On Classes of Convolutional Codes that are not Asymptotically Catastrophic

Kjell J. Hole,* *Member, IEEE*

April 23, 1999

Abstract

Denote by w_0 the minimum average weight per edge over all nonzero cycles in the state diagram for a convolutional code, and assume that a technique is available for generating canonical parity-check matrices for codes with increasing degree m. The obtained class of codes is asymptotically catastrophic if w_0 approaches zero for large m. In this paper we prove the existence of convolutional code classes that are not asymptotically catastrophic by providing explicit constructions of codes with nonzero w_0 for all m.

Index Terms—asymptotically catastrophic, constructions, convolutional codes.

I. Introduction

In this paper we prove the existence of convolutional code classes that are not asymptotically catastrophic by providing explicit constructions and an existence proof. To obtain our results, we use some of the definitions and notations established in [10]. Hence, a convolutional code with block length n, dimension $n-r$, code degree m, and free distance d_{free} is denoted by

*The author is with the Department of Informatics, University of Bergen, HiB, N-5020 Bergen, Norway. Internet: `Kjell.Hole@ii.uib.no`. This work was supported by the Norwegian Research Council, NFR.

$(n, n-r, m, d_{free})$.[1] A code is represented by a $r \times n$ canonical parity-check matrix or an $(n-r) \times n$ canonical generator matrix [10, Sec. 3 and 7].[2] We only consider codes with symbols in the finite field F (=GF(2)).

It is well known that a state diagram representation for a code may be obtained from the parity-check matrix or the generator matrix. A path of length $\tau\,(\geq 1)$ in the state diagram consists of τ consecutive edges. A cycle of length τ is a path of length τ that starts and ends in the same state s and whose other states are different from each other and s. The average weight per edge of a cycle is equal to the total Hamming weight of all labels on the edges divided by τ.

Denote by w_0 the *minimum average weight per edge* [7] over all cycles in a state diagram, excluding the cycle of length one and weight zero that starts and ends in the zero state. Consider a class of $(n, n-r, m)$ convolutional codes for which canonical parity-check (or generator) matrices can be generated by applying some systematic method. The class is said to be *asymptotically catastrophic* [3] if w_0 approaches zero when the method is used to generate matrices with increasing degree m.

Hemmati and Costello [3] proved that $w_0 \leq 2^{m-2}/(3 \cdot 2^{m-2} - 1)$ for a class of $(2, 1, m)$ convolutional codes. Hole and Hole [5] showed that the bound is in fact valid for a large class of $(n, n-1, m)$ codes with $n \geq 2$. Recently, Hole [4] derived even tighter upper bounds on w_0 for smaller classes of $(n, n-1, m)$ codes. The classes in [4] are asymptotically catastrophic since the bounds approach zero for large m.

Wyner and Ash [10, Sec. 7] have given a construction for $(2^m, 2^m-1, m, 3)$ codes. Hole and Hole [5] have shown that—since the 1×2^m canonical parity-check matrices for these codes all contain the polynomial $1 + D^m$—the codes have minimum possible $w_0 = 1/m$. Because w_0 approaches zero for large m, the class of Wyner-Ash codes is asymptotically catastrophic. The same is true for the class of $(2^{m-1}, 2^{m-1} - 1, m, 4)$ extended Wyner-Ash codes.

There is clearly a need for classes of convolutional codes with w_0 bounded away from zero. To determine such classes, we first show in Section II that

[1]The code degree m is called the (overall) constraint length by many authors.

[2]A $k \times n$ polynomial matrix is canonical if the gcd of its $k \times k$ minors is 1 and the greatest degree of the minors is m.

w_0 for a convolutional code $\mathcal{C}$ is lower bounded by $d_{\mathcal{E}}/e_{max}$ where $d_{\mathcal{E}}$ is the minimum distance of a block code embedded in $\mathcal{C}$, and e_{max} is the maximum row degree of a canonical parity-check matrix for $\mathcal{C}$. The bound is then used to identify classes of codes with large w_0.

In Section III we construct a class of $(2^{2m}-2^m, 2^{2m}-2^m-m, m, 3)$ convolutional codes with $w_0 \geq 1$ for $m \geq 1$. The construction is modified to obtain a class of $(2^{2m-1}, 2^{2m-1}-m, m, 4)$ codes with $w_0 \geq 1$ for $m \geq 2$.

Section IV shows how high-rate convolutional codes with large $d_{free}\,(> 4)$ and large $w_0\,(> 1)$ can be obtained from block codes. The construction technique is applied to binary BCH codes in Section V to construct classes of $(2^m-1, 2^m-1-(t-1)m, m)$ convolutional codes with $d_{free} \geq 2t+1$ and $w_0 \geq 2t-3$ for $2 \leq t < 2^{\lceil m/2 \rceil -1}+1$ and $m \geq 3$. We also obtain classes of $(2^m-1, 2^m-2-(t-1)m, m)$ codes with $d_{free} \geq 2t+2$ and $w_0 \geq 2t-2$.

It is shown in Section VI that there exist classes of low-rate codes whose free distances increase with m and whose minimum average weights w_0 are large for all m. As an example, we construct a class of $(m+1, 1, m)$ codes with $d_{free} \geq 3m$ and $w_0 \geq 1$ for all $m \geq 2$.

A summary and a discussion are given in Section VII.

II. Convolutional codes with large w_0

Let an $(n, n-r, m)$ binary convolutional code $\mathcal{C}$ be defined by a canonical parity-check matrix, and let e_i denote the maximum degree of the polynomials in the ith row of the matrix, $1 \leq i \leq r$. The maximum degrees e_i are the *row degrees*, or Forney indices, of the canonical parity-check matrix. Assume that the rows of the parity-check matrix are ordered such that the row degrees satisfy $e_1 = e_2 = \cdots = e_\gamma = 0$ for some γ, $0 \leq \gamma < r$, and $1 \leq e_{\gamma+1} \leq e_{\gamma+2} \leq \cdots \leq e_r$. All ordered canonical parity-check matrices for a code have the same row degrees and the sum $e_1 + e_2 + \cdots + e_r$ of the row degrees is equal to the code degree m [10, Sec. 3].

