EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

SIGNAL DETECTION OF DELTA-CODED SPEECH

B.G. Taylor

(to be published in Proceedings IEE)

EP Division

Geneva,
March 1976

SIGNAL DETECTION OF DELTA-CODED SPEECH

B.G. Taylor, B.Sc., Ph.D., C.Eng., M.I.E.E.

Indexing terms: Delta modulation, Signal detection, Speech transmission

Abstract

A signal detection method for delta-coded speech is presented which exploits the redundancy of the source to reduce the error rate due to additive Gaussian noise in the channel. The method implements a Bayes' strategy with dual decision boundaries selected by a 6th-order digital predictor. Fixed decision boundary pairs are established which are appropriate for a wide range of input s.n.r. This easily implemented technique is shown to result in an error rate reduction of about 40% in coherent antipodal and coherent and noncoherent orthogonal communications.

List of symbols

α = component error probability

b = detector bias level

B_n = detected nth binary signal element

β = component error probability

Δ = quantization interval

E_n = binary prediction of nth signal element

$E[\ ]$ = expected value of variant

θ_i = signal transmitted for symbol i

I_0 = zero-order modified Bessel function of the first kind

K = decision threshold

Λ = likelihood ratio

M = mean signal power

μ = difference between expected values of x_c for θ_0 and θ_1

ν = argument of modified Bessel function

P_s = predictor success probability

P_T = total error probability

P_{Tf} = fixed bias error probability

P_{Tv} = variable bias error probability

ρ = signal set cross-correlation coefficient

SNR_i = received signal average signal-to-noise ratio

σ = r.m.s. channel noise level

σ_c = standard deviation of conditional distribution of correlator output at end of integration period

τ = signal element interval

x = received signal

x_c = correlator output at end of integration period

x_d = decision boundary for coherent signalling

x_{dn} = decision boundary for noncoherent signalling

x_n = envelope detector output signal

Abbreviations

a.g.c. = automatic gain control

ΔM = delta modulation

f.m. = frequency modulation

f.s.k. = frequency shift keying

s.n.r. = signal-to-noise ratio

1 Introduction

The optimal fixed-structure predictors for speech encoded by delta modulation have been determined(1) for Markov process approximations to the source of order 1 to 7. Their performance has been evaluated with a view to implementing the classical channel encoder operation of reduction of the 'natural' source redundancy by predictive coding(2), followed by the addition of 'artificial' redundancy in the form of the check digits of an error-correcting code.

As an alternative application, it is interesting to consider exploiting the a priori information made available by such a predictor at a receiver to optimise the signal detection of a normal ΔM message. While the performance gains attainable in this way are less than when channel encoding is implemented, the method can be very simply applied to existing delta-coded speech circuits such as troposcatter links(3) to secure an error rate reduction without any change of the coding and transmission procedure.

A ΔM receiver improvement has been proposed by Tanaka et al(4), who show that for a sine wave signal an error rate reduction is achieved if the channel signal is replaced by a simple 1-element prediction (the complement of the preceding element) when the former is in the vicinity of the detector decision threshold. In this paper, a receiver structure is presented which employs selection of the decision threshold itself by a 6-element predictor to achieve optimal (minimum error probability) detection of speech messages.

2 Optimal receiver structure

For a ΔM receiver performing bit-by-bit detection of signal elements from channel signals impaired by additive Gaussian noise, a Bayes' strategy of minimizing the average risk is appropriate. Equal costs may reasonably be assigned to the errors occurring when the received signal x is judged to have been caused by transmitted signal θ_0 or θ_1

