A new upper bound on the first-event error probability for maximum-likelihood decoding of fixed binary convolutional codes

Cedervall, Mats; Johannesson, Rolf; Zigangirov, Kamil

Published in:
IEEE Transactions on Information Theory

1984

[3] Y. Sugiyama, M. Kasahara, S. Hirasawa, and T. Namekawa, "A method for solving key equation for decoding Goppa codes," *Inform. Contr.*, vol. 29, pp. 173–180, Jan. 1975.
[4] D. M. Mandelbaum, "A method for decoding of generalized Goppa codes," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 137–140, Jan. 1977.
[5] I. S. Reed, R. A. Scholtz, T. K. Truong, and L. R. Welch, "The fast decoding of Reed–Solomon Codes using Fermat theoretic transforms and continued fractions," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 100–106, Jan. 1978.
[6] L. R. Welch and R. A. Scholtz, "Continued fractions and Berlekamp's algorithm," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 19–27, Jan. 1979.
[7] D. M. Mandelbaum, "On a class of arithmetic codes and a decoding algorithm," *IEEe Trans. Inform. Theory*, vol. IT-22, pp. 85–88, Jan. 1976.
[8] F. Barsi and P. Maestrini, "Improved decoding algorithm for arithmetic residue codes," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 640–643, Sept. 1978.
[9] D. M. Mandelbaum, "Further results on decoding arithmetic residue codes," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 643–644, Sept. 1978.
[10] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*. London: Oxford, 1960.
[11] S. S. Yau and Y. C. Liu, "Error correction in redundant residue number systems," *IEEE Trans. Computers*, vol. C-22, pp. 5–11, Jan. 1973.
[12] N. S. Szabo and R. I. Tanaka, *Residue Arithmetic and its Applications to Computer Technology*. New York: McGraw-Hill, 1967.

A New Upper Bound on the First-Event Error Probability For Maximum-Likelihood Decoding of Fixed Binary Convolutional Codes

MATS CEDERVALL, MEMBER, IEEE, ROLF JOHANNESSON, MEMBER, IEEE, AND KAMIL SH. ZIGANGIROV

***Abstract*—An upper bound on the first-event error probability for maximum-likelihood decoding of fixed binary convolutional codes on the binary symmetric channel is derived. The bound is evaluated for rate 1/2 codes, and comparisons are made with simulations and with the bounds of Viterbi, Van de Meeberg, and Post. In particular, the new bound is significantly better than Van de Meeberg's bound for rates above R_{comp}.**

I. Introduction

Several authors have addressed the problem of analyzing the performance of a maximum likelihood (ML) decoder for convolutional codes. Viterbi [1] made clever use of signal flowchart techniques to derive his famous upper bound on the first-event error probability of ML decoding of fixed convolutional codes. Van de Meeberg [2] obtained a significant improvement of Viterbi's bound for large signal to noise ratios, while Post [3] using quite a different technique, derived a bound that is slightly better than these two bounds for low signal to noise ratios. Furthermore, Post [4] showed how to calculate the union bound exactly. Finally, Schalkwijk *et al.* [5] presented a method for *exact* calculation of the event error probability. However, this method is feasible only for short codes.

Manuscript received April 4, 1983; revised October 20, 1983. This research was supported in part by the Royal Swedish Academy of Sciences and the Royal Swedish Academy of Engineering Sciences in liasion with the Academy of Sciences of the USSR. This work was presented at the IEEE International Symposium on Information Theory, June 21–25, 1982, Les Arcs, France.
M. Cedervall and R. Johannesson are with the Department of Computer Engineering, University of Lund, PO Box 725, S-220 07 Lund, Sweden.
K. Sh. Zigangirov is with the Institute for Problems of Information Transmission, USSR Academy of Sciences, 19 Ermolovoy Street, Moscow GSP-4, USSR.

In this correspondence we derive tighter upper bounds of both Viterbi- and Van de Meeberg-type for binary convolutional codes on the binary symmetric channel (BSC). Our bound is significantly better than Van de Meeberg's bound and generally also better than Post's bound for rates above the computational cut-off rate R_{comp}.