If $\gamma > 0$, then the γ first rows of the parity-check matrix for the convolutional code $\mathcal{C}$ define a parity-check matrix for an $[n, n-\gamma]$ binary block code, $\mathcal{E}$, called the *embedded block code* [6]. For $\gamma = 0$, we define $\mathcal{E}$ to be the trivial $[n, n]$ block code consisting of all possible binary n-tuples. Let $d_{\mathcal{E}}$ be the minimum distance of $\mathcal{E}$ and note that $d_{\mathcal{E}} = 1$ for $\gamma = 0$.

The labels on the edges in the state diagram for $\mathcal{C}$ are the codewords in $\mathcal{E}$ [6]. Observe that no nonzero label can have weight less than $d_{\mathcal{E}}$. As an example, the 2×3 canonical parity-check matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & D & 1+D \end{bmatrix}$$

has row degrees $e_1 = 0$ and $e_2 = 1$, i.e. $\gamma = 1$. The first row of the matrix defines a $[3,2]$ embedded block code $\mathcal{E}$ with $d_{\mathcal{E}} = 2$.

Consider the all-zero paths in the state diagram for $\mathcal{C}$. Since the diagram contains a cycle of length one and weight zero that starts and ends in the zero state, there exists at least one all-zero path of length τ for any $\tau \geq 1$. Let $e_{max} = \max_{1 \leq i \leq r}\{e_i\} = e_r$. Forney [2] has shown that the number of distinct all-zero paths of length τ is equal to $2^{d(\tau)}$ where

$$d(\tau) = \sum_{i:e_i \geq \tau} (e_i - \tau).$$

For $\tau \geq e_{max}$, we have $d(\tau) = 0$, and the unique path of length τ is obtained by cycling around the zero state cycle τ times. When $\tau < e_{max}$, there exist more than one path of length τ. Consequently, the longest all-zero path nowhere merged with the zero state has length $e_{max} - 1$, and any path of length e_{max} that does not contain the zero state (except perhaps as the ending state) has weight at least equal to $d_{\mathcal{E}}$. As an immediate consequence, we have:

Theorem 1: Any $(n, n-r, m)$ convolutional code defined by a canonical parity-check matrix with maximum row degree e_{max} has a state diagram representation with minimum average weight per edge $w_0 \geq d_{\mathcal{E}}/e_{max}$, where $d_{\mathcal{E}}$ is the minimum distance of the embedded block code.

For $e_{max} = 1$, it follows from the above discussion that an all-zero path can only be obtained by cycling around the zero state cycle. Hence, we have the following corollary.

Corollary 1: If an $(n, n-r, m)$ convolutional code is defined by a canonical parity-check matrix with maximum row degree $e_{max} = 1$, then all edges in the state diagram have weights at least equal to $d_{\mathcal{E}}$, except the edge constituting the zero state cycle, and $w_0 \geq d_{\mathcal{E}}$.

It is clear from Theorem 1 that a class of $(n, n-r, m)$ convolutional codes is not asymptotically catastrophic if all codes are defined by canonical parity-check matrices with small maximum row degrees e_{max}. In the next section we show that such classes can be obtained by letting $r = r(m)$ be an increasing function of m because, in this case, we can increase $m = \sum_{i=1}^{r(m)} e_i$ without changing e_{max}.

III. Two simple class constructions

Ytrehus [14] has shown how to construct parity-check matrices for $(n, n-r, m)$ convolutional codes with free distances 3 and 4 for any $r \geq 1$. We use a modified version of Ytrehus' construction technique for $r = m$. Unlike Ytrehus, we prove that the constructed parity-check matrices are canonical and that the obtained code classes are not asymptotically catastrophic. We also provide alternative proofs for the free distances.

A. Construction of codes with $d_{free} = 3$ and $w_0 \geq 1$

Let $a \leq 2^{2m} - 1$ be a positive integer with binary representation

$$B_a = (b_0 b_1 b_2 \cdots b_{2m-1})$$

where b_{2m-1} is the most significant bit. We require that a has a value such that not all bits in the even positions $\{b_0, b_2, b_4, \cdots, b_{2m-2}\}$ are equal to zero. Note that there are $2^{2m} - 2^m$ values of a for which this is true. The $2m$ bits of the vector B_a can be divided into m parts containing 2 bits each:

$$B_a = (b_0 b_1, b_2 b_3, \cdots, b_{2m-2} b_{2m-1}).$$

If we let the bits in each part be the coefficients of a binary polynomial in the indeterminate D, then we obtain a vector of m polynomials

$$B_a(D) = (b_0 + b_1 D, b_2 + b_3 D, \cdots, b_{2m-2} + b_{2m-1} D).$$

As an example, for $m = 3$ and $a = 9$ we have

$$B_9 = (10, 01, 00) \text{ and } B_9(D) = (1, D, 0).$$

Let $\mathbf{H}_m(D)$ denote an $m \times n$ canonical parity-check matrix for an $(n, n-m, m, 3)$ convolutional code. We use the set of vectors $\{B_a(D)\}$ to define a general expression for a matrix $\mathbf{H}_m(D)$ with maximum row degree that satisfies the requirement of Corollary 1. The matrix is given by

$$\mathbf{H}_m(D) = \left[\begin{array}{ccccc} B_1^T(D) & B_3^T(D) & B_4^T(D) & \cdots & B_{2^{2m}-1}^T(D) \end{array} \right] \tag{1}$$

where $B_a^T(D)$ denotes the transpose of the polynomial m-vector $B_a(D)$. Clearly, $\mathbf{H}_m(D)$ has $n = 2^{2m} - 2^m$ columns. For $m = 2$, the matrix in (1) is equal to

$$\left[\begin{array}{cccccccccccc} 1 & 1+D & 0 & 1 & D & 1+D & 1 & 1+D & 0 & 1 & D & 1+D \\ 0 & 0 & 1 & 1 & 1 & 1 & D & D & 1+D & 1+D & 1+D & 1+D \end{array} \right].$$

Theorem 2: The $m \times (2^{2m} - 2^m)$ canonical parity-check matrix in (1) defines a $(2^{2m} - 2^m, 2^{2m} - 2^m - m, m, 3)$ convolutional code with $w_0 \geq 1$ for all $m \geq 1$.