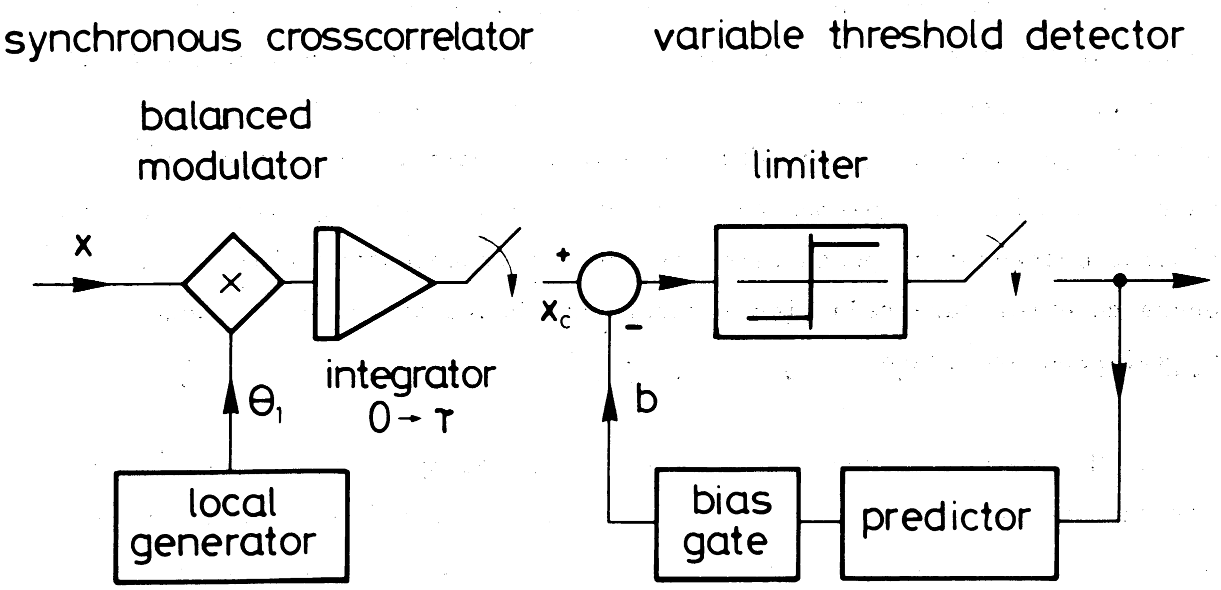

Fig. 1

Optimal coherent receiver structure (pulse transmission case) when the reverse is true, and the threshold K with which the likelihood ratio

$$\Lambda(x) = \frac{p(x|\theta_1)}{p(x|\theta_0)} \tag{1}$$

is to be compared is then simply

$$K = \frac{p(\theta_0)}{p(\theta_1)} \tag{2}$$

Since in this case the total probability of error is minimized, the strategy is also that of Siegert's 'ideal observer'[6].

In a typical coherent discrete communication system, the signals θ_0, θ_1, with cross-correlation coefficient ρ, are either orthogonal ($\rho = 0$) or antipodal ($\rho = -1$). Their replicas, either locally generated or stored as the impulse responses of two matched filters*, are cross-correlated in the Bayes' receiver structure with the signal delivered by the channel. To the difference x_c between the correlator outputs at the end of each signal element interval τ is added the negative of a bias level b corresponding to the Bayes' test decision boundary and the sum is limited and sampled to generate the detector output sequence.

Fig. 1 indicates this structure for the $\rho = 0$ case of on - off pulse transmission, in which case only one correlator is required as signal $\theta_0 = 0$, and the probability density distributions $p(x_c|\theta_i)$, i = 0, 1, are shown in Fig.2.

By considering average signal power M (i.e. 'mark' power = 2M), the performance evaluation applies equally for the case of equi-energy orthogonal signalling (e.g. coherent f.s.k.) in which a second correlator is required for θ_0 and the distributions for x_c are translated for symmetry about the origin.

---

* For white noise, the filters have impulse responses of the form of the signals θ_0, θ_1, run backwards in time from $t = \tau$.

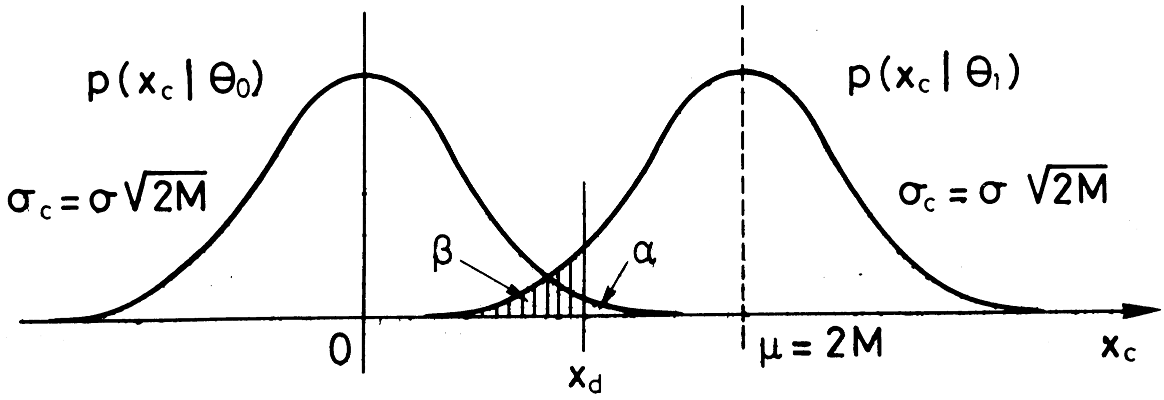

Fig. 2

Signal probability density distributions

$E[x_c]$ for θ_1 transmitted is then

$$\mu = 2M(1-\rho) \tag{3}$$

and for channel noise power σ^2 the conditional distributions, which remain Gaussian because of the linearity of the correlation detector, have standard deviation

