Active distances for convolutional codes

Höst, Stefan; Johannesson, Rolf; Zigangirov, Kamil; Zyablov, Viktor V.

Published in:
IEEE Transactions on Information Theory

DOI:
10.1109/18.749009

1999

Active Distances for Convolutional Codes

Stefan Höst, *Student Member, IEEE*, Rolf Johannesson, *Fellow, IEEE*,
Kamil Sh. Zigangirov, *Member, IEEE*, and Viktor V. Zyablov, *Associate Member, IEEE*

***Abstract*— A family of active distance measures for general convolutional codes is defined. These distances are generalizations of the extended distances introduced by Thommesen and Justesen for unit memory convolutional codes. It is shown that the error correcting capability of a convolutional code is determined by the active distances. The ensemble of periodically time-varying convolutional codes is defined and lower bounds on the active distances are derived for this ensemble. The active distances are very useful in the analysis of concatenated convolutional encoders.**

***Index Terms*—Active distances, cascaded convolutional codes, convolutional codes, extended distances.**

I. Introduction

THE column distance is often considered to be of fundamental importance when we study or construct convolutional codes [1], [2]. It has the well-known property that it will not increase any more when it has reached the free distance. In this paper we introduce a family of distances that stay "active" in the sense that we consider only those codewords which do not pass two consecutive zero encoder states. These distances determine the error correcting capability of the code and they are of particular importance when we consider concatenated convolutional encoders.

The active distances can be regarded as (nontrivial) generalizations to encoder memories $m > 1$ of the "extended" distances introduced for unit-memory convolutional codes by Thommesen and Justesen [3].

In Section II, we give definitions of the active distances for time-invariant convolutional codes. Some important properties of time-invariant convolutional codes are obtained via the active distances in Section III. After having introduced restricted sets of information sequences in Section IV, we define the active distances for the ensemble of periodically time-varying convolutional codes. Lower bounds on the active distances for the ensemble of periodically time-varying convolutional codes are derived in Section V. Finally, in Section VI we discuss various applications of the active distances.

Manuscript received December 3, 1996; revised March 13, 1998. This work was supported in part by the Royal Swedish Academy of Sciences in cooperation with the Russian Academy of Sciences and in part by the Swedish Research Council for Engineering Sciences under Grant 94-83. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Ulm, Germany, June 27–July 4, 1997.

S. Höst, R. Johannesson, and K. Sh. Zigangirov are with the Department of Information Technology, Lund University, S-221 00 Lund, Sweden (e-mail: {stefanh; rolf; kamil}@it.lth.se).

V. V. Zyablov is with the Institute for Problems of Information Transmission of the Russian Academy of Science, GSP-4, Moscow, 101447 Russia (e-mail: zyablov@iitp.su).

Communicated by N. Seshadri, Associate Editor for Coding Techniques.
Publisher Item Identifier S 0018-9448(99)01383-8.

II. Definitions of Active Distances for Time-Invariant Convolutional Codes

Consider a binary, rate $R = b/c$ convolutional code with a rational generator matrix $G(D)$ of memory m. The causal information sequence

$$\boldsymbol{u}(D) = \boldsymbol{u}_0 + \boldsymbol{u}_1 D + \boldsymbol{u}_2 D^2 + \cdots \tag{1}$$

is encoded as the causal codeword

$$\boldsymbol{v}(D) = \boldsymbol{v}_0 + \boldsymbol{v}_1 D + \boldsymbol{v}_2 D^2 + \cdots \tag{2}$$

where

$$\boldsymbol{v}(D) = \boldsymbol{u}(D)G(D). \tag{3}$$

For simplicity, we sometimes write $\boldsymbol{u} = \boldsymbol{u}_0\boldsymbol{u}_1\cdots$ and $\boldsymbol{v} = \boldsymbol{v}_0\boldsymbol{v}_1\cdots$ instead of $\boldsymbol{u}(D)$ and $\boldsymbol{v}(D)$, respectively. When we consider sequences of length $n+1$ we use the notation $\boldsymbol{x}_{[0,\,n]} = \boldsymbol{x}_0\boldsymbol{x}_1\cdots\boldsymbol{x}_n$.

Let the binary m-dimensional vector of b-tuples $\boldsymbol{\sigma}_t$ be the encoder state at depth t of a realization in controller canonical form of the generator matrix and let $\boldsymbol{\sigma}_t^{(i)} = (\sigma_{t1}^{(i)}\sigma_{t2}^{(i)}\cdots\sigma_{tb}^{(i)})$ be the b-tuple representing the contents of position i of the shift registers (counted from the input connections) (see Fig. 1). (When the jth constraint length $\nu_j < m$ for some j, then we set the jth component of $\boldsymbol{\sigma}_t^{(i)}$ to be 0.) Then we have $\boldsymbol{\sigma}_t = \boldsymbol{\sigma}_t^{(1)}\boldsymbol{\sigma}_t^{(2)}\cdots\boldsymbol{\sigma}_t^{(m)}$. To the information sequence $\boldsymbol{u} = \boldsymbol{u}_0\boldsymbol{u}_1\cdots$ corresponds the state sequence $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0\boldsymbol{\sigma}_1\cdots$.

Let $\mathcal{S}_{[t_1,\,t_2]}^{\boldsymbol{\sigma}_1,\,\boldsymbol{\sigma}_2}$ denote the set of state sequences $\boldsymbol{\sigma}_{[t_1,\,t_2]}$ that start at depth t_1 in state $\boldsymbol{\sigma}_1$ and terminate at depth t_2 in state $\boldsymbol{\sigma}_2$ and do not have two consecutive zero states in between, i.e.,

$$\mathcal{S}_{[t_1,\,t_2]}^{\boldsymbol{\sigma}_1,\,\boldsymbol{\sigma}_2} \stackrel{\text{def}}{=} \{\boldsymbol{\sigma}_{[t_1,\,t_2]} = \boldsymbol{\sigma}_{t_1} = \boldsymbol{\sigma}_1,\, \boldsymbol{\sigma}_{t_2} = \boldsymbol{\sigma}_2 \text{ and } \boldsymbol{\sigma}_i,\, \boldsymbol{\sigma}_{i+1} \text{ not both } = \boldsymbol{0},\, t_1 \le i < t_2\}. \tag{4}$$

Definition: Let $\mathcal{C}$ be a convolutional code encoded by a rational generator matrix $G(D)$ of memory m which is realized in controller canonical form. The *jth-order active row distance* is

$$a_j^r \stackrel{\text{def}}{=} \min_{\mathcal{S}_{[0,\,j+1]}^{\boldsymbol{0},\,\boldsymbol{\sigma}},\, \boldsymbol{\sigma}_{j+1+i}^{(1,\,i)} = \boldsymbol{0},\, 1 \le i \le m} \{w_{\mathrm{H}}(\boldsymbol{v}_{[0,\,j+m]})\} \tag{5}$$

where $\boldsymbol{\sigma}$ denotes any value of the state $\boldsymbol{\sigma}_{j+1}$ such that $\boldsymbol{\sigma}_{j+1}^{(1)} \neq \boldsymbol{0}$, and $\boldsymbol{\sigma}_{j+1+i}^{(1,\,i)}$ denotes the i first positions of the shift registers (counted from the input connections), i.e.,

$$\boldsymbol{\sigma}_{j+1+i}^{(1,\,i)} = \boldsymbol{\sigma}_{j+1+i}^{(1)}\boldsymbol{\sigma}_{j+1+i}^{(2)}\cdots\boldsymbol{\sigma}_{j+1+i}^{(i)}. \qquad \square$$

Let $\nu_{\min}$ be the minimum of the constraint lengths ν_i, $i = 1, 2, \cdots, b$, of the generator matrix $G(D)$ of memory m, i.e.,

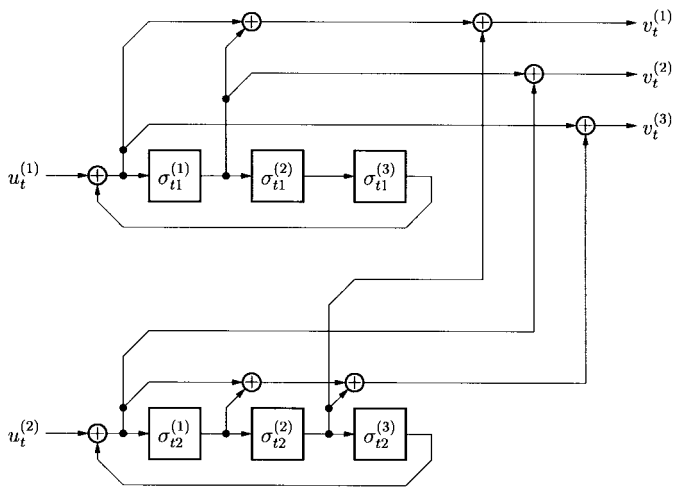

Fig. 1. The controller canonical form of the rate $R = 2/3$ generator matrix $G(D) = \begin{pmatrix} \frac{1+D}{1+D^3} & \frac{D}{1+D^3} & \frac{1}{1+D^3} \\ \frac{D^2}{1+D^3} & \frac{1}{1+D^3} & \frac{1+D+D^2}{1+D^3} \end{pmatrix}$.

$\nu_{\min} = \min_i \{\nu_i\}$ and $m = \max_i \{\nu_i\}$. Then the active row distance of order j is the minimum weight of paths that diverge from the zero state at depth 0, possibly "touches" the all-zero path only in nonconsecutive zero states at depth k, where $1 + \nu_{\min} \le k \le j$, and, finally, reemerges with the all-zero path at depth ℓ, where $j + 1 + \nu_{\min} \le \ell \le j + 1 + m$.