In the following we prove Theorem 2 by establishing two lemmas. We first show that the parity-check matrix is canonical, i.e., the gcd of the $m \times m$ minors is 1 and their greatest degree is equal to the code degree m [10, Sec. 3]. We then prove that the codes have $d_{free} = 3$ and $w_0 \geq 1$.

Lemma 1: The $m \times (2^{2m} - 2^m)$ parity-check matrix defined by (1) is canonical.

Proof: It is sufficient to show that $\mathbf{H}_m(D)$ contains one $m \times m$ submatrix with determinant 1 and another submatrix with determinant $(1+D)^m$. The matrix in (1) contains the $m \times m$ minor

$$\left| \begin{array}{cccc} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{array} \right|$$

where the columns are given by $B_i^T(D)$, $i \in \{2^j \mid j = 0, 2, 4, \cdots, 2m-2\}$, and the empty spaces denote zeros. Clearly, the determinant is 1. The $m \times m$ minor with columns $B_k^T(D)$, $k \in \{2^l + 2^{l+1} \mid l = 0, 2, 4, \cdots, 2m-2\}$,

$$\left| \begin{array}{cccc} 1+D & & & \\ & 1+D & & \\ & & \ddots & \\ & & & 1+D \end{array} \right|$$

has determinant $(1+D)^m$. Q.E.D.

Lemma 2: The $(2^{2m}-2^m, 2^{2m}-2^m-m, m)$ convolutional code defined by the canonical parity-check matrix in (1) has $d_{free}=3$ and $w_0 \geq 1$.

Proof: There can be no codeword of weight 1 since no column of $\mathbf{H}_m(D)$ contains only zeros. A word of weight 2 would correspond to an equation of the form

$$D^{c_1} B_i^T(D) + D^{c_2} B_j^T(D) = 0(D)$$

where c_1 and c_2 are integers and $0(D)$ denotes a column vector of m polynomials whose coefficients are all zero. Assume, without loss of generality, that $c_1 \geq c_2$. In this case we have $D^{c_1-c_2} B_i^T(D) = B_j^T(D)$. The maximum possible degree of a polynomial in $B_j^T(D)$ is 1, and we must therefore have $c_1 - c_2 \in \{0,1\}$. We cannot have $c_1 - c_2 = 0$ since the parity-check matrix does not contain two equal columns. For $c_1 - c_2 = 1$, we get $B_i^T(D) = D^{-1} B_j^T(D)$ which cannot be true since $B_j^T(D)$ contains at least one polynomial with constant term 1.

To see that there exists a codeword of weight 3, note that the parity-check matrix in (1) contains the columns $B_1^T(D)$ and $B_3^T(D)$ for all $m \geq 1$, and that $(1+D) \cdot B_1^T(D) + 1 \cdot B_3^T(D) = (1+D)(1,0,\cdots,0)^T + (1+D,0,\cdots,0)^T = (0,\cdots,0)^T$.

Since $\mathbf{H}_m(D)$ in (1) has row degrees $e_1 = \cdots e_m = 1$, we have $e_{max} = 1$ and $d_{\mathcal{E}} = 1$ $(\gamma = 0)$. From Corollary 1, we now get $w_0 \geq 1$. Q.E.D.

B. Construction of codes with $d_{free} = 4$ and $w_0 \geq 1$

Let $q \leq 2^{2m}-1$ be a positive *odd* integer with binary representation

$$E_q = (1 b_1 b_2 \cdots b_{2m-1})$$

where b_{2m-1} is the most significant bit. The $2m$ bits of the vector E_q can be divided into m parts with one part containing 1 bit, $m-2$ parts containing 2 bits each, and one part containing 3 bits:

$$E_q = (1, b_1 b_2, \cdots, b_{2m-5} b_{2m-4}, b_{2m-3} b_{2m-2} b_{2m-1}).$$

As before, we let the bits in each part be the coefficients of a binary polynomial in the indeterminate D:

$$E_q(D) = (1, b_1 + b_2 D, \cdots, b_{2m-5} + b_{2m-4} D, b_{2m-3} + b_{2m-2} D + b_{2m-1} D^2).$$

For example, $m = 3$ and $q = 47$ result in the vectors

$$E_{47} = (1, 11, 101) \text{ and } E_{47}(D) = (1, 1 + D, 1 + D^2).$$

Theorem 3: The $m \times 2^{2m-1}$ canonical parity-check matrix

$$\mathbf{H}_m^*(D) = \left[\begin{array}{ccccc} E_1^T(D) & E_3^T(D) & E_5^T(D) & \cdots & E_{2^{2m}-1}^T(D) \end{array} \right] \tag{2}$$

defines a $(2^{2m-1}, 2^{2m-1} - m, m, 4)$ convolutional code with $w_0 \geq 1$ for all $m \geq 2$.

