On the Use and Performance of Error-Voiding and Error-Marking Codes

By E. O. ELLIOTT

(Manuscript received May 25, 1966)

In contrast to payroll or inventory data, which must reach the recipient in its entirety, there is another class of data that includes radar-tracking data, remote-sensory data or control data, etc., for which the requirement of completeness is not so stringent. Error control for this class of data may be accomplished by forward-acting error-correcting codes which void or mark any detected errors that they do not correct. In order to evaluate these error-voiding methods, the error rates for such codes are estimated in this paper using the error statistics of the Alexander-Gryb-Nast study.

A class of 18 (about 50 percent redundant) cyclic codes capable of correcting from one to five errors and having block lengths from 15 to 47 bits is examined. Only bounded-distance decoding is evaluated, but each code is assigned each possible decoding radius up to the maximum permissible radius determined by the capability of the code. Since interleaving generally reduces error rates, the error rates for this class of codes are estimated for interleaving constants from 50 to 300 in steps of 50.

It is concluded that:

(i) If voids are permissible (at a rate of about 10^{-4}) then low undetected-error rates may be achieved by a code capable of correcting many errors but used to correct only two or three errors. Such a code might be about 50 percent redundant and have a block length between 25 and 50 bits.

(ii) It is impractical to obtain low void rates. If voids are not tolerable, then retransmission is required to obtain low error rates.

(iii) Interleaving is more effective with codes correcting three (or more) errors than with those correcting only single or double errors.

I. INTRODUCTION

In contrast to payroll or inventory data, which must reach the recipient in its entirety, there is another class of data that includes radar-tracking data, remote-sensory data or control data, etc., for which the requirement of completeness is not so stringent. The distinction between these two classes of data is fundamental in the classification of customer requirements in data transmission and the selection of appropriate error-control methods.

If the complete data message is required at the receiving station then error control must either be carried out by error detection and retransmission or by forward-acting error correction. Of these two methods the former is the more economical to achieve low error rates. However, if completeness of the transmitted message is not essential and receipt of say 99.9 percent would be satisfactory, then very inexpensive error-voiding techniques may be employed to achieve the desired low error rates. With these techniques an error-detecting code is used to detect and then void (or mark) all detectable errors. If lower void rates are desired some error correction may be introduced and the remaining error-detection capability of the code used to void or mark errors. In order to evaluate these error-voiding methods, the error rates for such codes are estimated in this paper using the error statistics of the Alexander-Gryb-Nast study.[1]

Data for which completeness is an important requirement would include payrolls, inventories, orders, sales and banking records, etc. Since accuracy is a very important factor for this type of data, an automatic retransmission error-control system would probably be required. However, one can imagine cases in which manual retransmission would suffice. Errors could be marked or voided by the error-detecting code and when errors are so indicated in a message the recipient could initiate steps to obtain the missing data. This might be tolerable in some situations if only a small fraction of the messages required this special handling.

At the other extreme are messages which need not be received in their entirety to be effective. Among these we might list radar tracking data, remote sensory data, and remote control data. Again detected errors would be marked or voided. In some cases, the discarded data might be restored by some extrapolation or interpolation with neighboring blocks of the presumed error-free data. In other cases, it might suffice to operate with just the nonvoided blocks of data. For these procedures to work it is of course necessary that the void rate be low enough. The void rate itself however is not the sole factor determining the feasibility of the system. The time distribution of voids may also be very important. For example, with radar-tracking data a void rate of 10^{-4} (words/word) might be tolerable in itself but if it were realized on a channel on which the voids tended to occur in runs or bunches it might not be tolerable since the tracking system cannot operate if too long a stretch of data is missing. This paper is concerned only with void rates and does not treat the time distribution of voids.

The splitting of a code's function to achieve both error correction and detection is accomplished by noting the distance of the received word from the nearest code word. If this distance is less than or equal to a given number which is called the employed correction radius then the received word is decoded as that nearest code word, otherwise it is not decoded and a detected error is announced. This is the method of bounded-distance decoding. Although there are other methods of combining error correction and detection, this one is considered here because several practical decoding algorithms conform to it.

