# REPORTS
# IN
# INFORMATICS

ISSN 0333-3590

## Upper-Bounding the Minimum Average Cycle Weight of Classes of Convolutional Codes

Eirik Rosnes

REPORT NO 265 January 2004

*Department of Informatics*

UNIVERSITY OF BERGEN

*Bergen, Norway*

This report has URL http://www.ii.uib.no/publikasjoner/texrap/ps/2004-265.ps

Reports in Informatics from Department of Informatics, University of Bergen, Norway, is available at
http://www.ii.uib.no/publikasjoner/texrap/.

Requests for paper copies of this report can be sent to:

Department of Informatics, University of Bergen, Høyteknologisenteret,
P.O. Box 7800, N-5020 Bergen, Norway

# Upper-Bounding the Minimum Average Cycle Weight of Classes of Convolutional Codes

Eirik Rosnes*
Department of Informatics
University of Bergen
Norway
eirik@ii.uib.no

**Abstract**

An upper bound on the *minimum average cycle weight per branch* over all cycles in a minimal state diagram of a convolutional code, excluding the all-zero cycle around the all-zero state, is derived. We generalize the approach by Hole and Hole and derive bounds for large code classes within the class of rate $(n-r)/n$, $r \geq 2$, codes. The results of an exhaustive computer search indicate that the derived upper bound is almost tight as the code degree, or the overall constraint length, grows.

## 1 Introduction

Let $w_0$ denote the *minimum average cycle weight per branch* over all cycles in a minimal state diagram of a convolutional code, excluding the all-zero cycle around the all-zero state. Codes with low $w_0$ contain long codewords of low weight. These codes are susceptible to long error events when used with either maximum likelihood (Viterbi) or sequential decoding [1], [2]. Further, the *active distances* for convolutional codes [3], describing which error patterns are guaranteed to be corrected under a maximum likelihood decoding assumption, are lower bounded by a linearly increasing function with *slope* $w_0$. In this sense, $w_0$ determines the code's error correcting capability.

When working with concatenated codes, e.g., serial concatenated convolutional codes, decoded using iterative decoding schemes, simulation results indicate that outer codes with large $w_0$ and small degree compare favorably with other choices for the outer code [4].

Huth and Weber [5] derived a general upper bound on $w_0$ for rate $1/n$ convolutional codes. Recently, Jordan *et al.* [6] generalized the bound to rate $(n-r)/n$, $r \geq 1$, codes. The new upper bound applies to codes having a canonical generator matrix containing a delay-free $(n-r)\times(n-r)$ minor of degree equal to the code degree. A general lower bound on $w_0$ can be found in Hole and Hole [2].

In [1], Hemmati and Costello derived a tight upper bound on $w_0$ for a special class of rate $1/2$ convolutional codes. By generalizing the approach in [1], the authors of [2] showed that the bound in [1] applies to a particular class of rate $(n-1)/n$, $n \geq 2$, codes as well.

This paper is organized as follows: In Section 2, a convenient matrix notation for convolutional codes is described. A new upper bound on the minimum average cycle weight per branch is then derived in Section 3. The bound applies to classes of rate $(n-r)/n$, $r \geq 1$, codes, and is based on the argument given by Hole and Hole in [2] for rate $(n-1)/n$ codes. In Section 4, the results of an exhaustive computer search are given, verifying that the derived upper bound is almost tight as the code degree grows. Conclusions are presented in Section 5.

*This work was supported by Nera Research and the Norwegian Research Council (NFR) Grant 137783/221.

## 2 Convolutional code preliminaries

We use some of the notation and definitions introduced in [7], as well as a convenient matrix notation established in [8], [9], and [10].

A linear $(n, n-r, \nu)$ convolutional code $\mathcal{C}$ is an $(n-r)$-dimensional subspace in an $n$-dimensional vector space $F(D)^n$, where $F(D)$ is the field of rational functions in the indeterminate $D$ over the field $F$ [7, Def. 2.4, p. 1073]. The code degree, or the overall constraint length, is denoted by $\nu$. In this paper the convolutional code symbols are taken from the binary field $F = GF(2)$. A convolutional code can be defined by an $r \times n$ polynomial parity check matrix $\mathbf{H}(D)$. We assume in general a canonical parity check matrix [7, Def. 3.5, p. 1080]. Let the $j$th polynomial in the $i$th row of $\mathbf{H}(D)$ be denoted by $h_j^{(i)}(D) = h_{j,0}^{(i)} + h_{j,1}^{(i)}D + \cdots + h_{j,\nu_i}^{(i)}D^{\nu_i} \in F[D]$, where $F[D]$ is the ring of all polynomials in $D$ with coefficients in $F$. The maximum degree of the polynomials in the $i$th row is the $i$th row degree, denoted by $\nu_i$. Every canonical parity check matrix of a given convolutional code has the same set of row degrees with $\nu = \sum_{i=1}^{r} \nu_i$ [7, Theorem 3.10, p. 1081].

The coefficients of $h_j^{(i)}(D)$ define a column vector $\mathbf{h}_j^{(i)}$ with $h_{j,0}^{(i)}$ as its topmost element. The $n$ polynomials in the $i$th row of the parity check matrix $\mathbf{H}(D)$ give rise to a $(\nu_i + 1) \times n$ matrix $\mathbf{H}^{(i)} = (\mathbf{h}_1^{(i)}, \ldots, \mathbf{h}_n^{(i)})$ over the field $F$. Furthermore, let $\mathbf{H}$, referred to as a *combined parity check matrix*, be defined as

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}^{(1)} \\ \mathbf{H}^{(2)} \\ \vdots \\ \mathbf{H}^{(r)} \end{pmatrix} = (\mathbf{h}_1, \ldots, \mathbf{h}_n) . \tag{1}$$

Note that the combined parity check matrix is a $(\nu + r) \times n$ matrix over the field $F$. The $D$-transform of $\mathbf{H}$ is $\mathbf{H}(D)$, where the polynomial in the $i$th row and $j$th column is the previously defined polynomial $h_j^{(i)}(D)$.

Let $\mathbf{x} = (x_a x_{a+1} \cdots x_b)^T$ be a finite dimensional column vector, where $(\cdot)^T$ denotes the transpose of its argument. The $l$th shift of $\mathbf{x}$, denoted by $\mathbf{x}^{\overset{l}{\leftarrow}}$, is defined as $\mathbf{x}^{\overset{l}{\leftarrow}} = (x_{a+l} \cdots x_b 0 \cdots 0)^T$, $l \geq 0$. The last $l$ coordinates in $\mathbf{x}^{\overset{l}{\leftarrow}}$ are equal to zero.

A codeword in $\mathcal{C}$ is a semi-infinite sequence of $n$-tuples, or column vectors. An arbitrary codeword sequence is denoted as $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2, \ldots)$, where $\mathbf{v}_t$, $t > 0$, is an $n$-tuple (or *label*) given as $\mathbf{v}_t = (v_t^1 \cdots v_t^n)^T$. We now define the $i$th *syndrome vector* $\mathbf{s}_t^{(i)} = (s_{t,0}^{(i)} \cdots s_{t,\nu_i}^{(i)})^T$ of dimension $(\nu_i + 1)$ at time $t$, recursively as follows ($t > 0$, $1 \leq i \leq r$):

$$\mathbf{s}_t^{(i)} = (\mathbf{s}_{t-1}^{(i)})^{\overset{1}{\leftarrow}} + \mathbf{H}^{(i)}\mathbf{v}_t, \tag{2}$$

with $\mathbf{s}_0^{(i)}$ equal to a fixed vector, e.g., the all-zero vector. Furthermore, we define the $i$th *syndrome sequence* as $(s_{0,0}^{(i)} s_{1,0}^{(i)} \cdots)$ where $s_{t,0}^{(i)}$, $t \geq 0$, is the zeroth element in the $i$th syndrome vector $\mathbf{s}_t^{(i)}$. The code consists of all semi-infinite sequences $\mathbf{v}$ such that all the corresponding syndrome sequences are equal to the all-zero sequence.

