
© 1990 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
TCM on Frequency-Selective Fading Channels: a Comparison of Soft-Output Probabilistic Equalizers

Peter Hoher

German Aerospace Research Establishment (DLR)
Institute for Communications Technology
D-8031 Oberpaffenhofen, West-Germany
Tel. ++49/8153/28-804, Fax ++49/8153/28-243

Abstract

We investigate suboptimal, low-complexity nonlinear equalizer structures derived from the probabilistic symbol-by-symbol MAP algorithm and the M-ary Soft-Output Viterbi Algorithm (SOVA). Both algorithms deliver reliability information for each symbol. The complexity of both algorithms in their simplest form is on the order of the conventional reduced-state Viterbi equalizer (VE) with hard outputs.

Simulation results are given for a terrestrial time-varying frequency-selective fading channel. Realistic channel estimation at high Doppler speeds is included. The coding gain of the investigated 8-state trellis-coded 8PSK scheme is about 6 dB at $10^{-9}$ BER. The gain by making use of soft-decisions is about 4-5 dB at $10^{-9}$ BER.

1 Introduction

Trellis Coded Modulation (TCM) is a power and bandwidth efficient modulation scheme [1]. The ability to combat time-selective (frequency-flat) fading when supported by interleaving makes it an interesting candidate for low-rate mobile reception. In this case performance is comparable with a D-th order diversity scheme [2, 3], where D depends on the length of the shortest path deviating from the minimum distance path. Applications for flat fading channels were given for coherent and differential coherent detection. In either case, ML-decoding is done with a soft-decision Viterbi decoder [4].

For data rates on the order of the coherence bandwidth of the channel, or larger, an equalizer must be used to compensate for intersymbol interference. Two powerful equalizer algorithms are known [5]: The probabilistic symbol-by-symbol MAP algorithm provides the MAP-probabilities for each individual channel symbol [4, Ref. 36-40 and Appendix], [5, Section 6.6], [6], [7], whereas the Viterbi algorithm (VA) [8, 4, 5] is a sequence estimator and delivers the ML-(channel)path. Both algorithms gather energy from all channel tap-gains without suffering from noise enhancement or error propagation, which is important in mobile reception, because the echoes may be stronger than the direct signal.

However, there are two disadvantages in this context: first, complexity in terms of the number of equalizer-states, $M^L$, for both algorithms increases exponentially with channel memory L, where M is the cardinality of the TCM signal space, which is twice that of the uncoded reference system. Secondly, the Viterbi equalizer (VE) outputs the most likely survivor, which appears as a hard input for the TCM-decoder, whereas the symbol-by-symbol algorithm is able to provide a soft output, but is more complex.

The first disadvantage can be reduced by making use of state reduction techniques [9] or by sequential decoding. Since a sequential decoder tends to break down on bursty channels, we favor state reduction techniques here. The second disadvantage can be avoided by delivering reliability information together with the decisions. This is the key issue of this paper.

Extracting reliability information can be performed as follows:

- By making use of channel state information (CSI) only,
- by restarting the inner VE whenever the outer TCM-decoder detects an unallowed sequence ("Generalized VA, GVA" [10]), or
- by using a “soft-output equalizer” [11, 12, 13], which may be the SOVA or the symbol-by-symbol MAP algorithm.
Using CSI to "soften" the hard equalizer output is a one-dimensional operation, which leads to a degradation if \(M > 2\). Furthermore, this simple procedure is not symbol-selective (irrespective of the CS, some channel symbols are always more reliable than others).

GVA and SOVA work just opposite: the SOVA delivers a hard decision about the path (namely the ML-path), but a soft decision about the symbols along this selected path, whereas the GVA delivers a set of paths with hard symbols along these paths. In contrast, the symbol-by-symbol MAP algorithm naturally provides no decisions at all (although most often a hard decision on the output is performed [4]).

In the following, we restrict our investigation to algorithms of the third category. We present the transmission system under consideration, summarize the known algorithms and give simplifications, and finally show simulation results including real (open-loop) channel estimation, which appear to be superior to a competitive system [14] at high Doppler speeds.

