TRUNCATION EFFECTS IN VITERBI DECODING

Robert J. McEliece and Ivan M. Onyszchuk

California Institute of Technology, Pasadena, CA 91125

Abstract

Practical Viterbi decoders often fall significantly short of full maximum likelihood decoding (MLD) performance because of survivor truncation effects. In this paper, we study the tradeoff between truncation length and performance loss for the two most common variations of Viterbi's algorithm: *best state decoding* (BSD) and *fixed state decoding* (FSD). We find that FSD survivors should be about twice as long as BSD survivors for comparable performance.

1. Introduction

Viterbi decoding is in principle a maximum likelihood decoding (MLD) algorithm for convolutional codes. However, in practice, Viterbi decoding falls short of full MLD performance for several reasons, one of the most important being the need to truncate survivors. In order to minimize the hardware complexity of the path memory section in a Viterbi decoder, one must use the shortest possible survivor truncation length that does not seriously compromise the decoder's performance. Thus in order to design high-performance and cost-effective Viterbi decoders, it is essential to know the precise tradeoff between truncation length T and performance loss. In this paper, we will study this tradeoff for the two most common variations of Viterbi's algorithm: *best state decoding* (BSD) and *fixed state decoding* (FSD).

For each bit produced by a BSD algorithm, the decoder finds the state with the best accumulated metric, and outputs the oldest bit in the survivor corresponding to this state, whereas in FSD algorithms, the decoder always outputs the oldest bit in the survivor of some fixed state (say the all-zeroes state). It is clear that for a given T, a BSD algorithm should outperform a FSD algorithm. However, finding the best state may be prohibitively difficult in high-speed decoders or in decoders with many states. In such cases, FSD is preferred despite its need for a larger T. For example, FSD is used in NTT's one-chip, 20 Mbit/sec decoder for the NASA standard K=7, rate 1/2 code [2], and in JPL's new K=15, rate 1/4, multichip decoder for the *Galileo* deep-space mission [3].

This work was partially supported by Caltech's Jet Propulsion Laboratory, and by the Air Force Office of Scientific Research, under contract AFOSR-88-0247.

In the next section, a review of Hemmati and Costello's results [1] for best state decoding shows that if T is too small, the code's free distance is decreased. Otherwise, MLD performance is achieved on "asymptotically quiet" channels such as an AWGN channel with very large bit signal-to-noise ratio E_b/N_0 because there is no loss in d_{free}. A truncation length T_b^* was proposed [1] that results in no asymptotic loss and a small loss with respect to MLD for low channel noise. We show that there is a truncation length $T_b > T_b^*$, such that if $T \geq T_b$, BSD results in a negligible loss from MLD on the unquantized AWGN channel (AGC) when the decoded bit error rate (BER) is $\leq 10^{-5}$. In section 3, we present similar results for FSD : a theoretical truncation length T_f^* which causes no asymptotic loss on the BSC and actually results in near MLD performance. We also estimate a slightly larger practical value T_f, usually about twice T_b, such that for $T \geq T_f$, FSD performs very close to MLD on the AGC. In section 4, we tabulate various truncation lengths for many rate $1/n$ codes and verify the theoretical results by analyzing several simulations.

2. The Hemmati-Costello Results for Best State Viterbi Decoding

Assume that all zeroes are transmitted by a memory m, rate $1/n$, binary convolutional encoder. A truncation length T decoder stores for each state only the most recent T survivor bits. Let ℓ be the trellis path associated with the survivor for the state with best accumulated metric at time $t + T$. A best state decoder will output the oldest bit of this survivor. Let $L_i^{(T)}$ be the set of all *long* ($\geq T$ branches) trellis paths from state 0, never returning to state 0, into state i (Figure 1). $L_0^{(T)}$ is defined to be an empty set.

