X-520-67-360

NASA TM X-55900

A COMPARISON OF DECISION SYSTEMS WITH DATA REJECTION

J. C. MORAKIS

JULY 1967

GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

X-520-67-360

A COMPARISON OF DECISION SYSTEMS WITH DATA REJECTION

J. C. Morakis

July 1967

GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland

ABSTRACT

The advent of convolutional codes provides us with extremely low error rates at the expense of data rejection. These codes are compared to a perfect code and to a decision system capable of variable deletion rates. The deletion rates can be adjusted to values equal to those of convolutional codes resulting thus in a fair comparison.

CONTENTS

A COMPARISON OF DECISION SYSTEMS WITH DATA REJECTION

1. INTRODUCTION

It has been shown (References 1 and 2), that convolutional codes (References 3 and 4) can reduce error rates to extremely low values. The only drawback is the necessity to reject data. This deletion rate, P_{del}, is a function of E/N_0 and the resulting probability of error. The above results are shown in Figure 1. Unfortunately this type of graph does not offer a fair comparison between the convolutional codes and the algebraic ones, or other detection systems due to the fact that the curves representing the algebraic codes do not suffer from any rejection of the data. For example in the case of the biorthogonal $n = 8$ we see that for $E/N_0 = 4$ db

$$P_e = 7.214 \times 10^{-4} \text{ and } P_c = 1 - P_e = .9992786$$

On the other hand for the convolutional code and the same signal to noise ratio

$$P_e = 3 \times 10^{-4} \text{ and } P_{del} = .03$$

thus

$$P_c = 1 - P_e - P_{del} = .9697$$

The above simple calculation shows that although the convolutional code offers a smaller probability of error, the algebraic code offers a higher probability of correct detection.

The performance of one code over another can be described by the average risk function $B(P_e, P_c, C_1, C_2, C_3)$ where C_1, C_2, C_3 are cost functions, and B can be simply expressed as

$$B = C_1 P_e + C_2 P + C_3 P_{deletion}$$

$$= C_1 P_e + C_2 P_c + C_3 (1 - P_c - P_e)$$

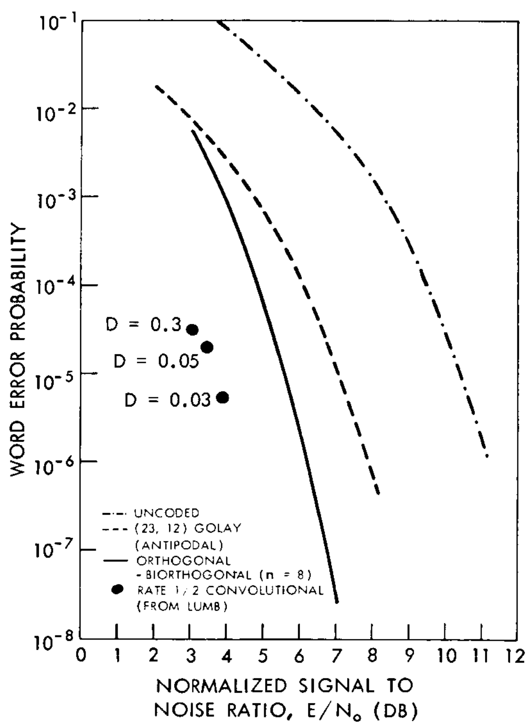

Figure 1—Comparative Performance of Communication Systems.

Thus unless C_1, C_2 and C_3 are known one cannot pass a judgment on the relative performance of two systems. However if one of either P_e, P_c, or $P_{deletion}$ could be made the same for the two systems in question (of course under equivalent E/N_0), then a comparison is possible without the knowledge of the C_i's.

For example if the probability of deletion (D) is made the same for both systems then the two B's become

$$B_1 = C_1 P_{e1} + C_2 P_{c1} + C_3 D$$

$$B_2 = C_1 P_{e2} + C_2 P_{c2} + C_3 D$$

then

$$B_1 - B_2 = C_1 \left(P_{e1} - P_{e2}\right) + C_2 \left(P_{c1} - P_{c2}\right) \tag{1}$$

but

$$P_{ci} = 1 - P_{ei} - D \ . \tag{2}$$

Substituting (2) in (1) we get

$$\begin{aligned} B_1 - B_2 &= C_1 \left(P_{e1} - P_{e2}\right) + C_2 \left(1 - P_{e1} - 1 + P_{e2}\right) \\ &= \left(C_1 - C_2\right) \left(P_{e1} - P_{e2}\right) \ . \end{aligned}$$

Since $C_1 - C_2 > 0$* then system 2 is better than system 1 if $P_{e2} < P_{e1}$. Thus we have showed that one way to compare the two systems is to keep E/N_0 and D the same for both systems, and then examine P_e for each of the systems; the system with smallest P_e gives a better performance.

There are at least two ways to lower the probability of error of a system and simultaneously suffer deletions. This error to deletion tradeoff can be accomplished by changing from a strictly error correcting code, (let us say it corrects up to and including e_1 errors), to a combination error correction and detection code (say it corrects up to and including e_2 errors and detects up to and including d errors). In general for the same redundancy the sum of the number of the errors corrected and the errors detected for system 2 is larger than the number of errors corrected in system 1, or

$$e_2 < e_1 < e_2 + d_1 \ .$$

*$C_2 < C_1$ for all cases where the error is undesirable and correct detection is desirable, i.e., the error cost is larger than the correct detection cost.