Viterbi used the union bound in his derivation. While the union bound is quite tight when there are few channel symbols in error, it is rather loose when we have many errors among the channel symbols. Thus, let us separate the error event into two disjoint events corresponding to few (F) and many (M) errors, respectively. If we let E denote the event that the first information symbol is erroneously decoded by maximum-likelihood decoding on the binary symmetric channel (BSC), we have

$$\begin{aligned} P[E] &= P[E|F]P[F] + P[E|M]P[M] \\ &\le P[E|F]P[F] + P[M] \\ &= P[E,F] + P[M]. \end{aligned} \tag{1}$$

In Section II we will use a random-walk argument to upper bound the probability that we have many channel errors, $P[M]$. The union bound is used in Section III to obtain a good upper bound on the probability that the first information symbol is erroneously decoded and that we have few channel errors, $P[E,F]$. These two bounds are combined in Section IV to a tighter Viterbi-type bound, and in Section V we give an improved Van de Meeberg-type bound. Finally, in Section VI we discuss some numerical results.

II. Many Channel Errors—Random Walk

To obtain an upper bound on the probability that we have many channel errors $P[M]$, we use a random-walk argument that is based on a lemma given by Gallager [6, p. 312].

Let $z_1, z_2, \cdots$ be a sequence of statistically independent identically distributed discrete random variables. Then for any $\lambda \le 0$ such that

$$E[2^{\lambda z_i}] \le 1, \tag{2}$$

and for any f,

$$P\left[\min_n \sum_{i=1}^{n} z_i \le -f\right] \le 2^{\lambda f}. \tag{3}$$

Let us accrue a metric s_e when we have a channel error and a metric s_c when the channel symbol is correctly received. Then the cumulative metric along the correct path is a random walk with $P[z_i = s_e] = p$ and $P[z_i = s_c] = q = 1 - p$, where p is the crossover probability of the BSC.

Suppose we choose metrics

$$s_e = \log \frac{p}{a}, \tag{4}$$

and

$$s_c = \log \frac{q}{1-a}, \tag{5}$$

where $p < a < 1$ is a parameter to be chosen later. Then condition (2) will be satisfied for those λ such that

$$p\left(\frac{p}{a}\right)^{\lambda} + q\left(\frac{q}{1-a}\right)^{\lambda} \le 1. \tag{6}$$

Noting that $\lambda = -1$ causes (6) to be satisfied, we have from (3) that

$$P\left[\min_n \sum_{i=1}^{n} z_i \le -f\right] \le 2^{-f}. \tag{7}$$

0018-9448/84/0900-0762$01.00 ©1984 IEEE

Let S_k denote the cumulative metric for the first k channel symbols, and suppose we have j_k errors among them. Then

$$S_k = j_k s_e + (k - j_k) s_c. \tag{8}$$

Now we can more precisely state what we mean by "few" and "many" errors. Those error patterns for which S_k stays above the barrier at $-f$ contain few errors, and those error patterns for which the cumulative metric hits or crosses the barrier contain many errors. Few errors, i.e., $\min_k S_k > -f$, is equivalent to

$$j_k s_e + (k - j_k) s_c > -f, \qquad \text{for all } k \tag{9}$$

or

$$j_k < \frac{f}{s_c - s_e} + k\frac{s_c}{s_c - s_e} \triangleq r_k, \qquad \text{for all } k. \tag{10}$$

In Fig. 1 we illustrate inequality (10).

From inequality (7) we have

$$P[M] = P\left[\min_k S_k \le -f\right] \le 2^{-f}, \tag{11}$$

where f is a parameter to be chosen later. The bound (11) will be exploited in Section IV.

It is interesting to notice that our choice of the metrics s_e and s_c is closely related to the Fano metric [7]. To see this, let φ_0 be the solution of

$$\varphi R = E_0(\varphi), \tag{12}$$

where $E_0(\varphi)$ is the Gallager function. If we choose a particular value of a, viz.

$$a = p^{1/(1+\varphi_0)}/\left(p^{1/(1+\varphi_0)} + q^{1/(1+\varphi_0)}\right) \tag{13}$$

(c.f. [6, p. 146]) we obtain from (4) and (5)

$$s_e = \varphi_0/(1+\varphi_0)(\log 2p - R) \tag{14}$$

and

$$s_c = \varphi_0/(1+\varphi_0)(\log 2q - R), \tag{15}$$

which differ by a factor exactly $\varphi_0/(1+\varphi_0)$ from the corresponding Fano metrics.

