# A SYSTEMATIC METHOD FOR COMPUTING THE MINIMIM FREE DISTANCE OF A CONVOLUTIONAL CODE

**Y.Z. Boutros, G.P. Fiani, E.S. Looka**
Alexandria University, Alexandria, Egypt.

ABSTRACT

**The minimum free distance of a convolutional code is the most important factor affecting its error correcting capability. In this paper we present a systematic method, easy to implement,for computing the minimum free distance. The computer search is minimized by first excluding catastrophic codes,then applying a matrix technique on different code patterns to locate the best codes.**

*Keywords*: Convolutional Code (CC), Free distance, Catastrophic code.

## 1. INTRODUCTION

The development of Convolutional code (CC) theory has been different from that of block codes.With block codes algebraic properties are very important in constructing good classes of codes and in developing decoding algorithms. This is not the case with CC which require extensive computer search to locate good codes.

The minimum free distance ($d_{free}$) is a good indicator of the performance of a CC. A number of techniques to construct good CC have been presented by Massey [1], Berlekamp and [2] Larsen [3].

The algorithm to compute $d_{free}$ proposed by Bahl[4] and modified by Larsen is limited to small values of the constraint length k and to code rates of 1/n. For large values of k,the number of storage locations becomes unacceptably large.

In this paper we present a systematic and easy to implement method for computing $d_{free}$ of a general CC (m,n,k). Although this method uses computer search, the search is minimized by first excluding catastrophic codes [6] and then by applying a matrix technique on different code patterns to locate the best codes. Good CC are tabulated for constraint lengths up to 8 and code rates 1/2, 1/3 and 2/3. 2/3.Our computer results are in agreement with the previous published results [7-10] when available.

## 2. BASIC DEFINITIONS

In this section, we review some of the basic definitions and properties of CC used in the computations of $d_{free}$.

### *2.1. Convolutional Codes*

In general a CC is denoted by (m,n,k) where

n is the number of input bits fed into the encoder

m is the number of output bits that can be obtained

k is the memory length called constraint length.

The ratio n/m is called the code rate (R). The transmitted sequence is generated by convolving the source sequence with a fixed binary function. The encoder complexity depends only on the code rate n/m and the shift register length k. A CC encoder (m,n,k) can be represented by m shift register(s) of lengths $\leq k$ and n modul o-2 adders (implemented by XOR gates) as shown in Figure(1)

### *2.2. State representation and state diagram*

A convolutional encoder is a linear time invariant finite state machine. The state of a CC encoder is the contents of its shift registers. For a binary encoder of constraint length k there are $2^k$ states. The number of links emanating from each node is equal to the number of possible input patterns (two in case of binary). The connection representation and state diagram of a (2,1,2) CC is shown in Figure (2). The 'D' labels on the branches denote the Hamming distance (HD) between the output of the branch transition and the all-zeros path, L denotes one branch unit and denotes that the transition is caused by a '1' input.

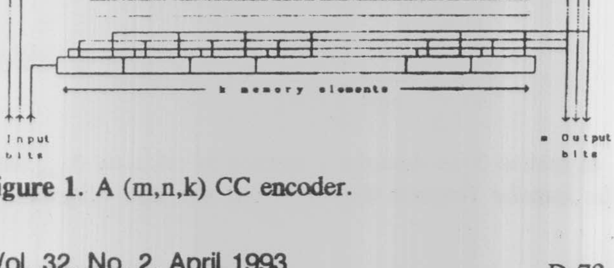
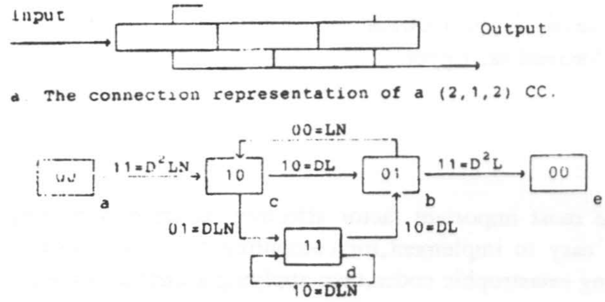



**Figure 1.** A (m,n,k) CC encoder.




a. The connection representation of a (2,1,2) CC.

b. The state digram representation of a (2,1,2) CC.