A compact form of (2) is obtained using the combined parity check matrix. The *combined syndrome vector* at time $t$, $\mathbf{s}_t = ((\mathbf{s}_t^{(1)})^T, \ldots, (\mathbf{s}_t^{(r)})^T)^T = (s_{t,0}^{(1)} \cdots s_{t,\nu_1}^{(1)}, \ldots, s_{t,0}^{(r)} \cdots s_{t,\nu_r}^{(r)})^T$ is given by

$$\mathbf{s}_t = (\mathbf{s}_{t-1})^{\overset{1}{\leftarrow}} + \mathbf{H}\mathbf{v}_t, \tag{3}$$

where $(\mathbf{s}_{t-1})^{\overset{1}{\leftarrow}} = (((\mathbf{s}_{t-1}^{(1)})^{\overset{1}{\leftarrow}})^T, \ldots, ((\mathbf{s}_{t-1}^{(r)})^{\overset{1}{\leftarrow}})^T)^T$, i.e., the shift operator should be applied to each component individually. The combined syndrome vectors in (3) have dimension $\nu + r$. Assuming that the syndrome vectors are computed from codewords, the set of possible combined syndrome vectors $\mathcal{V}_t = \{\mathbf{s}_t : (\mathbf{v}_1, \mathbf{v}_2, \ldots) \in \mathcal{C}\}$ is a vector space of dimension $\nu$ (after an initial transient).

**Example 1** *Consider a* $(4, 2, 5)$ *binary convolutional code with free distance* 6 *defined by the polynomial canonical parity check matrix*

$$\mathbf{H}(D) = \begin{pmatrix} 1 + D + D^2 & 1 + D^2 & 1 + D^2 & 1 + D^2 \\ 0 & D & D^2 & 1 + D + D^3 \end{pmatrix} . \tag{4}$$

*The corresponding binary combined parity check matrix is*

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}^{(1)} \\ \mathbf{H}^{(2)} \end{pmatrix} = \left( \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right), \tag{5}$$

*where the solid line separates the two component matrices* $\mathbf{H}^{(1)}$ *and* $\mathbf{H}^{(2)}$*. Further, let* $\mathbf{s}_{t-1} = (010, 0110)^T$*,* $\mathbf{v}_t = (0001)^T$*, and* $\mathbf{v}_{t+1} = (0000)^T$*. From (3) we have*

$$\mathbf{s}_t = (001, 0001)^T \quad \textit{and} \quad \mathbf{s}_{t+1} = (010, 0010)^T. \tag{6}$$

*The binary column vectors* $\mathbf{v}_t$ *and* $\mathbf{v}_{t+1}$ *are contained in a codeword since* $s_{t,0}^{(1)} = s_{t,0}^{(2)} = s_{t+1,0}^{(1)} = s_{t+1,0}^{(2)} = 0$*. Note that when the label is the all-zero label, the combined syndrome vector at the next time instant is obtained by simply shifting each component of the previous combined syndrome vector individually.*

A code $\mathcal{C}$ may be represented by a state diagram, where each state represents a combined syndrome vector. Obviously, the number of states in the state diagram is equal to $2^\nu$. The set of possible transitions between states is determined by the equation in (3). An important observation is that a transition from state $\mathbf{s}$ on a branch with weight zero will lead to the state $\mathbf{s}^{\overset{1}{\leftarrow}}$. (The weight of a branch is the Hamming weight of the label on the branch.)

A path of length $p$ in a state diagram consists of $p$ consecutive branches. A *cycle* is a path returning back to the state where it started, in which the intermediate states are distinct and different from the starting state.

**Lemma 1** *Let* $\mathbf{H}$ *denote a combined parity check matrix defining a code* $\mathcal{C}$ *with free distance* $d_{\text{free}}$*. Then the following holds:*

1. $d_{\text{free}} = 1$ *if and only if* $\mathbf{H}$ *contains the all-zero combined column.*
2. *If* $\mathbf{H}$ *contains two equal combined columns, then* $d_{\text{free}} = 2$*.*

The proof of Lemma 1 is trivial, and is omitted for brevity. In this work only codes with $d_{\text{free}} \geq 3$ are considered. Thus, using Lemma 1, we will assume without loss of generality that every combined parity check matrix has distinct combined columns different from the all-zero combined column.

Two combined parity check matrices are said to be *d-equivalent* if the two corresponding codes have the same free distance $d_{\text{free}} \geq 2$.

In [11], Paaske described some equivalence relations on the set of polynomial parity check matrices for $r = 1$. In fact only parity check matrices where all polynomials have constant terms different from zero need to be considered. In [8], Ytrehus generalized some of these relations to $r \geq 1$. The following lemma was proved in [8]:

**Lemma 2** *Every combined parity check matrix* $\mathbf{H}$*, of constraint length* $\nu$*, is d-equivalent to some combined parity check matrix* $\tilde{\mathbf{H}}$*, of constraint length* $\leq \nu$*, in which, for every column index* $j$*,* $1 \leq j \leq n$*, there is some row index* $i$*,* $1 \leq i \leq r$*, such that* $\tilde{h}_{j,0}^{(i)} = 1$*.*

Lemmas 1 and 2 imply the following.

**Corollary 1** *Without loss of generality, only combined parity check matrices* $\mathbf{H}$ *with distinct combined columns containing no combined column* $\mathbf{h} = (h_0^{(1)} \cdots h_{\nu_1}^{(1)}, h_0^{(2)} \cdots h_{\nu_2}^{(2)}, \ldots, h_0^{(r)} \cdots h_{\nu_r}^{(r)})^T$ *of the form*

$$h_0^{(1)} = h_0^{(2)} = \cdots = h_0^{(r)} = 0 \tag{7}$$

*need to be considered when looking at codes with free distance $\geq 3$, where the $i$th row degree of* $\mathbf{H}$ *is denoted by* $\nu_i$, $1 \leq i \leq r$.

Note that Corollary 1 reduces to the equivalence relation proposed by Paaske in [11] for $r = 1$.

# 3 Upper bound on the minimum average cycle weight

A *zero segment* of length $p$ is a path consisting of $p$ branches of Hamming weight zero such that any extension of the path in either direction has nonzero weight. The all-zero cycle around the all-zero state is excluded. The zero segments in a state diagram from a canonical parity check matrix are disjoint, finite length paths whose states are different from the all-zero state [12]. To obtain an upper bound on $w_0$ we consider zero segments and cycles containing zero segments.

**Definition 1** *Let* $\mathcal{H}_n^{(r)}(\nu_1, \ldots, \nu_r)$, $r \geq 1$, $\nu_i \geq 2$ *for all* $i$, $1 \leq i \leq r$, *and* $n \geq 2r$, *be the class of all canonical* $(\nu + r) \times n$ *combined parity check matrices with ordered row degrees, i.e., with* $\nu_1 \leq \cdots \leq \nu_r$, *satisfying the following constraints: There exists an ordered set* $J = \{1 \leq j_1 < \cdots < j_{2r} \leq n\}$ *of cardinality* $2r$ *such that*

$$h_{j,0}^{(i)} = h_{j,\nu_i}^{(i)}, \quad 1 \leq i \leq r, \quad j \in J, \tag{8}$$

*and*

$$\mathcal{V}_{\text{low}} = \left\{ (h_{j,0}^{(1)} h_{j,1}^{(1)}, \ldots, h_{j,0}^{(r)} h_{j,1}^{(r)})^T : j \in J \right\} \tag{9}$$

$$\mathcal{V}_{\text{high}} = \left\{ (h_{j,\nu_1 - 1}^{(1)} h_{j,\nu_1}^{(1)}, \ldots, h_{j,\nu_r - 1}^{(r)} h_{j,\nu_r}^{(r)})^T : j \in J \right\} \tag{10}$$

*are sets of* linearly independent *column vectors spanning* $F^{2r}$. *In addition, only combined parity check matrices with combined columns that do not satisfy (7) are considered.*

**Example 2** *Choose* $r = 2$, $\nu_1 = 2$, *and* $\nu_2 = 3$. *The binary combined parity check matrix in (5) satisfies the constraint in (8) with* $J = \{1, 2, 3, 4\}$. *Further,*

$$\mathcal{V}_{\text{low}} = \{(11, 00)^T, (10, 01)^T, (10, 00)^T, (10, 11)^T\} \tag{11}$$

$$\mathcal{V}_{\text{high}} = \{(11, 00)^T, (01, 00)^T, (01, 10)^T, (01, 01)^T\} \tag{12}$$

*are sets of linearly independent column vectors. Thus, the canonical parity check matrix (4) from the previous example is contained in the class* $\mathcal{H}_4^{(2)}(2, 3)$.