III. Few Channel Errors—Union Bound

To upper bound the probability that we make an error when we decode the first information symbol and have few channel symbol errors, $P[E,F]$, we use the union bound and obtain

$$P[E,F] \le \sum_k \sum_i P[E_{ki}, F], \tag{16}$$

where E_{ki} is the event that a path of length k and weight i is causing a decoding error.

A path has few channel errors only if it stays below the barrier in Fig. 1 for *all* k. If we take all paths of length k with $j_k < r_k$ channel errors we will get all paths with few channel errors together with the paths with many channel errors that take one or more detours above the barrier. Hence we have

$$P[E_{ki}, F] \le P[E_{ki}, j_k < r_k], \tag{17}$$

where

$$P[E_{ki}, j_k < r_k] = \sum_{j_k < r_k} \binom{k}{j_k} p^{j_k} q^{k-j_k} P[E_{ki} | j_k]. \tag{18}$$

If we multiply the sum in (18) by $\eta^{-(r_k - j_k)}$, $0 < \eta \le 1$, we obtain the inequality

$$P[E_{ki}, j_k < r_k] \le \eta^{-r_k} \sum_{j_k < r_k} \binom{k}{j_k} (\eta p)^{j_k} q^{k-j_k} P[E_{ki} | j_k]. \tag{19}$$

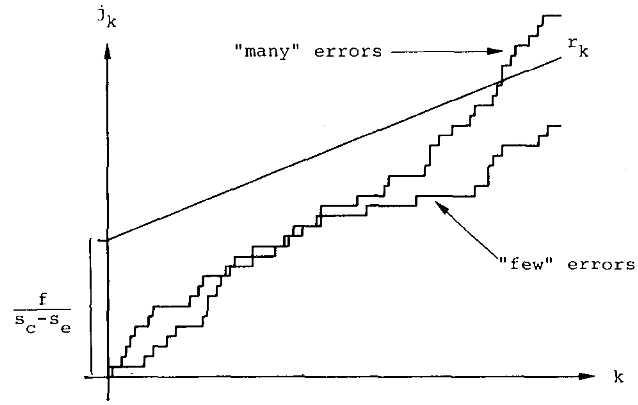

Fig. 1. Barrier at r_k which separates paths with many and few errors.

Let us introduce

$$p_0 \triangleq \frac{\eta p}{\eta p + q} \le p \tag{20}$$

and

$$q_0 \triangleq 1 - p_0 = \frac{q}{\eta p + q} \ge q. \tag{21}$$

Substituting (20) and (21) into (19) and rearranging (19) we get

$$P[E_{ki}, j_k < r_k] \le \left(\frac{pq_0}{p_0 q}\right)^{r_k} \left(\frac{q}{q_0}\right)^k \sum_{j_k < r_k} \binom{k}{j_k} p_0^{j_k} q_0^{k-j_k} P[E_{ki} | j_k]. \tag{22}$$

Overbounding by summing over all $0 \le j_k \le k$, we have

$$P[E_{ki}, j_k < r_k] \le \left(\frac{pq_0}{p_0 q}\right)^{r_k} \left(\frac{q}{q_0}\right)^k P[E_{ki}]_{p_0}, \tag{23}$$

where $P[E_{ki}]_{p_0}$ is the probability that a decoding error is caused by a path of length k and weight i on an improved BSC with crossover probability p_0.

Using the Bhattacharyya bound [8] we obtain from (23)

$$P[E_{ki}, j_k < r_k] \le \left(\frac{pq_0}{p_0 q}\right)^{r_k} \left(\frac{q}{q_0}\right)^k \left(2\sqrt{p_0 q_0}\right)^i. \tag{24}$$

Using the definition of r_k given in (14) and rearranging (24) we get

$$P[E_{ki}, j_k < r_k] \le \left(\frac{pq_0}{p_0 q}\right)^{f/(s_c - s_e)} L^k D^i, \tag{25}$$

where

$$L = \frac{q}{q_0}\left(\frac{pq_0}{p_0 q}\right)^{s_c/(s_c - s_e)}, \tag{26}$$

and

$$D = 2\sqrt{p_0 q_0}. \tag{27}$$

Finally, we combine (16) and (17) with (25) and obtain the following upper bound on the probability of having few channel symbol errors and making an error when decoding the first information symbol:

