the distance or output weight,

L represents the length of the path,

I represents the input weight,

a is the Hamming weight of the output symbol, and

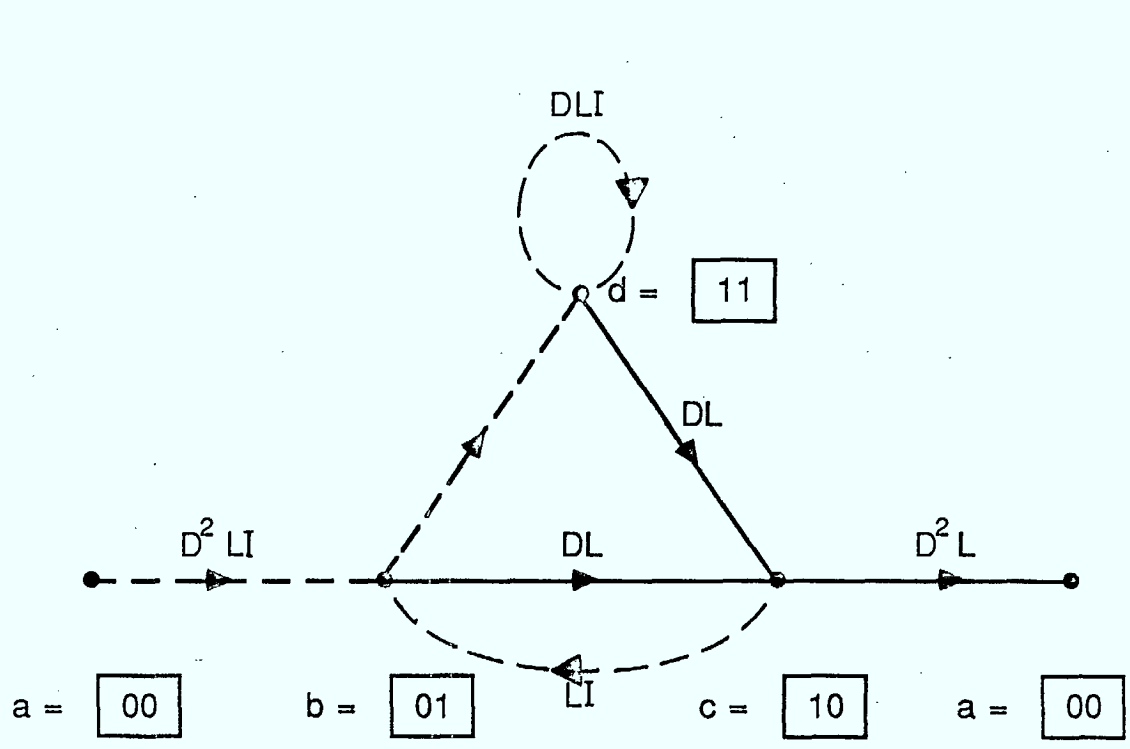
b is the Hamming weight of the input symbol.

Using signal flow graph theory, the state equations of the diagram are solved to yield the generating function T(D,L,I). This function has the form of an infinite sum of products of D, L, and I. The coefficient of each term represents the number of paths with distance of the exponent of D, input weight of the exponent of I, and length of the exponent of L. The sum may also be represented as a fraction of polynomials in D, L, and I. To obtain the weight profile, L and I are made equal to unity. The resulting function T(D) yields the number of paths at various distances from the correct path. The redrawn state diagram and the code generating function of the example code are shown in Figure 2.7. The free distance (d_f) is the minimum distance of any error path from the correct path, and is a good indicator of how well the code will perform.

The codes used in this study were obtained from other papers which investigated optimal codes [15, 16, 17]. The best codes had maximum free distance (d_f) for given code rate and constraint length. Only relatively short constraint length codes were used so that Viterbi decoding could be utilized. The optimal rate 1/2 codes were used in Case 1A [15]. They were also combined to form a rate 2/4 code for Case 2B and a rate 3/6 code for Case 3B. A dual-3 rate 1/2 code was also considered in Case 3B [17]. Rate 2/3 codes with maximum d_f were employed in Cases 1B, 2A, and 2C while rate 3/4 codes comprised Case 3A [16]. The various code generator matrices are shown in Table 2.1 and a complete description of each case appears in Table 3.1.

2.2.2 Maximum Likelihood Decoding

There has been much published work on the decoding of convolutional codes [12, p. 235; 18, 19, 20]. The Viterbi algorithm for maximum likelihood sequence estimation is



$$T(D,L,I) = \frac{D^5L^3I}{1 - DL(1 + L)I}$$

$$= D^5L^3I + D^6L^4(1 + L)I^2 + \cdots + D^{5+k}L^{3+k}(1 + L)^kI^{1+k} + \cdots$$

$$T(D) = \frac{D^5}{1 - 2D}$$

$$= D^5 + 2D^6 + \cdots + 2^kD^{k+5} + \cdots$$

Figure 2.7 Generating Function

TABLE 2.1

Convolutional Code Generator Matrices

Rate	Constraint Length	Free Distance	Generator Matrix
1/2	2	5	[11 01 11]
1/2	6	10	[11 01 11 11 00 10 11]
2/3	4	5	$\begin{bmatrix} 101\ 100\ 110 \\ 011\ 101\ 011 \end{bmatrix}$
2/3	6	7	$\begin{bmatrix} 101\ 111\ 010\ 101 \\ 011\ 111\ 101\ 011 \end{bmatrix}$
2/4	6	6	$\begin{bmatrix} 1100\ 1100\ 0100\ 1100 \\ 0011\ 0011\ 0001\ 0011 \end{bmatrix}$
3/4	5	5	$\begin{bmatrix} 1001\ 1111\ 0000 \\ 0101\ 0101\ 1001 \\ 0011\ 0100\ 0011 \end{bmatrix}$
3/6	3	6	$\begin{bmatrix} 100100\ 100101 \\ 010010\ 010100 \\ 001001\ 001010 \end{bmatrix}$
3/6	6	5	$\begin{bmatrix} 110000\ 010000\ 110000 \\ 001100\ 000100\ 001100 \\ 000011\ 000001\ 000011 \end{bmatrix}$

the most popular decoding technique for codes of short constraint length. This technique was investigated as being the most likely to be employed for a digital satellite channel. A brief summary of the decoding operation follows.

The basic problem can be most easily visualized as selecting the best path through the code trellis, based on information from the receiver. The maximum a posteriori probability for the path is used as the selection criterion. A decision metric for each branch is computed based on the received signal in each signalling interval. The metric is proportional to the logarithm of the likelihood function, in keeping with the coding literature [12, p. 238; 21, p. 188]. This metric is maximized by the decoder. The metric may also be considered as a distance measure between code vectors, in which case the negative log likelihood would be used and the decoder would perform minimization of the metric. Proceeding with the convention adopted, we have

$$p(\mathbf{r} \mid \mathbf{x}^k) = \prod_i p(r_i \mid x_i^k) \tag{2.12}$$

for a memoryless channel, where

$\mathbf{r}$ is the vector of receiver outputs with components r_i and

$\mathbf{x}^k$ is the code symbol vector for the kth trellis path with components x_i^k.

The metrics are obtained as

$$M_k = \ln p(\mathbf{r} \mid \mathbf{x}^k) = \sum_i m_i^k \tag{2.13}$$

where

$$m_i^k = \ln p(r_i \mid x_i^k) \tag{2.14}$$

M_k is the metric for the kth path, and

m_i^k is the metric for the ith code symbol on the kth path.

To accumulate a metric for every possible path through the trellis would be prohibitive, as the number of paths grows exponentially at each stage. The Viterbi algorithm makes use of the fact that paths remerge into each node in groups of 2^b at each stage in the trellis. It is necessary to keep track of only one optimum path leading into each state. The metric for each branch (branch metric) into a given state is added to the accumulated metric (state metric) for the best path into the previous state from which the branch originated. These 2^b metrics are compared, and the best path is retained as the **survivor** into that state. This process is repeated at each stage through the trellis. The maximum likelihood path will never be discarded by this method because none of the paths originating from a given state can accumulate a better metric than the survivor.

The decoder must select a single most likely path in order to deliver an output symbol. If a sufficiently long path history is kept, the 2^ν survivor paths (one for each state) will share a common stem, and the oldest bits corresponding to all the paths will be the same. For the sake of reduced complexity, it is desirable to truncate the path history at some fixed decoding depth. Several authors [12, p. 258; 14, p. 261] have shown that little degradation from optimum performance occurs when the decoding

depth is chosen to be from five to ten times ν, depending on the code rate. The output bits are chosen from the path that currently has the best accumulated metric.

The actual metric used by this decoder may come from either a hard or soft decision receiver. The nature of the metric employed does not affect the operation of the decoder, although it will influence the error performance. For a hard decision receiver, bit decisions are made at the demodulator output, prior to decoding. The negative Hamming distance between the received symbol and the output symbol of each branch is used as the branch metric in this case. The Hamming distance between codewords is generally proportional to the distance between the corresponding signals in the signal space. Thus minimizing the Hamming distance between code vectors, or equivalently maximizing the negative Hamming distance, is the usual hard decision metric. If the added complexity of a soft decision demodulator can be tolerated, then more information is available and the branch metric is proportional to the logarithm of the likelihood function. This results in better error performance.

2.3 Signal Space Mapping

The assignment of codewords to signal points has an important influence on the overall system performance. Previous work on the design of optimal codes has used the Hamming distance between codewords as the distance measure between paths [15, 16]. The Euclidean distance in the signal space between points in the signal constellation actually determines the probability of a transmission error. Ideally, the codewords would be assigned to signal points with Euclidean distance spacings proportional to the Hamming distance of the codewords, so that erroneous decoding into a near neighbour would result in few errors.

Trellis coding [7, 8] considers the actual paths through the trellis when assigning signals. As previously stated, the signal constellation is divided, by a method termed set partitioning, into subsets with maximum spacing between signal points. The signals are assigned to branches in the trellis according to a set of rules that attempts to maximize the Euclidean distance between signal error paths. These rules are as follows

1. all signals should occur with equal frequency and with a fair amount of regularity and symmetry;

2. parallel transitions between states are assigned signals from the subset with maximum signal spacing;

3. transitions originating from the same state are assigned signals from a subset with maximum possible spacing;

4. transitions ending in the same state are assigned signals from a subset with maximum possible spacing.

The papers on trellis coding [7, 8] also recommend that the signal set be expanded to twice the number of points in the uncoded signal set in order to achieve the maximum coding gain without unnecessary complexity.

The multiple tone signal sets which use constant energy orthogonal tones (ie. variable signal energy), form a rectangular constellation as described previously. If each bit of a codeword is mapped onto a different tone, the Euclidean distance between signal points is proportional to the square root of the Hamming distance between codewords.

This situation makes the mapping of codewords onto signal points straightforward, in order to obtain optimum coded performance.

The constant signal energy constellations have an irregular signal spacing, which varies with the number of tones in common between signals. The Euclidean distance between signals cannot be easily mapped onto the Hamming distance between codewords, and so the rules for trellis coding were applied. Set partitioning could not be done in the usual way because of the irregular signal spacings. The signal sets were divided into groups of 2^b (for a rate = b/n code) with maximum spacing between all the members of each subset. All of the codes used have a similar structure, so that the different signal sets may be subdivided the same way. None of the codes have parallel transitions, so that the second rule of trellis coding may be disregarded. The branches in the various trellises diverge and remerge in groups of 2^b according to the rate of the code. For all the optimum codes used, the same group of codewords are associated with the branches that diverge from and remerge into a given state. This allows for the partitioning of the signal sets into groups of 2^b , to produce a desirable mapping onto the codewords. The one exception to this situation is the dual-3 rate 1/2 code used in Case 3B. The same groups of codewords do not appear on the diverging and remerging branches, so that trellis coding rules three and four cannot be satisfied simultaneously. No alternate mapping could be found to improve the spacings between error paths, so the same mapping scheme as for the other rate 3/6 code was employed.

The actual assignment of codewords to signals for the constant signal energy multiple tone constellations are shown in Table 2.2. The codewords are shown as decimal numbers, grouped according to the branch assignments, and the signals are represented by ones and zeroes. Each digit of the signal representation corresponds to

TABLE 2.2

Signal Space Mappings

Case 1B: Rate=2/3, 2 tones, Constant Energy Signals (signalling over 2T)

0	01 01	1	01 11
3	10 01	2	11 01
5	01 10	4	11 10
6	10 10	7	10 11

Case 2C: Rate=2/3, 4 tones, Constant Energy Signals

0	0001	1	0110
3	0010	2	0011
5	0100	4	1001
6	1000	7	1100

Case 3C: Rate=3/6, 8 tones, Constant Energy Signals

0	0001 0000	1	0010 0000	4	0100 0000	5	1000 0000
3	1110 0000	2	1101 0000	7	1011 0000	6	0111 0000
12	1000 0011	13	0100 0011	8	0010 0011	9	0001 0011
15	0100 1001	14	1000 1001	11	0001 1001	10	0010 1001
48	0010 0101	49	0001 0101	52	1000 0101	53	0100 0101
51	1000 1100	50	0100 1100	55	0010 1100	54	0001 1100
60	0100 0110	61	1000 0110	56	0001 0110	57	0010 0110
63	0010 1010	62	0001 1010	59	1000 1010	58	0100 1010
16	0000 0001	17	0000 0010	20	0000 0100	21	0000 1000
19	0000 1110	18	0000 1101	23	0000 1011	22	0000 0111
28	0011 1000	29	0011 0100	24	0011 0010	25	0011 0001
31	1001 0100	30	1001 1000	27	1001 0001	26	1001 0010
32	0101 0010	33	0101 0001	36	0101 1000	37	0101 0100
35	1100 1000	34	1100 0100	39	1100 0010	38	1100 0001
44	0110 0100	45	0110 1000	40	0110 0001	41	0110 0010
47	1010 0010	46	1010 0001	43	1010 1000	42	1010 0100

a tone, and a one indicates a tone which is sent as part of the signal. The use of the signal space mapping described here was compared to an arbitrary mapping scheme for the four tone signal constellation. An improvement of approximately 1.5 dB (E_b/N_0) was obtained for a rate 2/3 code with constraint length of 6.

2.4 Error Performance

The performance of the coded system depends on the structure of the code, the decoding metric, and the coding channel. The coding channel is the effective channel as seen by the encoder and decoder. It includes the properties of the modulator and demodulator, as well as the actual transmission channel. Further details of the following development of error bounds may be found in the references [12, p. 242; 14, p. 243; 21, p. 192; 22].

The first step in obtaining the error performance of the system is to determine the pairwise error probability between two transmitted code vectors. This is the probability that the metric for the error path is larger than the correct path metric for a given received signal sequence, and is expressed by

$$p(\mathbf{x} \rightarrow \hat{\mathbf{x}}) = p\left\{\sum_n m(r_n, \hat{x}_n) \geq \sum_n m(r_n, x_n)\right\} \tag{2.15}$$

The Chernoff bound may be applied to the pairwise error probability to obtain

$$p(\mathbf{x} \rightarrow \hat{\mathbf{x}}) \leq \prod_n E\left\{ \exp(\lambda[m(r_n, \hat{x}_n) - m(r_n, x_n)]) \mid x_n\right\} \tag{2.16}$$

For most metrics of interest, the expected value in (2.16) has the form

$$D(\lambda) = E\left\{ \exp(\lambda[m(r_n, \hat{x}_n) - m(r_n, x_n)]) \mid x_n \right\} \tag{2.17}$$

so that

$$p(\mathbf{x} \rightarrow \hat{\mathbf{x}}) \leq [D(\lambda)]^{w(\mathbf{x}, \hat{\mathbf{x}})} \tag{2.18}$$

where $w(\mathbf{x}, \hat{\mathbf{x}})$ is the Hamming distance between $\mathbf{x}$ and $\hat{\mathbf{x}}$ or the number of bits which differ in the two sequences. For the case of an arbitrary metric, the parameter D is given by

$$D = \min_{\lambda \geq 0} D(\lambda) = \min_{\lambda \geq 0} E\left\{ \exp(\lambda[m(r, \hat{x}) - m(r, x)]) \mid x \right\} \Big|_{x \neq \hat{x}} \tag{2.19}$$

When the maximum likelihood metric is used for decoding, ie.

$$m(r, x) = \ln p(r \mid x) \tag{2.20}$$

then one may use the Bhattacharyya parameter

$$Z = \sum_r \sqrt{p(r \mid x)\, p(r \mid \hat{x})} \Big|_{x \neq \hat{x}} \tag{2.21}$$

for the value of D.

To obtain the probability of an error event, the union bound is used to give

$$P_E \leq \sum_j a(j) D^j \tag{2.22}$$

where

P_E is the probability of an error event

$a(j)$ is the number of error paths of distance j, and

D^j is the pairwise error probability for a path of distance j.

Since the convolutional codes employed are linear, the all zeroes path may be considered as the transmitted sequence, and the coefficients a(j) are obtained from the generating function T(D) to yield

$$P_E \leq T(D) \tag{2.23}$$

with the value of D determined by the coding channel.

The bound on the probability of bit error can be obtained in a similar fashion, and is given by

$$P_B \leq \frac{1}{b} \sum_i \sum_j i\, a(i,j)\, D^j \tag{2.24}$$

where

b is the number of information bits per code symbol,

i is the information weight of the path (ie. the number of bit errors), and

a(i,j) is the number of paths of weight j with information weight i.

The coefficients a(i,j) are the same as in the augmented generating function T(D,I), and the values of i are the same as the exponent of I in each term of the function. The augmented generating function is obtained by equating L to unity in the function T(D,L,I), such as that shown in Figure 2.7. The bit error bound can be written incorporating this function as

$$P_B \leq \frac{1}{b} \frac{\partial T(D, I)}{\partial I}\Big|_{I=1} \tag{2.25}$$

The error bounds presented in this section are used to verify the simulation results for the cases where the coding channel can be easily characterized. The multiple tone signals with variable signal energy form a rectangular signal constellation which can be readily analyzed for coherent reception. The probability of error between two adjacent signals is obtained from the error function

$$p = Q\left(\frac{d}{\sqrt{2N_0}}\right) \tag{2.26}$$

where d is the Euclidean distance between the two signal points. The probability of an error event for a rectangular signal set of n dimensions (ie. n tones) is

$$P_E = 1 - (1 - p)^n \tag{2.27}$$

For the case of hard decision decoding, we may use the Bhattacharryya parameter for an M-ary symmetric channel as the value for D

$$D = \left(\frac{M-2}{M-1}\right)P_E + 2\sqrt{\frac{P_E(1-P_E)}{M-1}} \tag{2.28}$$

The rectangular signal set is not strictly an M-ary symmetric channel, since all of the cross-over probabilities between signals are not equal. However, the error bound still applies in this case.

To obtain an error bound for soft decision decoding, we note that the probability that one code vector is decoded as another code vector is given by

$$P_E(w) = Q\left(\frac{d}{\sqrt{2N_0}}\right) \tag{2.29}$$

where d is again the Euclidean distance between code vectors. The error function may be upper bounded by an exponential function as follows

$$Q\left(\frac{d}{\sqrt{2N_0}}\right) \leq \frac{1}{2} \exp - \frac{d^2}{4N_0} \tag{2.30}$$

For the rectangular signal constellation, the Euclidean distance is proportional to the square root of the Hamming distance, and so we may write

$$d = \sqrt{wE_S} \tag{2.31}$$

where

w is the Hamming distance between the code vectors, and

E_S is the energy of a tone.

Thus the error bound may be rewritten as

$$P_E(w) \leq \frac{1}{2}\left[\exp - \frac{E_S}{4N_0}\right]^w \tag{2.32}$$

To obtain the bound for the overall probability of an error event, we sum the individual components over w to yield

$$\begin{aligned} P_E &= \sum_w P_E(w) \\ &\leq \sum_w \frac{1}{2} a(w) \left[\exp - \frac{E_S}{4N_0} \right]^w \\ &= \frac{1}{2} T(D) \Big|_{D = \exp - \frac{E_S}{4N_0}} \end{aligned} \tag{2.33}$$

These two values for the parameter D may also be used to find the bound on the bit error probability for coherent demodulation of coded rectangular signal sets. Equation (2.24) or (2.25) may be used, with a multiplying factor of one half in the soft decision case. The bit error bounds were computed for the various codes used with multiple tone signals with fixed signal energy per orthogonal component (ie. constant energy tones). These bounds are plotted with the corresponding simulation results in the appropriate subsections of the next chapter. The performance of coding with constant signal energy constellations and coherent demodulation, and the performance of all the coded systems employing noncoherent reception are not easily evaluated. Just as the error probability for uncoded transmission was difficult to obtain, the evaluation of the bound parameter D is not tractable for these cases.

CHAPTER THREE

RESULTS

3.1 Overview

3.1.1 Program Description

The simulation program used to model the communication system was written in FORTRAN, and run on both VAX 11/750 and IBM 3081 computers. The program is divided into a mainline routine, which handles the initialization, input, and output tasks, and several subroutines which correspond to the various components of the system. The input file for the simulation contains the parameters of the code, which are the rate, the constraint length, and the generator matrix. A look-up table of output symbols corresponding to various branches in the trellis is generated by an initialization subroutine. This table is used by the encoding and decoding subroutines. The type of demodulation (coherent/noncoherent), and whether hard or soft decisions are to be made, are also indicated in the input file. The different signal space mappings are obtained by using different versions of the channel subroutine, which will be described later.

After the initialization procedures are complete, the main program commences by generating a random bit stream. This is done with a library subroutine for pseudorandom number generation. The information symbols are passed to the encoder subroutine which returns the encoded data symbols. These data symbols are passed to the channel subroutine which simulates the modulator, the AWGN channel, and the

demodulator. In each signalling interval, the subroutine calculates and returns a metric for each member of the signal space. These metrics are then passed to the decoding subroutine, which constructs the received bit stream. The main program compares the received data to the original information bits, and keeps track of the errors. Error events, symbol errors, and bit errors are all tabulated. The signal to noise ratio (SNR) starts at one dB and is incremented in steps of one dB. One hundred thousand data symbols are simulated at each level of SNR. After each set of data points is processed, the three error probabilities are calculated and sent to a data file along with the SNR value. The SNR is then incremented and the procedure repeats until the error event count is less than ten in one hundred thousand data points at the given SNR. The output data file is used to plot error performance curves which appear later in this chapter.

The encoder subroutine preserves the previous input bits which determine the state of the encoder. The current information symbol then determines the transition to the next state and the corresponding output symbol. This is accomplished by using the present state and input bits as an index to the table generated during initialization. The output symbol is contained in the table, and is returned to the mainline by the encoder subroutine.

The channel subroutine receives the data symbol and several parameters of the modulation scheme. The SNR level, the type of demodulation (coherent/noncoherent), and the type of decoding metric (hard/soft) are all passed from the mainline. Different versions of the channel subroutine are used to accommodate the various signal space mappings in the simulations of the different cases. Given the transmitted signal and the SNR value, a decision variable is computed for each member of the signal space.

Samples of Gaussian noise are calculated from uniformly distributed pseudorandom numbers, according to the polar formula [6].

$$X_1 = V_1 \sqrt{-\frac{2 \ln S}{S}} \tag{3.1}$$

and

$$X_2 = V_2 \sqrt{-\frac{2 \ln S}{S}} \tag{3.2}$$

are two independent samples of a zero mean, unit variance, normally distributed random variable, where

$$S = V_1^2 + V_2^2 \tag{3.3}$$

and S must be less than unity.

$$V_1 = 2U_1 - 1 \tag{3.4}$$

and

$$V_2 = 2U_2 - 1 \tag{3.5}$$

where U_1 and U_2 are samples of a random variable, uniformly distributed between zero and one. A random phase angle is added to the transmitted signal for the case of noncoherent reception. The angle ($\emptyset$) is assumed to be uniformly distributed between zero and two pi radians, and is obtained from a pseudorandom number generated by the library subroutine. For soft decisions, the actual decision variables are returned to the main program as decoder metrics. The data symbol corresponding to the largest decision variable is determined in order to calculate the hard decision metric. The negative Hamming distance between the demodulated symbol and the branch symbol is

returned as the branch metric in this case. Details of the demodulation and metric calculation are given in Chapter two.

The decoder subroutine implements the Viterbi algorithm. A bit history and path metric are maintained for each state of the code. At each signal interval, the metric for the signal corresponding to each transition in the trellis is added to the path metric for the originating state of that branch. The paths entering each state are compared, and the one with the largest metric is retained as the survivor. This process makes use of the look-up table of trellis states and transitions, and the information symbols for the bit histories also come from the table. The output bits are taken from the bit history corresponding to the path with the largest metric at each step. The length of the history maintained by the decoder varies with the code constraint length and code rate according to the formula

$$L = 5b\nu \tag{3.6}$$

where

L is the length of the bit history,

ν is the code constraint length, and

b is the information bit rate (code rate = b/n).

This decoder has negligible performance degradation due to path history truncation. Clark and Cain [14, p. 262] suggest that path histories for near optimal decoder operation should be 5ν for rate 1/2 codes, 8ν for rate 2/3 codes, and 10ν for rate 3/4 codes.

TABLE 3.1

Case Descriptions

CASE	INFORMATION RATE	CODE RATE	CONSTRAINT LENGTHS	NUMBER OF TONES	NUMBER OF SIGNALS	E_{peak}/E_{av}
1A	1 bit/T	1/2	2, 6	2	4	2.0
1B	1 bit/T	2/3	4, 6	2	8	1.0
2A	2 bits/T	2/3	4, 6	3	8	2.0
2B	2 bits/T	2/4	6	4	16	2.0
2C	2 bits/T	2/3	4, 6	4	8	1.0
3A	3 bits/T	3/4	5	4	16	2.0
3B	3 bits/T	3/6	3, 6	8	64	1.0

NUMBER OF TONES = BANDWIDTH $\times$ 2T

3.1.2 Case Summary

There are seven different coding and modulation schemes considered in this report. The characteristics of each scenario, and the numbering scheme used to disinguish them, are shown in Table 3.1. They are divided into three groups, and the number of the case indicates the number of information bits transmitted per signalling interval. There is also a letter associated with each case to differentiate the different code rates and modulation schemes employed.

Four of the cases (1A, 2A, 2B, 3A) use signals which may contain any combination of tones. These signals have a variable signal energy, depending on the number of tones in the signal. They are compared to the baseline system according to average signal energy. For example, let E_S be the energy of a single tone. For the four tone case, there are sixteen signals with average signal energy given by

$$E_{av} = \frac{1(0) + 4(E_S) + 6(2E_S) + 4(3E_S) + 1(4E_S)}{16} = 2E_S \tag{3.7}$$

Therefore, the signal energy of this system is reduced by 1/2 for comparison with the baseline system, which has a signal energy of E_S. The other variable signal energy cases are also scaled appropriately. The remaining three cases (1B, 2C, 3B) have constant energy signals, with signal energy equal to the energy of a single tone, for direct comparison with the baseline system.

Each of the seven cases has several sets of results. The various systems were all simulated with both coherent and noncoherent demodulation, both hard and soft decoding metrics, and different code constraint lengths. All of the simulation results for each case appear on two figures, one for coherent reception and one for noncoherent

reception, each with the corresponding baseline system performance curve. The measure of performance on the graphs is the bit error rate (BER) plotted against the energy per bit to noise spectral density ratio (E_b/N_0). As mentioned previously, the error events and symbol errors as well as the signal energy to noise ratio were tabulated by the simulation program, but the bit error performance was chosen as the most suitable criterion for comparison. Table 3.2 shows the required E_b/N_0 to give a BER of 10^{-4} for every case simulated. These values were obtained from the figures shown later in this chapter, with some extrapolation required on some of the curves. For data transmission, system performance at a BER of 10^{-5} or less is usually of most interest. However, the simulation run time to obtain reliable data in this region is prohibitive. Some curves display irregular behaviour for the last data point (ie. the lowest BER point). This occurs because the BER is calculated using a very small number of error events, and the random occurrence of a single error causes a large displacement of the point on the graph. Since error performance curves are known to behave smoothly at low BER values, these points are neglected when extrapolating the curves. All performance values quoted in this chapter will refer to E_b/N_0 in dB at the reference BER of 10^{-4}.

As a verification of the simulation program performance, several test cases were run and compared to theoretical results. The baseline systems (2, 4, and 8-ary FSK) were simulated for both coherent and noncoherent demodulation, and the results matched closely with the theoretical performance. The modulation scheme which uses all possible tone combinations to obtain a rectangular signal set of 2^M signals from M orthogonal tones was also simulated for the various values of M. The performance of this system, with no coding and coherent demodulation, reflected the theoretical error probabilities for rectangular signal constellations. In both cases, the error performance curves for the simulation matched the theoretical results to within one-half dB over the entire curve. In

TABLE 3.2

Coherent Results

CASE	BASELINE PERFORMANCE	HARD DECISION PERFORMANCE	SOFT DECISION PERFORMANCE
1A	11.5	9.7 ($\nu = 2$), 8.3 ($\nu = 6$)	8.2 ($\nu = 2$), 6.4 ($\nu = 6$)
1B	11.5	12.6 ($\nu = 4$), 12.1 ($\nu = 6$)	9.7 ($\nu = 4$), 9.4 ($\nu = 6$)
2A	8.8	9.2 ($\nu = 4$), 9.3 ($\nu = 6$)	7.5 ($\nu = 4$), 7.0 ($\nu = 6$)
2B	8.8	9.9 ($\nu = 6$)	7.4 ($\nu = 6$)
2C	8.8	9.7 ($\nu = 4$), 9.5 ($\nu = 6$)	7.3 ($\nu = 4$), 6.8 ($\nu = 6$)
3A	7.4	9.3 ($\nu = 5$)	7.2 ($\nu = 5$)
3B	7.4	11.7 ($\nu = 3$), 9.9 ($\nu = 6$)	7.7 ($\nu = 3$), 5.0 ($\nu = 6$)

Noncoherent Results

CASE	BASELINE PERFORMANCE	HARD DECISION PERFORMANCE	SOFT DECISION PERFORMANCE
1A	12.3	10.8 ($\nu = 2$), 9.4 ($\nu = 6$)	9.2 ($\nu = 2$), 7.7 ($\nu = 6$)
1B	12.3	12.8 ($\nu = 4$), 12.4 ($\nu = 6$)	10.5 ($\nu = 4$), 10.0 ($\nu = 6$)
2A	9.6	10.1 ($\nu = 4$), 9.4 ($\nu = 6$)	8.1 ($\nu = 4$), 7.5 ($\nu = 6$)
2B	9.6	10.1 ($\nu = 6$)	8.0 ($\nu = 6$)
2C	9.6	10.0 ($\nu = 4$), 9.9 ($\nu = 6$)	7.6 ($\nu = 4$), 6.9 ($\nu = 6$)
3A	8.2	9.6 ($\nu = 5$)	7.7 ($\nu = 5$)
3B	8.2	12.0 ($\nu = 3$), 10.0 ($\nu = 6$)	7.8 ($\nu = 3$), 5.8 ($\nu = 6$)

coded systems which employ the rectangular signal constellations, error bounds were calculated as described in section 2.4. The coherent simulation results with their respective error bounds are shown in the appropriate subsections of this chapter.

3.2 Simulation Results

3.2.1 Case 1A: Rate = 1/2, 2 tones, Variable Signal Energy

This case has the best results of those investigated. All of the coded systems outperform the baseline by between 1.5 and 5.1 dB at the reference BER of 10^{-4}. The number of points in the signal space is doubled by the use of multiple tone signals so that optimal rate 1/2 codes can be employed. This provides a redundant coded bit for each bit of information transmitted, and thus good error correction ability. The two codes used in the simulation have constraint lengths of 2 and 6 with free distances of 5 and 10 respectively. The longer code is quite common, and has been used in other satellite applications [19].

The results for coherent reception are shown in Figure 3.1. The improvement over the baseline system ranges from 1.8 dB for the short code with hard decisions, to 5.1 dB for the long code with soft decisions. The use of soft decisions provides a gain of 1.5 dB over hard decisions for the short code, yielding a margin of 3.3 dB over the baseline. A 1.9 dB improvement, from 3.2 dB better than the reference with hard decisions to a 5.1 dB advantage with soft decisions, is obtained for the longer code. For hard decisions, the effect of increasing the constraint length from 2 to 6 yields 1.4 dB improvement, and 1.8 dB is gained in the soft decision case.

The performance of the noncoherent receiver appears in Figure 3.2. The results are very similar to the coherent case, with 1.5 to 4.6 dB of improvement over the baseline system. The short constraint length code provides a 1.5 dB advantage with hard decisions, and an additional 1.6 dB gain for soft decisions. Increasing the code constraint length enhances performance by 1.4 dB for a 2.9 dB improvement on the baseline with hard decisions, and by 1.5 dB in the soft decision case.

The theoretical error bounds for the two codes are also shown. Figure 3.3 shows the bound for the shorter code, and the longer code appears in Figure 3.4.

3.2.2 Case 1B: Rate = 2/3, 2 tones, Constant Signal Energy (signalling over 2T)

Case 1B uses a rate 2/3 optimal code with the three bit codewords sent over two signalling intervals. There is less redundant information in the data stream, and so the performance is understandably worse than Case 1A. The advantage to this scenario is that the signals have constant energy. In this case, the computation of the decision variables is simplified. However, the Euclidean distance between the signals is smaller in this case than in the previous modulation scheme, which is detrimental to error performance. Codes with constraint lengths of 4 and 6 and corresponding free distances of 5 and 7 were simulated.