The above equation indicates that although system 2 will correct less errors than system 1, system 2 is able to reject data if errors are detected and, since in general $e_2 + d > e_1$, the number of undetected and uncorrected errors is less than that of system 1. Consequently the error rate in system 2 is smaller than that of system 1. The second method is a specific case of the more general case discussed in Reference 5. This technique is much simpler than error detection and correction. Under this technique the largest of the correlator outputs is detected. Up to this point the procedure is identical to the Maximum Likelihood Decision Scheme. The value of the ratio

$$B_1 = \frac{x_{(N)} - \mu}{\sigma}^{*}$$

is measured and the data is rejected if this value is smaller than a predetermined threshold B_t. If the value of the Statistic B_1 is larger than B_t the decision is that $x_{(N)}$ is the useful signal.

2. PERFORMANCE OF THE GOLAY (23, 12) CODE WITH ERROR CONTROL

This code was chosen because it is a perfect code and will correct all three or less errors.

The word error probability is given by

$$P_w = \sum_{i=4}^{n} \binom{n}{i} P^i (1-P)^{n-i} \tag{3}$$

where P is the bit error probability and is given by

$$P = \int_{\sqrt{(1-\rho) E/N_0}}^{\infty} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt . \tag{4}$$

E/N_0 is the Energy per bit to Noise power per unit bandwidth. For antipodal signals $1 - \rho = 2$; if E/N_0 is the energy per information bit to N_0 then the energy per bit to N_0 is k/n (E/N_0) because there are k information and n bits per word. Thus for the Golay (23, 12) Code

$$P = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{24E/23N_0}}^{\infty} e^{-t^2/2} dt \tag{5}$$

*$x_{(N)}$ is the largest output, μ is the sample mean and σ is the sample standard deviation.

(3) and (5) were used to plot the performance curve of this code (shown as the dotted line) in Figure 2. Equation 4 is a special case of the generalized equation.

$$P = \frac{1}{2\pi}\int_{B_t}^{\infty} e^{-x^2/2}\,dx \int_{-\infty}^{x} e^{-(s-A)^2/2}\,ds \tag{6}$$

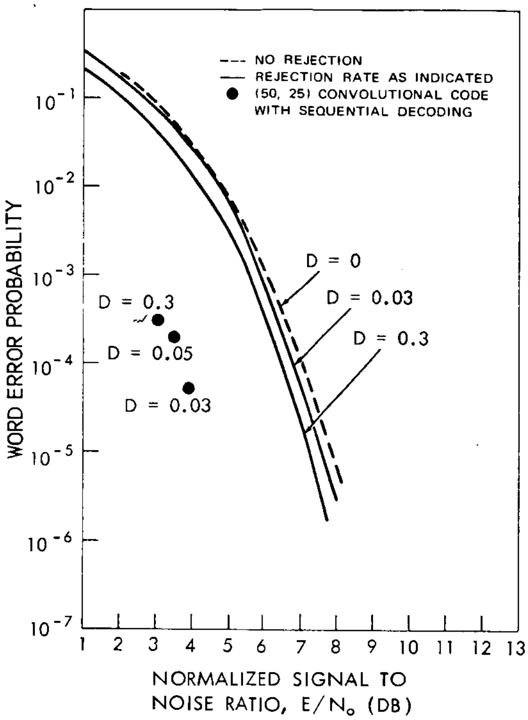

Figure 2—Performance of the Golay (23, 12, 3) code.

with $B_t = -\infty$ and $A = \sqrt{2E(1-\rho)/N_0}$; when $B_t = -\infty$, $y = s - A$

$$P = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-x^2/2}\,dx \int_{-\infty}^{x-A} e^{-y^2/2}\,dy$$

$$= \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(-\frac{A}{2}\right) = \frac{1}{2} - \frac{1}{2}\,\mathrm{erf}\left(\frac{A}{2}\right)$$

but

$$\mathrm{erf}\left(\frac{A}{2}\right) = \frac{2}{\sqrt{\pi}}\int_0^{A/2} e^{-t^2}\,dt$$

now if

$$t = x/\sqrt{2}$$

$$\mathrm{erf}\left(\frac{A}{2}\right) = \sqrt{\frac{2}{\pi}}\int_0^{A/\sqrt{2}} e^{-x^2/2}\,dx$$

then

$$P = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{2}{\pi}}\int_0^{A/\sqrt{2}} e^{-x^2/2}\,dx + \left[\frac{1}{2} - \frac{1}{2}\right]$$

$$= 1 - \left[\frac{1}{2} + \frac{1}{\sqrt{2\pi}}\int_0^{A/\sqrt{2}} e^{-x^2/2}\,dx\right]$$

$$= 1 - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{A/\sqrt{2}} e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}}\int_{A/\sqrt{2}}^{\infty} e^{-x^2/2}\,dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{\sqrt{E(1-\rho)/N_0}}^{\infty} e^{-x^2/2}\,dx\ .$$

This is clearly the same as Equation (2). Equation (4) gives the probability of error under a Maximum Likelihood Decision provided the largest signal exceeds a given threshold B_t. If it does not exceed B_t no decision is made (the data is erased). Under this decision scheme one can reduce the probability of error to any desired value at the expense of higher rejection rates.