**Figure 2.** A (2,1,2) CC.

A catastrophic error is defined as an event where by a finite number of code symbol errors cause an in finite number of decoded data bit errors. Massey & Sain [6] have proved that a necessary and sufficient condition for a CC to be catastrophic is that the generators have a common polynomial factor. Catastrophic error propagation exists if there is at least one closed loop of zero Hamming Distance (zero weight) in the state diagram [11].

A systematic convolutional code [12] is one in which the input n-tuple appears as part of the output branch word m-tuple associated with that n-tuple. In each branch the first n symbols are exact replicas to the input n symbols followed by m-n parity or coded symbols. Systematic codes are not catastrophic[11].

### *2.3. The minimum free distance*

The minimum free distance of a CC, $d_{free}$,is defined as follows [13]:

$$d_{free} = \min \{HD(V', V''): U' \neq U''\},$$

where V'' and V″ are the codewords corresponding to the information sequences U′ and U″ respectively (after adding zeros to the shorter one to become of equal lengths). Because a CC is a linear code, i.e., the modulo-2 sum of any two codewords is also a codeword; $d_{free}$, is also the minimum distance between all codewords sequences and the all-zeros sequence. It is the minimum-weight codeword of any length produced by a non-zero information sequence.

Heller has derived a relatively simple upper bound on $d_{free}$ of a (m,1,k) CC, namely

$$d_{free} \le \min_{r \ge 1} \left[ \frac{2^{r-1}}{2^r - 1} (K + r - 1) m \right].$$

In section 3 we describe a method to calculate $d_{free}$ using the transfer function derived from the state diagram. In section 4 we present a matrix formulation technique for the transfer function T(D) and in section 5 we propose a systematic method for calculating $d_{free}$ only by matrix multiplication using the formulation in section 4.

## 3. MINIMUM FREE DISTANCE CALCULATION FROM THE TRANSFER FUNCTION.

The transfer function T(D,L,N), or generating function, of the code is expressed as

[T(D,L,N) = final state initial state.

Consider the state diagram in Figure(2). The state equations are written as follows:

$$c = D^2 L N + L N b$$

$$b = D L c + D L d$$

$$d = D L N c + D L N d$$

$$e = D^2 L b$$

where 'a' denotes the initial state and 'e' the final state. Solving these equations, we obtain the transfer function:

$$T(D,L,N) = \frac{D^5 L^3 N}{1 - D(L)(1+L)N}$$

and using graph theory,

$$T(D,L,N) = D^5 L^3 N + D^6 L^4 (1+L) N^2 + D^7 L^5 (1+L)^2 N^3 + \dots$$

which gives the following meanings[15]:

1. There is only one path of distance 5 length 3 caused by one '1'.
2. There are two paths of distance 6, one of length 4, the other of length 5 and both are caused by two 1'nns.

This information is the weight structure of the code that determines the code performance.

$d_{free}$ is the minimum weight of all paths in the state diagram that diverge from and remerge with the all-zeros state [16]. $d_{free}$ is the lowest power of D in the transfer function. The error correcting capability, t, of a CC is,

$$t = \left[ \frac{d_{free} - 1}{2} \right]$$

Now since we are interested in $d_{free}$, we ignore all parameters not related to the distance, namely we assign a value of 1 to both L and N. Thus we obtain the transfer function T(D),

$$T(D) = \frac{D^5}{1-2D}$$

$$= D^5 + 2D^6 + 4D^7 + \ldots$$

the minimum free distance of this code is 5, the lowest power of D in the transfer function.

## 4. MATRIX FORMULATION TECHNIQUE FOR T(D).

### 4.1. *The next state and output matrices*

The state diagram can be represented by two matrices: The next state matrix (NS) and the output matrix (OP). The states constitute the rows of both matrices and the different input patterns constitute the columns. For an (m,n,k) binary CC both matrices have dimensions $2^k \times 2^n$. The next state matrix and the output matrix for the (2,1,2) CC of Figure (2) are shown in Table I.