$$\sigma_c = \sigma\sqrt{2M(1-\rho)} . \tag{4}$$

3 Decision boundary loci

From (1) and (2), the decision boundary x_d is found by setting

$$K = \Lambda(x_d) = \frac{p(x_d \mid \theta_1)}{p(x_d \mid \theta_0)} = \frac{\frac{1}{\sqrt{2\pi}\,\sigma_c}\exp\left[-\frac{(x_d-\mu)^2}{2\sigma_c^2}\right]}{\frac{1}{\sqrt{2\pi}\,\sigma_c}\exp\left[-\frac{x_d^2}{2\sigma_c^2}\right]}$$

$$= \exp\left[\frac{\mu}{\sigma_c^2}\left(x_d - \frac{\mu}{2}\right)\right] \tag{5}$$

from which

$$x_d = \frac{\mu}{2} + \frac{\sigma_c^2}{\mu}\log_e K \tag{6}$$

By means of a ΔM signal processor, state transition probability matrices for Markov process approximations to ΔM speech have been generated, and a representative example is shown in Table 1. The processor, which is implemented with a hybrid configuration of dual analog computers interfaced to a digital minicomputer, operates on real speech messages which have been scaled in frequency by an f.m. magnetic tape system. The vector state of the source at any time is defined by a finite number of the symbols which have been generated prior to that time. The operand and transform states are concisely identified in Table 1 by the corresponding octal numbers, while the conditional probabilities are decimal fractions. For an unconstrained random signal

Table 1

4TH - ORDER TRANSITION PROBABILITY MATRIX FOR DELTA - CODED SPEECH

Sampling frequency = 96 kHz   Pk - pk message amplitude = 128Δ

↓	00	01	02	03	04	05	06	07	10	11	12	13	14	15	16	17
00	0.895	0	0	0	0	0	0	0	0.468	0	0	0	0	0	0	0
01	0.105	0	0	0	0	0	0	0	0.532	0	0	0	0	0	0	0
02	0	0.963	0	0	0	0	0	0	0	0.965	0	0	0	0	0	0
03	0	0.037	0	0	0	0	0	0	0	0.035	0	0	0	0	0	0
04	0	0	0.437	0	0	0	0	0	0	0	0.077	0	0	0	0	0
05	0	0	0.563	0	0	0	0	0	0	0	0.923	0	0	0	0	0
06	0	0	0	0.832	0	0	0	0	0	0	0	0.813	0	0	0	0
07	0	0	0	0.168	0	0	0	0	0	0	0	0.187	0	0	0	0
10	0	0	0	0	0.198	0	0	0	0	0	0	0	0.149	0	0	0
11	0	0	0	0	0.802	0	0	0	0	0	0	0	0.851	0	0	0
12	0	0	0	0	0	0.923	0	0	0	0	0	0	0	0.579	0	0
13	0	0	0	0	0	0.077	0	0	0	0	0	0	0	0.421	0	0
14	0	0	0	0	0	0	0.036	0	0	0	0	0	0	0	0.037	0
15	0	0	0	0	0	0	0.964	0	0	0	0	0	0	0	0.963	0
16	0	0	0	0	0	0	0	0.523	0	0	0	0	0	0	0	0.094
17	0	0	0	0	0	0	0	0.477	0	0	0	0	0	0	0	0.906

source, each operand state column vector of the matrix contains two non-zero elements of value 0.5, while the greater and lesser transition probabilities recorded for ΔM speech messages reflect the redundancy of these signal sequences.

On the basis of such a characterisation of the message source, it is possible to process groups of past detector output elements to generate predictions $p(\theta_0)$, $p(\theta_1)$ for the following θ_i. Using this additional a priori information, the detector bias level can be switched so that the decision boundary location for each element corresponds to comparison of the likelihood ratio with the optimal threshold.

From (3), (4) and (6), the normalised boundary location for the orthogonal case is given by

$$\frac{x_d}{\mu} = \frac{1}{2} + \frac{1}{2\,SNR_i} \log_e \frac{p(\theta_0)}{p(\theta_1)} \qquad (7)$$

in which the input average signal-to-noise ratio

$$SNR_i = \frac{M}{\sigma^2} \qquad (8)$$

and decision boundary loci for a range of $p(\theta_0)$ and SNR_i are given in Fig. 3.