The codes are also evaluated over a range of degrees of interleaving, because, if a given code is interleaved on a burst channel its performance improves. Interleaving may be thought of as though it were accomplished by reading the encoded data into a rectangular array row by row and then sending it on line column by column. The length of a row is the block length of the code. The number of rows is the interleaving constant t; two adjacent bits of an originally encoded block would be sent on line with $t - 1$ other bits between them.

This memorandum examines the effect that block length, redundancy, correction radius, and interleaving have on undetected error rates and void rates over the switched telephone network. Specifically, a class of 18 cyclic codes with block lengths ranging from 15 to 47 bits is examined. Among these are codes capable of correcting from one to five errors. Although a variety of redundancies is represented, codes with about 50 percent redundancy predominate. Using data from the Alexander-Gryb-Nast study, the error rates for this class of codes are estimated for each permissible correction radius with no interleaving and with interleaving with constants from 50 to 300 in steps of 50. A number of practical means for implementing many of these error-control systems are available. For this reason, the present investigation aims at giving a useful qualitative insight into the role of bounded-distance decoding for error control on the switched telephone network.

For a more complete understanding of error control, it would be necessary to consider alternative methods of decoding such as burst decoding, threshold decoding, etc. Also error statistics for other channels and modes of communication should be considered. This awaits further development of analytical techniques and the availability of additional error data.

II. ESTIMATING ERROR RATES FOR BOUNDED-DISTANCE DECODING

A group code is commonly specified by the pair (n,k) where n is the block length of the code and k is the number of information bits in a code word. When bounded-distance decoding is employed, the correction radius a is added to this pair and the code is specified by the triplet (n,k,a). Of course, a is less than or equal to the maximum error-correcting capability e of the code. Thus, if x is the transmitted word and y is the received word then if y is at distance less than or equal to a from some code word z, y is decoded as z. If $z \neq x$ then an undetected error results, and if y is not within distance a from any code word a detected error results. To estimate the probabilities P_E and P_D of these two events (undetected error and detected error) the method of Ref. 2 is employed as it was in Ref. 3 to study permutation decoding which is a special case of bounded distance decoding.

Because of the perfect distance symmetries between the words of a group code, the probabilities P_E and P_D do not depend on which code word is transmitted. Therefore, to calculate P_E and P_D and simplify matters we assume the all zero word θ is transmitted. Let $C_a(m)$ be the total number of words of weight m which are at a distance less than or equal to a from some code word. As in Ref. 2 let $P(m,n)$ be the total probability that m errors occur in a transmitted block of n bits so that $P(m,n)/\binom{n}{m}$ is an approximation to the probability that a particular pattern of m errors occur. Then, assuming θ is transmitted we see that, as an approximation,

$$P_E \cong \sum_{m=a+1}^{n} C_a(m) \frac{P(m,n)}{\binom{n}{m}}. \tag{1}$$

Similarly, if $D_a(m)$ is the total number of words of weight m at a distance greater than a from any code word then

$$P_D \cong \sum_{m=a+1}^{n} D_a(m) \frac{P(m,n)}{\binom{n}{m}}. \tag{2}$$

Clearly $D_a(m) = \binom{n}{m} - C_a(m)$ and now the problem is to obtain $C_a(m)$.

In Ref. 3 a formula is given for $C(m,j)$, the number of words of weight m which are at a distance j ($j \leqq e$) from some code word, and since

$$C_a(m) = \textstyle\sum_{j=0}^{a} C(m,j)$$

the desired numbers are thus obtainable. Unfortunately, the formula in Ref. 3 involves a summation with terms of alternating sign which requires triple precision programming to obtain satisfactory accuracy on the computer. As a consequence of this requirement, the following alternate and more direct formula for $C(m,j)$ was derived. With it, double precision FORTRAN programming suffices.

$$C(m,j) = \sum_{i=0}^{j} w(m + 2i - j) \binom{m + 2i - j}{i} \binom{n - m - 2i + j}{j - i} \tag{3}$$

where $w(r)$ represents the number of code words of weight r. It is obtained from simple combinatorial considerations as follows.