The probability of correct detection for binary antipodal signals is given by

$$Q = \frac{1}{2\pi} \int_{B_t}^{\infty} e^{-(s-A)^2/2} \, ds \int_{-\infty}^{s} e^{-x^2/2} \, dx \tag{7}$$

with $A = \sqrt{4E/N_0}$. P_e is given by (4) and the erasure rate E_b is

$$E_b = 1 - P_c - P_e . \tag{8}$$

This detection scheme can be added to coding in the following manner.

The n coded bits are detected according to this scheme with P_e, P_c, and E as given by Equations (6), (7), and (8).

If the number of erasures in the word of n bits is larger than r we reject the data.

Then the word deletion rate (from now on referred to as deletion rate) is given by

$$\begin{aligned} D(E_b, r) &= \sum_{j=r+1}^{n} \binom{n}{j} E_b^{\,j} (1 - E_b)^{n-j} \\ &= 1 - \sum_{j=0}^{r} \binom{n}{j} E_b^{\,j} (1 - E_b)^{n-j} \end{aligned} \tag{9}$$

under this scheme the word error probability for an (n, k, e) perfect code* with the erasures assumed as errors is

$$P_w' = \sum_{i=0}^{r} \sum_{j=e+1-i}^{n-i} \binom{n}{i} \binom{n-i}{j} E_b^{\,i} P^j Q^{n-i-j} \tag{10}$$

where $P = P_e$ and $Q = P_c = 1 - P$. Equation (10) is an upperbound on the word error probability because under this scheme all erasures are assumed to be errors. To find the exact word error probability let us establish a decision rule as follows. If the number of bits that do not satisfy the

*(n, k, e) → n bits per word; k information bits; corrects all e or less errors.

threshold conditions* is larger than a specified number r, the word is deleted. If it is less than r then a Maximum Likelihood decision is made on all n bits. Under this decision rule the deletion rate is still given by Equation (9).

The expression for the word error probability is found in this manner. There are two types of bits in the n-bit word. Type-one bits satisfy the threshold requirements and have P and Q given by

$$P = \text{prob}\left(x_{(N)} > B_t;\ \theta\epsilon\omega\right)^{\dagger} \tag{11}$$

$$Q = \text{prob}\left(x_{(N)} > B_t;\ \theta\epsilon\Omega - \omega\right) \tag{12}$$

Type-two bits do not satisfy the threshold requirements. The probability of error for those bits is

$$\begin{aligned} E_p &= \text{prob}\left(x_{(N)} \leq B_t;\ \theta\epsilon\omega\right) \\ &= \text{prob}\left(\theta\epsilon\omega\ \forall x_{(N)}\right) - \text{prob}\left(x_{(N)} > B_t;\ \theta\epsilon\omega\right) \\ &= P_{MLDS} - P^{\ddagger}. \end{aligned} \tag{13}$$

Similarly the probability of correct decision for the type-two bits is

$$E_q = Q_{MLDS} - Q \tag{14}$$

and the rejection rate is

$$E_b = E_p + E_q\ ; \tag{15}$$

using Equation 11-15 and

$$Q_{MLDS} = 1 - P_{MLDS}$$

then

$$E_b = E_p + E_q = P_{MLDS} + Q_{MLDS} - P - Q = 1 - P - Q$$

*Those bits were erasures in the previous scheme.
†ω is the noise parameter space and $\Omega - \omega$ is the signal parameter space.
‡MLDS stands for Maximum Likelihood Decision System.

an obvious result. The word error probability for ℓ type-two correct bits, i type-two incorrect bits, j incorrect type-one bits, and $n-\ell-i-j$ correct type-one bits in an e error correcting n-bit code is given by

$$P_w = \sum_{\ell=0}^{r} \sum_{i=0}^{r-\ell} \sum_{j=e+1-i}^{n-\ell-i} \binom{n}{\ell} \binom{n-\ell}{i} \binom{n-\ell-i}{j} E_q^{\ell} E_p^{i} P^{j} Q^{n-\ell-i-j} \tag{16}$$

with $\ell + i \leq r$. Since from Equation (9) the deletion rate is a function of E_b and r, different combinations of E_b and ℓ will result in the same D; the above results are shown in Figure 3.

For deletion rates .03 and .3, E_b and the corresponding r were read from Figure 3 as shown in Tables 1 and 2.