For a polynomial generator matrix realized in controller canonical form we have the following equivalent formulation:

$$a_j^r = \min_{\boldsymbol{u}_j \neq 0,\, \mathcal{S}_{[0,\,j+1]}^{\mathbf{0},\,\boldsymbol{\sigma}}} \{w_{\mathrm{H}}(\boldsymbol{u}_{[0,\,j]} \boldsymbol{G}_j^r)\} \tag{6}$$

where $\boldsymbol{\sigma}$ denotes any value of the state $\boldsymbol{\sigma}_{j+1}$ with $\boldsymbol{\sigma}_{j+1}^{(1)} = \boldsymbol{u}_j$ and

$$\boldsymbol{G}_j^r = \begin{pmatrix} G_0 & G_1 & \cdots & G_m & & & \\ & G_0 & G_1 & \cdots & G_m & & \\ & & \ddots & \ddots & & \ddots & \\ & & & G_0 & G_1 & \cdots & G_m \end{pmatrix} \tag{7}$$

is a $(j+1)\times(j+1+m)$ truncated version of the semi-infinite matrix

$$\boldsymbol{G} = \begin{pmatrix} G_0 & G_1 & \cdots & G_m & & \\ & G_0 & G_1 & \cdots & G_m & \\ & & \ddots & \ddots & & \ddots \end{pmatrix}. \tag{8}$$

Notice that the active row distance sometimes can decrease but, as we shall show in Section V, in the ensemble of convolutional codes encoded by periodically time-varying generator matrices there exists a convolutional code encoded by a generator matrix such that its active row distance can be lower-bounded by a linearly increasing function.

From the definition follows immediately

Triangle Inequality: Let $G(D)$ be a rational generator matrix with $\nu_{\min} = m$. Then its active row distance satisfies the triangle inequality

$$a_j^r \le a_i^r + a_{j-i-1-m}^r \tag{9}$$

where $j > i + m$ and the sum of the lengths of the paths to the right of the inequality is

$$i + m + 1 + (j - i - m - 1) + m + 1 = j + m + 1 \tag{10}$$

i.e., equal to the length of the path to the left of the inequality. □

Furthermore, we have immediately the following important

Theorem 1: Let $\mathcal{C}$ be a convolutional code encoded by a noncatastrophic generator matrix. Then

$$\min_j \{a_j^r\} = d_{\mathrm{free}}. \tag{11}$$

□

The following simple example shows that the triangle inequality (9) would not hold if we did not include state sequences that contain isolated inner zero states in the definition of $\mathcal{S}_{[t_1,\,t_2]}^{\boldsymbol{\sigma}_1,\,\boldsymbol{\sigma}_2}$.

Example 1: Consider the memory $m = 1$ encoding matrix

$$G(D) = (1 \quad D). \tag{12}$$

The code sequences corresponding to the state sequences $(0, 1, 0, 1, 0)$ and $(0, 1, 1, 1, 0)$ are $(10, 01, 10, 01)$ and $(10, 11, 11, 01)$, respectively. It is easily verified that $a_0^r = 2$, $a_1^r = 4$, and $a_2^r = 4$, which satisfy the triangle inequality

$$a_2^r \le a_0^r + a_0^r. \tag{13}$$

If we consider only state sequences without isolated inner zero states the lowest weight sequence of length four would pick up distance 6 and exceed the sum of the weight for the two length two sequence, which would still be four, in violation of the triangle inequality. □

Remark: If we consider the ensemble of periodically time-varying generator matrices $\boldsymbol{G}$ (or $G(D)$) to be introduced in Section IV and require that the corresponding code sequences consist of only randomly chosen code symbols (i.e., we do not allow transitions from the zero state to itself), then for a given length the set of state sequences defined by $\mathcal{S}_{[t_1,\,t_2]}^{\boldsymbol{\sigma}_1,\,\boldsymbol{\sigma}_2}$ is as large as possible.

Next we shall consider an "active" counterpart to the column distance.

Definition: Let $\mathcal{C}$ be a convolutional code encoded by a rational generator matrix $G(D)$ of memory m realized in controller canonical form. The *jth-order active column distance* is

$$a_j^c \stackrel{\text{def}}{=} \min_{\mathcal{S}_{[0,\,j+1]}^{\mathbf{0},\,\boldsymbol{\sigma}}} \{w_{\mathrm{H}}(\boldsymbol{v}_{[0,\,j]})\} \tag{14}$$

where $\boldsymbol{\sigma}$ denotes any encoder state. □

For a polynomial generator matrix we have the following equivalent formulation:

$$a_j^c = \min_{\mathcal{S}_{[0,\,j+1]}^{\mathbf{0},\,\boldsymbol{\sigma}}} \{w_{\mathrm{H}}(\boldsymbol{u}_{[0,\,j]}\boldsymbol{G}_j^c)\} \tag{15}$$

where $\boldsymbol{\sigma}$ denotes any encoder state and

$$\boldsymbol{G}_j^c = \begin{pmatrix} G_0 & G_1 & \cdots & G_m & & & \\ & G_0 & G_1 & \cdots & G_m & & \\ & & \ddots & \ddots & & \ddots & \\ & & & G_0 & G_1 & \cdots & G_m \\ & & & & G_0 & & G_{m-1} \\ & & & & & \ddots & \vdots \\ & & & & & & G_0 \end{pmatrix} \tag{16}$$

is a $(j+1)\times(j+1)$ truncated version of the semi-infinite matrix $\boldsymbol{G}$ given in (8).

It follows from the definitions that

$$a_j^c \le a_{j-k}^r \tag{17}$$

where $k \le \min\{j, \nu_{\min}\}$ and, in particular, if $\nu_{\min} = m \le j$, then

$$a_j^c \le a_{j-m}^r. \tag{18}$$

From (17) it follows that when $j \ge \nu_{\min}$ the active column distance of order j is upper-bounded by the active row distance of order $j-\nu_{\min}$, i.e., by the minimum weight of paths of length $j+1$ starting at a zero state and terminating at a zero state without passing consecutive zero states in between.

The active column distance a_j^c is a nondecreasing function of j but, as we shall show in Section V, in the ensemble of convolutional codes encoded by periodically time-varying generator matrices there exists a convolutional code encoded by a generator matrix such that its active column distance can be lower-bounded by a linearly increasing function.

Definition: Let $\mathcal{C}$ be a convolutional code encoded by a rational generator matrix $G(D)$ of memory m. The *jth-order active reverse column distance* is

$$a_j^{rc} \stackrel{\text{def}}{=} \min_{\mathcal{S}_{[m,\,m+j+1]}^{\boldsymbol{\sigma},\,\mathbf{0}}} \{w_{\mathrm{H}}(\boldsymbol{v}_{[m,\,j+m]})\}, \tag{19}$$

where $\boldsymbol{\sigma}$ denotes any encoder state. □

For a polynomial generator matrix we have the following equivalent formulation to (19):

$$a_j^{rc} = \min_{\mathcal{S}_{[m,\,m+j+1]}^{\boldsymbol{\sigma},\,\mathbf{0}}} \{w_{\mathrm{H}}(\boldsymbol{u}_{[0,\,j+m]}\boldsymbol{G}_j^{rc})\} \tag{20}$$

where $\boldsymbol{\sigma}$ denotes any encoder state and

$$\boldsymbol{G}_j^{rc} = \begin{pmatrix} G_m & & & \\ G_{m-1} & G_m & & \\ \vdots & G_{m-1} & \ddots & \\ G_0 & \vdots & & G_m \\ & G_0 & & G_{m-1} \\ & & \ddots & \vdots \\ & & & G_0 \end{pmatrix} \tag{21}$$

is a $(j+m+1)\times(j+1)$ truncated version of the semi-infinite matrix $\boldsymbol{G}$ given in (8).

The active reverse column distance a_j^{rc} is a nondecreasing function of j but, as we shall show in Section V, in the ensemble of convolutional codes encoded by periodically time-varying generator matrices there exists a convolutional code encoded by a generator matrix such that its active reverse column distance can be lower-bounded by a linearly increasing function.

Furthermore, the active reverse column distance of a polynomial generator matrix $G(D)$ is equal to the active column distance of the *reciprocal generator matrix*

$$\mathrm{diag}\,(D^{\nu_1}D^{\nu_2}\cdots D^{\nu_b})G(D^{-1}).$$

Definition: Let $\mathcal{C}$ be a convolutional code encoded by a rational generator matrix $G(D)$ of memory m. The *jth-order active segment distance* is

$$a_j^s \stackrel{\text{def}}{=} \min_{\mathcal{S}_{[m,\,m+j+1]}^{\boldsymbol{\sigma}_1,\,\boldsymbol{\sigma}_2}} \{w_{\mathrm{H}}(\boldsymbol{v}_{[m,\,j+m]})\} \tag{22}$$

where $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ denote any encoder states. □

For a polynomial generator matrix we have the following equivalent formulation:

$$a_j^s = \min_{\mathcal{S}_{[m,m+j+1]}^{\boldsymbol{\sigma}_1,\,\boldsymbol{\sigma}_2}} \{w_{\mathrm{H}}(\boldsymbol{u}_{[0,\,j+m]}\boldsymbol{G}_j^s)\} \tag{23}$$

where $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ denote any encoder states, and $\boldsymbol{G}_j^s = \boldsymbol{G}_j^{rc}$.