Suppose x is a code word of weight r and y is a word of weight m and let i be the number of bit positions in which x is 1 and y is 0, and l be the number of bit positions in which x is 0 and y is 1. Let $j = i + l$ so j is the distance between x and y. Then $m = (r - i) + l = r + j - 2i$ and $i = (r + j - m)/2$. The total number of possible y's of weight m is then given by

$$\binom{r}{i} \binom{n - r}{j - i}.$$

Considering that each code word of weight $m + 2i - j$ $(= r)$ therefore has

$$\binom{m + 2i - j}{i} \binom{n - m - 2i + j}{j - i}$$

distinct words of weight m at distance j $(j \leqq e)$ from it and clearly $m - j \leqq r \leqq m + j$, i.e., $0 \leqq i \leqq j$, (3) then follows.

III. ERROR RATES FOR A SAMPLE COLLECTION OF CODES

Cyclic codes or shortened cyclic codes have received a great deal of attention in the field of error control because of the ease in their implementation. For this reason a collection of 18 cyclic codes for which the spectral functions $w(r)$ were readily available was chosen for this study. Most of these codes and spectra are given in Ref. 3. Those which were not in Ref. 3 were included so that a wider range of redundancies would be represented. The codes are listed in Table I which gives their block length n, dimension k, minimum distance d, and maximum error-correcting capability e.

Using (1) and (2) and the $P(m,n)$ values from the Alexander-Gryb-Nast study, the probabilities P_E and P_D were estimated for each of these codes. Samples of the results are shown in Figs. 1, 2, and 3.

TABLE I—LIST OF CYCLIC CODES

n	k	d	e
15	11	3	1
15	10	4	1
15	7	5	2
15	6	6	2
15	5	7	3
15	4	8	3
15	2	10	4
17	9	5	2
17	8	6	2
21	12	5	2
21	11	6	2
23	12	7	3
23	11	8	3
31	21	5	2
31	20	6	2
31	16	7	3
31	15	8	3
47	24	11	5

Fig. 1 shows the effect which the correction radius has on error rates for two of the codes. The undetected error rate is noted to decrease about one order of magnitude for each unit decrease in correction radius. Also, the detected error rate, which is about 10^{-4}, is rather insensitive to the correction radius.

In Fig. 2 the undetected error rate is plotted as a function of the efficiency of a 15-bit code. Each order of magnitude improvement in the error rate requires an increase of three redundant bits (which is 20 percent of the block length). An examination of the error rates for the codes with block length 31 (not shown here) reveals that the same change of three redundant bits is required to achieve an order of magnitude improvement with this longer block length code.

Error rates of double-error-correcting codes of about 50 percent redundancy are presented in Fig. 3 as a function of block length. Again the detected error rate is not a very sensitive parameter while the undetected error rate ranges over many orders of magnitude. Quite acceptable error rates are attainable with block lengths not much greater than 25 bits.

IV. THE EFFECT OF INTERLEAVING ON ERROR RATES

Through interleaving (with constant t), the bits of a code word are separated when transmitted on line so each pair of adjacent bits have