A *truncation error*, defined to be a bit error not made by MLD, can occur if $\ell \in \bigcup_{i\neq 0} L_i^{(T)}$. Using the union-Bhattacharyya bound, the BSD probability of truncation error [1] on a binary-input discrete memoryless channel with output alphabet Y is

$$\begin{aligned} P_{\text{best}}^{(T)} &\le \Pr(\ell \in \bigcup_{i\neq 0} L_i^{(T)}) \\ &\le \sum_{i=1}^{2^m-1} \Pr(\ell \in L_i^{(T)}) \\ &\le \sum_{i=1}^{2^m-1} G_i^{(T)}(D)\ \Big|_{D=\gamma} \end{aligned}$$

where $G_i^{(T)}(D)$ is the generating function of all paths in $L_i^{(T)}$ and $\gamma = \sum_{y\in Y}\sqrt{p(y|0)p(y|1)}$. Now including MLD errors, the BER of a truncation length T, best state decoder is

$$P_e^{(T)} \le \frac{\partial G(D,N)}{\partial N}\Big|_{N=1,D=\gamma} + P_{\text{best}}^{(T)}.$$

The least exponent of D in the above expression for $P_{\text{best}}^{(T)}$ is the *truncation distance* [1]

$$d_{\text{best}}^{(T)} = \min_{1\le i\le 2^m-1}\ \min_{\ell\in L_i^{(T)}} \mathrm{w}(\ell)$$

where $\mathrm{w}(\ell)$ is the number of ones in ℓ. Let T_b^* be the least value of T such that $d_{\text{best}}^{(T)} > d_{\text{free}}$. Hemmati and Costello recommend using truncation length T_b^* because for small γ, $P_e^{(T)}$ will be dominated by the first term due to MLD errors. Since the union bound is sharp for low channel noise, if $d_{\text{best}}^{(T)} < d_{\text{free}}$, there will be an E_b/N_0 loss of $10\log_{10}(d_{\text{free}}/d_{\text{best}}^{(T)})$ dB with respect to MLD.

3. New Results for Fixed State Viterbi Decoders

In FSD, the decoder always outputs the oldest bit of the survivor corresponding to a fixed state i. Let $S_i^{(T)}$ be the set of *all* (short) trellis paths having less than T branches from state 0 into state i. A truncation error may occur if some path into state i, of length $\ge T$ trellis branches, has a better metric $\mu(\ell)$ than the metric of every $s \in S_i^{(T)}$. This event may be written as

$$\begin{aligned} \Omega_i^{(T)} &= \bigcup_{\ell\in L_i^{(T)}} \bigcap_{s\in S_i^{(T)}} \{\mu(\ell)\le\mu(s)\}. \\ \Pr(\Omega_i^{(T)}) &\le \sum_{\ell\in L_i^{(T)}} \Pr(\bigcap_{s\in S_i^{(T)}} \{\mu(\ell)\le\mu(s)\}) \\ &\le \sum_{\ell\in L_i^{(T)}} \Pr\{\mu(\ell)\le\mu(s)\} \end{aligned}$$

for any $s \in S_i^{(T)}$. Therefore,

$$\begin{aligned} \Pr(\Omega_i^{(T)}) &\le \min_{s\in S_i^{(T)}} \sum_{\ell\in L_i^{(T)}} \Pr\{\mu(\ell)\le\mu(s)\} \\ &\le \min_{s\in S_i^{(T)}} \sum_{\ell\in L_i^{(T)}} 2^{d^{01}(\ell,s)}\gamma^{\mathrm{w}(\ell)-\mathrm{w}(s)} \end{aligned}$$

on the BSC by the lemma in the Appendix. Let $P_{\text{fixed}}^{(T)}$ be the probability of the analogous event to $\Omega_i^{(T)}$ when a random codeword is transmitted instead of zeroes.

Claim. *If zeroes and ones are equally likely in the input data to the encoder, then*

$$P_{\text{fixed}}^{(T)} = \frac{1}{2^m}\sum_{j=1}^{2^m-1} \Pr(\Omega_j^{(T)}).$$