**Definition 2** *Let* $B = \{(b_0 b_1 \cdots b_{n-1})^T \in F^n : b_l = 0, l \notin J$ *and* $b_l \in \{0, 1\}, l \in J\}$ *be a set of branch labels.*

In Fig. 1, a zero segment with ending state $\mathbf{t}$ is given. The zero segment is connected to an internal state $\mathbf{q}$ of some zero segment via a branch with label from the set $B$ of branch labels. This figure is used repeatedly in the proofs of the lemmas below, and is included to make the understanding easier.

**Lemma 3** *For codes in the class* $\mathcal{H}_n^{(r)}(\nu_1, \ldots, \nu_r)$, *let* $\mathbf{t}$ *be the ending state of some zero segment. There exists a unique label* $\mathbf{b} \in B$ *on a branch from* $\mathbf{t}$ *to the starting state of some zero segment.*

*Proof:* The proof for the lemma is 3-fold. First we prove that there exists a label $\mathbf{b} \in B$ on a transition from the ending state $\mathbf{t}$ of some zero segment to a valid state $\mathbf{q}$ such that $\mathbf{b}$ may be followed by the all-zero label $\mathbf{o} = (0 \cdots 0)^T$. The starting state of the zero segment is denoted by $\mathbf{s}$ and its length by $p$. Secondly, we prove that $\mathbf{q}$ cannot be the all-zero state, i.e., the zero segment cannot be connected to the self-loop of weight zero around the all-zero state. In Fig. 1, $\mathbf{q}$ is depicted as an internal state of some zero segment. In the last part of the proof we show that this cannot be the case, i.e., $\mathbf{q}$ has to be the starting state of some zero segment.

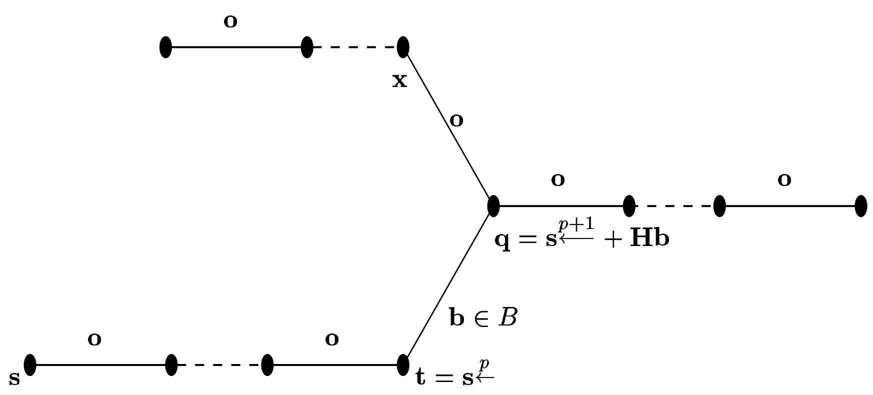


Fig. 1. A zero segment with starting state $\mathbf{s}$, length $p$, and ending state $\mathbf{s}^{\xleftarrow{p}}$. The zero segment is connected to an internal state $\mathbf{q}$ of some zero segment via a branch with label from $B$.

Part I: Consider the expression for $\mathbf{q}$ given by

$$\mathbf{q} = (q_0^{(1)} \cdots q_{\nu_1}^{(1)}, \ldots, q_0^{(r)} \cdots q_{\nu_r}^{(r)})^T = \mathbf{t}^{\xleftarrow{1}} + \mathbf{Hb} = (\mathbf{s}^{\xleftarrow{p}})^{\xleftarrow{1}} + \mathbf{Hb} = \mathbf{s}^{\xleftarrow{p+1}} + \mathbf{Hb}, \tag{13}$$

where

$$\mathbf{s}^{\xleftarrow{p+1}} = (s_{p+1}^{(1)} \cdots s_{\nu_1}^{(1)} 0 \cdots 0, \ldots, s_{p+1}^{(r)} \cdots s_{\nu_r}^{(r)} 0 \cdots 0)^T. \tag{14}$$

Since $\mathbf{s}$ is the starting state of a zero segment of length $p$, there exists at least one $j$, $1 \leq j \leq r$, such that $s_{p+1}^{(j)} = 1$. Further, from (9) there exists a unique nonzero label $\mathbf{b} \in B$ such that $q_0^{(1)} = q_1^{(1)} = \cdots = q_0^{(r)} = q_1^{(r)} = 0$, i.e., state $\mathbf{q}$ has an outgoing branch of weight zero since $\mathbf{q}^{\xleftarrow{1}}$ is a valid state.

Part II: At least the last two elements in every component of $\mathbf{s}^{\xleftarrow{p+1}}$ are zero. Further, we know that there exists at least one $j$, $1 \leq j \leq r$, such that $s_{p+1}^{(j)} = 1$, i.e., the first two elements cannot be both zero in every component of $\mathbf{s}^{\xleftarrow{p+1}}$. From (9), the selected label from $B$ in (13) is not equal to the all-zero label $\mathbf{o}$. Consequently, from (10) there will exist a $j'$, $1 \leq j' \leq r$, such that at least one of the last two elements in the $j'$ component of $\mathbf{q}$ is nonzero, i.e., $\mathbf{q}$ is not equal to the all-zero state.

Part III: We need to show that $\mathbf{q}$ is the starting state of some zero segment, and not an intermediate state. We will prove the statement by showing that the assumption of $\mathbf{q}$ being an intermediate state will lead to a contradiction. If $\mathbf{q}$ is an intermediate state of some zero segment, then the state, denoted by $\mathbf{x}$, leading to $\mathbf{q}$ via a branch with the all-zero label will be of the form $\mathbf{x} = (00x_2^{(1)} \cdots x_{\nu_1}^{(1)}, \ldots, 00x_2^{(r)} \cdots x_{\nu_r}^{(r)})^T$. An outgoing branch of weight zero will lead to the state $\mathbf{x}^{\xleftarrow{1}} = (0x_2^{(1)} \cdots x_{\nu_1}^{(1)} 0, \ldots, 0x_2^{(r)} \cdots x_{\nu_r}^{(r)} 0)^T$. We know that there exists at least one $j$, $1 \leq j \leq r$, such that the first element in the $j$th component of $\mathbf{s}^{\xleftarrow{p+1}}$ is equal to 1. The label $\mathbf{b}$ in (13) is chosen such that $q_0^{(j)} = 0$. The requirement in (8) implies that $q_{\nu_j}^{(j)} = 1$, and a contradiction follows by comparison with $\mathbf{x}^{\xleftarrow{1}}$. □

**Lemma 4** *For codes in the class $\mathcal{H}_n^{(r)}(\nu_1, \ldots, \nu_r)$, let $\mathbf{s}^E$ be the ending state of a particular zero segment, and let $\mathbf{s}_1^S$ and $\mathbf{s}_2^S$ be the starting states of two distinct zero segments. If there exists a label in $B$ on a transition from $\mathbf{s}^E$ to $\mathbf{s}_1^S$, then there is no label in $B$ on a transition from $\mathbf{s}^E$ to $\mathbf{s}_2^S$ for any $\mathbf{s}_2^S$.*

*Proof:* The situation is depicted in Fig. 2, and the statement follows immediately from the fact that every ending state of a zero segment after a left shift has specific values in the first two elements in each component. The purpose of a label in $B$ is to "zero-out" these elements. From (9), there exists only one label in $B$ for this purpose. □

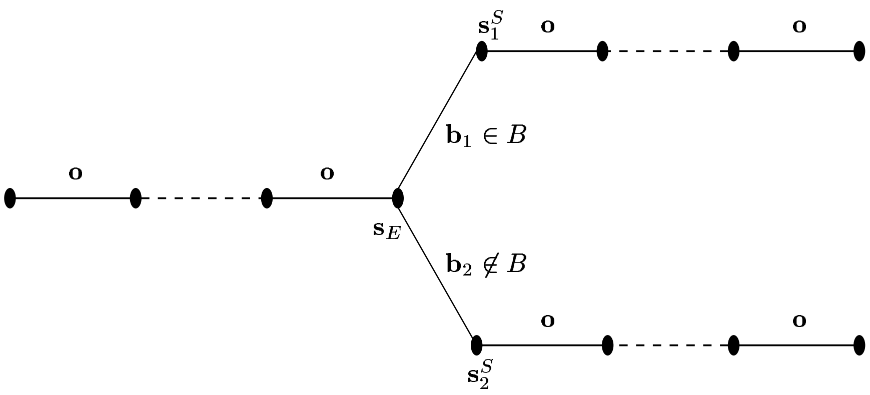


Fig. 2. A zero segment with ending state $\mathbf{s}_E$ connected via single branches to two zero segments with starting states $\mathbf{s}_1^S$ and $\mathbf{s}_2^S$, respectively. Only one of the branch labels can be taken from the set $B$.