If we consider the segment distances for two sets of consecutive paths of lengths $i+1$ and $(j-i-1)+1$, respectively, then the terminating state of the first path is not necessarily identical to the starting state of the second path. Hence, the active segment distance for the set of paths of the total length $j+1$ does not necessarily satisfy the triangle inequality. However, we have immediately the following

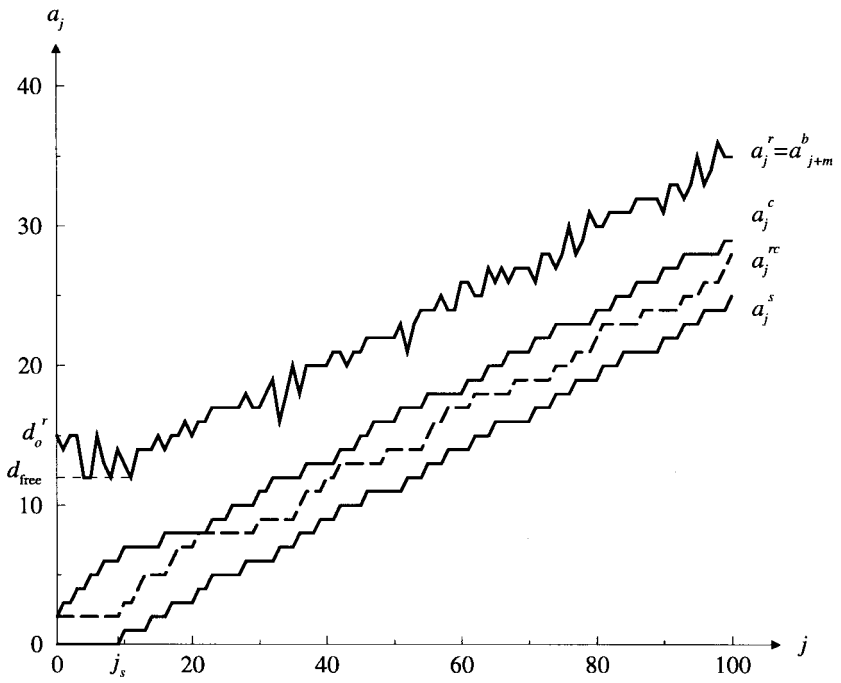

Fig. 2. The active distances for the encoding matrix in Example 2.

Theorem 2: Let $G(D)$ be a generator matrix of memory m. Then its active segment distance satisfies the inequality

$$a_j^s \geq a_i^s + a_{j-i-1}^s \tag{24}$$

where $j > i$ and the sum of the lengths of the paths to the right of the inequality is

$$i+1+j-i-1+1 = j+1 \tag{25}$$

i.e., equal to the length of the path to the left of the inequality. □

The active segment distance a_j^s is a nondecreasing function of j but, as we shall show in Section V, in the ensemble of convolutional codes encoded by periodically time-varying generator matrices there exists a convolutional code encoded by a generator matrix such that its active segment distance can be lower-bounded by a linearly increasing function.

The *start* of the active segment distance is the largest j for which $a_j^s = 0$ and is denoted j_s.

The jth-order active row distance is characterized by a fixed number of almost freely chosen information tuples, $j+1$, followed by a varying number, between $\nu_{\min}$ and m, of zero-state driving information tuples ("almost" since we have to avoid consecutive zero states $\boldsymbol{\sigma}_i\boldsymbol{\sigma}_{i+1}$ for $0 \leq i < j+1$ and assure that $\boldsymbol{\sigma}_{j+1}^{(1)} \neq \mathbf{0}$). Sometimes we find it useful to consider a corresponding distance between two paths of fixed total length, $j+1$, but with a varying number of almost freely chosen information tuples. Hence, we introduce the following (final) active distance.

Definition: Let $\mathcal{C}$ be a convolutional code encoded by a rational generator matrix $G(D)$ of memory m. The *jth-order active burst distance* is

$$a_j^b \stackrel{\text{def}}{=} \min_{S_{[0,\,j+1]}^{\mathbf{0},\,\mathbf{0}}} \{w_{\text{H}}(\boldsymbol{v}_{[0,\,j]})\} \tag{26}$$

where $j \geq \nu_{\min}$. □

For a polynomial generator matrix we have the following equivalent formulation:

$$a_j^b \stackrel{\text{def}}{=} \min_{S_{[0,\,j+1]}^{\mathbf{0},\,\mathbf{0}}} \{w_{\text{H}}(\boldsymbol{u}_{[0,\,j]}\boldsymbol{G}_j^c)\} \tag{27}$$

where $\boldsymbol{G}_j^c$ is given in (16).

The active row and burst distances are related via the following inequalities:

$$\begin{cases} a_j^b \geq \min\limits_i \{a_{j-\nu_i}^r\} \\ a_j^r \geq \min\limits_i \{a_{j+\nu_i}^b\}. \end{cases} \tag{28}$$

Clearly, when $\nu_{\min} = m$, we have

$$a_j^b = \begin{cases} \text{undefined}, & 0 \leq j < m \\ a_{j-m}^r, & j \geq m. \end{cases} \tag{29}$$

For a noncatastrophic generator matrix we have

$$\min_j \{a_j^b\} = d_{\text{free}}. \tag{30}$$

From the definition it follows that the active burst distance satisfies the triangle inequality.

Example 2: In Fig. 2 we show the active distances for the encoding matrix $G(D) = (1+D+D^2+D^3+D^7+D^8+D^9+D^{11} \quad 1+D^2+D^3+D^7+D^8+D^9+D^{11})$. Notice that the active row distance of the zeroth order, a_0^r, is identical to the row distance of the zeroth order, $d_0^r = 15$, which upper-bounds $d_{\text{free}} = 12$, and the start $j_s = 9$. □

From the definitions follow that the active distances are encoder properties, not code properties. However, it also follows that the active distances are invariant over the set of minimal-basic [4] (or canonical if rational) [5] encoding matrices for a code $\mathcal{C}$. Hence, when we in the sequel consider active distances for convolutional codes it is understood that

these distances are evaluated for the corresponding minimal-basic (canonical) encoding matrices.

III. Properties of Convolutional Codes Via the Active Distances

We define the *correct path* through a trellis to be the path determined by the encoded information sequence and we call the (encoder) states along the correct path *correct states*. Then we define an *incorrect segment* to be a segment starting in a correct state $\boldsymbol{\sigma}_{t_1}$ and terminating in a correct state $\boldsymbol{\sigma}_{t_2}$, $t_1 < t_2$, such that it differs from the correct path at some but not necessarily all states within this interval. Let $e_{[k,\ell)}$ denote the number of errors in the error pattern $\boldsymbol{e}_{[k,\ell)}$, where $\boldsymbol{e}_{[k,\ell)} = \boldsymbol{e}_k\boldsymbol{e}_{k+1}\cdots\boldsymbol{e}_{\ell-1}$.

For a convolutional code $\mathcal{C}$ with a generator matrix of memory m consider any incorrect segment between two arbitrary correct states, $\boldsymbol{\sigma}_{t_1}$ and $\boldsymbol{\sigma}_{t_2}$. A minimum-distance (MD) decoder can output an incorrect segment between $\boldsymbol{\sigma}_{t_1}$ and $\boldsymbol{\sigma}_{t_2}$ only if there exists a segment of length $j+1$ c-tuples, $\nu_{\min} \le j < t_2 - t_1$, between these two states such that the number of channel errors $e_{[t_1,t_2)}$ within this interval is at least $a_j^b/2$. Thus we have the following.

Theorem 3: A convolutional code $\mathcal{C}$ encoded by a rational generator matrix of memory m can correct all error patterns $\boldsymbol{e}_{[t_1,t_2)}$ that correspond to incorrect segments between any two correct states, $\boldsymbol{\sigma}_{t_1}$ and $\boldsymbol{\sigma}_{t_2}$, and satisfy

$$e_{[t_1+k,\,t_1+1+i)} < a_{i-k}^b/2 \tag{31}$$

for $0 \le k \le t_2 - t_1 - \nu_{\min} - 1$, $k + \nu_{\min} \le i \le t_2 - t_1 - 1$. □

We have immediately the following.

Corollary 4: A convolutional code $\mathcal{C}$ encoded by a rational generator matrix of memory m and smallest constraint length $\nu_{\min} = m$ can correct all error patterns $\boldsymbol{e}_{[t_1,t_2)}$ that correspond to incorrect segments between any two correct states, $\boldsymbol{\sigma}_{t_1}$ and $\boldsymbol{\sigma}_{t_2}$, and satisfy

$$e_{[t_1+k,\,t_1+1+i)} < a_{i-k-m}^r/2 \tag{32}$$

for $0 \le k \le t_2 - t_1 - m - 1$, $k + m \le i \le t_2 - t_1 - 1$. □

Both the active column distance and the active reverse column distance are important parameters when we study the error correcting capability of a convolutional code. A counterpart to Theorem 3 follows.