$$\begin{aligned} P[E,F] &\le \left(\frac{pq_0}{p_0 q}\right)^{f/(s_c - s_e)} \sum_k \sum_i a_{ki} L^k D^i \\ &= \left(\frac{pq_0}{p_0 q}\right)^{f/(s_c - s_e)} T(D, L), \end{aligned} \tag{28}$$

where a_{ki} is the number of paths of length k and weight i

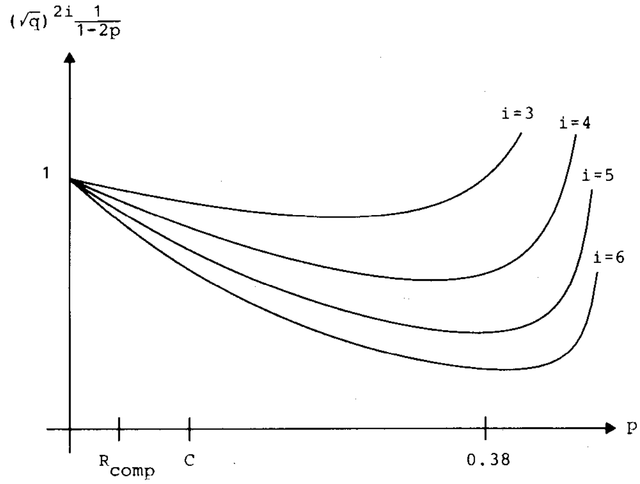

Fig. 2. Factor of Van de Meeberg's bound.

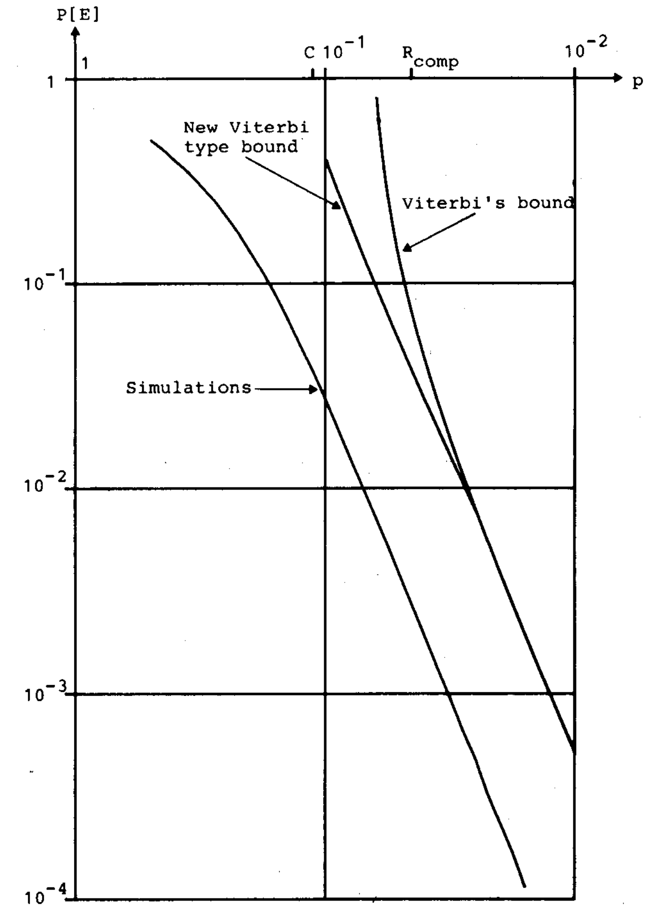

Fig. 3. New Viterbi type bound compared with Viterbi's bound and simulations for $R = 1/2$, $m = 2$, ODP code with $G^{(1)} = 5$ and $G^{(2)} = 7$.

stemming from the root node. $T(D, L)$ is the generating function for the output sequence weights and lengths.