**Lemma 5** *For codes in the class $\mathcal{H}_n^{(r)}(\nu_1,\ldots,\nu_r)$, let $\mathbf{s}_1$ and $\mathbf{s}_2$ be the starting states of two distinct zero segments, and let $\mathbf{q}$ be the starting state of some zero segment. If there exists a label in $B$ on a transition from the ending state of the first zero segment to $\mathbf{q}$, then there is no label in $B$ on a transition from the ending state of the second zero segment to $\mathbf{q}$, for any $\mathbf{s}_2$.*

*Proof:* Assume that the statement is false, i.e., there exist two distinct zero segments having single branch connections in $B$, denoted by $\mathbf{b}_1$ and $\mathbf{b}_2$, respectively, to the starting state $\mathbf{q}$ of some zero segment. Let the starting states of the two zero segments be denoted by $\mathbf{s}_1$ and $\mathbf{s}_2$, respectively. The two zero segments are depicted in Fig. 3. The state $\mathbf{q}$ is given by the two expressions

$$\mathbf{q} = (\mathbf{s}_1)^{\overset{p+1}{\longleftarrow}} + \mathbf{H}\mathbf{b}_1 \tag{15}$$

$$\mathbf{q} = (\mathbf{s}_2)^{\overset{r+1}{\longleftarrow}} + \mathbf{H}\mathbf{b}_2 \tag{16}$$

where $p$ and $r$ are the lengths of the two zero segments, respectively. At least two of the last elements in every component of both $(\mathbf{s}_1)^{\overset{p+1}{\longleftarrow}}$ and $(\mathbf{s}_2)^{\overset{r+1}{\longleftarrow}}$ are equal to zero. We have to consider two cases: (A) $(\mathbf{s}_1)^{\overset{p+1}{\longleftarrow}}$ and $(\mathbf{s}_2)^{\overset{r+1}{\longleftarrow}}$ differ somewhere in the first two elements of at least one component. In this case $\mathbf{b}_1 \neq \mathbf{b}_2$, which from (10) implies that there is a difference somewhere in the last two elements in at least one component in the two expressions above for $\mathbf{q}$. (B) $(\mathbf{s}_1)^{\overset{p+1}{\longleftarrow}}$ and $(\mathbf{s}_2)^{\overset{r+1}{\longleftarrow}}$ do not differ somewhere in the first two elements of any component. In this case $\mathbf{b}_1 = \mathbf{b}_2$, but the two expressions for $\mathbf{q}$ above cannot be equal since $(\mathbf{s}_1)^{\overset{p+1}{\longleftarrow}}$ and $(\mathbf{s}_2)^{\overset{r+1}{\longleftarrow}}$ are unequal by assumption. □

Let $\mathcal{O}$ denote an initially empty set of cycles. Start with some zero segment $\underline{\alpha}$. There exists a unique label $\mathbf{b} \in B$ connecting the ending state of $\underline{\alpha}$ to a zero segment $\underline{\tilde{\alpha}}$ (it may happen that $\underline{\alpha} = \underline{\tilde{\alpha}}$). If a cycle is generated, then add the cycle to $\mathcal{O}$ and continue with another zero segment different from the ones considered. Otherwise, continue with the next zero segment. Repeat the procedure until the set of zero segments is exhausted.

**Corollary 2** *For codes in the class $\mathcal{H}_n^{(r)}(\nu_1,\ldots,\nu_r)$, the cycle set $\mathcal{O}$ is exclusive, i.e., there is no pair of cycles in $\mathcal{O}$ with a common state.*

*Proof:* This follows immediately from the construction and the lemmas above. □

**Lemma 6** *For codes in the class $\mathcal{H}_n^{(r)}(\nu_1,\ldots,\nu_r)$, let $L$ and $W$ be the accumulated length and Hamming weight of all cycles in $\mathcal{O}$, respectively. Then the minimum average cycle weight per branch, $w_0$ in the state diagram from a parity check matrix in $\mathcal{H}_n^{(r)}(\nu_1,\ldots,\nu_r)$ is upper bounded by $W/L$.*

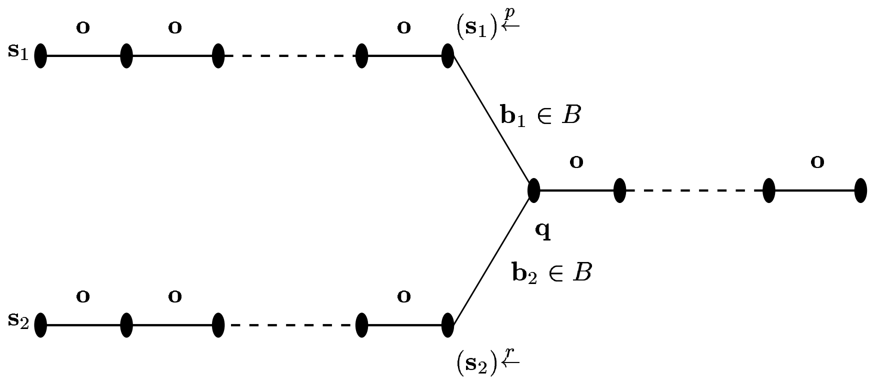


Fig. 3. Two zero segments of length $p$ and $r$ with ending states $(\mathbf{s}_1)^{\stackrel{p}{\leftarrow}}$ and $(\mathbf{s}_2)^{\stackrel{r}{\leftarrow}}$, respectively. The two zero segments have single branch connections with labels from $B$ to the starting state $\mathbf{q}$ of some zero segment.

Lemma 6 was first proved by Hemmati and Costello in [1]. For completeness, the argument is repeated below.

*Proof:* The minimum average cycle weight per branch, $w_0$ is upper bounded by the average cycle weight per branch of some cycle in the set of all cycles in the state diagram, and in particular by the average cycle weight per branch of some cycle in $\mathcal{O}$. Even further, $w_0$ cannot be larger than the average weight per branch of all the cycles in some cycle set. □

**Lemma 7** *Consider an $(n, n-r)$, $r \geq 1$, convolutional code $\mathcal{C}$ defined by a canonical combined parity check matrix $\mathbf{H}$ with row degrees $\nu_i$, $1 \leq i \leq r$. The number, $n_p$ of zero segments of length $p$, $p \geq 1$, in the state diagram from $\mathbf{H}$ is*

$$n_p = 2^{d(p)} + 2^{d(p+2)} - 2^{d(p+1)+1}, \tag{17}$$

*where*

$$d(p) = \sum_{i:1 \leq i \leq r,\ \nu_i \geq p} (\nu_i - p). \tag{18}$$

*If $\nu_i \geq 2$ for all $i$, $1 \leq i \leq r$, then the total number of zero segments is*

$$N_0 = (2^r - 1)2^{\nu - 2r}. \tag{19}$$

*In the state diagram from a parity check matrix in the class $\mathcal{H}_n^{(r)}(\nu_1, \ldots, \nu_r)$, $2 \leq \nu_1 \leq \cdots \leq \nu_r$, defining an $(n, n-r)$, $r \geq 1$, code $\mathcal{C}$, the accumulated length $L$ of all cycles in $\mathcal{O}$ is*

$$L = (2^{r+1} - 1)2^{\nu - 2r} - 1. \tag{20}$$

The proof of Lemma 7 is given in the Appendix.