2 Transmission System

![Transmission System Diagram](image)

Figure 1: Transmission system including an equivalent discrete-time white-noise channel model and a soft-output equalizer.

The baseband-equivalent transmission system under investigation is shown in Fig. 1. An \((M/2)\)-ary i.i.d. bit stream \(u_n\) is fed into the TCM encoder. The complex-valued \(M\)-ary output symbol stream \(\{t_n\}\) will be interleaved (at least the \(M\) labels). The scrambled output sequence \(\{I_n\}\) again is i.i.d. when interleaving is perfect\(^1\). \(I_k\) is transmitted via the discrete-time equivalent channel. Irrespective of using a receiver with an adaptive whitening matched filter [8, 5] or with a fixed receiving filter, the observed signal is [18]

\[
  z_k = \sum_{l=-L}^{L} f^{(l)} I_{k-l} + \eta_k; \quad L := L_- + L_+,
\]

where the \(L + 1\) complex tap-gains \(f^{(l)}\) represents the transmitter- and receiving filter and the time-varying

\(^1\)Unfortunately, it is impossible to deliver TCM-subset information to the channel when restriction is made to a recursive equalizer algorithm [18].

3 Algorithms

3.1 Symbol-by-Symbol MAP Equalizer

Suppose for the moment that the symbols \(I_k := M(i_k)\) are transmitted in finite blocks of length \(N + 1\), where \(i_k \in M\) and \(M\) is the mapping. The task is to compute the MAP probabilities \(P(i_k | I_{0}^{N})\), \(0 \leq k \leq N\), given the observations \(I_{0}^{N} := \{I_0, \ldots, I_k, \ldots, I_N\}\).

Let \(\mu_k\) denote the \(S = M^K\) states at time \(k\), which are considered by a reduced-state algorithm \((K \leq L_+)\), and let \(T(i_{k-1})\) be the set of all states, for which the transmitted symbol is \(i_{k-1}\). Then, we have

\[
P(i_{k-1} | I_{0}^{N}) = \frac{\sum_{\mu_k \in T(i_{k-1})} P(\mu_k, I_{0}^{N})}{\sum_{\mu_k} P(\mu_k, I_{0}^{N})}.
\]

(2) \(P(\mu_k, I_{0}^{N})\) can be factorized

\[
P(\mu_k, z_{0}^{N}) = P(\mu_k, z_{0}^{k-1}) \cdot P(z_{k}^{N} | \mu_k),
\]

(3) and each term can be solved recursively [4, Appendix]

\[
P(\mu_k, z_{0}^{N}) = \left( \sum_{\mu_{k-1}} P(\mu_{k-1}, z_{0}^{k-2}) \cdot e^{(-\lambda(I_{k-1}))} \right) \cdot \left( \sum_{\mu_{k+1}} P(z_{k}^{N} | \mu_{k+1}) \cdot e^{(-\lambda(I_{k}))} \right),
\]

(4)

where

\[
e^{(-\lambda(I_{k}))} := P(\mu_{k+1} | \mu_k) \cdot P(z_k | \mu_k) \sim \exp \left( \frac{\left[ z_k - \sum_{l=K}^{K} f^{(l)} I_{k-l} - \sum_{l=K+1}^{L} f^{(l)} I_{k-l} \right]^2}{2\sigma_k^2} \right)
\]

(5)

is the exponentiated metric increment from \(k\) to \(k + 1\), \(\hat{I}_k\) are the (trial) symbols according to the transitions

\(^2\)This imposes certain conditions on the signals in the case of a fixed receiving filter, which are fulfilled in the following [18].
\( \xi_k := (\mu_k - \mu_{k+1}), \hat{I}_k \) are previous symbols along the path history (see [9]), and \( \sigma^2 \) is the instantaneous noise variance. Finally, we obtain the \( M \)-ary vector [4]

\[
(- \log P(i_k | z_0^N); 1 \leq i_k \leq M)
\]

(6)

for each \( k \), where \( 0 \leq k \leq N \). Initialization depends on whether the trellis is tailed or not. Note that the first factor in (4) represents a forward recursion, which corresponds to the survivor extension in a conventional VA [4]. The second factor is simply time-reversed because of the symmetry of the trellis and represents a backward recursion, which is the counterpart of the back-search procedure in a VA.