For the coherent receiver, the worst case coded system is 1.1 dB worse than the baseline performance, as shown in Figure 3.5. This is the short constraint length code with hard decisions. Increasing the constraint length yields only 0.5 dB improvement but soft decisions provide larger gains. The short constraint length code with soft decisions gives 1.8 dB improvement over the baseline system, which is 2.9 dB better than the hard

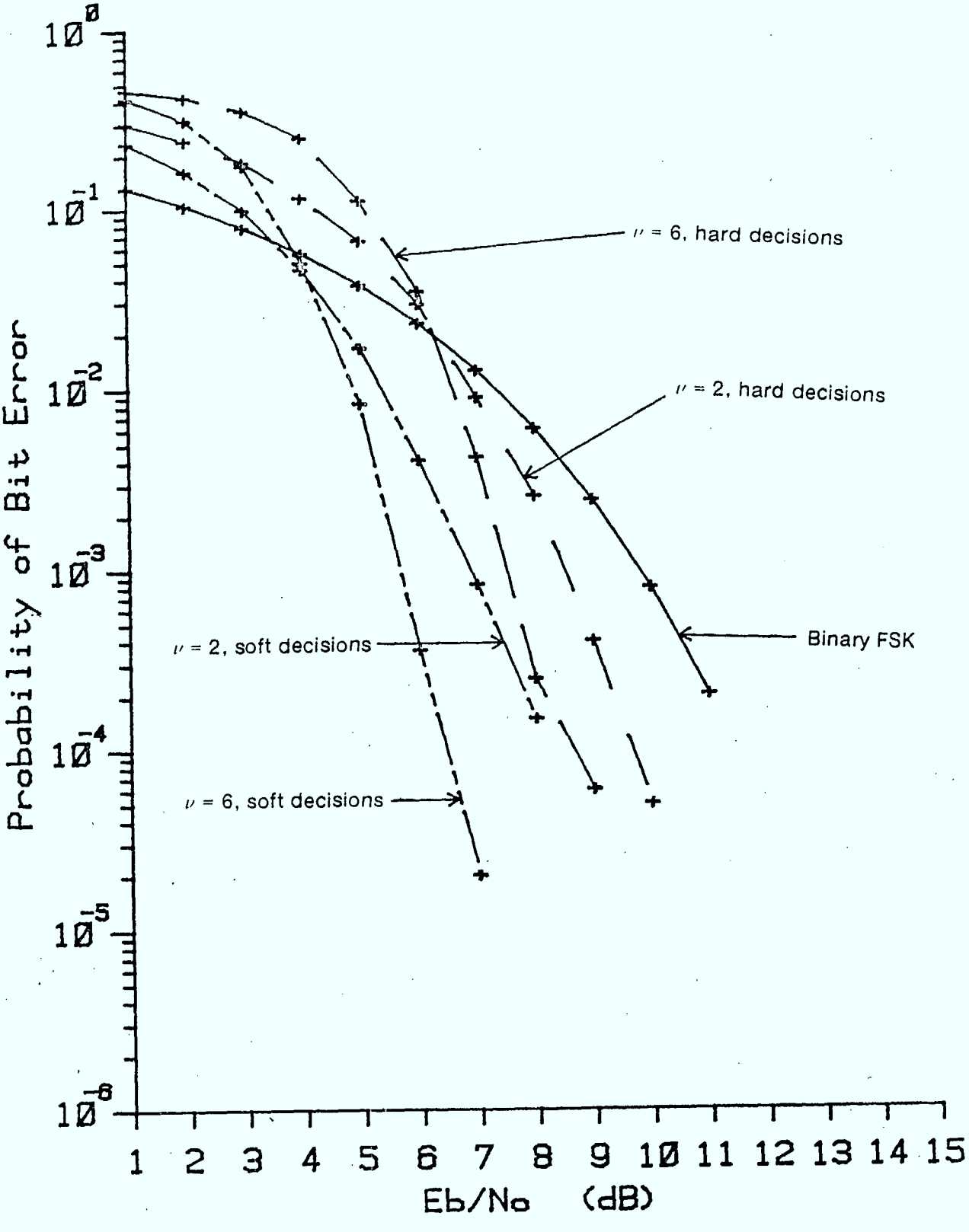


Figure 3.1 Case 1A Simulation Results, Coherent Reception

Rate = 1/2, 2 tones, Variable Signal Energy

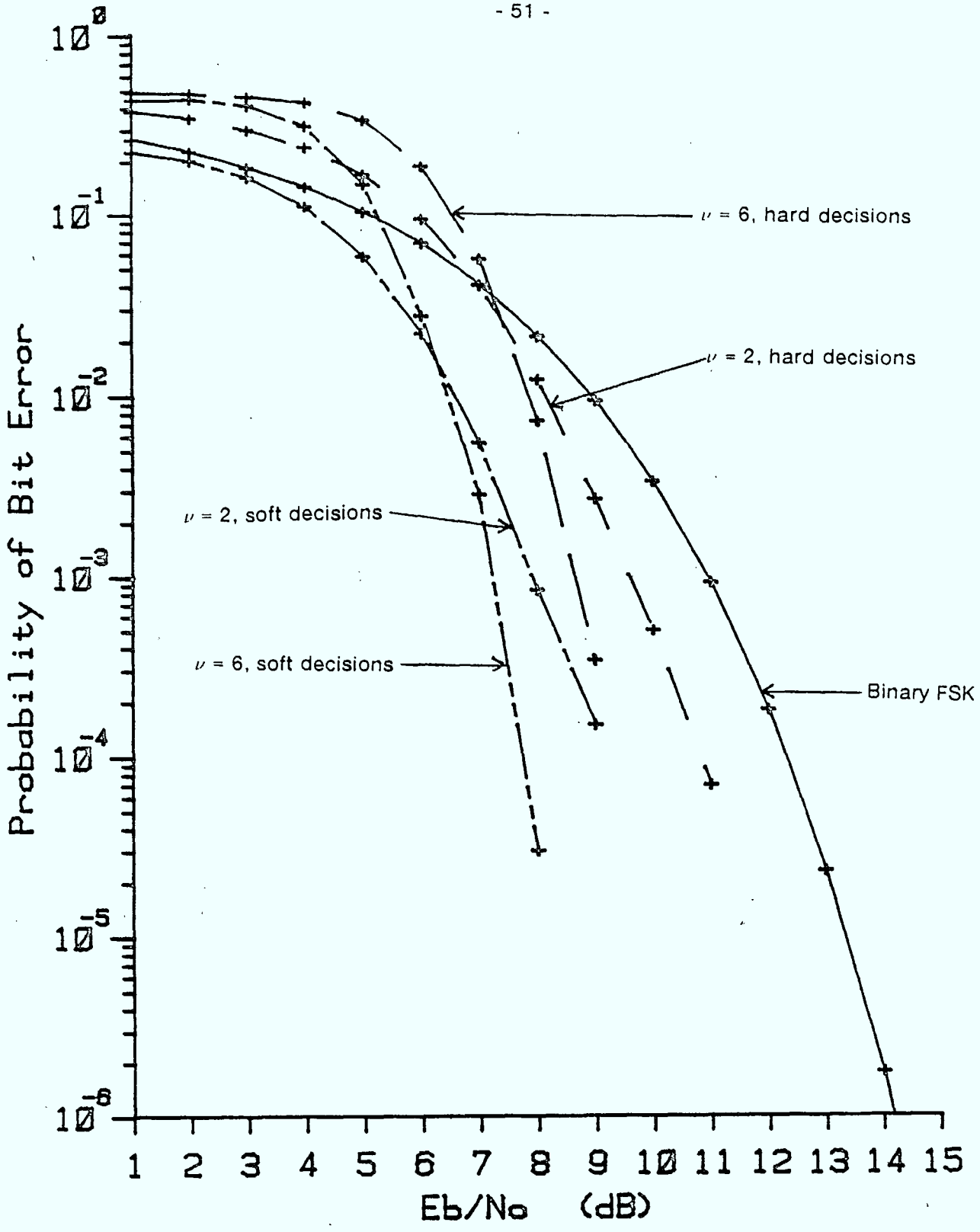


Figure 3.2 Case 1A Simulation Results, Noncoherent Reception

Rate = 1/2, 2 tones, Variable Signal Energy

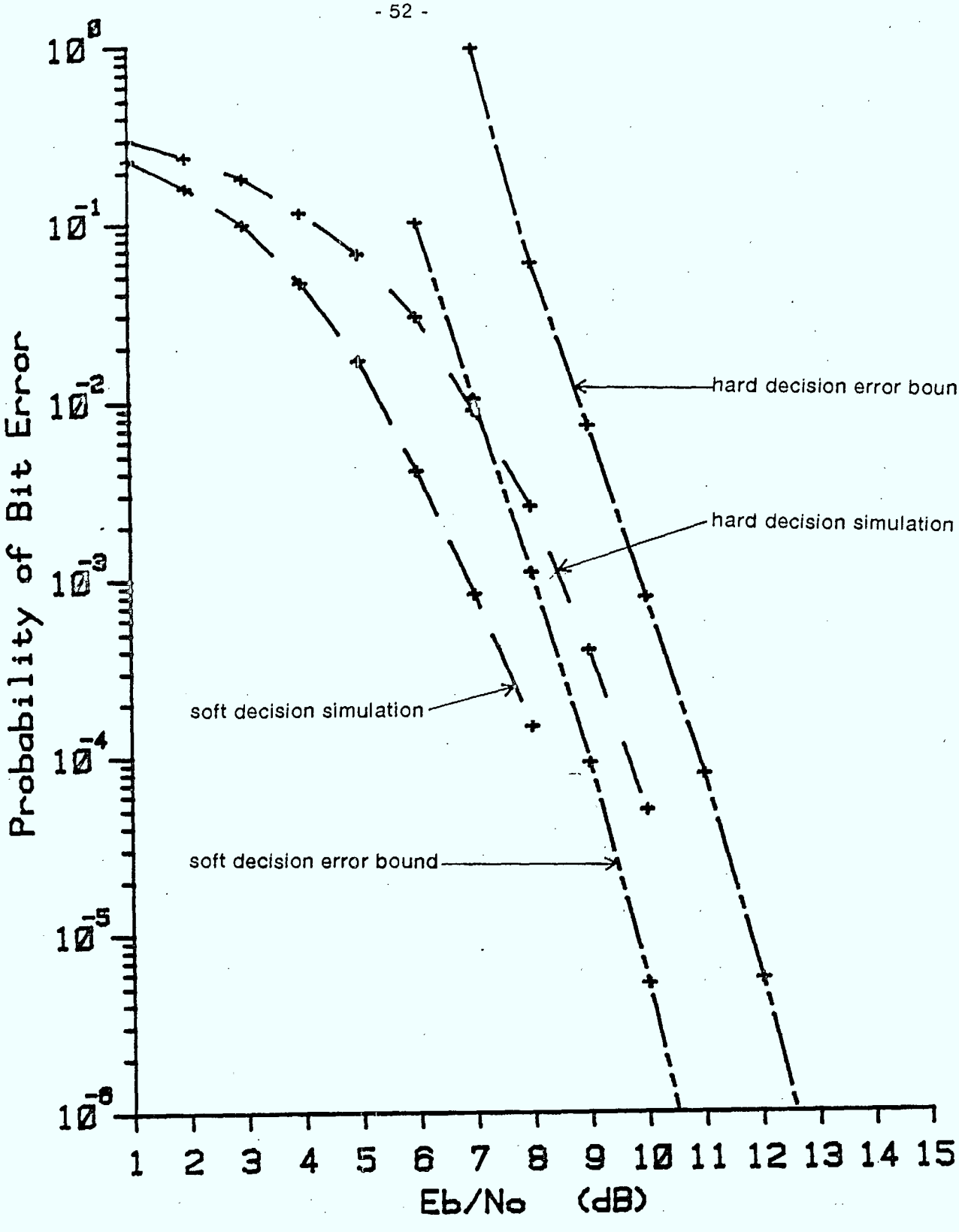


Figure 3.3 Case 1A Code Error Bounds (ν = 2), Coherent Reception

Rate = 1/2, 2 tones, Variable Signal Energy

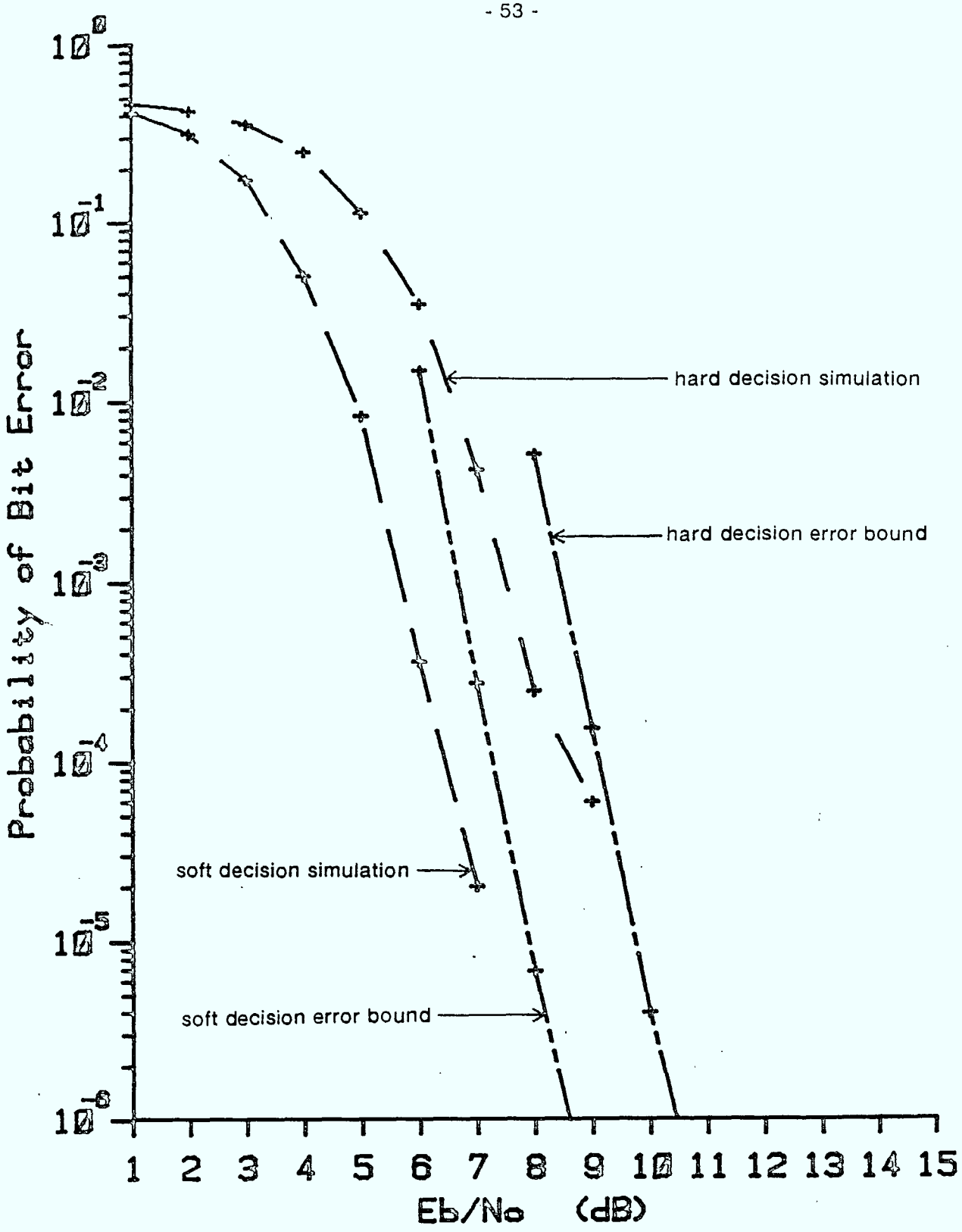


Figure 3.4 Case 1A Code Error Bounds (ν = 6), Coherent Reception

Rate = 1/2, 2 tones, Variable Signal Energy

decision case. The longer constraint length again provides a very modest additional gain of 0.3 dB.

The noncoherent system has performance curves very similar to the coherent case, and they are shown in Figure 3.6. The hard decision simulations are worse than the baseline by 0.5 and 0.1 dB for the short and long constraint length codes respectively. Using soft decisions yields 2.3 to 2.4 dB of improvement over hard decisions. This results in 1.8 dB improvement over the baseline for the shorter code, and 2.3 dB gain for the longer code.

In all situations, the coded system performance curves are steeper than that of the baseline. This implies that the hard decision performance will approach or even surpass the baseline system at higher SNR levels. As well, the soft decision improvement will increase as the curves diverge. This additional improvement will not be very large, probably less than one dB. The difference in slope of the curves also means that the baseline system surpasses the performance of the soft decision coded system at low SNR values. This happens below the point where the curves intersect, known as the cross-over point. The cross-over point occurs at E_b/N_0 of 7.8 dB for coherent reception and 8.5 dB in the noncoherent case.

3.2.3 Case 2A: Rate = 2/3, 3 tones, Variable Signal Energy

This is the first case which has an information rate of two bits per signalling interval, corresponding to a baseline system of 4-ary FSK. A rate 2/3 code is used, which requires an eight point signal space. Three tones are employed, with all combinations of tones allowed, to provide eight variable energy signals. The two codes used in the

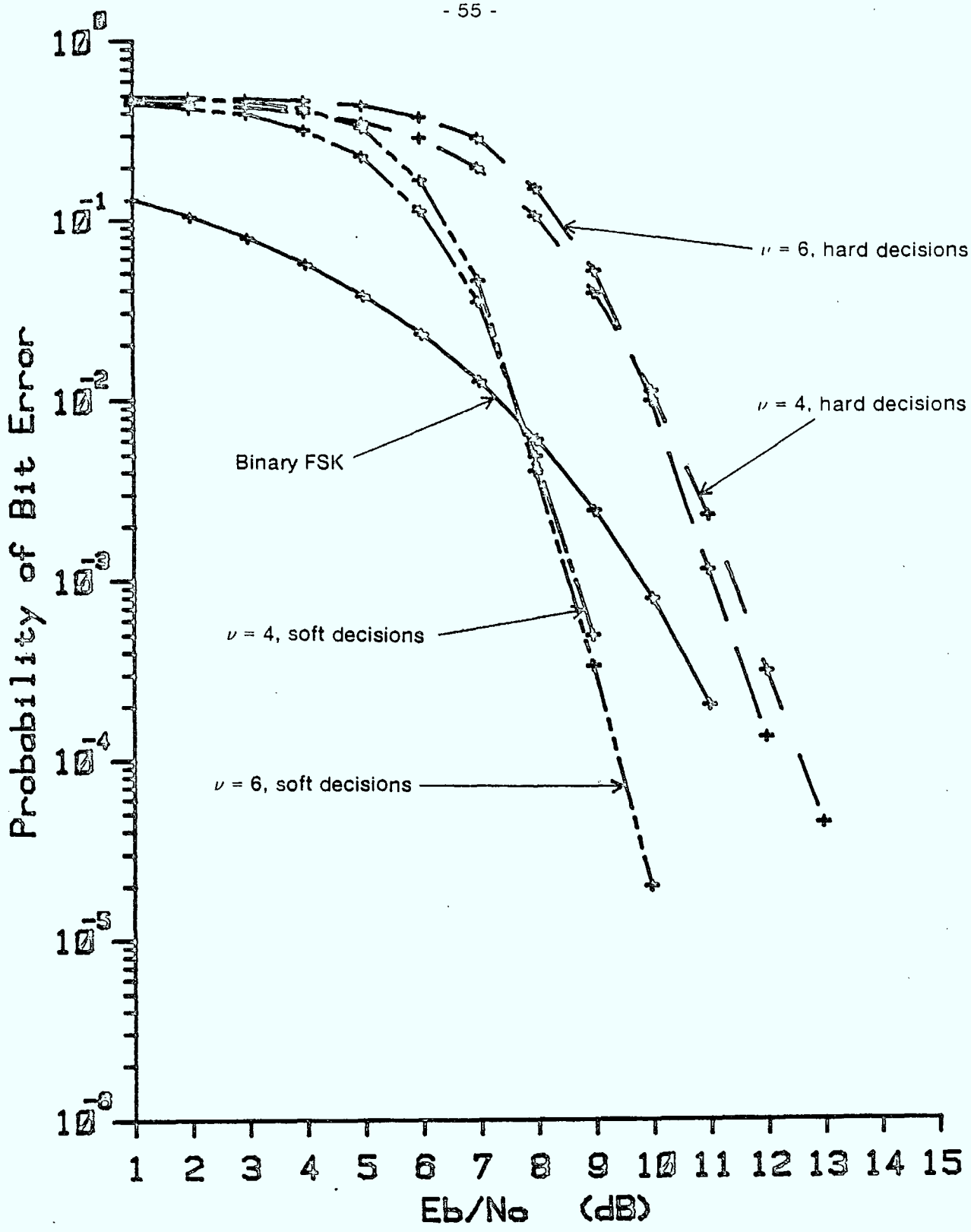


Figure 3.5 Case 1B Simulation Results, Coherent Reception

Rate = 2/3, 2 tones, Constant Signal Energy (signalling over 2T)

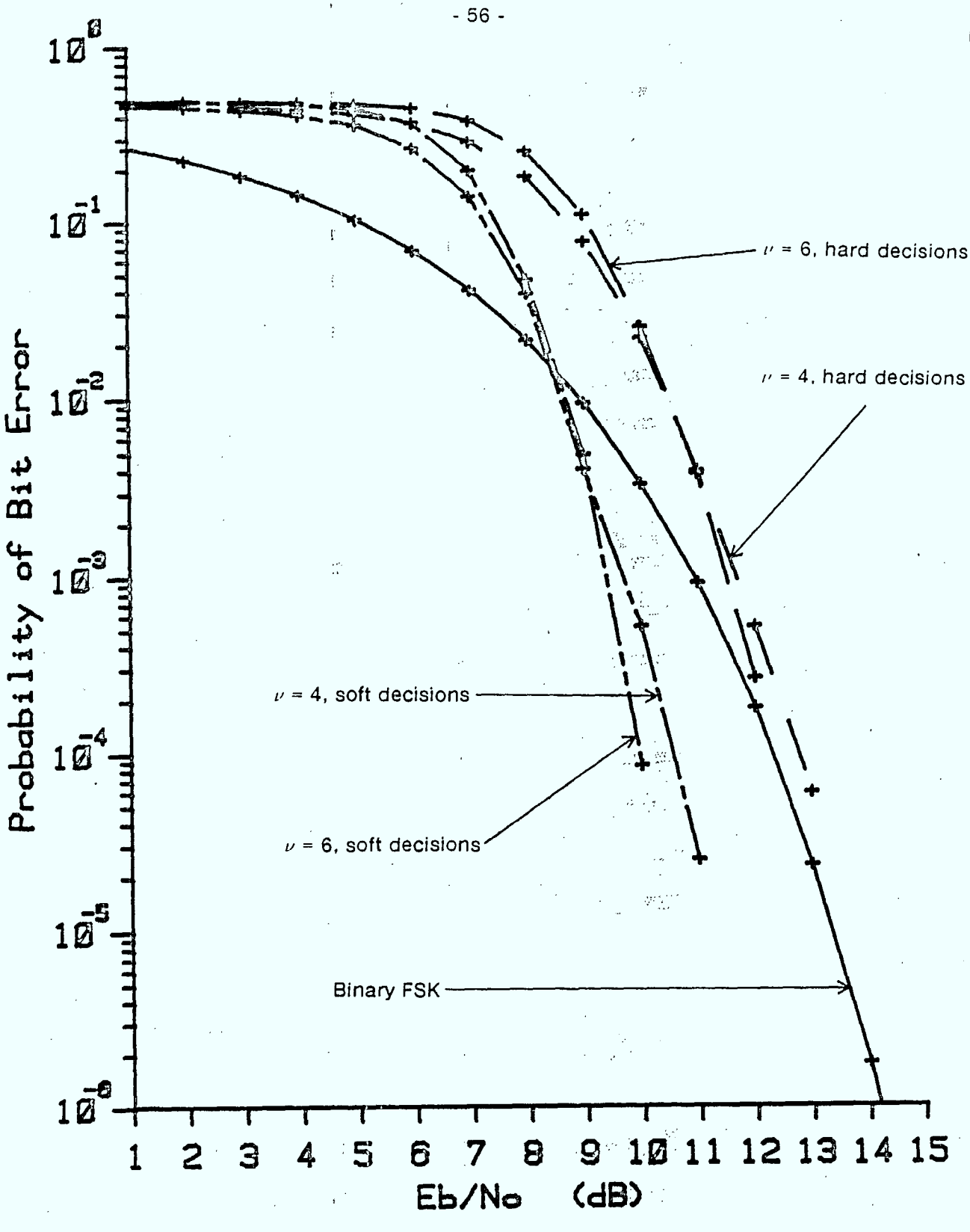


Figure 3.6 Case 1B Simulation Results, Noncoherent Reception

Rate = 2/3, 2 tones, Constant Signal Energy (signalling over 2T)

simulation are the same as those in the previous case, with constraint lengths of 4 and 6.

The results for coherent reception are shown in Figure 3.7. The performance of both codes is virtually identical with hard decision decoding, and is about 0.5 dB worse than the baseline system. Soft decision decoding yields 1.3 dB improvement over the reference system for the short constraint length code, and an additional 0.5 dB for the longer code. The cross-over points for the soft decision curves are between 5.5 and 6.0 dB (E_b/N_0).

Figure 3.8 displays the performance curves for noncoherent reception of Case 2A. The short constraint length code with hard decisions is 0.5 dB worse than the baseline, while the longer code is 0.2 dB better than the reference system. The use of soft decisions yields about 2 dB of improvement over the hard decision case. The performance gain over the baseline with the soft decision decoder is 1.5 dB for the shorter code, and 2.1 dB for the longer constraint length code. Both soft decision curves intersect the baseline curve at 6.3 dB on the horizontal axis.

This case achieves performance improvements with a reduction in required bandwidth. The spacing between orthogonal tones is the same for the three tone coded system and the four tone baseline system. The coded signals therefore require only 3/4 of the bandwidth, although a higher peak power transmitter is neccessary for the multiple tone signals. It is also significant to note the free distances of the optimal rate 2/3 codes. The shorter code ($\nu = 4$) has a free distance of 5, while the longer code ($\nu = 6$) has a free distance of 7. These parameters will be used for comparison between the different cases at this information rate.

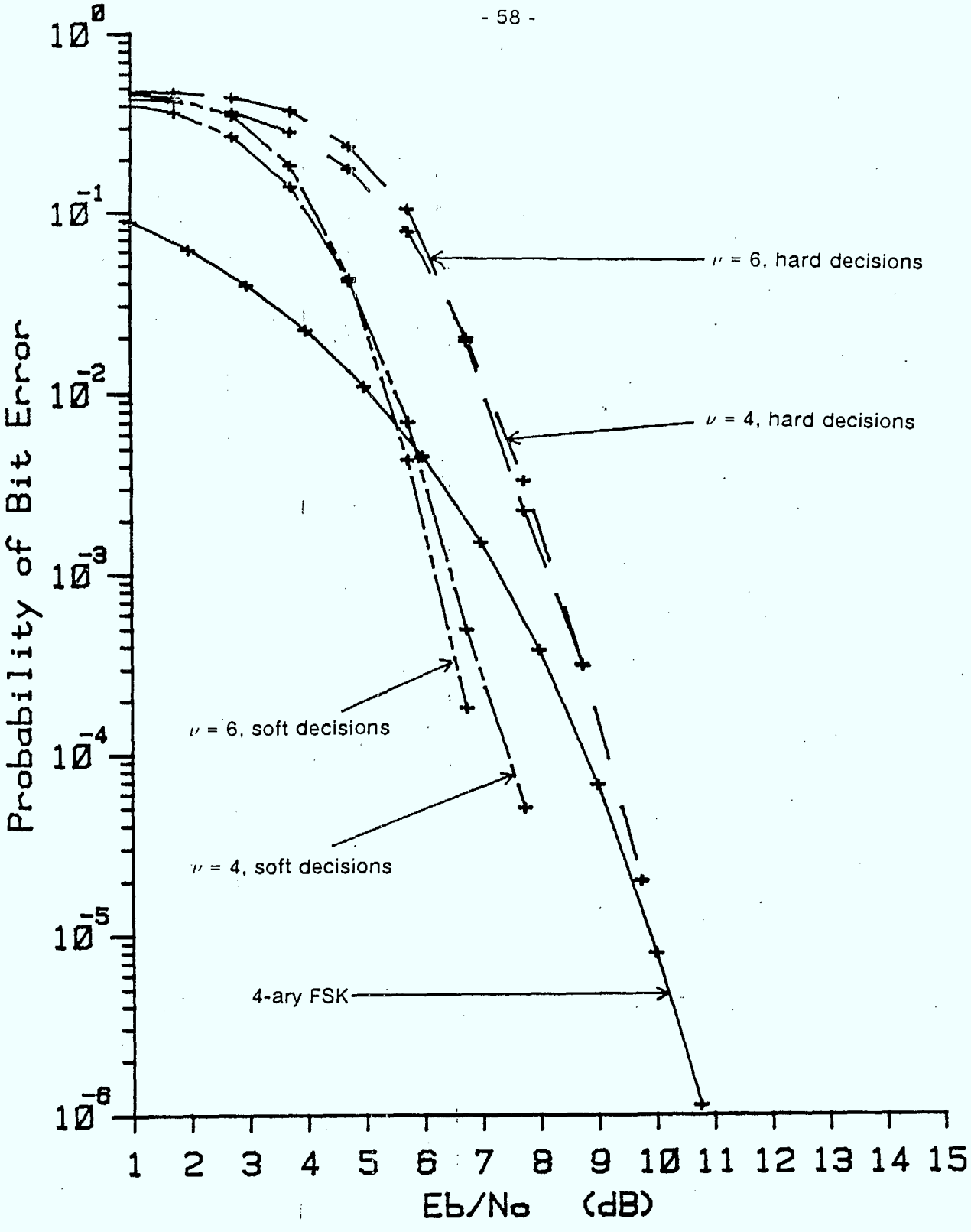


Figure 3.7 Case 2A Simulation Results, Coherent Reception

Rate = 2/3, 3 tones, Variable Signal Energy

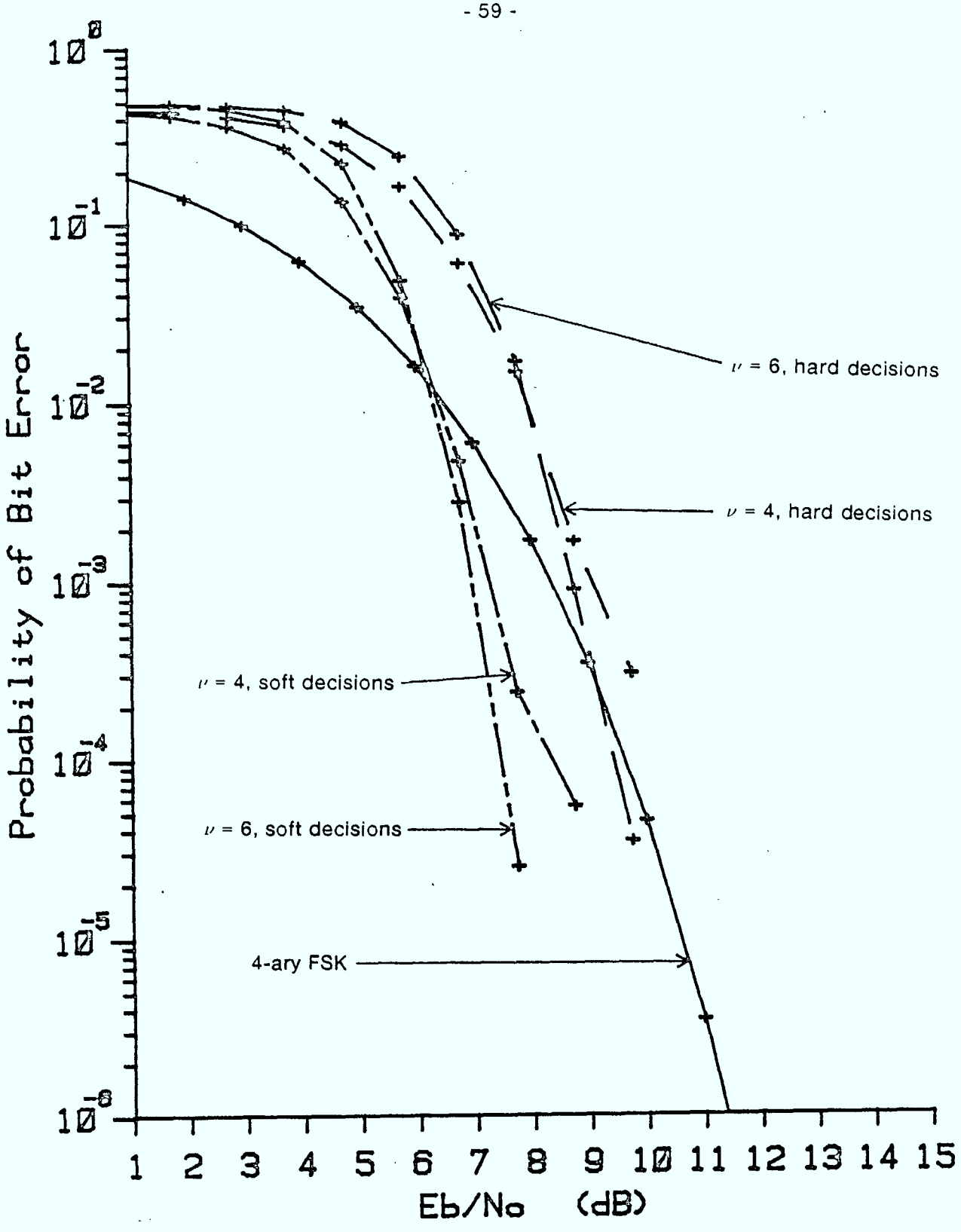


Figure 3.8 Case 2A Simulation Results, Noncoherent Reception

Rate = 2/3, 3 tones, Variable Signal Energy

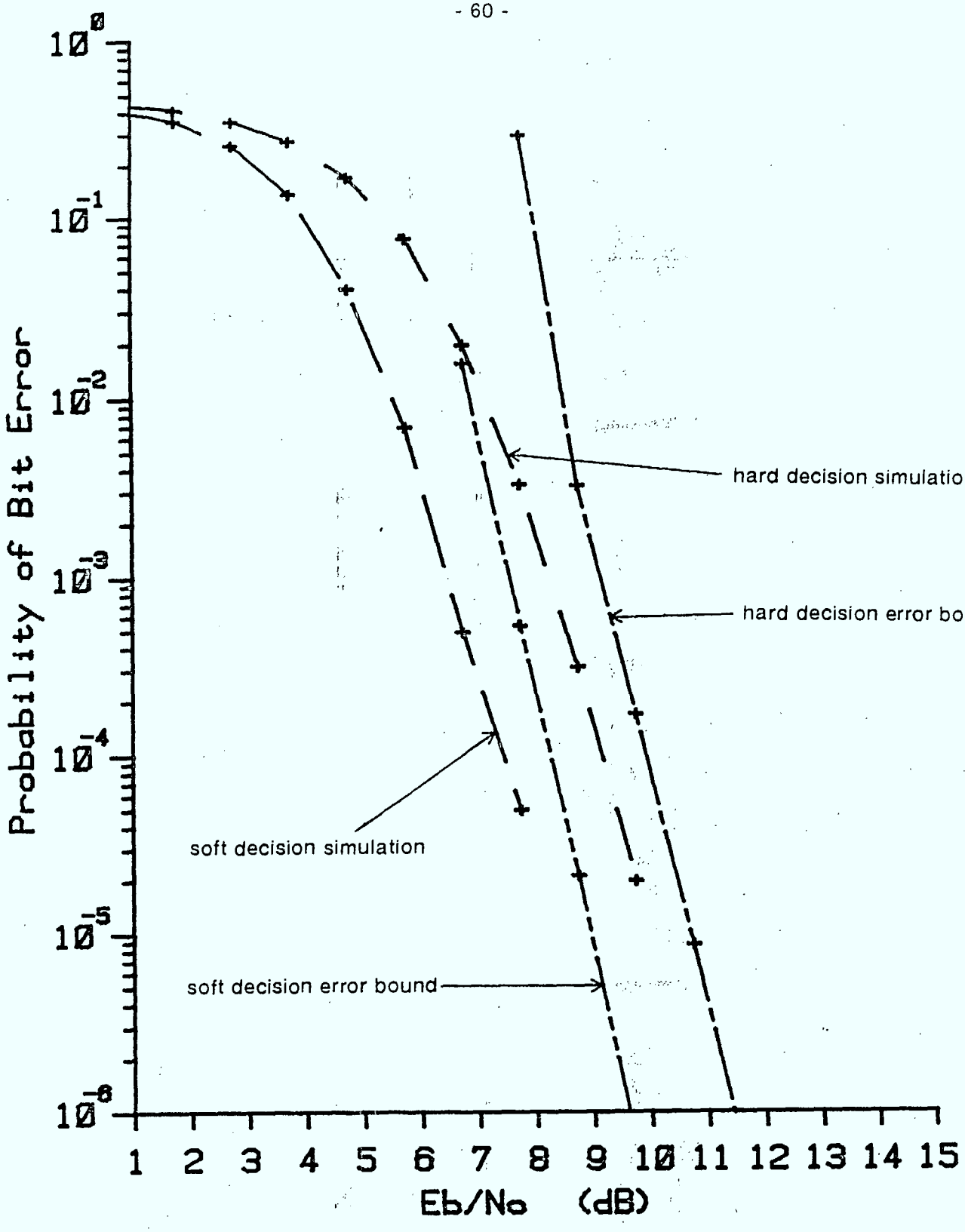


Figure 3.9 Case 2A Code Error Bounds (ν = 4), Coherent Reception

Rate = 2/3, 3 tones, Variable Signal Energy

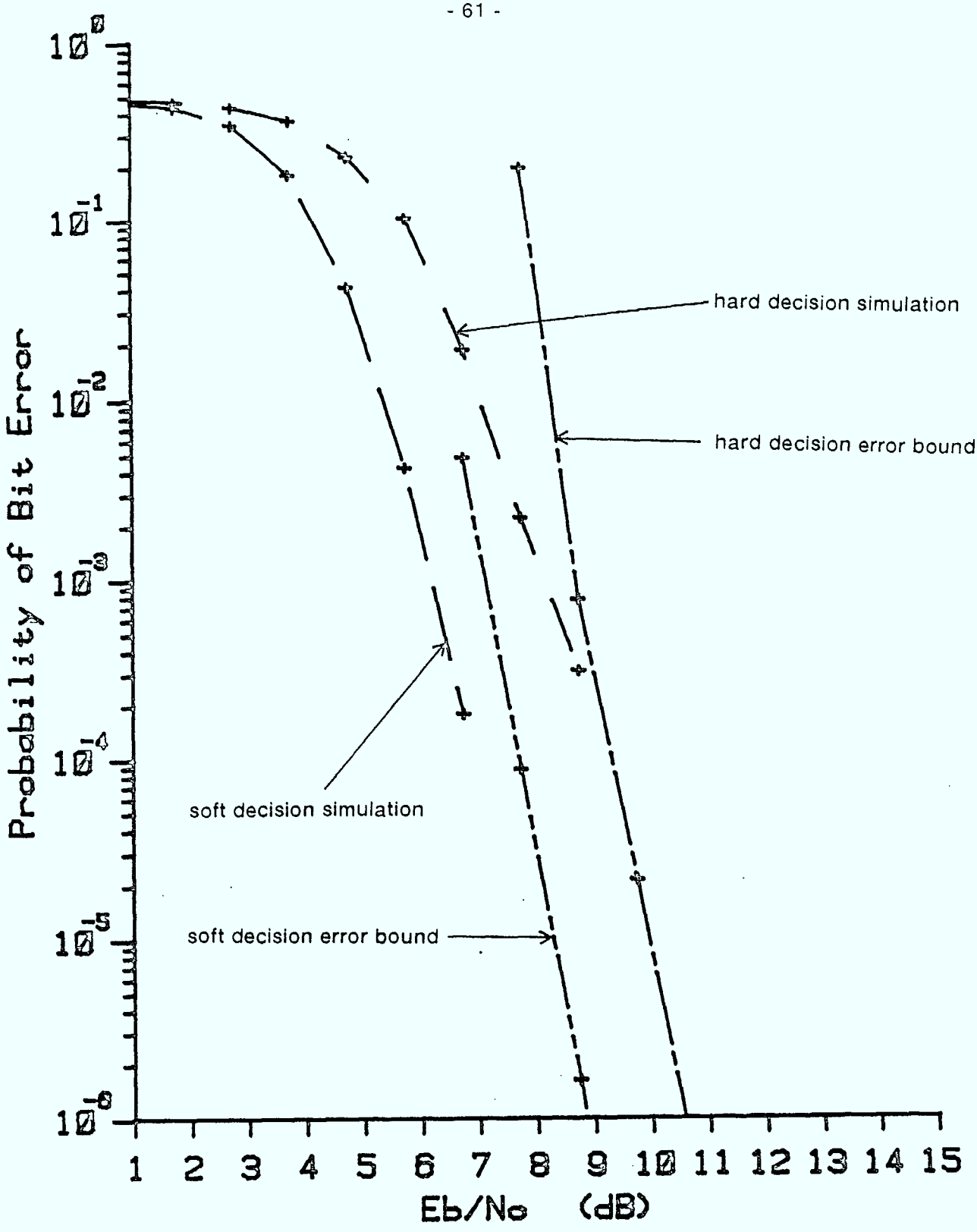


Figure 3.10 Case 2A Code Error Bounds (ν = 6), Coherent Reception

Rate = 2/3, 3 tones, Variable Signal Energy

There are two figures which display the theoretical error bounds for coded performance with rectangular signal constellations, as described in Chapter two. Figure 3.9 shows the bound for the shorter code and Figure 3.10 contains the results for the longer constraint length code.