Theorem 5: Let $\mathcal{C}$ be a convolutional code encoded by a rational generator matrix of memory m and let $\boldsymbol{e}_{[t_1,t_2)}$ be an error sequence between the two correct states $\boldsymbol{\sigma}_{t_1}$ and $\boldsymbol{\sigma}_{t_2}$. A minimum-distance decoder will output a correct state $\boldsymbol{\sigma}_t$ at depth t, $t_1 < t < t_2$, if

$$\begin{cases} e_{[i,\,t)} < a_{t-i-1}^c/2, & t_1 \le i < t \\ e_{[t,\,j)} < a_{j-t-1}^{rc}/2, & t < j \le t_2. \end{cases} \tag{33}$$

□

Proof: Assume without loss of generality that the correct path is the all-zero path. The weight of any path of length $t - i$ diverging from the correct path at depth i, $i < t$, and not having two consecutive zero states is lower-bounded by a_{t-i-1}^c. Similarly, the weight of any path of length $j-t$, $j > t$, reemerging with the correct path at depth j and not having two consecutive zero states is lower-bounded by a_{j-t-1}^{rc}. Hence, if $e_{[i,\,t)} < a_{t-i-1}^c/2$ and $e_{[t,\,j)} < a_{j-t-1}^{rc}/2$, then $\boldsymbol{\sigma}_t$ must be correct. □

Since

$$a_{t-i-1}^c + a_{j-t-1}^{rc} \le a_{j-i-1}^b \tag{34}$$

it follows that we can regard Theorem 3 as a corollary to Theorem 5.

Example 3: Assume that the binary, rate $R = 1/2$, memory $m = 2$ convolutional encoding matrix

$$G(D) = (1 + D + D^2 \quad 1 + D^2)$$

is used to communicate over a binary-symmetric channel (BSC) and that we have the following error pattern:

$$\boldsymbol{e}_{[0,\,20)} = 10000100000000010000000010000000000100001 \tag{35}$$

or, equivalently,

$$\begin{aligned} \boldsymbol{e}_{[0,\,20)}(D) = {} & (10) + (01)D^2 + (01)D^7 \\ & + (10)D^{12} + (10)D^{17} + (01)D^{19}. \end{aligned} \tag{36}$$

The active distances for the encoding matrix is given in Fig. 3. From Theorem 3 it is easily seen that if we assume that $\boldsymbol{\sigma}_0$ is a correct state and that there exists a $t' \ge 20$ such that $\boldsymbol{\sigma}_{t'}$ is a correct state then, despite the fact that the number of channel errors $e_{[0,\,20)} = 6 > d_{\text{free}} = 5$, the error pattern (35) is corrected by a minimum-distance decoder. The error pattern

$$\boldsymbol{e}'_{[0,\,20)} = 10100100000000000000000000000000000101001 \tag{37}$$

or, equivalently,

$$\begin{aligned} \boldsymbol{e}'_{[0,\,20)}(D) = {} & (10) + (10)D + (01)D^2 \\ & + (10)D^{17} + (10)D^{18} + (01)D^{19} \end{aligned} \tag{38}$$

contains also six channel errors but with a different distribution; we have three channel errors in both the prefix and suffix 101001. Since $\nu_{\min} = m = 2$ and the active row distance $a_0^r = 5$, the active burst distance $a_2^b = 5$; hence, Theorem 3 does not imply that the error pattern (37) is corrected by a minimum-distance decoder; the states $\boldsymbol{\sigma}_1$, $\boldsymbol{\sigma}_2$, $\boldsymbol{\sigma}_{18}$, and $\boldsymbol{\sigma}_{19}$ will in fact be erroneous states. However, from Theorem 5 follows that if $\boldsymbol{\sigma}_0$ is a correct state and if there exists a $t' \ge 20$ such that $\boldsymbol{\sigma}_{t'}$ is a correct state, then at least $\boldsymbol{\sigma}_{10}$ is also a correct state. □

We will now study the set of code sequences corresponding to encoder state sequences that do not contain two consecutive zero states. From the properties of the active segment distance it follows that such code sequences can contain at most $j_s + 1$ zero c-tuples, where j_s is the start of the segment distance. Lower bounds on the number of nonzero code symbols between two bursts of zeros are given in the following.

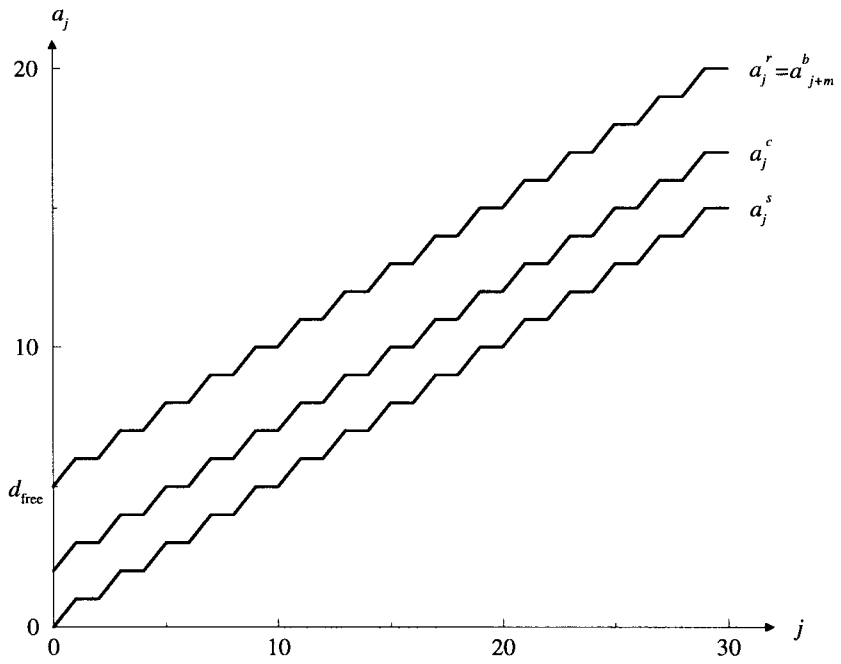

Fig. 3. The active distances for the encoding matrix in Example 3.

Theorem 6: Consider a binary, rate $R = b/c$ convolutional code and let $\boldsymbol{v}_{[0,\,j']}^{c}$, $\boldsymbol{v}_{[0,\,j']}^{rc}$, and $\boldsymbol{v}_{[m,\,j'+m]}^{s}$ denote code sequences corresponding to state sequences in $\mathcal{S}_{[0,\,j'+1]}^{\mathbf{0},\boldsymbol{\sigma}}$, $\mathcal{S}_{[0,\,j'+1]}^{\boldsymbol{\sigma},\mathbf{0}}$, and $\mathcal{S}_{[m,\,m+j'+1]}^{\boldsymbol{\sigma}_1,\boldsymbol{\sigma}_2}$, respectively, where $\boldsymbol{\sigma}$, $\boldsymbol{\sigma}_1$, and $\boldsymbol{\sigma}_2$ denote any encoder states.

i) Let w_j^c denote the number of ones in (the weight of) a code sequence $\boldsymbol{v}_{[0,\,j']}^{c}$ counted from the beginning of the code sequence to the first burst of j consecutive zero c-tuples. Then w_j^c satisfies

$$w_j^c \geq a_{j+\lceil w_j^c/c\rceil-1}^{c}. \tag{39}$$

ii) Let w_j^{rc} denote the number of ones in (the weight of) a code sequence $\boldsymbol{v}_{[0,\,j']}^{rc}$ counted from the last burst of j consecutive zero c-tuples to the end of the code sequence. Then w_j^{rc} satisfies

$$w_j^{rc} \geq a_{j+\lceil w^{rc}/c\rceil-1}^{rc}. \tag{40}$$

iii) Let $w_{j_1,\,j_2}^{s}$ denote the number of ones in (the weight of) a code sequence $\boldsymbol{v}_{[m,\,j'-m]}^{s}$ counted between any two consecutive bursts of j_1 and j_2 consecutive zero c-tuples, respectively. Then $w_{j_1,\,j_2}^{s}$ satisfies

$$w_{j_1,\,j_2}^{s} \geq a_{j_1+j_2+\lceil w_{j_1,\,j_2}^{s}/c\rceil-1}^{s}. \tag{41}$$

□

Proof:

i) The subsequence up to the beginning of the first burst of j consecutive zero c-tuples consists of at least $\lceil w_j^c/c\rceil$ c-tuples. Thus the length of the subsequence that includes the first burst of j consecutive zero c-tuples is at least $j+\lceil w_j^c/c\rceil$ c-tuples and, hence, w_j^c must satisfy (39).

ii) Analogously to the proof of i).

iii) Since $w_{j_1,\,j_2}^{s}$ is the weight of the subsequence between the two bursts of j_1 and j_2 consecutive zeros, respectively, the total length including these bursts of zeros is at least $j_1+\lceil w_{j_1,\,j_2}^{s}/c\rceil+j_2$. Clearly, the weight of a subsequence of this length is lower-bounded by the corresponding active segment distance, which completes the proof. □

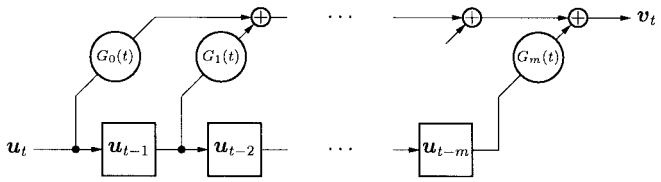

Fig. 4. A general time-varying polynomial convolutional encoder.

IV. Active Distances for Time-Varying Convolutional Codes

So far we have considered only *time-invariant* or *fixed* convolutional codes, i.e., convolutional codes encoded by time-invariant generator matrices. When it is too difficult to analyze the performance of a communication system using time-invariant convolutional codes, we can often obtain powerful results if we study time-varying convolutional codes instead.

Assuming polynomial generator matrices we have

$$\boldsymbol{v}_t = \boldsymbol{u}_t G_0 + \boldsymbol{u}_{t-1} G_1 + \cdots + \boldsymbol{u}_{t-m} G_m \tag{42}$$

where G_i, $0 \leq i \leq m$, is a binary $b \times c$ time-invariant matrix.