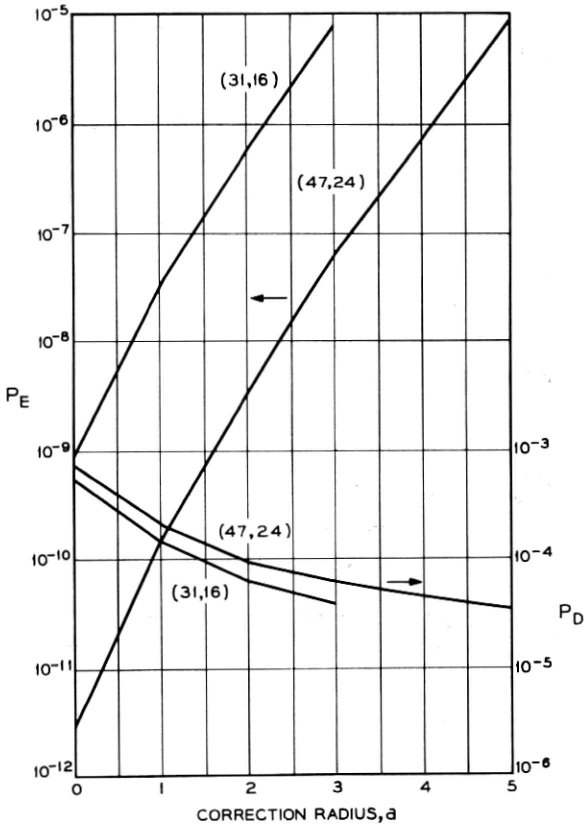

Fig. 1 — Error rates vs correction radius (no interleaving).

$t - 1$ bits from other code words between them. The effect of this is to decrease the error vulnerability dependence between the bits of a code word so that the interleaved channel is less of a burst channel and more like a memoryless channel. To examine the effect interleaving has on error control with random error-correcting codes, the $P_t(m,n)$ probabilities were approximated for interleaving constants $t = 50$, 100, 150, 200, 250, and 300, and the error rates for the 18 codes were estimated as in the previous section. To approximate the $P_t(m,n)$ values first the error autocorrelation function $a_t(k)$ of the interleaved channels is obtained from a smoothed version of the autocorrelation function $a(s)$ of the Alexander-Gryb-Nast data through the relation

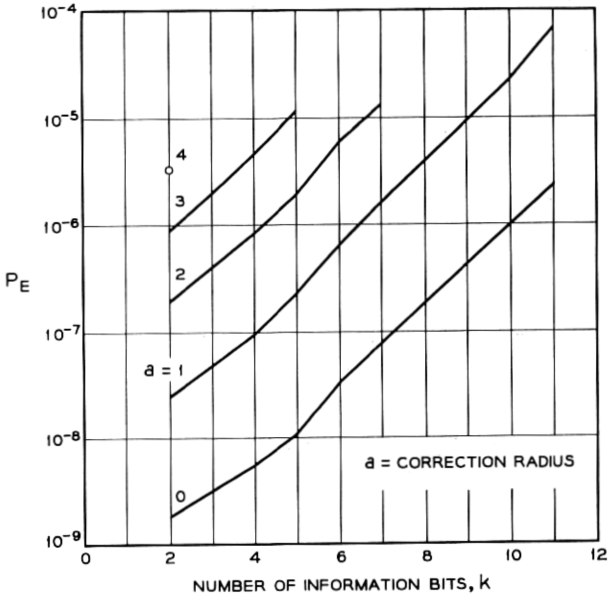

Fig. 2 — Error rates of $(15, k)$ codes.