3.2.4 Case 2B: Rate = 2/4, 4 tones, Variable Signal Energy

This case employs all four tones in the coded system that are used in the baseline FSK modulation. All tone combinations are used to provide sixteen signals with signal energies that depend on the number of constituent tones. The code for this case has rate 2/4 to provide a four bit codeword for two bits of information in each signalling interval. An optimum distance profile code of this rate could not be found in the literature, and so two identical rate 1/2 encoders in parallel were used. Only one code was simulated, with a constraint length of 6, composed of two rate 1/2 codes of constraint length 3.

The performance of the coherent reception system is shown in Figure 3.11. With hard decisions, the coded simulation is 1.1 dB worse than the baseline. Soft decisions yield a 2.5 dB gain to provide 1.4 dB improvement over the reference system. Figure 3.12 displays the performance curves for the noncoherent receiver. The soft decision decoder is superior to the baseline system by 1.6 dB while the hard decision curve is 0.5 dB worse than 4-ary FSK.

The performance of the coded system in this case is virtually the same as the previous case in spite of the lower rate code and increased signal bandwidth. The reason for this is that the code structure employed does not take full advantage of the signal space. The code formed by two cascaded rate 1/2 encoders has double the constraint length of a single component code yet only achieves the same free distance.

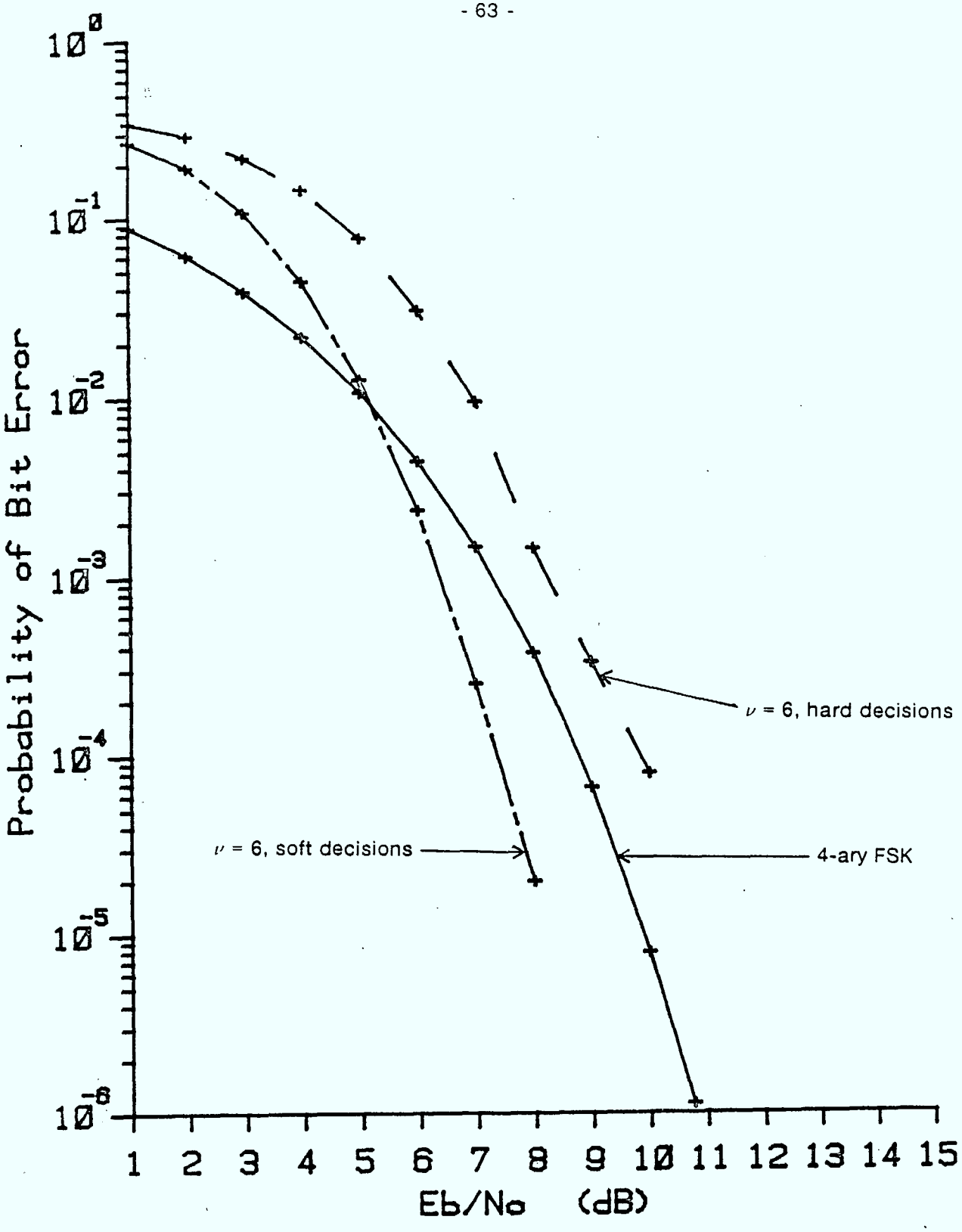


Figure 3.11 Case 2B Simulation Results, Coherent Reception

Rate = 2/4, 4 tones, Variable Signal Energy

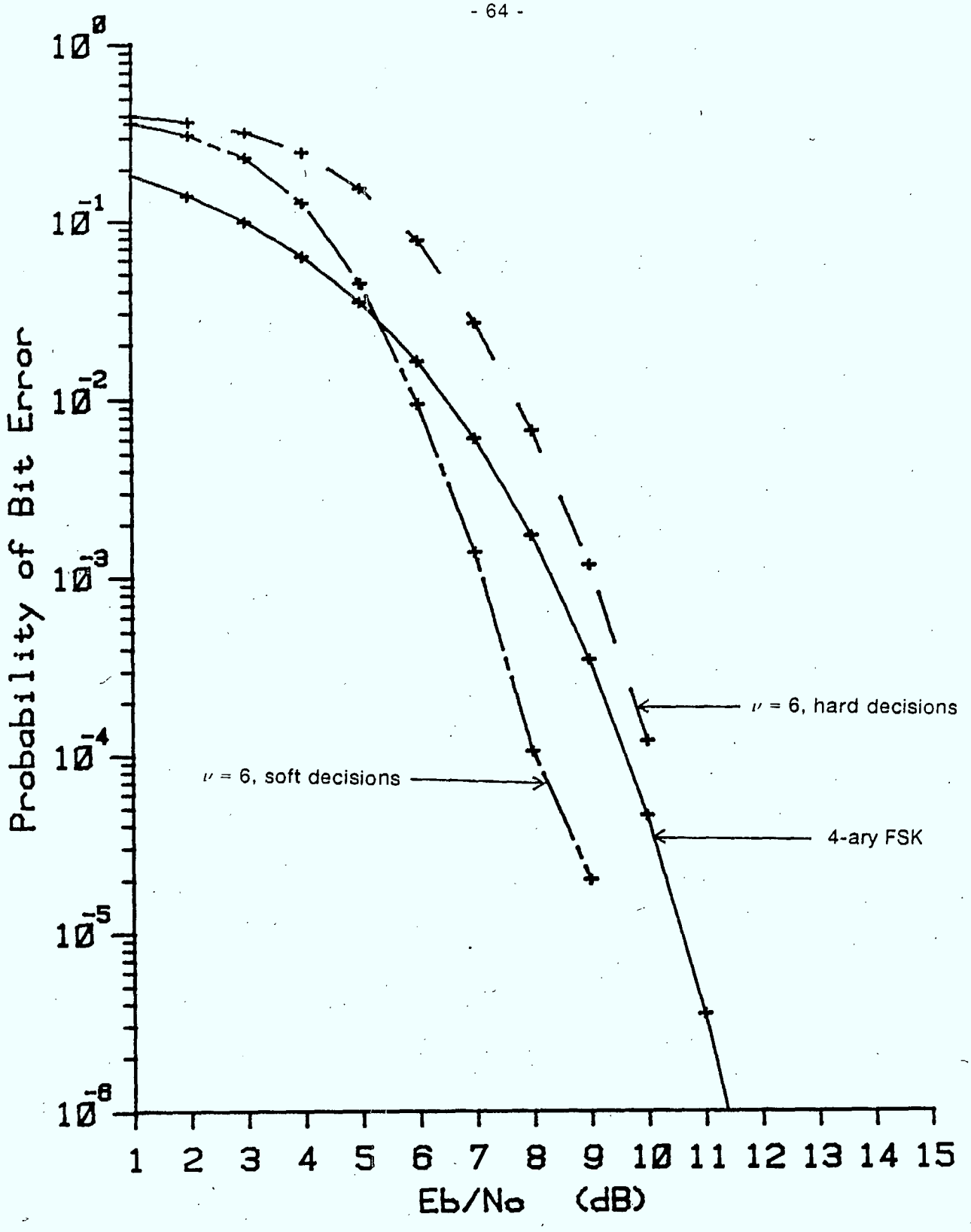


Figure 3.12 Case 2B Simulation Results, Noncoherent Reception

Rate = 2/4, 4 tones, Variable Signal Energy

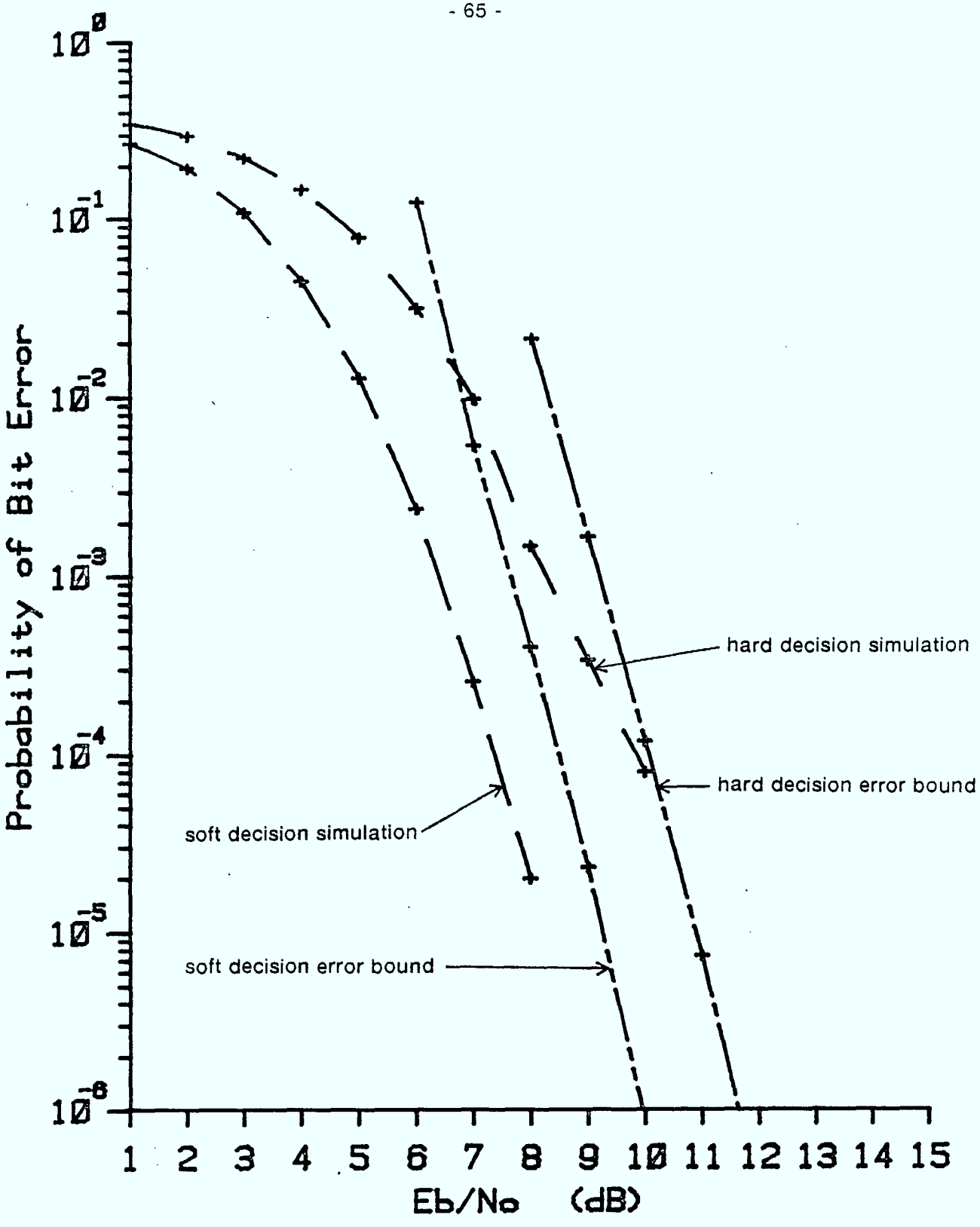


Figure 3.13 Case 2B Code Error Bounds (ν = 6), Coherent Reception

Rate = 2/4, 4 tones, Variable Signal Energy

The code used in this simulation case has a free distance of 6 which is approximately the same as the rate 2/3 codes in Case 2A. Therefore there is no appreciable performance gain for this system. The theoretical error bound for this code is shown in Figure 3.13.

3.2.5 Case 2C: Rate = 2/3, 4 tones, Constant Signal Energy

This simulation employs the same rate 2/3 convolutional codes as Case 2A. In this case, the eight signals required are constructed from four orthogonal tones. Single and double tone signals with constant signal energy are used. The codewords are mapped onto the signals to optimize the error performance as described in Chapter two.

The simulation results for coherent detection are plotted in Figure 3.14. The hard decision decoder performs worse than the baseline by 0.7 dB and 0.9 dB for the short and long constraint length codes respectively. Soft decisions give improved performance with a margin of 1.5 dB for the short constraint length code over the baseline. The longer code provides an additional gain of 0.5 dB for a 2.0 dB improvement on the reference system.

The noncoherent performance curves appear in Figure 3.15. The baseline system is again superior to hard decision decoding for both codes. The short constraint length code has a 0.4 dB disadvantage while the longer code has 0.3 dB worse performance. With soft decision decoding, the short constraint length coded system improves by 2.0 dB. The longer code shows 3.0 dB gain over the hard decision case for a 2.7 dB improvement over the reference system.

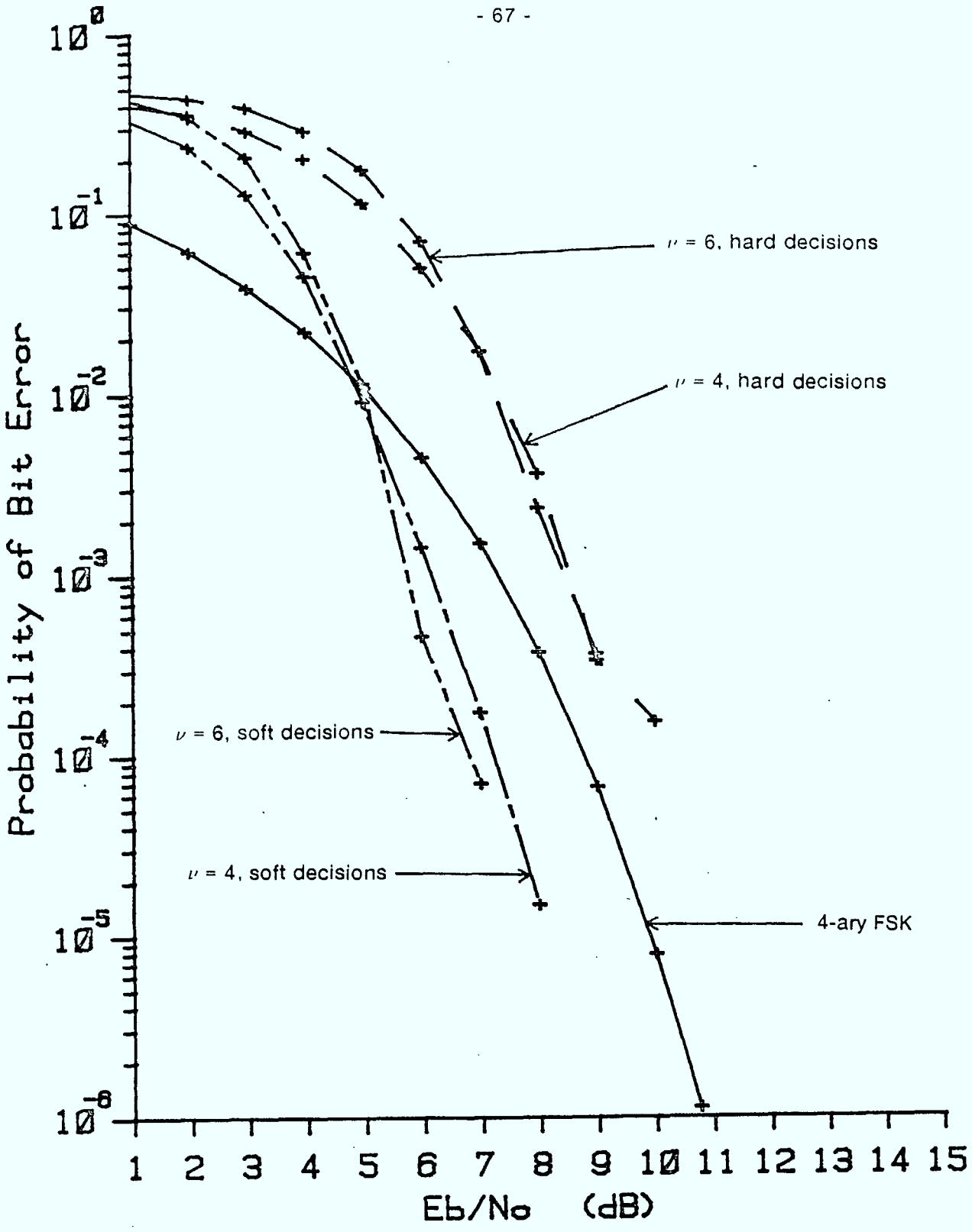


Figure 3.14 Case 2C Simulation Results, Coherent Reception

Rate = 2/3, 4 tones, Constant Signal Energy

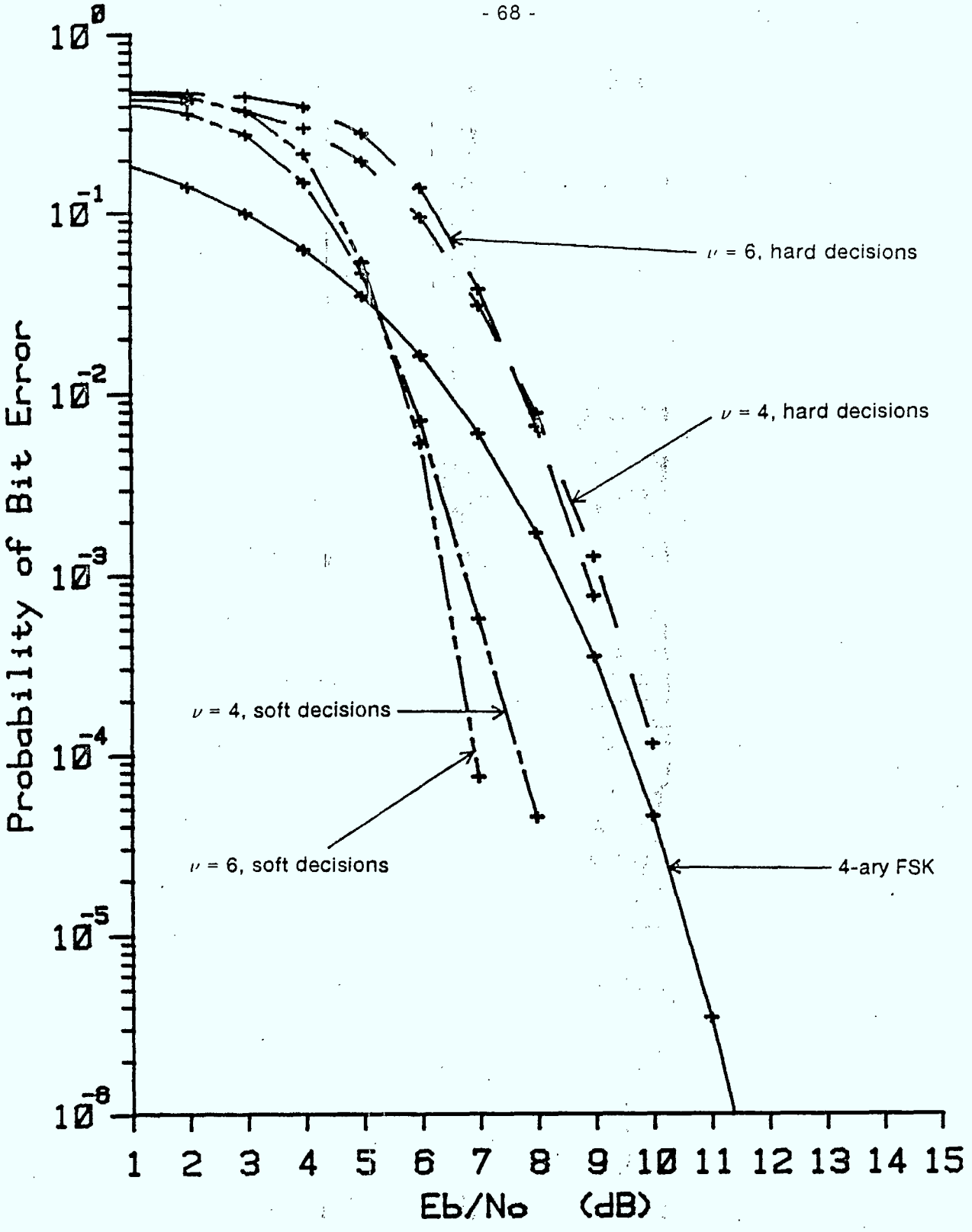


Figure 3.15 Case 2C Simulation Results, Noncoherent Reception

Rate = 2/3, 4 tones, Constant Signal Energy

This case has the best performance of the three simulations at this information rate. The additional advantage of constant energy signals imposes less stringent requirements on the actual transmitter and receiver.

3.2.6 Case 3A: Rate = 3/4, 4 tones, Variable Signal Energy

The final two cases have an information rate of three bits per signalling interval, which corresponds to a baseline system of 8-ary FSK. Case 3A employs a rate 3/4 code and a signal space of sixteen points. The signals are comprised of all possible combinations of four orthogonal tones, which require only one half the bandwidth of the baseline system. The simulation was run with one code which has a constraint length of 5 and a free distance of 5.

Figure 3.16 shows the results of the simulation for coherent reception. With hard decisions, the coded system is 1.9 dB worse than the baseline. Soft decisions yield a 2.1 dB improvement, resulting in a 0.2 dB advantage over FSK signalling. The performance of the noncoherent receiver appears in Figure 3.17. The hard decoder exceeds the reference system performance by 1.4 dB while soft decisions provide a 1.9 dB gain. This results in a modest 0.5 dB improvement over the baseline for the coded system with soft decision decoding.

This case does not provide significant performance gain over 8-ary FSK. The relatively high rate code provides a free distance of 5 and the signalling bandwidth is reduced by one half from the reference system. Figure 3.18 shows the error performance bounds for this code with a rectangular signal constellation. The signals also have variable energy, requiring increased receiver complexity.

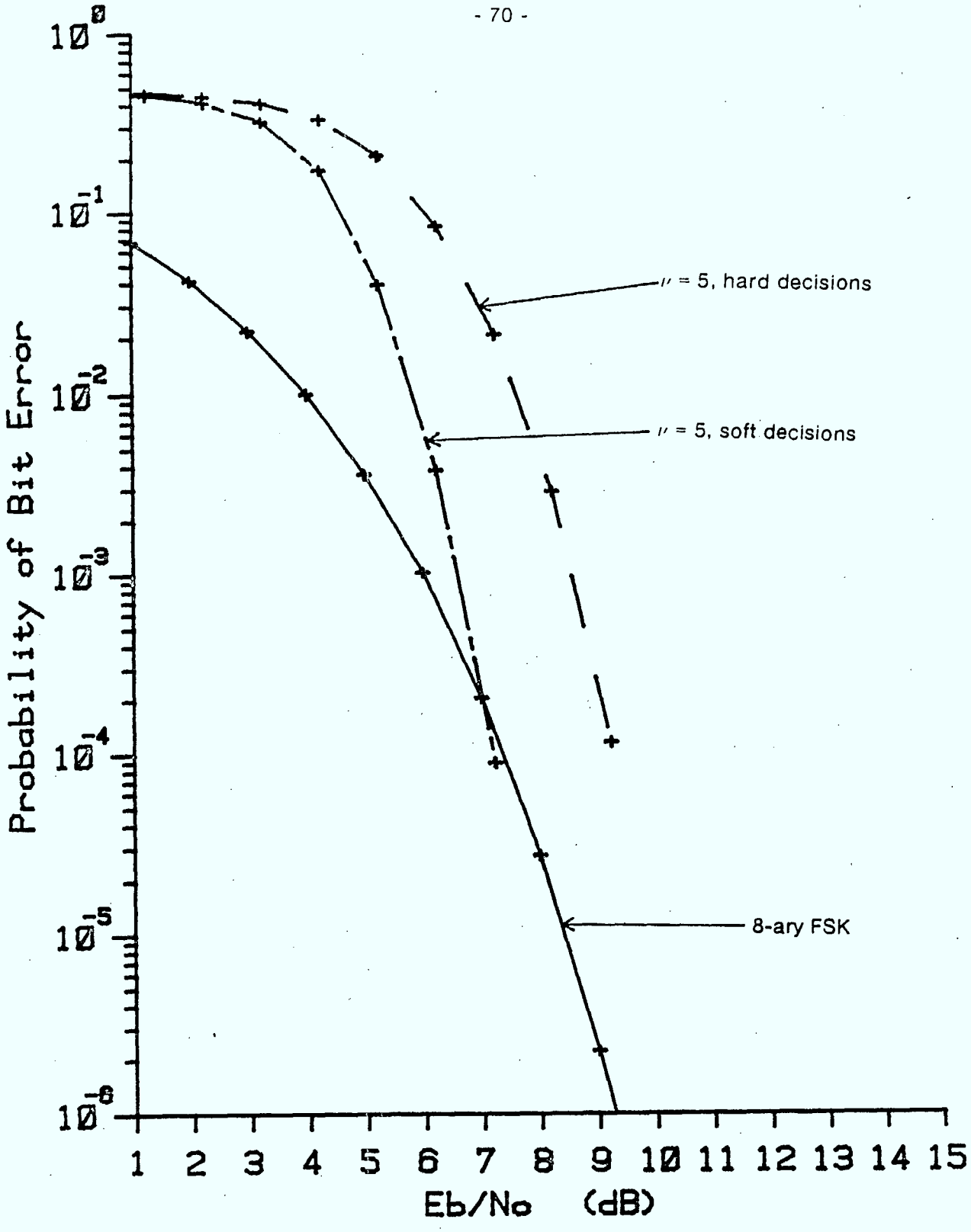


Figure 3.16 Case 3A Simulation Results, Coherent Reception

Rate = 3/4, 4 tones, Variable Signal Energy

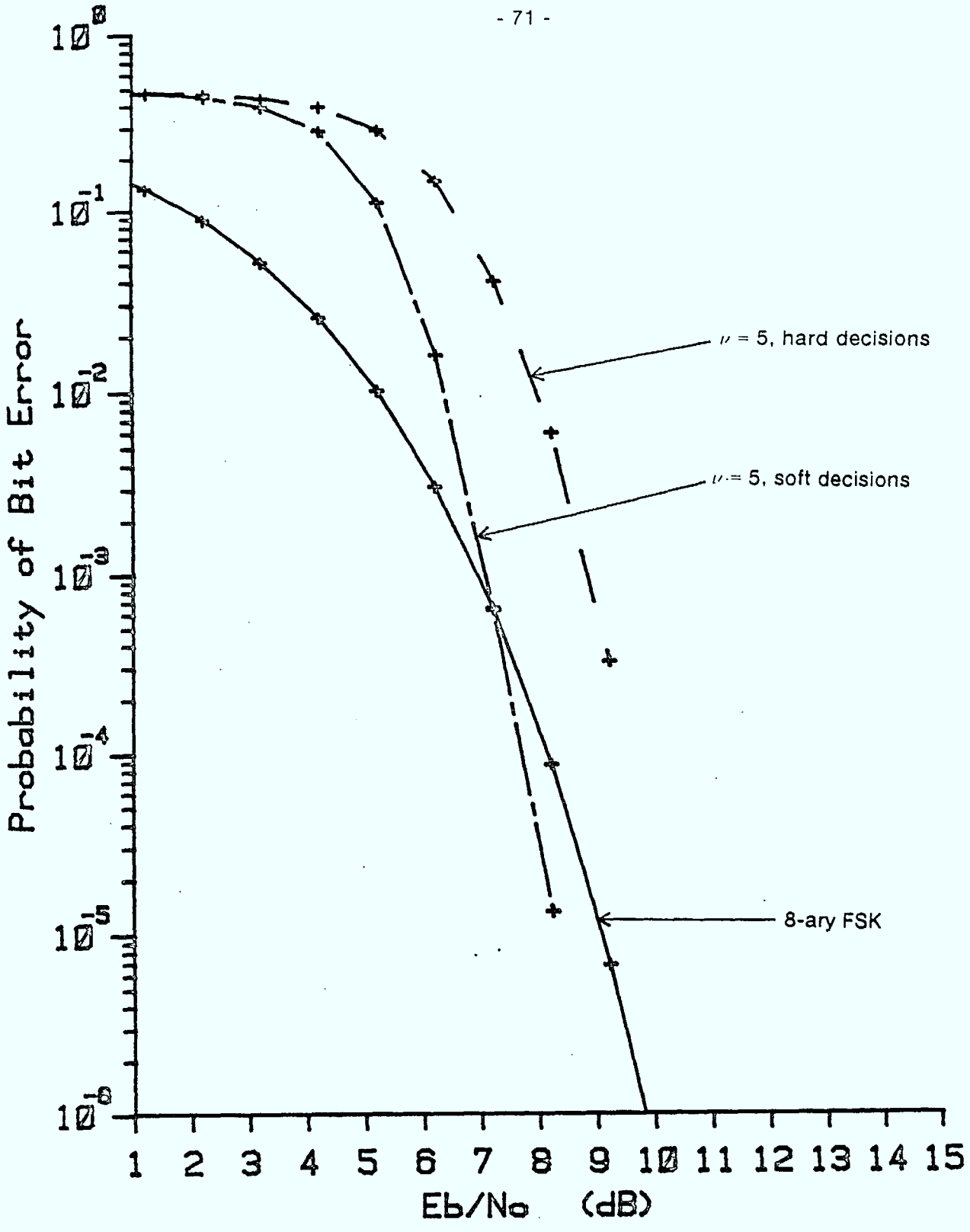


Figure 3.17 Case 3A Simulation Results, Noncoherent Reception

Rate = 3/4, 4 tones, Variable Signal Energy

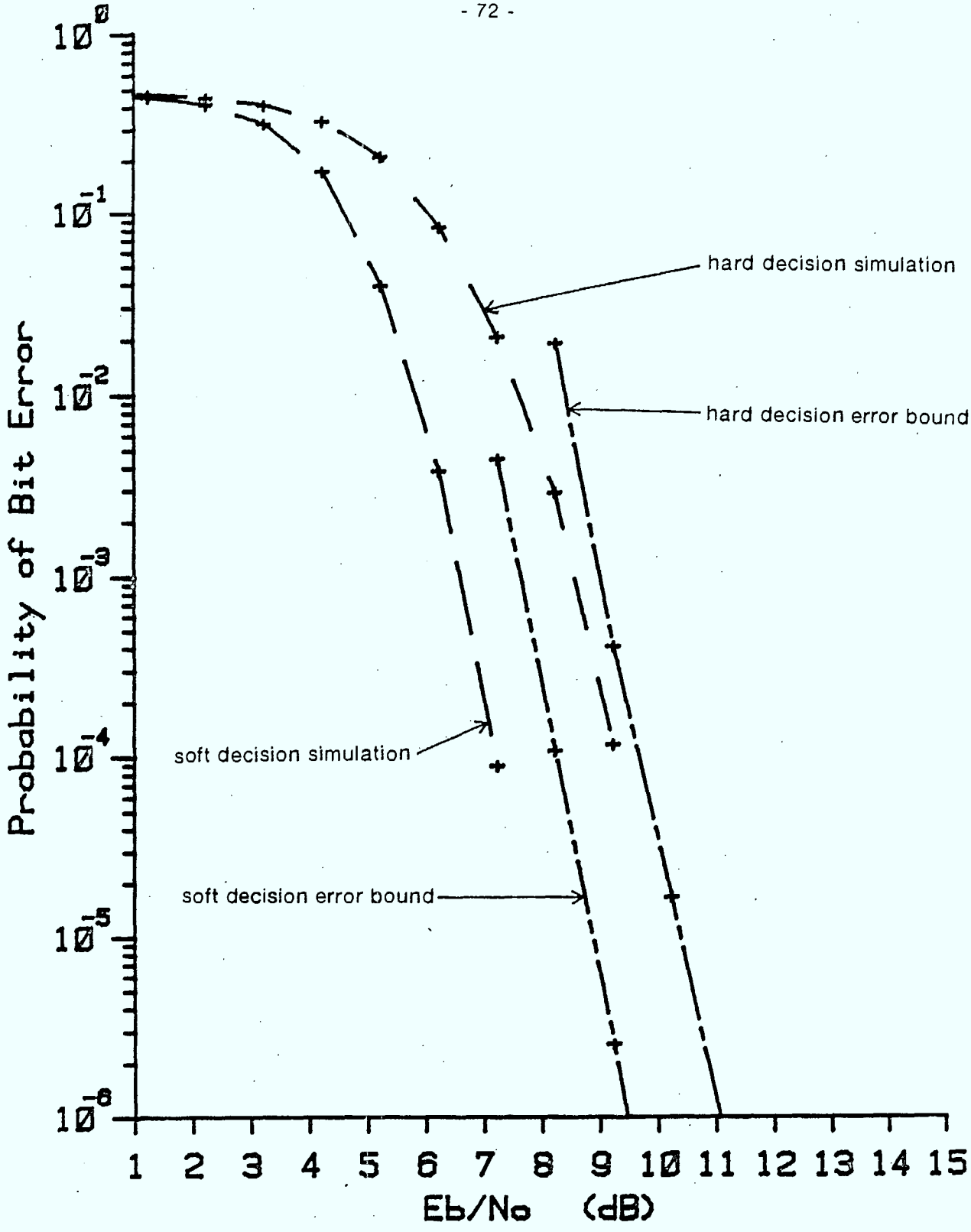


Figure 3.18 Case 3A Code Error Bounds (ν = 5), Coherent Reception

Rate = 3/4, 4 tones, Variable Signal Energy

3.2.7 *Case 3B: Rate = 3/6, 8 tones, Constant Signal Energy*

The codes for this case have a rate of 3/6 and constraint lengths of 3 and 6. The short code is a dual-3 rate 1/2 convolutional code, and the longer code is comprised of three identical cascaded rate 1/2 encoders, each with $\nu = 2$. The sixty-four signals required are composed of single tones and triples of tones selected from a set of eight orthogonal tones. The signals have constant energy and are assigned to the codewords as previously described.