**Definition 3** *Let $\mathcal{Z} = \mathcal{Z}(\mathbf{H})$ denote the set of zero segments in the state diagram from a canonical combined parity check matrix $\mathbf{H}$ with row degrees $\nu_i \geq 2$ for all $i$, $1 \leq i \leq r$, defining an $(n, n-r)$, $r \geq 1$, code $\mathcal{C}$. Define the (non-injective) mapping*

$$\begin{aligned} \varphi : \mathcal{Z} &\to F^{2r} \\ \underline{\alpha} &\mapsto (s_1^{(1)} s_2^{(1)}, \ldots, s_1^{(r)} s_2^{(r)})^T, \end{aligned} \tag{21}$$

*where $\mathbf{s} = (s_0^{(1)} \cdots s_{\nu_1}^{(1)}, \ldots, s_0^{(r)} \cdots s_{\nu_r}^{(r)})^T$ denotes the ending state of $\underline{\alpha}$. Further, in general, $(s_1^{(1)} s_2^{(1)}, \ldots, s_1^{(r)} s_2^{(r)})^T$ is called the* image *of $\underline{\alpha}$ under $\varphi$ if and only if $\underline{\alpha}\varphi = (s_1^{(1)} s_2^{(1)}, \ldots, s_1^{(r)} s_2^{(r)})^T$.*

**Definition 4** *The* multiplicity *of* $\mathbf{x} \in F^{2r}$ *is the cardinality of the largest subset* $\tilde{\mathcal{Z}}$ *of* $\mathcal{Z}$*, such that* $\tilde{\mathcal{Z}}\varphi = \mathbf{x}$.

**Lemma 8** *Given a canonical combined parity check matrix* $\mathbf{H}$ *with row degrees* $\nu_i \geq 2$ *for all* $i$*,* $1 \leq i \leq r$*, the cardinality of the image of* $\mathcal{Z} = \mathcal{Z}(\mathbf{H})$ *under* $\varphi$*,* $\mathcal{Z}\varphi$ *is* $2^{r-\gamma}(2^r - 1)$*, where* $\gamma$ *is the number of rows with row degree 2. Further, the multiplicity of* $\mathbf{x} \in F^{2r}$ *is either zero or equal to some fixed positive integer for all* $\mathbf{x} \in F^{2r}$*, i.e.,* $\mathcal{Z}$ *can be partitioned into equally sized subsets,* $\mathcal{Z}_j$*,* $1 \leq j \leq 2^{r-\gamma}(2^r - 1)$*, such that* $\underline{\alpha}\varphi \neq \underline{\delta}\varphi$ *for every pair* $(\underline{\alpha}, \underline{\delta}) \in \mathcal{Z}_i \times \mathcal{Z}_j$*,* $i \neq j$*.*

*Sketch of Proof:* The first part of the lemma is a simple combinatorial exercise. In general, the image of $\mathcal{Z}$ under $\varphi$ contains column vectors of the form

$$\mathbf{x} = \left(x_0^{(1)}x_1^{(1)}, \ldots, x_0^{(\gamma)}x_1^{(\gamma)}, x_0^{(\gamma+1)}x_1^{(\gamma+1)}, \ldots, x_0^{(r)}x_1^{(r)}\right)^T, \tag{22}$$

where one and only one of the following two constraints apply (both constraints can not be satisfied):

1) $(\gamma > 0) : x_1^{(i)} = 0$ for all $i$, $1 \leq i \leq \gamma$, and there exists at least one $i$, $1 \leq i \leq \gamma$, such that $x_0^{(i)} = 1$.

2) $(\gamma < r) : x_0^{(i)} = x_1^{(i)} = 0$ for all $i$, $1 \leq i \leq \gamma$, and there exists at least one $i$, $\gamma + 1 \leq i \leq r$, such that $x_0^{(i)} = 1$.

We will illustrate this below in Example 3. The number of column vectors $\mathbf{x} \in F^{2r}$ satisfying condition 1) is $2^{2(r-\gamma)}(2^\gamma - 1)$. This is the case since (A) all possibilities are allowed in the last $2(r-\gamma)$ coordinates, (B) $x_1^{(1)} = \cdots = x_1^{(\gamma)} = 0$, and (C) $(x_0^{(1)} \cdots x_0^{(\gamma)})^T$ should be different from the all-zero vector. Using the same type of argument, the number of column vectors $\mathbf{x} \in F^{2r}$ satisfying condition 2) can be shown to be $2^{r-\gamma}(2^{r-\gamma} - 1)$. The cardinality of $\mathcal{Z}\varphi$ is

$$|\mathcal{Z}\varphi| = 2^{2(r-\gamma)}(2^\gamma - 1) + 2^{r-\gamma}(2^{r-\gamma} - 1) = 2^{r-\gamma}(2^r - 1), \tag{23}$$

and the first part of the lemma follows.

For the second part, let $G$ be the subset of states in the state diagram from $\mathbf{H}$ of the form

$$\left(0s_1^{(1)} \cdots s_{\nu_1-1}^{(1)}0, \ldots, 0s_1^{(r)} \cdots s_{\nu_r-1}^{(r)}0\right)^T. \tag{24}$$

Obviously, $G$ is a group under component-wise addition. Let $H^{(\gamma)}$ be the subgroup of $G$ consisting of every state $\mathbf{s}$ of the form

$$\mathbf{s} = \left(\mathbf{0}_3, \ldots, \mathbf{0}_3, \mathbf{0}_3 s_3^{(\gamma+1)} \cdots s_{\nu_{\gamma+1}-1}^{(\gamma+1)}0, \ldots, \mathbf{0}_3 s_3^{(r)} \cdots s_{\nu_r-1}^{(r)}0\right)^T, \tag{25}$$

where $\mathbf{0}_3 = (000)$. In the following we consider cosets of $G$ of the form $\mathbf{g} + H^{(\gamma)}$, where $\mathbf{g}$ is any vector in $G$ of the form

$$\left(0g_1^{(1)}g_2^{(1)}0 \cdots 0, \ldots, 0g_1^{(r)}g_2^{(r)}0 \cdots 0\right)^T, \tag{26}$$

such that $(g_1^{(1)}g_2^{(1)}, \ldots, g_1^{(r)}g_2^{(r)})^T \in \mathcal{Z}\varphi$. Every element in the coset $\mathbf{g} + H^{(\gamma)}$ is the ending state of a unique zero segment. In more detail, there is a one-to-one correspondence between the zero segments in $\mathcal{Z}_j$ for some $j$, $1 \leq j \leq 2^{r-\gamma}(2^r - 1)$, and the elements in $\mathbf{g} + H^{(\gamma)}$. In fact, $(\mathcal{Z}_j)\varphi = (g_1^{(1)}g_2^{(1)}, \ldots, g_1^{(r)}g_2^{(r)})^T$. From abstract algebra we know that every coset has the same number of elements, and the result follows. □

**Example 3** *Choose* $r = 3$*,* $\gamma = 1$ *(which implies that* $\nu_1 = 2$*), and* $\nu_2 = \nu_3 = 3$*. In general, the starting state* $\mathbf{s}$ *of some zero segment* $\underline{\alpha}$ *of length* $p$ *is of the form*

$$\mathbf{s} = \left(00s_2^{(1)}, 00s_2^{(2)}s_3^{(2)}, 00s_2^{(3)}s_3^{(3)}\right)^T. \tag{27}$$

*Note that all* $2^5$ *possibilities will not result in valid starting states. We consider the cases* $p = 1$ *and* $p = 2$ *separately.*

*(A)* $(p=1)$ : *Set* $s_2^{(1)}=1$. *In this case* $\mathbf{s}$ *is a valid starting state for any values of* $s_2^{(2)}$, $s_3^{(2)}$, $s_2^{(3)}$, *and* $s_3^{(3)}$. *The image of* $\underline{\alpha}$ *under* $\varphi$ *is the column vector*

$$\underline{\alpha}\varphi = \left(10, s_2^{(2)}s_3^{(2)}, s_2^{(3)}s_3^{(3)}\right)^T. \tag{28}$$

