On the error exponent for woven convolutional codes with outer warps

Zyablov, Viktor V.; Shavgulidze, Sergo; Skopintsev, Oleg; Höst, Stefan; Johannesson, Rolf

Published in:
IEEE Transactions on Information Theory

DOI:
10.1109/18.771237

1999

On the Error Exponent for Woven Convolutional Codes with Outer Warp

Victor Zyablov, *Associate Member, IEEE*, Sergo Shavgulidze, Oleg Skopintsev, Stefan Höst, and Rolf Johannesson, *Fellow, IEEE*

***Abstract*—In this correspondence the error exponent and the decoding complexity of binary woven convolutional codes with outer warp and with binary convolutional codes as outer and inner codes are studied. It is shown that an error probability that is exponentially decreasing with the product of the outer and inner code memories can be achieved with a nonexponentially increasing decoding complexity.**

***Index Terms*—Concatenated convolutional codes, decoding complexity, woven convolutional codes, woven error exponents.**

I. Introduction

Binary woven convolutional codes were introduced in [1], where it was shown that such codes have large free distance and, therefore, appear to be attractive for use in communication situations where low error probabilities are needed. In this correspondence we investigate the probabilistic characteristics (the behavior of the error exponent) of woven convolutional codes with outer warp.

The same problem for similar constructions was considered in [2] and [3]. In [2], we considered concatenated codes with many inner binary block codes and many outer nonbinary convolutional codes and in [3] concatenated codes with many inner binary unit-memory convolutional codes and many outer Reed–Solomon codes.

A peculiarity of the codes discussed in this paper is that they are completely based on binary convolutional codes which significantly simplifies their decoding procedure.

The woven convolutional codes with outer warp are briefly described in Section II. In Section III it is shown that the decoding error probability decreases exponentially with the product of the outer and inner code memories and that at the same time the decoding complexity increases exponentially only with the square root of the same product. We show that the value of the error exponent is nonzero in the whole range of rates between zero and the channel capacity. Viterbi decoding of convolutional codes yields a better error exponent but does not allow a nonexponential growth (with memory) of the decoding complexity.

Manuscript received March 19, 1998; revised October 26, 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 97-723. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Cambridge, MA, August 16–21, 1998.

V. Zyablov was with the Department of Information Technology, Information Theory Group, Lund University, P.O. Box 118, S-22100 Lund, Sweden. He is now with the Institute for Problems of Information Transmission, Russian Academy of Science, GSP-4, Moscow, 101447 Russia (e-mail: zyablov@iitp.ru).

S. Shavgulidze was with the Department of Information Technology, Information Theory Group, Lund University, P.O. Box 118, S-22100 Lund, Sweden. He is now with the Department of Digital Communication Theory, Georgian Technical University, Tbilisi, 380075 Georgia (e-mail: sergo_130@hotmail.com).

O. Skopintsev is with the Institute for Problems of Information Transmission, Russian Academy of Science, GSP-4, Moscow, 101447 Russia (e-mail: skopintsev@orc.ru).

S. Höst and R. Johannesson are with the Department of Information Technology, Information Theory Group, Lund University, P.O. Box 118, S-22100 Lund, Sweden (e-mail: stefanh@it.lth.se; rolf@it.lth.se).

Communicated by E. Soljanin, Associate Editor for Coding Techniques.

Publisher Item Identifier S 0018-9448(99)04173-5.

II. Woven Convolutional Codes with Outer Warp

Consider the block diagram of a communication system with *woven convolutional codes with outer warp* [1] given in Fig. 1.

The encoding structure is given in Fig. 2. It consists of l_o rate $R_o = b_o/c_o$, outer, binary convolutional encoders, all of memory m_o and one rate $R_i = b_i/c_i$, inner, binary convolutional encoder of memory m_i.

The information sequence is divided into subblocks of $l_o b_o$ information symbols each. These subblocks are fed into l_o parallel outer encoders. The c_o code sequences from each encoder are serialized and written row-wise into a buffer consisting of l_o rows—the *warp*. The binary code symbols are read column-wise and used as inputs to the inner encoder—the *woof*.

