Ascetic convolutional codes

Øyvind Ytrehus

University of Bergen, Department of Informatics,
Høyteknologisenteret, N-5020 Bergen, Norway
E-mail: oyvind@ii.uib.no
Supported by the Norwegian Research Council (NFR).

Acknowledgement

The “Applications” section of this paper presents an overview of research carried out at the Department of Informatics at the University of Bergen. Kjell Jørgen Hole, Marianne Fjelltveit Hole, and Rolf Erstad made significant contributions to this work.

Abstract

Ascetic convolutional codes contain fewer edge labels than other convolutional codes. This is advantageous in several practical situations. Selecting an ascetic code over a non-ascetic code carries no severe penalty, in terms of trade-off of minimum distance, code rate and complexity.

1. Introduction

One day at the dawn of time, when convolutional codes were created, all the little convolutional code children were assigned the following task: They were to assign binary labels to their so far naked trellis edges, in order to achieve a Large Free Hamming Distance.

Every little $(n = 8, k = 6, \nu = 1)$ convolutional code distributed vectors from all of F_2^n over their trellis edges. Ah, well, not quite everyone. One little code only selected even-weight labels. All the other little codes chided her, “How can you *possibly* expect to achieve a Large Free Hamming Distance when you don’t use more than half of the available edge labels? Here, try (01111111)”. But she answered, stubbornly, “No. I need no odd-weight vectors”.

She was an *ascetic convolutional code*.

Judgement day arrived, and the little codes had to submit their (now edge labelled) trellis to the Modified Viterbi Algorithm. Some clever codes had achieved Free Hamming Distances of one or two or even three, some careless ones had constructed catastrophic trellises and were forever doomed to Loss of Dimension or even to an existence as mere block codes. “Was this the last one?”, asked the Modified Viterbi Algorithm, preparing to leave, (it had been a long day). The little codes answered, “Well, there’s that little ascetic code, but she only used half of the edge labels anyway”. As usual in this kind of story, the little ascetic code still had a go with the Modified Viterbi Algorithm. What was her Free Hamming Distance? The answer is postponed until Theorem 3.5.

(But she did live happily ever after).

2. Preliminaries

The notation introduced in this section is kept at a minimum. In particular, we shall neither need nor define the concepts of *encoder*, *basic encoder*, *minimal encoder*, or the procedure for obtaining a minimal encoder from the parity check matrix. For a comprehensive treatment of the theory of convolutional codes the reader is referred to [1], or to the pioneering work in [2, 3, 4].

Let $\mathbf{h} = (h_1, \ldots, h_t)^T$ (for some $t > 0$) be a binary (column) vector. The *l-th shift* of $\mathbf{h}$ is the sequence $\mathbf{h}^{\xrightarrow{l}} = \left(\overbrace{0, \ldots, 0}^{l}, h_1, h_2, \ldots, h_t, 0, \ldots\right)^T$, starting at position one and extending to infinity, in which $\mathbf{h}$ can be found as the subvector starting in the $(l+1)$-th position and all other entries are zero.

Let similarly $\mathbf{H}$ be a binary $(\nu + 1) \times n$ matrix. The l-th shift of $\mathbf{H}$ is the matrix $\mathbf{H}^{\xrightarrow{l}}$ of n columns and an infinite number of rows, in which $\mathbf{H}$ can be found as the submatrix starting in the $(l+1)$-th row and all other entries are zero. Let further

$$
\begin{aligned}
\mathbf{H}^* =& \left(\mathbf{H}^{\xrightarrow{0}} \quad \mathbf{H}^{\xrightarrow{1}} \quad \mathbf{H}^{\xrightarrow{2}} \quad \cdots\right) \\
=& (\mathbf{h}_1^* \quad \mathbf{h}_2^* \quad \cdots \quad \mathbf{h}_n^* \quad \mathbf{h}_{n+1}^* \quad \cdots),
\end{aligned}
$$

where for $i \geq 1$, the column $\mathbf{h}_i^* = \mathbf{h}_{((i-1) \bmod n)+1}^{\xrightarrow{\lfloor (i-1)/n \rfloor}}$ is the $\lfloor (i-1)/n \rfloor$-th shift of the $(((i-1) \bmod n) + 1)$-th column of $\mathbf{H}$.