*It follows that there will be exactly* $2^4=16$ *distinct images under* $\varphi$, *each with multiplicity 1. Note that* $\underline{\alpha}\varphi$ *has the form in (22) under condition 1). Secondly, set* $s_2^{(1)}=0$. *Since the length of* $\underline{\alpha}$ *is 1,* $s_2^{(2)}$ *and* $s_2^{(3)}$ *can not both be zero. Further,* $s_3^{(2)}$ *and* $s_3^{(3)}$ *can not both be zero, because* $\mathbf{s}$ *is the starting state of a zero segment. The possible images of* $\underline{\alpha}$ *under* $\varphi$ *is the set of column vectors*

$$\begin{array}{l}\{\ (00,01,10)^T, (00,01,11)^T, (00,10,01)^T,\\ \ \ (00,10,11)^T, (00,00,11)^T, (00,11,00)^T,\\ \ \ (00,11,01)^T, (00,11,10)^T, (00,11,11)^T\ \},\end{array} \tag{29}$$

*all of which are different from the previous 16. Thus, we have* $16+9=25$ *distinct images, each with multiplicity 1. Note that* $\underline{\alpha}\varphi$ *has the form in (22) under condition 2).*

*(B)* $(p=2)$ : *In this case every starting state* $\mathbf{s}$ *of* $\underline{\alpha}$ *is of the form*

$$\mathbf{s} = \left(000, 000s_3^{(2)}, 000s_3^{(3)}\right)^T, \tag{30}$$

*with not both* $s_3^{(2)}$ *and* $s_3^{(3)}$ *equal to zero. If* $s_3^{(2)}=1$, *then* $\mathbf{s}$ *is a valid starting state for any value of* $s_3^{(3)}$. *Secondly, if* $s_3^{(3)}=1$, *then again* $\mathbf{s}$ *is a valid starting state for any value of* $s_3^{(2)}$. *Consequently, there are 3 distinct starting states which result in the following 3 images under* $\varphi$

$$\{\ (00,10,00)^T, (00,00,10)^T, (00,10,10)^T\ \}, \tag{31}$$

*all of which are different from the previous 25. Note that* $\underline{\alpha}\varphi$ *has the form in (22) under condition 2).*

*To summarize we have 28 images each with multiplicity 1. Evaluating* $2^{r-\gamma}(2^r-1)$ *for* $r=3$ *and* $\gamma=1$ *gives the same number.*

**Definition 5** *Let* $\mathbf{v}_i$, $1\le i\le 2r$, *denote the ith basis vector in* $\mathcal{V}_{\rm low}$. *Further, write* $(\mathcal{Z}_j)\varphi$, $1\le j\le|\mathcal{Z}\varphi|$, *as a linear combination of the basis vectors in* $\mathcal{V}_{\rm low}$ *as follows:*

$$(\mathcal{Z}_j)\varphi = \alpha_1^{(j)}\mathbf{v}_1 + \alpha_2^{(j)}\mathbf{v}_2 + \cdots + \alpha_{2r}^{(j)}\mathbf{v}_{2r}, \tag{32}$$

*where the column vector* $(\alpha_1^{(j)}\cdots\alpha_{2r}^{(j)})^T$ *consists of the coefficients in the linear combination of the basis vectors in* $\mathcal{V}_{\rm low}$ *representing* $(\mathcal{Z}_j)\varphi$. *Define* $w_j$ *as the Hamming weight of the jth column of the* $2r\times|\mathcal{Z}\varphi|$ *matrix*

$$\begin{pmatrix}\alpha_1^{(1)} & \alpha_1^{(2)} & \cdots & \alpha_1^{(|\mathcal{Z}\varphi|)}\\ \alpha_2^{(1)} & \alpha_2^{(2)} & \cdots & \alpha_2^{(|\mathcal{Z}\varphi|)}\\ \cdots & \cdots & \cdots & \cdots\\ \alpha_{2r}^{(1)} & \alpha_{2r}^{(2)} & \cdots & \alpha_{2r}^{(|\mathcal{Z}\varphi|)}\end{pmatrix}, \tag{33}$$

*and let* $A_\gamma \triangleq w_1+\cdots+w_{|\mathcal{Z}\varphi|}$.

**Lemma 9** *In the state diagram from a parity check matrix in the class* $\mathcal{H}_n^{(r)}(\nu_1,\ldots,\nu_r)$, $2\le\nu_1\le\cdots\le\nu_r$, *defining an* $(n,n-r)$, $r\ge1$, *code* $\mathcal{C}$, *the accumulated weight* $W$ *of all cycles in* $\mathcal{O}$ *is*

$$W = A_\gamma\cdot 2^{\nu-3r+\gamma}, \tag{34}$$

*where* $A_\gamma$ *is defined in Definition 5.*

*Proof:* In general,

$$W = |\mathcal{Z}_1|w_1 + |\mathcal{Z}_2|w_2 + \cdots + |\mathcal{Z}_{|\mathcal{Z}\varphi|}|w_{|\mathcal{Z}\varphi|}, \tag{35}$$

where $w_j$, $1 \leq j \leq |\mathcal{Z}\varphi|$, is defined in Definition 5. From Lemmas 7 and 8,

$$|\mathcal{Z}_1| = |\mathcal{Z}_2| = \cdots = |\mathcal{Z}_{|\mathcal{Z}\varphi|}| = \frac{N_0}{|\mathcal{Z}\varphi|}, \tag{36}$$

and the result follows from the fact that $A_\gamma = w_1 + \cdots + w_{|\mathcal{Z}\varphi|}$, $|\mathcal{Z}\varphi| = 2^{r-\gamma}(2^r - 1)$, and $N_0 = (2^r - 1)2^{\nu-2r}$. □

**Theorem 1** *Let $\mathcal{C}$ be a code defined by a parity check matrix in the class $\mathcal{H}_n^{(r)}(\nu_1, \ldots, \nu_r)$, $2 \leq \nu_1 \leq \cdots \leq \nu_r$. The minimum average cycle weight per branch in the state diagram for $\mathcal{C}$, excluding the all-zero cycle around the all-zero state, is upper bounded by*

$$w_0 \leq W/L = \frac{A_\gamma 2^{\nu-3r+\gamma}}{(2^{r+1} - 1)2^{\nu-2r} - 1}, \tag{37}$$

*where $\gamma$ is the number of rows with row degree 2 in the canonical parity check matrix from $\mathcal{H}_n^{(r)}(\nu_1, \ldots, \nu_r)$ defining $\mathcal{C}$.*

*Proof:* The result follows immediately from Lemmas 6, 7, and 9. □

## 3.1 Codes with $r = 1$

In this subsection we will show that our derived upper bound in Theorem 1 with $r = 1$ reduces to the main theorem in [2]. We first consider Definition 1, which reduces to:

*Let $\mathcal{H}_n^{(1)}(\nu)$, $n \geq 2$, be the class of all canonical $(\nu + 1) \times n$ combined parity check matrices with $\nu \geq 2$, satisfying the following constraints: There exists an ordered set $J = \{1 \leq j_1 < j_2 \leq n\}$ of cardinality 2 such that*

$$h_{j,0} = h_{j,\nu}, \quad j \in J, \tag{38}$$

*and*

$$\mathcal{V}_{\text{low}} = \left\{(h_{j_1,0}h_{j_1,1})^T, (h_{j_2,0}h_{j_2,1})^T\right\} \tag{39}$$

$$\mathcal{V}_{\text{high}} = \left\{(h_{j_1,\nu-1}h_{j_1,\nu})^T, (h_{j_2,\nu-1}h_{j_2,\nu})^T\right\} \tag{40}$$

*are sets of* linearly independent *column vectors spanning $F^2$. In addition, every polynomial in the parity check matrices should have constant terms different from zero.*

Actually, $\mathcal{H}_n^{(1)}(\nu)$ is the class of all canonical parity check matrices with $\nu \geq 2$ that satisfy

$$\begin{aligned} h_{1,0} = h_{2,0} = \cdots = h_{n,0} = 1 \\ h_{j_1,1} \neq h_{j_2,1} \\ h_{j_1,\nu-1} \neq h_{j_2,\nu-1} \\ h_{j_1,\nu} = h_{j_2,\nu} = 1, \end{aligned} \tag{41}$$

which is equal to the class considered in [2]. Further, we have to consider two cases ($\nu = 2$ and $\nu \geq 3$):

1. $\nu = 2$ implies $\gamma = 1$. In this case $\mathcal{Z}\varphi = \{(10)^T\}$ and $A_\gamma = A_1 = 1$. From (37), $w_0 \leq 2^{\nu-2}/(3 \cdot 2^{\nu-2} - 1)$.