In general, a rate $R = b/c$, binary convolutional code can be *time-varying.* Then (42) becomes

$$\boldsymbol{v}_t = \boldsymbol{u}_t G_0(t) + \boldsymbol{u}_{t-1} G_1(t) + \cdots + \boldsymbol{u}_{t-m} G_m(t) \tag{43}$$

where $G_i(t)$, $i = 0, 1, \cdots, m$, is a binary $b \times c$ time-varying matrix. In Fig. 4 we illustrate a general time-varying polynomial convolutional encoder. As a counterpart to the

$$\boldsymbol{G}_t = \begin{pmatrix} G_0(t) & G_1(t+1) & \cdots & G_m(t+m) & & \\ & G_0(t+1) & G_1(t+2) & \cdots & G_m(t+1+m) & \\ & & \ddots & \ddots & & \ddots \end{pmatrix}. \tag{44}$$

semi-infinite matrix $\boldsymbol{G}$ given in (8) we have (44) at the top of this page.

Remark: With a slight abuse of terminology we call for simplicity a time-varying polynomial transfer function matrix a *generator matrix* although it might not have full rank.

We have the general *ensemble of binary, rate* $R = b/c$*, time-varying convolutional codes* with generator matrices of memory m in which each digit in each of the matrices $G_i(t)$ for $0 \le i \le m$ and $t = 0, 1, 2, \cdots$ is chosen independently and is equally likely to be 0 and 1.

As a special case of the ensemble of time-varying convolutional codes we have the ensemble of binary, rate $R = b/c$*, periodically* time-varying convolutional codes encoded by a polynomial generator matrix $\boldsymbol{G}_t$ (44) of memory m and *period* T, in which each digit in each of the matrices $G_i(t) = G_i(t+T)$ for $0 \le i \le m$ and $t = 0, 1, \cdots, T-1$, is chosen independently and is equally likely to be 0 and 1. We denote this ensemble $\mathcal{E}(b, c, m, T)$.

Before we define the active distances for periodically time-varying convolutional codes encoded by time-varying polynomial generator matrices we introduce the following sets of information sequences, where we always assume that $t_1 \le t_2$.

Let $\mathcal{U}^r_{[t_1-m,\,t_2+m]}$ denote the set of information sequences $\boldsymbol{u}_{t_1-m}\boldsymbol{u}_{t_1-m+1}\cdots\boldsymbol{u}_{t_2+m}$ such that the first m and the last m subblocks are zero and such that they do not contain $m+1$ consecutive zero subblocks, i.e.,

$$\begin{aligned}\mathcal{U}^r_{[t_1-m,\,t_2+m]} \stackrel{\text{def}}{=} \{&\boldsymbol{u}_{[t_1-m,\,t_2+m]} | \boldsymbol{u}_{[t_1-m,\,t_1-1]} = \boldsymbol{0} \\ &\boldsymbol{u}_{[t_2+1,\,t_2+m]} = \boldsymbol{0}, \text{ and } \boldsymbol{u}_{[i,\,i+m]} \ne \boldsymbol{0},\ t_1 - m \le i \le t_2\}.\end{aligned} \tag{45}$$

Let $\mathcal{U}^c_{[t_1-m,\,t_2]}$ denote the set of information sequences $\boldsymbol{u}_{t_1-m}\boldsymbol{u}_{t_1-m+1}\cdots\boldsymbol{u}_{t_2}$ such that the first m subblocks are zero and such that they do not contain $m+1$ consecutive zero subblocks, i.e.,

$$\begin{aligned}\mathcal{U}^c_{[t_1-m,\,t_2]} \stackrel{\text{def}}{=} \{&\boldsymbol{u}_{[t_1-m,\,t_2]} | \boldsymbol{u}_{[t_1-m,\,t_1-1]} = \boldsymbol{0}, \\ &\text{and } \boldsymbol{u}_{[i,\,i+m]} \ne \boldsymbol{0},\ t_1 - m \le i \le t_2 - m\}.\end{aligned} \tag{46}$$

Let $\mathcal{U}^{rc}_{[t_1-m,\,t_2+m]}$ denote the set of information sequences $\boldsymbol{u}_{t_1-m}\boldsymbol{u}_{t_1-m+1}\cdots\boldsymbol{u}_{t_2+m}$ such that the last m subblocks are zero and such that they do not contain $m+1$ consecutive zero subblocks, i.e.,

$$\begin{aligned}\mathcal{U}^{rc}_{[t_1-m,\,t_2+m]} \stackrel{\text{def}}{=} \{&\boldsymbol{u}_{[t_1-m,\,t_2+m]} | \boldsymbol{u}_{[t_2+1,\,t_2+m]} = \boldsymbol{0}, \\ &\text{and } \boldsymbol{u}_{[i,\,i+m]} \ne \boldsymbol{0},\ t_1 - m < i \le t_2\}.\end{aligned} \tag{47}$$

Let $\mathcal{U}^s_{[t_1-m,\,t_2]}$ denote the set of information sequences $\boldsymbol{u}_{t_1-m}\boldsymbol{u}_{t_1-m+1}\cdots\boldsymbol{u}_{t_2}$ such that they do not contain $m+1$ consecutive zero subblocks, i.e.,

$$\mathcal{U}^s_{[t_1-m,\,t_2]} \stackrel{\text{def}}{=} \{\boldsymbol{u}_{[t_1-m,\,t_2]} | \boldsymbol{u}_{[i,\,i+m]} \ne \boldsymbol{0},\ t_1 - m < i < t_2 - m\}. \tag{48}$$

Next we introduce the $(j+m+1) \times (j+1)$ truncated, periodically time-varying generator matrix of memory m and period T

$$\boldsymbol{G}_{[t,\,t+j]} = \begin{pmatrix} G_m(t) & & & \\ G_{m-1}(t) & G_m(t+1) & & \\ \vdots & G_{m-1}(t+1) & \ddots & \\ G_0(t) & \vdots & \ddots & G_m(t+j) \\ & G_0(t+1) & & G_{m-1}(t+j) \\ & & \ddots & \vdots \\ & & & G_0(t+j) \end{pmatrix} \tag{49}$$

where $G_i(t) = G_i(t+T)$ for $0 \le i \le m$.

We are now well-prepared to generalize the definitions of the active distances for convolutional codes encoded by polynomial generator matrices to time-varying convolutional codes encoded by polynomial time-varying generator matrices:

Definition: Let $\mathcal{C}$ be a periodically time-varying convolutional code encoded by a periodically time-varying polynomial generator matrix of memory m and period T.

The *jth-order active row distance* is

$$a^r_j \stackrel{\text{def}}{=} \min_t \min_{\mathcal{U}^r_{[t-m,\,t+j+m]}} \{w_{\mathrm{H}}(\boldsymbol{u}_{[t-m,\,t+j+m]}\boldsymbol{G}_{[t,\,t+j+m]})\}. \tag{50}$$

The *jth-order active column distance* is

$$a^c_j \stackrel{\text{def}}{=} \min_t \min_{\mathcal{U}^c_{[t-m,\,t+j]}} \{w_{\mathrm{H}}(\boldsymbol{u}_{[t-m,\,t+j]}\boldsymbol{G}_{[t,\,t+j]})\}. \tag{51}$$

The *jth-order active reverse column distance* is

$$a^{rc}_j \stackrel{\text{def}}{=} \min_t \min_{\mathcal{U}^{rc}_{[t-m,\,t+j]}} \{w_{\mathrm{H}}(\boldsymbol{u}_{[t-m,\,t+j]}\boldsymbol{G}_{[t,\,t+j]})\}. \tag{52}$$

The *jth-order active segment distance* is

$$a^s_j \stackrel{\text{def}}{=} \min_t \min_{\mathcal{U}^s_{[t-m,\,t+j]}} \{w_{\mathrm{H}}(\boldsymbol{u}_{[t-m,\,t+j]}\boldsymbol{G}_{[t,\,t+j]})\}. \tag{53}$$

□

For a periodically time-varying convolutional code encoded by a periodically time-varying, noncatastrophic, polynomial generator matrix with active row distance a^r_j we define its free distance by a generalization of (11)

$$d_{\text{free}} \stackrel{\text{def}}{=} \min_j \{a^r_j\}. \tag{54}$$

In the following section, we will derive lower bounds on the active distances. There we need the following.