IV. A Tighter Viterbi-Type Bound

We now show how to combine the bounds (1), (15), and (28) to obtain a new upper bound on the first-event error probability,

$$P[E] \le \left(\frac{pq_0}{p_0 q}\right)^{f/(s_c - s_e)} T(D, L) + 2^{-f}. \tag{29}$$

The bound (29) is valid for all f. By taking the derivative of the

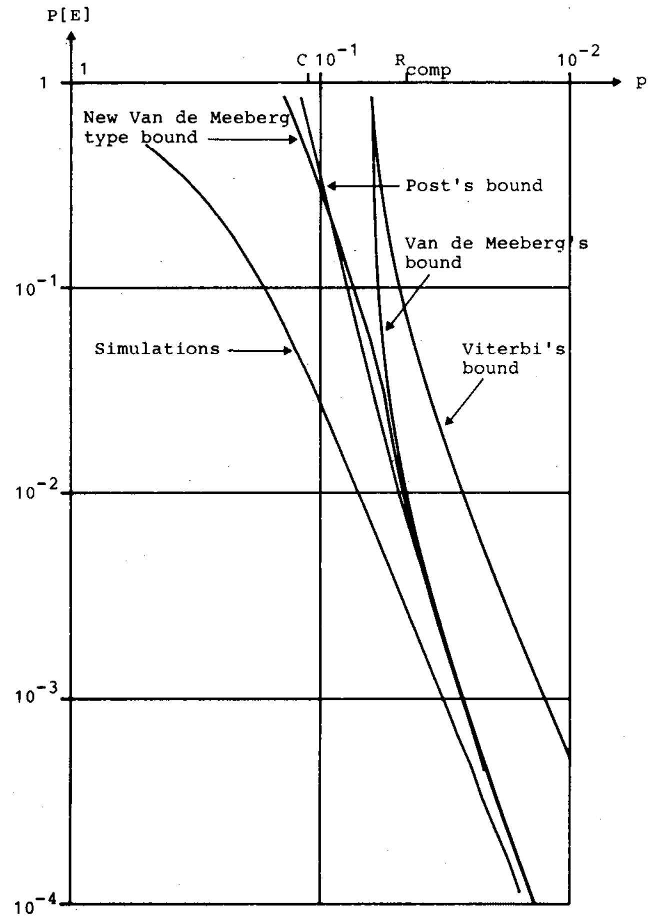

Fig. 4. New Van de Meeberg type bound compared with Van de Meeberg's and Post's bounds and with simulations for $R = 1/2$, $m = 2$, ODP code with $G^{(1)} = 5$ and $G^{(2)} = 7$.

right side of (29) we find that its minimum is obtained for

$$f_0 = \frac{(s_c - s_e)\left(\log(s_c - s_e) - \log T(D, L) - \log\log \dfrac{pq_0}{p_0 q}\right)}{\log \dfrac{pq_0}{p_0 q} + s_c - s_e}. \tag{30}$$

Inserting (30) and rearranging (29) give the upper bound

$$P[E] \le 2^{h(\gamma)} T(D, L)^{\gamma}, \tag{31}$$

where $h(\gamma)$ is the binary entropy function and

$$\gamma^{-1} = 1 + \frac{\log \dfrac{pq_0}{p_0 q}}{s_c - s_e}. \tag{32}$$

Finally, we use (4) and (5) and obtain a new Viterbi-type bound,

$$P[E] \le \inf_{0 < p_0 \le p} \; \inf_{p < a < 1} 2^{h(\gamma)} T(D, L)^{\gamma}, \tag{33}$$

where

$$\gamma^{-1} = 1 + \frac{\log \dfrac{pq_0}{p_0 q}}{\log \dfrac{qa}{p(1-a)}} \tag{34}$$

