Concise Papers

Error Detection in the Presence of Synchronization Loss

MARTIN E. HELLMAN, MEMBER, IEEE

Abstract—**Cyclic codes are very attractive for error detection because of their low cost encoding and decoding circuits, and because of their high guaranteed minimum distance, but they suffer from very poor protection when word frame synchronization is lost. This note suggests starting the encoder and decoder circuits in an essentially random state, instead of the usual all-zero state. Under this condition of operation it is shown that the undetected error rate against synchronization loss is 2^{-p} where p is the number of parity check bits.**

I. INTRODUCTION

As noted by Bennett and Davey [1, p. 297], cyclic codes offer very poor protection against loss of synchronization. However, because of their low complexity encoding and decoding circuitry, and because of their high guaranteed minimum distance, they are frequently used. Additional protection for detecting synchronization loss is then required.

In all that follows, we assume that the cyclic code of interest is systematic; that it has k information bits, $u = u_1,u_2,\cdots,u_k$, and p parity bits, for a total blocklength of $n = k + p$; that the generator polynomial of the cyclic code [5, pp. 221–223] is $g(D)$; and that the p-stage encoding [5, p. 225] circuit is used. It should be emphasized that in this note we are concerned only with error detection, as opposed to error correction.

First, let us see why cyclic codes offer such poor protection against loss of synchronization. By definition, a cyclic code is one in which any cyclic shift of a codeword is yet another codeword. That is, if $x = x_1,x_2,x_3,\cdots,x_{n-1}, x_n$ is a codeword, so is $x_2,x_3,\cdots,x_{n-1},x_n,x_1$. Now suppose that timing is retarded by 1 bit (synch loss equals $+1$). Assuming no transmission errors occur, we thus think $x_2,x_3,\cdots,x_n$ are the first $n - 1$ transmitted bits. The last bit is due to random channel noise and is denoted by G for garbage. Our supposed codeword is thus $x_2,x_3,\cdots,x_{n-1},x_n,G$. If $G = x_1$ (which occurs 50 percent of the time), then the received sequence is a codeword and the error detection circuitry indicates that the transmission was good. A similar argument shows that synch loss of $\pm m$, with no additive transmission errors, causes an undetected error rate of 2^{-m}. Since we are frequently designing for undetected error rates in the range 10^{-6} to 10^{-12}, synch loss can become the dominant problem, even if it occurs as rarely as one time in 10^5.

Bennett and Davey suggest a simple method for alleviating this shortcoming of cyclic codes. By inverting the parity bits prior to transmission, the cyclic nature of the code is destroyed. If synch is maintained, this does not affect the error control capabilities since the receiver reinverts the received parity checks to produce the original cyclic code. If, however, synch loss is $+1$, then the receiver's inversion produces $x_2 + e_2$, $x_3 + e_3$, $\cdots$, $x_k + e_k$, $\bar{x}_{k+1} + e_{k+1}$, $x_{k+2} + e_{k+2}$, $\cdots$, $x_n + e_n$, $\bar{G}$, where e_i denotes the ith additive transmission error and addition is modulo-two. Now, even if no transmission errors occur ($e_i = 0$ for all i), and even if $\bar{G} = x_1$ (which occurs half the time), there is still a single error in what is thought to be the kth bit, and this is detected by the decoder. Bennett and Davey thus suggest this as a remedy [1, p. 299]. However, it is easily seen that the single error $e_{k+1} = 1$ now goes undetected 50 percent of the time.

Stiffler [2], [3] finds conditions under which modified cyclic codes will be able to detect a combination of synch loss up to $\pm r$ and additional errors provided they number less than d. However, as with most error detection systems, guaranteed performance is frequently much inferior to actual performance.

This note shows that by starting the encoder and decoder in the same, but random, initial state the probability of synch loss going undetected is 2^{-p}, where p is the number of parity bits.

II. RANDOM COSET CODES

Letting $p = n - k$ denote the number of parity check bits, 2^{-p} is the fraction of received sequences which are codewords. Therefore, on a totally noisy channel the undetected error rate is less than 2^{-p} [4]. Since 2^{-p} is quite small for reasonable value of p and, as we shall show, random coset codes make synch loss look like a totally noisy channel, such codes are attractive for detecting loss of synchronization.