Sketch of proof: To see that the matrix in (2) is canonical, we observe that there exists an $m \times m$ minor

$$\left| \begin{array}{ccccc} 1 & 1 & \cdots & 1 & 1 \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ D & D & \cdots & D & 1+D \end{array} \right|$$

whose columns are given by $E_i^T(D)$, $i \in \{1 + 2^{2m-2}\} \bigcup \{1 + 2^j + 2^{2m-2} \mid j = 1, 3, 5, \cdots, 2m - 3\}$. If the first row is multiplied by D and added to the last row, then we obtain a triangular matrix with determinant 1. The $m \times m$ minor with columns $E_k^T(D)$, $k \in \{1\} \bigcup \{1 + 2^l \mid l = 2, 4, 6, \cdots, 2m - 4\} \bigcup \{1 + 2^{2m-1}\}$, has determinant D^m.

Since all columns of $\mathbf{H}_m^*(D)$ are distinct, the free distance is at least 3, and since the first row is the all-one vector and $e_1 = 0$, any nonzero edge label has even weight ($d_{\mathcal{E}} = 2$). Consequently, $d_{free} \geq 4$. On the other hand, $d_{free} \leq 4$ since $\sum_{i=1,3,5,7} E_i^T(D) = 0(D)$ for all $m \geq 2$.

The row degrees of $\mathbf{H}_m^*(D)$ are $e_1 = 0$, $e_2 = \cdots e_{m-1} = 1$, and $e_m = 2$. Since $e_{max} = 2$ and $d_{\mathcal{E}} = 2$, it follows from Theorem 1 that $w_0 \geq 1$. Q.E.D.

IV. A simple construction technique

It is well known that generator matrices for block codes may be combined to obtain a non-catastrophic generator matrix for a convolutional code (see [1] for a thorough analysis of this technique). We show how to combine parity-check matrices for block codes to obtain a canonical parity-check matrix for a convolutional code. Our simple construction technique makes it possible to build code classes that are not asymptotically catastrophic since the technique produces codes with $w_0 \geq d_{\mathcal{E}} \geq 1$ for all m. The technique is defined by two propositions whose proofs may be found in an appendix.

The *high order coefficient matrix* associated with a polynomial parity-check matrix $\mathbf{H}(D)$ is the $r \times n$ F-matrix whose ith row contains the coefficients of the maximum degree terms D^{e_i}. The matrix $\mathbf{H}(D)$ is canonical if the gcd of the $r \times r$ minors is 1 and the high order coefficient matrix has full rank r [10, Sec. 3].

Let $\mathbf{A}$ and $\mathbf{B}$ be two $r \times n$ F-matrices. It is assumed that all rows of $\mathbf{A}$ are nonzero while the γ, $0 \leq \gamma < r$, first rows of $\mathbf{B}$ contain only zeros.

Proposition 1: The $r \times n$ matrix $\mathbf{A} + D\mathbf{B}$ is a canonical parity-check matrix for an $(n, n-r, r-\gamma)$ convolutional code if (i) $\mathbf{A}$ has full rank r, (ii) the $r - \gamma$ nonzero rows of $\mathbf{B}$ are linearly independent, and (iii) the rows of $\mathbf{A}$ are linearly independent of the nonzero rows of $\mathbf{B}$.

Let the matrices $\mathbf{A}$ and $\mathbf{B}$ satisfy the requirements of Proposition 1 and assume that the code dimension $n - r$ is greater than the code degree $r - \gamma$, i.e., the code contains words with only one nonzero label.[3] Let d_A be the minimum distance of the $[n, n-r]$ block code with parity-check matrix $\mathbf{A}$, let d_B be the minimum distance of the $[n, n-r+\gamma]$ block code with parity-check matrix defined by the $r-\gamma$ nonzero rows of $\mathbf{B}$, and let d_{AB} be the minimum distance of the $[n, n-2r+\gamma]$ block code with parity-check matrix given by the r rows of $\mathbf{A}$ and the $r - \gamma$ nonzero rows of $\mathbf{B}$.

Proposition 2: The canonical parity-check matrix $\mathbf{A} + D\mathbf{B}$ defines an $(n, n-r, r-\gamma)$, $n > 2r-\gamma$, convolutional code with free distance $d_{free} = d_{AB}$ if $d_{AB} \leq d_A + d_B$, else $d_{free} \geq d_A + d_B$. Furthermore, $w_0 \geq d_{\mathcal{E}}$ where $d_{\mathcal{E}}$ is the minimum distance of the $[n, n-\gamma]$ embedded block code whose parity-check matrix is given by the γ first rows of $\mathbf{A}$.

---

[3] The words with only one nonzero label are represented by parallel edges starting and ending in the zero state of the code's state diagram. One of the edges has an all-zero label while the other edges have labels containing ones.

V. Classes of BCH convolutional codes

Rosenthal and York [13] have used binary BCH codes and techniques from linear systems theory to construct classes of binary convolutional codes with designed free distances. It is not known whether or not the classes are asymptotically catastrophic. In the following we use the construction technique introduced in the previous section to build two classes of convolutional codes from BCH codes. We show that the obtained canonical parity-check matrices define convolutional codes with minimum average weights w_0 close to the designed free distances.

A. Codes with $d_{free} \geq 2t+1$ and $w_0 \geq 2t-3$

There exists a $[2^m-1, 2^m-1-tm]$ binary BCH code with designed minimum distance $2t+1$ for $2 \leq t < 2^{\lceil m/2 \rceil - 1} + 1$ and $m \geq 3$ [9, Ch. 9, Cor. 8]. The binary parity-check matrix, denoted by $\mathbf{H}_{BCH}$, has tm rows and $2^m - 1$ columns. To obtain $\mathbf{H}_{BCH}$, we first form a $t \times (2^m-1)$ matrix containing powers of a primitive element α in the field $\mathrm{GF}(2^m)$ and then replace each entry in the matrix by a binary m-tuple (to save space, we will represent the m-tuples by powers of α). Let $\mathbf{A}_m$ be the $(t-1)m \times (2^m-1)$ binary matrix given by the $(t-1)m$ first rows of $\mathbf{H}_{BCH}$:

$$\mathbf{A}_m = \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^{2^m-2} \\ 1 & (\alpha^3) & (\alpha^3)^2 & \cdots & (\alpha^3)^{2^m-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & (\alpha^{2t-3}) & (\alpha^{2t-3})^2 & \cdots & (\alpha^{2t-3})^{2^m-2} \end{bmatrix}$$

and let $\mathbf{B}_m$ be the $(t-1)m \times (2^m-1)$ binary matrix

$$\mathbf{B}_m = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ 1 & (\alpha^{2t-1}) & (\alpha^{2t-1})^2 & \cdots & (\alpha^{2t-1})^{2^m-2} \end{bmatrix}$$

where the $(t-2)m$ first rows contain only zeros and the m last rows are the m last rows of $\mathbf{H}_{BCH}$.