**Table I.** The next state and the output matrices of a (2,1,2) CC.

| INPUT → / ↓ STATE | | | 0 | 1 |
|---|---|---|---|---|
| a | 1 | 00 | 1 | 3 |
| b | 2 | 01 | 1 | 3 |
| c | 3 | 10 | 2 | 4 |
| d | 4 | 11 | 2 | 4 |

Next state matrix (NS)

| INPUT → / ↓ STATE | | | 0 | 1 |
|---|---|---|---|---|
| a | 1 | 00 | 1 | $D^2$ |
| b | 2 | 01 | $D^2$ | 1 |
| c | 3 | 10 | D | D |
| d | 4 | 11 | D | D |

Output matrix (OP)

Let NST denote the number of states ($NST=2^{k+n}$) and NIN the number of elements of the input alphabet ($NIN = 2^n$), then, NS (i,j) = next state reached from state i in case the input is j;

$$1 \le 1 \le NST; \; 1 \le j \le NIN;$$

and OP (i,j) = is the output generated with the transition NS(i,j).

These matrices can be automatically generated by a computer program given a prespecified encoder connections.

### 4.2. *The transfer function in matrix form*

A technique used to describe T(D) is by rewriting the state equations as follows [13]:

$$x_3 = D^2 + x_2$$

$$x_2 = D\,x_3 + D\,x_4$$

$$x_4 = D\,x_3\,D\,x_4$$

$$x_1 = D^2\,x_2$$

where $x_1$, $x_2$, $x_3$ and $x_4$ denote the states a,b,c, and d respectively. Now, separating state 00 into an originating and a terminating state, we can write in matrix form:

$$\begin{matrix}10\\01\\11\\00\end{matrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} = \begin{bmatrix}0 & D^2 & 0 & 0\\0 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} + \begin{bmatrix}0\\0\\D^2\\0\end{bmatrix}$$

$$[X] = [A]\,[X] + [X_o]$$

$$[X] = [I-A]^{-1}\,[X_o]$$

where $[I-A]^{-1}$ exists for non catastrophic codes[13]. Therefore,

$$[X] = \{\,[I] + [A] + [A^2] + [A^3] \ldots\ldots\,\}\,[X_o].$$

This expression is equivalent to T(D)and isuseful for calculating the transfer function in case of large number of states. It is clear that the matrix A is sparse and for (2,1,2)CC the number of nonzero elements in A are 6 while the total number of elements in A is 16. In general for an (m.n.k) binary CC, the matrix A consists of $2^k \times 2^k$ elements, of which at maximum $2^k \times 2^n - n$ are non-zero. The ratio is approximately $1/(2^{k-n})$.

The matrix A can be constructed in as systematic manner as follows:

$$A(i, NS(i,j)) = OP(i, NS(i,j)) \; ; \; i=2 \rightarrow NST, \; j=1 \rightarrow NIN$$

and

$$X_o\,(NS(1,j)) = OP(1,j); \; j=2 \rightarrow NIN.$$

### 4.3. *Minimum free distance calculation.*

It is required to calculate the least power of D in the transfer function T(D). In order to perform computer calculations, we substitute for D the value 2 in the matrices A and $X_o$, thus,

$$(I-A)^{-1} = \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} + \begin{bmatrix}0&4&0&0\\0&0&2&2\\0&1&0&0\\0&0&2&2\end{bmatrix} + \begin{bmatrix}0&4&0&0\\0&0&2&2\\0&1&0&0\\0&0&2&2\end{bmatrix}^2 + \begin{bmatrix}0&4&0&0\\0&0&2&2\\0&1&0&0\\0&0&2&2\end{bmatrix}^3 + \ldots$$

$$(I-A)^{-1}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}+\begin{bmatrix}0&4&0&0\\0&0&2&2\\0&1&0&0\\0&0&2&2\end{bmatrix}+\begin{bmatrix}0&0&8&8\\0&2&4&4\\0&0&2&2\\0&0&4&4\end{bmatrix}+\begin{bmatrix}0&8&6&6\\0&4&2&2\\0&2&4&4\\0&4&2&2\end{bmatrix}+\dots$$

and $X_o$ is $[0\ 4\ 0\ 0]^T$.

Now taking only two terms of the expansion, i.e. I+A, gives $x_1 = 0$.

Taking three terms gives $x_1 = 32$ which is $2^5$, the equivalent of $D^5$. Thus $d_{free}$ for this code is 5.