Let $\mathcal{F}$ be the set of infinite binary sequences $\mathbf{u} = (u_1, u_2, \ldots)$ starting at time 1. $\mathbf{H}$ is a parity-check matrix for a convolutional code $\mathcal{C} = \{\mathbf{u} \in \mathcal{F} | \mathbf{H}^* \mathbf{u}^T = \mathbf{h}_1 u_1 + \mathbf{h}_2 u_2 + \ldots + \mathbf{h}_i u_i + \ldots = \mathbf{0}\}$. The elements of $\mathcal{C}$ are called code words.

(***Remark.*** Strictly speaking, $\mathcal{C}$ should be defined as a vector space over the field of Laurent series over $F = GF(2)$ (see e. g. [1]). The current definition was chosen to enhance simplicity and since in practice nonzero code words will start at or after some fixed time.)

$\mathcal{C}$ has *block length* n, *dimension* $n-1$ and constraint length at most ν. The constraint length is equal to ν if there exists no other parity check matrix $\mathbf{H}'$ for the same code $\mathcal{C}$ with smaller row dimension; in this case $\mathbf{H}$ is a minimal parity check matrix. In the following we will assume that parity check matrices are minimal.

Let $\mathcal{H}$, referred subsequently to as a *combined* matrix, be a collection of r matrices $\mathbf{H}^{(i)}, i = 1, \ldots, r$, over F where $\mathbf{H}^{(i)}$ has $\nu_i + 1$ rows and n columns.

The n columns of $\mathbf{H}^{(i)}$ are denoted $\mathbf{h}_j^{(i)} = (h_{j,0}^{(i)}, h_{j,1}^{(i)}, \ldots, h_{j,\nu_i}^{(i)})^T$, $j = 1, \ldots, n$, where the last row of $\mathbf{H}^{(i)}$ has at least one nonzero entry. The binary column vector $\mathbf{h}_j = (h_{j,0}^{(1)}, \ldots, h_{j,\nu_1}^{(1)}, h_{j,0}^{(2)}, \ldots, h_{j,0}^{(r)}, \ldots, h_{j,\nu_r}^{(r)})^T \in F^{\nu+r}$ will be called the j-th *combined column* of $\mathcal{H}$. Thus $\mathcal{H}$ can be regarded as a binary $\left(\sum_{i=1}^{r} (\nu_i + 1)\right) \times n$ matrix.

Example 1. *Let*

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ - & - & - & - & - & - & - & - \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ - & - & - & - & - & - & - & - \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ - & - & - & - & - & - & - & - \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

The code $\mathcal{C}_1$ defined by this matrix has length 8 and dimension 4.

(A more common notation would be to represent $\mathbf{H}$ in terms of its D-*transform*).

Example 1 (Cont.). *The D-transform of $\mathbf{H}$ is*

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ - & - & - & - & - & - & - & - \\ 1+D & D & 1 & D & D & 1+D & 1 & 0 \\ - & - & - & - & - & - & - & - \\ D & 1 & 1 & D & 0 & 1+D & 0 & 1+D \\ - & - & - & - & - & - & - & - \\ 1 & D & 1 & 1 & 0 & D & D & 1+D \end{pmatrix}.$$

Each matrix $\mathbf{H}^{(r)}$ is a parity-check matrix for some convolutional code $\mathcal{C}^{(i)}$ of dimension $(n-1)$. We call $\mathcal{H}$ the parity-check matrix for the convolutional code $\mathcal{C} = \cap_{i=1}^{r} \mathcal{C}^{(i)}$. $\mathcal{C}$ has block length n, dimension at least $n-r$, and *constraint length* (or *memory*) $\nu = \sum_{i=1}^{r} \nu_i$. The *free distance* $d_{\text{free}} = d_{\text{free}}(\mathcal{C})$ is the minimum Hamming weight of any nonzero code word in $\mathcal{C}$. $\mathcal{C}$ is said to be an $(n, n-r, \nu, d_{\text{free}})$ binary convolutional code. If we want to emphasize the distribution of the ν_is, the code will be referred to as an $(n, n-r, \underline{\nu}, d_{\text{free}})$ code, where $\underline{\nu} = (\nu_1, \ldots, \nu_r)$ is the constraint length vector.