The woven convolutional code with outer warp has the overall rate $R_w = b/c$, where $b = l_o b_o$ and $c = l_o c_o / R_i$. Hence we have

$$R_w = R_o R_i. \tag{1}$$

III. Error Exponent for Woven Convolutional Codes with Outer Warp

We focus on the case when the codewords of the woven convolutional codes with outer warp are transmitted over the binary symmetric channel (BSC). Consider the decoding procedure where we first carry out the hard-decisions Viterbi decoding of the inner code. Then the estimated information symbols of the inner code are fed into the l_o parallel outer decoders (see Fig. 1), where we again use hard-decisions Viterbi decoding. The estimated information symbols from all outer decoders are delivered as the output of the woven communication system.

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 $m+1$ consequtive zero subblocks, i.e.,

$$\begin{aligned}\mathcal{U}^r_{[t_1-m,t_2+m]} &\stackrel{\text{def}}{=} \big\{\boldsymbol{u}_{[t_1-m,t_2+m]} \,\big|\, \boldsymbol{u}_{[t_1-m,t_1-1]} = \boldsymbol{0},\\ &\qquad \boldsymbol{u}_{[t_2+1,t_2+m]} = \boldsymbol{0}, \text{ and } \boldsymbol{u}_{[i,i+m]} \neq \boldsymbol{0}, t_1 - m \le i \le t_2\big\}.\end{aligned} \tag{2}$$

The decoder output error sequence $\boldsymbol{e}_{[t-m-1,t+j+m+1]}$ is called an *error burst of length* $j+1$ starting at time t and ending at time $t+j+1$, if

$$\boldsymbol{e}_{t-m-1} = \boldsymbol{0}, \boldsymbol{e}_{[t-m,t+j+m]} \in \mathcal{U}^r_{[t-m,t+j+m]}$$

and $\boldsymbol{e}_{t+j+m+1} = \boldsymbol{0}$.

Consider periodically time-varying, rate $R = b/c$ convolutional codes encoded by polynomial, periodically time-varying generator matrices of memory m and period T. Assume that the transmitted sequence is the all-zero sequence. In order to upper-bound the distribution of the length of an error burst starting at time t we

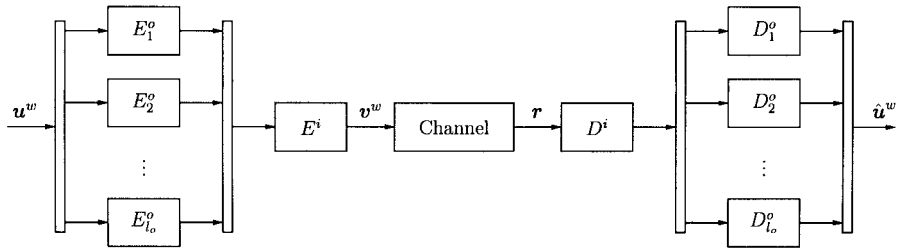

Fig. 1. Block diagram of a woven coded system with outer warp.

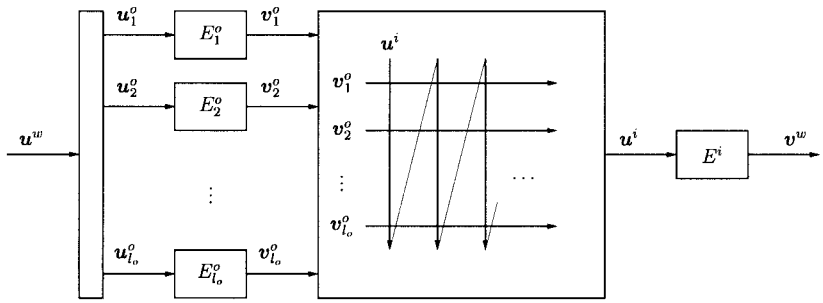

Fig. 2. Encoding structure of woven convolutional codes with outer warp.

consider the block code $\mathcal{B}_t(j)$, where

$$\mathcal{B}_t(j) = \left\{ \boldsymbol{v}_{[t,t+j+m]} \,\middle|\, \boldsymbol{v}_{[t,t+j+m]} = \boldsymbol{0} \text{ or } \boldsymbol{v}_{[t,t+j+m]} = \boldsymbol{u}_{[t-m,t+j+m]} \boldsymbol{G}_{[t,t+j+m]} \right\} \tag{3}$$

where

$$\boldsymbol{u}_{[t-m,t+j+m]} \in \mathcal{U}^r_{[t-m,t+j+m]} \text{ and}$$
$$\boldsymbol{G}_{[t,t+j+m]} = \begin{pmatrix} G_m(t) & & & \\ G_{m-1}(t) & G_m(t+1) & & \\ \vdots & G_{m-1}(t+1) & \ddots & \\ G_0(t) & \vdots & \ddots & G_m(t+j+m) \\ & G_0(t+1) & & G_{m-1}(t+j+m) \\ & & \ddots & \vdots \\ & & & G_0(t+j+m) \end{pmatrix} \tag{4}$$

is a $(j+2m+1) \times (j+m+1)$ truncated, time-varying generator matrix and where $G_i(t), i = 0, 1, \cdots$, are binary $b \times c$ time-varying matrices.

