Upper Bound on the Frame Error Probability of Terminated Trellis Codes

Giuseppe Caire, *Member, IEEE*, and Emanuele Viterbo, *Member, IEEE*

***Abstract*— The frame-error rate (FER) of trellis codes with trellis termination can be *approximated* by considering the length of the error events negligible with respect to the frame length. In this letter we prove that this approximation is actually a true upper bound. This may be useful to orientate the design of frame-based transmission systems directly in terms of FER instead of the more common but less significant bit-error rate figure.**

***Index Terms*— Trellis-coded modulation performance.**

I. Introduction and Motivations

In many modern communication systems analog speech, audio and video signals are compressed and transmitted in digital form. Source encoders are normally frame-oriented and produce frames of encoded data. In order to cope with poor quality channels, each frame is then channel-encoded, interleaved and transmitted over the channel.

When trellis codes are used, a way to preserve the frame structure of transmission is *trellis termination* i.e., the encoder is forced to start from a known state at the beginning of each frame and it is driven to a known state at the end of the frame.[1]

At the receiver, various error concealment techniques can be employed and subjective tests are required to establish the quality of the system as function of the frame-error rate (FER) at the channel decoder output. This motivates the interest in evaluating the FER of terminated trellis codes (TTC).

II. System Model

We consider a general time-invariant trellis code that admits a minimal (canonical) feedforward encoder [2].

The minimal encoder is a finite-state machine with k input rails that accept symbols from a q-ary input alphabet $\mathcal{A}$. Each rail is connected to a shift register of length $\nu_\ell \geq 0$, for $\ell = 1 \cdots k$. The encoder state space (i.e., the set of encoder states) is $\mathcal{S}$ of cardinality $|\mathcal{S}| = \prod_{\ell=1}^{k} q^{\nu_\ell}$. At each trellis step, the encoder accepts a k-tuple of input symbols and produces a symbol belonging to a D-dimensional signal constellation. In the following, sequences of symbols are denoted by boldface characters and can be viewed as real vectors of length ND. Each code sequence $\mathbf{x}$ corresponds to a path in the trellis diagram of the encoder. The number of branches outgoing from each state and merging into each state is, at each step, $B = q^k$. Starting from state $s_0 \in \mathcal{S}$ at step 0, all states can be reached in at most $\nu_{\max} = \max(\nu_\ell)$ steps. Trellis termination is obtained by adding to the encoder input sequence a tail of $\sum_{\ell=1}^{k} \nu_\ell$ known symbols.

A TTC $\mathcal{C}$ spanning $N \geq \nu_{\max}$ trellis steps can be viewed as a block code over the reals of length ND and size $|\mathcal{C}| = B^N/|\mathcal{S}|$. We assume that each nth real component of the transmitted code sequence is sent through an additive white Gaussian noise (AWGN) channel with possibly time-varying real amplitude gains g_n, so that the n-th received sample is $y_n = g_n x_n + z_n$ for $n = 1, \ldots ND$, where $z_n \sim \mathcal{N}(0, N_0/2)$. For the sake of generality, we do not impose any statistical property on the channel gains g_n.

In this letter we derive an upper bound on $P_f(e|\mathbf{g})$, the conditional FER of TTC's, when the decoder has perfect knowledge of the channel gains $\mathbf{g}$ [perfect channel state information (CSI)].

In the case of the standard stationary AWGN channel we just substitute $g_n = 1$ for all n, while in the case of a fading channel the result of the bound should be averaged over the joint distribution of the g_n's. For example, with independent Rayleigh fading the g_n are i.i.d. Rayleigh distributed.

III. Analysis

In principle, the standard Union Bound for block codes can be used to upperbound $P_f(e|\mathbf{g})$ as

$$P_f(e|\mathbf{g}) \leq \frac{1}{|\mathcal{C}|} \sum_{\mathbf{x} \in \mathcal{C}} \sum_{\widehat{\mathbf{x}} \neq \mathbf{x}} P(\mathbf{x} \to \widehat{\mathbf{x}}|\mathbf{g}) \quad (1)$$

where $P(\mathbf{x} \to \widehat{\mathbf{x}}|\mathbf{g})$ is the conditional pairwise-error probability (PEP) of choosing $\widehat{\mathbf{x}} \in \mathcal{C}$ instead of $\mathbf{x}$ given $\mathbf{g}$, as if $\mathbf{x}$ and $\widehat{\mathbf{x}}$ were the only two possible decoder outcomes [3]. Unfortunately, bound (1) turns out to be very loose and computationally too complex to evaluate because of the large number of code words in $\mathcal{C}$.

