communication systems," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 113–119, Apr. 1971.
[16] A. J. Viterbi, "Phase-locked loop dynamics in the presence of noise by Fokker–Planck techniques," *Proc. IEEE*, vol. 51, pp. 1737–1753, Dec. 1963.
[17] R. W. Lucky, J. Salz, and E. J. Weldon, *Principles of Data Communication*. New York: McGraw-Hill, 1968.
[18] D. D. Falconer and R. D. Gitlin, "Bounds on error-pattern probabilities for digital communications systems," *IEEE Trans. Commun.*, vol. COM-20, pp. 132–139, Apr. 1972.
[19] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2. New York: Wiley, 1966.
[20] R. Matyas and P. J. McLane, "Data-aided tracking loops for channels with phase jitter and intersymbol interference," in *Proc. 1973 Int. Communications Conf.*, pp. 33-8–33-13.

★

Robert Matyas (S'70–M'73) was born in Montreal, P. Q., Canada, on December 10, 1949. He received the B.Eng. degree in electrical engineering from McGill University, Montreal, in 1971 and the M.Sc. degree in electrical engineering from Queen's University, Kingston, Ont., Canada, in 1973.

Concurrent with his graduate studies at Queen's University, he was a Teaching and Research Assistant with the Department of Electrical Engineering. His area of research as a graduate student centered on the effect of phase error on digital communications. Since joining the Earth Station Engineering Group, Telesat Canada, Ottawa, Ont., Canada, in 1973, he has been involved in the implementation of a time division multiple access (TDMA) link in the Canadian domestic satellite system.

★

Peter J. McLane (S'68–M'69) was born in Vancouver, B. C., Canada, on July 6, 1941. He received the B.A.Sc. degree from the University of British Columbia, Vancouver, in 1965, the M.S.E. degree from the University of Pennsylvania, Philadelphia, in 1966, and the Ph.D. degree from the University of Toronto, Toronto, Ont., Canada, in 1969, all in electrical engineering. At the University of Pennsylvania he held a Ford Foundation Fellowship and at the University of Toronto he held a National Research Council of Canada Scholarship.

From 1966 to 1967 he was a Junior Research Officer with the National Research Council of Canada, Ottawa, Ont., Canada. He held summer positions with this organization in 1965 and 1966 and with the Defence Research Board of Canada in 1964. Since 1969 he has been a faculty member with the Department of Electrical Engineering, Queen's University, Kingston, Ont., Canada, where he is currently an Associate Professor. His research interests are in signal processing for communications and radar and in vehicle control in transportation systems. He has served as a consultant on research problems with the Canadian Department of Communications and the Canadian Institute of Guided Ground Transport at Queen's University.

Coherent and Noncoherent Detection of CPFSK

WILLIAM P. OSBORNE, MEMBER, IEEE, AND MICHAEL B. LUNTZ

Abstract—**Continuous phase frequency shift keying (CPFSK) is potentially an attractive modulation scheme for use on channels whose performance is limited by thermal noise. In this paper results for the performance available with CPFSK are given for coherent detection and noncoherent detection with arbitrary modulation indices and arbitrary observation intervals.**

This work serves two purposes. First, it provides interesting, new results for the noncoherent detection of CPFSK which indicate that the performance of such a system can be better than the performance of coherent PSK. Secondly, it provides a complete analysis of the performance of CPFSK at high SNR as well as low SNR and thereby unifies and extends the results previously available.

INTRODUCTION

IN SEVERAL recent papers the performance gain available by multiple bit detection of continuous phase frequency shift keying (CPFSK) signals has been discussed. Pelchat *et al.* have discussed the distance properties and, hence, high SNR performance of coherently detected CPFSK waveforms for two and three bit observation intervals [1]. In addition, this paper discusses optimum coherent demodulation with infinite observation interval. DeBuda [2] has discussed the performance of coherent CPFSK with a modulation index of 0.5 and given a self-synchronizing receiver structure for this case. Forney [3] has discussed the use of the Viterbi algorithm for detection of coherent CPFSK and, in particular, the modulation index 0.5 case studied by DeBuda is examined. Pelchat and Adams [4] have discussed the minimum probability of bit error noncoherent receiver for the three-bit observation interval and they have shown that the low SNR performance can be estimated by the average matched filter concept. In this paper receiver structures which minimize the probability of bit error for both coherent and noncoherent detection for arbitrary observation intervals are presented. The performance of both the coherent and the noncoherent demodulators is bounded employing the average matched filter concept at low SNR and employing the union bound at high SNR. This combination of bounds forms a performance bound which is a good estimate of the performance available with these receivers at all SNR's.

Paper approved by the Associate Editor for Communication Theory of the IEEE Communications Society for publication without oral presentation. Manuscript received October 4, 1973.
The authors are with Radiation, Inc., Melbourne, Fla. 32901.

The paper is organized in three major sections. These are coherent detection, noncoherent detection, and a summary. In the first two sections the receivers are presented followed by low and high SNR bounds. In the final section the results of the first two sections are discussed. In addition, the realizability of the various demodulators is discussed briefly.

COHERENT DETECTION OF CPFSK

The detection problem to be addressed in this paper consists of observing n bits of a CPFSK waveform and producing an optimum decision on one bit. In the coherent case, the decision is made on the first bit by observing the waveform during this bit time and $n-1$ additional bit times. The data are assumed to be random ± 1's and the interference is additive white Gaussian noise.

The CPFSK waveform during the first bit interval can be expressed as

$$s(t) = \cos\left(\omega_c t + \frac{a_1 \pi h t}{T} + \theta_1\right) \qquad 0 \le t \le T \tag{1}$$

where a_1 is the data, θ_1 is the phase of the RF carrier at the beginning of the observation interval, and h, the modulation index, is the peak-to-peak frequency deviation divided by the bit rate. In accord with the continuity of phase, the waveform during the ith bit time of the observation interval can be written as

$$s(t) = \cos\left(\omega_c t + \frac{a_i \pi h (t-(i-1)T)}{T} + \sum_{j=1}^{i-1} a_j \pi h + \theta_1\right) \qquad (i-1)T \le t \le iT. \tag{2}$$

The objective is to design a receiver which observes n bit times of data and uses the fact that the carrier phase during the ith bit time depends upon the data in the first bit time to minimize the probability of bit error. For the case of coherent detection to be treated in this section, θ_1 is assumed known and set to zero with no loss of generality. In the next section the noncoherent case is treated where in θ_1 is assumed to be a random variable uniformly distributed between $\pm\pi$.

Let the signal waveform during the observation interval be denoted by $s(t,a_1,A_k)$ where A_k represents a particular data sequence, i.e., it represents the $n-1$ tuple $\{a_2,a_3\cdots,a_n\}$, and the actual waveform is again given by (2). The detection problem is then to observe $s(t,a_1,A_k)$ in noise and produce an optimum decision as to the polarity of a_1. The problem stated in this manner is the composite hypothesis problem treated in [5] and other texts. This solution is known to be the likelihood ratio test and for the CPFSK waveform the likelihood ratio, l, can be expressed as

$$l = \frac{\displaystyle\int_A \exp\left(\frac{2}{N_0}\int_0^{nT} r(t)s(t,1,A)\,dt\right) f(A)\,dA}{\displaystyle\int_A \exp\left(\frac{2}{N_0}\int_0^{nT} r(t)s(t,-1,A)\,dt\right) f(A)\,dA} \tag{3}$$

where the integral $\int_A dA$ is taken to mean the $n-1$ fold integral

$$\int_{a_2}\int_{a_3}\cdots\int_{a_n} da_2\, da_3 \cdots da_n.$$

The density of A is given by, $f(A) = f(a_2)\,f(a_3)\cdots f(a_n)$, where $f(a_i)$ is the density function of the ith data bit, and the data bits are assumed to be independent. The density function of the random data bits is given by,

$$f(a_i) = \tfrac{1}{2}\delta(a_i - 1) + \tfrac{1}{2}\delta(a_i + 1). \tag{4}$$

Using (4) in (3) and carrying out the integration, the likelihood ratio becomes

$$l = \frac{\exp\left(\frac{2}{N_0}\int_0^{nT} r(t)s(t,1,A_1)\,dt\right) + \cdots + \exp\left(\frac{2}{N_0}\int_0^{nT} r(t)s(t,1,A_m)\,dt\right)}{\exp\left(\frac{2}{N_0}\int_0^{nT} r(t)s(t,-1,A_1)\,dt\right) + \cdots + \exp\left(\frac{2}{N_0}\int_0^{nT} r(t)s(t,-1,A_m)\,dt\right)} \tag{5}$$

where

$$m = 2^{n-1}.$$

The receiver structure defined by (5) is shown in block diagram in Fig. 1. The receiver correlates the received waveform with each of the m possible transmitted signals beginning with data 1, then forms the sum of $\exp(c_j)$ where c_j is the correlation of the received waveform with the jth signal waveform beginning with a data 1. A similar operation of correlating and summing for the m possible waveforms beginning with a data -1 is performed and the decision is based on the polarity of the difference in the two sums.