To simplify the implementation, we can consider the use of a predictor which makes only a binary estimate E_n of the more probable nth symbol. For example, a 6th-order predictor with average success probability $P_s = 0.898$ is defined[1] by

$$E_n = B_{n-6}\cdot B_{n-5}\cdot B_{n-4}\cdot B_{n-2} + B_{n-4}\cdot \overline{B}_{n-1} + B_{n-2}\cdot \overline{B}_{n-1}\cdot (B_{n-3}+B_{n-5})$$

$$+ B_{n-4}\cdot B_{n-3}\cdot (B_{n-2}+B_{n-5}) + B_{n-3}\cdot \overline{B}_{n-2}\cdot \overline{B}_{n-1}\cdot (B_{n-6}+B_{n-5}) \qquad (9)$$

in which the order of reception of the binary digits is ...B_{n-2}, B_{n-1}, B_n... In this case the bias level is switched between values

$$\frac{b}{\mu} = \frac{1}{2} \pm \frac{1}{2\,SNR_i} \log_e \frac{P_s}{1 - P_s} \qquad (10)$$

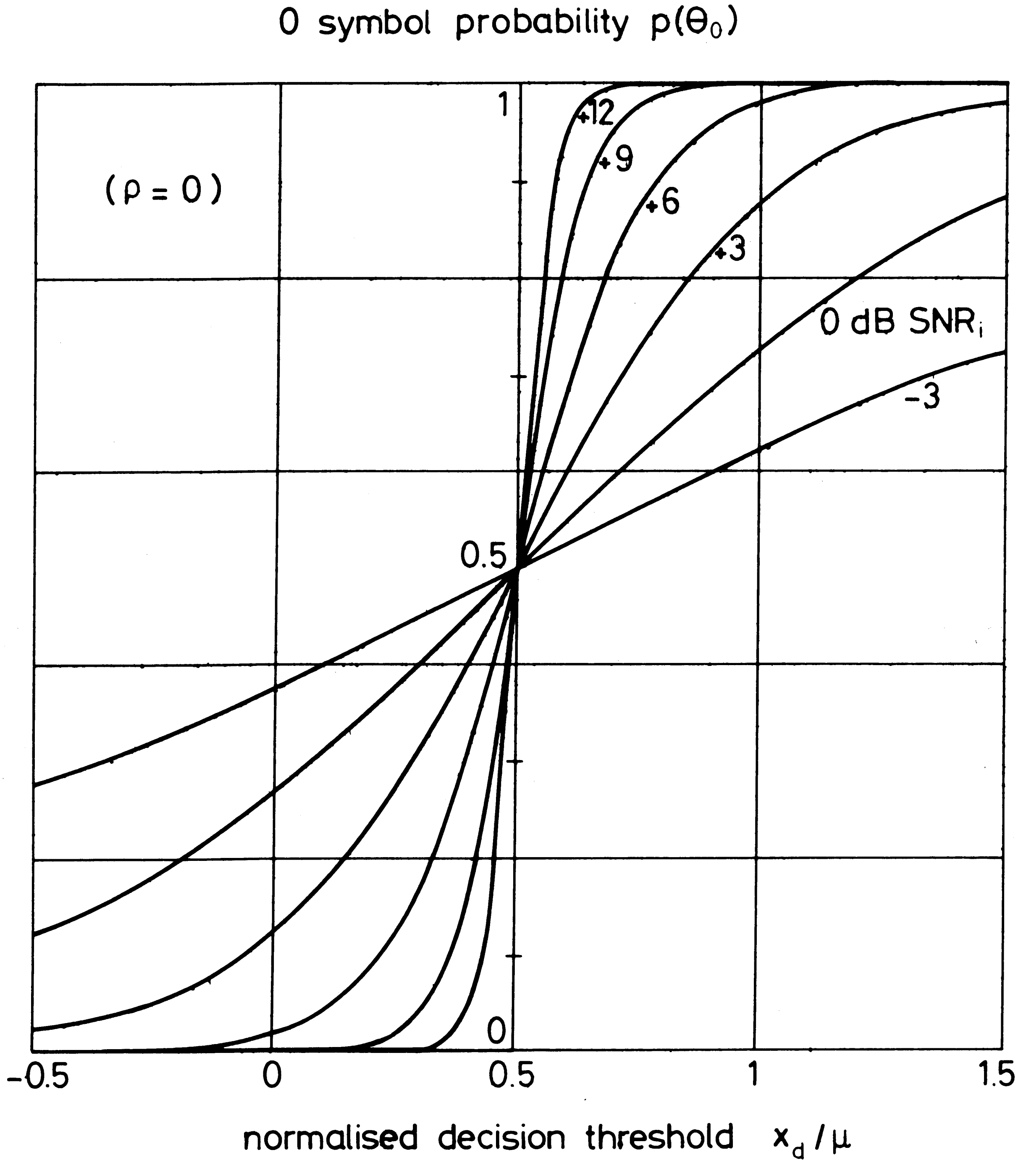

Fig. 3

Decision boundary loci

with sign + or - according as a 0 or 1 symbol is predicted. The bias characteristics for $P_s = 0.898$ are shown in Fig. 4, in which the apparent rather critical dependence of the optimum levels on SNR_i is removed by considerations which follow.

4 Error rates

For transmitted signal θ_0, signal detection errors occur when $x_c > x_d$, which occurs with probability