Example 1 (Cont.) *The code $\mathcal{C}_1$ has constraint length vector (0,1,1,1) and constraint length 3. The free Hamming distance of $\mathcal{C}_1$ is, incidentally, 8. Thus, it is possible to find 8 columns of the associated doubly semi-infinite matrix $\mathbf{H}^*$ which are linearly dependent, but any 7 or less columns of $\mathbf{H}^*$ are linearly independent.*

Assumption. Without loss of generality, we can assume that $\nu_1 \leq \ldots \leq \nu_r$. For a particular code, the parity check matrix is not unique, but the sequence $\nu_1, \ldots, \nu_r$ is [1].

Definition. $N(r, (\nu_1, \ldots, \nu_r), d_{\text{free}})$ is the largest n such that there exists a convolutional code of block length n, dimension $n-r$, constraint length vector $(\nu_1, \ldots, \nu_r)$, and free distance d_{free}. Similarly, $N(r, \nu, d_{\text{free}}) = \max_{(\nu_1+\ldots+\nu_r=\nu)} N(r, (\nu_1, \ldots, \nu_r), d_{\text{free}})$. We assume that $d_{\text{free}} \geq 3$; then $N(r, \nu, d_{\text{free}})$ is finite.

Definition. An $(n, n-r, \underline{\nu}, d_{\text{free}})$ convolutional code $\mathcal{C}$ is *ascetic* if not all binary vectors in F_2^n can be found as edge labels in the trellis for $\mathcal{C}$. It is easy to see, and shown in [5], that the set of edge labels is a vector space of dimension $n - |\{i, 1 \leq i \leq r : \nu_i = 0\}|$. Thus, in particular, ascetic codes have $\nu_1 = 0$.

3. The Free Hamming distance of ascetic codes

In this section we discuss Hamming distance properties of ascetic codes, and compare these codes to non-ascetic codes.

The best known upper bounds on $N(r, \nu, d_{\mathrm{free}})$ were established by considering the block codes obtained by truncation of the convolutional code. The following theorem is a slight generalization ([6]) of a result appearing in [7].

Theorem 3.1 ([7]). *Let $\mathcal{C}$ be an $(n, n-r, \nu, d_{\mathrm{free}})$ binary convolutional code defined by the combined parity check matrix $\mathcal{H}$ with combined columns $\mathbf{h}_j, j = 1, \ldots, n$. Let $\varepsilon_j, 1 \le j \le n$, be defined by $\varepsilon_j = max\{s | h^{(i)}_{j,\nu_i - l} = 0$ for $1 \le i \le r, 0 \le l \le s\}$. Then*

$$d_{\mathrm{free}} \le \min \left\{ d(N, K) \mid N = mn + \sum_{j=1}^{n} \varepsilon_j,\ 1 \le K = N - (\nu + rm),\ m \ge 0 \right\}, \qquad (1)$$

where $d(N, K)$ is the largest minimum Hamming distance of any linear binary block code of block length N and dimension K. (Note: A table of bounds on $d(N, K)$ can be found in [8]).

Lemma 3.2 ([6]) $N(1, \nu, 4) = 2^{\nu-1}$.

Similarly, we have the following lemma which applies to codes with $r > 1$:

Lemma 3.3. *For $r \ge 2$, $N(r, \nu, 4) \le 2^{\nu+r-1}$ ([6]).*

Simple Proof. If there existed a $(2^{\nu+r-1} + 1, 2^{\nu+r-1} + 1 - r, \nu, 4)$ code, it could be truncated (as in the Heller bound) to a $[2^{\nu+r-1} + 1, 2^{\nu+r-1} + 1 - (\nu + r), 4]$ block code, which in turn could be shortened to a $[2^{\nu+r-1}, 2^{\nu+r-1} - (\nu + r), 3]$ block code. This contradicts the Hamming bound.