However, there are some cases in which the matrix A needs to be raised to a power more than two to give a correct result. Consider another (2,1,2) CC with the following matrices A and $X_o$.

$$A=\begin{bmatrix}0&2&0&0\\0&0&4&2\\0&2&0&0\\0&0&1&2\end{bmatrix}\quad X_o=\begin{bmatrix}0\\0\\2\\0\end{bmatrix}$$

In this case, the expansion up to $A^2$ gives the value $x_1=16$, while it gives the value $x_1$-8 if the expansion is truncated after $A^3$. Thus we must take in consideration not to truncate the infinite series unless we obtain a non-decreasing value of $x_1$. Extensive computer simulations showed that if x has non-decreasing values for two successive iterations, it will not decrease any more.

## 5. PROPOSED SYSTEMATIC METHOD FOR MIN FREE DISTANCE CALCULATION.

A systematic method for calculating $d_{free}$ consists of the following steps:

1. Given a (m,n,k) CC, construct the NS matrix.
2. Generate all possible m output connections (their number is $2^{(k+n)m}$).
3. Discard all output connections that constitute a catastrophic code. Any method can be used to test whether a given code is catastrophic or not. The efficient technique for detecting a catastrophic (m,n,k) CC that we have previously presented[17] is recommended.
4. Calculate the OP matrix for each output combination.
5. Use the NS and OP matrices to form the sparse matrix A and the vector $X_o$ corresponding to the transfer function T(D).
6. Calculate the powers of the matrix A and each time calculate $AX_o$ until the value of $x_1$ does not decrease any more.
7. $d_{free}$ for that code with the specified output connections is calculated by the expansion $[I-A]^{-1}\ X_o$.
8. Repeat for all output combinations at step 4.
9. The minimum free distance of the code is the least $d_{free}$ obtained in the cycle of steps 4 through 9.

**Table II.** The number of catastrophic for CC of different rates and constraint lengths.

| Code | Catastrophic | Non Catast. | Total Number | Ratio |
|---|---|---|---|---|
| (2,1, 2) | 12 | 16 | 28 | 0.429 |
| (2,1, 3) | 56 | 64 | 120 | 0.467 |
| (2,1, 4) | 240 | 256 | 496 | 0.484 |
| (2,1, 5) | 992 | 1024 | 2016 | 0.492 |
| (2,1, 6) | 4032 | 4096 | 8128 | 0.496 |
| (2,1, 7) | 16256 | 16384 | 32640 | 0.498 |
| (2,1, 8) | 65281 | 65536 | 130817 | 0.499 |
| (2,1, 9) | 261638 | 262144 | 523782 | 0.500 |
| (2,1,10) | 1047580 | 1048576 | 2096156 | 0.500 |
| (2,1,11) | 4192376 | 4194304 | 8386680 | 0.500 |
| (2,1,12) | 16773620 | 16777220 | 33550830 | 0.500 |
| (2,1,13) | 67102690 | 67108860 | 134211600 | 0.500 |
| (3,1, 2) | 20 | 64 | 84 | 0.238 |
| (3,1, 3) | 168 | 512 | 680 | 0.247 |
| (3,1, 4) | 1360 | 4096 | 5456 | 0.249 |
| (3,1, 5) | 10912 | 32768 | 43680 | 0.250 |
| (3,1, 6) | 87360 | 262144 | 349504 | 0.250 |
| (3,1, 7) | 699008 | 2097152 | 2796160 | 0.250 |
| (3,1, 8) | 5592321 | 16777220 | 22369540 | 0.250 |
| (3,1, 9) | 44739070 | 134217700 | 178956800 | 0.250 |
| (3,1,10) | 357913700 | 1073742000 | 1431655000 | 0.250 |
| (3,1,11) | 2863310000 | 8589934000 | 11453250000 | 0.250 |
| (3,1,12) | 22906490000 | 68719480000 | 91625970000 | 0.250 |
| (4,1, 2) | 30 | 256 | 286 | 0.105 |
| (4,1, 3) | 420 | 4096 | 4516 | 0.093 |
| (4,1, 4) | 6120 | 65536 | 71656 | 0.085 |
| (4,1, 5) | 92752 | 1048576 | 1141328 | 0.081 |
| (4,1, 6) | 1441440 | 16777220 | 18218660 | 0.079 |
| (4,1, 7) | 22717760 | 268435500 | 291153200 | 0.078 |
| (4,1, 8) | 360704700 | 4294968000 | 4655672000 | 0.077 |
| (4,1, 9) | 5748970000 | 68719480000 | 74468450000 | 0.077 |
| (4,1,10) | 91804820000 | 1099512000000 | 1191316000000 | 0.077 |
| (4,1,11) | 1467446000000 | 17592190000000 | 19059630000000 | 0.077 |
| (4,1,12) | 23467690000000 | 281475000000000 | 304942700000000 | 0.077 |
| (5,1, 2) | 42 | 1024 | 1066 | 0.039 |
| (5,1, 3) | 924 | 32768 | 33692 | 0.027 |
| (5,1, 4) | 23256 | 1048576 | 1071832 | 0.022 |
| (5,1, 5) | 649264 | 33554430 | 34203700 | 0.019 |
| (5,1, 6) | 19315300 | 1073742000 | 1093057000 | 0.018 |
| (5,1, 7) | 595205500 | 34359740000 | 34954950000 | 0.017 |
| (5,1, 8) | 18684490000 | 1099512000000 | 1118196000000 | 0.017 |
| (5,1, 9) | 592143700000 | 35184370000000 | 35776520000000 | 0.017 |
| (5,1,10) | 18856710000000 | 1125900000000000 | 1144757000000000 | 0.016 |