The performance curves for coherent reception are shown in Figure 3.19. The coded simulations with hard decision decoding are significantly worse than the baseline system. The margins are 4.3 dB for the short code and 2.5 dB for the longer code. The performance gains for soft decision decoding are also quite large so that the short constraint length code is 0.3 dB worse than the baseline with soft decison decoding. A gain of 2.4 dB over the reference system is achieved by the long code when soft decoder metrics are employed.

The results for noncoherent reception are quite similar to the coherent case as shown in Figure 3.20. The short code with hard decisions is 3.8 dB worse than the baseline, while the long code has a 1.8 dB disadvantage. The use of soft decision decoding provides the short code with a 0.4 dB improvement on the baseline performance. The long code outperforms the baseline by 2.4 dB with soft decisions.

The long constraint length code with soft decision decoding gives good performance in this case. The relatively poor results of the other simulations are due mainly to two factors. The structure of the shorter code is such that the codewords cannot be mapped onto the signal set to provide maximum distance between error paths.

This accounts for the large disparity between the different constraint length codes. The codes used do not have large free distances, and the signal space mapping does not provide proportionality between Hamming distance and Euclidean distance. Therefore the hard decision decoding metric has a large disadvantage compared to the soft decision case.

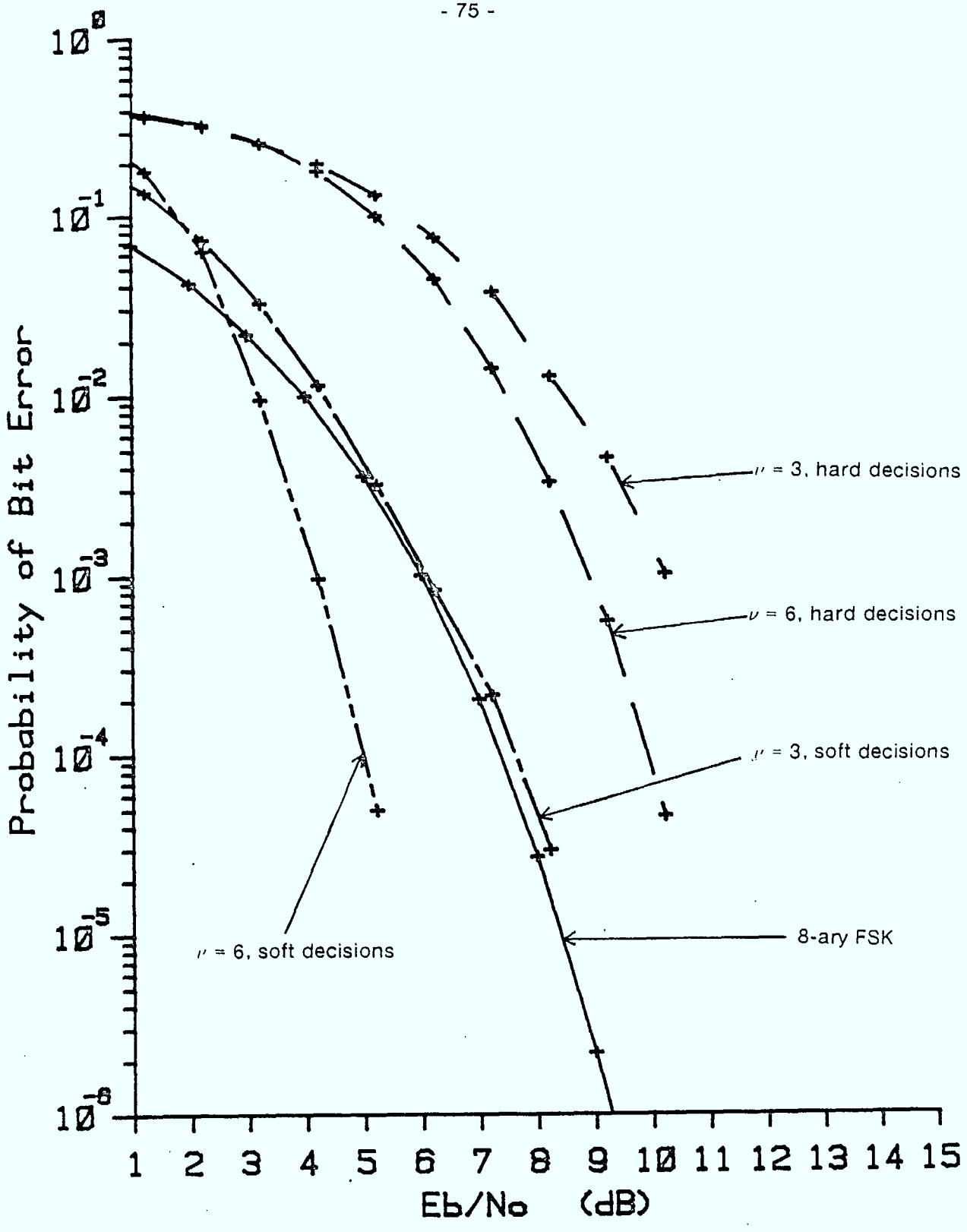


Figure 3.19 Case 3B Simulation Results, Coherent Reception

Rate = 3/6, 8 tones, Constant Signal Energy

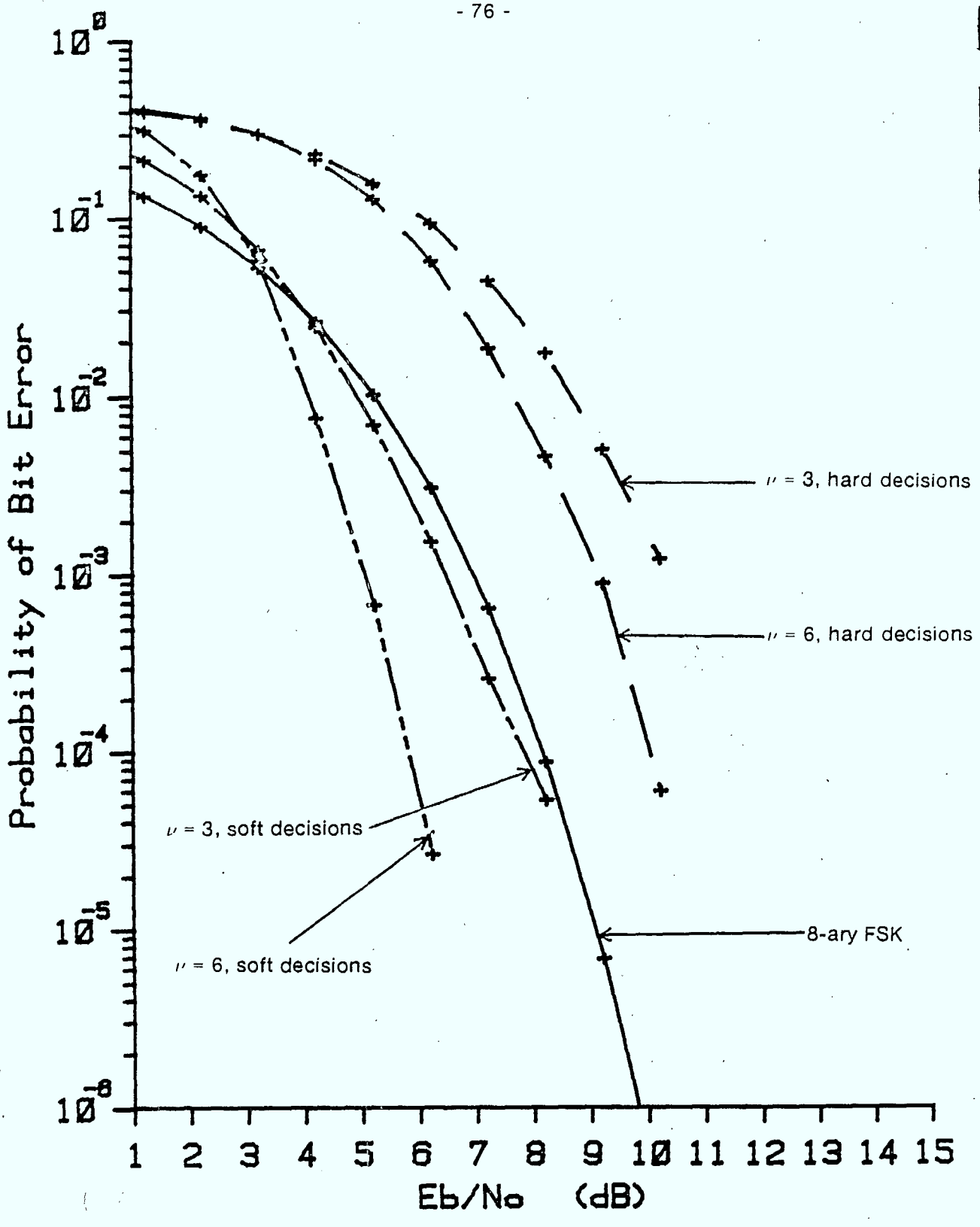


Figure 3.20 Case 3B Simulation Results, Noncoherent Reception

Rate = 3/6, 8 tones, Constant Signal Energy

CHAPTER FOUR

CONCLUSIONS

4.1 Findings of the Study

The major result of this study is quite obvious from the performance curves in the previous chapter. The use of convolutional coding with multiple tone signal sets gives significant improvement over FSK when soft decision decoding is employed. This performance gain is on the order of 2.0 dB (E_b/N_0) at a BER of 10^{-4} for at least one case at each information rate. There are numerous parameters of both the coding and modulation schemes which influence the results of the simulations.

The most readily apparent effect is caused by the type of decoding metric employed. The simulations with soft decision metrics indicate performance superior to those with hard decisions by a margin of from 1.5 to 4.9 dB. This situation would be expected because some information about the received signal is lost when a hard decision is made in the receiver. The difference is greater for schemes which have less correspondence between the Hamming distance of the codewords and the Euclidean distance between the signals, namely the constant signal energy constellations.

The performance of the coherent and noncoherent systems is very similar. In absolute terms, the coherent receiver gives superior performance due to the exact knowledge of the phase of the transmitted signal. However, the relative improvements for the coded systems and the gains for soft decision decoding are virtually identical for both types of receivers. Although noncoherent FSK is usually used in actual systems,

coherent results were presented for completeness. Also, theoretical error performance bounds can be calculated for coherent reception, and they provide verification of the simulation results.

The difference between the constant signal energy constellations and those with variable energy signals is difficult to accurately derive from the cases under consideration. The variable energy case proves superior for the transmission of a single bit per signalling interval. Case 1A shows the largest gains of all simulations, with gains of up to 5.1 dB. However, the use of constant energy signals with two tones involves a rather drastic reduction in signal spacing. The number of signals is also reduced significantly, and requires that a higher rate code be employed. These factors combine to cause the relatively poorer performance of Case 1B. The transmission of two bits per period provides the best opportunity for comparison of the constellations. The codes and the size of the signal sets are fairly similar in Cases 2A, 2B, and 2C. Under these conditions, neither of the two types of constellations appears significantly superior. The performance results for each simulation are within one dB for all three cases. When three bits are transmitted per signalling interval, the constant signal energy situation provides the best performance. However, Case 3A has a higher rate code and uses only one half the number of tones as in Case 3B. In spite of this, it actually has better performance for hard decision decoding, due to the signal space mapping.

There is another difference between constant and variable energy signal constellations that has more significance when considering actual implementation of the system. With variable energy signals, the transmitter output power varies according to the number of tones to be sent. A high peak to average power requirement might prove to be a disadvantage since the transmitter will be operating at less than its full peak

power capability for a large percentage of the time. The zero energy signal might also prove undesirable to implement if channel fading or drop-outs were possible. The receiver for unequal energy signals has increased complexity because of the necessity to compute an exponential function and a modified Bessel function in order to calculate each decision variable. However, the mapping of codewords onto the signal set is straightforward for this situation. In the constant signal energy case, the signal generator in the transmitter must adjust the amplitude of the signal tones to ensure the signal energy remains constant. The computation of decision variables in the receiver is simplified, but the mapping of codewords onto the signal space must be incorporated in both the transmitter and receiver.

4.2 Suggestions for Further Work

The results of this study raise some questions that may provoke further research. Although significant performance improvements have been discovered, further improvements may be obtained from different codes and signal constellations. One example of this might be codes with better distance profiles than the cascaded rate 1/2 codes used in this study, to improve hard decision performance. The use of larger signal spaces with lower rate codes is possible for eight tone and even sixteen tone signalling. Only 64 of 256 possible tone combinations were used in the signal set of Case 3B, and 16 tones provide 65,536 possible signals from which to construct a constellation. It is noteworthy that trellis coding in coherent communications systems obtains larger coding gains with larger signal sets [8]. Excessive computing time requirements for the simulation program used in this study precluded further exploration of larger signal sets.

Another subject for further investigation is the performance of these schemes in an anti-jamming system in the presence of jammer interference. Other studies [3, 4] have shown good results for convolutional coding with conventional FSK signalling, and so this is a promising area. The implementation of soft decision decoding becomes more complex in a jamming environment. Side information regarding the presence of a jamming signal is required by the decoder or else soft decision decoding seriously degrades in performance.

4.3 Summary

The use of convolutional codes with multiple tone signals has been shown to provide performance improvements over M-ary FSK signalling. The coded system does not require additional bandwidth to maintain the same information rate. Indeed in some cases a smaller modulation bandwidth and hence higher processing gain appears possible. The coding is suitable for implementation in a frequency hopped spead spectrum anti-jam system with a usual spectrum analyzing receiver followed by a Viterbi decoder. Increased performance gains through different codes and signal sets, as well as performance in a jamming environment are topics of further interest.

REFERENCES

1. Stark, Wayne E., "Coding for Frequency Hopped Spread Spectrum Communication with Partial Band Interference", **IEEE Transactions on Communications**, Vol. COM-33, No. 10 (Oct. 1985), pp. 1036-1057.

2. Pursley, Michael B. and Stark, Wayne E., "Performance of Reed-Solomon Coded Frequency Hopped Spread Spectrum Communications in Partial Band Interference", **IEEE Transactions on Communications**, Vol. COM-33, No. 8 (Aug. 1985), pp. 767-774.

3. Ma, Howard H. and Poole, Margaret A., "Error-Correcting Codes Against the Worst-Case Partial-Band Jammer", **IEEE Transactions on Communications**, Vol. COM-32, No. 2 (Feb. 1984), pp. 124-133.

4. Torrieri, Don J., "Frequency Hopping with Multiple Frequency Shift Keying and Hard Decisions", **IEEE Transactions on Communications**, Vol. COM-32, No. 5 (May 1984), pp. 574-582.

5. Pawula, R. F. and Golden, R., "Simulations of Convolutional Coding/Viterbi Decoding with Noncoherent CPFSK", **IEEE Transactions on Commmunications**, Vol. COM-29, No. 10 (Oct. 1981), pp. 1522-1526.

6. Keightley, R. J., "A Simulation of Frequency-Dehopped Binary and 4-ary NC-FSK Signals", Communications Research Centre, Ottawa, Canada, Technical Note No. 722, April 1984.

7. Ungerboeck, Gottfried, "Channel Coding with Multilevel/Phase Signals", **IEEE Transactions on Information Theory**, Vol. IT-28, No. 1 (Jan. 1982), pp. 55-66.

8. Ungerboeck, Gottfried, "Trellis-Coded Modulation with Redundant Signal Sets", **IEEE Communications Magazine**, Vol. 25, No. 2 (Feb. 1987), pp. 5-21.

9. Calderbank, A. Robert and Mazo, James E., "A New Description of Trellis Codes", **IEEE Transactions on Information Theory**, Vol. IT-30, No. 6 (Nov. 1984), pp. 784-791.

10. Thapar, Hemant K., "Real-Time Application of Trellis Coding to High-Speed Voiceband Data Transmission", **IEEE Transactions on Selected Areas in Communications**, Vol. SAC-2, No. 5. (Sept. 1984), pp. 648-658.

11. Wozencraft, John M. and Jacobs, Irwin Mark, **Principles of Communication Engineering**, John Wiley and Sons Inc., 1965.

12. Viterbi, Andrew J. and Omura, Jim K., **Principles of Digital Communication and Coding**, McGraw-Hill Book Co., 1979.

13. Golomb, S. W. (editor), **Digital Comunications with Space Applications**, Prentice-Hall Inc., 1964.

14. Clark, George C. and Cain, J. Bibb, **Error Correction Coding for Digital Communications**, Plenum Press, 1981.

15. Larsen, Knud J., "Short Convolutional Codes with Maximal Free Distance for Rates 1/2, 1/3, and 1/4", **IEEE Transactions on Information Theory**, Vol. IT-19, No. 3 (May 1973), pp. 371-372.

16. Paaske, Erik, "Short Binary Convolutional Codes with Maximal Free Distance for Rates 2/3 and 3/4", **IEEE Transactions on Information Theory**, Vol. IT-20, No. 5 (Sept. 1974), pp. 683-689.

17. Odenwalder, J. P., **Dual-k Convolutional Codes for Noncoherently Demodulated Channels**, Proc. IEEE Int. Telem. Conf., 1976.

18. Forney, G. David, "The Viterbi Algorithm", **Proceedings of the IEEE**, Vol. 61, No. 3 (March 1973), pp. 268-278.

19. Heller, Jerrold A. and Jacobs, Irwin Mark, "Viterbi Decoding for Satellite and Space Communication", **IEEE Transactions on Communications Technology**, Vol. COM-19, No. 5 (Oct. 1971), pp. 835-848.

20. Viterbi, Andrew J., "Orthogonal Tree Codes for Comunication in the Presence of White Gaussian Noise", **IEEE Transactions on Communications Technology**, Vol. COM-15, No. 2 (April 1967), pp. 238-242.

21. Simon, Marvin K., Omura, Jim K., Scholtz, Robert A., and Levitt, Barry K., **Spread Spectrum Communications, Volume I**, Computer Science Press Inc., 1985.

22. Omura, Jim K. and Levitt, Barry K., "Coded Error Probability Evaluation for Antijam Communication Systems", **IEEE Transactions on Communications**, Vol. COM-30, No. 5 (May 1982), pp. 896-903.

Interim Progress Report

Part II

HIGH DATA RATE SPREAD SPECTRUM SYSTEMS WITH BAND-EFFICIENT MODULATIONS

by

Y.M. Lam,

P.H. Wittke.

ABSTRACT

In this report final results on the error performance of a high data rate, hopped spread spectrum transmission system that employs band efficient modulations that are coherent during each hop, are presented. Earlier reports provided results on the signalling spectrum, optimum receivers and system complexity. In this report period, upper bounds on the average bit error probability of the hop-by-hop sequence estimation noncoherent receiver for MSK, DMSK and TFM with h=0.5 and with both rectangular and raised cosine pulse shapings have been evaluated. To indicate the tightness of the bounds, a computer simulation study has been carried out for the hop-by-hop receiver with MSK modulation, for different lengths of hop interval.

1. INTRODUCTION

A slow frequency-hopped, high data rate spread spectrum system [1], with a bandwidth efficient modulation in each hop interval, namely continuous phase modulation (CPM) [2], was proposed in [3,4]. Power density spectra of the frequency-hopped signals with various correlative encoding schemes [5,6] and baseband pulse shapings were calculated for various hop lengths in [4]. In general, as the length of the hop interval increases the power density spectrum of the frequency-hopped CPM becomes more compact with lower sidelobes and approaches that of the CPM signal without hopping. The use of higher order correlative encoding schemes and baseband pulse shaping is found to be effective in bandwidth and sidelobe reduction only for long hop intervals. For short hop lengths, the usual techniques for reducing bandwidth by using suitable higher order correlative encoding schemes and pulse shaping do not yield much in bandwidth reduction. Hence, for systems with short frequency hopping intervals it would be more advantageous to use simple CPM schemes such as minimum shift keying (MSK) [7,8] allowing a simpler receiver implementation.

Three algorithm-based, sequence estimating noncoherent receivers were proposed for decoding the frequency-hopped correlative encoded CPM signals in [4]. Except for the first receiver, which decodes the transmitted sequence on a hop-by-hop basis, the other two receivers both require the signal to noise ratio (SNR) to be known by the receivers in order to calculate the metric across the hops. In this report, we shall concentrate on the performance of the hop-by-hop sequence estimation receiver which appears the most promising design among the three receivers proposed.

Upper bounds on the bit error probability of the hop-by-hop sequence estimation receiver will be presented. The upper bound calculation is based on the union bound approach with some approximation necessary when hop intervals are long. The bit error probability bound will be presented for various correlative encoding schemes with rectangular and raised cosine pulse shapings.

A computer simulation study of the receiver is also carried out to obtain a more precise error performance evaluation and to indicate the accuracy of the bounds. The simulation also allows the effects on the receiver performance of modification of the decoding algorithm, to be easily observed. It is found that if the decoder keeps more than one survivor for each state of the modulation trellis, an improvement in the performance of the sequence estimation receiver is obtained.

2. APPROXIMATE UPPER BOUND ON BIT ERROR PROBABILITY

The optimum maximum likelihood noncoherent hop-by-hop detection receiver is derived in [3,4]. Given that $\underline{\alpha} = \{ \alpha_0, \alpha_1, \ldots, \alpha_{N-1} \}$ is the transmitted sequence, the receiver decides on the sequence $\underline{\alpha}' = \{ \alpha_0', \alpha_1', \ldots, \alpha_{N-1}' \}$ which gives the maximum equivalent likelihood [4] given by

$$\ell' (\underline{\alpha}, \underline{\alpha}') = \ell_c^{\,2}(\underline{\alpha}, \underline{\alpha}') + \ell_s^{\,2}(\underline{\alpha}, \underline{\alpha}') \qquad (2.1)$$

where $\ell_c(\underline{\alpha}, \underline{\alpha}')$ and $\ell_s(\underline{\alpha}, \underline{\alpha}')$ are the inphase and quadrature correlations [4] which are given by

$$\ell_c(\underline{\alpha},\underline{\alpha}') = \int_o^{NT} r(t) \cos [2\pi f_c t + \psi(t, \underline{\alpha}')] \, dt \tag{2.2a}$$

$$\ell_s(\underline{\alpha}, \underline{\alpha}') = \int_o^{NT} r(t) \sin [2\pi f_c t + \psi(t, \underline{\alpha}')] \, dt \tag{2.2b}$$

where $r(t)$ is the received signal and $\psi(t, \underline{\alpha}')$ is the information carrying phase function as defined in the previous report [4].

Errors occur whenever the detected sequence $\underline{\alpha}'$ differs from $\underline{\alpha}$ in one or more places. The probability of detecting an incorrect sequence $\underline{\alpha}' \neq \underline{\alpha}$ is simply the probability that $\ell'(\underline{\alpha}, \underline{\alpha}')$ is greater than $\ell'(\underline{\alpha}, \underline{\alpha})$, which is just the probability of one Rician variable exceeding another for which the expression is known [4,10,12,13]. Since the probability of error depends on only the difference between $\underline{\alpha}'$ and $\underline{\alpha}$ rather than on the individual $\underline{\alpha}$ and $\underline{\alpha}'$ [4,10], we let

$$\underline{\gamma} = \underline{\alpha}' - \underline{\alpha} \tag{2.3}$$

The union bound on the average bit error probability is simply given by averaging the bit error probability over all the possible transmitted sequences as

$$P_e \leq \frac{1}{\mu} \sum_{\substack{\underline{\gamma} \\ \underline{\gamma} \neq o}} \frac{e(\underline{\gamma})}{N} W(\underline{\gamma}) P(\underline{\gamma}) \tag{2.4}$$

where $\mu = 2^N$ is the total number of possible transmitted sequences over the hop interval of length NT. $e(\underline{\gamma})$ is the number of nonzero sequence elements of $\underline{\gamma}$, which corresponds to the number of bit errors when $\underline{\alpha}'$ is detected given that $\underline{\alpha}$ is transmitted. $W(\underline{\gamma})$ is the total number of pairs of $\underline{\alpha}'$ and $\underline{\alpha}$ corresponding to the

difference sequence $\underline{\gamma}$. Since a difference sequence $\underline{\gamma}$ having $e(\underline{\gamma})$ nonzero elements would result from $2^{N-e(\underline{\gamma})}$ different pairs of $\underline{\alpha}'$ and $\underline{\alpha}$, $W(\underline{\gamma})$ is given as

$$W(\gamma) = 2^{N-e(\underline{\gamma})} \tag{2.5}$$

$P(\underline{\gamma})$ is the probability of an error by detecting $\underline{\alpha}'$ rather than $\underline{\alpha}$ as the transmitted sequence and is given by

$$P(\gamma) = \frac{1}{2} \left[1 - Q\left(\sqrt{b}, \sqrt{a}\right) + Q\left(\sqrt{a}, \sqrt{b}\right) \right] \tag{2.6}$$

where $Q(\ ,\ )$ is the Marcum Q-function [12,13]. Techniques for recursive evaluation of Q are given in [14,15]. The parameters a and b are given by

$$\begin{Bmatrix} a \\ b \end{Bmatrix} = \frac{N E_b}{2 N_o} \left[1 \mp \left\{ 1 - \left| \rho(\gamma) \right|^2 \right\}^{1/2} \right] \tag{2.7}$$

where $\rho(\gamma)$ is the complex correlation and $|\rho(\gamma)|^2$ is given by

$$\left| \rho(\underline{\gamma}) \right|^2 = \left[\frac{1}{NT} \int_o^{NT} \cos \psi\ (t, \underline{\gamma})\ dt \right]^2 + \left[\frac{1}{NT} \int_o^{NT} \sin \psi\ (t, \underline{\gamma})\ dt \right]^2 \tag{2.8}$$

For binary transmission γ_n can take on 3 possible values namely 0, + 2 and - 2. Hence there can be $3^N - 1$ different difference sequences $\underline{\gamma}$, for a hop length of NT in Eq. (2.4). For any $\underline{\gamma}$ there is always another sequence with opposite signs in all the sequence elements. The number of terms in Eq. (2.4) can be reduced by a factor of 2 since $P(-\underline{\gamma}) = P(\underline{\gamma})$ as can be easily seen from Eqs. (2.6) to (2.8). The number of difference sequences to be considered in Eq. (2.4) is then $\frac{1}{2}(3^N - 1)$.

If we represent 2, 0 and -2 as 0, 1 and 2 respectively, then each difference sequence element γ_n can be represented as a base-3 digit, and the difference sequence can then be represented by N base-3 digits. Each $\underline{\gamma}$ to be

considered in the error-rate calculation Eq. (2.4), can be represented by a base-3 number of value from 0 to $\frac{1}{2}(3^N-1)-1$. Numbers from 0 to $\frac{1}{2}(3^N-1)-1$ are decoded sequentially as base-3 digits, which are then converted to 0, +2 or -2 as the difference sequence elements, to be used in evaluating the union bound as given by Eq. (2.4).

$|\rho(\underline{\gamma})|^2$ required to calculate $P(\underline{\gamma})$ in Eq. (2.4), can be rewritten as

$$|\rho(\gamma)|^2 = \{\sum_{k=o}^{N-1} [\cos \zeta_k C(d_k) - \sin \zeta_k S(d_k)]\}^2 + \{\sum_{k=o}^{N-1} [\cos \zeta_k S(d_k) + \sin \zeta_k C(d_k)]\}^2 \quad (2.9)$$

where

$$C(d_k) = \int_o^T \cos 2\pi h d_k q(t)\, dt \quad (2.10a)$$

$$S(d_k) = \int_o^T \sin 2\pi h d_k q(t)\, dt \quad (2.10b)$$

$$\zeta_k = \pi h \sum_{n=o}^{k-1} d_n \quad (2.11)$$

d_k is the correlated difference sequence given by

$$d_k = \frac{1}{C} \sum_{\ell=o}^{m} k_\ell \gamma_{k-\ell} \quad (2.12)$$

where C is the normalizing constant and k_ℓ's are the coefficients of the correlative encoding polynomial as defined in [4].

To speed up the calculation of $|\rho(\underline{\gamma})|^2$, all possible integrals of the cosine and sine of the possible phase over a symbol interval, $C(d_k)$ and $S(d_k)$ are initially calculated and stored in an array. This avoids repeated numerical integration to obtain $C(d_k)$ and $S(d_k)$ when the baseband pulse is not rectangular.

When N is large, evaluation of the union bound as given by Eq. (2.4) becomes computationally infeasible, since the number of difference sequences to be considered is $\frac{1}{2}(3^N-1)$, which grows exponentially with N. Approximation of the union bound is then necessary for large N. At high SNR, the probabilities of those error paths, which have large differences [4,9,10,11], are small compared with those having small differences. The paths having large differences can then be neglected. The upper found on the bit error probability is approximated as

$$P_e \leq \frac{1}{\mu} \sum_{\substack{\underline{\gamma}\in S \\ \underline{\gamma}\neq o}} \frac{e(\underline{\gamma})}{N} W(\underline{\gamma})\, P(\underline{\gamma}) \tag{2.13}$$

where S denotes the set of difference sequences where the difference is not large.

To determine S the set of $\underline{\gamma}$'s, we notice that there will be complete error events during the hop and also incomplete error events starting near the end of a hop. For the hop-by-hop detection receiver, the difference sequences $\underline{\gamma}$ corresponding to an error event $e_i^{k_i}$ of length ℓ_i starting at time $k_1 T$, will have $\gamma_i=0$ for $i<k_1$ and $i>k_1+\ell_i$. The complex correlation as given in Eq. (2.8) becomes

$$\begin{aligned} |\rho(\underline{\gamma})|^2 &= \left\{ \frac{1}{NT} \left[k_1 T + \int_{k_1 T}^{(k_1+\ell_i)T} \cos\psi\ (t,\underline{\gamma})\, dt + (N-k_1-\ell_i)\, T \right] \right\}^2 \\ &\quad + \left\{ \frac{1}{NT} \int_{k_1 T}^{(k_1+\ell_i)T} \sin\psi\,(t,\underline{\gamma})\, dt \right\}^2 \\ &= \left\{ \frac{1}{NT} \left[(N-\ell_i)T + \int_{k_1 T}^{(k_1+\ell_i)T} \cos\psi\,(t,\underline{\gamma})\, dt \right] \right\}^2 \\ &\quad + \left\{ \frac{1}{NT} \int_{k_1 T}^{(k_1+\ell_i)T} \sin\psi\,(t,\underline{\gamma})\, dt \right\}^2 \end{aligned} \tag{2.14}$$

where $k_1 + \ell_i \leq N$

Hence for error events having a particular phase difference path segment from $k_1 T$ to $(k_1 + \ell_i)T$, $|\rho(\gamma)|^2$ will be the same, independent of k_1 as long as $o<k_1<N-\ell_i$, and hence the $P[e_i]$ will be the same. It is then sufficient to determine only the segment of $\underline{\gamma}$ from 0 to $\ell_i - 1$ corresponding to an error event starting at time 0. There will be two $\underline{\gamma}$'s corresponding to the error event e_i starting from t = 0 to $(N-\ell_i)T$ and having the same error probability.

For a correlative encoded CPM with frequency pulse of length LT, the minimum length of a complete error event is (L+1)T. Difference sequences of various lengths $\ell_i \geq L+1$ are generated to find the complete error event paths starting at time 0. Once a complete error event path, which has $\gamma_o \neq 0$ and $\gamma_\ell = o$ (i.e. the difference phase path deviates from zero at time 0 and merges back to zero at time $\ell_i T$) is determined, the correlation is calculated by Eq. (2.14). Since it is computationally infeasible to include all complete error events for large N and also since long error events have large differences, hence only difference sequence segments of length $L+1 < \ell_i < \ell_{max}$ are included.

If an error event starts at some time (N-L)T or later, before the end of a hop, then an incomplete error event will occur. These incomplete error events starting near the end of a hop have the estimated phase path differing from the true path in only a few places and thus are likely to occur. The $\frac{1}{2}(3^L-1)$ possible difference sequences corresponding to the incomplete error events at the end of a hop are generated and their probabilities evaluated.

To summarize, the bit error probability upper bound is approximated by

$$P_e \leq \frac{1}{\mu} \{ \sum_{e_i \varepsilon S_c} 2(N-\ell_i+1)\, 2^{N-e_i} \frac{e_i}{N} P[e_i] + \sum_{e_j \varepsilon S_I} 2^{N-e_j+1} \frac{e_j}{N} P[e_j] \} \qquad (2.15)$$

where S_c is the set of complete error events having small equivalent distances and S_I is the set of incomplete error events occurring near the end of a hop.

The upper bound on the error performance of the hop-by-hop noncoherent receiver has been evaluated for a variety of correlative encoding and baseband pulse shaping schemes when the frequency hopping interval is 4, 16 and 64 symbol intervals. For the hop length of 4 intervals, union bounds are evaluated while for the longer hop intervals 16 and 64, approximated upper bounds are evaluated.

Minimum shift keying (MSK), duobinary MSK (DMSK), and tamed FM (TFM) encoding schemes with h=0.5 and rectangular pulse shaping are compared as shown in Figs. 1 to 3 for N=4, 16 and 64 respectively. It can be seen that MSK performs much better than the other two schemes for the various hop lengths. When raised cosine pulse shaping is used, the bit error probability bounds for the three encoding schemes are as shown in Figs. 4 to 6 for N=4, 16 and 64 respectively. Again, the higher order correlative encoding schemes have higher probability of error.

Next, the effect of the length of a hop on the error performance of the receiver is illustrated in Figs. 7 to 12 for each modulation. As expected, the performance of the receiver improves as the hop interval lengthens. However, for the same degree of increase in hop length, the improvement in error performance is less for MSK and is more pronounced for the higher order correlative encoding schemes such as TFM.

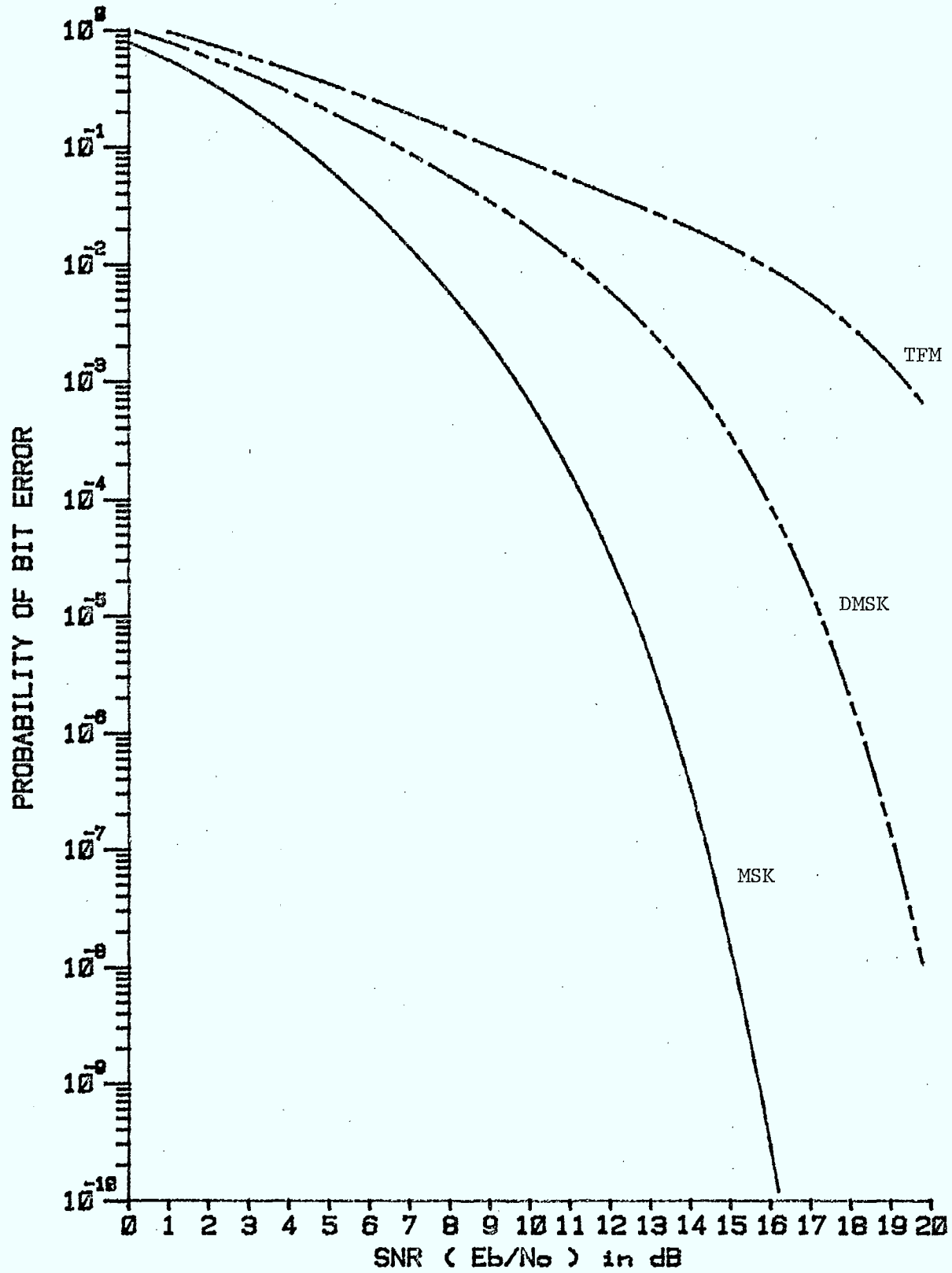


Figure 1: N=4, union bounds on P_e for MSK, DMSK and TFM with rectangular pulse shaping.