PERFORMANCE OF THE COHERENT DEMODULATOR

The performance of the optimum demodulator shown in Fig. 1 cannot be computed analytically. However, its performance can be bounded by two bounds. One bound is tight at high SNR and the other is tight at low SNR. These bounds taken as a single bound are a reasonably good performance bound at all values of SNR.

Upper Bound on Performance—Low SNR

The receiver presented in the previous section computes sums of random variables of the form

$$x_{1k} = \exp\left(\frac{2}{N_0}\int_0^{nT} r(t)s(t,1,A_k)\,dt\right). \tag{6}$$

At low values of E_b/N_0 the random variable x_{1k} can be approximated by

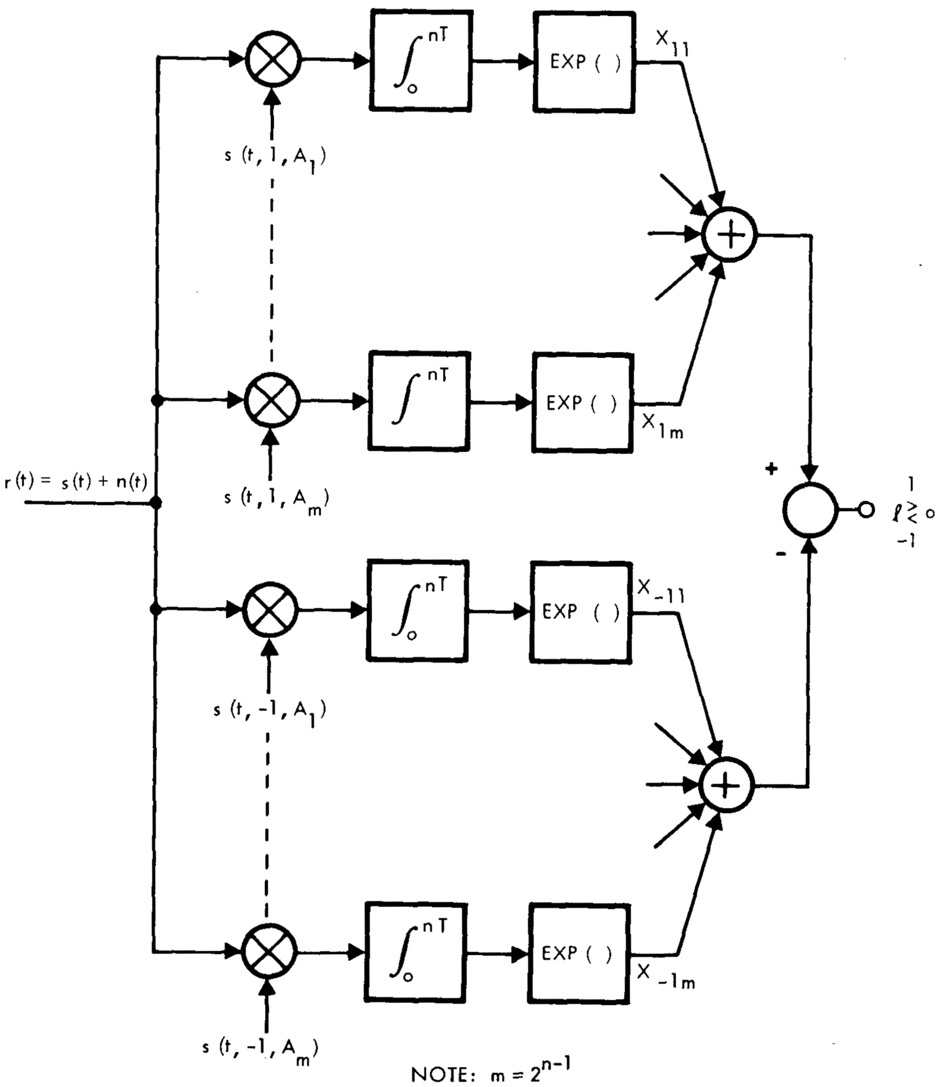

Fig. 1. Block diagram of optimum coherent receiver.

$$x_{1k} \approx 1 + \frac{2}{N_0} \int_0^{nT} r(t) s(t,1,A_k)\, dt. \tag{7}$$

Using the approximation of (7) in (5) the receiver operation at low SNR becomes,

$$\int_0^{nT} r(t) \left(\sum_{k=1}^{m} s(t,1,A_k)\right) \overset{1}{\underset{-1}{\gtrless}} \int_0^{nT} r(t) \left(\sum_{k=1}^{m} s(t,-1,A_k)\right). \tag{8}$$

The receiver described by (8) can be thought of as a device which correlates with the average waveform given a data 1 and average waveform given a data -1 and makes a decision based on the largest of these correlations. A block diagram of this receiver is shown in Fig. 2.

The decision variable, Λ, for the low SNR receiver shown in Fig. 2, is a Gaussian random variable and, hence, the probability of error is computed using only the mean and variance of Λ. The mean of Λ will depend on the transmitted signal and therefore the probability of error will also depend on the transmitted sequence. First the probability of error given a particular transmitted sequence will be computed.

The mean of Λ given a particular sequence is given by[1]

$$E(\Lambda \mid s(t,1,A_j)) = \int_0^{nT} s(t,1,A_j) (\bar{s}(t,1) - \bar{s}(t,-1))\, dt \tag{9}$$

where

$$\bar{s}(t,1) = \sum_{k=1}^{m} s(t,1,A_k).$$

The variance of Λ is independent of a particular transmitted sequence and is given by,

$$\mathrm{Var}(\Lambda) = N_0/2 \int_0^{nT} (\bar{s}(t,1) - \bar{s}(t,-1))^2\, dt. \tag{10}$$

The probability of error, given this sequence, is given by

$$\Pr(\epsilon \mid s(t,1,A_j)) = Q\left[\frac{E(\Lambda \mid s(t,1,A_j))}{(\mathrm{Var}(\Lambda))^{1/2}}\right] \tag{11}$$

[1] $E(x)$ is used to denote the expected value of x.

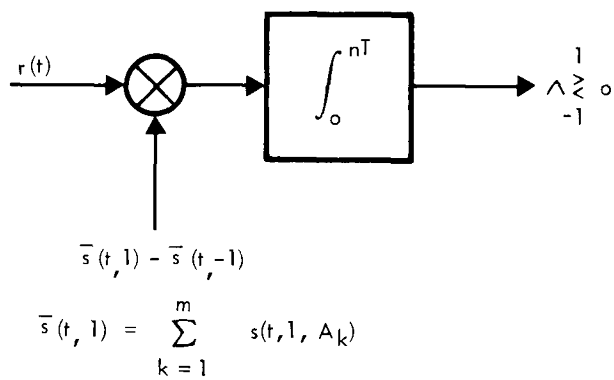

Fig. 2. Low SNR approximation to optimum coherent CPFSK receiver.

where

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

The probability of bit error is given by averaging over the possible transmitted sequences, i.e.,

$$P(\epsilon) = \sum_{j=1}^{m} P(\epsilon \mid s(t,1,A_j))P(A_j). \quad (12)$$

The probability of error can be determined using (11) and (13) if expressions are available for the mean and variance of Λ. The variance of Λ is given by (10), however, to evaluate (11) requires an expression for $\bar{s}(t,1)$ and $\bar{s}(t,-1)$. During the kth bit time, $\bar{s}(t,1)$ is given by the sum of the 2^{k-1} possible waveforms which can be transmitted using 2^{k-2} possible starting phases. Following this line of reasoning, the average waveform, given a data 1 in the first bit, during the kth bit time, can be expressed as

$$\bar{s}_k(t,1) = \frac{1}{2^{k-2}} \sum_{i=1}^{2^{k-2}} \operatorname{Cos}\left(w_c t + \frac{\pi h(t-(k-1)T)}{T} + \theta_i\right) + \operatorname{Cos}\left(w_c t - \frac{\pi h(t-(k-1)T)}{T} + \theta_i\right). \quad (13)$$

where θ_i is the ith possible starting phase for the kth bit.