$$\alpha = \int_{x_d}^{\infty} p(x_c \mid \theta_0)\, dx_c = \int_{\frac{\mu}{2}\left(1 + \frac{1}{SNR_i} \log_e \frac{p(\theta_0)}{p(\theta_1)}\right)}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma_c} \exp\left(-\frac{x_c^2}{2\sigma_c^2}\right) dx_c$$

$$= \frac{1}{2}\left[1 - \operatorname{erf}\left(\frac{\sqrt{SNR_i}}{2} + \frac{1}{2\sqrt{SNR_i}} \log_e \frac{p(\theta_0)}{p(\theta_1)}\right)\right] \tag{11}$$

While for θ_1, error probability

$$\beta = \frac{1}{2}\left[1 - \operatorname{erf}\left(\frac{\sqrt{SNR_i}}{2} - \frac{1}{2\sqrt{SNR_i}} \log_e \frac{p(\theta_0)}{p(\theta_1)}\right)\right] \tag{12}$$

The total error probability is then

$$P_T = p(\theta_0)\alpha + p(\theta_1)\beta \tag{13}$$

For the case of fixed bias detection of a ΔM signal with $p(\theta_0)/p(\theta_1) = 1$, (13) becomes

$$P_{Tf} = \frac{1}{2}\left[1 - \operatorname{erf}\left(\frac{\sqrt{SNR_i}}{2}\right)\right] \tag{14}$$

which is shown in Fig. 5, together with the corresponding error characteristic for θ_i antipodal ($\rho = -1$ in (3) and subsequently).

Employing bias switching by (10), the total error probability becomes

$$P_{Tv} = \frac{1}{2}\left[1 - P_s \operatorname{erf}\left(\frac{\sqrt{SNR_i}}{2} + \frac{1}{2\sqrt{SNR_i}} \log_e \frac{P_s}{1-P_s}\right) + (P_s - 1)\operatorname{erf}\left(\frac{\sqrt{SNR_i}}{2} - \frac{1}{2\sqrt{SNR_i}} \log_e \frac{P_s}{1-P_s}\right)\right] \tag{15}$$