2. $\nu \geq 3$ implies $\gamma = 0$. In this case $\mathcal{Z}\varphi = \{(11)^T, (10)^T\}$ and $A_\gamma = A_0 = 2$. From (37), $w_0 \leq 2^{\nu-2}/(3 \cdot 2^{\nu-2} - 1)$.

In both cases, we get the upper bound provided in the main theorem in [2].

### 3.2 Codes with $r = 2$

The upper bound in Theorem 1 depends on $A_\gamma$, which again depends on $\mathcal{V}_{\text{low}}$. As an example we will consider $r = 2$ codes. The number of zero segments of length $p$, $p \geq 1$, can be written out explicitly using (17) and (18) as illustrated by the following examples.

**Example 4** *Consider codes with $\nu_1 = \nu_2 = \nu/2$, $\nu$ even. Using Lemma 7, we can find the number of zero segments of length p, $1 \leq p \leq \nu/2 - 1$, which is*

$$n_p = \begin{cases} 9 \cdot 2^{\nu-2(p+2)} & \text{if } 1 \leq p \leq \nu/2 - 2, \\ 3 & \text{if } p = \nu/2 - 1, \end{cases} \tag{42}$$

*for all $\nu \geq 4$, $\nu$ even.*

**Example 5** *Consider codes with $2 \leq \nu_1 < \nu_2$. Using Lemma 7, we can find the number of zero segments of length p, $1 \leq p \leq \nu_2 - 1$, which is*

$$n_p = \begin{cases} 9 \cdot 2^{\nu-2(p+2)} & \text{if } 1 \leq p \leq \nu_1 - 2, \\ 5 \cdot 2^{\nu_2-p-2} & \text{if } p = \nu_1 - 1, \\ 2^{\nu_2-p-2} & \text{if } \nu_1 \leq p \leq \nu_2 - 2, \\ 1 & \text{if } p = \nu_2 - 1, \end{cases} \tag{43}$$

*for all $2 \leq \nu_1 < \nu_2$.*

Theorem 1 reduces to the following corollary when $r = 2$.

**Corollary 3** *Let $\mathcal{C}$ be a code defined by a parity check matrix in the class $\mathcal{H}_n^{(2)}(\nu_1, \nu_2)$, $2 \leq \nu_1 \leq \nu_2$. The minimum average cycle weight per branch in the state diagram for $\mathcal{C}$, excluding the all-zero cycle around the all-zero state, is upper bounded by*

$$w_0 \leq W/L = \begin{cases} \frac{A_2}{7 \cdot 2^{\nu-4}-1} & \text{if } \nu_1 = 2 \text{ and } \nu_2 = 2, \\ \frac{A_1 \cdot 2^{\nu-5}}{7 \cdot 2^{\nu-4}-1} & \text{if } \nu_1 = 2 \text{ and } \nu_2 \geq 3, \\ \frac{A_0 \cdot 2^{\nu-6}}{7 \cdot 2^{\nu-4}-1} & \text{otherwise.} \end{cases} \tag{44}$$

**Remark 1** *To use Corollary 3, we have to determine $A_2$, $A_1$, and $A_0$ using (32) and (33). To do that, we have to write out the image $\mathcal{Z}\varphi$ explicitly for $\gamma = 2$, $\gamma = 1$, and $\gamma = 0$. For $\gamma = 2$, $|\mathcal{Z}\varphi| = 3$, and from (22) under condition 1)*

$$\mathcal{Z}\varphi = \{(10,00)^T, (10,10)^T, (00,10)^T\}. \tag{45}$$

*Further, for $\gamma = 1$, $|\mathcal{Z}\varphi| = 6$, and from (22) under conditions 1) and 2)*

$$\mathcal{Z}\varphi = \{(10,00)^T, (10,10)^T, (00,10)^T, (10,01)^T, (10,11)^T, (00,11)^T\}. \tag{46}$$

*At last, for $\gamma = 0$, $|\mathcal{Z}\varphi| = 12$, and from (22) under condition 2)*

$$\mathcal{Z}\varphi = \begin{array}{l} \{ (10,00)^T, (10,01)^T, (11,01)^T, (10,10)^T, (10,11)^T, (11,10)^T, \\ \quad (11,11)^T, (00,10)^T, (00,11)^T, (11,00)^T, (01,10)^T, (01,11)^T \}. \end{array} \tag{47}$$

*For any given parity check matrix in the class $\mathcal{H}_n^{(2)}(\nu_1, \nu_2)$, $2 \leq \nu_1 \leq \nu_2$, determine the basis vectors in $\mathcal{V}_{\text{low}}$. Then, use (45), (46), (47), and $\mathcal{V}_{\text{low}}$ to determine the matrix in (33), from which $A_2$, $A_1$, and $A_0$ are easily computed as the corresponding Hamming weight of the matrix.*

**Example 6** *Define the subclass of canonical combined parity check matrices from $\mathcal{H}_n^{(2)}(\nu_1, \nu_2)$, $2 \leq \nu_1 \leq \nu_2$, with*

$$\mathcal{V}_{\text{low}} = \left\{(11,00)^T, (10,01)^T, (10,00)^T, (10,11)^T\right\}. \tag{48}$$

*The column vectors in $\mathcal{V}_{\text{low}}$ are linearly independent. Using $\mathcal{V}_{\text{low}}$, it is a simple task to verify that $A_2 = 6$, $A_1 = 10$, and $A_0 = 26$. The code in (4) is within this subclass with $n = 4$, $\nu_1 = 2$, and $\nu_2 = 3$, and it is the tightest code of free distance $d_{\text{free}} \geq 3$ within the class $\mathcal{H}_4^{(2)}(\nu_1, \nu_2)$, $\nu_1 + \nu_2 = 5$.*

Table 1. Canonical parity check matrices in decimal compressed form of the tightest rate 2/4 codes found versus code degree $\nu$.

| $\nu$ | $(\nu_1, \nu_2)$ | Parity check matrix | $w_0$ | $w_0^{\text{bound}}$ | $w_0^{\text{bound}} - w_0$ |
|---|---|---|---|---|---|
| 4 | (2,2) | (5,7,21,40) | 1/2 | 0.667 | 0.167 |
| 5 | (2,3) | (7,21,37,93) | 2/3 | 0.769 | 0.103 |
| 6 | (2,4) | (7,21,173,253) | 2/3 | 0.741 | 0.074 |
| 7 | (2,5) | (21,197,247,445) | 21/31 | 0.727 | 0.050 |
| 8 | (2,6) | (21,621,829,847) | 31/49 | 0.721 | 0.088[1] |

## 4 Computer search

A computer search was carried out to determine the tightness of the derived upper bound. The search was performed over all canonical parity check matrices in the class $H_4^2(\nu_1, \nu_2)$ with $d_{\text{free}} \geq 3$ and $\nu_1 + \nu_2 = \nu$, $2 \leq \nu_1 \leq \nu_2$. The results are tabulated in Table 1 for different code degrees. Note that the difference between the upper bound and $w_0$ is vanishingly small as the code degree gets large. The parity check matrices are tabulated in *decimal compressed form*. As an example, the matrix in (4) is tabulated as $(7, 21, 37, 93)$. Karp's algorithm [13] was used to determine $w_0$. Note that the asymptotic complexity of Karp's algorithm is $\Theta(2^{2\nu} \cdot \min(2^{n-r}, 2^{\nu}))$, making it hard to search for codes with large dimension and code degree.

## 5 Conclusion

We have generalized the approach by Hole and Hole in [2] deriving upper bounds on the minimum average cycle weight per branch of special classes of $(n, n-r)$, $r \geq 2$, convolutional codes.

We looked in particular at $r = 2$ codes. The results of an exhaustive computer search indicate that the derived upper bound is almost tight as the code degree increases. For $r \geq 3$ a partial search seems to indicate that the derived bounds are less tight.