Theorem 7: Consider a periodically time-varying, rate $R = b/c$, polynomial generator matrix of memory m and period T represented by $\boldsymbol{G}_t$, where $\boldsymbol{G}_t$ is given in (44).

i) Let the information sequences be restricted to the set $\mathcal{U}^r_{[t-m,\,t+j+m]}$. Then the code symbols in the segment $\boldsymbol{v}_{[t,\,t+j+m]}$ are mutually independent and equiprobable over the ensemble $\mathcal{E}(b,\,c,\,m,\,T)$ for all j, $0 \le j < T$.
ii) Let the information sequences be restricted to the set $\mathcal{U}^c_{[t-m,\,t+j]}$. Then the code symbols in the segment $\boldsymbol{v}_{[t,\,t+j]}$ are mutually independent and equiprobable over the ensemble $\mathcal{E}(b,\,c,\,m,\,T)$ for all j, $0 \le j < \max\{m+1,\,T\}$.
iii) Let the information sequences be restricted to the set $\mathcal{U}^{rc}_{[t-m,\,t+j]}$. Then the code symbols in the segment $\boldsymbol{v}_{[t,\,t+j]}$ are mutually independent and equiprobable over the ensemble $\mathcal{E}(b,\,c,\,m,\,T)$ for all j, $0 \le j < \max\{m+1,\,T\}$.
iv) Let the information sequences be restricted to the set $\mathcal{U}^s_{[t-m,\,t+j]}$. Then the code symbols in the segment $\boldsymbol{v}_{[t,\,t+j]}$ are mutually independent and equiprobable over the ensemble $\mathcal{E}(b,\,c,\,m,\,T)$ for all j, $0 \le j < T$. □

Proof: It follows immediately that for $0 \le j < T$ the code tuples $\boldsymbol{v}_i$, $i = t,\, t+1,\, \cdots,\, t+j$, are mutually independent and equiprobable in all four cases. Hence, the proof of iv) is complete. In cases ii) and iii) it remains to show that the statements hold also for $T \le j \le m$ when $m \ge T$.

ii) Consider the information sequences in the set $\mathcal{U}^c_{[t-m,\,t+j]}$, where $0 \le j \le m$. Let $t \le i \le t+j$, then, in the expression

$$\boldsymbol{v}_i = \boldsymbol{u}_i G_0(i) + \boldsymbol{u}_{i-1} G_1(i) + \cdots + \boldsymbol{u}_{i-m} G_m(i) \tag{55}$$

there exists a k, $0 \le k \le m$, such that at least one of the b-tuples $\boldsymbol{u}_{i-k}$ is nonzero and all the previous b-tuples $\boldsymbol{u}_{i-k'}$, $k < k' \le m$, are zero. Hence, $\boldsymbol{v}_i$ and $\boldsymbol{v}_{i'}$, $t \le i < i' \le t+j$, are mutually independent and equiprobable. This completes the proof of ii).

iii) Consider the information sequences in the set $\mathcal{U}^{rc}_{[t-m,\,t+j]}$, where $0 \le j \le m$. Let $t \le i \le t+j$, then, in (55) at least one of the b-tuples $\boldsymbol{u}_{i-k}$, $0 \le k \le m$, is nonzero and all the following b-tuples $\boldsymbol{u}_{i-k'}$, $0 \le k' < k$, are zero. Hence, $\boldsymbol{v}_i$ and $\boldsymbol{v}_{i'}$, $t \le i < i' \le t+j$, are mutually independent and equiprobable.

i) For the information sequences in $\mathcal{U}^r_{[t-m,\,t+j+m]}$ it remains to show that $\boldsymbol{v}_i$ and $\boldsymbol{v}_{i'}$ are mutually independent and equiprobable also for $T \le i' - i < T+m$. From the definition of $\mathcal{U}^r_{[t-m,\,t+j+m]}$ it follows that $\boldsymbol{u}_{[t-m,\,t-1]} = \boldsymbol{0}$, $\boldsymbol{u}_t \ne \boldsymbol{0}$, $\boldsymbol{u}_{t+j} \ne \boldsymbol{0}$, and $\boldsymbol{u}_{[t+j+1,\,t+j+m]} = \boldsymbol{0}$. For $j = T$, we can choose, e.g.,

$$\boldsymbol{u}_{[t-m,\,t+m]} = \boldsymbol{u}_{[t+T-m,\,t+T+m]} \in \mathcal{U}^r_{[t-m,\,t+T+m]}$$

which implies that $\boldsymbol{v}_{[t,\,t+m]} = \boldsymbol{v}_{[t+T,\,t+T+m]}$. However, for $T - m \le j < T$, $\boldsymbol{v}_i$, $t \le i < t+m$, and $\boldsymbol{v}_{i'}$, $t+j < i' \le t+j+m$, are mutually independent and equiprobable. □

From Theorem 7 follows immediately.

Corollary 8: Consider a rate $R = b/c$ polynomial generator matrix of memory m represented by $\boldsymbol{G}$, where $\boldsymbol{G}$ is given in (8).

i) Let the information sequences be restricted to the set $\mathcal{U}^c_{[t-m,\,t+j]}$. Then the code symbols in the segment $\boldsymbol{v}_{[t,\,t+j]}$ are mutually independent and equiprobable over the ensemble $\mathcal{E}(b,\,c,\,m,\,1)$ for all j, $0 \le j \le m$.
ii) Let the information sequences be restricted to the set $\mathcal{U}^{rc}_{[t-m,\,t+j+m]}$. Then the code symbols in the segment $\boldsymbol{v}_{[t,\,t+j]}$ are mutually independent and equiprobable over the ensemble $\mathcal{E}(b,\,c,\,m,\,1)$ for all j, $0 \le j \le m$. □

V. Lower Bounds on the Active Distances for Time-Varying Convolutional Codes

In this section we shall derive lower bounds on the active distances for the ensemble of periodically time-varying convolutional codes. First we consider the active row distance and begin by proving the following.

Lemma 9: Consider the ensemble $\mathcal{E}(b,\,c,\,m,\,T)$ of binary, rate $R = b/c$, periodically time-varying convolutional codes encoded by polynomial generator matrices of memory m. The fraction of convolutional codes in this ensemble whose jth-order active row distance a^r_j, $0 \le j < T$, satisfies

$$a^r_j \le \hat{a}^r_j < (j+m+1)c/2 \tag{56}$$

does not exceed

$$T 2^{([(j+1)/(j+m+1)]R + h(\hat{a}^r_j/(j+m+1)c) - 1)(j+m+1)c}$$

where $h(\ )$ is the binary entropy function. □

Proof: Let

$$\boldsymbol{v}_{[t,\,t+j+m]} = \boldsymbol{u}_{[t-m,\,t+j+m]} \boldsymbol{G}_{[t,\,t+j+m]} \tag{57}$$

where $\boldsymbol{u}_{[t-m,\,t+j+m]} \in \mathcal{U}^r_{[t-m,\,t+j+m]}$ and assume that

$$\hat{a}^r_j < (j+m+1)c/2. \tag{58}$$

Then, it follows from Theorem 7 that

$$\begin{aligned} &P(w_{\mathrm{H}}(\boldsymbol{v}_{[t,\,t+j+m]}) \le \hat{a}^r_j) \\ &\quad = \sum_{i=0}^{\hat{a}^r_j} \binom{(j+m+1)c}{i} 2^{-(j+m+1)c} \\ &\quad < 2^{(h(\hat{a}^r_j/(j+m+1)c)-1)(j+m+1)c}, \qquad 0 \le j < T-m \end{aligned} \tag{59}$$

where the last inequality follows from the standard inequality

$$\sum_{i=0}^{k} \binom{n}{i} < 2^{h(k/n)n},\ k \le n/2. \tag{60}$$

(Notice that we need the denominator "2" in the right inequality in (60) in order to be able to apply inequality (60).) Using

$$2^{(j+1)b} = 2^{(j+1)Rc} \tag{61}$$

as an upper bound on the cardinality of $\mathcal{U}^r_{[t-m,\,t+j+m]}$, we have

$$P\left(\min_{\mathcal{U}^r_{[t-m,\,t+j+m]}}\{w_{\mathrm{H}}(\boldsymbol{v}_{[t,\,t+j+m]})\}\le \hat{a}^r_j\right) < 2^{(j+1)Rc}2^{(h(\hat{a}^r_j/(j+m+1)c)-1)(j+m+1)c} = 2^{([(j+1)/(j+m+1)]R+h(\hat{a}^r_j/(j+m+1)c)-1)(j+m+1)c} \quad (62)$$

for each t, $0\le t<T$. Using the union bound completes the proof. □

For a given f, $0\le f<1$, let j_0 be the smallest integer j satisfying the inequality

$$\left(1-\frac{j+1}{j+m+1}R\right)(j+m+1)c\ge \log\frac{T^2}{1-f}. \quad (63)$$

For large memories m such a value always exists. Let $\hat{a}^r_j$

$$0<\hat{a}^r_j<(j+m+1)c/2 \quad (64)$$

denote the largest integer that for given f, $0\le f<1$, and j, $j\ge j_0$, satisfies the inequality

$$\left(\frac{j+1}{j+m+1}R+h\left(\frac{\hat{a}^r_j}{(j+m+1)c}\right)-1\right)(j+m+1)c \le -\log\frac{T^2}{1-f}. \quad (65)$$

Then, from Lemma 9 follows that for each j, $j_0\le j<T$, the fraction of convolutional codes with jth-order active row distance satisfying (56) is upper-bounded by

$$T2^{-\log[T^2/(1-f)]}=\frac{1-f}{T}. \quad (66)$$

Hence, we use the union bound and conclude that the fraction of convolutional codes with active row distance $a^r_j\le\hat{a}^r_j$ for at least one j, $j_0\le j<T$, is upper-bounded by

$$\sum_{j=j_0}^{T-m-1}\frac{1-f}{T}<1-f. \quad (67)$$

Thus we have proved the following.

Lemma 10: In the ensemble $\mathcal{E}(b, c, m, T)$ of periodically time-varying convolutional codes, the fraction of codes with active row distance

$$a^r_j>\hat{a}^r_j,\qquad j_0\le j<T \quad (68)$$

is larger than f, where for a given f, $0\le f<1$, j_0 is the smallest integer satisfying (63) and $\hat{a}^r_j$ the largest integer satisfying (65). □

By taking $f=0$, we have immediately

Corollary 11: There exists a binary, periodically time-varying, rate $R=b/c$, convolutional code encoded by a polynomial generator matrix of period T and memory m such that its jth-order active row distance for $j_0\le j<T$ is lower-bounded by $\hat{a}^r_j$, where $\hat{a}^r_j$ is the largest integer satisfying

$$\left(\frac{j+1}{j+m+1}R+h\left(\frac{\hat{a}^r_j}{(j+m+1)c}\right)-1\right)(j+m+1)c \le -2\log T \quad (69)$$

and j_0 is the smallest integer satisfying

$$\left(1-\frac{j+1}{j+m+1}R\right)(j+m+1)c\ge 2\log T. \quad (70)$$

□

In order to get a better understanding of the significance of the previous lemma we shall study the asymptotical behavior of the parameters j_0 and $\hat{a}^r_j$ for large memories.