The rate of the block code $\mathcal{B}_t(j)$ is upper-bounded by

$$r(j) = \frac{j+1}{j+m+1} R. \tag{5}$$

(This is an *upper bound* since we have imposed a restriction on $\boldsymbol{u}_{[t-m,t+j+m]}$.)

Let $\mathcal{L}_t(j)$ denote the event that an error burst starting at time t has length $j+1$. A necessary—but not sufficient—condition for $\mathcal{L}_t(j)$ to occur is that the block code $\mathcal{B}_t(j)$ will be erroneously decoded. Thus we have

$$P(\mathcal{L}_t(j)) \le P\left(\mathcal{E}_t(j) \,\middle|\, \boldsymbol{u}_{[t-m,t+j+m]} = \boldsymbol{0}\right) \tag{6}$$

where $\mathcal{E}_t(j)$ denotes the event that $\mathcal{B}_t(j)$ is erroneously decoded.

In [4] we showed the existence of convolutional codes for which the probability of an error burst of length $j+1$ is upper-bounded by the inequality

$$P(\mathcal{L}_t(j)) \le 2^{-(E_G(r(j))+o(1))(j+m+1)c} \tag{7}$$

where $E_G(*)$ is the Gallager error exponent for block codes used over the BSC [5]:

$$E_G(r) = \begin{cases} -\rho \log_2(2\sqrt{p(1-p)}), & 0 \le r < R_{\exp} \\ 1 - \log_2(1 + 2\sqrt{p(1-p)}) - r = R_{\text{comp}} - r, & R_{\exp} \le r < R_{\text{crit}} \\ \rho \log_2 \frac{\rho}{p} + (1-\rho) \log_2 \frac{1-\rho}{1-p}, & R_{\text{crit}} \le r < C \end{cases} \tag{8}$$

where

$$\rho = h^{-1}(1-r) \tag{9}$$

is the Gilbert–Varshamov parameter, $h(*)$ is the binary entropy function, R_{comp} is the computational cutoff rate, and p is the crossover probability for the BSC. The critical rate R_{crit} and the expurgation rate $R_{\exp}$ are defined as

$$R_{\text{crit}} = 1 - h\left(\frac{\sqrt{p}}{\sqrt{p} + \sqrt{1-p}}\right) \tag{10}$$

and

$$R_{\exp} = 1 - h\left(\frac{2\sqrt{p(1-p)}}{1 + 2\sqrt{p(1-p)}}\right) \tag{11}$$

respectively. For a detailed proof of (7) the reader is referred to [6]. It is well known that the Gallager error exponent $E_G(r)$ is a lower bound on the reliability exponent for a block code of rate r.

It is convenient to introduce the normalized error burst length $\ell = (j+1)/m$. Then, from (7) follows that [6]

$$P(\mathcal{L}_t(j)) \le 2^{-(L(\ell)-\varepsilon)mc}, \qquad \varepsilon > 0 \tag{12}$$