## Appendix

The proof of Lemma 7 is given below. In the first part of the proof we derive the expression for the number of zero segments of length $p$, $p \geq 1$. Then we use this expression to calculate the accumulated length of all cycles in $\mathcal{O}$, and the total number of zero segments.

Part I: According to Forney [12], the number of distinct all-zero paths of length $p$, $p \geq 1$, in a minimal state diagram of rate $(n-r)/n$, $r \geq 1$, convolutional codes defined by a canonical parity check matrix with row degrees $\nu_i$, $1 \leq i \leq r$, is $2^{d(p)}$, where $d(p)$ is defined in (18). For convenience the definition is restated below:

$$d(p) = \sum_{i:1 \leq i \leq r,\, \nu_i \geq p} (\nu_i - p). \tag{49}$$

In [1], Hemmati and Costello used this result to establish a closed form expression for the number of zero segments of length $p$ in a minimal state diagram of a rate 1/2 convolutional code. Their proof is extended to the general case as follows:

Let $\nu_{\max}$ denote the maximum row degree. From (49), it follows that there is one and only one all-zero path of length $p$ for $p \geq \nu_{\max}$, namely the all-zero path around the all-zero state. Thus, there are no zero segments of length $\geq \nu_{\max}$. The number of all-zero paths of length $\nu_{\max} - 1$ is $2^{d(\nu_{\max}-1)}$. After excluding the all-zero path around the all-zero state, the remaining all-zero paths have to be zero segments of length $\nu_{\max} - 1$. If this is not the case, then it is possible to extend any

[1]The search is not exhaustive.

of these all-zero paths in at least one direction. After the extension we will have an all-zero path of length $\nu_{\max}$. From (49) the number of all-zero paths of length $\geq \nu_{\max}$ is 1, which is the all-zero path around the all-zero state. Since we have already excluded this path from consideration, we have a contradiction. Hence, there are $n_{\nu_{\max}-1} = 2^{d(\nu_{\max}-1)} - 1$ zero segments of length $\nu_{\max} - 1$.

From (49), there are $2^{d(\nu_{\max}-2)}$ all-zero paths of length $\nu_{\max} - 2$. To count the number of zero segments of length $\nu_{\max} - 2$, we must exclude the 2 subsequences of length $\nu_{\max} - 2$ of each zero segment of length $\nu_{\max} - 1$ as well as the all-zero path around the all-zero state. Thus,

$$n_{\nu_{\max}-2} = 2^{d(\nu_{\max}-2)} - 2n_{\nu_{\max}-1} - 1. \tag{50}$$

The same type of argument leads to the general expression

$$n_p = 2^{d(p)} - 2n_{p+1} - 3n_{p+2} - \cdots\cdots - (\nu_{\max} - p)n_{\nu_{\max}-1} - 1. \tag{51}$$

Substituting $p+1$ for $p$ in the recursive relation in (51) gives

$$n_{p+1} = 2^{d(p+1)} - 2n_{p+2} - 3n_{p+3} - \cdots\cdots - (\nu_{\max} - p - 1)n_{\nu_{\max}-1} - 1. \tag{52}$$

Taking the difference between $n_p$ from (51) and $n_{p+1}$ from (52) yields

$$n_p - n_{p+1} = 2^{d(p)} - 2^{d(p+1)} - 2n_{p+1} - n_{p+2} - n_{p+3} - \cdots - n_{\nu_{\max}-1}, \tag{53}$$

which implies

$$n_p = 2^{d(p)} - 2^{d(p+1)} - n_{p+1} - \cdots - n_{\nu_{\max}-1} = 2^{d(p)} - 2^{d(p+1)} - n_{p+1} - \sum_{i=p+2}^{\nu_{\max}-1} n_i. \tag{54}$$

Further,

$$n_{p+1} = 2^{d(p+1)} - 2^{d(p+2)} - \sum_{i=p+2}^{\nu_{\max}-1} n_i. \tag{55}$$

Combining (54) and (55) gives the general expression

$$n_p = 2^{d(p)} + 2^{d(p+2)} - 2^{d(p+1)+1}, \tag{56}$$

valid for all $p \geq 1$.

Part II: We use the general expression for $n_p$ in (56).

$$\begin{aligned} L &= \sum_{p=1}^{\nu_{\max}} (1+p)n_p = \sum_{p=1}^{\nu_{\max}} (1+p)2^{d(p)} + \sum_{p=1}^{\nu_{\max}} (1+p)2^{d(p+2)} - 2\sum_{p=1}^{\nu_{\max}} (1+p)2^{d(p+1)} \\ &= \sum_{p=1}^{\nu_{\max}} (1+p)2^{d(p)} + \sum_{p=-1}^{\nu_{\max}} (1+p)2^{d(p+2)} - 2^{d(2)} - 2\sum_{p=0}^{\nu_{\max}} (1+p)2^{d(p+1)} + 2^{d(1)+1}. \end{aligned} \tag{57}$$

In the second summation in (57) we now change the summation index from $p$ to $q = p+2$, and in the third summation we change the summation index from $p$ to $q = p+1$. It follows (using the fact that $d(\nu_{\max}+1) = d(\nu_{\max}+2) = 0$) that

$$\begin{aligned} L &= \sum_{q=1}^{\nu_{\max}} (1+q)2^{d(q)} + \sum_{q=1}^{\nu_{\max}+2} (q-1)2^{d(q)} - 2\sum_{q=1}^{\nu_{\max}+1} q2^{d(q)} - 2^{d(2)} + 2^{d(1)+1} \\ &= \nu_{\max}2^{d(\nu_{\max}+1)} + (\nu_{\max}+1)2^{d(\nu_{\max}+2)} - 2(\nu_{\max}+1)2^{d(\nu_{\max}+1)} - 2^{d(2)} + 2^{d(1)+1} \\ &= 2^{d(1)+1} - 2^{d(2)} - 1. \end{aligned} \tag{58}$$

Using the facts that $d(1) = \nu - r$ and $d(2) = \nu - 2r$ (since $\nu_i \geq 2$ for all $i$, $1 \leq i \leq r$), it follows that

$$L = (2^{r+1} - 1)2^{\nu-2r} - 1. \tag{59}$$

Part III: We can use the same argument as above to derive the total number of zero segments

$$N_0 = \sum_{p=1}^{\nu_{\max}} n_p = (2^r - 1)2^{\nu-2r}. \tag{60}$$

## Acknowledgment

The author wishes to thank Kjell J. Hole and Øyvind Ytrehus for helpful comments and suggestions.

## References

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

[2] 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.

[3] S. Höst, R. Johannesson, K. S. Zigangirov, and V. V. Zyablov, "Active distances for convolutional codes," *IEEE Trans. Inform. Theory*, vol. 45, pp. 658–669, Mar. 1999.

[4] R. Jordan, V. Pavlushkov, and V. V. Zyablov, "Maximum slope convolutional codes." *IEEE Trans. Inform. Theory*, submitted for publication, 2001.

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

[6] R. Jordan, V. Pavlushkov, and V. V. Zyablov, "An upper bound on the slope of convolutional codes," in *Proc. IEEE Int. Symp. on Inform. Theory (ISIT)*, (Lausanne, Switzerland), p. 424, July 2002.

[7] R. J. McEliece, "The algebraic theory of convolutional codes," in *Handbook of Coding Theory* (V. S. Pless and W. C. Huffman, eds.), ch. 12, pp. 1065–1138, North-Holland, Amsterdam: Elsevier, 1998.

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

[9] P. CharnKeitKong, H. Imai, and K. Yamaguchi, "On classes of rate $k/(k+1)$ convolutional codes and their decoding techniques," *IEEE Trans. Inform. Theory*, vol. 42, pp. 2181–2193, Nov. 1996.

[10] K. J. Hole, "On classes of convolutional codes that are not asymptotically catastrophic," *IEEE Trans. Inform. Theory*, vol. 46, pp. 663–669, Mar. 2000.

[11] E. Paaske, "Short convolutional codes with maximal free distances for rates 2/3 and 3/4," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 683–689, Sept. 1974.

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

[13] R. M. Karp, "A characterization of the minimum cycle mean in a digraph," *SIAM J. Discr. Math.*, vol. 23, pp. 309–311, 1978.