A random coset code is one in which a randomly chosen binary sequence $r = r_1,r_2,\cdots,r_n$ is added to each codeword in the cyclic code. The receiver subtracts off r to regain the cyclic code when synchronization is maintained. Then the error control characteristics are the same as for the cyclic code. If, however, synch loss is $+1$ and no transmission errors occur, when the receiver subtracts off r it produces the sequence $x_2 + r_2 - r_1$; $x_3 + r_3 - r_2$, $\cdots$, $x_n + r_n - r_{n-1}$, $G - r_n$. This can be rewritten as $x_2 + r_2'$, $x_3 + r_3'$, $\cdots$, $x_n + r_n'$, $x_1 + r_1'$ where the sequence $r_1',r_2',\cdots,r_n'$ can be shown to be random. Since the receiver sees the codeword $x_2,x_3,\cdots,x_n,x_1$ plus a totally noisy sequence, the probability that this synch loss goes undetected is 2^{-p}. Similar reasoning shows that this error rate is obtained when there are transmission errors and when synch loss takes on any nonzero value.

This 2^{-p} undetected error rate is an average over the 2^n values for r. Using the usual reasoning, this implies that there must be at least one value of r which has an error rate of less than 2^{-p}, and that 99 percent of the values of r have error rates of at most $100(2^{-p})$. But there are clearly certain choices, such as $r = 0$, which should be avoided. So should any choice which allows a small synch loss and a single additive error to go undetected. For example, if r has 1's in its last p positions, but 0's in its first $k = n - p$ positions, we merely invert the parity checks as suggested by Bennett and Davey. However, the above reasoning coupled with simulation, allows designs of this type to be used with confidence. A similar statement applies to the random coding arguments which follow.

Note that the encoder and decoder circuits are the same as those for the original cyclic code, except for the addition of n bits of memory to store r. There are several techniques which can be used to reduce this additional hardware complexity, but care must be exercised. For example, using a pseudorandom, as opposed to a truly random, sequence for r may yield very different performance. In particular, if a maximal length shift register sequence (MLSRS) is used for r, then r is a codeword in a new cyclic code (the MLSRS code). It is possible for the MLSRS code to be a subset of the original cyclic code. In this case the "coset code" generated by the addition of r is the original cyclic code, and no protection against synch loss has been added.

Paper approved by the Associate Editor for Communication Theory of the IEEE Communications Society for publication without oral presentation. Manuscript received December 9, 1974.

The author is with the Department of Electrical Engineering, Stanford University, Stanford, Calif. 94305 and is a Technical Consultant to Vidar Corporation, Mountain View, Calif. 94040.

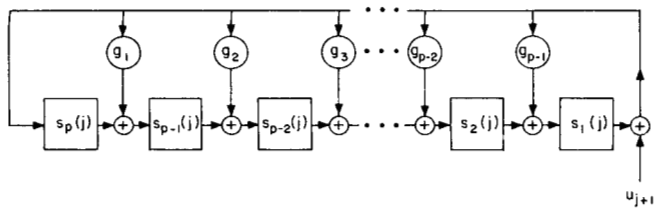

Fig. 1. p-stage encoder.

We summarize two more useful techniques in the following theorems.

Theorem 1: Any coset code can be obtained by adding a p bit sequence to the parity bits of the cyclic codeword while leaving the information bits unchanged. Further, if the p bit sequence is chosen randomly, then the undetected error rate is 2^{-p} when synchronization is lost.

Theorem 2: Any coset code can be obtained by starting the encoder in a nonzero initial state. If the initial state is chosen at random, then the undetected error rate is 2^{-p} when synchronization is lost. We assume that the encoder is the p-stage encoder shown in Fig. 1, and whose operation is described in detail in Gallager [5, p. 225].

Remark: Theorem 1 demonstrates that the additional hardware can be reduced to a p bit memory. Theorem 2 shows that alternatively, no additional hardware is needed if a suitable means exists for driving the encoder circuit to some nonzero state prior to encoding the first information bit. Of course the decoder must be driven to the same nonzero state prior to reception of the first information bit.

Proof of Theorem 1: It is easily seen that two values of $\mathbf{r}$ whose mod-two sum is a codeword in the cyclic code generate the same coset code. Therefore, if

$$r(D) \equiv r_1 D^{n-1} + r_2 D^{n-2} + \cdots + r_n$$
$$= u'(D)g(D) + r'(D),$$

$\mathbf{r}$ and $\mathbf{r}'$ generate the same coset code. Since

$$r'(D) = \text{rem}\,[r(d)/g(D)],$$

it is of degree $p - 1$ and corresponds to adding zeroes to the information bits and a generally nonzero sequence to the parity bits. We thus see that there are only 2^p distinct coset codes, one for each value of $r'(D)$. If all 2^n values of $\mathbf{r}$ are equally likely, we know that synch loss goes undetected with probability 2^{-p} even if there are additional, additive errors. It is easily shown that this distribution induces a uniform distribution on all 2^p values for $r'(D)$. Therefore, choosing $r'(D)$ at random also causes an undetected error rate of 2^{-p} in the presence of synch loss. Q.E.D.