where

$$L(\ell) = E_G\left(\frac{\ell}{\ell+1} R\right)(\ell+1) \tag{13}$$

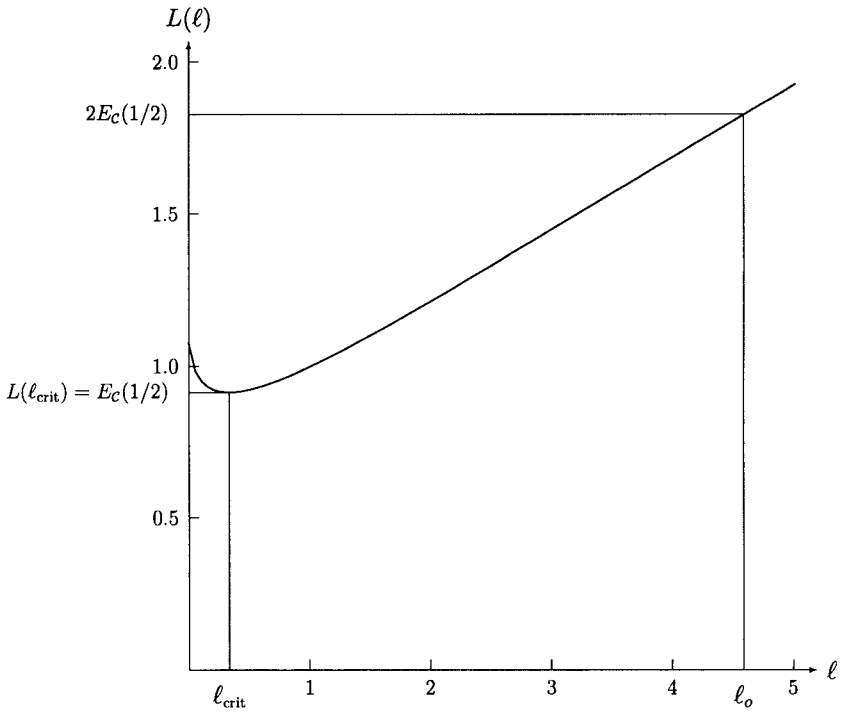

Fig. 3. Error burst length exponent $L(\ell)$ for rate $R = 1/2$.

is the *error burst length exponent*. For each rate R there is a normalized value $\ell_{\text{crit}} = (j_{\text{crit}} + 1)/m$ that minimizes the exponent $L(\ell)$. Forney called the value j_{crit} the *critical length* of the error event [7].

Asymptotically $(m \to \infty)$ the contribution to the error probability of a convolutional code will be dominated by the error bursts of the most likely normalized length, ℓ_{crit}. Let $E_{\mathcal{C}}(R)$ be the error exponent for convolutional codes used over the BSC [6], [8]

$$E_{\mathcal{C}}(R) = \frac{R\log_2(2\sqrt{p(1-p)})}{\log_2(2^{1-R}-1)}, \qquad 0 < R < R_{\text{comp}} \tag{14}$$

and

$$\begin{cases} E_{\mathcal{C}}(R) = E_0(\rho), & 0 \le \rho \le 1 \\ R = (1-\varepsilon)\frac{E_0(\rho)}{\rho}, & R_{\text{comp}} \le R < C, \ \varepsilon > 0 \end{cases} \tag{15}$$

where

$$E_0(\rho) = \rho - (1+\rho)\log_2\left(p^{\frac{1}{1+\rho}} + (1-p)^{\frac{1}{1+\rho}}\right) \tag{16}$$

and C is the channel capacity. Thus we conclude that

$$E_{\mathcal{C}}(R) = L(\ell_{\text{crit}}). \tag{17}$$

In Fig. 3 we give the error burst length exponent $L(\ell)$ as a function of the normalized error burst length ℓ for rate $R = 1/2$. Here $E_{\mathcal{C}}(1/2)$ is the value of the error exponent for convolutional codes of rate $R = 1/2$.

We choose as our inner code a convolutional code for which the probability of the normalized error burst length ℓ satisfies (12). The "channel" on which any outer decoder operates has memory due to the fact that the inner decoder may produce a burst of errors of arbitrary lengths $(\geq m_i + 1)$. We consider the conditions when this outer channel can be estimated using the model of a channel with independent errors where the probability of errors of any multiplicity is greater than or equal to the probability of the same event for the channel with memory. In order to satisfy this condition we have to choose the number of outer encoders l_o in such a manner that the probability of error in a code symbol of the outer code (regardless of the previous symbol) can be upper-bounded by the value p_o, where the crossover probability p_o for the outer channel is the first event error probability in case of Viterbi decoding of the inner code. If two consecutive symbols of an outer code are not located in the same burst of error (caused by the inner decoder) then it is clear that these errors are independent. But if these symbols are located in the same burst then the probability of this event does not exceed the probability of the event that the burst has length $(l_o + 1)$. Then we choose l_o such that for $\ell_o = (l_o + 1)/m$ we have

$$L(\ell_o) = 2E_{\mathcal{C}}(R_i) = 2L(\ell_{\text{crit}}). \tag{18}$$

From (12) and (18) it follows that

$$P(\mathcal{L}(l_o)) \le p_o^2 \tag{19}$$

holds. If we consider i, $i \geq 3$, errors in an outer code, then it follows from Fig. 3 that

$$P(\mathcal{L}(j)) \le p_o^i, \tag{20}$$

where $j = (i-1)l_o$. If condition (18) is satisfied, then the channel for each outer code can be regarded as a channel with independent errors with crossover probability p_o.