## 6. RESULTS AND DISCUSSION

The method described above is based on a computer search. However this search is minimized by first excluding all catastrophic codes using the proper technique presented in the our paper[17] then proceeding to the matrix construction and multiplication steps. In Table II we present a list of the number of catastrophic codes for selected CC. The ratio of the number of catastrophic codes to the total number of possible CC is about 50% for rate 1/2, 25% for rate 1/3, 7% for rate 1/4 and 1% for rate 1/5. It is useful to know from the beginning that all the codes having even decimal equivalents of all the generator polynomials are catastrophic. Also, for the special case (m,1,k), if each adder in the encoder has an even number of connections, the code is catastrophic. Thus, it is very useful to discard all catastrophic codes before attempting to calculate $d_{free}$, especially for rate 1/2 CC. It should be noted that the calculation of $d_{free}$ for an (m,n,k) CC is equivalent to calculating $d_{free}$ for a (m,1,n+k-1) CC.

There is also a remarkable reduction in the total number of computations required for matrix multiplication. Instead of performing additions and multiplications of order $n_3$ to multiply two arbitrary matrices of order n, the matrix A, being a sparse matrix, requires much less effort for multiplication. However the resultant matrices $A_2, A_3$ ,..tend to lose this property. Fortunately, the number of times A should be raised to calculate $d_{free}$ is relatively small. Extensive computer search showed that the matrix A needs not to be raised to a power more than 2 or 3 in almost all cases for the calculation of $d_{free}$. Table III lists the best codes for selected systematic and non systematic CC. Output connections are listed in their decimal form. It should be noted that making a CC systematic reduces the maximum possible distance for a system constraint length and rate. However, its advantages are that no inverting circuit (decoder) is needed for recovering the information sequence from the codeword and that the encoder realization requires fewer connections as compared to non-systematic codes.

Our results are in agreement with previously published results[11,12,16].

**Table III**. The best CC and their relative $d_{free}$.

a. Best Nonsystematic CC.

| CC | $d_{free}$ | Connections in decimals |
|---|---|---|
| (2,1,2) | 5 | 5, 7 |
| (2,1,3) | 6 | 13, 15 |
| (2,1,4) | 7 | 19, 29 |
| (2,1,5) | 8 | 43, 61 |
| (2,1,6) | 10 | 91, 121 |
| (2,1,7) | 10 | 167,249 |
| (2,1,8) | 12 | 369,491 |
| (3,1,2) | 8 | 5, 7, 7 |
| (3,1,3) | 10 | 11, 13, 15 |
| (3,1,4) | 12 | 21, 27, 31 |
| (3,1,5) | 13 | 39, 43, 61 |
| (3,1,6) | 14 | 91, 117, 121 |
| (3,1,7) | 16 | 149, 216, 247 |
| (3,2,2) | 3 | 6,2 2,4 6,4 |
| (3,2,3) | 4 | 4,1 2,4 6,7 |
| (3,2,4) | 5 | 7,2 1,5 4,7 |
| (3,2,5) | 6 | 48,12 24,32 56,60 |
| (3,2,6) | 7 | 48,24 24,52 52,60 |
| (3,2,7) | 8 | 48,14 28,38 44,60 |

b. Best Systematic CC.