Expanding the terms of the form of Cos $(a \pm b)$, (13) reduces to

$$\bar{s}_k(t,1) = \frac{1}{2^{k-2}} \operatorname{Cos}\left(\frac{\pi h(t-(k-1)T)}{T}\right) \sum_{i=1}^{2^{k-2}} \operatorname{Cos}(w_c t + \theta_i). \quad (14)$$

The possible values of θ_i and number of times each one occurs can best be seen by referring to a diagram of the phase of the waveform with respect to the carrier as a function of time. Such a diagram is shown in Fig. 3. By inspection of Fig. 3, (14) can be rewritten as

$$\bar{s}_k(t,1) = \frac{1}{2^{k-2}} \operatorname{Cos}\left(\frac{\pi h(t-(k-1)T)}{T}\right) \cdot \sum_{\alpha=0}^{k-2} \operatorname{Cos}(w_c t + (k-2\alpha-1)\pi h). \quad (15)$$

By combining terms of the form $\operatorname{Cos}(w_c t + \pi h + n\pi h)$ and $\operatorname{Cos}(w_c t + \pi h - n\pi h)$ in the sum in (15) $\bar{s}_k(t,1)$ can be reduced to,

$$\bar{s}_k(t,1) = \operatorname{Cos}^{k-2}(\pi h) \operatorname{Cos}(w_c t + \pi h) \cdot \operatorname{Cos}\left(\frac{\pi h(t-(k-1)T)}{T}\right). \quad (16)$$

$\bar{s}_k(t,-1)$ can be computed by an analogous procedure and the result is

$$\bar{s}_k(t,-1) = \operatorname{Cos}^{k-2}(\pi h) \operatorname{Cos}(w_c t - \pi h) \cdot \operatorname{Cos}\left(\frac{\pi h(t-(k-1)T)}{T}\right). \quad (17)$$

Equations (16) and (17) only apply for $k \geq 2$, however, the waveform during the first bit time is obvious.

The variance of the decision variable, Λ, can be computed by using (16) and (17) in (10). The procedure for this calculation is to compute the contribution to the variance of the kth bit and sum over the n bits. The result of this calculation is,

$$\operatorname{Var}(\Lambda) = \frac{N_0 T}{2}\left[1 - \operatorname{Sinc}(2h) + \frac{0.5(1-\operatorname{Cos}(2\pi h))(1+\operatorname{Sinc}(2h))(\operatorname{Cos}^{2n-2}(\pi h) - 1)}{\operatorname{Cos}^2(\pi h) - 1}\right]. \quad (18)$$

In order to complete the computation of the probability of error, the mean output of the demodulator, given a particular input sequence, must be computed. Let the data sequence of the input signal be a_1, A_k, then the input signal is,

$$s(t,a_1,A_k) = \operatorname{Cos}\left(w_c t + \frac{a_i \pi h(t-(i-1)T)}{T} + \sum_{j=1}^{i-1} a_j \pi h\right)$$
$$(i-1)T \leq t \leq iT. \quad (19)$$

The procedure for computing the mean will be to compute the contribution of the ith input bit to the mean

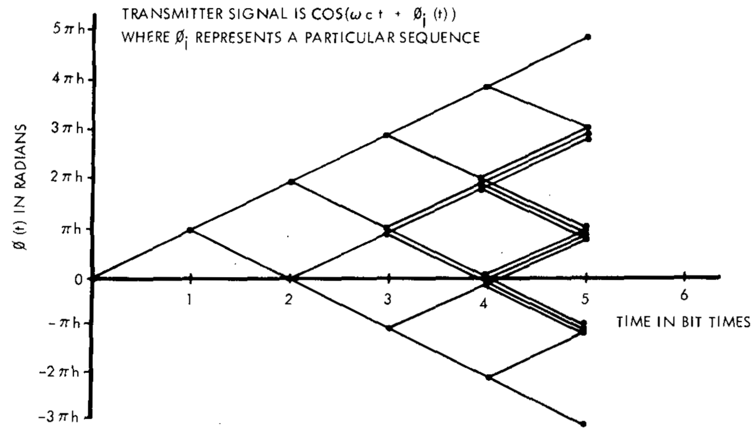

Fig. 3. Phase of CPFSK waveforms which are averaged to produce $\bar{s}(t, 1)$ versus time.

output and sum over the observation interval. From (9) and (16) the contribution to mean output due to the ith input bit, E_i, can be written as,

$$E_i = \int_0^T \mathrm{Cos}\left(w_c t + \frac{a_i \pi h t}{T} + \sum_{j=1}^{i-1} a_j \pi h\right) \cdot 2\,\mathrm{Sin}\,(\pi h)\,\mathrm{Cos}^{i-2}\,(\pi h) \cdot \mathrm{Cos}\,(\pi h t/T)\,\mathrm{Sin}\,(w_c t)\, dt. \tag{20}$$

Carrying out the integration in (20), E_i becomes,

$$E_i = \frac{T}{2}\,\mathrm{Sin}\,(\pi h)\,\mathrm{Cos}^{i-2}\,(\pi h)\left[\mathrm{Sin}\,(\theta_i) + \frac{a_i}{2\pi h} \cdot (\mathrm{Cos}\,(\theta_i) - \mathrm{Cos}\,(2\pi h a_i + \theta_i))\right] \tag{21}$$

where

$$\theta_i = \sum_{j=1}^{i-1} \pi h a_j.$$

Equation (21) is good only for $i > 1$ and for $i = 1$ E_1 is given by

$$E_1 = \frac{T}{2}\,(1 - \mathrm{Sinc}\,(2h)). \tag{22}$$

Upper Bound on Performance—High SNR

The equations presented above can be used to evaluate a bound on the performance of CPFSK at low SNR's. These will be used in conjunction with the union bound which is tight at high SNR's to provide the composite bound. The probability of error for the optimum receiver is overbounded by

$$\Pr\,(\epsilon) \leq \frac{1}{m} \sum_{l=1}^{m} \sum_{j=1}^{m} \Pr\,(x_{1l} \leq x_{-1j} \mid s(t,1,A_l)) \tag{23}$$

where x_{1l} is output of the correlator matched to the signal $s(t,1,A_l)$. Further,

$$\Pr\,(x_{1l} \leq x_{-1j} \mid s(t,1,A_l)) = Q\left[\left(\frac{nE_b}{N_0}\,(1 - \rho(l,j))\right)^{1/2}\right] \tag{24}$$

where

$$\rho(l,j) = \frac{1}{nE_b} \int_0^{nT} s(t,-1,A_l) s(t,1,A_j)\, dt. \tag{25}$$

The correlation coefficient $\rho(l,j)$ can be evaluated by using (19) for the signal waveforms, integrating one bit at a time, and summing the results over the observation interval. Carrying out this process, $\rho(l,j)$ can be written as

$$\rho(l,j) = \frac{1}{n} \sum_{k=1}^{n} \mathrm{Sinc}\left(\frac{h}{2}\,(a_k - b_k)\right) \cdot \cos\left[\frac{\pi h}{2}\,(a_k - b_k) + \sum_{j=1}^{k-1} \pi h (a_j - b_j)\right] \tag{26}$$

where the a's are the data bits A_l, the b's are the data bits A_j, and where $a_1 = 1$ and $b_1 = -1$.

Lower Bound on Performance

A lower bound on the performance of the coherent CPFSK receiver can be obtained by supposing that for each transmitted sequence the receiver needs only to decide between that sequence and its nearest neighbor. This receiver will perform at least as well as the receiver which does not know which of two sequences was transmitted but must compare with all possible sequences. The performance of this receiver is a lower bound to the performance of the optimum receiver presented in the previous section. This lower bound on the probability error in the CPFSK receiver can be written as,

$$\Pr\,(\epsilon) \geq \frac{1}{m} \sum_{l=1}^{m} Q\left(\left(\frac{nE_b}{N_0}\,(1 - \rho^*(l))\right)^{1/2}\right) \tag{27}$$

where $\rho^*(l)$ = maximum of $\rho(l,j)$ over all j.