In Fig. 4 we show the parameter ℓ_o as a function of R_i for $p_o = 0.01$.

We are now prepared to estimate the error exponent $E_w(R_w)$ for woven convolutional codes with outer warp. Our approach is

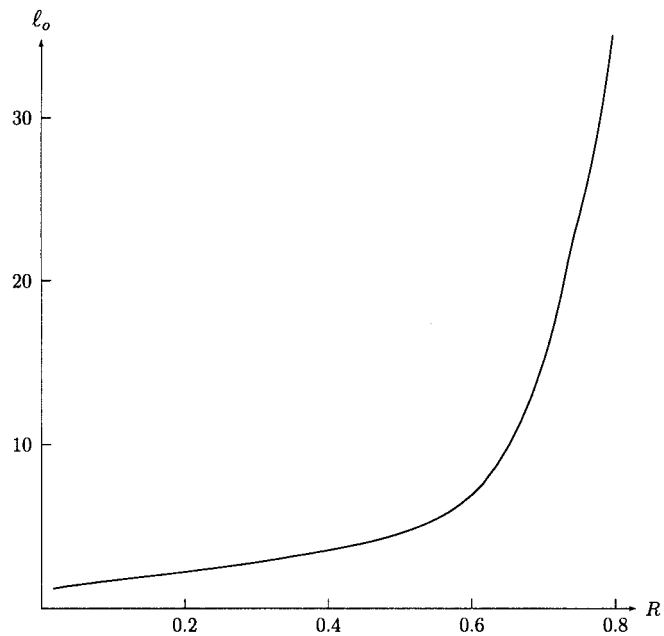

Fig. 4. Behavior of ℓ_o as a function of the inner code rate R_i.

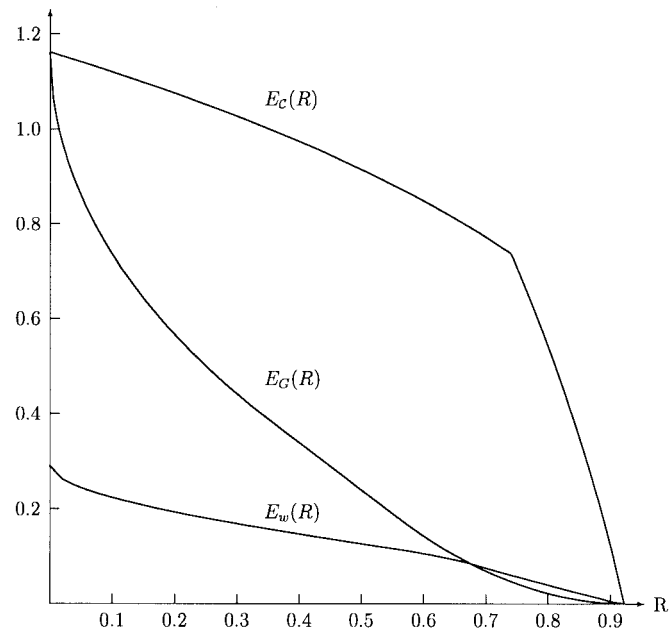

Fig. 5. Error exponents for block $(E_G(R))$, convolutional $(E_{\mathcal{C}}(R))$, and woven convolutional $(E_w(R))$ codes.

to generalize the relation between the Gilbert–Varshamov bound for binary block codes and the expurgated error exponent for randomly selected block codes for the BSC [5]. Thus we consider the relation between the Costello bound for binary convolutional codes [9], [6] and the expurgated error exponent for randomly chosen binary convolutional codes for the BSC (14). The latter can be expressed as follows:

$$E_{\mathcal{C}}^{\exp}(R) = \frac{R}{\log_2(2^{1-R}-1)} \log_2(2\sqrt{p(1-p)}), \quad 0 < R < R_{\text{comp.}} \tag{21}$$