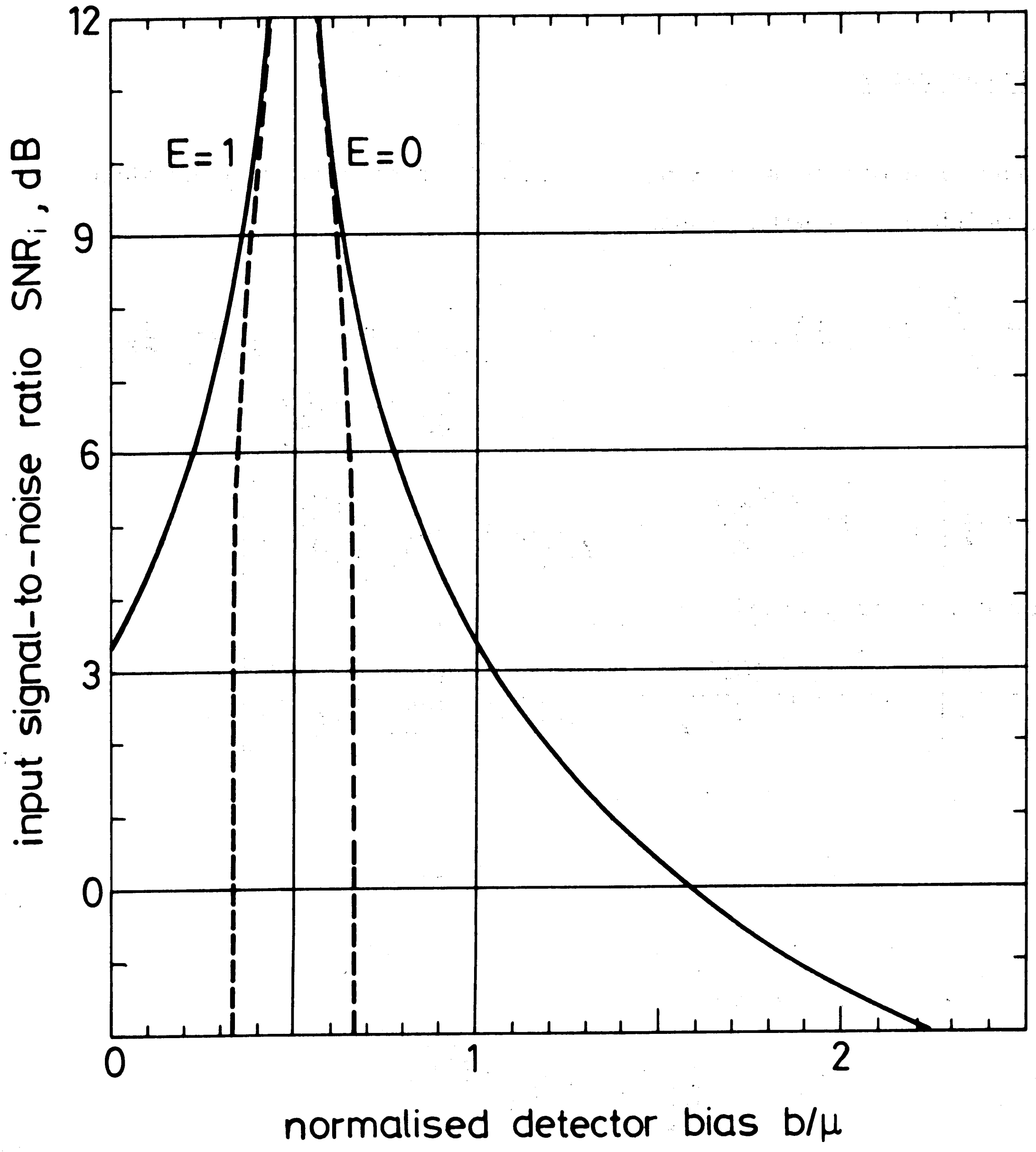

Fig. 4

Detector bias characteristics

———) P_s = 0.898 -------) 6th-order predictor

For the 6th-order predictor, P_s is a function of the data error rate which has been determined[1]. By a set of iterations for a range of channel SNR_i, bias levels and data error rates are computed from these data by (10) and (15) and the results are shown in Figs. 4 and 5. The use of bias switching is found to result in a lower error rate in detection for all values of SNR_i.

The effect of the degradation of P_s at low SNR_i is to make the normalised decision boundary locations near constant, and detector bias switching between levels

$$\frac{b}{\mu} = 0.34,\ 0.66 \qquad E = 1,\ 0 \tag{16}$$

is found to be near optimal for a wide range of SNR_i. With these levels, an error rate reduction of 39 - 43% results over the typical range in the case of both orthogonal and antipodal signalling.

5 Noncoherent signalling

In some digital communications situations noncoherent signalling is dictated by equipment instabilities or transmission path fluctuations. The optimal signal set is then restricted to an orthogonal pair, and typically on - off transmitter modulation is employed, while at the receiver band-pass filtering and envelope detection of the pulsed carrier is effected prior to the threshold comparator.

The probability density distribution for the envelope detector output x_n for a signal in noise has been studied by Rice[7] and

$$p(x_n|\theta_1) = \frac{x_n}{\sigma^2}\, I_0\!\left(\frac{\sqrt{2}\, x_n \sqrt{SNR_i}}{\sigma}\right) \exp\left[-\left(\frac{x_n^2}{2\sigma^2} + SNR_i\right)\right] \tag{17}$$

while

$$p(x_n|\theta_0) = \frac{x_n}{\sigma^2} \exp\left(-\frac{x_n^2}{2\sigma^2}\right) \qquad (I_0(0) = 1) \tag{18}$$