□

We observe that if a combined matrix contains a combined column $\mathbf{h}$ as well as some shift $\mathbf{h}^{\overset{l}{\rightarrow}}$ of $\mathbf{h}$, then the corresponding convolutional code will have minimum distance at most two. We will therefore avoid the use of combined columns $\mathbf{h} = (h^{(1)}_0, \ldots, h^{(1)}_{\nu_1}, \ldots, h^{(r)}_0, \ldots, h^{(r)}_{\nu_r})^T$ such that

$$h^{(i)}_0 = 0,\ \ \forall i \in \{1, \ldots, r\}. \qquad (2)$$

Theorem 3.4. *For $\nu \ge 1$ and $r \ge 1$, $N(r, \nu, 3) = (2^r - 1)2^\nu$.*

Proof. Wyner and Ash [9] noted that an $(n, n-1, \nu, 3)$ code can be constructed by selecting the parity check matrix $\mathbf{H}$ as the matrix consisting of all distinct $(\nu+1)$-dimensional vectors $\mathbf{h} = (h_0, \ldots, h_\nu)^T$ that have first entry $h_0 = 1$.
For $r > 1$, we can similarly choose $\mathcal{H}$ as the matrix in which the set of combined columns consists of all distinct $(\nu + r)$-dimensional vectors $\mathbf{h} = (h^{(1)}_0, \ldots, h^{(1)}_{\nu_1}, \ldots, h^{(r)}_0, \ldots, h^{(r)}_{\nu_r})$ except those on the form (2). The number of such vectors is $2^{\nu+r} - 2^\nu$.

□

Note that there is no restriction in this construction on the distribution of the ν_is. In particular, optimal codes of distance 3 can be constructed as “very ascetic”, i. e., $\nu_1 = \cdots = \nu_{r-1} = 0, \nu_r = \nu$.

Theorem 3.5. *For $\nu \ge 1$ and $r > 1$, $N(r, \nu, 4) = 2^{\nu+r-1}$.*

Proof. "$\leq$" was shown in Lemma 3.3.
"$\geq$:" A $(n, n-r, \nu, 4)$ code can be constructed as follows. Let the first element ν_1 of the constraint length vector be zero, the remainder of the constraint length vector can be chosen arbitrarily subject to $\sum_{j=2}^{r} \nu_j = \nu$. Select the combined parity check matrix $\mathcal{H}$ such that it consists of all distinct combined columns $\mathbf{h}$ that have "1" as their first entry. Since the columns are distinct the free distance is at least 3, and since the first row of $\mathcal{H}$ is the all-one vector and $\nu_1 = 0$, each code word has even Hamming weight in each block.

$\square$

We note that the codes described in Theorems 3.4 and 3.5 can be thought of as the convolutional code counterpart to the Hamming codes and its even weight subcodes.

Theorem 3.6 *For $\nu \geq 1$ and $r > 1$, non-ascetic codes of free Hamming distance 4 have length $n < 2^{\nu+r-1}$.*

Alternative Proof of Lemma 3.3 ([6]). Let $\mathcal{H}$ define an $(n, n-r, \nu, 4)$ binary convolutional code. Let $X_i, i = 1, \ldots, 2^r - 1$ be the set of combined columns $\mathbf{h}_j$ in $\mathcal{H}$ such that $(h_{j,0}^{(1)}, \ldots, h_0^{(r)})$ is the binary expansion of i, and let $x_i = |X_i|$. W. l. o. g., we can assume that $2^\nu \geq x_1 \geq x_i, 2 \leq i \leq 2^r - 1$. We first show that

$$\forall i : 1 \leq i \leq 2^{r-1} - 1 : \; x_{2i} + x_{2i+1} \leq 2^\nu. \tag{3}$$

If $x_{2i} = 0$, there is nothing to prove, so assume that $x_{2i} > 0$. Since the free distance is more than three, it follows that if $\mathbf{a} \in X_1$ and $\mathbf{b} \in X_{2i}$, then $(\mathbf{a} + \mathbf{b}) \notin X_{2i+1}$. There are x_1 choices for $\mathbf{a}$, hence $x_{2i+1} \leq 2^\nu - x_1$ implying (3). So

$$n = \sum_{j=1}^{2^r - 1} x_j = x_1 + \sum_{i=1}^{2^{r-1}-1} (x_{2i} + x_{2i+1}) \leq x_1 + \left(2^{r-1} - 1\right) 2^\nu \leq 2^{\nu+r-1}.$$