Numerical Results—Coherent Case

In the previous section three bounds on the performance of a coherent CPFSK system with observation interval of length of n were presented. The average matched filter bound is an upper bound on performance, which, by its construction, should be an approximation to the true performance at low SNR. The union bound is an upper bound

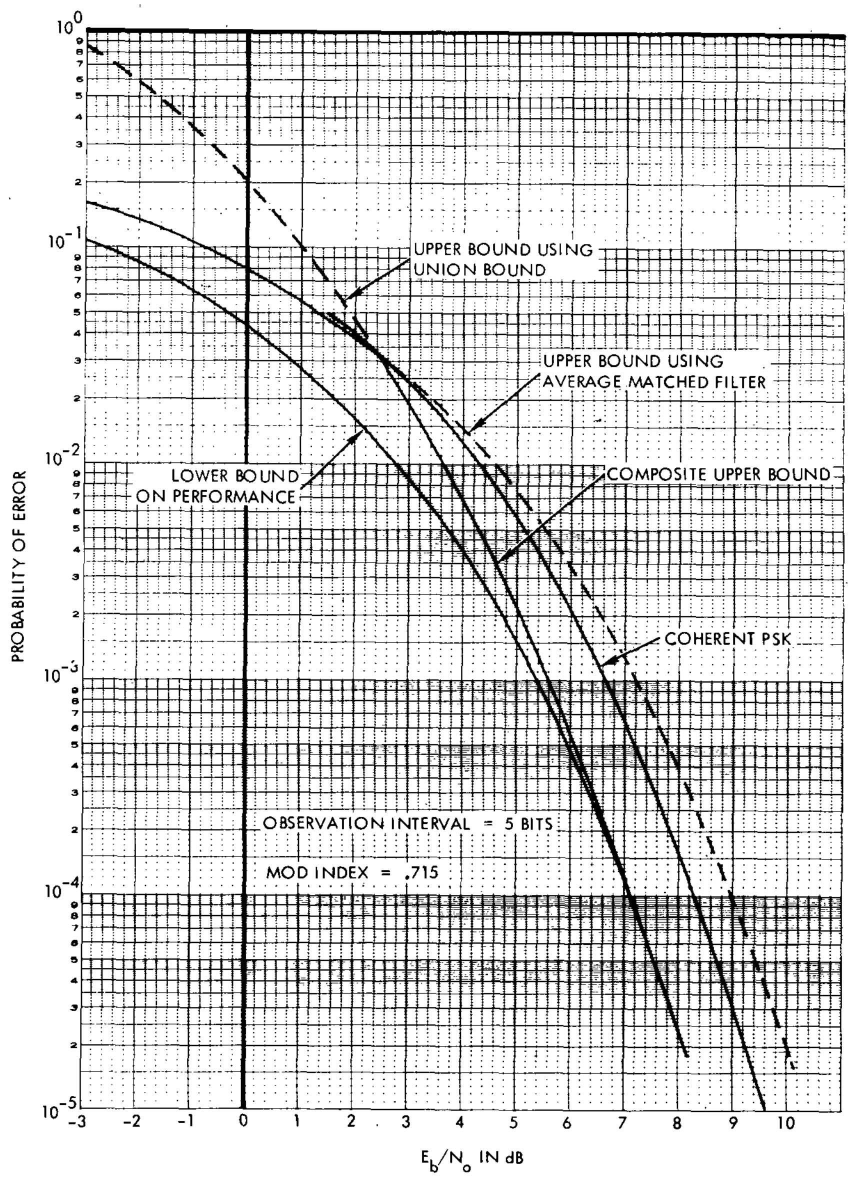

Fig. 4. Bounds on performance of CPFSK.

which is known to be tight at high SNR, and the bound given by (27) is a lower bound on performance at any SNR.

In order to illustrate the use of these bounds to estimate performance of coherent CPFSK systems all three were evaluated and plotted for an observation interval of five bits and modulation index, h, of 0.715. These results are plotted in Fig. 4. From Fig. 4 it can be seen that the composite upper bound constructed by taking the smaller of the average matched filter bound and the union bound converges to the lower bound at high SNR and, in fact, for error rates less than 10^{-3} these are essentially equal. The composite upper bound is within 1.5 dB of the lower bound at all SNR's showing that the composite bound is a good approximation to the true receiver performance at all SNR's and is tight at high SNR's. Fig. 4 illustrates the "goodness" of the three bounds only for one set of parameters, however, the author's use of these bounds in several cases has shown similar results, i.e., the composite upper bound is a good approximation to true receiver performance at all SNR's. Further evidence of this is shown in the noncoherent section in the form of a comparison of computer simulation results with this bound.

The modulation index of 0.715 was selected for evaluation because in [2] it was shown that the maximum value of the minimum distance over all transmitted words for a CPFSK signal was achieved by using this modulation index. In Fig. 5 the performance of CPFSK with this modulation index versus the length of the observation interval is illustrated. The curves in Fig. 5 are the composite upper bound results for the various observation intervals. The results show that little gain is available by using observation intervals longer than three bits at any SNR. Again, this behavior is a characteristic of CPFSK systems independent of modulation index, i.e., in other cases investigated the gain achieved by using an interval of more than three bits is very small. As has already been pointed out in [2], but is again illustrated in Fig. 4, CPFSK with a modulation index of 0.715 does perform better than coherent PSK.

NONCOHERENT DETECTION OF CPFSK

In this section the detection of CPFSK when the carrier phase is unknown will be discussed. Specifically, this section will discuss observing $2n + 1$ bits of a CPFSK waveform and making decisions on the $n + 1$st (middle) bit.[2]

The CPFSK waveform was described by (2) as

$$s(t) = (2P)^{1/2} \cos\left[\omega_c t + \frac{(a_i \pi h(t - (i-1)T)}{T} + \pi h \sum_{j=1}^{i=1} a_j + \theta_1\right] \qquad (i-1)T < t < iT$$

$$= (2P)^{1/2} \cos\left(\omega_c t + \frac{a_1 \pi h t}{T} + \theta_1\right) \qquad 0 < t < T. \tag{28}$$

Let the observed waveform be denoted by $s(t, a_{n+1}, A_k, \theta_i)$ where A denotes the $2n$ tuple $\{a_1, a_2, \cdots, a_n, a_{n+2} \cdots a_{2n+1}\}$. This notation is similar to that used in the previous section. It differs in that the initial phase θ_i is also an independent variable. Here, it is assumed that the a_i are equally probably to be ± 1 and are independent. The phase θ_1 is assumed uniformly distributed between $\pm\pi$. A receiver is to be designed to make decisions on the $n + 1$st bit, i.e., decide the polarity of a_{n+1}. It is desired to find the receiver structure which minimizes the probability of decision error. The statistic which must be computed for this composite hypothesis test is the likelihood ratio which can be expressed as,

$$l = \frac{\int_{\theta_1} \int_A f(A) f(\theta_1) \exp\left(\frac{2}{N_0} \int r(t) s(t, 1, A, \theta_1)\, dt\right) d\theta_1,\, dA}{\int_{\theta_1} \int_A f(A) f(\theta_1) \exp\left(\frac{2}{N_0} \int r(t) s(t, -1, A, \theta_1)\, dt\right) d\theta_1,\, dA}. \tag{29}$$

This likelihood ratio differs from the one given in (3) in that an additional expectation over the random initial phase must be taken and also in that the decision is performed on the middle rather than the first bit. Performing first the expectation over all transmitted sequences as was done to obtain (5), the likelihood ratio becomes,

$$l = \frac{\int_{\theta_1} \sum_{k=1}^{m'} \exp\left(\frac{2}{N_0} \int r(t) s(t, 1, A_k, \theta_1)\, dt\right) f(\theta_1)\, d\theta_1}{\int_{\theta_1} \sum_{k=1}^{m'} \exp\left(\frac{2}{N_0} \int r(t), s(t, -1, A_k, \theta_1)\, dt\right) f(\theta_1)\, d\theta_1} \tag{30}$$

where

$$m' = 2^{2n}.$$

The average over the random phase yields, as is well known, the zero-order modified Bessel function. After performing this average the likelihood function may be written

$$l = \sum_{i=1}^{m'} I_0\left(\frac{2}{N_0} z_{1i}\right) \Big/ \sum_{i=1}^{m'} I_0\left(\frac{2}{N_0} z_{-1i}\right) \tag{31}$$

where

$$z_{1i}^2 = \left(\int r(t) s(t, 1, A_i, 0)\, dt\right)^2 + \left(\int r(t) s\left(t, 1, A_i, \frac{\pi}{2}\right) dt\right)^2$$

and

$$z_{-1i}^2 = \left(\int r(t) s(t, -1, A_i, 0)\, dt\right)^2 + \left(\int r(t) s\left(t, -1, A_i, \frac{\pi}{2}\right) dt\right)^2.$$

The correlator references are the inphase and quadrature components of a constant amplitude waveform. Thus, for a data 1 in the middle bit interval, they may be written as

$$s(t, A_i, 0) + js\left(t, A_i, \frac{\pi}{2}\right) = (2P)^{1/2} \exp j[\omega_0 t + \phi_i(t)] \tag{32}$$

where $\phi_i(t)$ is the phase trajectory of the continuous phase waveform. Therefore, the quantity z_i^2 may be written in complex notation as

$$z_i^2 = 2P \left| \int r(t) \exp j[\omega_0 t + \phi_i(t)]\, dt \right|^2. \tag{33}$$

A similar expression may be written for a data -1 in the middle bit interval. A block diagram of the receiver which computes this likelihood ratio is shown in Fig. 6. This receiver correlates the received signal $r(t)$ with inphase

[2] It can be shown that the magnitude of the complex correlation between two CPFSK waveforms corresponding to data differing in only one bit is a minimum when the difference bit is in the middle.