$$D = 2\sqrt{p_0 q_0} \tag{35}$$

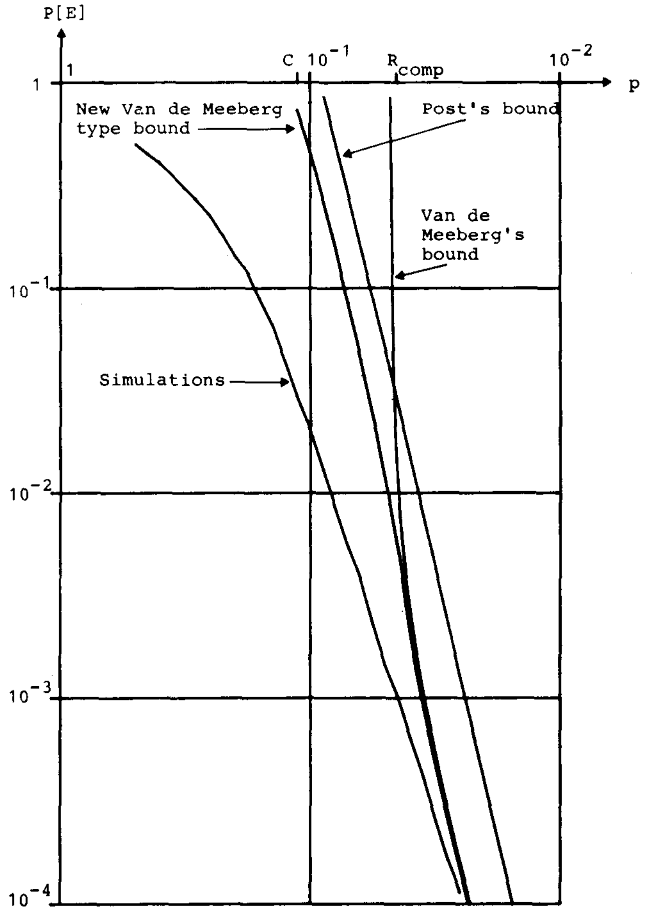

Fig. 5. New Van de Meeberg type bound compared with Van de Meeberg's and Post's bounds and with simulations for $R = 1/2$, $m = 4$ code with $G^{(1)} = 72$ and $G^{(2)} = 62$.

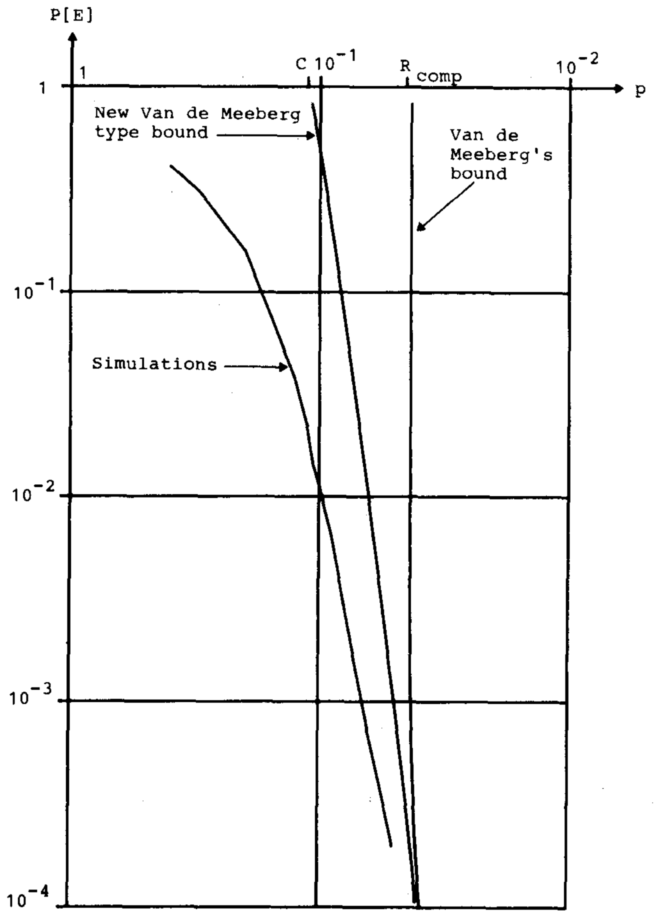

Fig. 6. New Van de Meeberg type bound compared with Van de Meeberg's bound and simulations for $R = 1/2$, $m = 8$, ODP code with $G^{(1)} = 557$ and $G^{(2)} = 751$.

and

$$L = \frac{q}{q_0}\left(\frac{pq_0}{p_0 q}\right)^{\log[q/(1-a)]/\log[qa/p(1-a)]}. \tag{36}$$

The new bound (33) is significantly better than Viterbi's original bound for rates $R_{\text{comp}} < R < C$. The latter bound can be obtained by choosing $p_0 = p$ rather than minimizing over p_0 on the right of (33).

We have derived the bound (33) only for the BSC, but our argument generalizes in a straightforward way to any symmetric binary-input discrete memoryless channel. For the BSC, however, the bound (33) can be improved.