A simple heuristic reasoning gives an approximation of $P_f(e|\mathbf{g})$. Denote by $(\mathbf{x}, \widehat{\mathbf{x}})_i^{i+L}$ an error event of length L starting at step i, i.e., a pair of code sequences that split at step i and whose first merge is L steps later. Assuming $L \ll N$ for all relevant error events and by neglecting trellis termination, we can write

$$P_f(e|\mathbf{g}) \simeq 1 - (1 - P(e|\mathbf{g}))^N \simeq NP(e|\mathbf{g}) \quad (2)$$

where $P(e|\mathbf{g})$ is the probability of having an error event starting from a given step (say step $i = 0$), given $\mathbf{g}$, which in

Manuscript received April 22, 1997. The associate editor coordinating the review of this letter and approving it for publication was Prof. H. V. Poor. This work was supported by the Italian Space Agency (ASI).

The authors are with the Dipartimento di Elettronica, Politecnico di Torino, I-10129 Turin, Italy (e-mail: viterbi@polito.it).

Publisher Item Identifier S 1089-7798(98)01286-1.

[1] In general, initial and final states may be different, although normally they are the same.

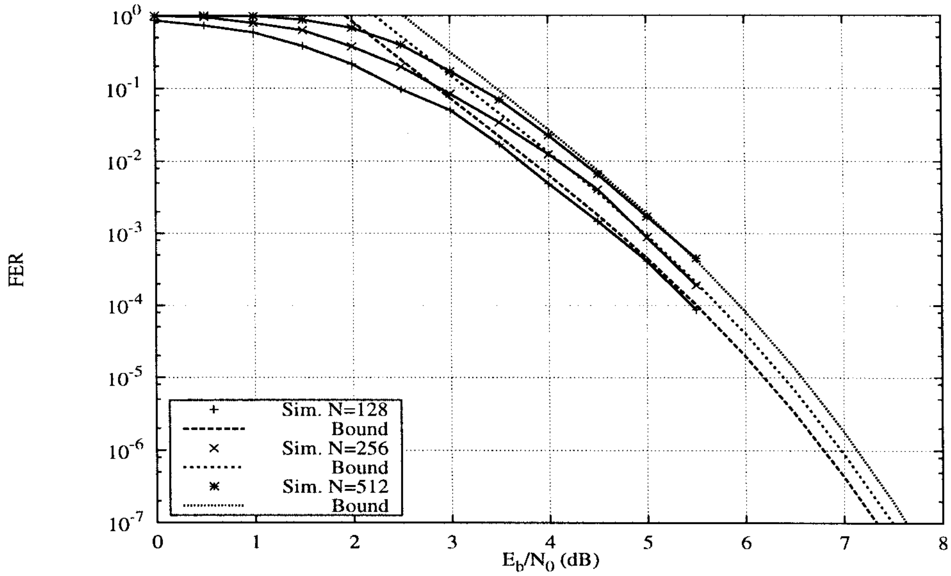

Fig. 1. FER: simulation versus bound.

turn can be upperbounded by the standard Union Bound [3]

$$P_{\rm ub}(e|\mathbf{g}) = \frac{1}{|\mathcal{S}|}\sum_{s_0\in\mathcal{S}}\sum_{L=1}^{\infty} B^{-L}\sum_{\mathbf{x}\in\mathcal{C}_L(s_0)}\sum_{(\mathbf{x},\widehat{\mathbf{x}})_0^L} P(\mathbf{x}\to\widehat{\mathbf{x}}|\mathbf{g}) \tag{3}$$

where $\mathcal{C}_L(s_0)$ is the subcode of all sequences of length L stemming from state $s_0 \in \mathcal{S}$ at step $i = 0$. Bound (3) can be computed by standard techniques based on the generalized transfer function of the encoder state diagram (see [3] for convolutional codes and [1] for more general trellis codes). Note also that expectations over the initial states s_0 and over the reference sequences $\mathbf{x}$ are not needed for geometrically uniform codes [2].

In the following we prove that the right-hand side of (2) is actually a true upper bound. The derivation is made in two steps: 1) we consider a trellis section of length N and compute the conditional FER $P_f(e|\mathbf{x},\mathbf{g})$ given the reference sequence $\mathbf{x}$, assuming that all the trellis sequences are admissible (starting and ending in any state) and 2) we consider the TTC consisting of all paths stemming from a given state s_0 and ending in a given state s_N after N steps. In order to take into account the case of nonuniform codes we average both over all $\mathbf{x}$ and over all pairs (s_0, s_N). There will be at least a TTC with conditional FER less or equal than average. In practice, for nonuniform codes the initial and the final states should be optimized. For uniform codes (for $N \geq \nu_{\max}$, which always holds in practical cases), the choice of s_0 and s_N is irrelevant and expectation is not needed, so that the bound holds for all TTC's.