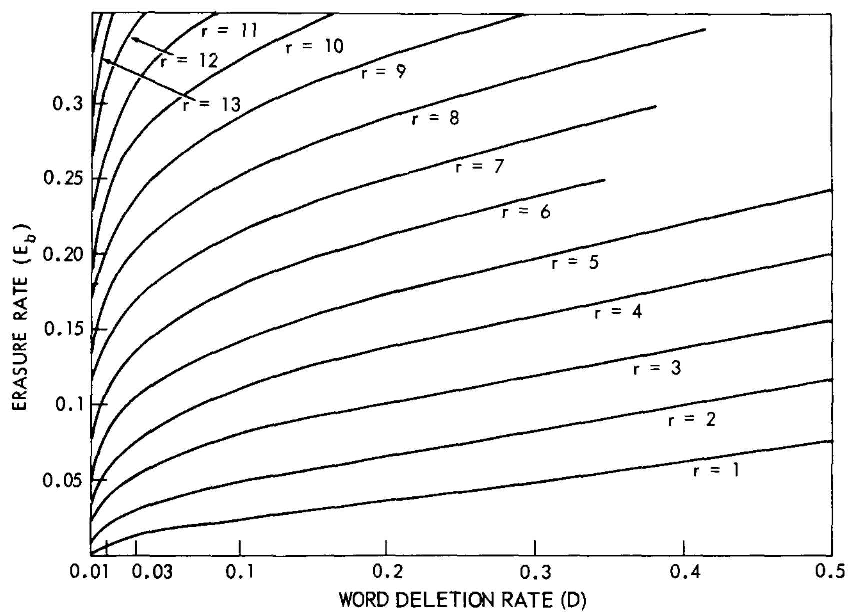

Figure 3—Erasure rate vs word deletion rate for n = 23.

Next $P(B_t)$ and $Q(B_t)$ were found for all B_t for binary antipodal signals by using Equations (6) and (7), $A = \sqrt{4(k/m)(E/N_0)}$; the P and Q corresponding to the required erasure rate E_b were found as shown on Tables 1 through 5 for the signal to noise ratios of interest; the results were plotted in Figure 2 together with the usual performance of the Golay Code (zero erasure rate). An examination of Figure 2 indicates that the performance of the convolutional decoding is superior than the above scheme of decoding.

Table 1

D = .3 E/N_0 = 3 db		Q_{MLDS} = .9289			
r	E	B_t	P	Q	P_w
1	.048	.53	.0608	.8912	.0451
2	.083	.79	.0547	.8623	.04423
3	.12	.97	.0483	.832	.04310
4	.159	1.13	.043	.798	.04227
5	.198	1.27	.038	.764	.04208
6	.24	1.40	.0325	.7275	.04103
7	.278	1.51	.0298	.6932	.04258
8	.32	1.62	.0248	.6552	.04304
9	.362	1.73	.0212	.6168	.04393
D = .03 E/N_0 = 3.5 db		Q_{MLDS} = .9289			
3	.053	.58	.059711	.887289	.072
5	.1055	.90	.050658	.843842	.0715
7	.170	1.17	.04133	.78867	.0727

Table 2

D = .05 E/N_0 = 3.5 db		Q_{MLDS} = .93962			
r	E	B_t	P	Q	P_w
1	.018	.26	.05637	.92563	.04184
3	.064	.76	.048	.888	.04153
5	.12	1.08	.0397	.8403	.04177
7	.187	1.31	.0316	.7814	.04245
9	.259	1.56	.0244	.7166	.04453
11	.337	1.77	.01818	.64782	.04745

Table 3

D = .03 E/N_0 = 4 db		Q_{MLDS} = .94962			
r	E	B_t	P	Q	P_w
1	.0135	.25	.0475	.9390	.0245
3	.053	.78	.04062	.90638	.02441
5	.1055	1.10	.0334	.8611	.02444
7	.170	1.39	.0263	.8037	.0248
9	.239	1.62	.0203	.7407	.0262
11	.315	1.84	.0150	.670	.0282
D = .3 E/N_0 = 4 db		Q_{MLDS} = .94962			
1	.048	.74	.041423	.910577	.0143
3	.012	1.18	.031693	.848307	.0134
5	.198	1.49	.023732	.778268	.0130
7	.278	1.73	.0174149	.704585	.0134

Table 4

D = .03 E/N_0 = 6 db Q_{MLDS} = .980256					
r	E	B_t	P	Q	P_w
3	.053	1.31	.01377	.93323	.000848
5	.1055	1.66	.01026	.88424	.000854
7	.170	1.94	.00729	.82271	.000903
D = .3 E/N_0 = 6 db Q_{MLDS} = .980256					
3	.12	.172	.0094945	.8705055	.0003462
5	.198	.205	.006299	.795701	.0003458
7	.278	.230	.00412	.71788	.0003897

Table 5

D = .03 E/N_0 = 8 db Q_{MLDS} = .995087					
r	E	B_t	P	Q	P_w
3	.053	2.02	.002494	.944506	.0000032
5	.1055	2.38	.001566	.892934	.0000035
7	.170	2.67	.000951	.829049	.0000042
D = .3 E/N_0 = 8 db Q_{MLDS} = .995087					
3	.12	2.43	.0014555	.8785445	$.83 \times 10^{-6}$
5	.198	2.78	.000772	.801228	$.99 \times 10^{-6}$
7	.278	3.04	.000432	.721568	1.44×10^{-6}

3. ORTHOGONAL CODES WITH DATA REJECTION

The performance of orthogonal signals when rejection is allowed has already been analyzed in Reference 5 for both the coherent and non-coherent cases, and for six different statistics. The simplest* statistic is

$$B_1 = \frac{x_{(N)} - \mu}{\sigma}$$

and for $\mu = 0$ and $\sigma = 1$ it becomes the familiar largest of the signals. The decision scheme is this. If the largest signal $x_{(N)}$, is larger than B_t the signal is accepted as the useful signal. If $x_{(N)}$ is smaller than B_t the data is rejected. Under this scheme the probability of correct decision is

$$P_c = \int_{B_t}^{\infty} f_s(s)\, ds \left[\int_{-\infty}^{s} f(x)\, dx \right]^{N-1} \tag{14}$$