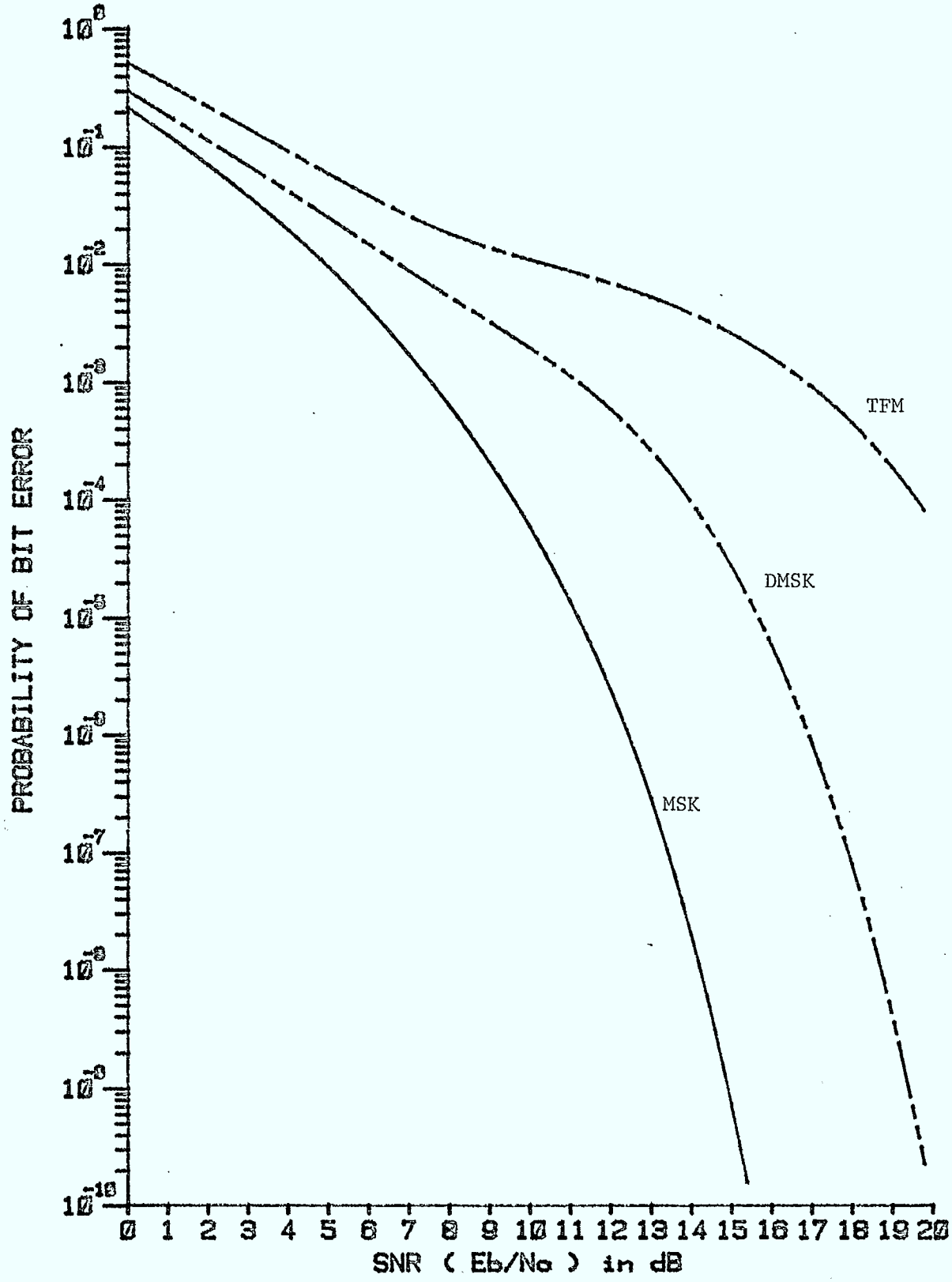


Figure 2: N=16, approximated upper bounds on P_e for MSK, DMSK and TFM with rectangular pulse shaping.

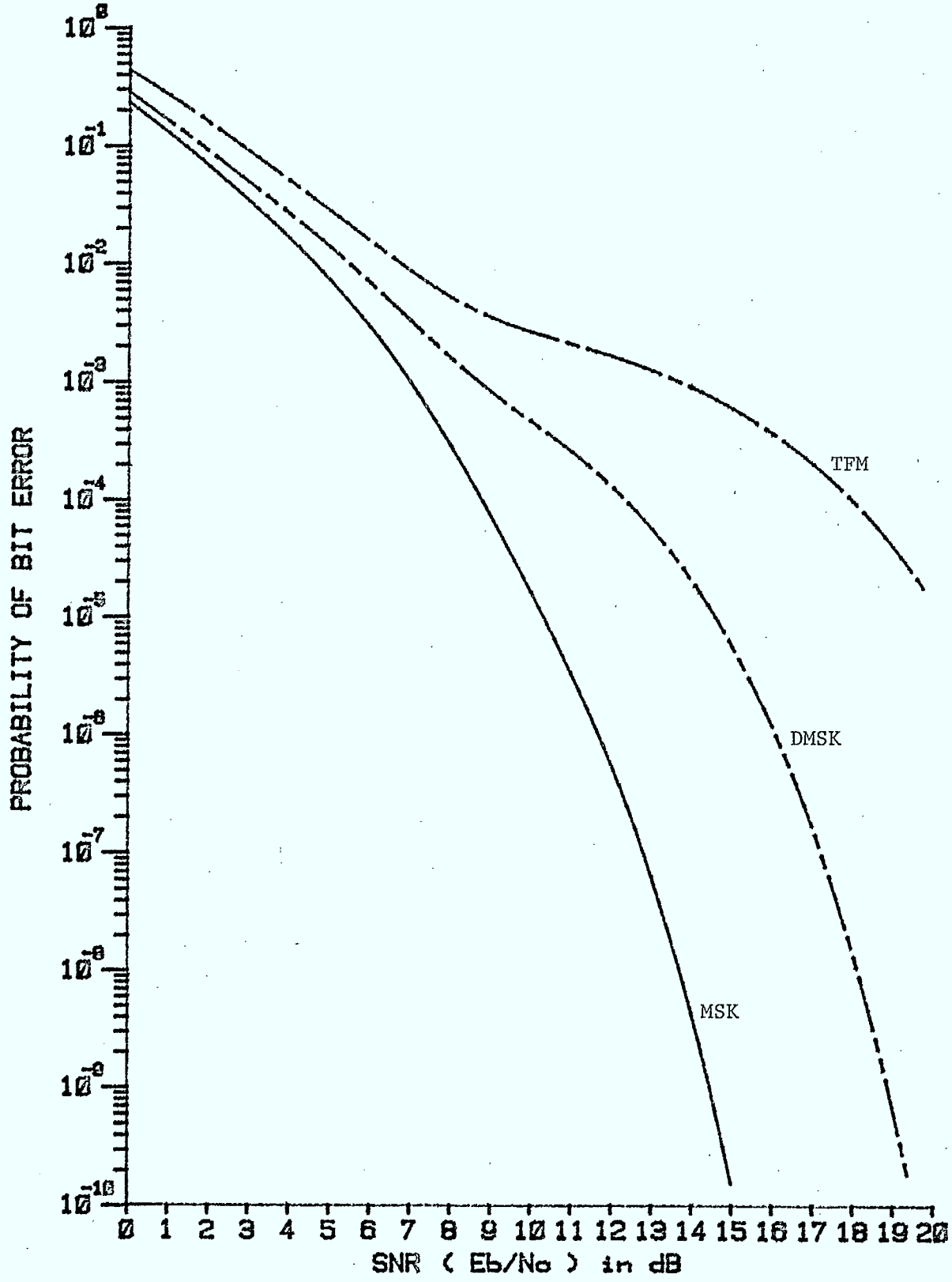


Figure 3: N=64, approximated upper bounds on P_e for MSK, DMSK and TFM with rectangular pulse shaping.

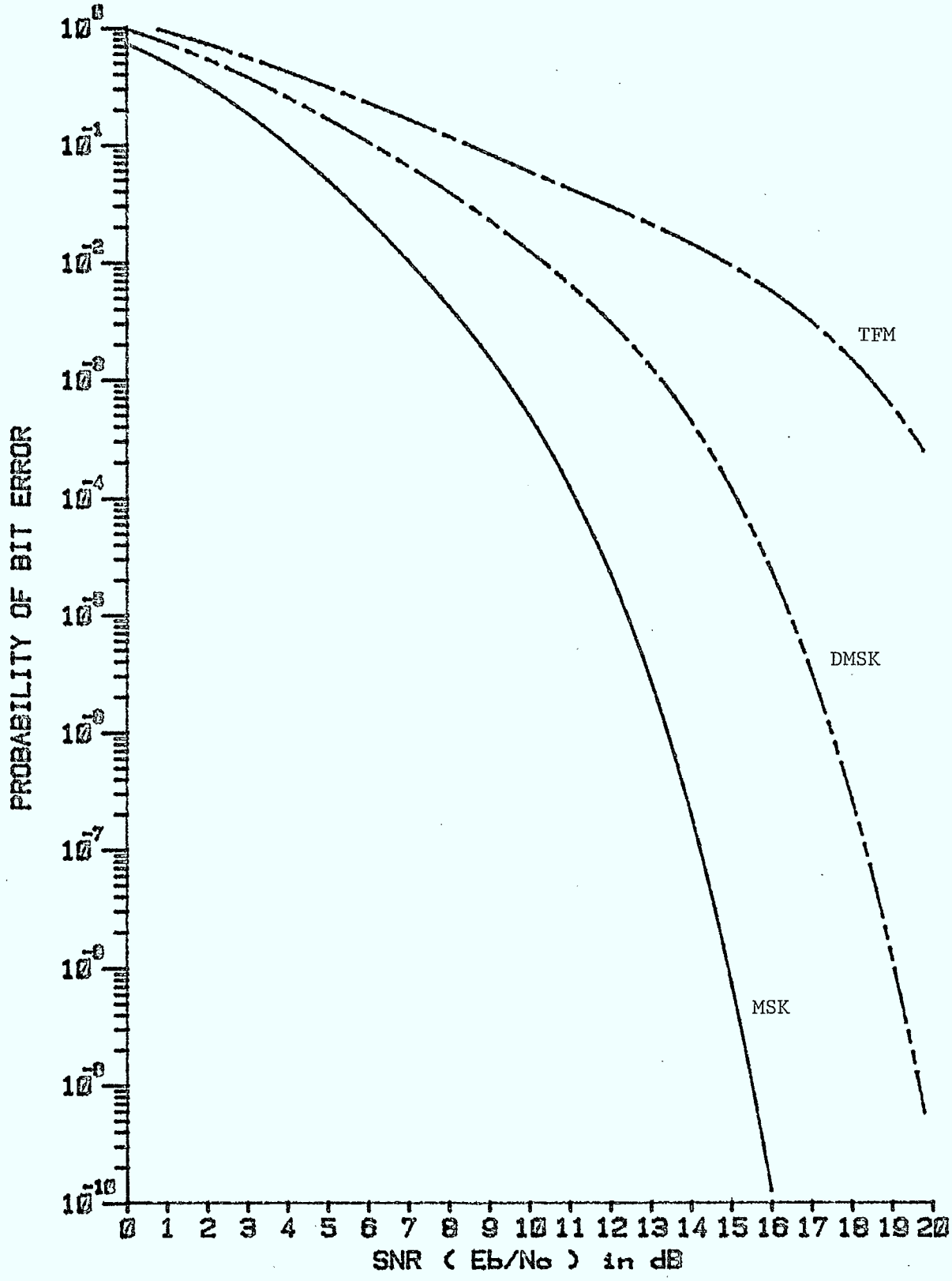


Figure 4: N=4, union bounds on P_e for MSK, DMSK and TFM with raised cosine pulse shaping.

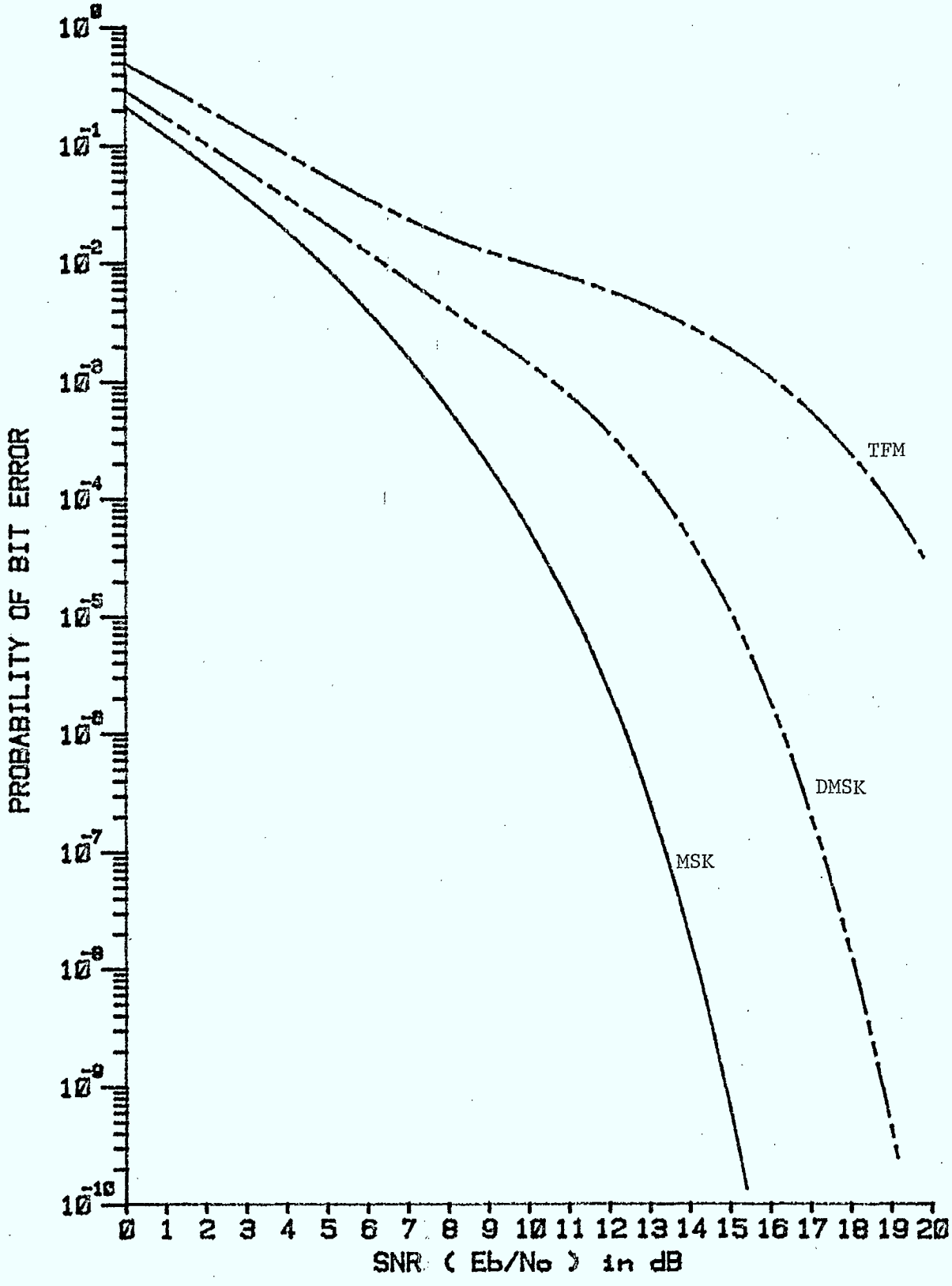


Figure 5: N=16, approximated upper bounds on P_e for MSK, DMSK and TFM with raised cosine pulse shaping.

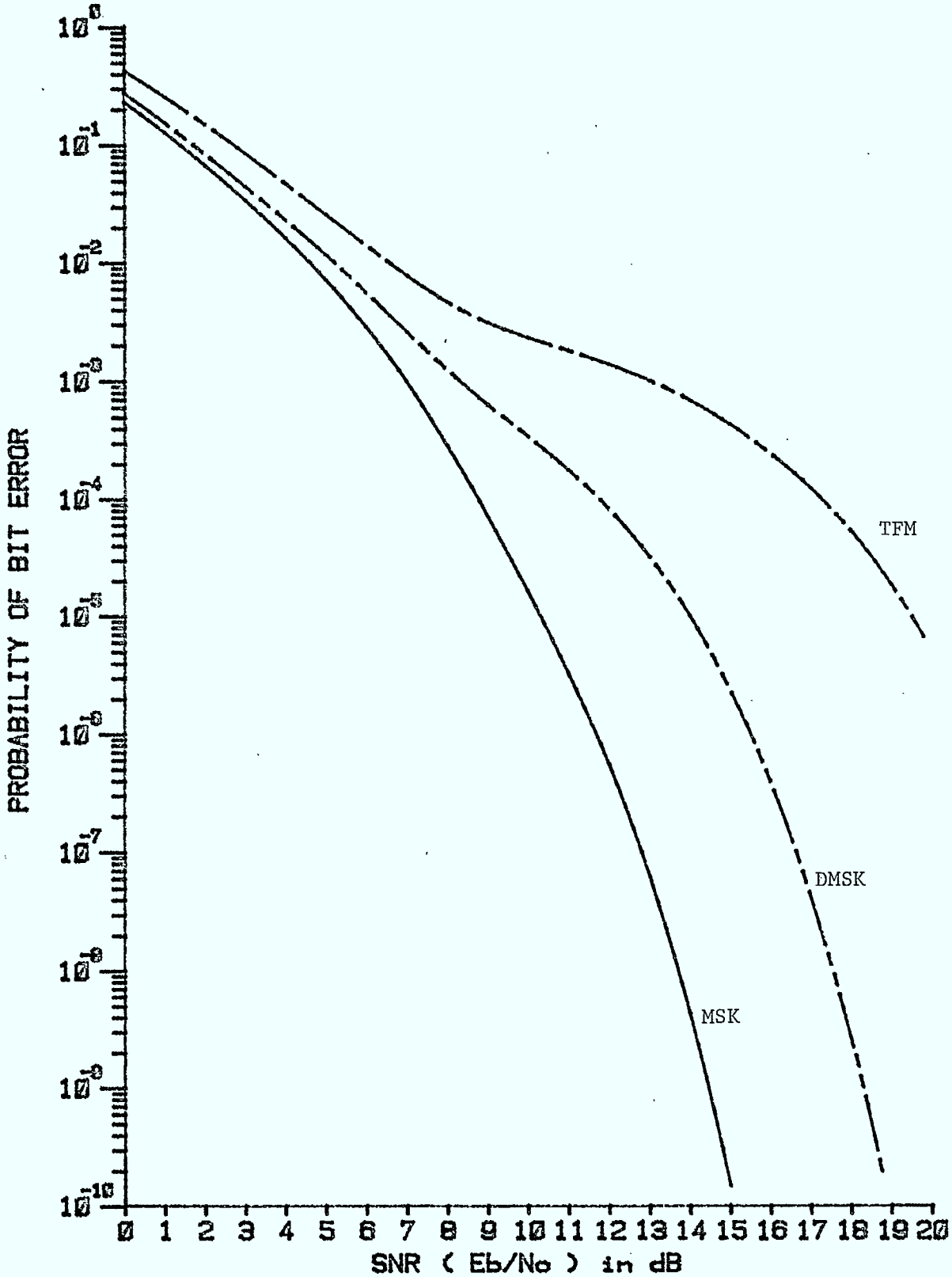


Figure 6: N=64, approximated upper bounds on P_e for MSK, DMSK and TFM with raised cosine pulse shaping.

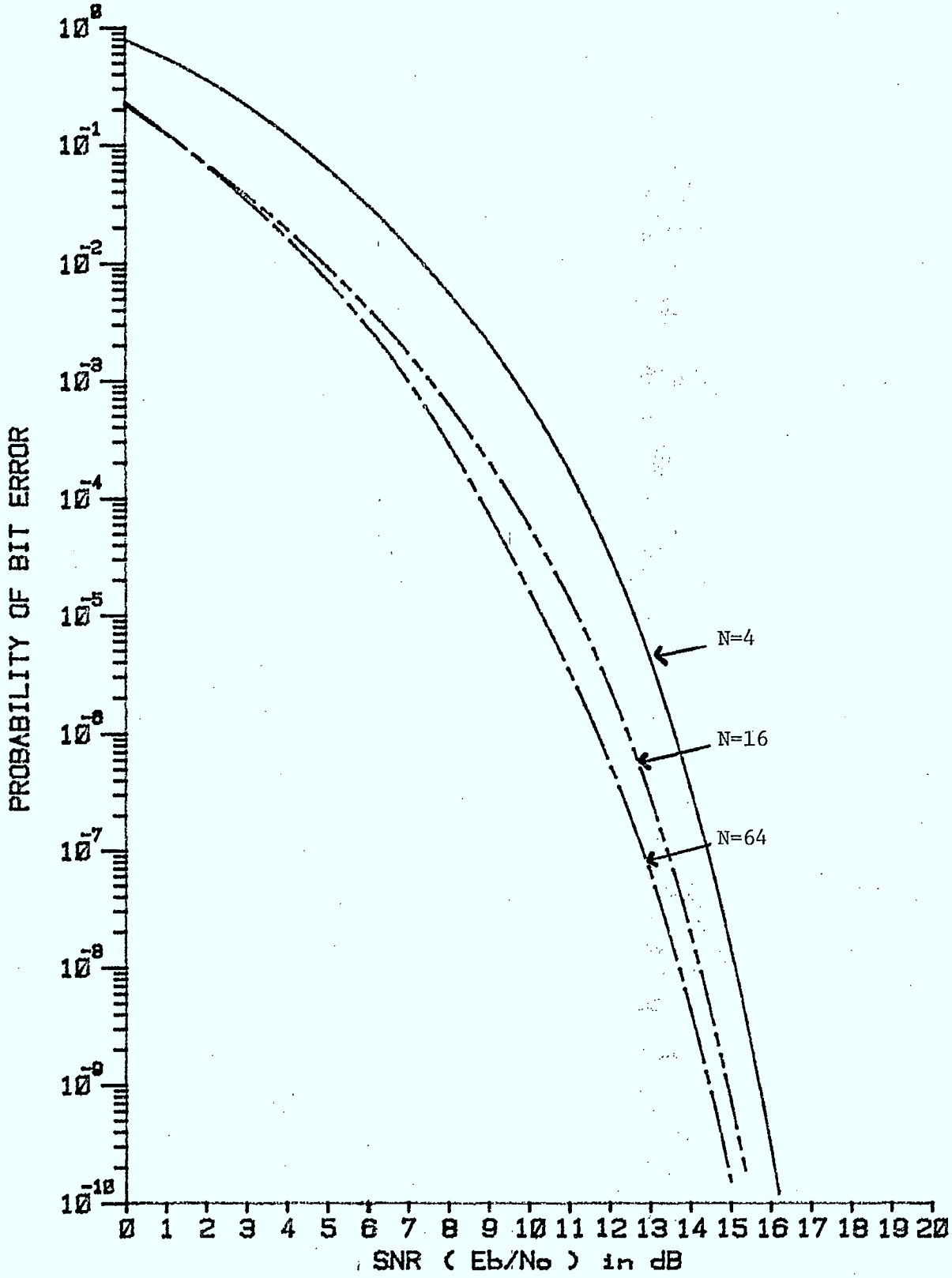


Figure 7: P_e bounds for MSK with rectangular pulse shaping.

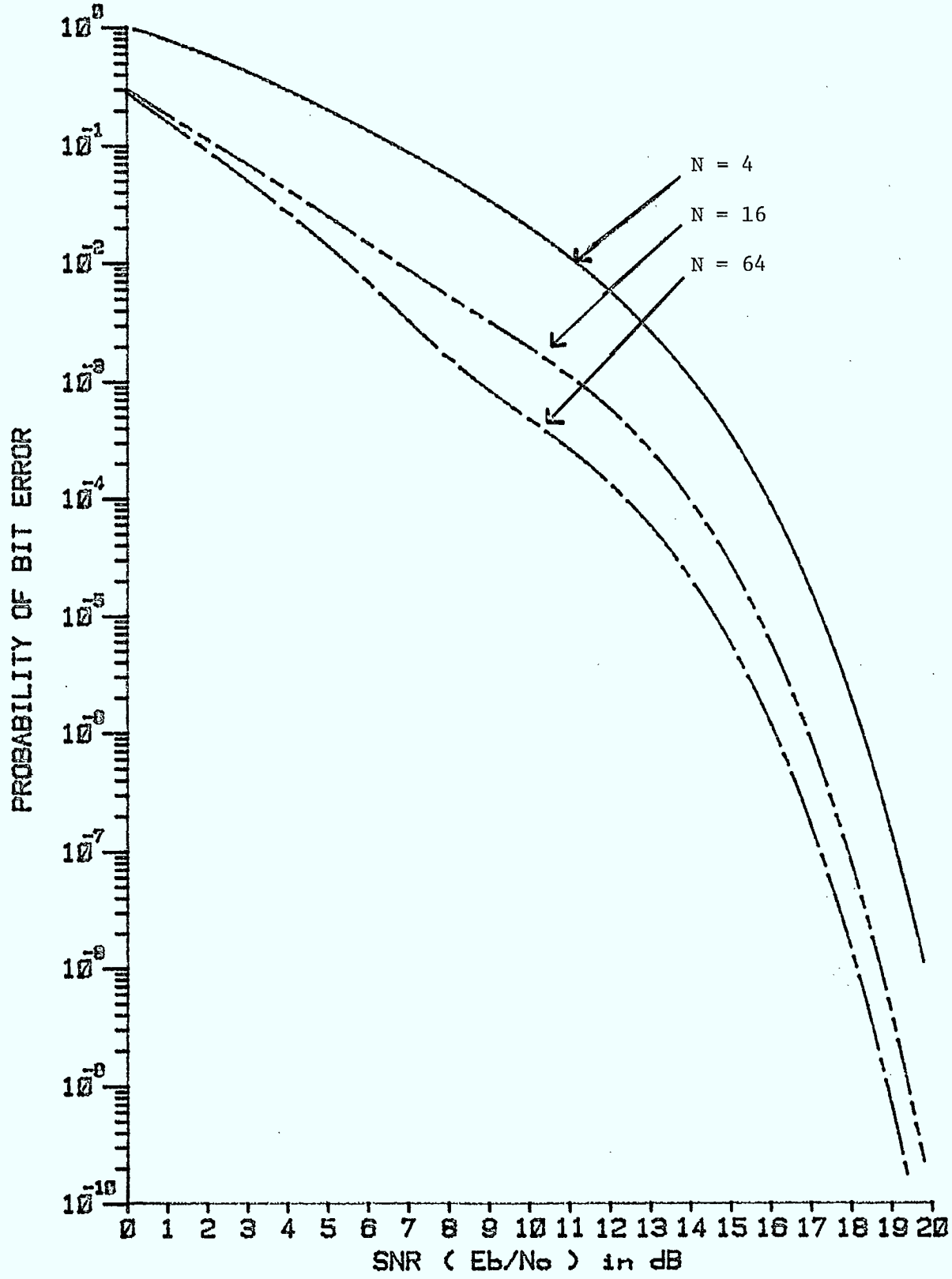


Figure 8: P_e bounds for DMSK with rectangular pulse shaping.

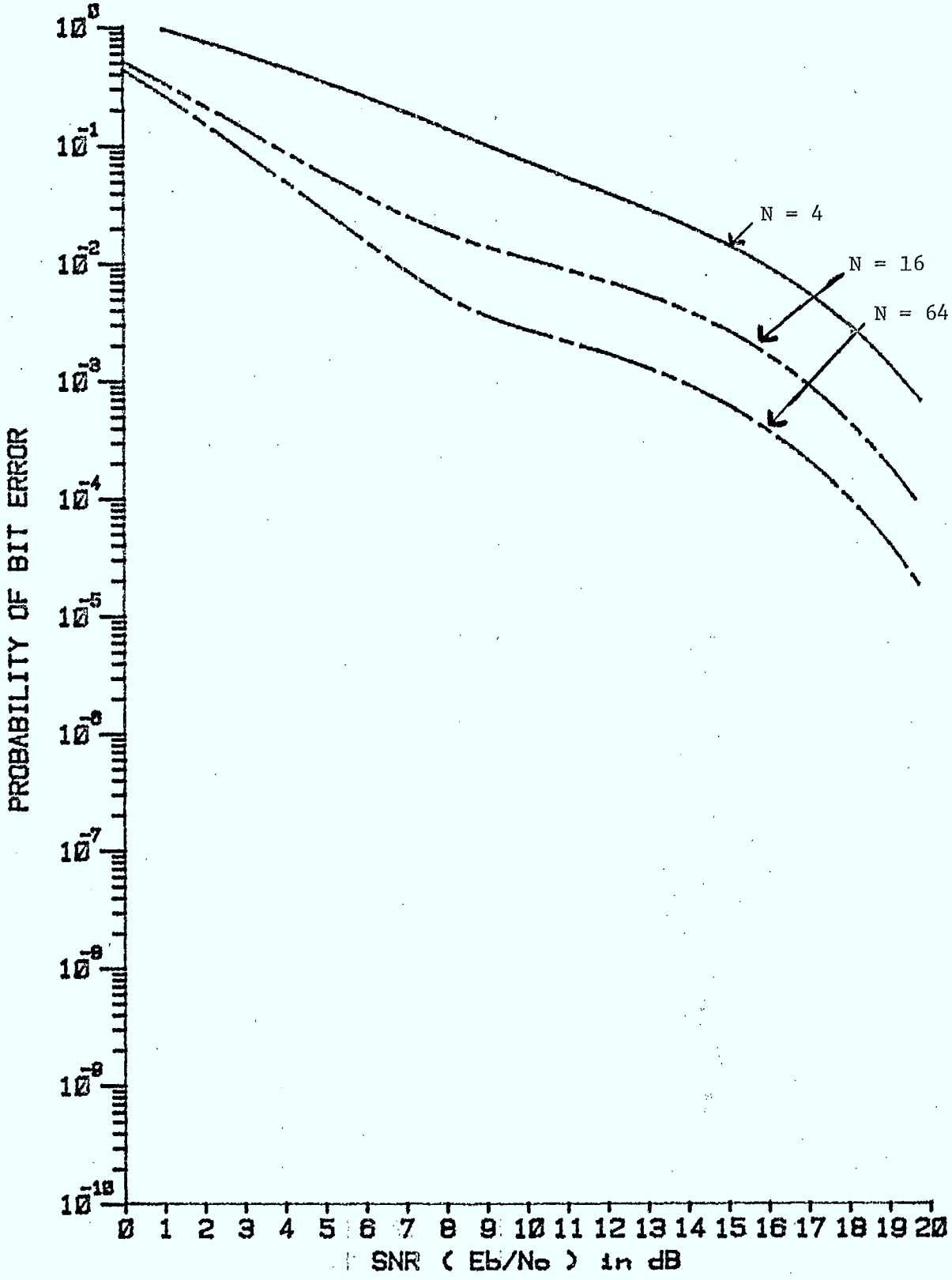


Figure 9: P_e bounds for TFM with rectangular pulse shaping.

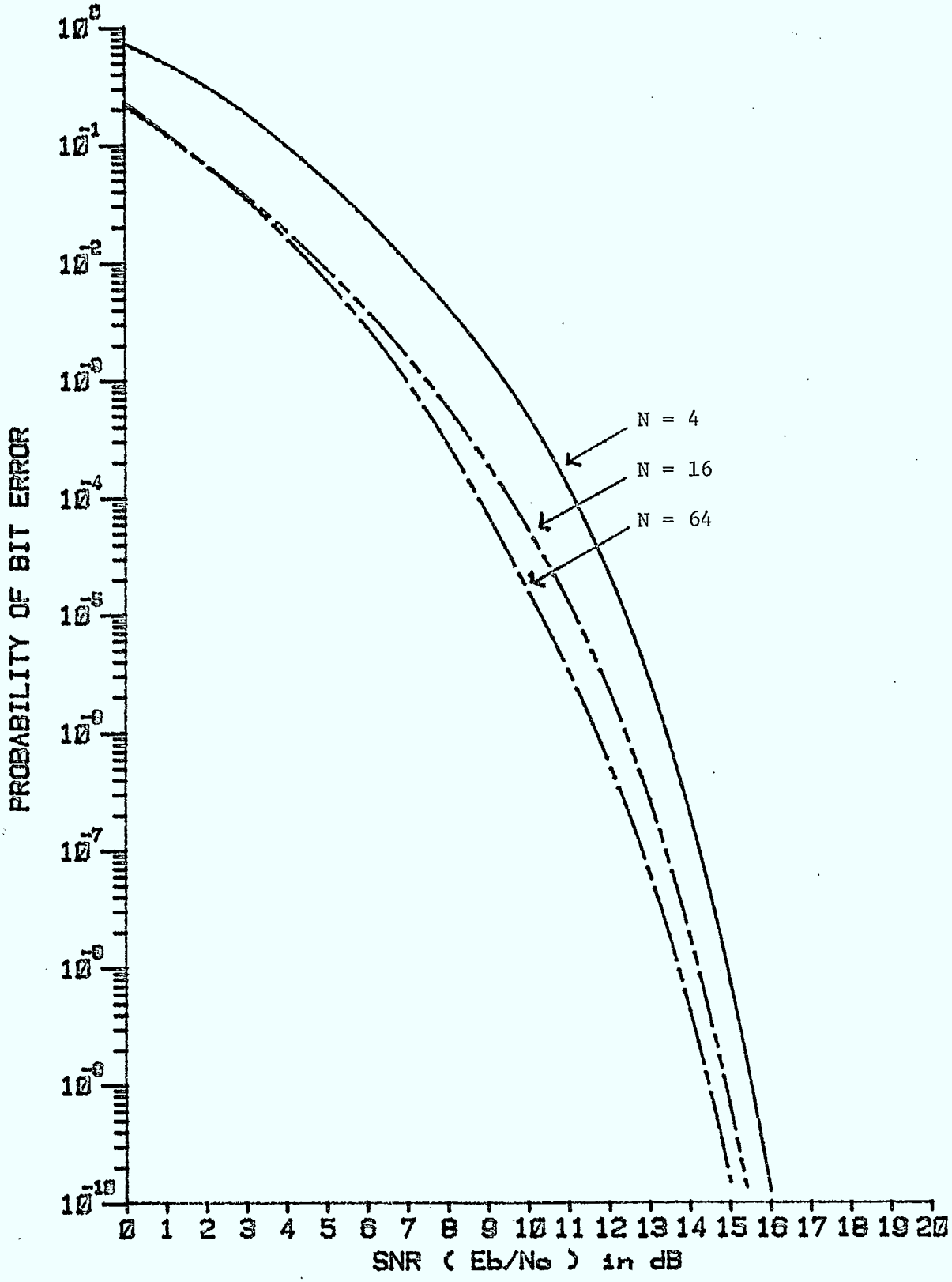


Figure 10: P_e bounds for MSK with raised cosine pulse shaping.