Proof: Suppose that the transmitted codeword c passes through state j at time $t+T$. This will occur with probability 2^{-m}. At this time, let $F_i^{(T)}$ be the set of trellis paths into (the fixed decoding) state i which merge with c only at time $\le t$ but not afterwards. Then $F_i^{(T)} + c = L_{i\oplus j}^{(T)}$. Similarly, if $E_i^{(T)}$ is the set of all trellis paths which go from c into state i in less than T branches, then $E_i^{(T)} + c = S_{i\oplus j}^{(T)}$. Therefore, there is a contribution of $\Pr(\Omega_{i\oplus j}^{(T)})$ to $P_{\text{fixed}}^{(T)}$ and thus

$$\begin{aligned} P_{\text{fixed}}^{(T)} &= \frac{1}{2^m}\sum_{j=0}^{2^m-1} \Pr(\Omega_{i\oplus j}^{(T)}) \\ &= \frac{1}{2^m}\sum_{j=1}^{2^m-1} \Pr(\Omega_j^{(T)}) \end{aligned}$$

because $L_0^{(T)}$ and therefore $\Omega_0^{(T)}$ are empty sets. ∎

$P_{\text{fixed}}^{(T)}$ is an upper bound on the probability of truncation error for FSD. Now including MLD errors, the BER of a truncation length T, fixed state decoder is

$$P_e^{(T)} \le \frac{\partial G(D,N)}{\partial N}\Big|_{N=1,D=\gamma} + P_{\text{fixed}}^{(T)}.$$

The least power of γ in the expression for $P_{\text{fixed}}^{(T)}$ is

$$\begin{aligned} d_{\text{fixed}}^{(T)} &= \min_{1\le i\le 2^m-1}\ \max_{s\in S_i^{(T)}}\ \min_{\ell\in L_i^{(T)}} \left[\mathrm{w}(\ell)-\mathrm{w}(s)\right] \\ &= \min_{1\le i\le 2^m-1}\left[\min_{\ell\in L_i^{(T)}} \mathrm{w}(\ell) - \min_{s\in S_i^{(T)}} \mathrm{w}(s)\right]. \end{aligned}$$

Let T_f^* be the least value of T such that $d_{\text{fixed}}^{(T)} > d_{\text{free}}$. Then for FSD with $T = T_f^*$, $P_e^{(T)}$ will be dominated by the first term due to MLD errors on a BSC with small crossover probability. For the AWGN channel, a similar calculation to that given in [4] for $\Pr\{\mu(\ell) \leq \mu(0)\} = p(\ell, 0)$ shows that

$$p(\ell, s) = Q\left(\sqrt{\frac{2[\mathrm{w}(\ell) - \mathrm{w}(s)]^2 E_s/N_0}{\mathrm{w}(\ell \oplus s)}}\right).$$

This means that $[\mathrm{w}(\ell) - \mathrm{w}(s)]$ in $d_{\text{fixed}}^{(T)}$ must be replaced by (usually smaller) $[\mathrm{w}(\ell) - \mathrm{w}(s)]^2/\mathrm{w}(\ell \oplus s)$ to obtain the analogous quantity for the AGC. Therefore, fixed state decoding requires a longer truncation length on the AGC than on the BSC to achieve the same E_b/N_0 losses with respect to MLD. This is completely unlike BSD because $d_{\text{best}}^{(T)}$ depends only on the code. Since we have been unable to compute $d_{\text{fixed}}^{(T)}$ exactly for the AGC, we recommend using T_f such that for the BSC, $d_{\text{fixed}}^{(T_f)} \geq d_{\text{best}}^{(T_b)}$. The simulations in the next section verify that using T_f^* for the BSC and T_f for the AGC result in negligible losses from MLD.

4. Truncation Lengths and Simulations

Table 1 contains truncation lengths for several typical rate $1/n$ codes whose generator polynomial coefficients are listed in octal notation (e.g. 15 means $x^3 + x^2 + 1$). T_b is the estimated truncation length of a best state decoder which requires an E_b/N_0 within 0.1 dB of that for MLD at BER $\leq 10^{-5}$ on the unquantized AWGN channel (AGC). The T_b values were obtained by calculating $P_e^{(T)}$ using the first ten terms of a generating function.

Since the m most recent bits of a survivor are the associated state number $0 \leq i \leq 2^m - 1$, there is no need to store these m bits. Therefore, only $L = T - m$ physical bits per survivor are required for truncation T decoding. If all encoder generator polynomials are reversed, then the MLD performance on a memoryless channel remains the same, but $P_{\text{best}}^{(T)}$ or $P_{\text{fixed}}^{(T)}$ may decrease. For example, code with octal generators 23 and 35 has $T_b^* = 15$ and $T_f^* = 26$ compared to $T_b^* = 17$ and $T_f^* = 29$ for the polynomials 27 and 31.