Let the period T grow as a power of m greater than one; choose $T=m^2$, say. Then, since j_0 is an integer, for large values of m we have $j_0=0$. Furthermore, the inequality (69) can be rewritten as

$$h\left(\frac{\hat{a}^r_j}{(j+m+1)c}\right)\le 1-\frac{j+1}{j+m+1}R+O\left(\frac{\log m}{m}\right) \quad (71)$$

or, equivalently, as[1]

$$\hat{a}^r_j\le h^{-1}\left(1-\frac{j+1}{j+m+1}R\right)(j+m+1)c+O(\log m). \quad (72)$$

Finally, we have proved

Theorem 12: There exists a binary, periodically time-varying, rate $R=b/c$, convolutional code encoded by a polynomial generator matrix of memory m that has a jth-order active row distance satisfying the inequality

$$a^r_j>h^{-1}\left(1-\frac{j+1}{j+m+1}R\right)(j+m+1)c+O(\log m), \quad (73)$$

for $j\ge 0$. □

The main term in (73) can also be obtained from the Gilbert–Varshamov bound for block codes using a geometrical construction that is similar to Forney's inverse concatenated construction [6].

Consider Gilbert–Varshamov's lower bound on the (normalized) minimum distance for block codes [7], viz.

$$\frac{d_{\min}}{N}\ge h^{-1}(1-R) \quad (74)$$

where N denotes the blocklength. Let

$$\hat{\delta}^r(j)=\frac{h^{-1}\left(1-\frac{j+1}{j+1+m}R\right)(j+1+m)c}{mc} \quad (75)$$

denote the main term of the right-hand side of (73) normalized by mc.

The construction is illustrated in Fig. 5 for $R=1/2$. The straight line between the points $(0,\hat{\delta}^r(j))$ and $(R,0)$ intersects $h^{-1}(1-R)$ in the point $(r, h^{-1}(1-r))$. The rate r is chosen to be

$$r=\frac{j+1}{j+1+m}R \quad (76)$$

i.e., it divides the line between $(0,0)$ and $(R,0)$ in the proportion $(j+1):m$. Then we have

$$\frac{\hat{\delta}^r(j)}{h^{-1}(1-r)}=\frac{j+1+m}{m} \quad (77)$$

[1] Here and hereafter we write $h^{-1}(y)$ for the *smallest* x such that $y=h(x)$.

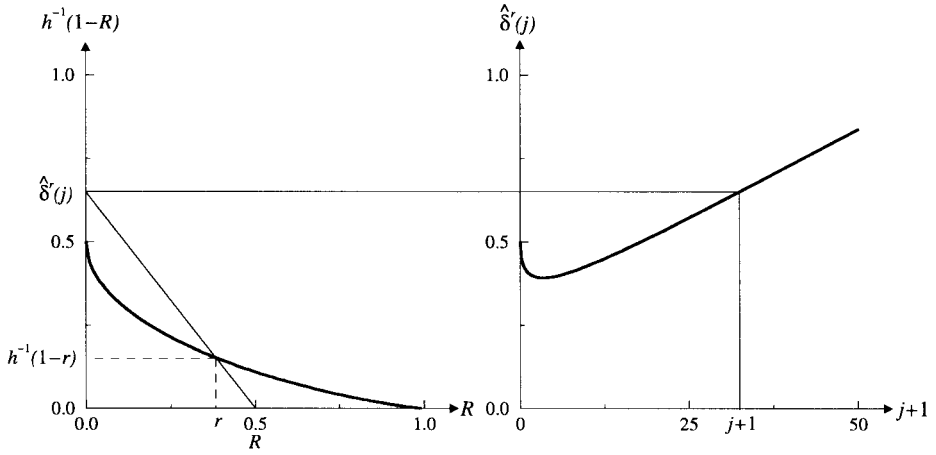

Fig. 5. Geometrical construction of the relationship between the lower bound on the active row distance for convolutional codes and the Gilbert–Varshamov lower bound on the minimum distance for block codes.

which is equivalent to (75). The relationship between r and j in Fig. 5 is given by (76).

We shall now derive a corresponding lower bound on the active column distance. Let

$$\boldsymbol{v}_{[t,\,t+j]} = \boldsymbol{u}_{[t-m,\,t+j]}\boldsymbol{G}_{[t,\,t+j]} \tag{78}$$

where $\boldsymbol{u}_{[t-m,\,t+j]} \in \mathcal{U}^c_{[t-m,\,t+j]}$ and let $\hat{a}^c_j$ be an integer satisfying the inequality

$$\hat{a}^c_j < (j+1)c/2. \tag{79}$$

Then, as a counterpart to (59) we have

$$P(w_{\mathrm{H}}(\boldsymbol{v}_{[t,\,t+j]}) \le \hat{a}^c_j) = \sum_{i=0}^{\hat{a}^c_j} \binom{(j+1)c}{i} 2^{-(j+1)c} < 2^{(h(\hat{a}^c_j/(j+1)c)-1)(j+1)c}, \qquad 0 \le j < T. \tag{80}$$

We use (61) as an upper bound on the cardinality of $\mathcal{U}^c_{[t-m,\,t+j]}$ and obtain

$$P\left(\min_{\mathcal{U}^c_{[t-m,\,t+j]}} \{w_{\mathrm{H}}(\boldsymbol{v}_{[t,\,t+j]})\} \le \hat{a}^c_j\right) < 2^{(j+1)Rc} 2^{(h(\hat{a}^c_j/(j+1)c)-1)(j+1)c} = 2^{(R+h(\hat{a}^c_j/(j+1)c)-1)(j+1)c} \tag{81}$$

for each t, $0 \le t < T$. Minimizing over $0 \le t < T$ and using the union bound complete the proof of the following.

Lemma 13: Consider the ensemble $\mathcal{E}(b, c, m, T)$ of binary, rate $R = b/c$, periodically time-varying convolutional codes encoded by polynomial generator matrices of memory m. The fraction of convolutional codes in this ensemble whose jth-order active column distance a^c_j, $0 \le j < T$, satisfies

$$a^c_j \le \hat{a}^c_j < (j+1)c/2 \tag{82}$$

does not exceed

$$T2^{(R+h(\hat{a}^c_j/(j+1)c)-1)(j+1)c}. \qquad \square$$

Next we choose j_0 to be the smallest integer j satisfying the inequality

$$(1-R)(j+1)c \ge \log T^2. \tag{83}$$

Let $\hat{a}^c_j$

$$0 < \hat{a}^c_j < (j+1)c/2 \tag{84}$$

denote the largest integer that for given j, $j \ge j_0$, satisfies the inequality

$$\left(R + h\left(\frac{\hat{a}^c_j}{(j+1)c}\right) - 1\right)(j+1)c \le -\log T^2. \tag{85}$$

Then, from Lemma 13 follows that for each j, $j_0 \le j < T$, the fraction of convolutional codes with a jth-order active column distance satisfying (84) is upper-bounded by

$$T2^{-\log T^2} = \frac{1}{T}. \tag{86}$$

Hence, we use the union bound and conclude that the fraction of convolutional codes with active column distance $a^c_j \le \hat{a}^c_j$ for at least one j, $j_0 \le j < T$, is upper-bounded by

$$\sum_{j=j_0}^{T-1} \frac{1}{T} < 1. \tag{87}$$

Thus we have proved the following.

Lemma 14: There exists a periodically time-varying, rate $R = b/c$, convolutional code encoded by a polynomial generator matrix of period T and memory m such that its jth-order active column distance for $j_0 \le j < T$ is lower-bounded by $\hat{a}^c_j$, where $\hat{a}^c_j$ is the largest integer satisfying

$$\left(R + h\left(\frac{\hat{a}^c_j}{(j+1)c}\right) - 1\right)(j+1)c \le -2\log T \tag{88}$$

and j_0 is the smallest integer satisfying

$$(1-R)(j+1)c \ge 2\log T. \tag{89}$$

$\square$

If, as before, we choose $T = m^2$, then $j_0 = O(\log m)$, and the inequality (88) can be rewritten as

$$h\left(\frac{\hat{a}^c_j}{(j+1)c}\right) \le 1 - R - \frac{4\log m}{(j+1)c} \tag{90}$$

for $j = O(m)$ or, equivalently, as

$$\hat{a}_j^c \le h^{-1}(1-R)(j+1)c + O(\log m). \tag{91}$$

Thus we have proved

Theorem 15: There exists a binary, periodically time-varying, rate $R = b/c$, convolutional code encoded by a polynomial generator matrix of memory m that has a jth-order active column distance satisfying the inequality

$$a_j^c > \rho(j+1)c + O(\log m) \tag{92}$$

for $j = O(m) > j_0 = O(\log m)$ and $\rho = h^{-1}(1-R)$ is the Gilbert–Varshamov parameter. □

Analogously we can prove

Theorem 16: There exists a binary, periodically time-varying, rate $R = b/c$, convolutional code encoded by a polynomial generator matrix of memory m that has a jth-order active reverse column distance a_j^{rc} which is lower-bounded by the right-hand side of the inequality (92) for all $j > j_0 = O(\log m)$. □

For the active segment distance we have the following.