V. A Tighter Van de Meeberg-Type Bound

Let P_i be the error probability for two binary codewords at distance i [1]. Van de Meeberg [2] used the fact that $P_{2i} = P_{2i-1}$ to tighten the Bhattacharyya bound and showed that

$$P_{2i} \le \binom{2\delta - 1}{\delta} 2^{-2\delta}\left(2\sqrt{p}\right)^{2i}, \tag{37}$$

where

$$\delta = \left\lfloor \frac{d_\infty + 1}{2} \right\rfloor. \tag{38}$$

Van de Meeberg used the inequality

$$\frac{q^i - p^i}{q - p} \le 1. \tag{39}$$

We notice that

$$\frac{q^i - p^i}{q - p} = q^i \frac{1 - (p/q)^i}{1 - 2p} < \left(\sqrt{q}\right)^{2i} \frac{1}{1 - 2p}. \tag{40}$$

For $p \le 0.38$ and $d_\infty > 4$ (see Fig. 2) the bound (40) is tighter than (39), and we have (for all practical values of p and d_∞ slightly improved) the Van de Meeberg-type bound

$$P_{2i} < \binom{2\delta - 1}{\delta} \frac{2^{-2\delta}}{1 - 2p}\left(2\sqrt{pq}\right)^{2i}. \tag{41}$$

Since $P_{2i} = P_{2i-1}$, $i \ge 1$, we can now rewrite our Viterbi-type bound (33) as

$$\begin{aligned} P[E] < \inf_{0 < p_0 \le p}\ \inf_{p < a < 1} 2^{h(\gamma)} &\left[\binom{2\delta - 1}{\delta}\frac{2^{-2\delta}}{1 - 2p_0}\right]^{\gamma} \\ &\cdot\left\{\frac{1}{2}[T(D,L) + T(-D,L)] \right. \\ &\left. + \frac{1}{2}D[T(D,L) - T(-D,L)]\right\}^{\gamma}, \end{aligned} \tag{42}$$

where γ, δ, D, and L are given in (34), (38), (35), and (36), respectively.

VI. Numerical Results

The generating function $T(D, L)$ is a formal power series in D and L, and for relatively short codes it is easily determined by solving linear equations for all states [8]. This is equivalent to

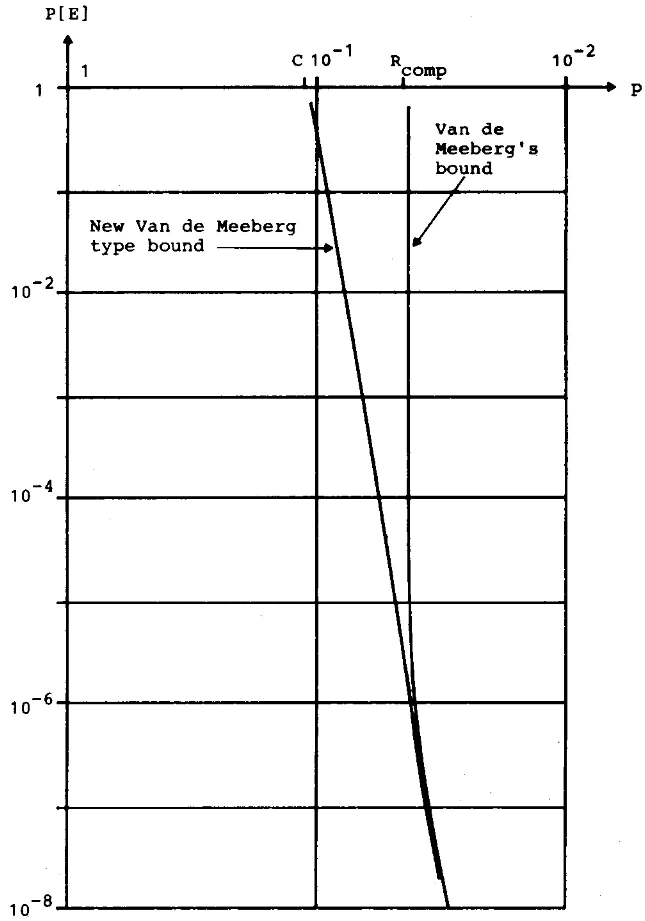

Fig. 7. New Van de Meeberg type bound compared with Van de Meeberg's bound for $R = 1/2$, $m = 15$, ODP code with $G^{(1)} = 467724$ and $G^{(2)} = 621134$.

finding an eigenvector of a sparse matrix. For codes of moderate memory, say $m = 10$, this is already not feasible. Therefore, we use the actual numerical values of D and L obtained from (35) and (36) and solve the corresponding *numerical* system of linear equations by iteration. The optimization is done by a straightforward numerical method.