*This statistic makes most of the assumptions as to the knowledge of μ and σ.

where $f_s(z)$ is the pdf of the useful signal and $f(z)$ is the pdf of the noise signals. The probability of error is

$$P_e = (N-1)\int_{B_t}^{\infty} f(x)\left[\int_{-\infty}^{x} f(t)\,dt\right]^{N-2} dx \int_{-\infty}^{x} f_s(s)\,ds \tag{15}$$

and the rejection rate is

$$D = 1 - P_c - P_e\,. \tag{16}$$

It should be noted that with $B_t = -\infty$

$$P_c = 1 - P_e$$

and

$$D = 0\,.$$

This last condition on B_t results in the Maximum Likelihood Decision Scheme. The results of the performance of the orthogonal signals are shown in Figures 4 to 9.

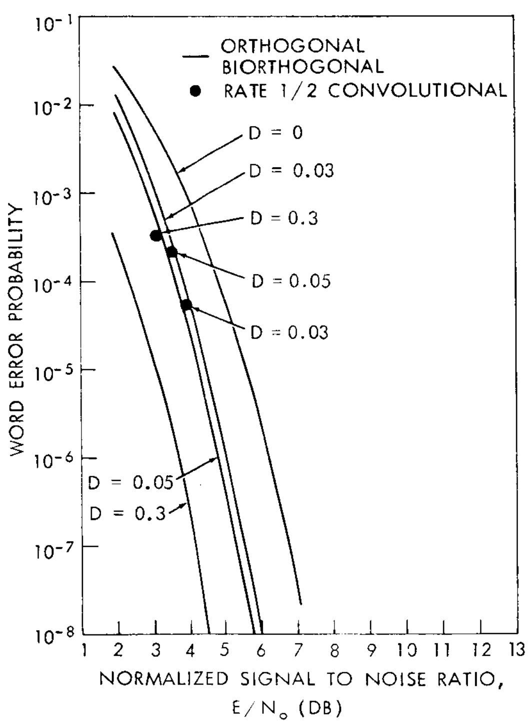

Figure 4—Performance of biorthogonal codes for n = 8 and deletion rates of 0, .03, .1, .3 (For this large n the performance of orthogonal and biorthogonal codes is indistinguishable for all practical purposes).

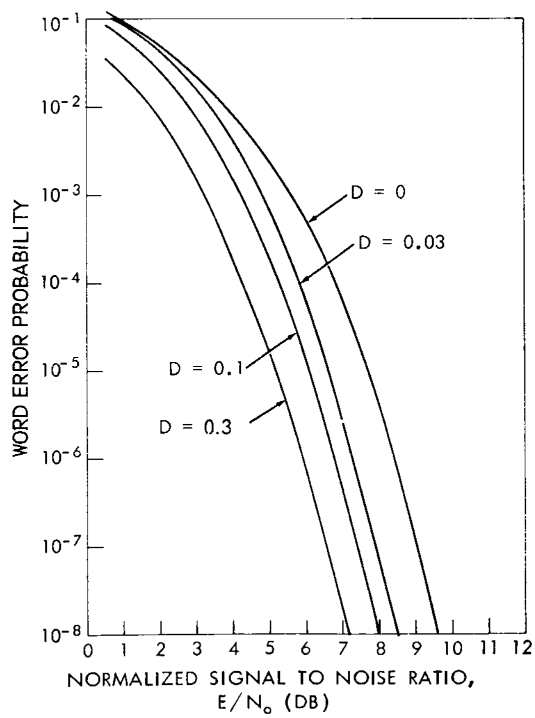

Figure 5—Performance of orthogonal and biorthogonal n = 4 and D = 0, .03, .1, .3.

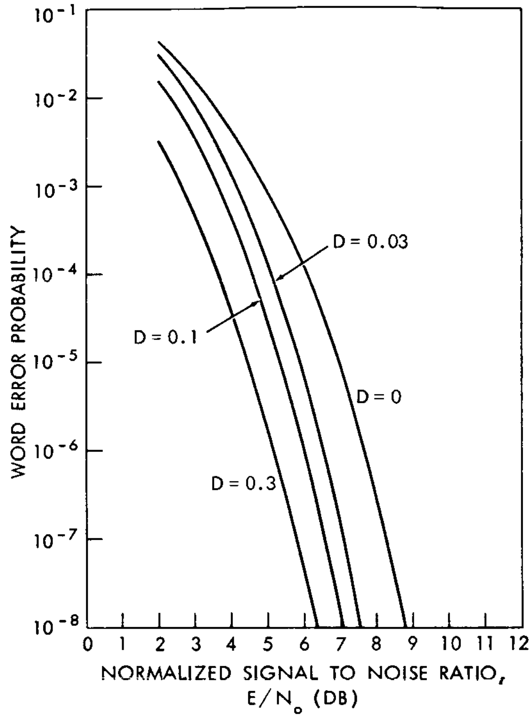

Figure 6—Performance of orthogonal and biorthogonal n = 5 and D = 0, .03, .1, .3.