Theorem 4: The canonical parity-check matrix $\mathbf{A}_m + D\mathbf{B}_m$ defines a $(2^m - 1, 2^m - 1 - (t-1)m, m)$ convolutional code with $d_{free} \geq 2t+1$ and $w_0 \geq 2t-3$ for any $2 \leq t < 2^{\lceil m/2 \rceil - 1} + 1$ and $m \geq 3$.

Proof: We use Proposition 1 to prove that the parity-check matrix is canonical. The rows of $\mathbf{A}_m$ and the nonzero rows of $\mathbf{B}_m$ together constitute a parity-check matrix for a binary BCH code of dimension $2^m - 1 - tm$ for any $2 \leq t < 2^{\lceil m/2 \rceil - 1} + 1$ and $m \geq 3$ [9, Ch. 9, Cor. 8]. Consequently, $\mathbf{A}_m$ has full rank, the m nonzero rows of $\mathbf{B}_m$ are linearly independent, and the rows of $\mathbf{A}_m$ are linearly independent of the nonzero rows of $\mathbf{B}_m$.

The matrix $\mathbf{A}_m$ defines a $[2^m - 1, 2^m - 1 - (t-1)m]$ BCH code with minimum distance $d_A \geq 2t-1$. If α^{2t-1} is a primitive element in $\mathrm{GF}(2^m)$, then the m last rows of $\mathbf{B}_m$ constitute a parity-check matrix for a $[2^m - 1, 2^m - 1 - m]$ Hamming code with minimum distance 3. When α^{2t-1} is not primitive, the submatrix contains two equal columns and it defines a code with minimum distance 2. Hence, we may assume that the code defined by $\mathbf{B}_m$ has minimum distance $d_B \geq 2$. Together, $\mathbf{A}_m$ and the nonzero rows of $\mathbf{B}_m$ define a $[2^m - 1, 2^m - 1 - tm]$ BCH code with distance $d_{AB} \geq 2t+1$. It now follows from Proposition 2 that the convolutional code has $d_{free} \geq 2t+1$.

The parity-check matrix for the embedded block code $\mathcal{E}$ consists of the $(t-2)m$ first rows of $\mathbf{A}_m$. For $t > 2$, $\mathcal{E}$ is a $[2^m - 1, 2^m - 1 - (t-2)m]$ BCH code with minimum distance $d_{\mathcal{E}} \geq 2t-3$, and for $t = 2$ the embedded block code is the trivial $[2^m - 1, 2^m - 1]$ block code with $d_{\mathcal{E}} = 1$. As a result, $w_0 \geq d_{\mathcal{E}} \geq 2t-3$ for $t \geq 2$. Q.E.D.

B. Codes with $d_{free} \geq 2t+2$ and $w_0 \geq 2t-2$

We can modify the matrices $\mathbf{A}_m$ and $\mathbf{B}_m$ to obtain two $((t-1)m+1) \times (2^m - 1)$ matrices

$$\mathbf{A}_m^* = \begin{bmatrix} 1 \;\; 1 \;\; \cdots \;\; 1 \\ \mathbf{A}_m \end{bmatrix}, \; \mathbf{B}_m^* = \begin{bmatrix} 0 \;\; 0 \;\; \cdots \;\; 0 \\ \mathbf{B}_m \end{bmatrix}$$

that add an extra parity-check to the labels of the convolutional code.

Theorem 5: The canonical parity-check matrix $\mathbf{A}_m^* + D\mathbf{B}_m^*$ defines a $(2^m - 1, 2^m - 2 - (t-1)m, m)$ convolutional code with $d_{free} \geq 2t+2$ and $w_0 \geq 2t-2$ for $2 \leq t < 2^{\lceil m/2 \rceil - 1} + 1$ and $m \geq 3$.

Proof: The tm binary rows of the parity-check matrix $\mathbf{H}_{BCH}$ are linearly independent of the all-one vector since otherwise all codewords of the BCH code have even weights. Consequently, the matrix $\mathbf{A}_m^*$ has full rank, the nonzero rows of $\mathbf{B}_m^*$ are linearly independent, and the rows of $\mathbf{A}_m^*$ are linearly independent of the nonzero rows of $\mathbf{B}_m^*$. From Proposition 1, the matrix $\mathbf{A}_m^* + D\mathbf{B}_m^*$ is canonical.

The matrix $\mathbf{A}_m^*$ is a parity-check matrix for a $[2^m-1, 2^m-2-(t-1)m]$ extended BCH code with minimum distance $d_A \geq 2t$, the m nonzero rows of $\mathbf{B}_m^*$ define a parity-check matrix for a code with minimum distance $d_B \geq 2$, and together the matrices $\mathbf{A}_m^*$ and $\mathbf{B}_m^*$ define parity-check equations for a $[2^m-1, 2^m-2-tm]$ extended BCH code with minimum distance $d_{AB} \geq 2t+2$. It follows from Proposition 2 that $d_{free} \geq 2t+2$. Since the embedded block code is a $[2^m-1, 2^m-2-(t-2)m]$ extended BCH code, with $d_{\mathcal{E}} \geq 2t-2$ for $t > 2$ and a $[2^m-1, 2^m-2]$ even weight code with $d_{\mathcal{E}} = 2$ for $t = 2$, we get $w_0 \geq d_{\mathcal{E}} \geq 2t-2$ for $t \geq 2$. Q.E.D.