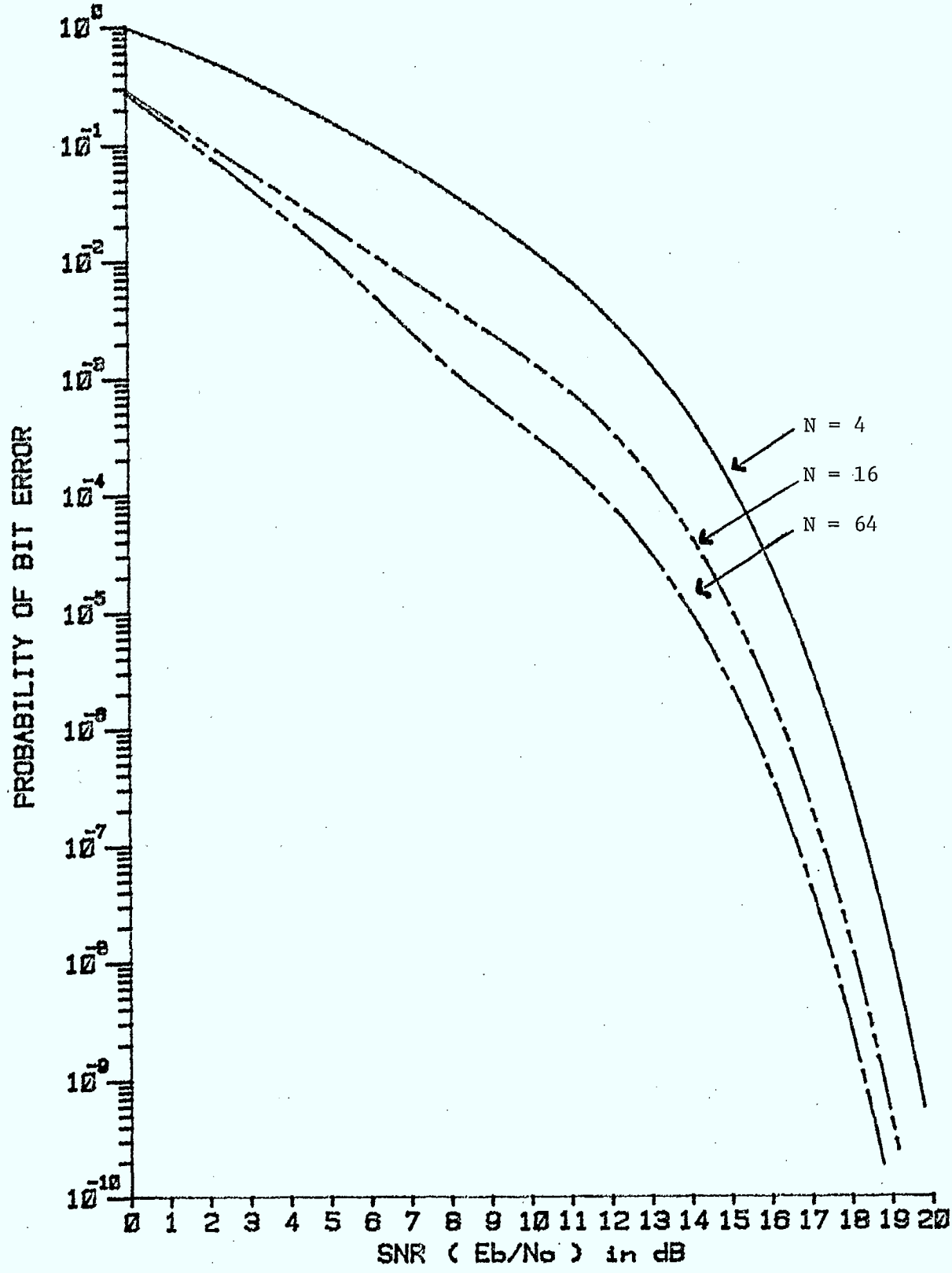


Figure 11: P_e bounds for DMSK with raised cosine pulse shaping.

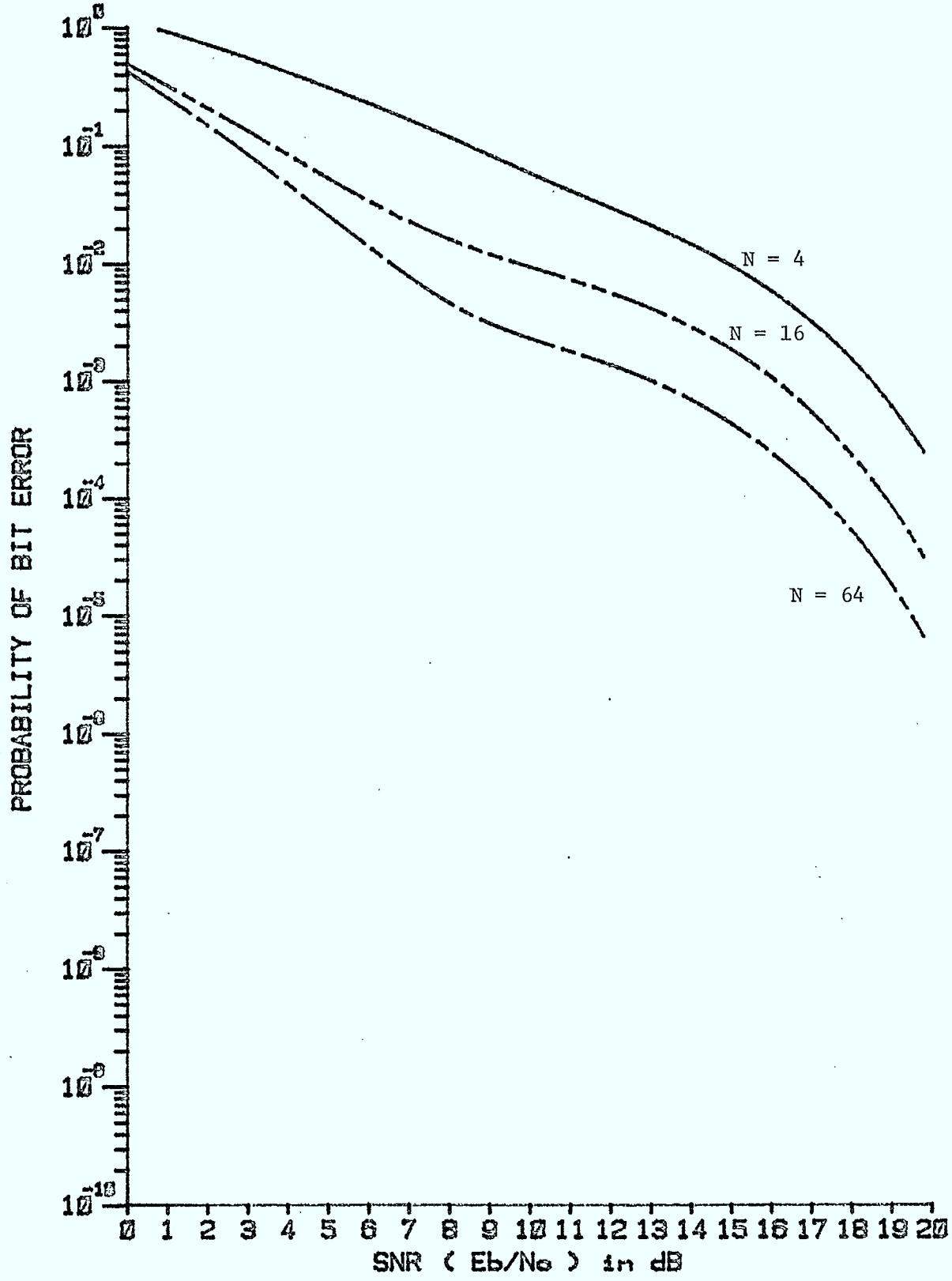


Figure 12: P_e bounds for TFM with raised cosine pulse shaping.

3. RECEIVER SIMULATION

A computer simulation study has been carried out for the hop-by-hop sequence estimating noncoherent receiver with MSK modulation. Standard Monte Carlo techniques are employed.

Recall that the maximum likelihood hop-by-hop noncoherent receiver decides on the sequence $\underline{\alpha}'$, which maximizes the equivalent likelihood as given by Eq.(2.1). The inphase and quadrature likelihoods as given by Eq.(2.2) can be rewritten as

$$\ell_c(\underline{\alpha}, \underline{\alpha}') = \sum_{k=0}^{N-1} \delta_{c,k}(\underline{\alpha}, \underline{\alpha}') \tag{3.1a}$$

$$\ell_s(\underline{\alpha}, \underline{\alpha}') = \sum_{k=0}^{N-1} \delta_{s,k}(\underline{\alpha}, \underline{\alpha}') \tag{3.1b}$$

where $\delta_{c,k}(\underline{\alpha}, \underline{\alpha}')$ and $\delta_{s,k}(\underline{\alpha}, \underline{\alpha}')$ denote the partial inphase and quadrature correlations over the k^{th} symbol interval and are given by

$$\delta_{c,k}(\underline{\alpha}, \underline{\alpha}') = \int_{kT}^{(k+1)T} r(t) \cos[2\pi f_c t + \psi(t,\underline{\alpha}')]dt \tag{3.2a}$$

$$\delta_{s,k}(\underline{\alpha}, \underline{\alpha}') = \int_{kT}^{(k+1)T} r(t) \sin[2\pi f_c t + \psi(t,\underline{\alpha}')]dt \tag{3.2b}$$

Both partial likelihoods can be further decomposed by expanding the cosine and sine terms in Eqs. (3.2a) and (3.2b) as in [4]. Without loss of information, we shall use Eq. (3.2) to find $\delta_{c,k}(\underline{\alpha}, \underline{\alpha}')$ and $\delta_{s,k}(\underline{\alpha}, \underline{\alpha}')$, which are needed for the simulation.

During the k^{th} symbol interval in the i^{th} hop, the dehopped received signal is given by

$$r(t) = \left\{\frac{2E}{T}\right\}^{1/2} \cos[2\pi f_c t + \psi(t,\underline{\alpha}) + \theta_i] + n(t) \tag{3.3}$$

$$iNT \leq t \leq (i+1)NT$$

where all the symbols used are the same as in [4]. $n(t)$ is the additive white Gaussian bandpass noise with zero mean and one-sided power spectral density No W/Hz which can be represented by [16]

$$n(t) = \sqrt{2}\, n_c(t) \cos 2\pi f_c t - \sqrt{2}\, n_s(t) \sin 2\pi f_c t \tag{3.4}$$

The baseband processes $n_c(t)$ and $n_s(t)$ are statistically independent, white and Gaussian with zero mean and one-sided power spectral density No W/Hz.

Substituting Eqs. (3.3) and (3.4) into Eqs. (3.2a) and assuming that $2\pi f_c t \gg 1$, we obtain

$$\delta_{c,k}(\underline{\alpha}, \underline{\alpha}') = \frac{1}{T}\int_{kT}^{(k+1)T} \cos[\psi(t,\underline{\alpha}') - \psi(t,\underline{\alpha}) - \theta_i]\, dt + \frac{1}{\sqrt{2}}\int_{kT}^{(k+1)T} n_c(t) \cos\psi(t,\underline{\alpha}')\, dt + \frac{1}{\sqrt{2}}\int_{KT}^{(k+1)T} n_s(t) \sin\psi(t,\underline{\alpha}')\, dt \tag{3.5}$$

Let

$$n_{c,k}(\alpha') = \frac{1}{T}\int_{kT}^{(k+1)T} n_c(t) \cos\psi(t,\underline{\alpha}')dt + \frac{1}{T}\int_{kT}^{(k+1)T} n_s(t) \sin\psi(t,\underline{\alpha}')dt \tag{3.6}$$

Eq. (3.5) can be rewritten as

$$\delta_{c,k}(\underline{\alpha}, \underline{\alpha}') = \frac{1}{2}\left\{\frac{2E}{T}\right\}^{1/2}\int_{kT}^{(k+1)T} \cos[\psi(t,\underline{\alpha}') - \psi(t,\underline{\alpha}) - \theta_i]dt + \frac{T}{\sqrt{2}}\, n_{c,k}(\underline{\alpha}') \tag{3.7}$$

Similarly, the quadrature correlation is obtained by substituting Eqs. (3.3) and (3.4) into (3.2b) and we obtain

$$\delta_{s,k}(\underline{\alpha},\underline{\alpha}') = \frac{1}{2}\left\{\frac{2E}{T}\right\}^{1/2}\int_{kT}^{(k+1)T} \sin[\psi(t,\underline{\alpha}') - \psi(t,\underline{\alpha}) - \theta_i]\, dt + \frac{T}{\sqrt{2}}\, n_{s,k}(\underline{\alpha}') \tag{3.8}$$

where

$$n_{s,k}(\underline{\alpha}') = \frac{1}{T}\int_{kT}^{(k+1)T} n_c(t)\sin\psi(t,\underline{\alpha}')dt - \frac{1}{T}\int_{kT}^{(k+1)T} n_s(t)\cos\psi(t,\underline{\alpha}')dt \qquad (3.9)$$

$n_{c,k}(\underline{\alpha}')$ and $n_{s,k}(\underline{\alpha}')$ as given by Eqs. (3.7) and (3.9) are both Gaussian with zero mean since $n_c(t)$ and $n_s(t)$ are zero mean and white Gaussian processes. Their variances can be shown to be

$$\text{VAR}\,[\,n_{c,k}(\underline{\alpha}')] = \frac{No}{2}\ \frac{1}{T}$$

$$\text{VAR}\,[\,n_{s,k}(\underline{\alpha}')] = \frac{No}{2}\ \frac{1}{T} \qquad (3.10)$$

Also, it can be shown that

$$E\,\{\,n_{c,k}\,(\underline{\alpha}')\,n_{s,k}\,(\underline{\alpha}')\,\} = 0 \qquad (3.11)$$

Hence, noise components of the inphase and quadrature partial likelihoods resulting from matched filtering of the received signal, matched to the particular signal corresponding to $\underline{\alpha}'$, are uncorrelated.

However, it can be shown that

$$E\,\{n_{c,k}(\underline{\alpha}')\,n_{c,k}(\underline{\alpha})\} = \frac{No}{2}\ \frac{1}{T}\ \frac{1}{T}\int_{kT}^{(k+1)T}\cos[\psi\,(t,\underline{\alpha}') - \psi\,(t,\underline{\alpha})]dt \qquad (3.12)$$

$$E\,\{n_{c,k}(\underline{\alpha}')\,n_{s,k}(\underline{\alpha})\} = -\frac{No}{2}\ \frac{1}{T}\ \frac{1}{T}\int_{kT}^{(k+1)T}\sin[\psi\,(t,\underline{\alpha}') - \psi\,(t,\underline{\alpha})]dt \qquad (3.13)$$

Hence, the noise component of the output of the filter matched to the inphase signal component for $\underline{\alpha}'$ may be correlated with that from the inphase matched filter for $\underline{\alpha}$ according to Eq. (3.12). The noise component of the output of the filter matched to the inphase component for $\underline{\alpha}'$ may also correlated with the noise component of the output of the quadrature matched filter for $\underline{\alpha}$ as indicated by Eq. (3.13).

The noise components of the outputs of the inphase and quadrature matched filters for different signals over a symbol interval as given by Eqs. (3.6) and (3.9) cannot be generated simply as independent Gaussian random variables. In order to generate in the simulation the filtered noise components at the inphase and quadrature matched filter outputs for different $\underline{\alpha}'$, the set of cosines and sines of the possible signal phases over an interval is first expanded using the Gram - Schmidt orthogonalization procedure [16,17]. The noise component from a filter matched to a particular inphase or quadrature signal component is then obtained by summing independent Gaussian random variates, which are weighed by their corresponding orthonormal coefficients.

Transmitted sequence $\underline{\alpha}$ is generated randomly one symbol at a time. The partial inphase and quadrature likelihoods for different possible transmitted symbols are computed according to Eqs. (3.7) and (3.8) with the noise components generated as outlined previously. The receiver then forms the equivalent likelihoods and estimates the corresponding maximum likelihood transmitted sequence $\underline{\alpha}'$ according to the decoding algorithm detailed in section 3.3 of previous report [4]. The random initial phase θ_i is also generated every hop as a random value uniformly distributed between 0 and 2π.

Although the decoding algorithm as described in [4] retains only one survivor for each state, simulation has shown that the performance of the receiver improves if the decoder keeps more than one survivor for each state during the decoding process. Only the MSK hop-by-hop noncoherent receiver has been simulated. Figs. 13 - 15 show how the error performance improves as the number of survivors kept for each state is varied from 1 to 4 for hop lengths 4, 16 and 64. It can be seen that keeping two survivors for each state would be sufficient. Keeping more than two survivors improves the error performance insignificantly. The simulation results for the hop-by-hop sequence estimation

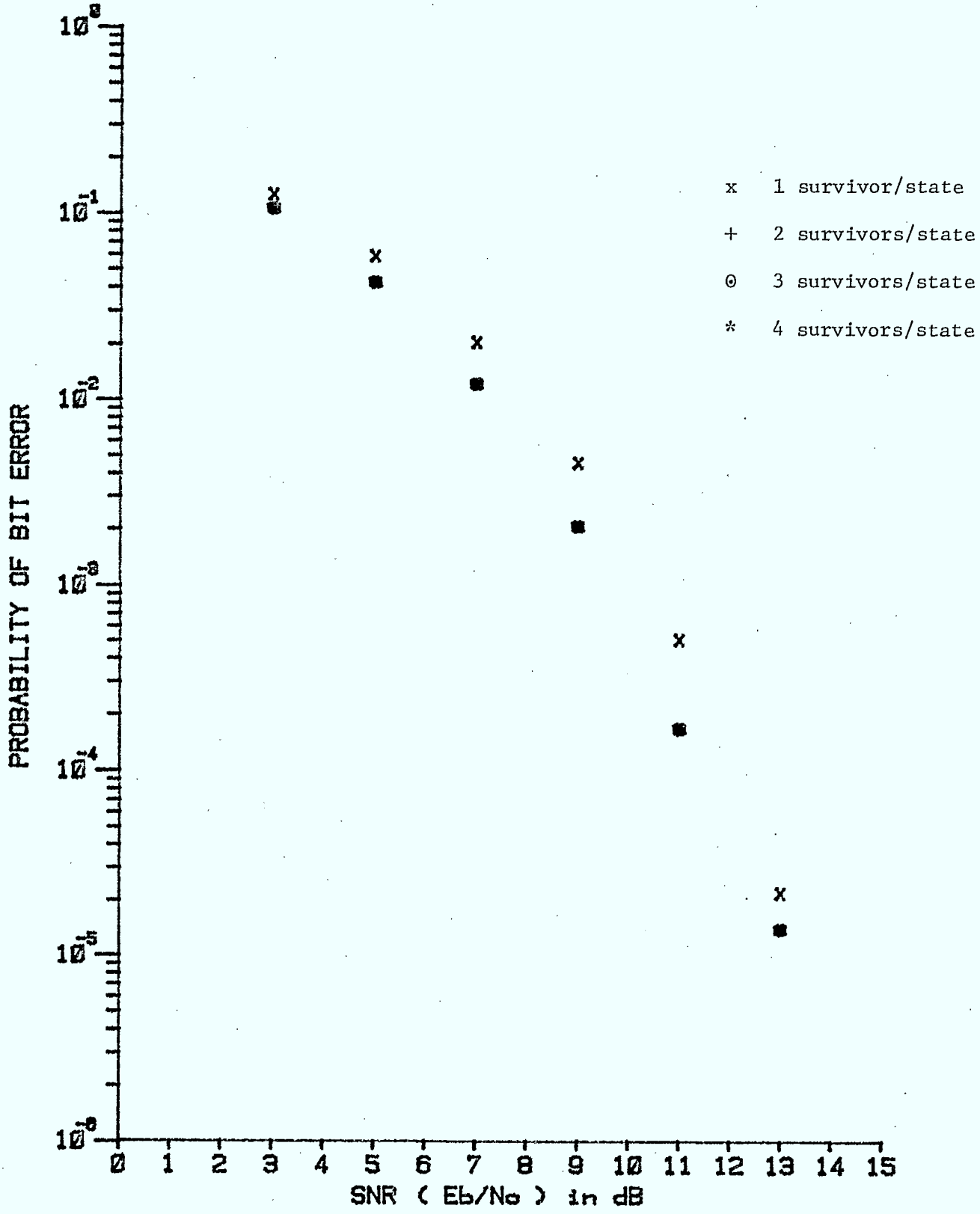


Figure 13: N=4, MSK simulation results.

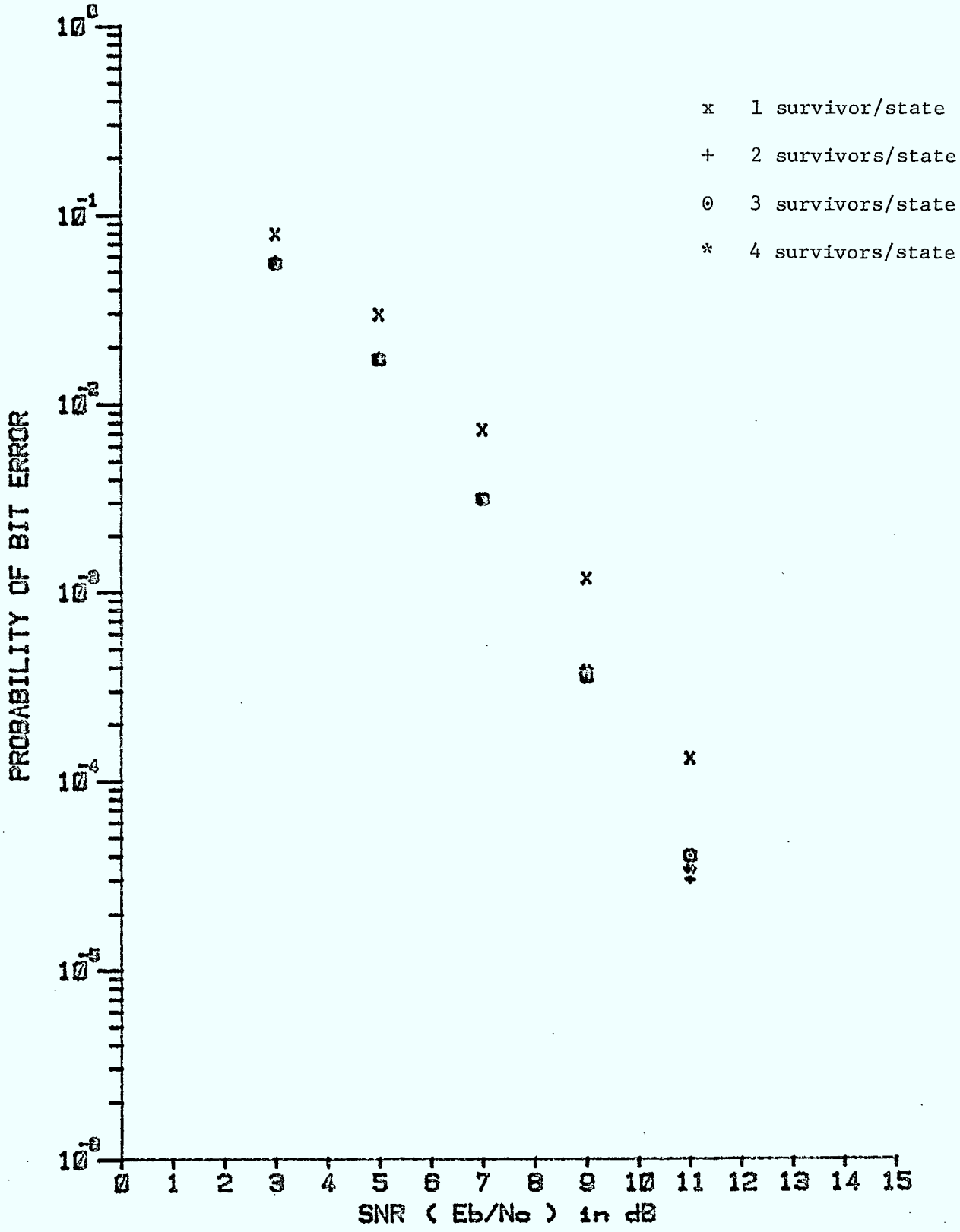


Figure 14: N=16, MSK simulation results.

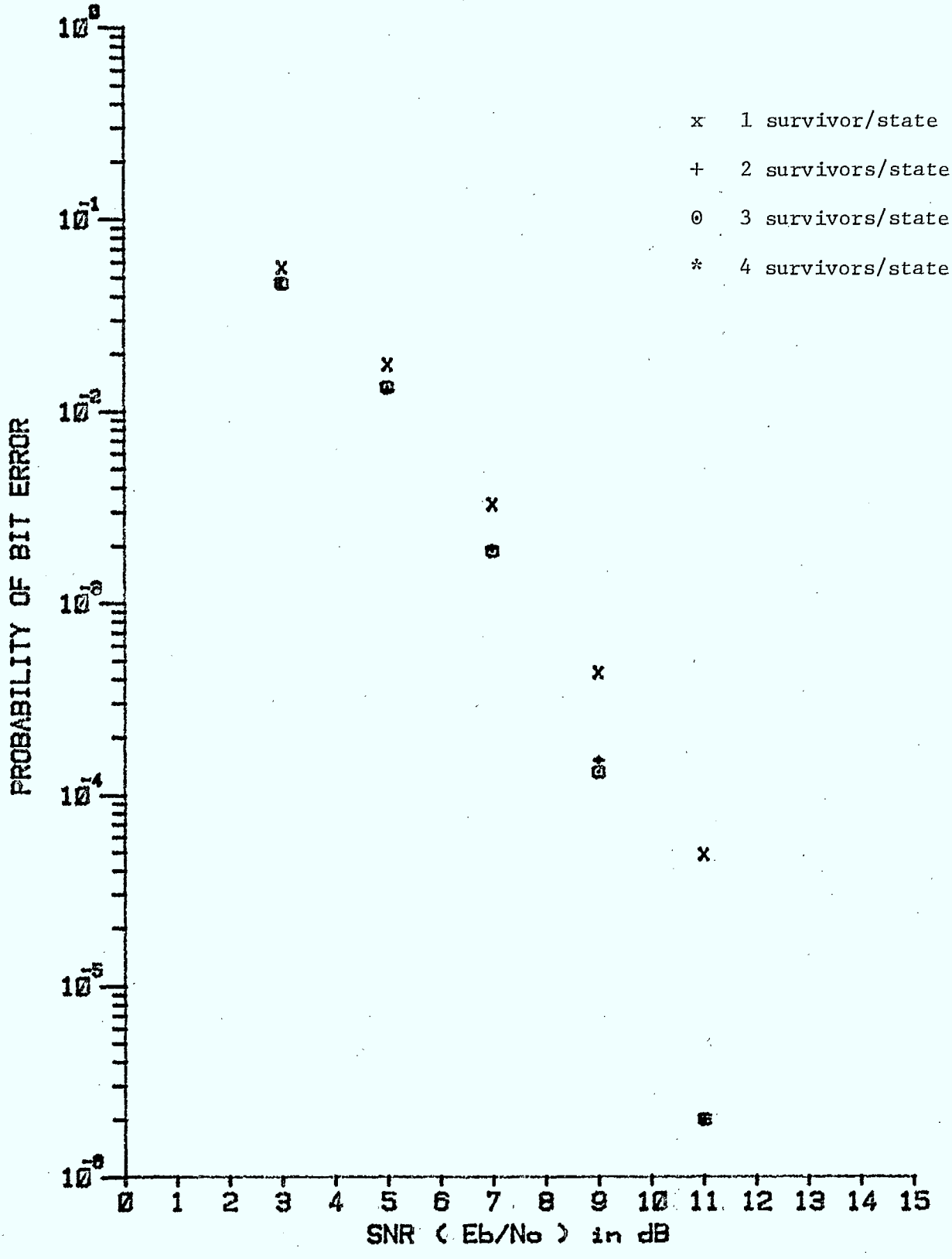


Figure 15: N=64 MSK simulation results.

receiver, which keeps two survivors for each state, is as shown in Fig. 16 for various hop interval lengths. The coherent MSK Viterbi Algorithm receiver performance is also shown. It can be seen that the error decreases with the increase in hop length but the error probability is always higher than the coherent MSK Viterbi receiver, as expected.

In Fig. 17, upper bounds evaluated for MSK with rectangular pulse shaping for N = 1, 4, 16 and 64, are plotted together with the simulation results in order to indicate the tightness of the upper bounds evaluated in the previous section. The curve for N=1 is actually the average probability of bit error rather than an error bound. It can be seen that the simulated error probability coincides with the theoretical result. At moderate to high SNR, the upper bound for N=4 is also tight, while for longer hop lengths as 16 and 64, the bounds are not as accurate.

4. CONCLUSIONS

In this report, upper bounds on the bit error probability have been evaluated for a variety of frequency-hopped correlative encoding schemes with rectangular and raised cosine baseband pulse shapings. It has been shown that the error performance of the hop-by-hop sequence estimation noncoherent receiver improves as the hop length increases. The simple MSK scheme has better performance than higher order correlative encoding schemes such as DMSK and TFM. Higher order correlative encoding schemes shows more pronounced error performance improvement for the same degree of increase in hop interval. A computer simulation study of the hop-by-hop sequence estimation noncoherent receiver has been carried out. It has been found that by keeping two survivors for each state, the error performance of the receiver achieves close to optimality using the same decoding algorithm described in the previous report. The simulation results indicate that the error bounds evaluated for short hop intervals are tight.

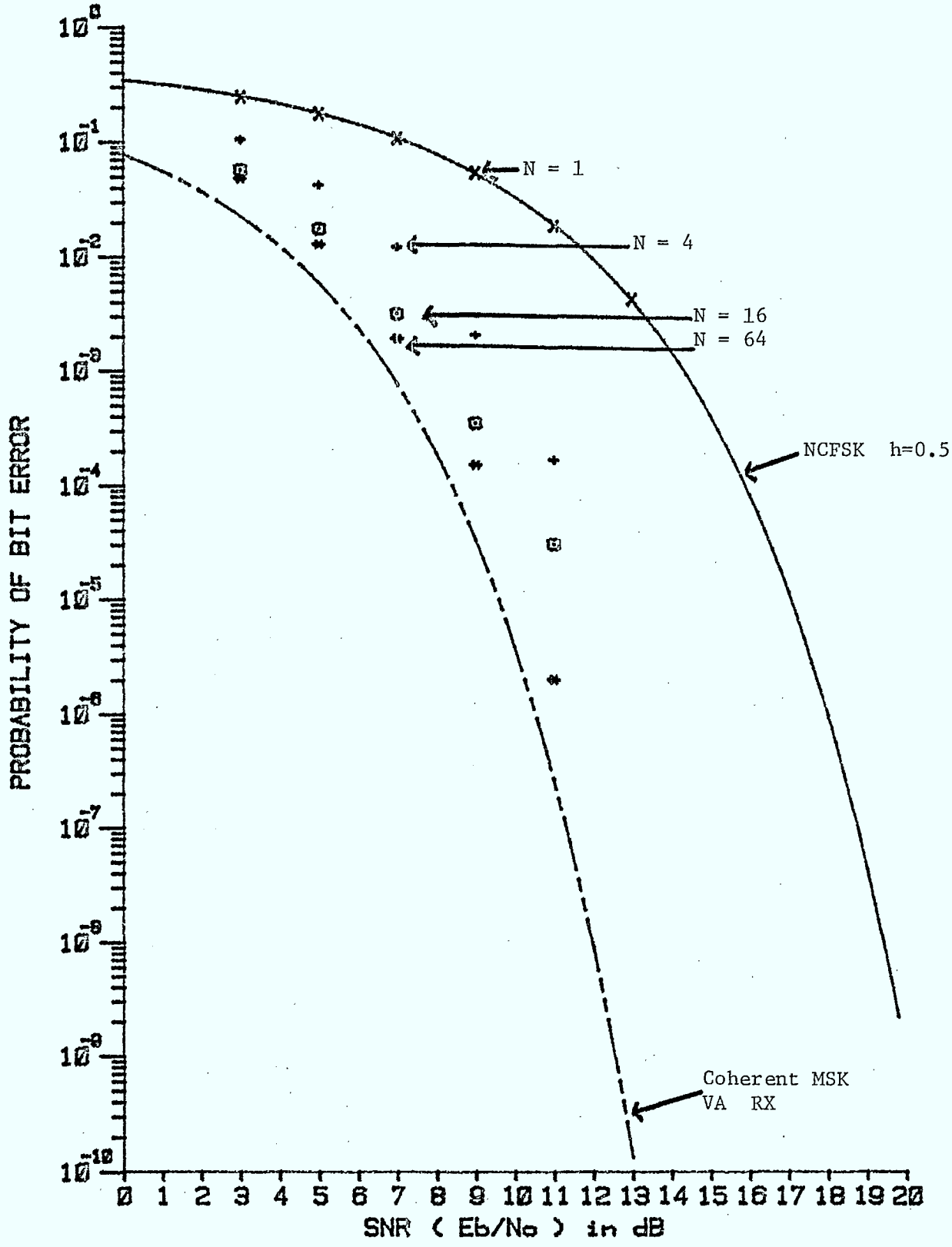


Figure 16: Simulation results for MSK

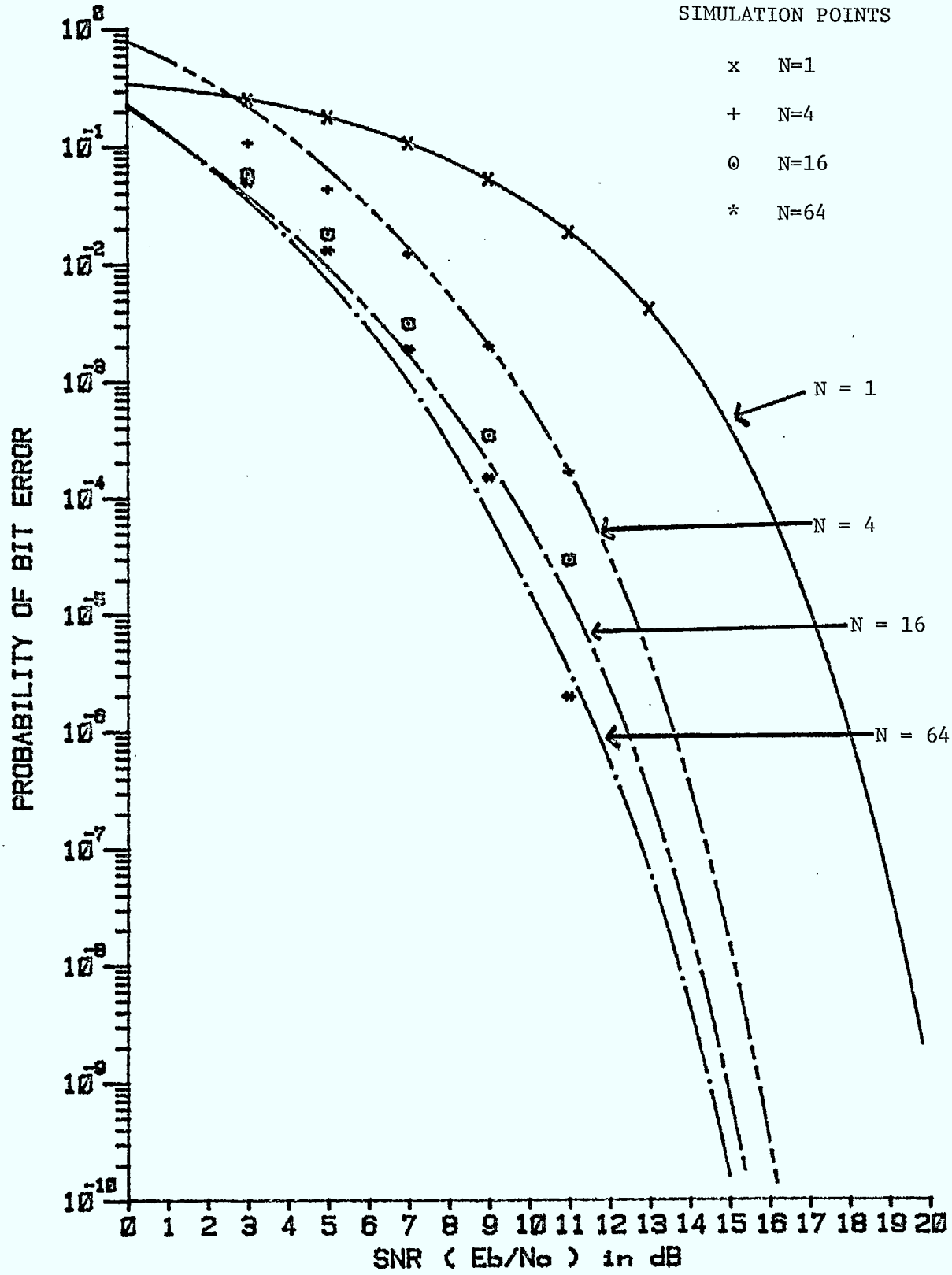


Figure 17: Upper bounds and simulation results for MSK.

REFERENCES

1. C.E. Cook and H.S. March, "An Introduction to Spread Spectrum", IEEE Communications Magazine, Vol. 21, No. 2, pp. 8-16, March 1983.
2. C.E. Sundberg, "Continuous Phase Modulation", IEEE Communications Magazine, Vol. 24, No. 4, pp. 25-38, April 1986.
3. P.H. Wittke, P.J. McLane, S.J. Simmons and Y.M. Lam, "A Bandwidth Efficient Frequency-Hopped Spread Spectrum Modulation Study", Dept. of Elect. Eng., Queen's University, Research Report No. 85-03, March 1985.
4. P.H. Wittke, P.J. McLane, Y.M. Lam and S.J. Simmons, "A Bandwidth Efficient Frequency-hopped Spread Spectrum Modulation Study", Dept. of Elect. Eng., Queen's University, Research Report No. 86-02, March 1986.
5. G.S. Deshpande and P.H. Wittke, "Correlative Encoded Digital FM", IEEE Trans. on Comm., Vol. COM-29, pp. 156-162, Feb. 1981.
6. P. Kabal and S. Pasupathy, "Partial-Response Signaling", IEEE Trans. on Comm., Vol. COM-23, pp. 921-934, Sept. 1975.
7. S. Pasupathy, "Minimum Shift Keying: A Spectrally Efficient Modulation", IEEE Communications Magazine, Vol. 17, No. 4, pp. 14-22, July 1979.
8. P.J. McLane, "The Viterbi Receiver for Correlative encoded MSK Signals", IEEE Trans. on Comm., Vol. COM-31, No. 2, pp. 290-295, Feb. 1983.
9. T. Aulin and C.E. Sundberg, "On Partially Coherent Detection of Digital Continuous Phase Modulation Signals", Technical Report TR-127, May 1979, Telecommunication Theory, University of Lund, Lund.
10. A. Svensson and C.E. Sundberg, "Error Probability Bounds for Noncoherently Detected CPM", Technical Report TR-183, August 1984, Telecommunication Theory, University of Lund,Lund.
11. S. Svensson and C.E. Sundberg, "On Error Probability for Several Types of Noncoherent Detection of CPM", IEEE GLOBCOM'84, Conference Record, pp. 22.5.1-22.5.7.