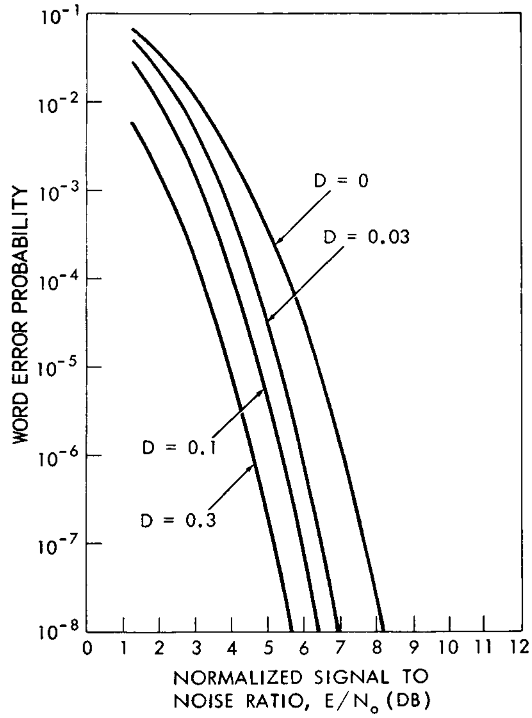

Figure 7—Performance of orthogonal and biorthogonal n = 6 and D = 0, .03, .1, .3.

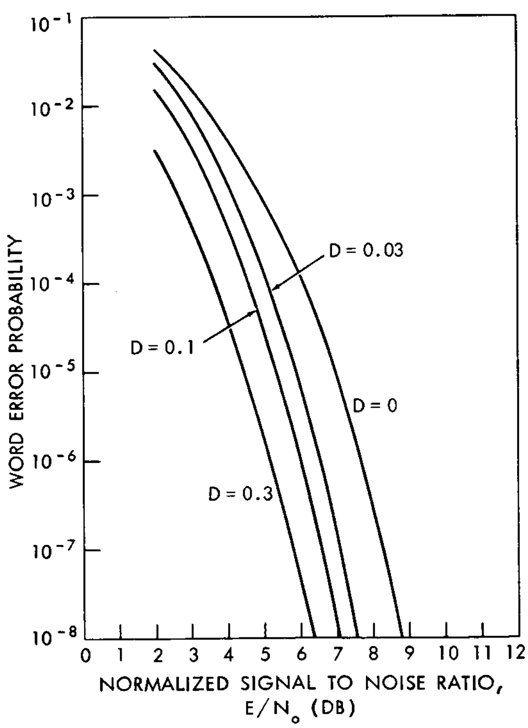

Figure 8—Performance of orthogonal and biorthogonal n = 7 and D = 0, .03, .1, .3.

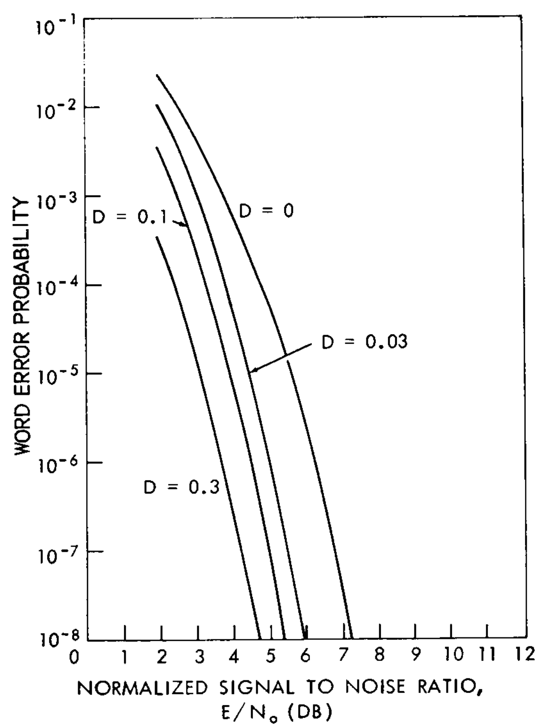

Figure 9—Performance of orthogonal n = 8 and D = 0, .03, .1, .3.

4. BIORTHOGONAL CODES WITH DATA REJECTION

Application of the procedure of the previous paragraph to biorthogonal signals* results in

$$P_c = \int_{B_t}^{\infty} f_s(s)\,ds \left[\int_{-s}^{s} f(x)\,dx\right]^{(N/2)-1}$$

similarly

$$P_e = \alpha(1-\alpha)^{\dagger}$$

where

$$\alpha = \left(\frac{N}{2} - 1\right) \int_{B_t}^{\infty} f(x) \left[\int_{-\infty}^{x} f(t)\,dt\right]^{(N/2)-2} \int_{-x}^{x} f_s(s)\,ds$$

and

$$D = 1 - P_c - P_e \,.$$

The results of this type of detection have been plotted in Figures 4 to 9.

5. CONCLUSIONS

Figure 2 shows how the performance of the Golay (23, 12, 3) code (dotted line) can be improved by allowing deletions for 3% or 30% of the data. This method of improving error rate at the expense of data rejection for algebraic codes is not unique. Consequently the amount of improvement shown is not necessarily optimum.