VI. Codes with $w_0 \geq d_{\mathcal{E}}$ and increasing d_{free}

This section shows that there exist classes of low-rate codes whose free distances increase with m and whose minimum average weights $w_0 \geq d_{\mathcal{E}} \geq 1$ for all m.

A. Lower bound on d_{free}

Let an (n, k, m) convolutional code be defined by a $k \times n$ canonical generator matrix $\mathbf{G}(D)$ with row degrees $f_1, f_2, \cdots, f_k$, and let $f_{min} = \min\{f_i\}$. Furthermore, let $\mathbf{g}_j(D) = (g_{1j}(D) \cdots g_{kj}(D))^T$ denote the jth column of $\mathbf{G}(D)$. The *low order coefficient matrix*, $[\mathbf{G}(D)]_l$, is the $k \times n$ F-matrix that consists of the constant terms of the polynomials in $\mathbf{G}(D)$. The high order coefficient matrix, $[\mathbf{G}(D)]_h$, is defined as for the parity-check matrix, i.e., $[\mathbf{G}(D)]_h$ is the $k \times n$ F-matrix containing the coefficients of the maximum degree terms D^{f_i}. The matrices $[\mathbf{G}(D)]_l$ and $[\mathbf{G}(D)]_h$ define $[n, k]$ block codes over F with minimum distances denoted by d_l and d_h, respectively.

Theorem 6: Let $\mathbf{G}_m(D)$ be a canonical generator matrix with $f_{min} > 0$, and let γ be the dimension of the F-vector space $V^\perp$ consisting of the n-tuples $(v_1 \cdots v_n)$, $v_j \in F$, satisfying $\sum_{j=1}^{n} v_j \mathbf{g}_j(D) = 0(D)$. If $\mathbf{G}_m(D)$ generates an $(n, n-\gamma-m, m)$ convolutional code, $\mathcal{C}_m$, with $n \leq 2m+\gamma$, then $w_0 \geq d_{\mathcal{E}}$ and

$$d_{free} \geq d_l + d_h + (f_{min} - 1)d_{\mathcal{E}}. \tag{3}$$

It is important to notice that the lower bound on the free distance in (3) increases with m if the least row degree f_{min} increases with m. To obtain a fast increasing f_{min}, all row degrees of a general $\mathbf{G}_m(D)$ matrix (valid for all m) should be nearly equal, i.e., $f_i \approx m/(n-\gamma-m)$, $1 \leq i \leq n-\gamma-m$. In the following we prove Theorem 6 by establishing two lemmas.

Lemma 3: A canonical parity-check matrix for $\mathcal{C}_m$ has row degrees $e_1 = e_2 = \cdots = e_\gamma = 0$ and $e_{\gamma+1} = e_{\gamma+2} = \cdots = e_{\gamma+m} = 1$.

Proof: View the canonical generator matrix $\mathbf{G}_m(D)$ as a canonical parity-check matrix for the $(n, \gamma+m, m)$ dual code $\mathcal{C}_m^\perp$. The polynomial codewords of degree zero in $\mathcal{C}_m^\perp$ constitute the F-vector space $V^\perp$ of binary row vectors $(v_1 \cdots v_n)$ satisfying $\sum_{j=1}^{n} v_j \mathbf{g}_j(D) = 0(D)$.

Forney [2, Appendix] (see also [10, Cor. 4.2]) has shown that the dimension of $V^\perp$ is given by the expression

$$\dim V^\perp = \sum_{i:1 \geq e_i} (1 - e_i),$$

where e_i, $1 \leq i \leq \gamma + m$, are the row degrees of a canonical generator matrix for $\mathcal{C}_m^\perp$, or equivalently, a canonical parity-check matrix for $\mathcal{C}_m$. Since $\dim V^\perp = \gamma$, there must be exactly γ row degrees equal to zero, and since the remaining m row degrees must sum to m, we have $e_1 = \cdots = e_\gamma = 0$ and $e_{\gamma+1} = \cdots = e_{\gamma+m} = 1$. Q.E.D.

From Corollary 1 and Lemma 3 we have that the code $\mathcal{C}_m$ has $w_0 \geq d_{\mathcal{E}}$ since $e_{max} = 1$. The value of $d_{\mathcal{E}}$ can be obtained by first determining a canonical parity-check matrix from the generator matrix $\mathbf{G}_m(D)$ [10, Appendix B], [12, Ch. 2]. The γ first rows of the ordered parity-check matrix constitute a parity-check matrix for the $[n, n-\gamma]$ embedded block code $\mathcal{E}$. Standard techniques can then be used to determine the minimum distance $d_{\mathcal{E}}$ of this code.

If the dimension $n-\gamma-m$ of $\mathcal{C}_m$ is larger than the code degree m, then there exist parallel edges between the states in the code trellis. When the dimension is no larger than m, the trellis contains parallel edges only when $f_{min} = 0$ (see [12, pp. 74–76] for an example). In Theorem 6 we have assumed that $\mathcal{C}_m$ has $n - \gamma - m \leq m$ and $f_{min} > 0$.

A *fundamental path* of length τ in the trellis for $\mathcal{C}_m$ is a path that starts in the zero state in some depth p and returns to the zero state for the first time in depth $p + \tau$, $\tau \geq 2$. Denote by $W_{min}(\tau)$ the minimum weight of a fundamental path with length τ.

Lemma 4: The code $\mathcal{C}_m$ has free distance $d_{free} \geq d_l + d_h + (f_{min} - 1)d_{\mathcal{E}}$.