12. M. Schwartz, W.R. Bennett and S. Stein, Communication Systems and Techniques, McGraw-Hill, 1966.

13. S. Stein, "Unified Analysis of Certain Coherent and Noncoherent Binary Communication Systems", IEEE Trans. on Information Theory, Vol. IT-10, pp. 43-51, Jan. 1964.

14. W.F. McGee, "Another Recursive Method of Computing the Q Function", IEEE Trans. on Information Theory, Vol. IT-16, pp. 500-501, July 1970.

15. S. Parl. "A New Method of Calculating the Generalized Q Function", IEEE Trans. on Information Theory, Vol. IT-26, pp. 121-124, Jan. 1980.

16. J.M. Wozencraft and I.M. Jacobs, Principles of Communication Engineering, John Wiley and Sons, 1967.

17. V.K. Bhargava, D. Haccoun, R. Matyas and P.P. Nuspl, Digital Communications by Satellite, John Wiley and Sons, Inc., 1981.

Interim Progress Report

Part III

SYNCHRONIZATION ASPECTS OF HOPPED SPREAD SPECTRUM SYSTEMS

by

S.J. Simmons

PART 3 - SYNCHRONIZATION ASPECTS

This part of the report deals primarily with the acquisition (and tracking) of parameters required to allow reliable data detection at the satellite, that is, uplink synchronization. Since this synchronization process will require feedback to the user, the issues of downlink format and synchronization need to be addressed, but to a lesser extent. High-level data formats and associated protocols are not considered.

We first define the system under consideration along with the assumptions that are employed. A possible downlink synchronization procedure is described. Various aspects of uplink synchronization are then outlined, and options for system implementations and strategies are presented. The "simplest possible" system (having lowest synchronization complexity) is then described.

1. SYSTEM UNDER CONSIDERATION

1.1 General

This is principally a point-to-point communications system that is based on circuit switching. Call setup is performed over control channels imbedded in the uplink/downlink data streams [1] that enjoy full antijam protection.

Channel assignments may be fixed or reconfigurable by a central controller overseeing demand assignment. To keep things simple, we assume a fixed assignment scheme with dedicated uplink/downlink slots for each user. More flexible systems simply imply longer waits for access to satellite resources for synchronization purposes. We assume that when user A is not engaged in a call, the satellite places the data detections from user A's uplink slot into user A's downlink slot. This provides the feedback which is crucial to initial synchronization.

1.2 The uplink

Uplink users are arranged in an FDMA format, with each user employing a non-coherent M-ary modulation. Users hop their transmit frequencies (as a group) over a very wide band. The frequency synthesizer is controlled by k-bit blocks (for 2^k different frequencies) produced by a pseudo-random sequence generator with a period which can be in excess of several days. The hopping rate provides several hops per bit for low rate users and creates many bits per hop for medium data rate users.

Uplink beam sharing is a possibility [2], but is not considered here as it poses no additional fundamental problems.

1.3 The downlink

The downlink follows a TDMA format and the composite signal may or may not be frequency spread. If spectrum spreading is employed here, the first step at the receiver is to synchronize to this spreading sequence and despread. This is a conventional synchronization problem and so is not considered further here. In any event, downlink spreading may not be a requirement if data is protected by encryption.

The downlink TDMA format is assumed as in Figure 1. One frame is composed of the intervals during which the downlink antenna beam hops to each of the coverage zones. While dwelling on a zone, a conventional sync pattern is followed by data slots for the users as well as a common information slot in which the satellite may transmit control messages or data which facilitates initial synchronization. Users look for the sync pattern, then "read" data from their assigned time slots.

2. DOWNLINK SYNCHRONIZATION

We assume again that the downlink TDMA signal is not spread. Each user receives signal energy only during the time that the downlink beam is dwelling on

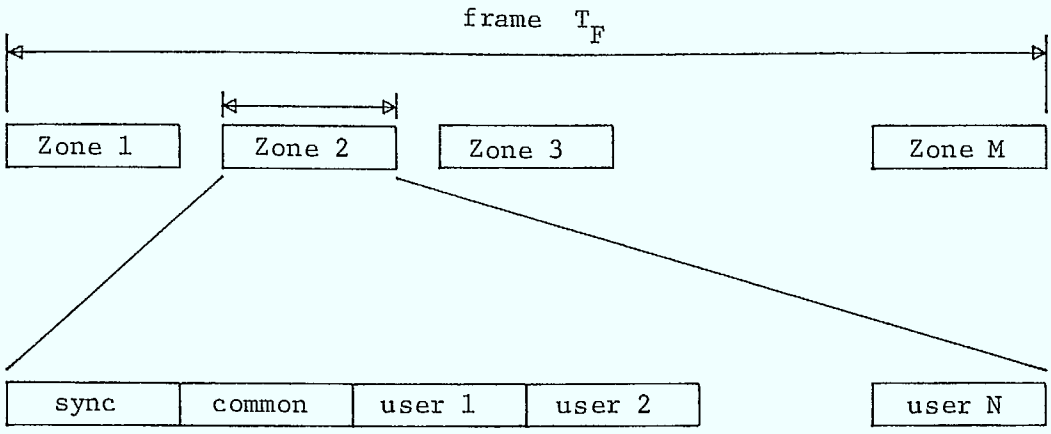


Figure 1. Downlink TDMA frame format.

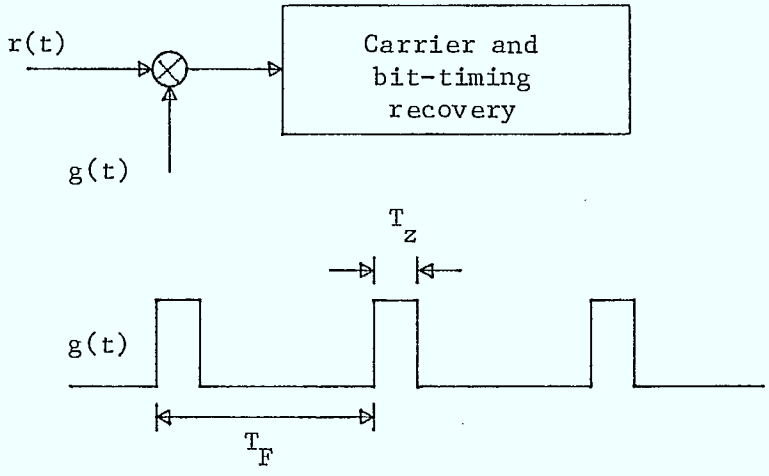


Figure 2. Gating for downlink synchronization.

the user's zone. For example, with 20 fully active zones, users see signal energy for about 5% of the time.

To perform initial acquisition, the following procedure may be employed. Users energy-detect over a sliding window of length T_z where T_z is the dwell time on the zone. Once received energy exceeds a threshold, the window is frozen and the windowed signal is passed to phase and bit-timing recovery circuits. Once these parameters are acquired, the window is expanded to $2T_z$ with the same centre position, and a search for the sync burst is begun using a sliding binary correlator [3]. When the sync burst is acquired, the user can identify his time slot and pick out the data destined for him. In addition, any common data supplied by the satellite can be recovered.

Note that unlike a conventional TDMA system having TDMA both on the uplink and on the downlink, we do not need bit-timing symbols dedicated to each user. All user data bursts are exactly aligned to common timing marks since these burst are formed by the satellite with a single timing reference. This also means that the signals from subsequent beam hops are easily combined in the carrier and bit-timing recovery loops. All that is needed is a simple gating waveform as in Figure 2.

3. UPLINK SYNCHRONIZATION: ACQUISITION

The following are identified as the parameters to be acquired (and tracked):

(a) Hopping sequence phase, i.e., the proper point in the long pseudo-random hopping pattern. (coarse sync)

(b) Hopping clock phase, i.e., the alignment of hop transitions of the uplink signal with those on board the satellite. (fine sync)

(c) Carrier frequency. There may be significant errors due to Doppler shifts with non-geosynchronous satellites, initial errors in the frequency synthesizers, and drift. (coarse and fine sync)

3.1 Hopping sequence phase/coarse carrier frequency

The acquisitions of these two parameters are tightly bound together. In effect, they create a two-dimensional search space over the regions of uncertainty. System implementation choices will determine the size of the uncertainty regions and ultimately the time for acquisition.

There appear to be two options with regard to initial uncertainty in hop sequence phase. We assume that the pseudo-random sequence is produced by a finite state machine realized as a clocked digital circuit. In the first option, this sequencer box is initially synchronized by physical connection to a master unit. Small errors in the clock frequency will cause phase error to accumulate from that point on (until the next synchronization with the satellite is achieved). There is, however, another attractive alternative. Since the state of the sequencer at any one time is implicit in the value of the internal storage elements (latches or flip-flops), it is possible to transmit this state to users on the downlink. By setting the latches or flip-flops to the same values in a replica sequencer circuit, a user acquires synchronization. Of course, the user must adjust for the delay in downlink transmission and for the uplink delay required for his signal to reach the satellite. If these delays are known within an uncertainty of T_p, then this determines the maximum initial error in the hopping sequence phase (error $< {}^{T_p}/_{T_h}$ hops where hopping rate $R_h = 1/T_h$). This uncertainty can be kept small by transmitting satellite ephemeris data in the "common" slot of the downlink frames to facilitate range calculations.

Such side information must be protected of course. This is ensured if the downlink is encrypted or if the sequencer circuit utilizes a secret key (bit-string) in the formation of the hopping pattern (for example a cipher-block chained encryption device [4]).

As far as carrier frequency uncertainty is concerned, this can be reduced by predictions of Doppler shift from satellite orbital calculations (for non-geosynchronous satellites). Accurate oscillator reference frequencies along with good stability will clearly also help.

3.1.1 Search Strategies

There are many search strategies which are possible [5, 6]. The two dimensional search space is divided into N cells separated in frequency by Δf_c and separated by one full hop time T_h.

A fast initial linear sweep over all cells with small dwell (observation) time at each can be performed, with possible acquisition detections being explored for longer times. If the initial sweep fails, it is repeated with the dwell time increased. As a variation, more time can be spent on those cells closest to the expected value of the sequence phase, and $\Delta f = 0$, since the likelihood of larger drifts and errors is correspondingly smaller. Alternatively, a more formal sequential probability ratio test (SPRT) can be employed. In the SPRT, we compute $\Gamma_K = p_s(\underline{r}_K)/p_n(\underline{r}_K)$ where p_s and p_n are, respectively, the distributions of received sequence $\underline{r}_K$ (of length K samples) given signal present, and noise-only present. This likelihood ration Γ_K is compared to upper and lower thresholds selected to produce desired values of probability of false alarm p_{fa} and of detection p_d. If the ratio falls between the thresholds, the test is repeated with the (K+1)'th sample r_{K+1} added. The choice of search strategy depends on the size of the initial uncertainty regions. If these are small, a single serial search with

detection threshold set for the desired p_{fa} and p_d would be used. For very large uncertainty regions, the SPRT would be desirable. For moderate uncertainty regions, a modified variable-dwell time search procedure would be sufficient.

A method for scanning through a range of hop sequence phases is now described. The arrangement is shown in Figure 3. The pseudo-random sequence generator (PNG) can be clocked by ϕ, the nominal clock, or by ϕ^{++}, a high speed clock. In addition, the clock input can be disabled so that the outputs of the PN generator do not change. A block of k bits is clocked separately into a buffer register that feeds the frequency synthesizer to select one of 2^k frequencies.

The procedure to be followed is now outlined. Assuming an initial uncertainty of $\pm H$ hops in hop sequence phase, we initially switch to ϕ^{++} to run the sequencer "ahead" by H hops. Since the nominal clock rate of the sequencer will be relatively low (at most kR_h) for the hop rates ($R_h \sim 20$ kHz) and number of frequencies ($k < 32$) of interest, a ϕ^{++} clock at ten times the frequency of ϕ should be well within the capabilities of modern digital circuits. While this run-up is being performed, the input register at the synthesizer is still being reloaded at the nominal hop rate R_h, so that the carrier frequency is still hopping. After the PNG has advanced H hops ahead of nominal, its clock input is returned to ϕ. Now we can drop back by one hop (relative to the satellite) by disabling the clock input for one hop, and then returning it to ϕ for an interval of I hops duration, the observation time at the current hop sequence phase. During this time, we observe the downlink return (which we have assumed is directed to us in a loopback mode) and either declare acquisition or continue the procedure to examine the next hop phase. The maximum observation interval to cover all 2H possible hop sequence phases is then $2H \cdot I \cdot T_h$ seconds. Of course, a

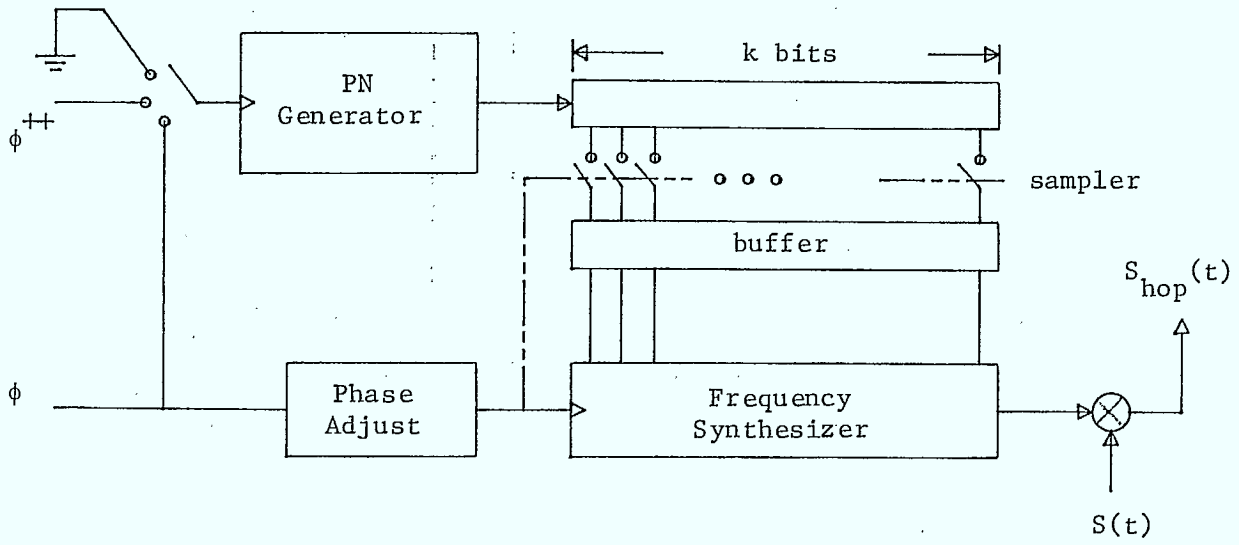


Figure 3. System for adjusting hopping sequence phase and clock.

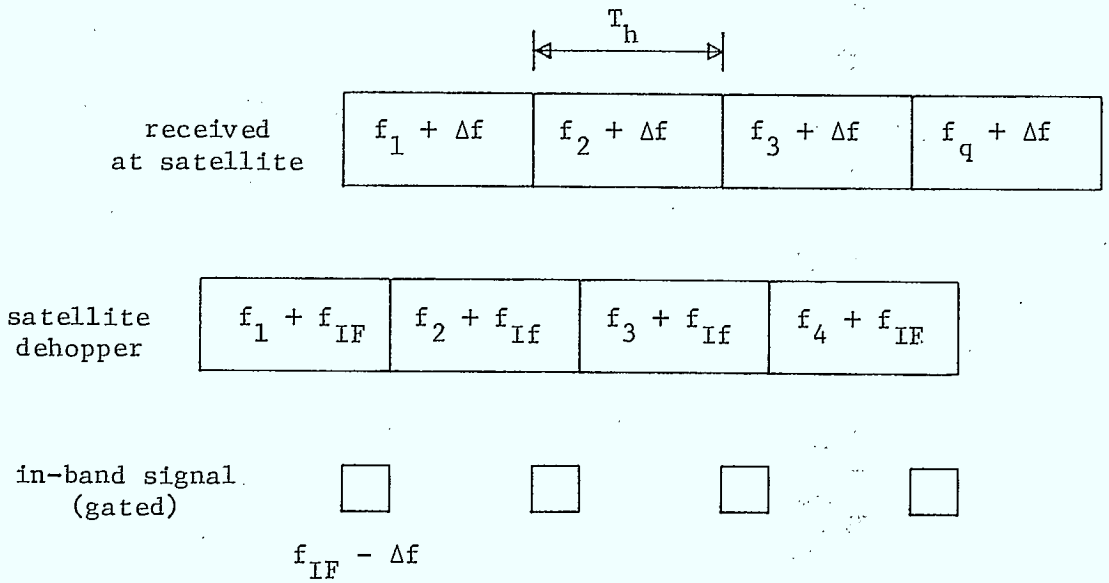


Figure 4. Effect of mis-aligned hop sequences.

choice of I is determined by the search strategy adopted. These I hops are assumed to span a search over several (or many) trial values for carrier centre frequency (separated by Δf_c).

3.1.2 Detection of Acquisition

At each trial value of hopping sequence phase and carrier centre frequency, we must decide if acquisition has been achieved. We concentrate on detection of hop sequence phase assuming initially that carrier centre frequency is accurate. Figure 4 depicts the alignment of the carrier hops with the satellite dehopping carrier, when the phases are within one hop. Notice that a serial search of possible hop sequence phases must produce an alignment at some trial phase that causes at least 1/2 of the signal energy to be properly dehopped by the satellite and pass through the satellite IF filters. In the worst case then, there will be a 3 dB loss of signal energy at proper alignment for this coarse acquisition phase.

Notice also that the transmitted user tone is "gated" before IF which will produce spectral spreading and some additional loss after filtering. Also, carrier centre frequency misalignment will reduce (or possibly eliminate) the signal component. We assume Δf_c is chosen to produce a 3 dB loss with worst case frequency alignment.

One option for the downlink return is to have the satellite transmit a hard decision (1 or 0) on the presence or absence of the agreed upon synchronizing tone (one of M tones). The decision threshold may be set to give some desired tone detection probability in the presence of maximum uplink fading and full band jamming. The ground user can then use these tone-detection decisions in a SPRT test, or, in a simpler search strategy, these decisions may simply be counted over an observation of I hops, and the count compared to a threshold.

As an example, the design threshold for this latter search might yield probabilities of tone detection, given tone present (with 3 + 3 = 6 dB loss due to hopping phase and carrier frequency misalignment) and tone absent, of $p(td|1) = .8$ and $p(td|0) = .1$, respectively. If we look for 5 or more detections in 10 hops, this will yield a false alarm probability of $P_{fa} \sim 10^{-3}$ and a probability of detection of acquisition of $P_d \sim .966$ (from standard binomial distribution calculations). Of course the detailed calculations of $p(td|0)$ and $p(td|1)$ in the presence of jamming will be complicated. With Gaussian noise only, these probabilities form a set of well known curves. The presence of partial-band jamming and frequency misalignment errors introduce further complications.

3.2 Hop clock phase fine carrier frequency

The procedures described above will yield a coarse synchronization to within one half of a hop and within $\Delta f_c/2$ carrier frequency error.

Achieving a finer alignment is complicated if the satellite sends only hard decisions about tone detections on the downlink. We wish to adjust the hop clock phase, and frequency synthesizer centre frequency, in small increments of Δt and Δf, and receive downlink data about the quality of the alignment. One possibility is to have the satellite send more than one bit per hop on the downlink where this data is a quality measure, for example, the output of an energy detector quantized to one of several levels. This may represent an expansion of the data rate normally assigned to a user's downlink data slot, but may be feasible since it is only needed for the brief and infrequent synchronization attempts.

Another possibility is to estimate the proper fine alignment from only hard-decision tone detection data as assumed above. In this case, the hop clock phase or carrier frequency may be stepped by Δ sequentially until a target rate of

missed detections (e.g., 8 of 10) is exceeded. The clock phase or carrier frequency may then be stepped in the opposite direction until missed detections again occur in excess of this target rate. Proper alignment may then be estimated to lie exactly half way between these two extremes. Of course, this procedure will take longer than an alternative employing multiple-bit alignment quality data transmitted by the satellite. The procedure is aided, however, by the fact that there should be a reasonably sharp threshold effect at which the rate of missed detections rises quickly with increased misalignment.

4. UPLINK SYNCHRONIZATION: TRACKING

Once acquired, parameters will drift toward loss of synchronization during normal data transmission. This may be avoided by tracking, that is, monitoring and adjusting the fine alignment. In data transmission mode, we can expect only hard decisions on the M tones to be available on the downlink, making tracking difficult. This contrasts sharply with the tracking problem in a conventional spread spectrum system in which we have full access to the signal emerging from the dehopper. In this conventional system, a tracking loop is employed [5].

It is, however, possible to imagine a tracking loop in the satellite system under consideration. To do this, low rate supplementary channels could be paired with each user's data slot in the downlink frames. In these channels, the satellite would provide alignment-quality data (e.g., energy detector output quantized to b bits) averaged over many hops. The ground terminal would use this data in a tau-dither scheme to make adjustments to hopping clock phase and carrier centre frequency. The time constant of the loop would include twice the propagation delay to the satellite. Because the loop time constant can be large compared to the hop duration T_h (assuming only slow drifts), the data rate of these supplementary channels need only be a fraction of the user data rate, and represent a small

overhead. This scheme does, however, require more complex electronics in the satellite.

One final alternative is to imbed parity check bits in the user data streams. If user A is talking to user B, and drifts toward loss of synchronization, this will show up as errors in the data. User B can then request that user A adjust his parameters to try for better alignment. This is also a feedback loop, but it has a much slower response time than the previous scheme. It also requires an extra level of synchronization in the data streams to identify which bits are data and which are check bits.

5. SIMPLEST SYSTEM

Given the previous discussion, it is now possible to identify a system that is in some sense ideal, and which has the "simplest" synchronization scheme in that searching is not necessary (at least for low rate users). This system has the following attributes:

(a) All users know the round-trip delay to the satellite to within a small fraction of a hop. At a hopping rate of 20 KHz, this implies range accuracy on the order of 1 km. Range determination is aided by satellite ephemeris data provided in the "common data" downlink slots.

(b) The internal state of the pseudo-random sequencer in the satellite is periodically transmitted on the downlink. This may be done with one bit per frame. Users can restart their replica sequencers and adjust for round-trip delay.

(c) The satellite hopping clock phase is aligned with the boundaries of the frames of the TDMA downlink. Note that this can only be exploited given (a).

(d) Satellite current centre frequency value is digitized and also sent in the common downlink slot (again 1 bit per frame may be sufficient). Users compare this to their own reference and adjust accordingly.

(e) Satellite motion is known accurately so that Doppler shifts can be calculated and compensation applied.

(f) As an alternative to (d) and (e), the satellite can provide continuous data on alignment quality of the carrier frequencies over low rate supplementary downlink slots for each user. This may be used both for acquisition and tracking.

This system is not intended as a proposal, it simply illustrates a baseline system from which strategic retreats can be made.

REFERENCES

[1] D.P. Kolba, "Generalized control and networking for EHF Satellite Communication Systems", AIAA 9th Communications Satellite Systems Conference, San Diego, CA., March 1982.

[2] D.P. Kolba, "System aspects of scanning beams for widely distributed users", International Telemetering Conference, San Diego, CA., October 1981.

[3] W.W. Wu, "Elements of Digital Satellite Communication: Volume I", Computer Science Press, Rockville, Maryland, 1984.

[4] D.E. Denning, "Cryptography and Data Security", Addison-Wesley, Reading, MA., 1982.

[5] R.E. Ziemer and R.L. Peterson, "Digital Communications and Spread Spectrum Systems", Macmillan, New York, N.Y., 1985.

[6] M.K. Simon, J.K. Omura, R.A. Scholtz and B.K. Levitt, "Spread Spectrum Communications: Volume III", Computer Science Press, Rockville, Maryland, 1985.

Interim Progress Report

Part IV

IDEAL PERFORMANCE OF INTERCEPTION RECEIVERS FOR FREQUENCY HOPPED SPREAD SPECTRUM SIGNALS

by

W. Hopkins,

P.J. McLane.

PART IV IDEAL PERFORMANCE OF INTERCEPTION RECEIVERS FOR FREQUENCY HOPPED SPREAD SPECTRUM SIGNALS

1. INTRODUCTION

Spread spectrum communications had been discussed with increasing frequency in the open literature in the last few years [1,2,3,4]. One of the attractive features of spread spectrum communications, from a user's point of view, is the difficulty of interception.

This report examines this feature from the viewpoint of an unfriendly interceptor. Given that an unknown spread spectrum signal is being transmitted, can it be intercepted by a receiver which has no knowledge of the sequence code? More precisely, what is the probability, given that a signal is being transmitted, that the receiver will be able to detect it? Furthermore, if the receiver has intercepted a spread spectrum signal, is it possible to determine if the source is moving? In other words, can a doppler shift be determined?

To answer the questions, it is necessary to determine the performance limits and optimum tradeoffs of such a receiver under ideal conditions. This appears to be a good starting point given that the interception problem is new to the present contract.

This report begins with a derivation of two interception receivers, assuming ideal conditions. It then establishes the performance limits of these receivers, first for only one or two transmission frequencies and then extends the results to a large number of transmission frequencies. Another version of the optimum receiver is developed and analyzed when a priori knowledge is available. The effects of frequency offset on the performance of the optimum receiver, which later will lead to a study of Doppler shift estimation, is also discussed. The report concludes with a discussion of future work.

2. DERIVATION OF INTERCEPTION RECEIVERS

2.1 Assumptions

In this section, two receivers will be discussed. For both, the following assumptions were initially made.

First, the bandwidth and duration of the transmitted signals are known exactly. Second, the receiver knows which frequencies are used by the transmitter and there is neither phase nor frequency distortion or offsets. That is, the receiver is perfectly coherent. The frequencies have a discrete distribution, are spaced an equal distance apart (the inverse of twice the signal duration) and arrive with equal probability. Resources are such that the receiver can contain a bank of filters, each one fixed to be centered at one of the transmission frequencies and matched to the signal bandwidth. To understand why the latter assumption is unrealistic, note that the transmission spectrum is typically 1 GHz wide and each signal has a bandwidth of 20 kHz. If there are 100 users, this implies that 500 filters will be required under conditions of group frequency hopping. Understandably, most interceptors will prefer to use far fewer filters. We note that a large number of filters can be represented using SAW Fourier processors. However, with these assumptions, best performance is obtained and represents the absolute limit of performance.

The only distortion to the signal is caused by additive white Gaussian noise.

Therefore the signal received is

$$H_1: r(t) = A \cos \omega_i t + n(t) \qquad 0 \leq t \leq T$$

if a signal is sent and (1)

$$H_o: r(t) = n(t) \qquad 0 \leq t \leq T$$

otherwise, where n(t) has zero mean and spectral height equal to $N_o/2$ W/Hz.

2.2 Maximum Likelihood Receiver

A maximum likelihood receiver, or a simple pulse detection system as discussed by Dillard [2], consists of a bank of filters, each followed by a decision making device. Each filter - decision pair can be viewed as a single detector (see Figure 1). Each detector makes a decision, declaring a signal in its band if the filter output is greater than a given threshold level, and none present otherwise. If any one of the detectors declares a signal present, the receiver decides that a signal has been received.

The analysis of its performance is quite simple. For the single frequency case, the probability of detection is:

$$Q_D = Q\left[\frac{\ell n\ \eta}{d} + \frac{d}{2}\right] \tag{2}$$

and the probability of false alarm is:

$$Q_F = Q\left[\frac{\ell n\ \eta}{d} - \frac{d}{2}\right] \tag{3}$$

where η is the decision threshold, $d = \left[\frac{2E_s}{N_o}\right]^{\frac{1}{2}}$, E_s is the energy of the received signal and

$$Q(x) = \int_x^{\infty} \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[-\frac{t^2}{2\sigma^2}\right] dt$$

When there is more than one transmission frequency, the case of interest, an analytical approach [4] gives

P_D = Probability of detection of the receiver

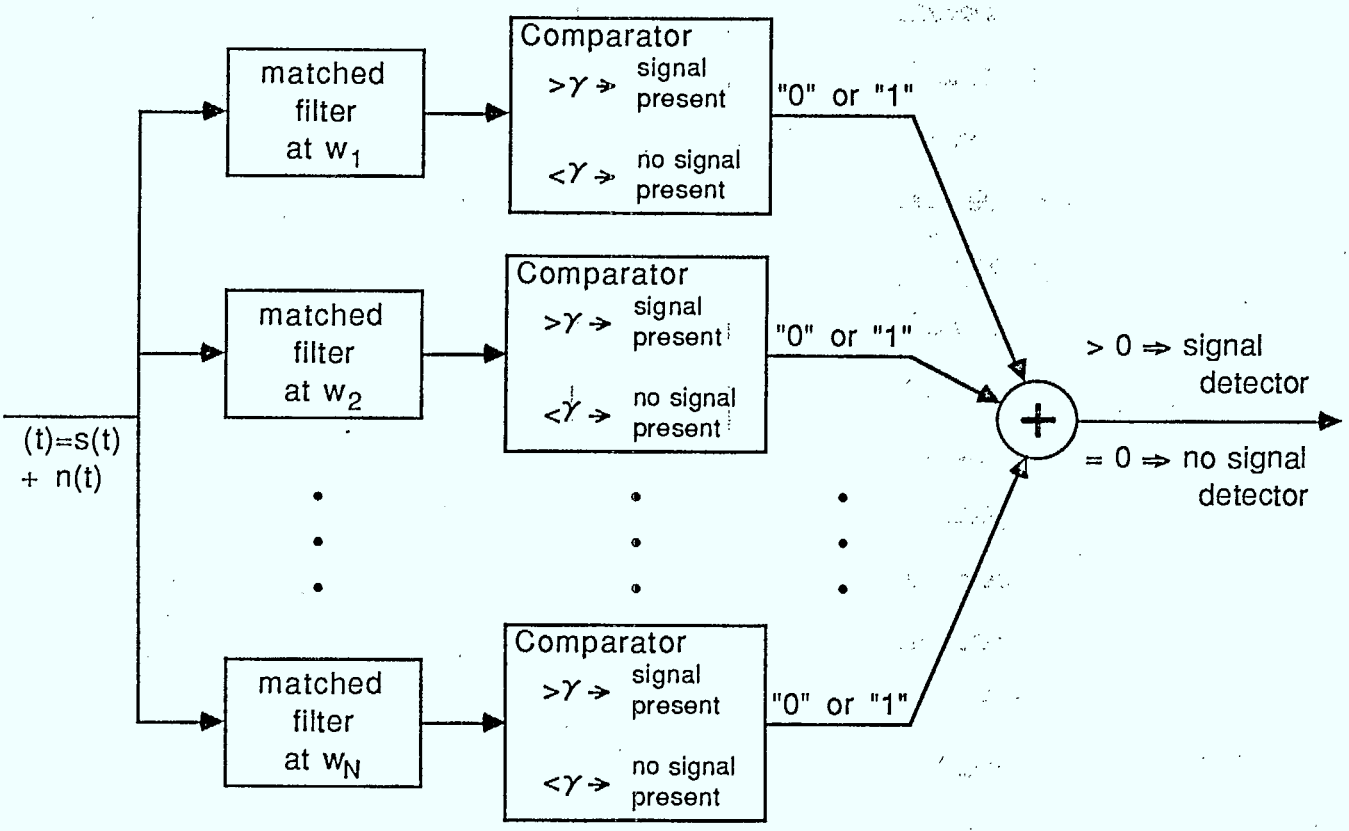
$$= 1 - (1 - Q_D)(1 - Q_F)^{N-1} \tag{4}$$

and

P_F = Probability of false alarm of the receiver

$$= 1 - (1 - Q_F)^N \tag{5}$$

where N is the number of detectors in the receiver.



Maximum Likelihood Detector

Figure 1: The maximum likelihood energy detector for one of N sinusoids in noise.

2.3 Optimum Detector

The optimum or average - likelihood receiver is based on a derivation of the generalized likelihood ratio test [8].

To determine the optimum detection test, two probability density functions must be found. The first is that of the received signal when no signal has been sent, which will be referred to as hypothesis 0 (H_o). The second is that of the received signal when a signal has been transmitted and will be referred to as hypothesis 1 (H_1).

The probability density function of the received signal under H_o is

$$P_r(R|H_o) = \frac{1}{(\pi N_o)^{1/2}} \exp\left[-\frac{\int_o^T r^2(t)\,dt}{N_o}\right] \tag{6}$$

The received signal, under H_1, has the probability density function

$$P_r(R|H_1) = \int_{\omega_{min}}^{\omega_{max}} \frac{P_\omega(\omega)}{(\pi N_o)^{1/2}} \exp\left[-\frac{1}{N_o}\left[\int_o^T r^2(t)\,dt - 2A\int_o^T r(t)\cos\omega dt\right]\right] d\omega \tag{7}$$

where $P_\omega(\omega) = \frac{1}{N}\sum_{i=1}^{N} \delta(\omega - \omega_i)$ is the probability density function of the transmission frequency under the assumptions made earlier. That is, the unknown signal frequency is taken to be discretely distributed over the hopped band.

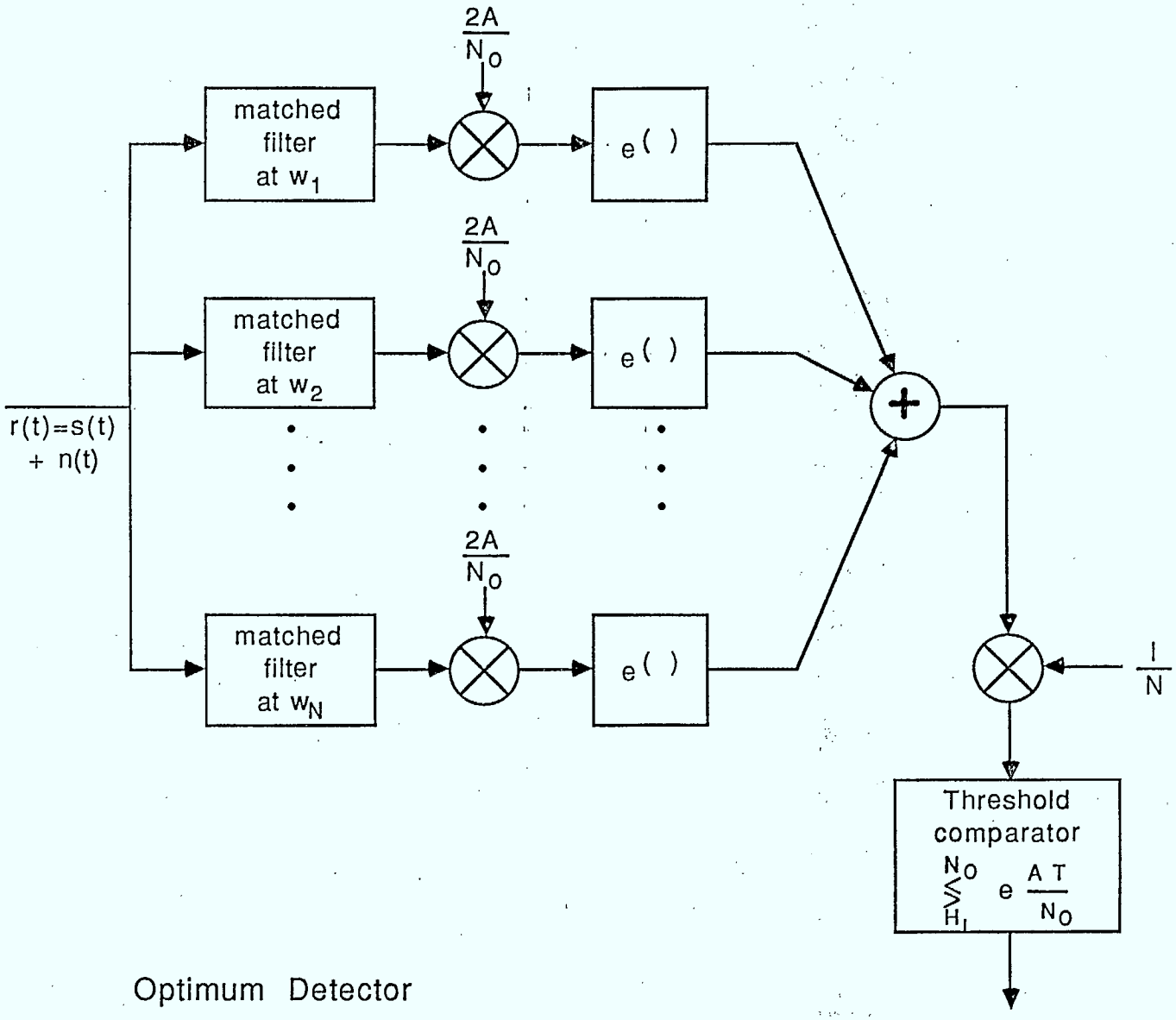
Taking the ratio of the two probability density functions gives the likelihood ratio test

$$L(r) = \frac{P_r(R|H_1)}{P_r(R|H_o)} \underset{H_o}{\overset{H_1}{\gtrless}} \eta \tag{8}$$

Performing the necessary manipulations, this becomes

$$L(r) = \frac{1}{N}\sum_{i=1}^{N} \exp\left[\frac{2A\int_o^T r(t)\cos\omega_i t dt}{N_o}\right] \underset{H_o}{\overset{H_1}{\gtrless}} \eta \exp\left[\frac{A^2T}{2N_o}\right] \tag{9}$$

the optimum receiver, therefore, takes the exponent of the filter outputs, sums them and compares the sum to the appropriate threshold as shown in Figure 2.