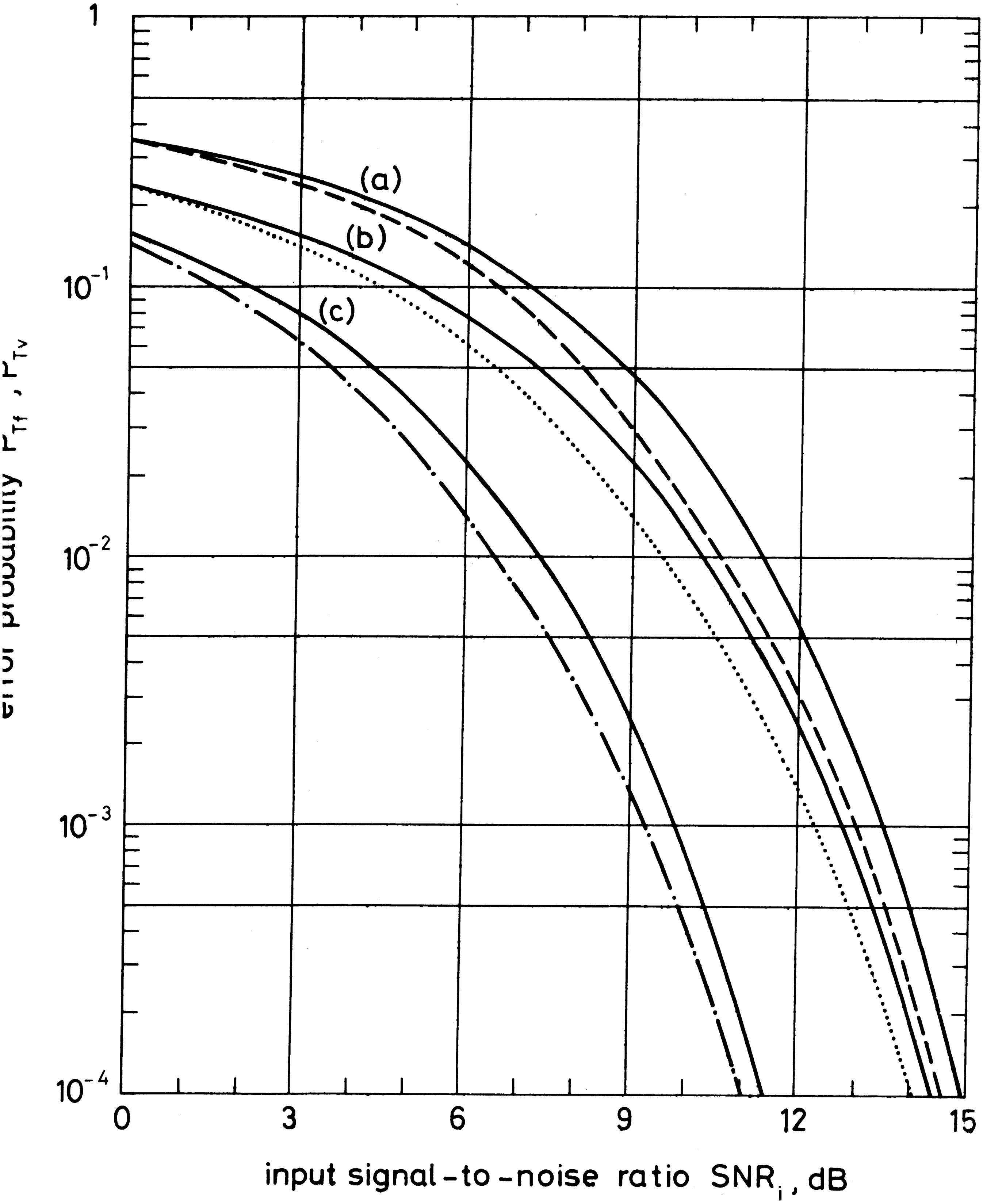

Fig. 5

Error probabilities

(a) Noncoherent (b) Coherent orthogonal (c) Coherent antipodal

Continuous curves: fixed bias P_{Tf} Broken curves: switched bias P_{Tv}

Hence for the Bayes' criterion decision boundary x_{dn},

$$I_0\left(\sqrt{2}\,\frac{x_{dn}}{\sigma}\sqrt{SNR_i}\right)\exp(-SNR_i) \quad = \quad \frac{p(\theta_0)}{p(\theta_1)} \tag{19}$$

from which x_{dn}/σ for switched bias detection may be determined by iterative solution of the approximation

$$\frac{\exp(\nu)}{\sqrt{2\pi\nu}}\left(1 + \frac{1}{8\nu} + \frac{9}{2(8\nu)^2}\right) \quad = \quad \frac{1 \pm (2P_s - 1)}{1 \pm (1 - 2P_s)}\exp(SNR_i) \tag{20}$$

in which ν is the argument of the zero-order modified Bessel function of the first kind in (19) and signs + or - apply for prediction E = 0 or 1 as before.