Clearly, the expurgated error exponent for convolutional code can be regarded as the product of two factors, viz., the normalized Costello bound for the free distance

$$\delta_{\mathcal{C}}(R) = \frac{R}{-\log_2(2^{1-R}-1)} \tag{22}$$

and the factor $-\log_2(2\sqrt{p(1-p)})$ which is related to the Bhattacharyya bound on the first event error probability P_{B2} when two codewords are used to communicate over the BSC, viz.,

$$P_{B2} < (2\sqrt{p(1-p)})^d \tag{23}$$

where d is the Hamming distance between the two codewords [6]. The first factor is determined by the outer code and the second by the inner code by choosing $p = p_o$, the first event error probability of the inner convolutional code, and, hence,

$$p_o = 2^{-(E_{\mathcal{C}}(R_i)-\varepsilon)m_i c_i}, \qquad \varepsilon > 0 \tag{24}$$

where $E_{\mathcal{C}}(R_i)$ is a function of the crossover probability for the BSC.

Let P_w be the first event error probability for the woven convolutional code, i.e.,

$$P_w = 2^{-(E_w(R_w)-\varepsilon)mc_o c_i} \tag{25}$$

where $E_w(R_w)$ is the error exponent for the woven convolutional code and $m = m_o m_i$. Then, P_w can be estimated by the error probability for the outer code and we have

$$P_w = 2^{-(E_w(R_w)-\varepsilon)mc_o c_i} = 2^{-(E_{\mathcal{C}}^{o}(R_o)-\varepsilon)m_o c_o} \tag{26}$$

where $E_{\mathcal{C}}^{o}(R_o)$ is the error exponent for the outer convolutional code using a BSC with crossover error probability p_o.

For large inner memories m_i all outer convolutional codes encounter very good channels $(p_o \to 0)$, which means that the computational cut-off rate is very close to the channel capacity and, as a consequence, for practically all rates between zero and the channel capacity we can use the expurgated error exponent as an estimate for the error exponent for the woven convolutional code. Hence, we have

$$\begin{aligned} E_w(R_w) &= -\delta_{\mathcal{C}}(R_o)\log_2(2\sqrt{p_o(1-p_o)})/m_i c_i \\ &= -\delta_{\mathcal{C}}(R_o)(1/2)(2+\log_2(p_o)+\log_2(1-p_o))/m_i c_i. \end{aligned} \tag{27}$$

Since p_o decreases exponentially with $m_i c_i$ (with the error exponent $E_{\mathcal{C}}(R_i)$) the terms 2 and $\log_2(1-p_o)$ have negligible influence compared to $\log_2 p_o$, and, thus we obtain

$$E_w(R_w) = (1/2)E_{\mathcal{C}}(R_i)\delta_{\mathcal{C}}(R_o). \tag{28}$$

Maximizing (28) with respect to R_i we obtain the following error exponent for woven convolutional codes with outer warp:

Theorem 1: The error exponent for rate $R_w = R_o R_i$ woven convolutional codes with outer warp is

$$E_w(R_w) = \max_{R_i}\{(1/2)E_{\mathcal{C}}(R_i)\delta_{\mathcal{C}}(R_w/R_i)\} \tag{29}$$

where $E_{\mathcal{C}}(R)$ is the error exponent for convolutional codes. □

The error exponents $E_w(R)$, $E_G(R)$, and $E_{\mathcal{C}}(R)$ are all given in Fig. 5 for the BSC with crossover probability $p = 0.01$. The exponent $E_w(R)$ for the woven convolutional codes is seen to be essentially smaller than the exponent $E_{\mathcal{C}}(R)$ for convolutional codes, but in the bound (25) it is multiplied by $mc_o c_i = m_o m_i c_o c_i$ which usually is much larger than the corresponding factor for ordinary convolutional codes.

Finally, some comments about the decoding complexity. Suppose that the decoder of woven convolutional codes with outer warp consists of one Viterbi decoder for the inner code and l_o Viterbi decoders for the outer codes. The decoding complexity is proportional to

$$\Gamma = 2^{m_i} + l_o 2^{m_o}. \tag{30}$$

Let us choose $m_i = m_o = \sqrt{m}$. Then, we have

Theorem 2: Suppose that a decoder for a woven convolutional code with outer warp consists of l_o Viterbi decoder for the outer convolutional codes and one Viterbi decoder for the inner convolutional code. If both the outer and inner convolutional codes have memory $\sqrt{m}$, then the complexity of the decoder is proportional to

$$\Gamma = (1 + l_o) 2^{\sqrt{m}}. \tag{31}$$

□

From Theorems 1 and 2 it follows, somewhat surprisingly, that for woven convolutional codes with $m_o = m_i = \sqrt{m}$ the decoding error probability decreases exponentially with m while the decoding complexity increases exponentially only with $\sqrt{m}$.