$\square$

Proof of Theorem 3.6 From the alternative proof of Lemma 3.3, it follows that if the block length of the code is $n < 2^{\nu+r-1}$, then $x_1 = 2^\nu$. Hence, the combined parity check matrix contains combined columns $\mathbf{x} = (1, 0, 0, 0, \ldots, 0)^T$ and $\mathbf{y} = (1, 1, 0, 0, \ldots, 0)^T$. If $\nu_1 > 0$, then $\mathbf{x} + \mathbf{y} + \mathbf{x}_0^{\overset{1}{\rightarrow}} = \mathbf{0}$, thus the minimum distance is at most 3.

$\square$

3.1. Computer search results

It is harder to *construct* convolutional codes of free Hamming distance larger than 4. This applies to ascetic as well as non-ascetic codes.

A computer search (described in [6]) has been used to find lower bounds (non-exhaustive "tabu" search) and exact values (exhaustive search) on $N(r, \nu, d_{\text{free}})$ and $N(r, \nu, d_{\text{free}})$ for $r \leq 4$ and moderate values of ν and d_{free}. Some results were presented in [10, 11]. As a general empirical observation, $N_A(r, \nu, d_{\text{free}}) = \max_{(\nu_1=0, \nu_2+\ldots+\nu_r=\nu)} N(r, (\nu_1, \ldots, \nu_r), d_{\text{free}})$ is usually close to or equal to $N(r, \nu, d_{\text{free}})$.

Sometimes surprisingly good "very ascetic" codes can be found.

Example 2. *The code $\mathcal{C}_1$ in Example 1 is ascetic and known to be optimal, in the sense that its block length is equal to $N(4,3,8)$. However, a more ascetic code exists. Let*

$$\mathbf{H}^{'} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ - & - & - & - & - & - & - & - \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ - & - & - & - & - & - & - & - \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ - & - & - & - & - & - & - & - \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

define the code $\mathcal{C}_2$. $\mathcal{C}_2$ is an (8,4,(0,0,0,3),8) code.

4. Applications

This section contains some applications of ascetic codes.

4.1. Zero-run lengths [12, 11]

In many communication and recording systems, the receiver derives symbol synchronization information from the transmitted or recorded sequence. If the symbol "zero" is represented as "absence of signal", a long zero-run (sequence of consecutive zeros) may cause the receiver to lose synchronization, and should be avoided. Hence, the maximum zero-run length, L_{max}, of a code is a design parameter which should be as small as possible.

Using a result by D. Forney, Hole [12] showed that for any coset of any convolutional code, $L_{max} \geq \nu_1$. Also, for any convolutional code there is at least one coset for which $L_{max} \leq (\nu_1 + 2)n - 2 - \epsilon$, where ϵ is a nonnegative integer depending on the code. Thus ascetic codes are preferable for these purposes, since $\nu_1 = 0$ and, thus, $L_{max} \leq 2n - 2 - \epsilon$. Note that this upper bound is independent on the constraint length ν. Further results on this problem was recently presented by Hole and Ytrehus [11].

4.2. Codes for precoded partial-response channels [5, 13, 14]

Wolf and Ungerboeck [15] suggested the use of cosets of convolutional code cosets for coding on precoded $1 - D$ partial response channels. In the absence of noise, on a precoded $1 - D$ channel, a binary zero input results in a zero output, while a "1" input results in a "+1" or a "-1" output, where the signs alternate. When Gaussian noise is added at the channel output, the relevant distance measure is the Euclidean squared distance, defined as $d^2(\mathbf{u}, \mathbf{v}) = \sum_i (u_i - v_i)^2$, where $\mathbf{u} = (u_0, u_1, \ldots)$ and $\mathbf{v} = (v_0, v_1, \ldots)$ are output sequences. Wolf and Ungerboeck noted that the free squared Euclidean distance of a code is lower bounded by its Hamming distance. They gave examples of codes, using $(n, n-1)$ convolutional codes with large free Hamming distance. However they were unable to provide examples with Euclidean distance larger than the lower bound.