The conditional probabilities are then

$$\alpha \quad = \int_{\frac{x_{dn}}{\sigma}}^{\infty} \frac{x_n}{\sigma}\exp\left(-\frac{x_n^2}{2\sigma^2}\right)\frac{dx_n}{\sigma} \quad = \quad \exp\left[-\frac{1}{2}\left(\frac{x_{dn}}{\sigma}\right)^2\right] \tag{21}$$

and

$$\beta \quad = \int_{0}^{\frac{x_{dn}}{\sigma}} \frac{x_n}{\sigma} I_0\left(\sqrt{2}\,\frac{x_n}{\sigma}\sqrt{SNR_i}\right)\exp\left[-\left(\frac{x_n^2}{2\sigma^2} + SNR_i\right)\right]\frac{dx_n}{\sigma} \tag{22}$$

in the computation of which the approximation $I_0(\nu) = \exp(-\nu^2/4)$ may be used for the range of integration resulting in $\nu \leq 1$ and the expansion of (20) for the remainder. (Both approximations incur an error of about 2% at $\nu = 1$).

For decision boundary iterative solutions of (20) for trial values of P_s, total error probabilities are computed from (21) and (22). These are used to revise the estimates by the known degradation of P_s with error rate, and by the repetitive application of this procedure the error characteristic for noncoherent signalling shown in Fig. 5 is determined. The improvement is found to approach the performance attainable by coherent reception of the same transmitted signals, with less complexity than that necessary for a phase-locked demodulator and without imposing the equipment and path stability requirements of the latter.

$E\left[x_n | \theta_1\right]$ in the noncoherent case is

$$\mu = d\sqrt{\frac{\pi}{2}}\left[\exp\left(-\frac{SNR_i}{2}\right)\right]\left[(1+SNR_i)\, I_0\left(\frac{SNR_i}{2}\right) + SNR_i\, I_1\left(\frac{SNR_i}{2}\right)\right], \tag{23}$$

and when the optimal x_{dn} are normalised to this level it is again found that the decision boundary locations are not a sensitive function of SNR_i, and remain close to

$$\frac{b}{\mu} = 0.50,\ 0.71 \qquad E = 1,\ 0 \tag{24}$$

over the typical range. Unlike those of (16), the bias levels are not symmetric about $\mu/2$ because the Rayleigh distribution of the detector output for θ_0 has non-zero mean. To maintain minimum error probability reception during received signal strength fluctuations, the levels may readily be tapped in the ratios (24) from a potential divider chain across which the mean detector output for θ_1 is developed by conventional gated a.g.c. methods.

6 Acknowledgments

The author is indebted to Prof. W.E.J. Farvis of the University of Edinburgh for valuable discussions of this work, and to Prof. S. Michaelson and the staff of Edinburgh Regional Computing Centre for their helpful co-operation. The paper presents material from a part of a Ph.D. dissertation entitled 'The Entropy of Delta-coded Speech' submitted by the author in May 1968 to the Faculty of Science of the University of Edinburgh. The research was supported by the Science Research Council and Ferranti Ltd.

7 References

1 TAYLOR, B.G.: 'The entropy of delta-coded speech', Proc. IEE, (to be published as Paper E2775)

2 ELIAS, P.: 'Predictive coding', IRE Trans., 1955, IT-1, pp. 16 - 33

3 BRAUN, W., ENGELBRECHT, L., MAGNUSKI, H., and STEEL, F.R., 'Delta-multiplex for long troposcatter links', IEEE Int. Conv. Rec., 1965, part 2, pp. 62 - 76

4 TANAKA, Y., YAMASHITA, K., and HOSOKAWA, S.: 'A new decoding method of delta modulation code, taking the redundancy of codes into consideration', Mem. Fac. Eng., Osaka City Univ., 1963, 5, pp. 43 - 55

5 HANCOCK, J.C., and WINTZ, P.A.: 'Signal detection theory', (McGraw-Hill, 1966)

6 LAWSON, J.L., and UHLENBECK, G.E.: 'Threshold signals', (MIT Rad. Lab. Series, 24, McGraw-Hill, 1950)

7 RICE, S.O.: 'Mathematical analysis of random noise, part 3', Bell Syst. Tech. J., 1945, 24, pp. 46 - 108