Proof: First, consider the zero state in some depth of the trellis for $\mathcal{C}_m$. The state has $2^{n-\gamma-m}$ outgoing edges whose labels are defined by the block code generated by $[\mathbf{G}(D)]_l$. Since the all-zero label is on the edge connecting the zero state in the current depth to the zero state in the next depth, the first edge on any fundamental path has weight at least equal to d_l. Similarly, the last edge on any fundamental path has weight no less than d_h.

Since, from Lemma 3, $e_{max} = 1$, it follows from Corollary 1 that the only all-zero labels in the trellis for $\mathcal{C}_m$ are the labels on the edges between the zero states. All other edges in the trellis have weights no less than than $d_{\mathcal{E}}$. As a result, $W_{min}(\tau) \geq d_l + d_h + (\tau - 2)d_{\mathcal{E}}$.

Next, observe that $d_{free} = \min_\tau\{W_{min}(\tau)\}$. Consider the direct-form (controller canonical form) realization of $\mathbf{G}_m(D)$ containing $n - \gamma - m$ shift registers where the ith register consists of f_i delay elements. Assume that the shift registers contain only zeros, i.e., the encoder is in the zero state. The shortest shift register has length f_{min}. Hence, the shortest fundamental path is generated by shifting a "1" followed by f_{min} "0"s into the shortest register while shifting only zeros into the other registers. Since the generated path has length $\tau = f_{min} + 1$, it follows that $d_{free} \geq d_l + d_h + (f_{min} - 1)d_{\mathcal{E}}$. Q.E.D.

B. Construction of codes with $d_{free} \geq 3m$ and $w_0 \geq 1$

It is a difficult open problem to determine a general expression for $\mathbf{G}_m(D)$ that defines codes with *maximum* obtainable free distances. In this subsection we give an example of a simple construction for $(m+1, 1, m)$ codes with $d_{free} \geq 3m$. The codes have systematic canonical generator matrices. We use the following corollary obtained by setting $\gamma = 0$ in Theorem 6.

Corollary 2: Let $\mathcal{C}_m$ be an $(n, n-m, m)$ convolutional code with $n \leq 2m$ defined by a canonical generator matrix $\mathbf{G}_m(D)$ with $f_{min} > 0$. If the n columns of $\mathbf{G}_m(D)$ are linearly independent over F, then $w_0 \geq 1$ and $d_{free} \geq d_l + d_h + f_{min} - 1$.

Let $g_j(D) = 1 + g_j^1 D + g_j^2 D^2 + \cdots + g_j^m D^m$, $g_j^i \in F$, denote the jth polynomial in a $1 \times (m+1)$ generator matrix $\mathbf{G}_m(D) = [g_j(D)]$. It is convenient to represent the coefficients of each polynomial by a column vector $\mathbf{g}_j = (1, g_j^1, g_j^2, \cdots, g_j^m)^T$. Let $\mathbf{G}_m(D)$ be a generator matrix whose polynomials have coefficients given by a $(m+1) \times (m+1)$ matrix of the form

$$\begin{bmatrix} \mathbf{g}_1 & \cdots & \mathbf{g}_{m+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 0 & 1 & ? & ? & \cdots & ? & ? \\ 0 & 0 & 0 & 1 & ? & \cdots & ? & ? \\ 0 & 0 & 0 & 0 & 1 & \cdots & ? & ? \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & ? \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & \cdots & 1 & 1 \end{bmatrix}. \tag{4}$$

Here the questionmarks denote random bits. Note that $g_1(D) = 1$ since only the first element in the first column of (4) is equal to one.

The systematic matrix $\mathbf{G}_m(D)$ defines an $(m+1, 1, m)$ convolutional code with $d_{free} \geq 3m$ and $w_0 \geq 1$ for all $m \geq 1$. To prove this, we only need to establish the requirements of Corollary 2. Clearly, $n = m + 1 \leq 2m$ for $m \geq 1$ and $f_{min} = m > 0$. Furthermore, the generator matrix is canonical since the gcd of the $m+1$ polynomials is 1 and their greatest degree is m. To see that the $m+1$ polynomials are linearly independent over F, we study the matrix of coefficients in (4). If the first row is added to the last row, then a cofactor expansion of the modified last row reveals that the determinant of the matrix is 1.

The low order coefficient matrix $[\mathbf{G}(D)]_l = [11 \cdots 1]$ and the high order coefficient matrix $[\mathbf{G}(D)]_h = [011 \cdots 1]$. Since the the block length is $m+1$, the minimum distances are equal to $d_l = m+1$ and $d_h = m$. As a result, $d_{free} \geq m + 1 + m + m - 1 = 3m$.

VII. Discussion and summary

Forney [2] showed that the row degrees are fundamental parameters of convolutional codes, and he suggested that the row degrees of the dual codes should be studied in more detail. We have followed Forney's suggestion and studied the row degrees of canonical parity-check matrices. Code classes that are not asymptotically catastrophic were constructed by determining general expressions for canonical parity-check matrices with small row degrees.

An $(n, n-r, m, d_{free})$ code is said to be *length optimal* if it has maximum length n, and hence maximum rate, for the given free distance d_{free}, redundancy r, and code degree m [10, Sec. 7]. It follows from results in [14] (see also [10, Sec. 7]) that the $(2^{2m}-2^m, 2^{2m}-2^m-m, m, 3)$ codes and the $(2^{2m-1}, 2^{2m-1}-m, m, 4)$ codes constructed in Section III are length optimal. Since the new classes are not asymptotically catastrophic, they are preferable to the asymptotically catastrophic classes of Wyner-Ash codes [10, Sec. 7], [5].

It is not always possible to obtain a code with both maximum d_{free} and maximum w_0 [7], [3]. The constructed BCH convolutional codes with $d_{free} \geq 2t+1$ $(2t+2)$ all have large $w_0 \geq 2t-3$ $(2t-2)$. It may well be that codes with larger d_{free} can be constructed by reducing w_0. Clearly, more work is needed to determine a good tradeoff between d_{free} and w_0.