Step 1: The conditional FER given the reference sequence $\mathbf{x} \in \mathcal{C}$ can be upperbounded as

$$P_f(e|\mathbf{x},\mathbf{g}) \leq \sum_{i=0}^{N-1}\sum_{L=1}^{N-i}\sum_{(\mathbf{x},\widehat{\mathbf{x}})_i^{i+L}} P(\mathbf{x}\to\widehat{\mathbf{x}}|\mathbf{g}). \tag{4}$$

Proof: The bound (4) can be obtained by expurgating the standard union bound $P_f(e|\mathbf{x},\mathbf{g}) \leq \sum_{\widehat{\mathbf{x}}\neq\mathbf{x}} P(\mathbf{x}\to\widehat{\mathbf{x}}|\mathbf{g})$ from all events $(\mathbf{x},\widehat{\mathbf{x}})$ obtained by concatenating two (or more) other events. All what we have to prove is that if an error event $(\mathbf{x},\widehat{\mathbf{x}})$ is the concatenation of two events $(\mathbf{x},\mathbf{x}')$ and $(\mathbf{x},\mathbf{x}'')$, with disjoint support[2] it can be elimitated from the Union Bound mantaining the inequality, irrespectively of the channel gain sequence $\mathbf{g}$. Let $\mathbf{G} = \mathrm{diag}(\mathbf{g})$ and let $\mathbf{z}$ and $\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{z}$ be the noise and the received sequences, respectively. The optimal decoder with perfect CSI decides according to the minimum Euclidean distance criterion. Hence, we need to show that if

$$||\mathbf{y} - \mathbf{G}\widehat{\mathbf{x}}||^2 \leq ||\mathbf{y} - \mathbf{G}\mathbf{x}||^2 \tag{5b}$$

$$(\widehat{\mathbf{x}} - \mathbf{x}) = (\mathbf{x}' - \mathbf{x}) + (\mathbf{x}'' - \mathbf{x}) \tag{5b}$$

then either $||\mathbf{y} - \mathbf{G}\mathbf{x}'||^2 \leq ||\mathbf{y} - \mathbf{G}\mathbf{x}||^2$ or $||\mathbf{y} - \mathbf{G}\mathbf{x}''||^2 \leq ||\mathbf{y} - \mathbf{G}\mathbf{x}||^2$. By manipulating (5a) and substituting (5b) we get the equivalent condition

$$\begin{aligned}||\mathbf{G}(\mathbf{x}' - \mathbf{x})||^2 + ||\mathbf{G}(\mathbf{x}'' - \mathbf{x})||^2 + 2(\mathbf{x}' - \mathbf{x})^T\mathbf{G}^T\mathbf{G}(\mathbf{x}'' - \mathbf{x})\\ \leq 2\mathbf{z}^T\mathbf{G}(\mathbf{x}' - \mathbf{x}) + 2\mathbf{z}^T\mathbf{G}(\mathbf{x}'' - \mathbf{x})\end{aligned} \tag{6}$$

Since $(\mathbf{x},\mathbf{x}')$ and $(\mathbf{x},\mathbf{x}'')$ have disjoint supports, the third term on the left handside of (6) vanishes.

Hence, either $||\mathbf{G}(\mathbf{x}' - \mathbf{x})||^2 \leq 2\mathbf{z}^T\mathbf{G}(\mathbf{x}' - \mathbf{x})$ or $||\mathbf{G}(\mathbf{x}'' - \mathbf{x})||^2 \leq 2\mathbf{z}^T\mathbf{G}(\mathbf{x}'' - \mathbf{x})$, otherwise (6) would be false. But the above implies the thesis. □

Going back to (4) we observe that the same pairwise error event of length L appears from step $i = 0$ to step $i = N - L$ for a total of $N - L + 1$ terms. Hence

$$P_f(e|\mathbf{x},\mathbf{g}) \leq \sum_{L=1}^{N}(N - L + 1)\sum_{(\mathbf{x},\widehat{\mathbf{x}})_0^L} P(\mathbf{x}\to\widehat{\mathbf{x}}|\mathbf{g}). \tag{7}$$

[2]We define the support of an error event $(\mathbf{x},\widehat{\mathbf{x}})_i^{i+L}$ as the ordered set of integers $\{i, i+1, \cdots, i+L-1\}$.