| CC | $d_{free}$ | Connections in decimals |
|---|---|---|
| (2,1,2) | 3 | 3 |
| (2,1,3) | 3 | 3, 7 |
| (2,1,4) | 4 | 7, 11 |
| (2,1,5) | 4 | 11,23 |
| (2,1,6) | 5 | 43,55 |
| (2,1,7) | 5 | 43,55 |
| (2,1,8) | 6 | 59,103 |
| (3,1,2) | 4 | 1, 3 |
| (3,1,3) | 5 | 3, 5 |
| (3,1,4) | 6 | 3, 13 |
| (3,1,5) | 7 | 7, 13 |
| (3,1,6) | 8 | 7, 45 |
| (3,1,7) | 9 | 71,45 |

## 7. CONCLUSION

The present method is efficient as compared to the traditional approach to derive $d_{free}$ form the transfer function as described in Section 3. Our systematic algorithm can be easily programmed on a digital computer to calculate $d_{free}$. The reduction of the number of calculations makes it possible to calculated $d_{free}$ for any (m,n,k) not restricted to short constraint lengths.

## REFERENCES


[1] J.L. Massey, "Threshold Decoding", *Res. Monograph 20*, Cambridge, Mass. MIT press, 1963.

[2] E.R. Berlekamp, "Notes on recurrent codes", *IEEE Trans. Info. Theory*, Vol. IT-10, pp. 257-258, July 1964.

[3] K.J.Larsen, "Short CC with maximal free distance for rates, 1/2, 1/3 and 1/4", *IEEE Trans. Info. Theory*, Vol. IT-1, pp. 1371-372, May 1973.

[4] L.R. Bahl, C.D. Cullum, W.D. Frazer and F.Jelinek, "An efficient, algorithm for computing free distance", *IEEE Trans. Info. Theory*, Vol. IT-18, pp. 437-439, May 1972.

[5] K.J. Larsen, "Comments on an efficient algorithm for computing, free distance", *IEEE Trans. Info. Theory*, pp. 577-579, July 1973.

[6] J.L. Massey and M.K. Sain, "Inverses of linear sequential, circuits", *IEEE Trans. Computers*, Vol. C17, pp. 330-337, April, 1968.

[7] J.P. Odenwalder, "Optimal Decoding of CC", *Ph.D. Dissertation*, Dept of systems Sciences, School of Engineering and Applied, Sciences, UCLA, 1970.

[8] E. Paaske, "Short Binary CC with maximal free distance for, rates 2/3 and 3/4", *IEEE Trans. Info Theory*, Vol. IT-20, pp. 683-689, Sep 1974.

[9] D.G. Daut, J.W. Modestino and L.D. Wismer, "New Short constraint,length CC construction for selected rational rate", *IEEE Trans. Info. Theory*, Vol. IT-28, pp. 793-799, Sept. 1982.

[10] P.J. Lee, "Further Results on Rate 1/N CC Construction with, minimum SNR criterion", *IEEE Trans. Comm.*, Vol. COM34, pp. 1396-399, April 1986.

[11] B. Sklar, *Digital Communication Fundamentals and, Applications*, Prentice Hall, Englewood Cliffs, NJ, 1988.

[12] A.J.Viterbi and J.K.Omura, Principles of *Digital Communications and Coding*, Mc-Graw Hill, 1979.

[13] S. Lin and D.J. Costello, *Error Control Coding*, Prentice Hall, NJ, 1983.

[14] A.J. Heller, "Short Constraint Length CC", Jet Propulsion, Laboratory California Institute of Technology, Pasadena, CA, *Space Program Summary* 37-54, Vol. 3, pp.171-174, December, 1968.

[15] G.C.Clark and J.B.Cain, *Error Correction Coding for Digital Comm.*, Plenum Press, NY, 1981.

[16] A.M. Michelson and A.H. Levesque, *Error Control Techniques for, Digital Communication*, John Wiley and Sons, NY, 1985.

[17] Y.Z. Boutros, G. Fiani and E. Looka, "An Efficient Technique for, Detecting Catastrophic CC", Submitted to *Alexandria, Engineering Journal*, Faculty of Engineering, Alexandria University.