The values $L_b = 7$ bits (only!) for the systematic $m = 6$ code (1,117) and $L_b = 29$ for JPL's $m = 14$ *Galileo* code result from trellis paths rapidly accumulating distance with length away from the all zeroes path. The values in Table 1 indicate that using $T_b = 4(m+1)$ or $5(m+1)$ is excessive in many cases and not matched to the particular code. Also, it is immediately clear that FSD requires about twice as many survivor bits than BSD.

The BER curves in Figure 2 for code with octal generators 5 and 7 show that BSD with $T_b^* = 8$ performs within 0.25 dB of MLD. All T_b curves for which $d_{\text{best}}^{(T_b)} \geq d_{\text{free}}$ will eventually join the MLD ($T = \infty$) curve. It is important to observe that there is a negligible gain in using any $T_b > 10$ or $T_f > 17$. For the AGC, the fact that $d_{\text{fixed}}^{(13)} \approx 3.6 < d_{\text{free}} = 5$ causes a large E_b/N_0 loss for $T_f = 13$. Figure 3 shows the the T_b curves are about the same distance away from the MLD curve on the BSC as on the AGC (in accordance with the bound on $P_{\text{best}}^{(T)}$). More importantly, Figure 3 verifies that FSD performs better on the BSC. For example, there is little loss at $T_f = 13$ (because $d_{\text{fixed}}^{(13)} = 6 > d_{\text{free}} = 5$) and $T_f = 15$ instead of 17 yields MLD performance. Note that the $T_b = 5$ decoder performs much worse than the $T_b = 6$ decoder because $d_{\text{best}}^{(5)} = 4$ whereas $d_{\text{best}}^{(6)} = 5 = d_{\text{free}}$.

Figure 4 shows the performance of the NASA code (generators 133,171) on the AGC for several truncation lengths and decoding methods. In accordance with the results in sections 2 and 3, a decoder with $T_b = 30$ or $T_f = 52$ performs within 0.1 dB of MLD for all BER $\leq 4 \times 10^{-5}$. Note that, the $T_b = 27$ loss is only 0.1 dB or less and the $T_b = 24$ loss is < 0.25 dB with respect to MLD for a BER $\leq 10^{-5}$.

FSD performs close to *worst state decoding* (in Figure 4 for $T = 40$) because the fixed state's accumulated metric is usually very far from that of the best state. One might expect that any truncation length 40, practical method of decoding final estimates of the information bits, such as a majority vote on all the survivors' oldest bits, would have a BER curve between the MLD $T = \infty$ one (virtually the same as for $T_b = 40$) and the worst state curve.

If FSD is performed with survivors being exchanged between decoder states as in the NTT chip, then memory cells near the end of the fixed state's survivor may be eliminated. By examining the code's trellis diagram, it is easliy verified that $\sum_{i=0}^{m-1}(2^m - 2^i) = 1 + (m-1)2^m$ cells can be deleted, making $L_f = T_f - 2m + 1$ on average.

5. References

[1] F. Hemmati and D.J. Costello, Jr., "Truncation Error Probability in Viterbi Decoding", *IEEE Trans. Comm.*, COM-25, pp. 530-532, May 1977.

[2] T. Ishitani, K. Tansho, N. Miyahara, and S. Kato, "A Scarce- State-Transition Viterbi-Decoder VLSI for Bit Error Correction", *IEEE Journal of Solid-State Circuits*, SC-22, No. 4, pp. 575-581, August 1987.

[3] J. Statman, G. Zimmerman, F. Pollara, and O. Collins, "A Long Constraint Length VLSI Viterbi Decoder for the DSN", *JPL TDA Progress Report*, Feb. 1988.

[4] A. J. Viterbi, "Convolutional Codes and their Performance in Communication Systems", *IEEE Trans. Comm.*, COM-19, pp. 751-772, Oct. 1971.