Proof of Theorem 2: The encoder is shown in Fig. 1. At times $j = 1,2,\cdots,k$, the jth information bit u_j is clocked in and the state of the circuit $\mathbf{s}(j)$ changes according to the equation

$$\mathbf{s}(j) = A\mathbf{s}(j-1) + Bu_j$$

where

$$A = \begin{bmatrix} g_{p-1} & 1 & 0 & 0 & \cdots & 0 \\ g_{p-2} & 0 & 1 & 0 & \cdots & 0 \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ g_1 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \qquad B = \begin{bmatrix} g_{p-1} \\ g_{p-2} \\ \cdot \\ \cdot \\ \cdot \\ g_1 \\ 1 \end{bmatrix}$$

and, of course, all arithmetic is done mod-two. From this equation we can easily show that $\mathbf{s}(k)$, the sequence of parity bits, is given by

$$\mathbf{s}(k) = A^k\mathbf{s}(0) + \sum_{i=1}^{k} A^{k-i}Bu(i).$$

If $\mathbf{s}(0) = \mathbf{0}$, then $\mathbf{s}(k) = \text{rem}\,[u(D)/g(d)]$ and we have a cyclic code [5, p. 255]. If $\mathbf{s}(0) \neq \mathbf{0}$, then the parity bits $s_1(k), s_2(k), \cdots, s_p(k)$ are the mod-two sum of the usual parity bits for the cyclic code $\sum_{i=1}^{k} A^{k-1}Bu_i$, and a term $A^k\mathbf{s}(0)$ which depends only on $\mathbf{s}(0)$. We thus see that starting the encoder in state $\mathbf{s}(0)$ is equivalent to adding the vector $\mathbf{R}' = A^k\mathbf{s}(0)$ to the parity bits. All we must do now is to show a one-to-one correspondence between the 2^p possible values for $\mathbf{s}(0)$ and the 2^p possible values of $A^k\mathbf{s}(0)$. But this follows from the fact that A, and hence A^k, is nonsingular. Q.E.D.

ACKNOWLEDGMENT

The author wishes to thank Dr. D. Spaulding of Vidar Corporation for several valuable discussions and suggestions.

REFERENCES

[1] W. R. Bennett and J. R. Davey, *Data Transmission*. New York: McGraw-Hill, 1965.
[2] J. J. Stiffler, "Synchronization of telemetry codes," *IRE Trans. Space Electron. Telem.*, vol. SET-8, pp. 112–117, June 1962.
[3] ——, "Comma-free error-correcting codes," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 107–112, Jan. 1965.
[4] M. Hellman, "Error detection made simple," in *1974 Int. Conf. Commun., Conf. Rec.*, pp. 9A1–9A4.
[5] R. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

Error Rate Lower Bounds for Digital Communication with Multiple Interferences

P. J. McLANE, MEMBER, IEEE

***Abstract*—This paper presents simple and general lower bounds for error rates in digital communication systems. The information bearing signal is taken to be impaired by additive interference and carrier phase jitter. The former is taken to be composed of two components; one component is peak-limited while the other is not. Intersymbol and cochannel interference are examples of peak-limited interference while additive thermal noise is an example of a nonpeak-limited interference. The novelty of the paper is in the generality of the results and in obtaining a simple error bound for transmission in the presence of cochannel interference and carrier phase jitter.**

I. INTRODUCTION

Obtaining upper and lower bounds for system error rate is now an established research area in digital communications [1]. In the past these bounds have been obtained in the presence of interysmbol interference and additive thermal noise [2]–[4]. In addition, upper bounds have recently been presented taking into account additive sinusoidal interference and thermal noise [5], [6]. (Sinusoidal interference is the aggregate of a number of randomly phased sinusoids in the system passband; as such it models both cochannel and adjacent channel interference). All of the studies noted above neglected the effects of carrier phase jitter on the system error rate.

In this work lower bounds on error rate are presented for three types of interference; peak-limited, nonpeak-limited and carrier phase jitter. Intersymbol and sinusoidal interference are representative of peak-limited interference while thermal noise is not peak-limited.

Paper approved by the Associate Editor for Communication Theory of the IEEE Communications Society for publication after presentation at the 7th Biennial Symposium on Communications, Queen's University, Kingston, Ont., Canada, May 30–31, 1974. Manuscript received June 13, 1974; revised December 9, 1974. This work was supported in part by the National Research Council of Canada under Grant A7389 and in part by the Canadian Federal Department of Communications under D.S.S. Contract OPJ2-0012.

The author is with the Department of Electrical Engineering, Queen's University, Kingston, Ont., Canada.