We calculate our Viterbi-type bound (33) for the standard rate $1/2$ convolutional code, viz. the memory $m = 2$ code with code generators $G^{(1)} = 5, G^{(2)} = 7$ (octal notation). In Fig. 3 we compare our bound both with simulations and with Viterbi's bound,

$$P[E] < T(D,L)|_{D=2\sqrt{pq},\, L=1}. \tag{43}$$

Our bound is significantly better than Viterbi's for rates above R_{comp}.

For the same code we compare (Fig. 4) our Van de Meeberg-type bound (42) with simulations and with both Van de Meeberg's and Post's bounds. Our bound is similar to Post's bound but significantly better than Van de Meeberg's bound. It is interesting to note that for this code there exists a region between R_{comp} and channel capacity C, where our new bound is slightly worse than Post's bound. For channels with small crossover probability our bound is, of course, equivalent to Van de Meeberg's bound. For a slightly longer code, viz. the memory $m = 4$ code with code generators $G^{(1)} = 72$ and $G^{(2)} = 62$ ($d_\infty = 7$), our bound is significantly better than not only Van de Meeberg's bound but also Post's bound (Fig. 5).

We close this correspondence by showing the dramatic improvement of Van de Meeberg's bound obtained for rates between R_{comp} and channel capacity C for two slightly longer codes, the optimum distance profile ODP codes [9] of memory $m = 8$ ($d_\infty = 12$) and $m = 15$ ($d_\infty = 18$). The curves are given in Figs. 6 and 7.

References

[1] A. J. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 751–772, Oct. 1971.
[2] L. Van de Meeberg, "A tightened upper bound on the error probability of binary convolutional codes with Viterbi decoding," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 389–391, May 1974.
[3] K. A. Post, "New bounds on the performance of binary convolutional codes using Viterbi decoding on a binary symmetric channel," *Proceedings, Symposium over Informatietheorie in de Benelux*, Zoetermeer, The Netherlands, May 29–30, 1980.
[4] K. A. Post, "Explicit evaluation of Viterbi's union bounds on convolutional code performance for the binary symmetric channel," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 403–404, May 1977.
[5] J. P. M. Schalkwijk, K. A. Post, and J. P. J. C. Aarts, "On a method of calculating the event error probability of convolutional codes with maximum likelihood decoding," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 737–743, Nov. 1979.
[6] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[7] R. M. Fano, "A heuristic discussion of probabilistic decoding," *IEEE Trans. Inform. Theory*, vol. IT-9, pp. 64–74, Apr. 1963.
[8] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.
[9] R. Johannesson, "Robustly-optimal rate one-half binary convolutional codes," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 464–468, July 1975.

Nonbinary Codes, Correcting Single Deletion or Insertion

GRIGORY TENENGOLTS

***Abstract*—In many digital communications systems, bursts of insertions or deletions are typical errors. A new class of nonbinary codes is proposed that correct a single deletion or insertion. Asymptotically, the cardinality of these codes is close to optimal. The codes can be easily implemented.**

I. Introduction

In communications systems, the disturbance of synchronization can cause digits to be deleted or inserted. Binary codes that correct single deletion or insertion were introduced by Sellers [1], Levenshtein [2], and Ullman [3]. Codes that correct multiple deletions or insertions were studied by Calabi and Harnett [4] and by Tanaka and Kasai [5].

Another class of binary codes correcting synchronization errors was given by Tenengolts [6]. These codes correct bit loss and substitution error in the preceding bit. Such errors are typical, for example, in tape perforation because of a fault in the tape transport mechanism.

In many practical systems, groups of symbols are inserted or deleted (see, for example, Raibman [7]). Therefore, a synthesis of nonbinary codes that correct deletions or insertions is of interest.

This correspondence considers a class of nonbinary codes, correcting single deletion or insertion. We will show that the cardinality (number of codewords) of these codes is close to asymptotically optimal. The codes can be easily implemented.

Manuscript received July 27, 1983; revised December 7, 1983. The results of this correspondence were presented at ESCON 77, Washington, DC, September 1977 and partially published in the Proceedings of this conference.

The author is with System Industries, 1855 Barber Lane, Milpitas, CA 95035.

0018-9448/84/0900-0766$01.00 ©1984 IEEE