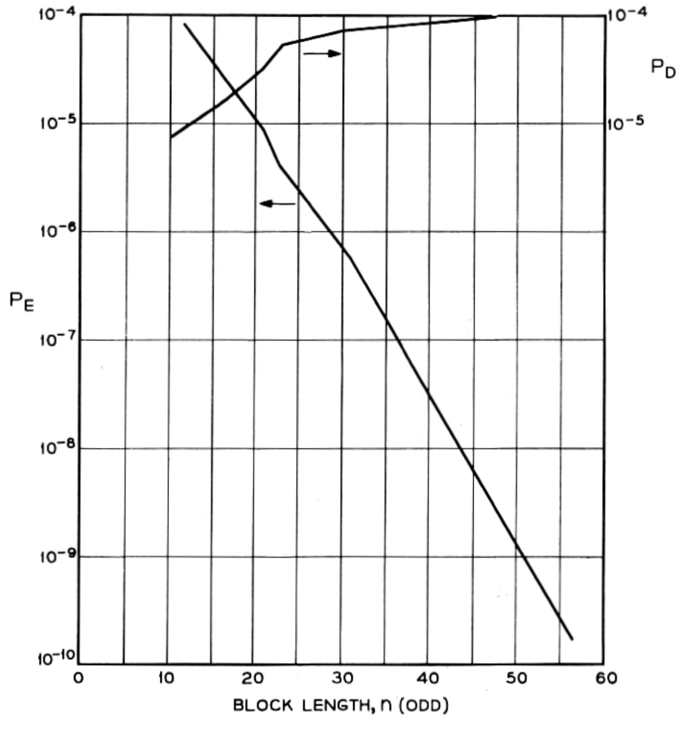

Fig. 3 — Error rates of double error-correcting $\{n, [(n + 1)/2]\}$ codes.

$$a_t(s) = a(t \cdot s).$$

Then the interleaved channels are assumed to have the property that the lengths of the error-free gaps before and after an error are independently distributed — i.e., the errors form a renewal process. (This seems to be a reasonable assumption since interleaving breaks down the memory in the error process. Its accuracy will be discussed below.) From the autocorrelation functions $a_t(s)$ the gap-length distributions $P_t(s)$ may then be calculated by the relations between them which are given in Ref. 4, and finally the $P_t(m,n)$ values are obtained from the $P_t(s)$ by the recursive methods of Ref. 4.

The undetected error rates of some codes used for forward acting error correction only are shown in Fig. 4 as a function of the interleaving constant t. There it is seen that interleaving is most effective on codes correcting four and five errors and is of only modest value on the codes correcting two or three errors. In fact, for the (31, 21) code it would

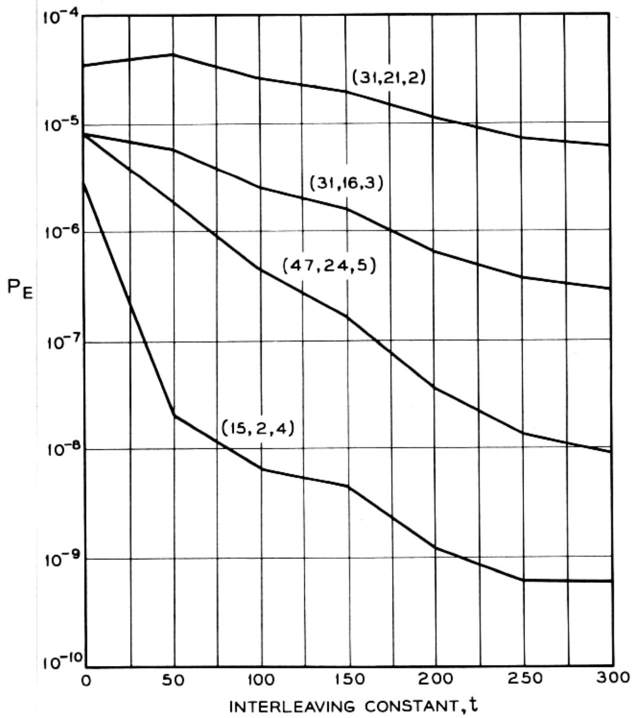

Fig. 4 — Error rate vs interleaving constant.

appear that the error rate increases in the range $t = 1 - 50$. This may result from the fact that the mathematical methods for obtaining $P_t(m,n)$ were different for $t = 1$ and $t = 50$ since the renewal assumption was not made in the case of no interleaving ($t = 1$). The renewal assumption would appear to be more appropriate the larger the interleaving constant t becomes so our estimates of error rates would be more accurate for the larger values of t.

The conclusion to be drawn from Fig. 4 is that it takes a powerful code interleaved extensively to provide a low error rate. The price, in redundancy, interleaving or decoding complexity, paid to do this is high and it would take special circumstances to justify it.

In Fig. 5, block-error rates are plotted against block length to further show the effect of interleaving. The relationship between error rates and block length is linear and the slope is determined by the amount of interleaving. The equivalent memoryless channel is also shown for comparison. It appears that a very considerable amount of interleaving would be required to approach the memoryless channel.