Step 2: Now we average $P_f(e|\mathbf{x},\mathbf{g})$ with respect to all $\mathbf{x} \in \mathcal{C}_N(s_0, s_N)$, a code consisting of all sequences of length N stemming from a state $s_0 \in \mathcal{S}$ at step $i = 0$ and ending in a state $s_N \in \mathcal{S}$ at step $i = N$. Moreover, we average the result with respect to all pairs (s_0, s_N). The total expectation gives (for $N \geq \nu_{\max}$):

$$\frac{1}{|\mathcal{S}|^2}\sum_{s_0\in\mathcal{S}}\sum_{s_N\in\mathcal{S}}\frac{|S|}{B^N}\sum_{\substack{\mathbf{x}\in\\ \mathcal{C}_N(s_0,s_N)}}\sum_{L=1}^{N}(N-L+1)\sum_{(\mathbf{x},\widehat{\mathbf{x}})_0^L}P(\mathbf{x}\to\widehat{\mathbf{x}}|\mathbf{g})$$

$$\overset{a}{=}\frac{1}{|\mathcal{S}|}\sum_{s_0\in\mathcal{S}}\frac{1}{B^N}\sum_{\substack{\mathbf{x}\in\\ \mathcal{C}_N(s_0)}}\sum_{L=1}^{N}(N-L+1)\sum_{(\mathbf{x},\widehat{\mathbf{x}})_0^L}P(\mathbf{x}\to\widehat{\mathbf{x}}|\mathbf{g})$$

$$\overset{b}{=}\frac{1}{|\mathcal{S}|}\sum_{s_0\in\mathcal{S}}\sum_{L=1}^{N}B^{-L}\sum_{\mathbf{x}\in\mathcal{C}_L(s_0)}(N-L+1)\sum_{(\mathbf{x},\widehat{\mathbf{x}})_0^L}P(\mathbf{x}\to\widehat{\mathbf{x}}|\mathbf{g})$$

where (a) follows from $\cup_{s_N\in\mathcal{S}}\mathcal{C}_N(s_0,s_N) = \mathcal{C}_N(s_0)$ and (b) follows from the fact that in $\mathcal{C}_N(s_0)$ there are exactly B^{N-L} paths $\mathbf{x}$ coinciding in the first L steps. We can further upper bound by replacing $N - L + 1$ by N and letting N go to ∞ in the sum

$$\begin{aligned} P_f(e|\mathbf{g}) &\leq \frac{N}{|\mathcal{S}|}\sum_{s_0\in\mathcal{S}}\sum_{L=1}^{\infty}B^{-L}\sum_{\substack{\mathbf{x}\in\\ \mathcal{C}_L(s_0)}}\sum_{(\mathbf{x},\widehat{\mathbf{x}})_0^L}P(\mathbf{x}\to\widehat{\mathbf{x}}|\mathbf{g}) \\ &= NP_{\mathrm{ub}}(e|\mathbf{g}). \end{aligned} \tag{8}$$

IV. RESULTS

As an example we consider the binary convolutional code of rate $R = 1/2$ and 16 states with generators $23, 35$ (in octal notation) transmitted with antipodal binary PAM over the stationary AWGN channel. In this case $\mathbf{g}$ is the all-one vector and the PEP is given by [3] $P(\mathbf{x} \to \widehat{\mathbf{x}}) = Q\left(\sqrt{2d_H(\mathbf{x},\widehat{\mathbf{x}})RE_b/N_0}\right)$ where $Q(z)$ is the Gaussian tail function, E_b is the average bit-energy and $d_H(\mathbf{x},\widehat{\mathbf{x}})$ is the componentwise Hamming distance between $\mathbf{x}$ and $\widehat{\mathbf{x}}$. Since this trellis code is uniform, the bound (8) applies to all TTC's, irrespective of the initial and final state.

Fig. 1 shows bound (8) for $N = 128\,256$ and 512. Monte Carlo simulation points are shown for comparison. We notice that (8) provides accurate results in the range of values interesting for applications (FER $\leq 10^{-2}$).

REFERENCES

[1] E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon, *Introduction to Trellis-Coded Modulation with Applications*. New York: MacMillan, 1991.

[2] G. D. Forney, Jr. and M. D. Trott, "The dynamics of group codes: State spaces, trellis diagrams and canonical encoders," *IEEE Trans. Inform. Theory*, vol. IT-39, pp. 1491–1513, Sept. 1993.

[3] J. G. Proakis, *Digital Communication*, 3rd ed. New York: McGraw-Hill, 1995.