Optimum Detector

Figure 2: The optimum energy detector for a sinusoid of unknown discrete frequency in noise.

Individual decisions are not made at each frequency as is the case in Figure 1. The objective of the average likelihood receiver is just to detect energy relative to the assumed frequency grid, $\{\omega_\ell\}$. Note that in both receivers coherent detection is assumed. This is an idealization in a frequency hopped system.

For the single frequency case, the probabilities of false alarm and detection are identical to those given in equations (2) and (3) for the maximum likelihood receiver, as would be expected. However, the analysis for N greater than one is far from simple and a closed form expression is not available. Note that

$\alpha_i = \frac{2A}{N_o} \int_o^T r(t) \cos\omega_i t dt$ is a normal random variable with probability density function

$$P_\alpha(\alpha_i|H_1) = \frac{1}{(2\pi d^2)^{1/2}} \exp\left[-\frac{\alpha_i^2}{2d^2}\right] \tag{10}$$

and

$$P_\alpha(\alpha_i|H_1) = \frac{1}{(2\pi d^2)^{1/2}} \exp\left[-\frac{(\alpha_i - d^2)^2}{2d^2}\right] \tag{11}$$

where, as before, $d^2 = 2E_s/No$.

Therefore, the problem becomes an attempt to find the distribution function of a sum of N lognormal variables.

2.4 Analysis of the ALT Receiver for 2 Transmission Frequencies

To begin the analysis of the optimum receiver (ALT), we start small and assume that there are only two transmission frequencies. Under this circumstances we want to evaluate the probability that $L = \frac{1}{2} \exp\left[-\frac{d^2}{2}\right] \sum_{i=1}^{2} e^{\alpha_i}$ is greater than some threshold η under each hypothesis. Equivalently, we wish to find the probability that

$$e^{\alpha_1} + e^{\alpha_2} \geq 2\eta \exp\left[-\frac{d^2}{2}\right] = b$$

There are two exact approaches to this problem. The first finds the conditional probability, $\Pr[e^{\alpha_1} + e^{\alpha_2} \geq b | \alpha_2]$, then averages over α_2 to find the desired probabilities. Under H_o, both α_i's have zero mean. Thus, the probability of false alarm is

$$\Pr[e^{\alpha_1} + e^{\alpha_2} \geq b | \alpha_2, H_o] = \begin{cases} 1 & \alpha_2 \geq \ell n b \\ Q\left[\frac{\ell n (b - e^{\alpha_2})}{d}\right] & \alpha_2 < \ell n b \end{cases} \quad (12)$$

and since $P_F = \int_{-\infty}^{\infty} \Pr[e^{\alpha_1} + e^{\alpha_2} \geq b | \alpha_2] \Pr[\alpha_2]\, d\alpha_2$, after some work, the following expression can be found.

$$P_F = Q\left[\frac{\ell n\, b}{d}\right] + \int_{-\infty}^{\ell n\, b} \frac{1}{(2\pi d^2)^{1/2}} \exp\left[-\frac{\alpha_2^2}{2d^2}\right] Q\left[\frac{\ell n (b - e^{\alpha_2})}{d}\right] d\alpha_2 \quad (13)$$

The probability of detection can be found in a similar manner. However, since one detector, say #1, will have $r(t) = A\cos\omega_1 t + n(t)$ and the other will have $r(t) = n(t)$ only the probability density function of α_2 will be unchanged. Since only the noise is a random variable,

$$P_{\alpha_1}(\alpha_1) = \frac{1}{(2\pi d^2)^{1/2}} \exp\left[-\frac{(\alpha_1 - d^2)^2}{2d^2}\right] \quad (14)$$

Therefore,

$$\Pr[e^{\alpha_1} + e^{\alpha_2} \geq b | \alpha_2, H_1] = \begin{cases} 1 & \alpha_2 \geq \ell n b \\ Q\left[\frac{\ell n (b - e^{\alpha_2})}{d} - d\right] & \alpha_2 < \ell n b \end{cases} \quad (15)$$

and so,

$$P_D = Q\left[\frac{\ell n\, b}{d}\right] + \int_{-\infty}^{\ell n\, b} \frac{1}{(2\pi d^2)^{1/2}} \exp\left[-\frac{\alpha_2^2}{2d^2}\right] Q\left[\frac{\ell n (b - e^{\alpha_2})}{d} - d\right] d\alpha_2 \quad (16a)$$

or, if we average over α_1 instead an alternate form is found:

$$P_D = Q\left[\frac{\ell n\ b}{d} - d\right] + \int_{-\infty}^{\ell n\ b} \frac{1}{(2\pi d^2)^{1/2}} \exp\left[-\frac{(\alpha_2 - d^2)^2}{2d^2}\right] Q\left[\frac{\ell n(b - e^{\alpha_1})}{d}\right] d\alpha_1 \qquad (16b)$$

The second approach attempts to find the distribution function of L under each hypothesis. Denote $\ell_i = e^{\alpha_i}$. Under H_o

$$P_l(\ell_l | H_o) = \frac{1}{\ell_i (2\pi d^2)^{1/2}} \exp\left[\frac{-(\ell n\ \ell_i)^2}{2d^2}\right] \qquad (17)$$

and $P_l(\ell | H_o) = P_l(\ell_1 | H_o) \otimes P_l(\ell_2 | H_o)$ where $\otimes$ denotes convolution. Finally,

$$P_F = \int_\eta^\infty P_l(\ell | H_o)\, d\ell \qquad (18a)$$

$$P_D = \int_\eta^\infty P_l(\ell | H_1)\, d\ell \qquad (18b)$$

Performing the necessary calculations, equations (13) and (16) were again obtained. Therefore, these expressions for the probabilities of detection and false alarm have been confirmed by alternate methods.

To evaluate this integral, Gauss Quadrature numerical integration was used. A plot of the probability of detection versus the SNR for different false alarm probabilities for a single transmission frequency and two transmission frequencies is given in Figure 3.

2.5 Comparison of Results to Maximum Likelihood Receiver:

The receiver operating characteristic for the maximum likelihood test was found to be very close to that for the ALT, showing a slightly suboptimum performance for low SNR and high P_F but matching the performance of the average likelihood test receiver for higher SNR and lower P_F. An example is given in Table 1 for SNR = 3 dB. The results for one frequency are also included.

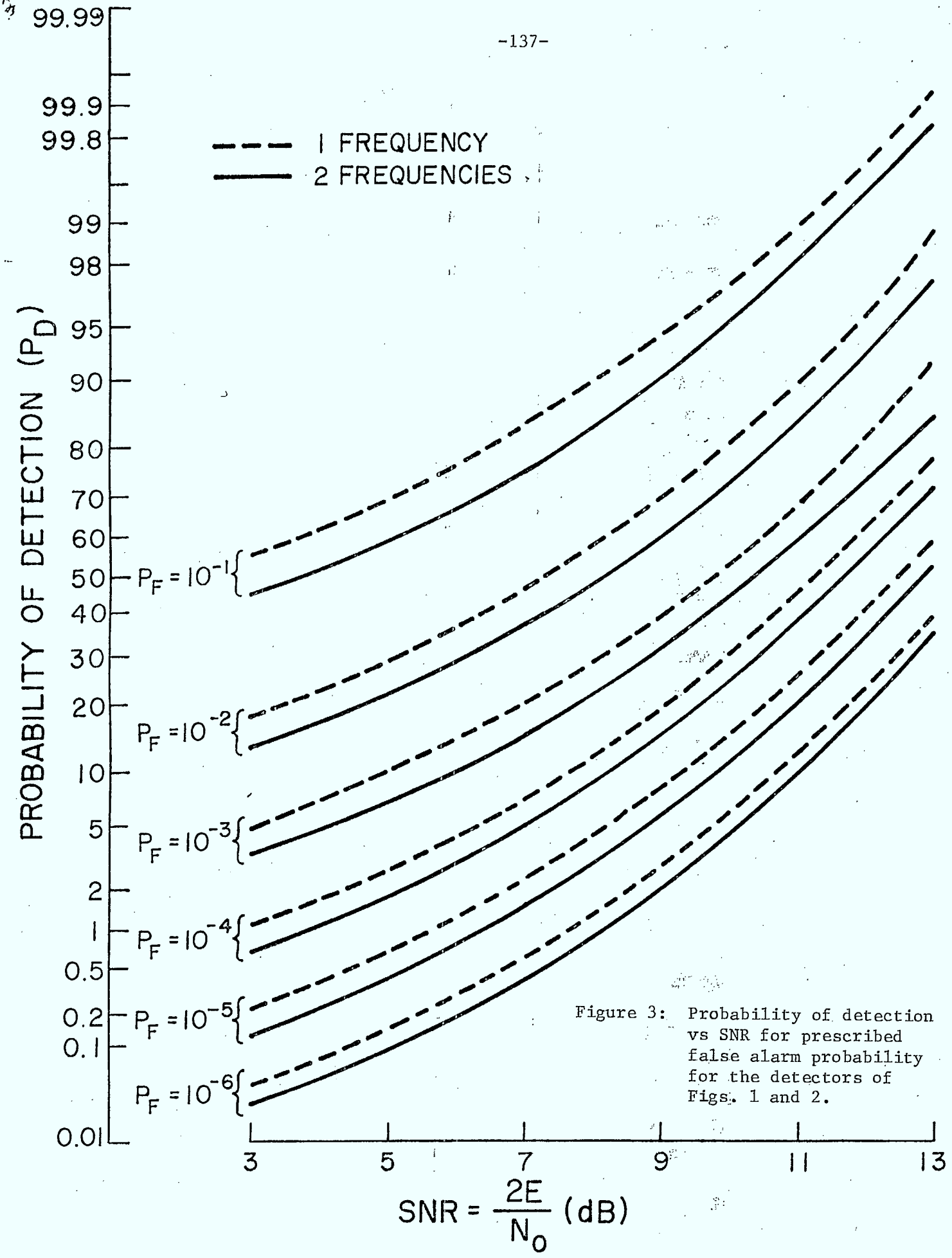


Figure 3: Probability of detection vs SNR for prescribed false alarm probability for the detectors of Figs. 1 and 2.

Table 1 Probability of detection for SNR = 3 dB

P_F	1 freq.	2 freq.	
		MLT	ALT
10^{-1}	.552	.443	.448
10^{-2}	.180	.127	.128
10^{-3}	.0467	.0307	.0308
10^{-4}	.0105	6.66×10^{-2}	6.66×10^{-2}
10^{-5}	2.17×10^{-3}	1.33×10^{-3}	1.33×10^{-3}
10^{-6}	5.96×10^{-4}	2.52×10^{-4}	2.52×10^{-4}

For the case of two frequencies, under ideal conditions, the two receivers appear to be roughly equivalent. From Figure 3, it can be seen that under ideal conditions, for low SNR, the probability of detection is quite low, particularly for $P_F = 10^{-3}$, the range of interest. It is necessary to have a SNR of 11 dB before the probability of detection exceeds 0.5. Under the assumed conditions, it would be possible to detect a transmitted signal only a small fraction of the time since, typically, we are concerned with low SNR values.

3. FREQUENCY UNCERTAINTY

It is of interest, since we wish to determine the effects of Doppler shift on the detector output, to consider the case of frequency offset. In the previous section, we assumed that the frequency separation is $\frac{1}{2T}$, and the received signal equals one of the filter center frequencies, $\omega_r = \omega_i$. Suppose instead that the received frequency, $\omega_r = \omega_i + \frac{\delta}{2T} = \frac{2\pi k_i + \delta}{2T}$ where $\delta << k_i$.

The probability of false alarm will be unchanged since it is independent of the transmission frequencies. Therefore, we need only examine the probability of

detection. Previously, under H_1, α_1 had a mean equal to the SNR, d^2, α_2 had a mean equal to zero and both had a variance equal to the SNR. Now, however,

$$\alpha_1 = \frac{2A}{N_o} \int_o^T r(t) \cos\omega_i t \, dt$$

$$= \frac{2A}{N_o} \int_o^T \left[A\cos \frac{\pi t}{T} (k_i + 2\delta) + n(t) \right] \cos \frac{k_i \pi t}{T} dt$$

$$= \frac{2A}{N_o} \left[\frac{\beta_1 AT}{2} + \int_o^T \eta(t) \cos\omega_i t \, dt \right] \qquad (19)$$

where $\beta_1 = \frac{2(k_1 + \delta)}{2k_1 + \delta} \frac{\sin\pi\delta}{\pi\delta}$. However, as stated previously, $k_1 >> \delta$ so $\beta_1 \cong \frac{\sin\pi\delta}{\pi\delta}$ $= \text{sinc}\delta$. Therefore, α_1 now is normal with a mean of $d^2 \text{sinc}\delta$.

Similarly, α_2 is found to be normal with a mean of $d^2 \text{sinc}(\delta-1)$. In Figure 4, for false alarm probability of 10^{-3}, the probability of detection is plotted against the frequency deviation δ. When $\delta = 0.0$ then $\omega_r = \omega_1$ and when $\delta = 1.0$, $\omega_r = \omega_2$. As expected, the probability of detection is a maximum at these points. When the received signal is outside the receiver bandwidth, the detection probability falls to that of false alarm.

4. A PRIORI KNOWLEDGE:

We assumed previously that each signal frequency is received with equal probability and the detection threshold was calculated accordingly. That is, the assumed discrete frequency grid for the average likelihood ratio test was taken to be uniformly distributed. Suppose, instead, that we have some knowledge of what frequency is sent and claim that one of the frequencies will be received with some probability, P_i. The likelihood ratio test must then be altered since

$$L(r) = \exp\left[-\frac{d^2}{2} \right] \sum_{i=1}^{N} P_i e^{\alpha_i} \qquad (20)$$

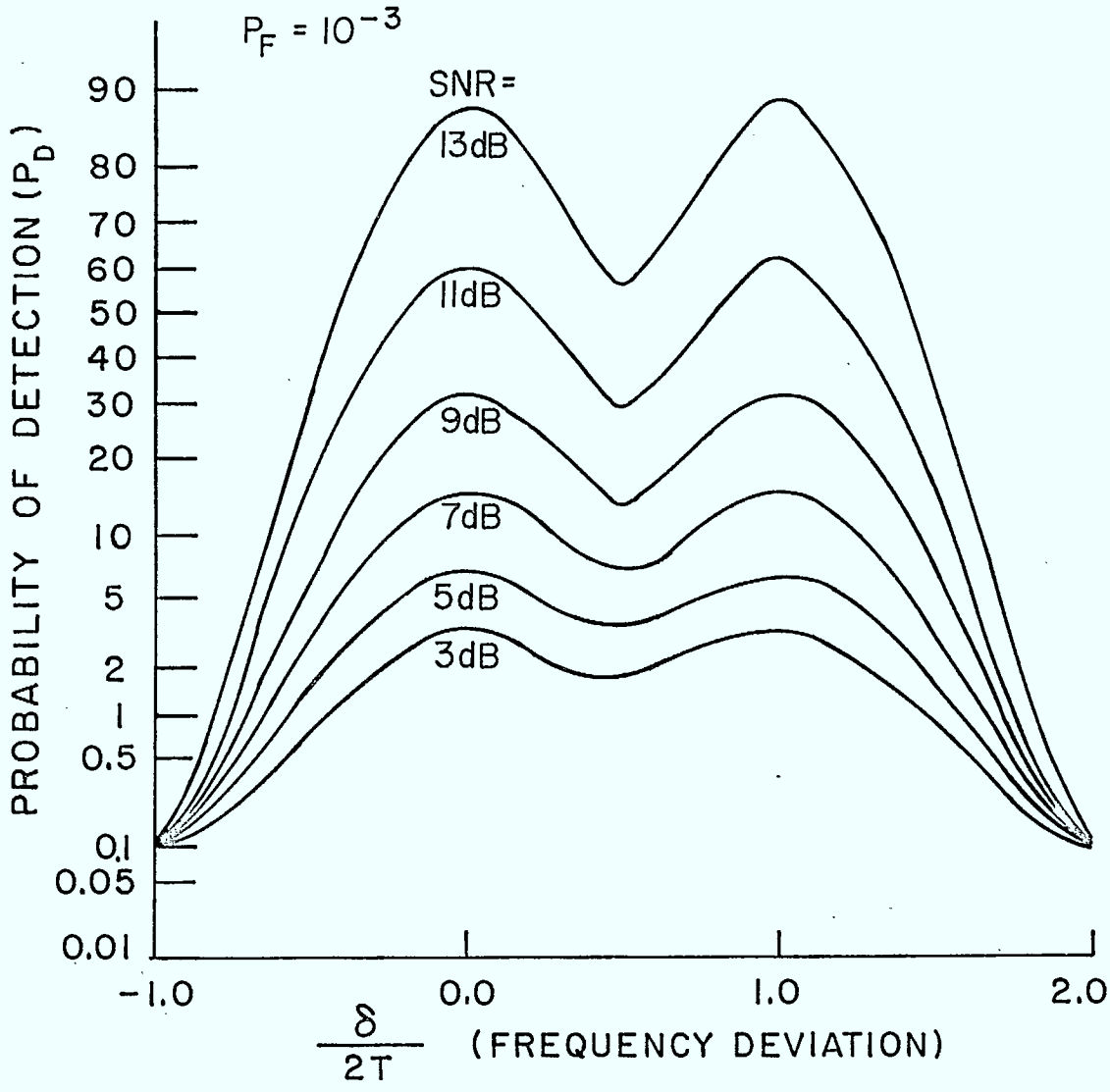


Figure 4: Frequency offset sensitivity of the optimum detector.

Therefore, for two transmission frequencies under these conditions:

$$L(r) = \exp\left[-\frac{d^2}{2}\right](P_1 e^{\alpha_1} + P_2 e^{\alpha_2}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{21}$$

Performing the analysis of the performance probabilities in the same manner as earlier, assuming once again that there is no frequency offset, a more general form of P_F is found. Defining $\ln b = \ln\left(\frac{\eta}{P_2}\right) + \frac{d^2}{2}$

$$P_F = Q\left[\frac{\ln\left[\frac{\eta}{P_2}\right] + \frac{d^2}{2}}{d}\right] \tag{22}$$

$$+ \int_{-\infty}^{\ln b} \frac{1}{(2\pi d^2)^{1/2}} \exp\left[-\frac{\alpha_2^2}{2d^2}\right] Q\left[\frac{\ln\left[\eta - P_2 \exp\left[\alpha_2 - \frac{d^2}{2}\right]\right] - \ln P_1}{d} + \frac{d}{2}\right] d\alpha_2$$

Since α_1 and α_2 are identically distributed, an alternate form of P_F can be found by switching P_1 and P_2.

Now, since the probability of arrival is now assumed higher for some frequencies than for others, the likelihood ratio test will be biased in their favor. This means the probability of detection is no longer the same at each detector. It is also possible that the interceptor is wrong and the actual arrival probabilities of the frequencies are different. Instead, assume frequency ω_i arrives with probability $\hat{P}_i$, and P_{Di} is the probability of detecting the signal at that frequency. Therefore,

$$P_D = P_{D_1}\hat{P}_1 + P_{D_2}\hat{P}_2$$

$$= P_{D_2} + \hat{P}_1 (P_{D_1} - P_{D_2}) \tag{23}$$

since $\hat{P}_1 = 1 - \hat{P}_2$.

Evaluation of P_{D_1} and P_{D_2} is similar to that of P_D done previously. So the following expressions for the detection probabilities are found

$$P_{D_1} = Q\left[\frac{\ell n\left[\frac{\eta}{P_1}\right] - \frac{d^2}{2}}{d_2}\right] \tag{24a}$$

$$+ \int_{-\infty}^{\ell n\, b} \frac{d\alpha_1}{(2\pi d^2)^{1/2}} \exp\left[-\frac{(\alpha_1 - d^2)^2}{2d^2}\right] Q\left[\frac{\ell n\left[\eta - P_1 \exp\left[\alpha_1 - \frac{d^2}{2}\right]\right] - \ell n P_2}{d} + \frac{d}{2}\right]$$

P_{D_2} is of similar form with P_2 substituted for P_1 and vice versa. Alternatively,

$$P_{D_1} = Q\left[\frac{\ell n\left[\frac{\eta}{P_2}\right] + \frac{d^2}{2}}{d}\right] \tag{24b}$$

$$+ \int_{-\infty}^{\ell n\, b} \frac{1}{(2\pi d^2)^{1/2}} \exp\left[-\frac{\alpha_2^2}{2d^2}\right] Q\left[\frac{\ell n\left[\eta - P_2 \exp\left[\alpha_2 - \frac{d^2}{2}\right]\right] - \ell n P_1}{d} - \frac{d}{2}\right] d\alpha_2$$

Again, P_{D^2} can be found by reversing the roles of P_1 and P_2. Table 2 gives values for P_D when $P_1 = 0.9$ for $P_F = 10^{-3}$.

Table 2: A priori knowledge effects on performance of the optimum receiver $P_F = 10^{-3}$

Probability of Detection

	$P_1 = 0.90$				$P_1 = 0.50$
SNR\$\hat{P}$	0.5	.75	0.9	1.0	0.5
3.0	.02419	.03542	.04212	.04664	.0308
5.0	.0505	.07242	.08561	.09438	.0658
7.0	.1147	.1551	.1792	.1954	.147
9.0	.2529	.3261	.3625	.3870	.319
11.0	.5457	.6059	.6419	.6661	.602
13.0	.8530	.8812	.8975	.9085	.880

The benefits of a priori knowledge are most significant at low SNR values, with 25% improvement in detection performance for SNR=3.0dB. It should be noted, however, that while the probability of detecting frequency has significantly

improved, there is much higher probability that the other frequency will be missed. Should the a priori information be wrong, the performance of the receiver will be degraded severely. Therefore, it is recommended, unless there is a great certainty about the a priori knowledge, that it not be used.

5. EXTENDING RESULTS TO MORE THAN TWO FREQUENCIES:

In trying to extend the results to more than two transmission frequencies it was necessary to find some method to approximate the probability distribution function of a sum of lognormal variables.

In a paper by Meyers [5], he suggests that the Gaussian Quadrature Method is applicable to this problem. However, while attempts to generate his curve for a single lognormal 0 dB (unit variance), zero-mean random variable were successful, there was found to be an extremely poor match for SNR values greater than this. Closer examination of this approach led to the conclusion that this method cannot follow the lognormal density function with any accuracy due to the exponential growth of its moments. It was, therefore, abandoned.

Two other approaches were found which give good approximations to the two frequency case, as indicated in Figure 5. They are discussed in a paper by Schwartz and Yeh [7]. The first, known as Wilkinson's approach, assumes that the sum of lognormal random variables is also lognormal. That is,

$$L = e^{Z} = \sum_{i=1}^{N} e^{z_i} \tag{25}$$

Now if the variance and mean of z are known then the probability distribution function can be found since

$$\Pr\left[e^{Z} > x\right] = Q\left[\frac{\ell n\, x}{\sigma_z} + m_z\right] \tag{26}$$

The variance and the mean can be found from the moments of L, which are calculated recursively [6]. Let $\mu_k(n-1)$ be the kth moment of a sum of n-1 lognormal random variables. Since the z_i's are independent,

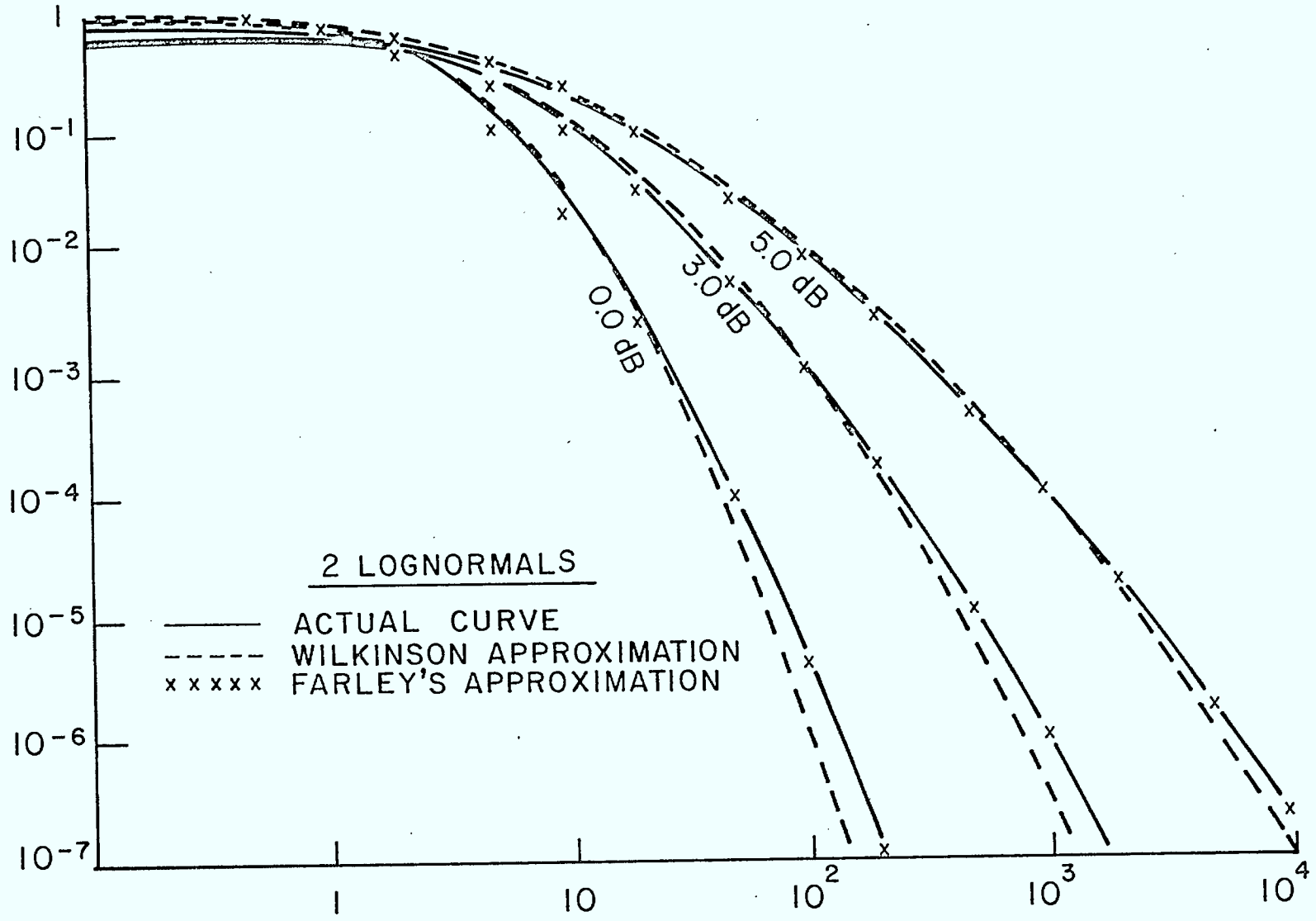


Figure 5: The probability distribution function, and two approximations, of a log normal random variable.

$$\mu_k(n) = \sum_{i=0}^{k} \binom{k}{i} \mu_i(1)\ \mu_{k-i}(n-1) \tag{27}$$

Now

$$\mu_k(1) = \exp\left[km_z + \frac{\sigma_z^2 k^2}{2}\right] \tag{28}$$

so,

$$\mu_1(n) = \exp\left[m_z + \frac{\sigma_z^2}{2}\right] \tag{29}$$

and

$$\mu_2(n) = \exp\left[2m_z + 2\sigma_z^2\right]$$

Solving the two equations gives

$$m_z = 2\ \ell n\ (\mu_1(n)) - 0.5\ \ell n\ (\mu_2(n)) \tag{30}$$

$$\sigma_z^2 = \ell n\ (\mu_2(n)) - 2\ \ell n\ (\mu_1(n)) \tag{31}$$

and so the distribution function of the sum can be found. The approximation to the actual curves, particularly for false alarm probabilities less than 10^{-2}, is extremely close, as can be seen in Figure 5.

The second approach, attributed to Farley, says that if the z_i's are identically distributed and independent, as $\sigma_y \to \infty$,

$$\Pr\ [L < \exp\ (\ \gamma\sigma_y + m_y)] = [1 - Q\ (\gamma)]^N \tag{32}$$

Where N is the number of lognormal variables in the random variable. This approach gives a better approximation than Wilkinson's, particularly for P_F greater than 10^{-2}.

An interesting observation about this approach is it implies, for large SNR, that the performance of both receivers presented in this report is the same. To see this, recall that for a single transmission frequency, $Q_F = Q\left[\frac{\ell n\ \eta}{d}\right]$, and for

the maximum likelihood receiver, $P_F = 1 - (1-Q_F)^N$. In this case, $\sigma_y = d^2$ and $m_y=0$. Letting $\eta = \exp(\gamma\sigma_y + m_y)$, we get

$$\Pr[L > \eta] = P_F(\text{ALT}) = 1 - (1-Q_F)^N = P_F(\text{MLT})$$

The problem with Farley's approach as opposed to Wilkinson's is that it is not immediately obvious how to extend these results to evaluate the probability of detection. This is currently being considered and it is suspected that the detection performance of the optimum receiver will be closely related to the performance of the maximum likelihood receiver.

6. FUTURE WORK

To date, results have been established for the most basic case of two different receivers. Three key areas are still under development.

The first is to find the probability of false alarm for large values of N. Results for 4 frequencies, along with the results of a Monte Carlo Simulation, are shown in Figure 6. The work in this area, as can be seen from the results presented, is nearly completed. The approximations used, i.e. Wilkinson or Farley, are not limited to N = 4.

The second is to extend the results to the probability of detection, particularly determining if the two receivers are indeed nearly equivalent in detection performance.†

The third area, and in some ways, the most important, is to analyze the effects of a Doppler shift on the receiver output and determine the feasibility of recognizing that the signal source is moving.

Finally, the effect of using a non-coherent receiver structure will be studied. The interception receiver structures have already been developed. Since no computational results are available, the theory was not included in this report.

† Note added in proof. The probability of detection study is now complete. The theme of the results are similar to that for the probability of false alarm.

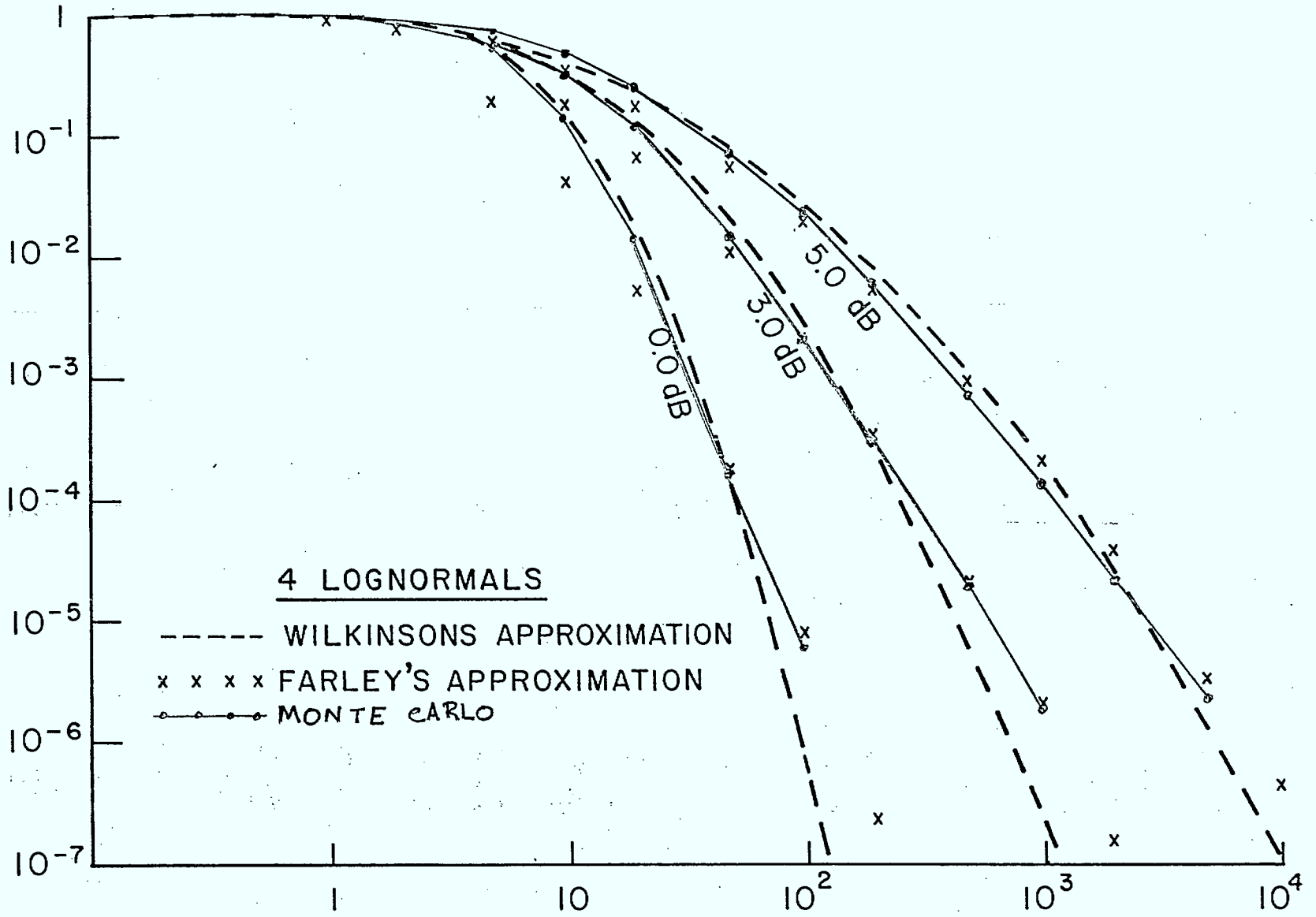


Figure 6: Approximations of the probability distribution function of a sum of 4 lognormal random variables.

7. CONCLUSION

Two coherent receivers for interception spread spectrum signals were presented and it was shown that they have near equivalent performance. A performance analysis of a receiver with only one or two frequencies was done. The results show a drop in performance as the number of frequencies increase and as the signal to noise ratio of the transmitted frequency decreases. A noticeable drop in performance was seen if the received signal was offset from the filter center frequencies. Approximation techniques for a large number of frequencies are available from the li terature.

REFERENCES

[1] M. Kavehrad, P.J. McLane, "Spread Spectrum for Indoor Digital Radio", IEEE Communications Magazine, Vol. 25, June 1987, pp.32-40

[2] R.A. Dillard, "Detectability of Spread-Spectrum Signals", IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-15, July 1979, pp.526-537.

[3] A.J. Viterbi, "Spread Spectrum Communications - Myths and Realities", IEEE Communications Magazine, Vol. 17, May 1979, pp.11-18.

[4] A.J. Viterbi, "When Not to Spread Spectrum - A Sequel", IEEE Communications Magazine, Vol. 23, April 1985, pp.12-17.

[5] M.H. Meyers, "Computing the Distribution of a Random Variable via Gaussian Quadrature Rules", Bell System Technical Journal, Vol. 61, Nov. 1982, pp.2245-2261.

[6] V.K. Prabhu, "Some Consideration of Error Bounds in Digital Systems", Bell System Technical Journal, Vol. 50, Dec. 1971, pp.3127-3151.

[7] S.C. Schwartz, Y.S. Yeh, "On the Distribution Function and Moments of Power Sums With Log-Normal Components", The Bell System Technical Journal, Vol. 61, September 1982, pp.1441-1462.

[8] H.L. Van Trees, Detection, Estimation, and Modulation Theory - Part I, John Wiley & Sons, New York: 1968.

[9] A. Papoulis, Probability, Random Variables, and Stochastic Processes, 2nd ed., McGraw-Hill Book Company, New York: 1984.