In Section VI we proved that there exist classes of low-rate codes with increasing d_{free} and $w_0 \geq 1$, and we constructed one such class. It is an interesting open problem to construct more and better classes.

Appendix

This appendix contains the proofs of the propositions given in Section IV. The proofs are based on observations made in [11, pp. 403–404] and [8].

Proof of Proposition 1: View the parity-check matrix as a generator matrix for the (n, r) dual code and assume that the matrix is used to encode a semi-infinite information sequence that is different from the all-zero sequence in infinitely many positions. It follows from (i)–(iii) that the generated codeword contains infinitely many nonzero symbols. In other words, the matrix

$\mathbf{A} + D\mathbf{B}$ is non-catastrophic, i.e., the gcd of its $r \times r$ minors is equal to D^l for some $l \in \{0, 1, 2, \cdots\}$.

Since, from (i), $\mathbf{A}$ has full rank, there exists a minor with constant term 1, and l must therefore be equal to zero. Consequently, the gcd of the $r \times r$ minors is 1. The high order coefficient matrix consists of the γ first rows of $\mathbf{A}$ and the $r - \gamma$ last rows of $\mathbf{B}$. It follows from (i)–(iii) that the matrix has full rank. Hence, the matrix $\mathbf{A} + D\mathbf{B}$ is canonical. Q.E.D.

Proof of Proposition 2: A semi-infinite binary parity-check matrix for the code is given by

$$\mathbf{H}_\infty = \begin{bmatrix} \mathbf{A} & & & \\ \mathbf{B} & \mathbf{A} & & \\ & \mathbf{B} & \mathbf{A} & \\ & & & \ddots \end{bmatrix},$$

and a semi-infinite codeword that "starts" at time $t = 0$ is represented by a sequence $\mathbf{v} = (\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \cdots)$ where $\mathbf{v}_t = (v_t^1 \cdots v_t^n)$ is a binary n-tuple (or label). Any codeword $\mathbf{v}$ must satisfy $\mathbf{H}_\infty \mathbf{v}^T = \mathbf{0}$.

Assume, without loss of generality, that $\mathbf{v}_0 \neq (0 \cdots 0)$. If $\mathbf{v}_t = (0 \cdots 0)$ for all $t > 0$, then $\mathbf{v}_0$ must have weight at least d_{AB} since the nonzero rows of the submatrix

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}$$

is the parity-check matrix for a block code with minimum distance d_{AB}, i.e., only d_{AB} or more columns in the submatrix can add to zero.

Consider a codeword $\mathbf{v}$ that contains at least one other n-tuple $\mathbf{v}_i \neq (0 \cdots 0)$ for some $i > 0$. In this case, $\mathbf{v}_0$ has weight no less than d_A since $\mathbf{A}$ is the parity-check matrix for a block code with minimum distance d_A, and $\mathbf{v}_i$ must have weight at least d_B since $\mathbf{B}$ is the parity-check matrix for a block code with minimum distance d_B.

The row degrees of $\mathbf{A} + D\mathbf{B}$ are $e_1 = \cdots e_\gamma = 0$, and $e_{\gamma+1} = \cdots e_r = 1$. Hence, $e_{max} = 1$ and it follows from Corollary 1 that $w_0 \geq d_{\mathcal{E}}$. Q.E.D.

References

[1] K. Abdel-Ghaffar, R. J. McEliece, and G. Solomon, "Some partial-unit-memory convolutional codes," TDA Progress Report 42-107, Jet Propulsion Lab., Pasadena, CA, pp. 57–72, Nov. 15, 1991.

[2] G. D. Forney, Jr., "Structural analysis of convolutional codes via dual codes," *IEEE Trans. Inform. Theory,* vol. IT-19, pp. 512–518, July 1973.

[3] F. Hemmati and D. J. Costello, Jr., "Asymptotically catastrophic convolutional codes," *IEEE Trans. Inform. Theory,* vol. IT-26, pp. 298–304, May 1980.

[4] K. J. Hole, "A note on asymptotically catastrophic convolutional codes of rate $(n-1)/n$," *IEEE Trans. Commun.*, vol. 45, pp. 1014–1016, Sept. 1997.

[5] M. F. Hole and K. J. Hole, "Tight bounds on the minimum average weight per branch for rate $(n-1)/n$ convolutional codes," *IEEE Trans. Inform. Theory,* vol. 43, pp. 1301–1305, July 1997.

[6] K. J. Hole and Ø. Ytrehus, "Cosets of convolutional codes with least possible maximum zero- and one-run lengths," *IEEE Trans. Inform. Theory,* vol. 44, pp. 423–431, Jan. 1998.

[7] G. K. Huth and C. L. Weber, "Minimum weight convolutional codewords of finite length," *IEEE Trans. Inform. Theory,* vol. IT-22, pp. 243–246, March 1976.

[8] G. S. Lauer, "Some optimal partial-unit-memory codes," *IEEE Trans. Inform. Theory,* vol. IT-25, pp. 240–243, March 1979.

[9] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* North-Holland Publishing Company, Amsterdam, 1977.

[10] R. J. McEliece, "The algebraic theory of convolutional codes," chapter in the *Handbook of Coding Theory.* Available WWW: http://www.systems.caltech.edu/EE/Faculty/rjm/

[11] W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes*, 2nd edition. MIT Press, Cambridge, MA, 1972.

[12] P. Piret, *Convolutional Codes.* MIT Press, Cambridge, MA, 1988.

[13] J. Rosenthal and E. V. York, "A construction of binary BCH convolutional codes," in *Proc. 1997 IEEE Int. Symp. Info. Theory,* Ulm, Germany, June 29–July 4, 1997, p. 291.

[14] Ø. Ytrehus, "A note on high rate binary convolutional codes," Tech. Report no. 68, Department of Informatics, University of Bergen, Aug. 11, 1992.