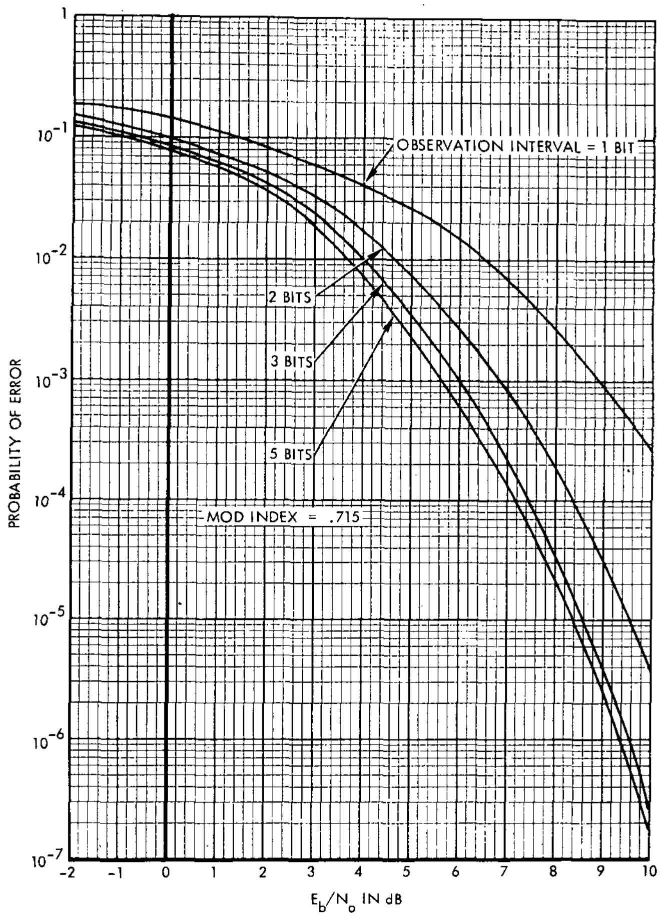

Fig. 5. Upper bounds on CPFSK performance.

and quadrature components of each of the possible transmitted signals. For each possible signal the receiver forms the root sum square of the inphase and quadrature components and weights this root with an $I_0(\cdot)$ nonlinearity. The sum of these numbers for all signals with a data one in the middle bit interval is compared with the sum for all signals with a data -1 in the middle bit interval.

Noncoherent Receiver Performance

No closed form analytical solution for the performance of the noncoherent receiver exists. However, as is the case for the coherent receiver, the performance of the receiver may be bounded. This bound, which is tight at high and low SNR may be determined analytically. The bounds on the performance of the noncoherent receiver are constructed in a manner similar to that used to analyze the coherent receiver in the previous section.

Low SNR Bound

The low SNR approximation to the optimum receiver makes use of the fact that for small arguments

$$I_0(X) \approx 1 + X^2/4. \tag{34}$$

Making this approximation in (31), describing the optimum processor, yields the low signal-to-noise processor

$$\sum_{i=1}^{m'} \left\{1 + \frac{1}{N_0^2} z_{1i}^2\right\} > \sum_{i=1}^{m'} \left\{1 + \frac{1}{N_0^2} z_{-1i}^2\right\} \tag{35}$$

implies a "1" was transmitted. Upon simplification this processor becomes

$$\sum_{i=1}^{m'} z_{1i}^2 \underset{\text{decide } -1}{\overset{\text{decide } 1}{\gtrless}} \sum_{i=1}^{m'} z_{-1i}. \tag{36}$$

It may be shown that the low SNR approximation processor described by (36) is mathematically equivalent to a pair of complex correlators. One correlator has as its reference the average of all transmitted waveforms containing a data 1 in the center bit interval The other correlator reference is the average of all waveforms with a data -1 in that bit interval. Thus, a test equivalent to (36) is

$$\left|\int r(t)\bar{s}(t,1)\,dt\right|^2 \underset{-1}{\overset{1}{\gtrless}} \left|\int r(t)\bar{s}(t,-1)\,dt\right|^2 \tag{37}$$

where

$$\bar{s}(t,1) = \sum_{i=1}^{m'} \exp\left[j\omega_0 t + \phi_{1i}(t)\right]$$

and

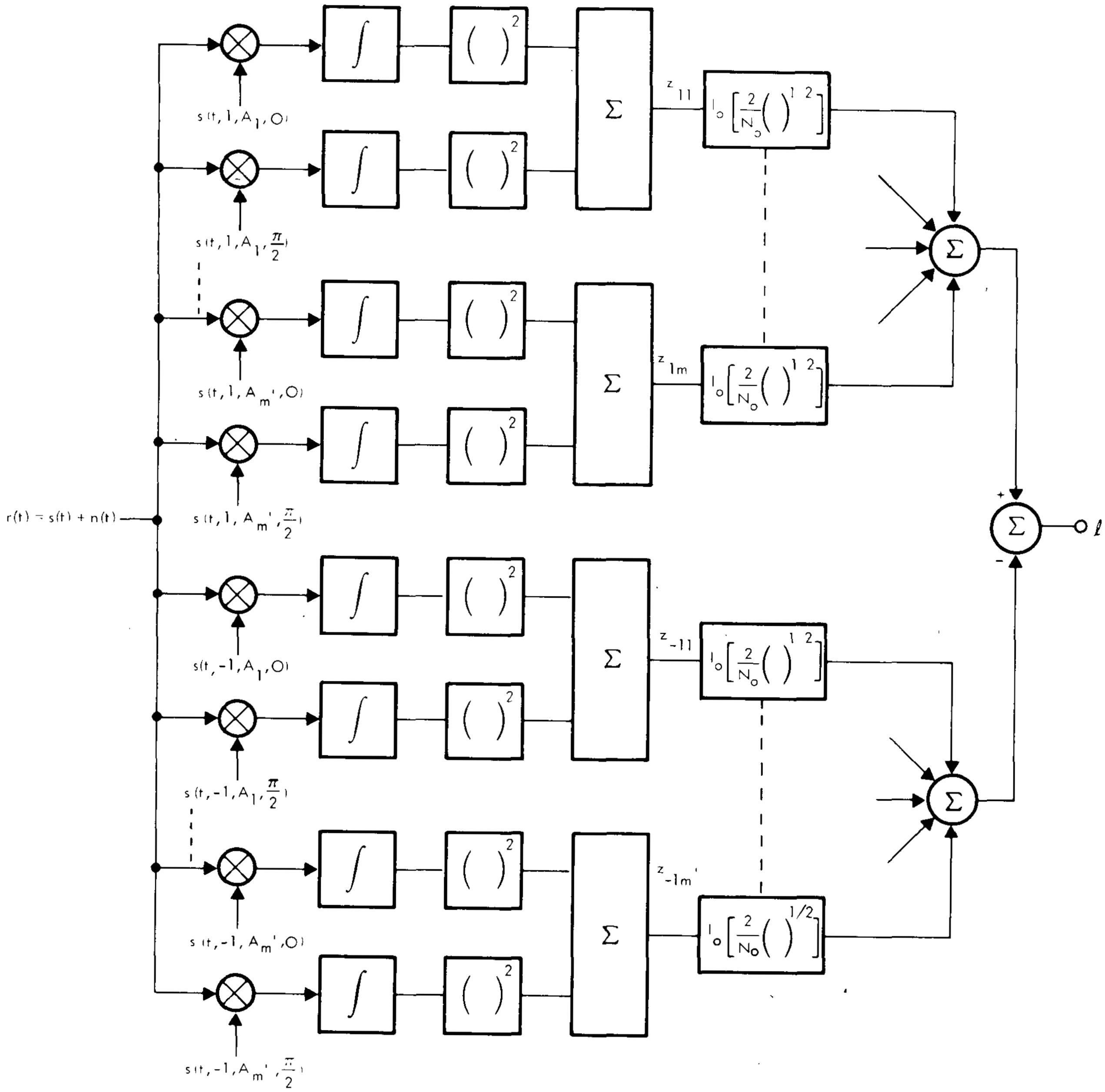

Fig. 6. Optimum noncoherent receiver.

$$\bar{s}(t,-1) = \sum_{i=1}^{m'} \exp\left[j\omega_0 t + \phi_{-1i}(t)\right].$$

The performance of this test may be computed by applying the results of Stein for the solution of the general binary noncoherent problem [6].