In the process of obtaining the curves of Figure 2, r and E_b were allowed to vary in each case to values that produced a constant deletion rate, D. This was done in order to examine the variation of the word error probability as a function of r and E_b. Observation of the Tables indicates that the variation of P_w is insignificant, although it reaches some minimum which corresponds to higher r's as the signal to noise ratio increases. Since higher values of r correspond to higher threshold values, the above result is reasonable.

*See appendix.

†There is a extra term which has been ignored because it is very small for the cases considered here (see Appendix).

Examination of Figure 4 shows that

(a) The performance of orthogonal and biorthogonal codes is indistinguishable for all curves ($D = 0, .03, .05, .3$).

(b) At 3 db the orthogonal and biorthogonal codes are superior to the convolutional codes for $D = .3$.

(c) At 3.5 db the orthogonal and biorthogonal codes are slightly better than the convolutional codes for $D = .05$.

(d) At 4 db, $D = .03$ the performance of the two different systems is about the same.

(e) The improvement in performance as D increases is much more pronounced in orthogonal and biorthogonal codes than in algebraic codes.

Figures 5 to 9 show the performance of orthogonal and biorthogonal codes for some other n's. It should be noticed that for relatively larger n orthogonal and biorthogonal codes perform equally well.

REFERENCES

1. Lumb, D. R., A Study of Codes for Deep Space Telemetry, IEEE International Convention and Exhibition, March 1967.

2. Massey, James L., Threshold Decoding. MIT Press, Cambridge, Mass., 1963.

3. Wozencraft, John M., and Jacobs, Irwin M., Principles of Communication Engineering. John Wiley and Sons, Inc., N. Y., 1965.

4. Wozencraft, J. M., and Reiffen, B., Sequential Decoding (Technology Press of Massachusetts Institute of Technology, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1961).

5. Morakis, J. C., "Statistical Tests for Signal Detection," NASA TR G802, 1967.

APPENDIX A

1. In the case of N biorthogonal signals the number of correlators is $N/2$ * and the probability of correct decision is the probability that the useful signal is larger than the absolute value of any of the $(N/2)-1$ remaining noise signals and larger than B_t, given that the sign of the useful signal has been correctly determined; thus

$$P_c = \text{prob}\left(|s| > |x|,\ \forall x;\ |s| > B_t\right)$$

or equivalently

$$P_c = \text{prob}\left(|s| > |x_{(L)}|;\ |s| > |x_{(1)}|;\ |s| > B_t\right) . \tag{A-1}$$

This statement is equivalent to the probability that the largest and smallest noise signals are within the range $(-s, s)$; for Gaussian signals with zero mean

$$\text{prob}\left(|s| > |x_{(L)}|\right) = \text{prob}\left(|s| \geq |x_{(1)}|\right) . \tag{A-2}$$

Thus, independently of the relative frequency of positive and negative useful signals.

$$\begin{aligned} P_c &= \text{prob}\left(|s| > |x_{(L)}|;\ |s| > B_t\right) \\ &= \int_{B_t}^{\infty} f_s(s)\,ds \left[\int_{-s}^{s} f(x)\,dx\right]^{(N/2)-1} \end{aligned} \tag{A-3}$$

$$B_t \geq 0 .$$

2. An error occurs whenever the absolute value of either the smallest or the largest noise signal is larger than the absolute value of the useful signal, or the useful signal is mistaken for its negative. If the above mentioned events are named events A, B and C respectively, then,

$$\begin{aligned} P_e = P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) \\ &\quad - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C) . \end{aligned} \tag{A-4}$$

*$N = 2^n$ is the number of correlators necessary for N orthogonal signals.

Since A, B, C are independent

$$P_e = P(A) + P(B) + P(C) - P(A)P(B) - P(A)P(C) - P(B)P(C)$$

$$+ P(A)P(B)P(C) \ . \qquad \text{(A-5)}$$

Now P(A), P(B), and P(C) are given by

$$P(A) = P\left(|x_{(1)}| > |s|;\ |x_{(1)}| > B_t\right) \qquad \text{(A-6)}$$

$$P(B) = P\left(|x_{(N)}| > |s|;\ |x_{(N)}| > B_t\right) \qquad \text{(A-7)}$$

$$P(C) = P\left(|-s| > x,\ \forall x;\ |-s| > B_t\right) \ .$$

P(B) is given by

$$P(B) = \int_{B_t}^{\infty} \left(\frac{N}{2} - 1\right) f(x) \left[\int_{-\infty}^{x} f(t)\, dt\right]^{(N/2)-2} \int_{-x}^{x} f_s(s)\, ds$$

$$B_t \geq 0 \ .$$

Again due to symmetries in $f\left(x_{(1)}\right)$ and $f\left(x_{(N)}\right)$, and symmetry of $f_s(s)$ for the word and its complement,

$$P(A) = P(B) = \alpha \ .$$

Finally

$$P(C) = \int_{B_t}^{\infty} f_s(-s)\, ds \left[\int_{-s}^{s} f(x)\, dx\right]^{(N/2)-1}$$

$$f_s(-s) = \frac{1}{\sqrt{2\pi}} e^{-(s+A)^2/2}$$

and P(C) has a negligible value for values of $s > 0$ ($B_t \geq 0$) and can be neglected for $A \geq 3$ and $n \geq 16$ in this paper. This can be seen in Figure A1 where $f_s(-s)$ and

$$F_L\left(|s|\right) = \left[\int_{-s}^{s} f(t)\, dt\right]^{(N/2)-1}$$

have been plotted.