Appendix

Lemma. *When all zeroes are sent over a binary symmetric channel, the probability that a ML decoder chooses codeword ℓ instead of s is*

$$p(\ell, s) \leq 2^{d^{01}(\ell,s)} \gamma^{w(\ell)-w(s)},$$

where $d^{01}(\ell, s) = |\{i : \ell_i = 0, s_i = 1\}|$ *and* $\gamma = 2\sqrt{p(1|0)p(1|1)}$.

Proof: Let E be the set of length t received vectors r such that $\Pr(r|\ell) \geq \Pr(r|s)$. Then when the all zeroes codeword is transmitted,

$$\begin{aligned} p(\ell, s) &\leq \sum_{r \in E} p(r|0) \\ &\leq \sum_{r \in E} \sqrt{\frac{p(r|\ell)}{p(r|s)}} p(r|0) \\ &\leq \sum_{r \in \{0,1\}^t} \prod_{i=1}^{t} \sqrt{\frac{p(r_i|\ell_i)}{p(r_i|s_i)}} p(r_i|0) \\ &= \prod_{i=1}^{t} \sum_{y \in \{0,1\}} \sqrt{\frac{p(y|\ell_k)}{p(y|s_k)}} p(y|0). \end{aligned}$$

For $d^{10}(\ell, s)$ values of i, $\ell_i = 1$ and $s_i = 0$, so the inner sum above is γ. For $d^{01}(\ell, s)$ values of i, $\ell_i = 0$ and $s_i = 1$, so the inner sum is

$$\begin{aligned} &\frac{p(0|0)^{3/2}}{\sqrt{p(0|1)}} + \frac{p(1|0)^{3/2}}{\sqrt{p(1|1)}} \\ &= \frac{p(0|0)^2 + p(1|0)^2}{\sqrt{p(0|1)p(1|1)}} \\ &= \frac{2(1 - 2p(1|0) + 2p(1|0)^2)}{2\sqrt{p(1|0)p(1|1)}} \leq \frac{2}{\gamma}. \end{aligned}$$

Otherwise $\ell_i = s_i$ and the inner sum is 1. Therefore,

$$\begin{aligned} p(\ell, s) &\leq 2^{d^{01}(\ell,s)} \gamma^{d^{10}(\ell,s) - d^{01}(\ell,s)} \\ &= 2^{d^{01}(\ell,s)} \gamma^{w(\ell)-w(s)}. \quad \blacksquare \end{aligned}$$

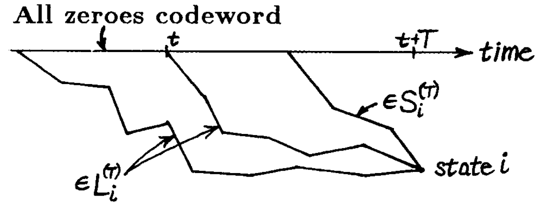

Figure 1. Long and Short Trellis Paths

TABLE 1. Truncation Lengths for Rate $1/n$ Codes

m	generators	d_f	T_b^*	T_b	L_b	T_f^*	T_f	L_f
2	5, 7	5	8	10	8	13	16	14
2	1, 7	4	5	5	3	10	10	8
3	15, 17	6	10	12	9	17	21	18
4	23, 35	7	15	18	14	26	32	28
5	65, 57	8	19	23	18	37	42	37
6	1, 117	6	12	13	7	27	29	23
6	133, 171	10	27	30	24	50	52	46
7	345, 237	10	28	32	25	52	55	48
8	561, 753	12	33	38	30	64	70	62
3	13, 15, 17	10	10	13	10	18	22	19
4	37, 33, 25	12	13	15	11	24	27	23
5	1, 75, 67	10	11	12	7	20	21	16
5	71, 65, 57	13	17	21	16	31	36	31
6	171, 165, 133	15	20	24	18	35	39	33
8	557, 663, 711	18	25	33	25	47	55	48
4	25, 33, 35, 37	13	9	12	8	17	20	16
7	275, 235 313, 357	18	18	25	18	34	40	33
14	46321, 51271 63667, 70535	35	35	43	29	67	75	61