Theorem 17: There exists a binary, periodically time-varying, rate $R = b/c$, convolutional code encoded by a polynomial generator matrix of memory m that has a jth-order active segment distance satisfying the inequality

$$a_j^s > h^{-1}\left(1 - \frac{j+m+1}{j+1}R\right)(j+1)c + O(\log m) \tag{93}$$

for $j = O(m) > j_s$, where

$$j_s < \frac{R}{1-R}m + O(\log m). \tag{94}$$

□

Proof: Consider the ensemble $\mathcal{E}(b, c, m, T)$. First we notice that the cardinality of $\mathcal{U}_{[t,\,t+j]}^s$ is upper-bounded by

$$2^{mb}2^{(j+1)b} = 2^{(j+m+1)Rc}. \tag{95}$$

Using (95) instead of (61) and repeating the steps in the derivation of the lower bound on the active column distance will give

$$h\left(\frac{\hat{a}_j^s}{(j+1)c}\right) \le 1 - \frac{j+m+1}{j+1}R - \frac{4\log m}{(j+1)c} \tag{96}$$

for all $j = O(m) > j_s$, or, equivalently,

$$\hat{a}_j^s \le h^{-1}\left(1 - \frac{j+m+1}{j+1}R\right)(j+1)c + O(\log m) \tag{97}$$

where

$$O < \hat{a}_j^s < (j+1)c/2 \tag{98}$$

instead of (90), (91), and (84), respectively, and the proof is complete. □

The parameter j_s is the start of the active segment distance (cf. Fig. 2).

For the ensemble of periodically time-varying convolutional code the active burst and active row distances are related through (29). Hence, we do not lower-bound the active burst distance separately.

Next we consider our lower bounds on the active distances, viz., (73), (92), and (93), and introduce the substitution

$$\ell = (j+1)/m \tag{99}$$

then we obtain asymptotically—for large memories m—the following lower bounds on the *normalized active distances*.

Theorem 18:

i) There exists a binary, periodically time-varying, rate $R = b/c$, convolutional code encoded by a polynomial generator matrix of memory m whose normalized active row distance asymptotically satisfies

$$\delta_\ell^r \stackrel{\text{def}}{=} \frac{a_j^r}{mc} \ge h^{-1}\left(1 - \frac{\ell}{\ell+1}R\right)(\ell+1) + O\left(\frac{\log m}{m}\right) \tag{100}$$

for $\ell \ge 0$.

ii) There exists a binary, periodically time-varying, rate $R = b/c$, convolutional code encoded by a polynomial generator matrix of memory m whose normalized active column distance (active reverse column distance) asymptotically satisfies

$$\delta_\ell^c \stackrel{\text{def}}{=} \frac{a_j^c}{mc}\left\{\delta_\ell^{rc} \stackrel{\text{def}}{=} \frac{a_j^{rc}}{mc}\right\} \ge h^{-1}(1-R)\ell + O\left(\frac{\log m}{m}\right) \tag{101}$$

for $\ell \ge \ell_0 = O(\log m/m)$.

iii) There exists a binary, periodically time-varying, rate $R = b/c$, convolutional code encoded by a polynomial generator matrix of memory m whose normalized active segment distance asymptotically satisfies

$$\delta_\ell^s \stackrel{\text{def}}{=} \frac{a_j^s}{mc} \ge h^{-1}\left(1 - \frac{\ell+1}{\ell}R\right)\ell + O\left(\frac{\log m}{m}\right) \tag{102}$$

for

$$\ell \ge \ell_s = \frac{R}{1-R} + O\left(\frac{\log m}{m}\right). \qquad \square$$

The typical behavior of the bounds in Theorem 18 is shown in Fig. 6. Notice that by minimizing the lower bound on the normalized active row distance (100) we obtain nothing but the main term in Costello's lower bound on the free distance [8], viz.,

$$\frac{R}{-\log(2^{1-R}-1)}.$$

VI. Comments

In this paper we have introduced a family of active distances for convolutional codes and shown that the error correcting capability of the code is to a large extent determined by these distances.

In [9] we used the active row distance to lower-bound the probability of the output error burst lengths for Viterbi decoding of periodically time-varying convolutional codes. From these lower bounds on the error burst lengths follow

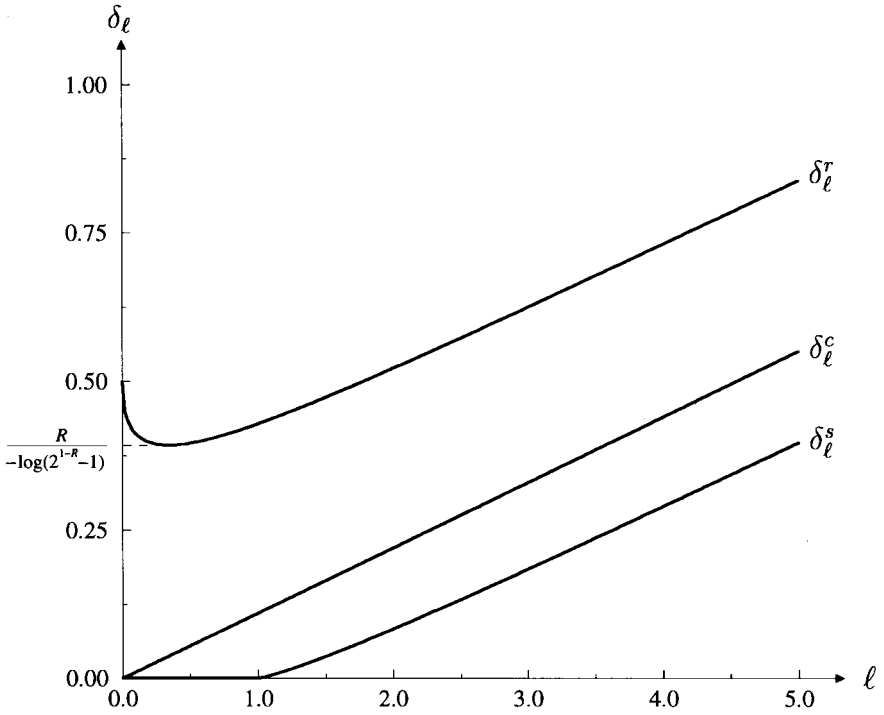

Fig. 6. Typical behavior of the lower bounds on the normalized active distances of Theorem 18.

easily the well-known upper error probability bounds for periodically time-varying convolutional codes [10].

The active distances were used in [11] to determine the free distances of two different constructions of binary concatenated convolutional codes, viz., woven convolutional codes with outer and inner warp, respectively. Both constructions have large free distances.

Concatenation is a both powerful and practical method to obtain constructions that are attractive for use in communication situations where very low error probabilities are needed. The simplest concatenated scheme with two convolutional encoders is a cascade without an interleaver but with matched rates, i.e., the outer convolutional code has rate $R_o = b_o/c_o$ and the inner convolutional code has rate $R_i = b_i/c_i$, where $b_i = c_o$. In [12] we have shown the existence of cascaded convolutional codes in the ensemble of periodically time-varying cascaded convolutional codes that have active distances with lower bounds similar to those derived in this paper. From the lower bound on the active row distance for the cascade it is shown that given only a restriction on the memory of the inner code, there exists a convolutional code, obtained as a simple cascade, with a free distance satisfying the Costello lower bound.

REFERENCES

[1] D. J. Costello, Jr., "A construction technique for random-error-correcting convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 631–636, 1969.

[2] J. L. Massey, "Coding and modulation in digital communications," in *Proc. Int. Zurich Seminar on Digital Communications*, 1974, pp. E2(1)–E2(4).

[3] C. Thommesen and J. Justesen, "Bounds on distances and error exponents of unit-memory codes," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 637–649, Sept. 1983.

[4] R. Johannesson and Z.-x. Wan, "A linear algebra approach to minimal convolutional encoders," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1219–1233, July 1993.

[5] G. D. Forney, Jr., R. Johannesson, and Z.-x. Wan, "Minimal and canonical rational generator matrices for convolutional codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1865–1880, Nov. 1996.

[6] G. D. Forney, Jr., "Convolutional codes II: Maximum-likelihood decoding," *Inform. Contr.*, vol. 25, pp. 222–250, 1974.

[7] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.

[8] D. J. Costello, Jr., "Free distance bounds for convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 356–365, July 1974.

[9] S. Höst, R. Johannesson, D. K. Zigangirov, K. Sh. Zigangirov, and V. V. Zyablov, "On the distribution of the output error burst lengths for Viterbi decoding of convolutional codes," in *Proc. IEEE Int. Symp. Information Theory* (Ulm, Germany, June 24–July 4, 1997), p. 108.

[10] R. Johannesson and K. Sh. Zigangirov, *Fundamentals of Convolutional Coding*. Piscataway, NJ: IEEE Press, Feb. 1999.

[11] S. Höst, R. Johannesson, and V. V. Zyablov, "A first encounter with binary woven convolutional codes," in *Proc. 4th Int. Symp. Communication Theory and Applications*, (Lake Districts, U.K., July 13–18), 1997.

[12] S. Höst, R. Johannesson, V. R. Sidorenko, K. Sh. Zigangirov, and V. V. Zyablov, "Cascaded convolutional codes," in *Communication and Coding*, M. Darnell and B. Honary, Eds. London/New York: Research Studies Press Ltd. and Wiley, 1998, pp. 10–29, in honor of P. G. Farrell on the occasion of his 60th birthday.