$$F_L\left(|s|\right) = \left(2F(s) - 1\right)^{(N/2)-1}$$

where

$$F(s) = \int_{-\infty}^{s} f(t)\,dt \;.$$

Substituting these results into (A-5) with $P(C) \approx 0$

$$\begin{aligned} P_e &= P(A) + P(B) - P(A)\,P(B) \\ &= \alpha(2-\alpha) \;. \end{aligned}$$

LIST OF SYMBOLS

n	the word length
N	the number of words
k	number of information bits per word
e	the maximum bit correction capability of a code
E_b	the bit erasure rate
D	the word deletion rate
P	the bit error probability
Q	the probability of correct detection (on a bit basis)
E_p	the probability that an erasure bit is in error
E_q	the probability that an erasure bit is correct ($E_p + E_q = E_b$)
P_{MLDS}	the error rate of a Maximum Likelihood Decision System
Q_{MLDS}	$= 1 - P_{MLDS}$; the probability of correct detection for MLDS
P_c	probability of correct detection (on a word basis)
P_w, P_e	word error rate
E	the signal energy per bit of information
N_0	noise power per unit bandwidth
ρ	correlation coefficient
$\cup$	the union of events
$\cap$	the intersection of events
$\forall$	for all
$x_{(L)}$	the largest of the noise signals.

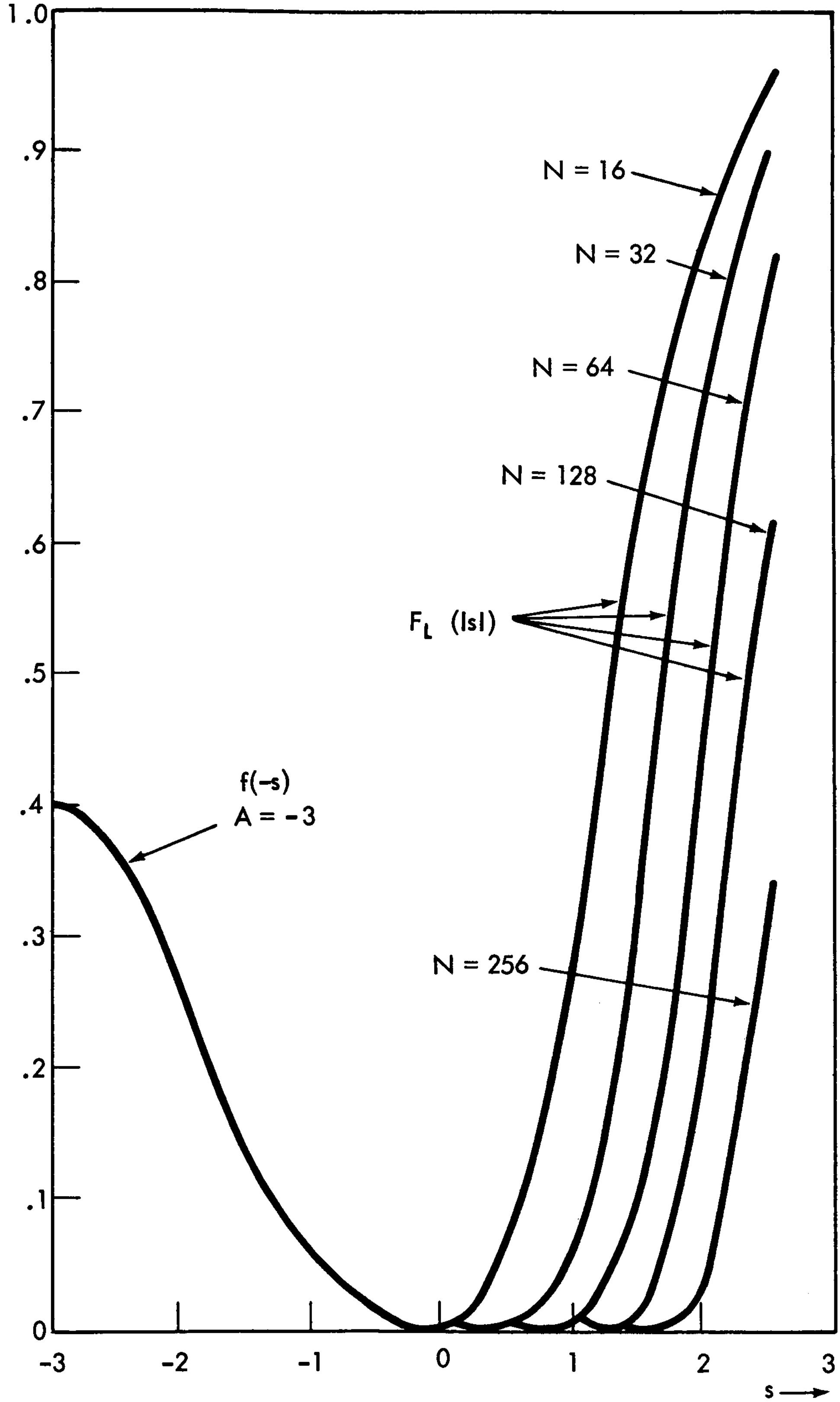

Figure A1