If z_1 and z_2 are two complex Gaussian variables with

$$M_1 = E(z_1)$$
$$M_2 = E(z_2)$$
$$\sigma^2 = \mathrm{Var}\,(z_1) = \mathrm{Var}\,(z_2)$$

and

$$\rho = \frac{1}{\sigma^2} E[(z_1 - M_1)^*(z_2 - M_2)]$$

then

$$\Pr\,(|z_2|^2 > |z_1|^2) = \tfrac{1}{2}[1 - Q(b^{1/2},a^{1/2}) + Q(a^{1/2},b^{1/2})] \tag{38}$$

where

$$\begin{Bmatrix} a \\ b \end{Bmatrix} = \frac{1}{2\sigma^2}\left[\frac{|M_1|^2 + |M_2|^2 - 2\,\mathrm{Re}\,(M_1 M_2^* \rho)}{1 - |\rho|^2} \pm \frac{|M_1|^2 - |M_2|^2}{(1 - |\rho|^2)^{1/2}}\right]. \tag{39}$$

The minus of the $\pm$ sign is used with a and the plus is used with b. The function $Q(x,y)$ is the Marcum Q function defined by

$$Q(x,y) = \int_y^\infty \exp\left(-\frac{x^2+w^2}{2}\right) I_0(xw) w\, dw \qquad (40)$$

where $I_0(\ )$ is the modified Bessel function. For a given input signal waveform, $\cos(\omega_0 t + \theta(t))$, the received signal is $r(t) = \cos(\omega_0 t + \theta(t)) + n(t)$. The variables z_1 and z_2, are the result of correlating $r(t)$ with $\bar{s}(t,-1)$ and $\bar{s}(t,1)$. Hence, the variables required to evaluate (39) become,

$$M_1 = E\left[\int \{\cos[\omega_0 t + \theta(t)] + n(t)\}\bar{s}(t,-1)\, dt\right]$$

$$= \int \cos[\omega_0 t + \theta(t)]\bar{s}(t,-1)\, dt$$

$$M_2 = E\left[\int \{\cos[\omega_0 t + \theta(t)] + n(t)\}\bar{s}(t,1)\, dt\right]$$

$$= \int \cos[\omega_0 t + \theta(t)]\bar{s}(t,1)\, dt,$$

$$\sigma^2 = E\left[\left\{\int \bar{s}(t,-1)n(t)\, dt\right\}^* \left\{\int \bar{s}(t,-1)n(t)\, dt\right\}\right]$$

$$= \frac{N_0}{2}\int |\bar{s}(t,-1)|^2\, dt,$$

and

$$\sigma^2\rho = E\left[\left\{\int \bar{s}(t,-1)n(t)\, dt\right\}^* \left\{\int \bar{s}(t,1)n(t)\, dt\right\}\right]$$

$$= \frac{N_0}{2}\int \bar{s}^*(t,-1)\bar{s}(t,1)\, dt. \qquad (41)$$

The complex correlator references $\bar{s}(t,-1)$ and $\bar{s}(t,1)$ may be found in a manner similar to that used in the coherent case. It is assumed that all possible signals have zero phase at the beginning of the middle bit and that time $t = 0$ corresponds to the beginning of the middle bit interval. For continuous phase FSK with a modulation index of h, during the middle bit the signal is $\exp(\pm j\pi ht)$ where the plus sign is used for a data one and the minus used for a data -1. During the next bit interval the average waveform is half the sum of the two possible waveforms, or

$$\bar{s}(t,1) = [\exp(j\pi h(t-T)) + \exp(-j\pi h(t-T))] \cdot \exp(j\pi h) \qquad T \le t \le 2T$$

for a data one in the middle interval and

$$\bar{s}(t,-1) = [\exp(j\pi h(t-T)) + \exp(-j\pi h(t-T)) \cdot \exp(-j\pi h) \qquad T \le t \le 2T$$

for a -1. In general, during the ith bit interval, after the middle bit, the average waveforms are

$$\left.\begin{array}{l}\bar{s}(t,1)\\ \\ \bar{s}(t,-1)\end{array}\right\} = \cos \pi h(t-iT)(\cos \pi h)^{i-1}\exp(\pm j\pi h)$$

$$T \le t \le (i+1)T \qquad (42)$$

where the $+$ sign in the exponential implies $\bar{s}(t,1)$.

Upon reversing the time axis, it is found by symmetry considerations that during the ith bit interval preceding the middle bit the average waveform is

$$\bar{s}(t,-1) = \bar{s}(t,1) = \cos \pi ht(\cos \pi h)^{i-1}$$

$$(i-1)T \le t \le iT. \qquad (43)$$

These equations may be used to compute σ^2 and ρ. Using (42) and (43) the integrals in (41) can be written as,

$$\sigma^2 = \frac{N_0}{2}\left[2\sum_{k=1}^{n}\int_0^1 \left\{\cos\frac{\pi ht}{T}(\cos \pi h)^{k-1}\right\}^2 dt + 1\right] \qquad (44)$$

and

$$\sigma^2\rho = \frac{N_0}{2}\left[\sum_{k=1}^{n}\int_0^1 (\cos \pi ht(\cos \pi h)^{k-1})^2\, dt + \int_0^1 \exp(-j2\pi ht) \cdot \sum_{k=1}^{n}\int_0^1 (\cos \pi ht(\cos \pi h)^{k-1})^2 \exp(-j2\pi h)\, dt\right].$$

Upon performing the indicated integrations and simplifying, it is found that

$$\sigma^2 = \frac{N_0}{2}\left\{(1 + \operatorname{sinc} 2h)\frac{1-\cos^{2n}(\pi h)}{1-\cos^2(\pi h)} + 1\right\} \qquad (45)$$

and

$$\rho = \frac{N_0}{\sigma^2}\left\{\tfrac{1}{2}(1 + \operatorname{sinc} 2h)(1 + \exp(-j2\pi h))\frac{1-\cos^{2n}(\pi h)}{1-\cos^2(\pi h)} + \exp(-j\pi h)\operatorname{sinc} h\right\}.$$

The mean outputs, M_1 and M_2, are dependent upon the input signal. Let the input bit sequence be $\{b_i\}$ with the index ranging from $-n$ to n. The middle bit is, therefore, b_0. Computation of M_1 and M_2 is performed in the same manner as before by computing the contribution due to each bit interval. Thus for $b_0 = 1$

$$M_2 = \sum_{i=1}^{n}\exp\left(-j\pi h\sum_{k=1}^{i-1} b_{-k}\right)\int_0^1 \cos \pi ht(\cos \pi h)^{i-1} \cdot \exp(-jb_{-i}\pi ht)\, dt + 1 + \sum_{i=1}^{n}\exp\left(-j\pi h\sum_{k=1}^{i-1} b_k\right) \cdot \int_0^1 \cos \pi ht(\cos \pi h)^{i-1}\exp(-jb_i\pi ht)\, dt \qquad (46)$$

and

$$M_1 = \sum_{i=1}^{n} \exp\left(-j\pi h \sum_{k=1}^{i-1} b_{-k}\right) \int_0^1 \cos \pi h t (\cos \pi h)^{i-1}$$

$$\cdot \exp\left(-j b_{-i} \pi h t\right) dt + \int_0^1 \exp\left(-j2\pi h t\right) dt$$

$$+ \exp\left(-j2\pi h\right) \sum_{i=1}^{n} \exp\left(-j\pi h \sum_{k=1}^{i-1} b_k\right)$$

$$\cdot \int_0^1 \cos \pi h t (\cos \pi h)^{i-1} \exp\left(-j b_i \pi h t\right) dt. \tag{47}$$

In the above equations, the sum $\sum_{i=1}^{i=0} x_i$ is defined to be zero. Evaluating the integrals of these equations yields

$$M_2 = A_1 + 1 + A_2$$

and

$$M_1 = A_1 + \exp\left(-j\pi h\right) \operatorname{sinc} h + \exp\left(-j2\pi h\right) \tag{48}$$

where

$$A_1 = \sum_{i=1}^{n} (\cos \pi h)^{i-1} \exp\left(-j\pi h \sum_{k=1}^{i-1} b_k\right) \left(1 + \operatorname{sinc}(h) \exp\left(-j\pi h b_{-i}\right)\right)$$

and

$$A_2 = \sum_{i=1}^{n} (\cos \pi h)^{i-1} \exp\left(-j\pi h \sum_{k=1}^{i-1} b_k\right) \left(1 + \operatorname{sinc}(h) \exp\left(-j\pi h b_i\right)\right). \tag{49}$$

When these equations are evaluated on a digital computer, a bound on the optimum receiver at low SNR is obtained. This bound is equivalent to the average matched filter bound shown in Fig. 4 for the coherent receiver. In the next section a union bound will be found which, when combined with the average matched filter bound, will yield a composite bound similar to that shown in Fig. 5.