Although it is not shown here, the detected error rates for codes

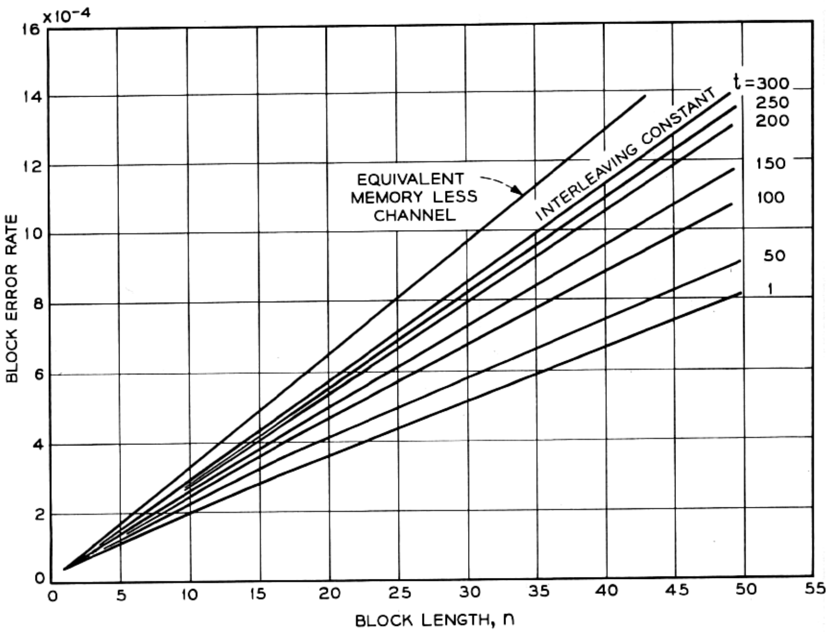

Fig. 5 — Block-error rate vs block length at various interleaving constant.

correcting several errors decrease a few orders of magnitude as the interleaving increases over the range considered here.

We have considered only random error-correcting codes. In Ref. 5, a burst-correcting code was studied at two levels of interleaving as part of an error-control experiment.

V. CONCLUSIONS

Bounded-distance decoding has been examined as a means of utilizing codes to do both error-correction and detection on the switched telephone network with existing data sets. Data containing detected errors would be either voided or marked.

If voids in the received data are permissible (at a rate of about 10^{-4}) then low undetected-error rates may be achieved by a code capable of correcting many errors but used to correct only two or three errors. Such a code might be about 50 percent redundant and have a block length between 25 and 50 bits.

The void rate is rather insensitive to correction radius, block length, and to a lesser extent, interleaving. It decreases with increasing correction radius, increases with increasing block length and decreases with increasing interleaving (for multi-error-correcting codes).

Interleaving is more effective with codes correcting three (or more) errors than those correcting only single or double errors.

If voids are not tolerable then retransmission is indicated as the means to obtain low error rates. A powerful interleaved and highly redundant error-correcting code is required to obtain low error rates. It would probably be called for in only very special cases.

Further work should be undertaken to investigate other methods of decoding codes, such as threshold or burst decoding, in order to gain a more complete insight in the realm of practical error control systems.

REFERENCES

1. Alexander, A. A., Gryb, R. M., and Nast, D. W., Capabilities of the Telephone Network for Data Transmission, B.S.T.J., *39,* May, 1960.
2. Elliott, E. O., Estimates of Error Rates for Codes on Burst-Noise Channels,
3. MacWilliams, Jessie, Permutation Decoding of Systematic Codes, B.S.T.J., *43,* January, 1964, p. 485.
4. Elliott, E. O., A Model of the Switched Telephone Network for Data Communications, B.S.T.J., *44,* January, 1965.
5. Weldon, E. J., Performance of a Forward-Acting Error-Control System on the Switched Telephone Network, B.S.T.J., *45,* May, 1966.