The reason for this is that cosets of non-ascetic codes contain long zero-runs. It is easy to see that the Euclidean squared distance between output sequences $\mathbf{u} = (u_0, u_1, \ldots)$ and

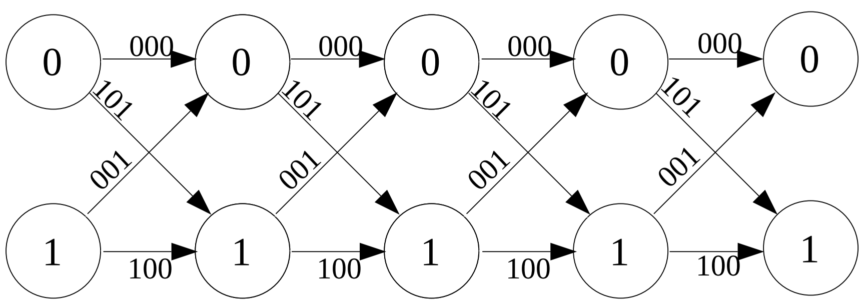
$\mathbf{0} = (0, 0, \ldots)$ is equal to $d^2(\mathbf{u}, \mathbf{0}) = \sum_i (u_i)^2 = \sum_i |u_i| =$ the Hamming distance between $\mathbf{u}$ and $\mathbf{0}$. On the other hand, many ascetic code cosets provide a free Euclidean squared distance which is much larger than the free Hamming distance of their corresponding codes.

Example 3.*Consider the code $\mathcal{C}_3$ defined by the parity check matrix*

$$\mathbf{H} = \begin{pmatrix} 0 & 1 & 0 \\ - & - & - \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

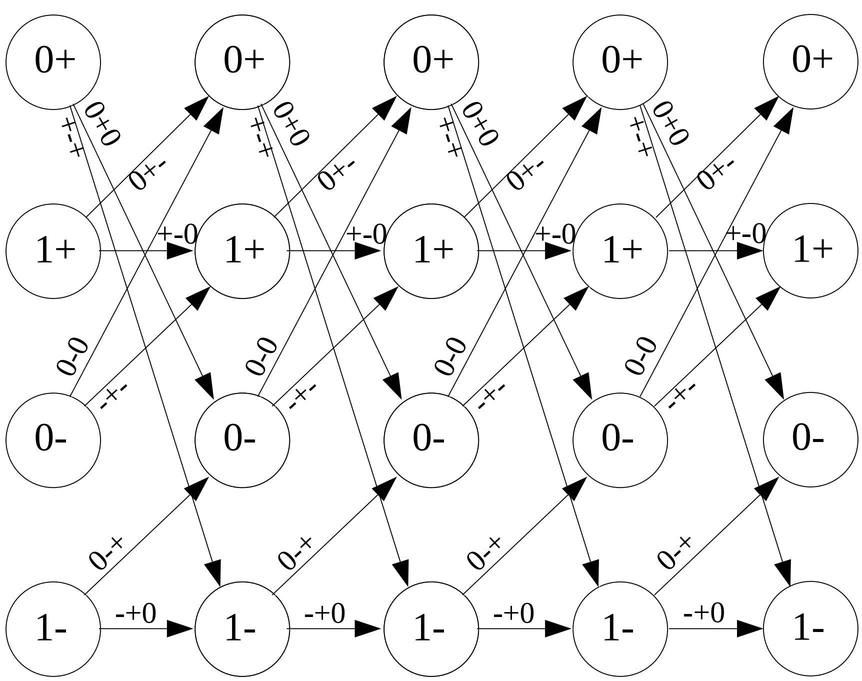
The trellis of $\mathcal{C}_3$ is shown in Figure 1. The code is poor with respect to free Hamming distance.

Figure 1. Trellis of $\mathcal{C}_3$



However, one of its cosets on the precoded $1 - D$ partial response channel can be described by the trellis in Figure 2. This code has free Euclidean squared distance of 14.

Figure 2. Trellis of $\mathcal{C}_3 + (010)$ on precoded partial response channel



More codes and details can be found in [5, 13, 14].

4.3. Simplified decoding [10, 16]

Hole and Ytrehus [10] presented a decoding algorithm for "complex" convolutional codes, such as PUM codes. For many codes, the number of operations required by an ordinary Viterbi decoder is dominated by calculating the branch metrics of the best edges between the states of the convolutional code trellis. The basic idea of the new algorithm is to decode in two steps: In step 1, all the edges are efficiently calculated by a special *coset trellis.* In step 2, the decoding results of step 1 are applied to a simplified convolutional code trellis.