High SNR Bound

The high SNR bound may be found by noting that for large arguments

$$\sum_i I_0(x_i) \approx I_0(x_l) \tag{50}$$

where x_l is the largest of the set $\{x_i\}$. With this approximation, the optimum detector described by (31) becomes

$$I_0\left(\frac{2}{N_0} z_{1l}\right) \underset{\text{Decide } -1}{\overset{\text{Decide } 1}{\gtrless}} I_0\left(\frac{2}{N_0} z_{-1k}\right) \tag{51}$$

where z_{1l} is the largest of $\{z_{1i}\}$ and z_{-1k} is the largest of $\{z_{-1i}\}$. Because $I_0(\ )$ is a monotonic function, (51) is equivalent to the test

$$z_{1l} \underset{\text{Decide } -1}{\overset{\text{Decide } 1}{\gtrless}} z_{-1k}. \tag{52}$$

The demodulator using the strategy of (52) could also choose the largest of all z_{ki} and then classify the largest as corresponding to a data 1 or a data -1. A decision error is made if, given a one was transmitted, one of the z_{-1i} was largest. Although an exact evaluation of the performance of this detector is not possible, the union bound will give a tight performance estimate at reasonably high SNR.

Suppose that a $2n + 1$ bit transmitted word is observed and that the middle bit is a data 1. The transmitted sequence, exclusive of the middle bit, is indicated by the index k so that an error is made if at least one of the $\{z_{-1j}\}$ is greater than z_{1k}. Then by the union bound

$$\Pr(\text{Error} \mid \text{Sequence } k \text{ Transmitted}) \leq \sum_{j=1}^{m'} \Pr(z_{-1j} > z_{1k}). \tag{53}$$

The average probability of error may now be computed by averaging over all transmitted sequences containing a one in the middle bit interval,

$$\Pr(\epsilon) = \frac{1}{m'} \sum_{k=1}^{m'} \Pr(\epsilon \mid \text{sequence } k \text{ was transmitted})$$

or

$$\Pr(\epsilon) \leq \frac{1}{m'} \sum_{k=1}^{m'} \sum_{j=1}^{m'} \Pr(z_{-1j} > z_{1k}). \tag{54}$$

In (54) the computation of the bounding performance of the detector described by (52) has been reduced to a binary error probability problem for which the solution is known [6]. For this situation the probability of error is

$$\Pr(z_{-1j} > z_{1k}) = 1/2[1 - Qb^{1/2}, a^{1/2}) + Q(a^{1/2}, b^{1/2})] \tag{55}$$

where

$$\begin{Bmatrix} a \\ b \end{Bmatrix} = \frac{S}{2N}\left[1 \pm (1 - |\rho|^2)^{1/2}\right]$$

and $S/2N$ is the SNR of z_{1k}. The value of ρ is the correlation between the transmitted waveforms corresponding to sequence j, with a data -1 in the middle bit interval, and sequence k, with a data 1 in the bit interval, and is given by

$$\rho = \frac{1}{2n+1} \sum_{k=1}^{2n+1} \exp\left(j \sum_{i=1}^{k-1} (a_i - b_i)\pi h\right) \cdot \exp\left(\frac{j\pi h}{2}(a_k - b_k)\right) \operatorname{sinc}\left(\frac{h}{2}(a_k - b_k)\right)$$

where $\{b_k\}$ is the kth bit sequence, with $b_{n+1} = 1$, and $\{a_k\}$ is the jth bit sequence, with $a_{n+1} = -1$.

Numerical Results—Noncoherent Case

The equations presented in this section for the bounding performance of the noncoherent receiver have been evaluated on a digital computer for three bit and five bit

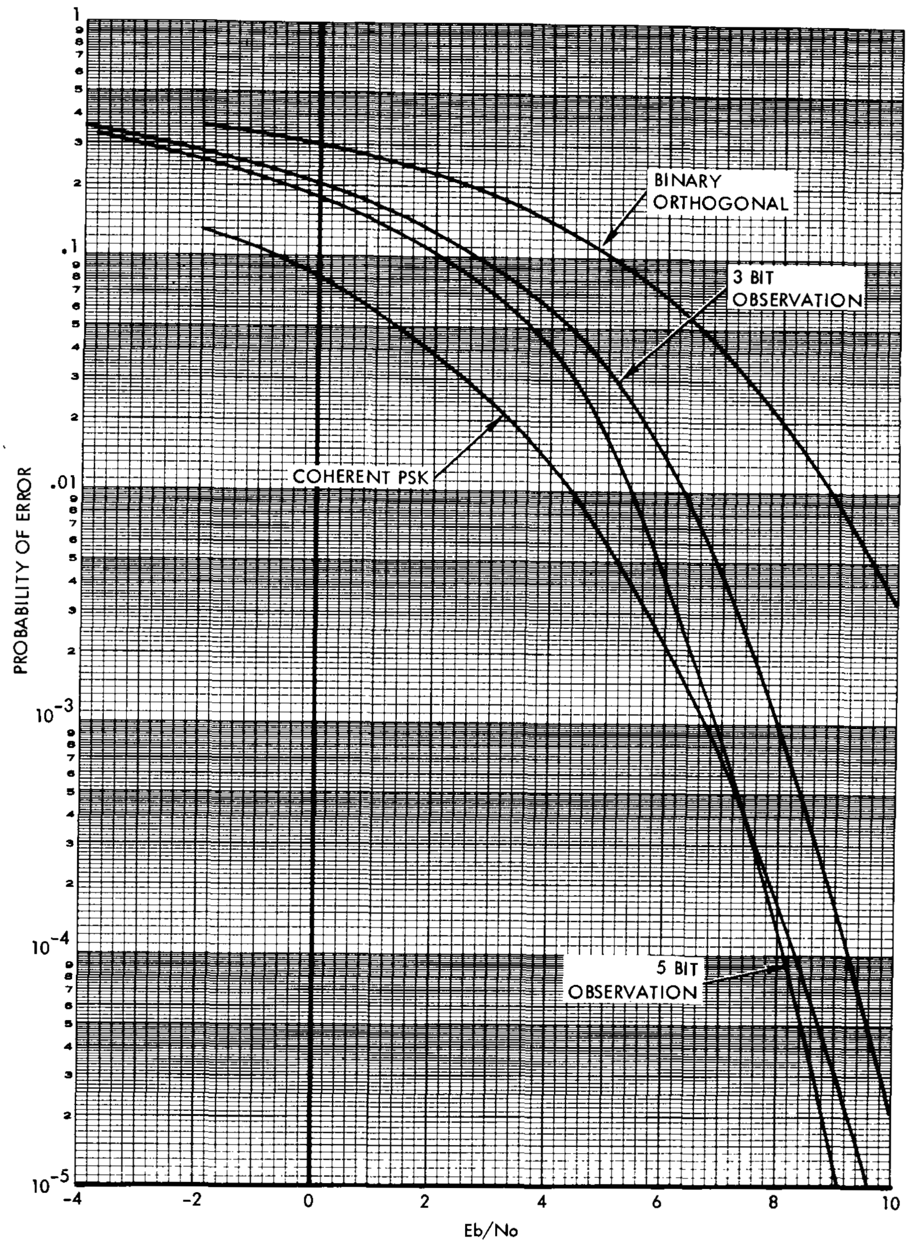

Fig. 7. Performance of noncoherent CPFSK receiver.

observation intervals for FSK with $h = 0.715$. The results are plotted in Fig. 7. Also plotted in Fig. 7 is the noncoherent detection performance for binary orthogonal signals and the coherent detection performance for antipodal signals. These two curves represent the best performance possible with single bit demodulation. Demodulation by observing five bits is seen to outperform PSK for E_b/N_0 greater than 8 dB. The performance of a demodulator observing three bits is within 1 dB of the performance of a coherent demodulator for probability of error less than 10^{-3}. In either case, five bit or three bit observation intervals, the demodulator performance significantly exceeds the performance of a single bit noncoherent demodulator.

The performance bounds for the multi-bit observation demodulator are tight at high and low SNR. It is felt that the bound is tight at all signal-to-noise ratios. In Fig. 8 the computed bound is compared with a digital computer simulation of the optimum receiver. The maximum difference between the bound and the simulation is about 1 dB at an E_b/N_0 of 4 dB. This demonstrates the quality of the bounding techniques employed for the analysis of CPFSK.