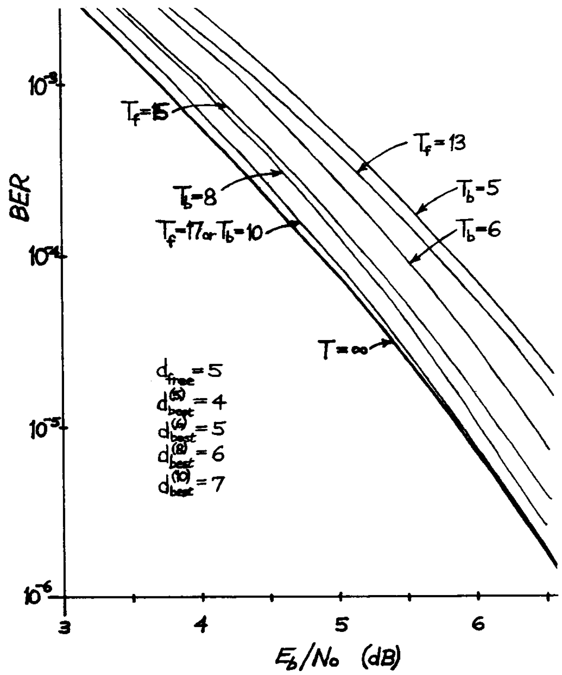

Figure 2. Performance of the (5,7) Code on the AWGN Channel

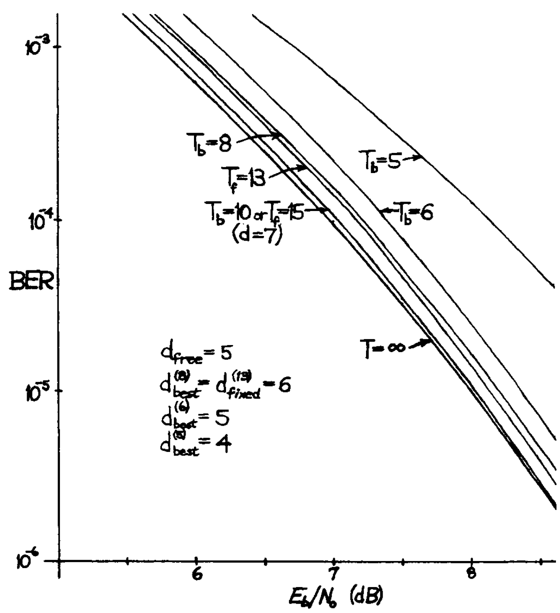

Figure 3. Performance of the (5,7) Code on the BSC

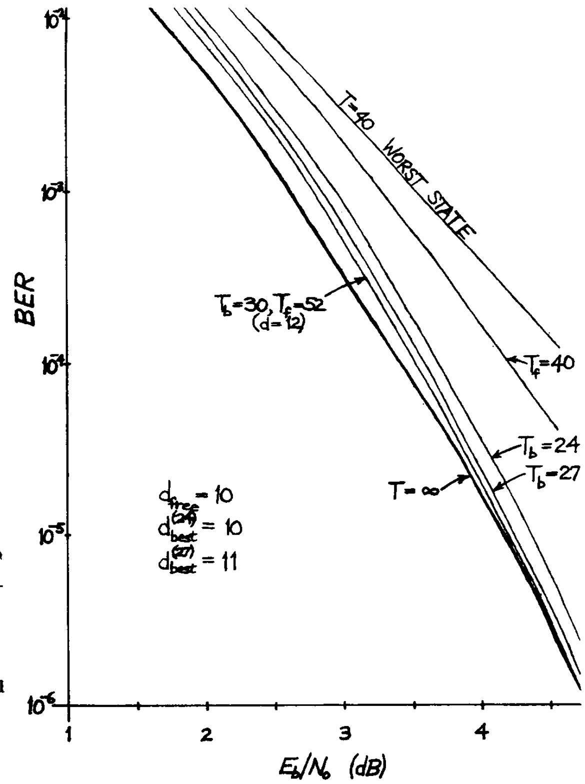

Figure 4. Performance of the (133,171) NASA Code on the Unquantized AWGN Channel