McEliece and Lin [17] recently described a decoding algorithm. As a particular example they showed that decoding of the code $\mathcal{C}_1$ in Example 1 can be carried out with a total of 104 additions per information bit.

Ascetic codes are particularly suited for the decoding algorithm in [10]. The reason is that since few edge labels are used, each label is used more often. The code $\mathcal{C}_1$ can be decoded with 141 operations (112 additions and 29) comparisons per decoded bit. However, decoding of the more ascetic code $\mathcal{C}_2$ requires only 64 operations (44 additions and 22 comparisons) per decoded information bit.

Hole and Hole [16] recently described the application of this decoding algorithm on the precoded $1 - D$ partial-response channel.

5. Conclusion

Ascetic codes have been presented. These codes are attractive for several applications. For many sets of parameters, these codes also have excellent Hamming or Euclidean distance properties.

References

[1] P. Piret, *Convolutional Codes - An Algebraic Approach*. The MIT Press, 1988.

[2] J. L. Massey and M. K. Sain, “Inverses of linear sequential circuits,” *IEEE Trans. Comput.*, vol. C-17, pp. 330–337, April 1968.

[3] G. D. Forney, Jr., “Convolutional codes I: Algebraic structure,” *IEEE Trans. on Information Theory*, vol. IT-16, pp. 720–738, Nov. 1970.

[4] G. D. Forney, Jr., “Structural analysis of convolutional codes via dual codes,” *IEEE Trans. on Information Theory*, vol. IT-19, pp. 512–518, July 1973.

[5] K. J. Hole and Ø. Ytrehus, “Improved coding techniques for precoded partial-response channels,” *IEEE Trans. on Information Theory*, vol. IT-40, pp. 482–493, March 1994.

[6] Ø. Ytrehus, “A note on high rate convolutional codes,” Department report 68, Department of Informatics, University of Bergen, August 1992.

[7] J. A. Heller, “Sequential decoding: Short constraint length convolutional codes,” space programs summary 37-54, Jet Propul. Lab., Calif. Inst. Tech., Pasadena, Dec. 1968.

[8] A. E. Brouwer and T. Verhoeff, “An updated table of minimum-distance bounds for binary linear codes,” *IEEE Trans. on Information Theory*, vol. IT-39, pp. 662–677, March 1993.

[9] A. D. Wyner and R. B. Ash, “Analysis of recurrent codes,” *IEEE Trans. on Information Theory*, vol. IT-9, pp. 143–156, 1963.

[10] M. F. Hole and Ø. Ytrehus, “Two-step trellis decoding of partial unit memory convolutional codes.” Submitted for publication.

[11] K. J. Hole and Ø. Ytrehus, “Further results on cosets of convolutional codes with short maximum zero-run lengths,” in *Proceedings of ISIT’95*, p. 146, September 1995.

[12] K. J. Hole, “Cosets of convolutional codes with short maximum zero-run lengths,” *IEEE Trans. on Information Theory*, vol. IT-41, pp. 1145–1150, July 1995.

[13] K. J. Hole and Ø. Ytrehus, “Trellis codes for precoded partial-response channels: Further improved search techniques,” in *Proceedings of ISITA’94*, pp. 475–479, November 1994.

[14] K. J. Hole, Ø. Ytrehus, R. Erstad, and M. F. Hole, “Convolutional codes for partial-response channels,” in *Proceedings of Photonics East’95 Conference on Coding and Signal Processing for Information Storage*, 1995.

[15] J. K. Wolf and G. Ungerboeck, “Trellis coding for partial-response channels,” *IEEE Trans. on Communication*, vol. COM-34, pp. 765–773, Aug. 1986.

[16] M. F. Hole and K. J. Hole, “Low complexity decoding of high-rate partial unit memory codes on precoded partial-response channels.” Submitted for publication.

[17] R. J. McEliece and W. Lin, “The trellis complexity of convolutional codes,” in *Proceedings of ISIT’95*, p. 131, September 1995.