SUMMARY

In the previous two sections the structure and the performance of coherent and noncoherent receivers for CPFSK which minimize probability of bit error have been presented. The performance of the receivers was overbounded by employing the concepts of the union bound and the average matched filter bound. In both cases the joint bound was shown to be a good estimate of the actual performance available with CPFSK systems. In particular for the noncoherent case a digital computer

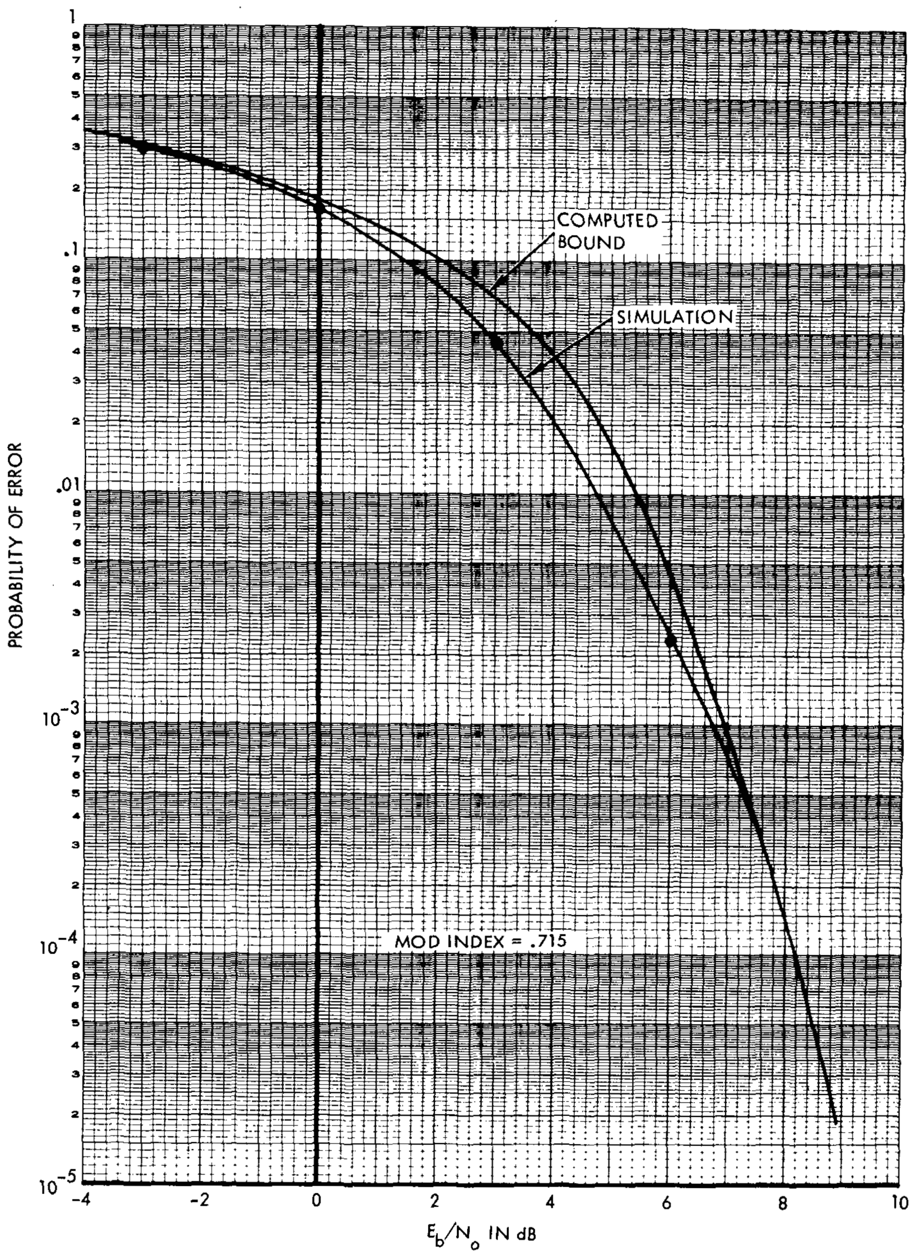

Fig. 8. Comparison of computed bound and simulation results.

simulation shows the differences between the bound and the actual performance to be less than 1 dB at any SNR and much less for most SNR's. Similarly, in the coherent case a lower bound was shown to differ from this composite upper bound by about 1.5 dB worst case and, hence, demonstrated the goodness of the bound for this case.

The equations presented serve to consolidate the performance calculations for CPFSK systems in that they provide a technique for computing performance which is applicable for all SNR's, all modulation indices and all observation intervals. The equations contain all of the previously published results and, in addition, allow the interested reader to investigate the performance of CPFSK systems with parameters for which previous results are not available.

The specific numerical results presented for a modulation index of 0.715 employing coherent and noncoherent detection serve to answer some questions about CPFSK and to demonstrate several points. In [1] the question of improving the performance of CPFSK at low SNR's by employing an observation longer than three bits was raised. This question is answered by the results in Fig. 5. There is improvement in performance at low SNR by allowing longer observation intervals, however, the improvement beyond a three bit interval is minor. In fact, for engineering purposes the three bit interval appears to be the optimum length for coherent receivers since the gain beyond this length is minimal and the complexity of the receiver grows rapidly with the length of the interval.[3]

The specific results presented for noncoherent detection of CPFSK show a new and rather interesting result. For a

[3] For modulation indices which are integers or integers plus a half, there is no gain beyond observation intervals of one and two bits, respectively.

modulation index of 0.715 and a five bit observation interval noncoherent CPFSK can outperform coherent PSK. This result raises the question of whether or not 0.715 is the best deviation ratio for noncoherent CPFSK systems. The answer is no; for example, the equations will show 2.7 to be slightly better. It does appear that 0.715 is the best compromise between bandwidth and performance, i.e., it is a local optimum for performance and it is the local optimum with the smallest modulation index.

The results presented above show that an observation interval of five bits is essentially optimum for noncoherent detection. A comparison of Fig. 4 and Fig. 7 shows that the coherent receiver and the five bit noncoherent receiver are equivalent in performance to within 0.5 dB. This shows that extending the observation interval to more than five bits cannot improve the performance of the noncoherent receiver by more than 0.5 dB.

This paper and previous results show that with coherent detection CPFSK can outperform coherent PSK which is at present the favored modulation technique for use on thermal noise limited channels. However, at this time there is no simple technique for obtaining the reference signals required for coherent detection of CPFSK from the received waveform for the modulation index (0.715) which produces the best performance or for any modulation indices except integers and integers plus one-half. This poses an interesting research question and limits the practical significance of the coherent results at present.

In this paper it was also shown that noncoherent detection of CPFSK can perform slightly better than coherent PSK. This result combined with the fact that CPFSK has a power spectrum which is superior to PSK in terms of percent power contained in a given bandwidth should make CPFSK with noncoherent detection an attractive modulation scheme for channels whose performance is limited by thermal noise. The noncoherent receiver does not possess the synchronization problems of the coherent structure and it is realizable using available technologies such as surface wave devices and digital filters.

ACKNOWLEDGMENT

The authors gratefully acknowledge numerous conversations with M. G. Pelchat on the general subject of continuous phase FM.

REFERENCES

[1] M. G. Pelchat, R. C. Davis, and M. B. Luntz, "Coherent demodulation of continuous phase binary FSK signals," in *Proc. Int. Telemetry Conf.*, Washington, D. C., 1971.
[2] R. DeBuda, "Coherent demodulation on frequency-shift keying with low deviation ratio," *IEEE Trans. Commun.* (Concise Paper), vol. COM-20, pp. 429–435, June 1972.
[3] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268–278, Mar. 1973.
[4] M. G. Pelchat and S. L. Adams, "Noncoherent detection of continuous phase binary FSK," in *Proc. Int. Conf. Communications*, Montreal, P.Q., Canada, 1971.
[5] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part I, Section 4.4. New York: Wiley, 1968.
[6] M. Schwartz, W. Bennett, and S. Stein, *Communications Systems and Techniques*, Section 8.2. New York: McGraw-Hill, 1966.

★

William P. Osborne (S'64–M'70) was born in Paducah, Ky., on February 2, 1944. He received the B.S.E.E. and M.S.E.E. degrees from the University of Kentucky, Lexington, in 1966 and 1967, respectively, and the Sc.D. degree from New Mexico State University, Las Cruces, in 1970.

During the summer of 1966 he worked for Collins Radio Company, Cedar Rapids, Iowa. From 1967 until 1970 he worked with the Communications Research Group at New Mexico State University studying various aspects of the Apollo Communications for NASA. Since 1970, he has been employed by Electronic Systems Division of Harris, formerly Radiation, Melbourne, Fla. He is presently the head of the Signal Processing Group in the Advanced Technology Department of the Electronic Systems Division. This group is working in the area of digital signal processing techniques.

★

Michael B. Luntz was born in Ponca City, Okla., on February 24, 1940. He received the B.S.E.E. degree from Rice University, Houston, Tex., in 1963.

From 1963 to 1964 and 1966 to 1968 he was employed at the Westinghouse Defense and Space Center, Baltimore, Md. While at Westinghouse, he participated in design and development of general-purpose computer peripheral and central processor hardware. Between 1964 and 1966 he served in the U. S. Army. Since 1968 he has been with Electronic Systems Division of Harris Corporation, formerly Radiation, Melbourne, Fla. As a member of the Advanced Technology Department at Radiation, he has been concerned with analysis of a wide range of statistical communications and signal processing problems.