References

[1] 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 District, U.K., July 13–18, 1997).

[2] V. V. Zyablov and S. A. Shavgulidze, "Error exponent of concatenated decoding of block-convolutional concatenated codes," in *Proc. Int. Workshop "Convolutional Codes; Multi-User Communication"* (Sochi, USSR, May 30–June 6, 1983).

[3] ——, "Generalized convolutional concatenated codes with unit memory," *Probl. Pered. Inform.*, vol. 22, no. 4, pp. 9–28, 1986.

[4] 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. 1997 IEEE Int. Symp. Information Theory* (Ulm, Germany, June 29–July 4, 1997).

[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

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

[7] G. D. Forney Jr., "Convolutional codes II: Maximum-likelihood decoding," *Inform. Contr.*, vol. 25, pp. 222–266, 1974.

[8] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.

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

New Rate $1/2$, $1/3$, and $1/4$ Binary Convolutional Encoders with an Optimum Distance Profile

Rolf Johannesson, *Fellow, IEEE*, and Per Ståhl, *Student Member, IEEE*

***Abstract*— Tabulations of binary systematic and nonsystematic polynomial convolutional encoders with an optimum distance profile for rate $1/2$, $1/3$, and $1/4$ are given. The reported encoders are found by computer searches that optimize over the weight spectra. The free distances for rate $1/3$ and $1/4$ are compared with Heller's and Griesmer's upper bounds.**

***Index Terms*—Convolutional encoders, free distance, optimum distance profile.**

The *distance profile* [1] $\boldsymbol{d} = [d_0, d_1, \cdots, d_m]$, where d_j is the jth-order *column distance* [2] and m is the memory of the convolutional encoder, is an important distance parameter for convolutional encoders. It is an encoder property but if we limit our interest to consider only encoding matrices $G(D)$ with $G(0)$ having full rank we can regard the distance profile as a code property [3]. When comparing codes with the same rate and memory, we say that a distance profile $\boldsymbol{d}$ is superior to a distance profile $\boldsymbol{d}'$ if $d_i > d_i'$ for the smallest $i, 0 \le i \le m$, where $d_i \ne d_i'$. The code with the superior $\boldsymbol{d}$ will generally require less computation with sequential decoding than the other code [1], [4].

In [5], extensive tables of rate $1/2$ convolutional encoders were given. In Tables I and II we give rate $1/2$ polynomial systematic and nonsystematic convolutional encoders, respectively, with an *optimum distance profile* (ODP encoders), i.e., with a distance profile equal to or superior to that of any other encoder. The generators are written in an octal form according to the convention introduced in [1]. For each value of the memory, we give the encoder with the largest free distance d_{free} among ODP encoders. (The free distance is the minimum Hamming distance between any two differing codewords.) Ties were resolved by comparing their weight spectra, i.e., by successively using the number of low-weight paths $n_{d_{\text{free}}+i}$ for $i = 0, 1, \cdots, 9$ as a further optimality criterion. The generators marked with "*" have better spectra than those given in [5].

In an earlier paper [6], systematic convolutional encoders of rate $1/3$ and $1/4$ were published together with a few short nonsystematic encoders of rate $1/3$. Only one spectral component, viz., the number of paths of weight d_{free}, was given. Here we give ten spectral components as well as extensive lists of nonsystematic encoders. We list rate $1/3$ and $1/4$ systematic as well as nonsystematic polynomial convolutional ODP encoders. The free distances are compared with Heller's and Griesmer's upper bounds on the free distances for nonlinear trellis and linear convolutional codes, respectively.

The free distance for any binary, rate $R = b/c$ convolutional code encoded by a polynomial, nonsystematic encoding matrix of memory

Manuscript received December 16, 1997; revised October 26, 1998. This research was supported in part by the Foundation for Strategic Research–Personal Computing and Communication under Grant PCC-9706–09.

The authors are with the Department of Information Technology, Information Theory Group, Lund University, S-221 00 Lund, Sweden.

Communicated by E. Soljanin, Associate Editor for Coding Techniques.
Publisher Item Identifier S 0018-9448(99)04174-7.