The basic advantage compared to the VA is the inherent soft output. The disadvantages are the computation in the probability domain instead of the logarithmic (metric) domain, the excessive storage, and the decision delay. In block-mode, two recursions are required instead of one recursion plus back-search or register exchange. When block-length \( N \) becomes large, “real-time” algorithms are needed, which are much more complex [7, 4, Appendix].

We solve the first disadvantage as follows: since the sum of \( S \) exponential terms \( \zeta \), is bounded by

\[
\max [\zeta_0, \ldots, \zeta_{S-1}] \leq \log_e \sum_{i=0}^{S-1} \exp(\zeta_i) \leq \log_e S + \max [\zeta_0, \ldots, \zeta_{S-1}],
\]

(7)

we get from (4)

\[
\begin{align*}
\max & \log_e P(\mu_{k-1}, z_0^{k-2}) - \lambda(\xi_{k-1}) + \\
+ & \max \log_e P(\mu_{k+1}^{N} | \mu_{k+1}) - \lambda(\xi_k) \\
\leq & \log_e P(\mu_k, z_0^N) = \log_e \left( \sum_{\mu_{k-1}} \ldots \right) + \log_e \left( \sum_{\mu_{k+1}} \ldots \right) \\
\leq & 2 \log_e S + \\
+ & \max \log_e P(\mu_{k-1}, z_0^{k-2}) - \lambda(\xi_{k-1}) + \\
+ & \max \log_e P(\mu_{k+1}^{N} | \mu_{k+1}) - \lambda(\xi_k).
\end{align*}
\]

(8)

In most cases the degradation is small, when the lower bound is selected. Now, computation is in the logarithmic domain, which allows fixed-point implementation. Another advantage is that the forward recursion now provides a survivor for each state. Hence, a further hard decision for ISI-truncation [9] or for synchronization purposes is not needed.

The second and third disadvantages are solved as follows: when the equivalent discrete-time channel model is minimum phase or nearly so (this can be controlled by an adaptive receiving filter), and when the channel is not pathological, maximum distance gain (on the average) is achieved when the trellis paths diverge, whereas Euclidean distance is smaller when the paths merge (the latter point does not apply to good channel codes). However, then the contribution of the second factor in (4) is small. Hence, we reduce backward extension to a limited range. Two cases are of special interest:

1. We omit the backward extension. Hence, the simple forward recursion leaves:

\[
\begin{align*}
\log_e P(\mu_{k+1}, z_0^k) & \approx \max_{\mu_{k+1}} \left[ \log_e P(\mu_k, z_0^{k-1}) - \lambda(\xi_k) \right] \\
& \approx \max_{\mu_{k+1}} \left[ \log_e P(\mu_k, z_0^{k-1}) - \\
& \quad - \frac{1}{2 \sigma^2} |z_k - \sum_{t=0}^{L-1} f(k) \hat{I}_k (t)^2 |. \right.
\end{align*}
\]

(9)

This formula was also mentioned by Koch [16]. Costs in terms of complexity, memory and delay are even less than for the conventional VA, but multipath diversity cannot be achieved.

2. Improvement is possible by additionally performing back-search over one interval:

\[
\begin{align*}
\log_e P(\mu_k, z_0^k) & \approx \max_{\mu_k} \left[ \log_e P(\mu_{k-1}, z_0^{k-2}) - \lambda(\xi_{k-1}) \right] - \lambda(\xi_k). \right.
\end{align*}
\]

(10)

This formula is of special interest, because complexity is also less than for the conventional VA. Compared to case 1, no additional memory is required, but multipath diversity is achievable.

We show simulation results in Section 4.

3.2 Soft-Output Viterbi Equalizer

The symbol-by-symbol MAP equalizer makes no attempt to restore a valid path sequence. This is in contrast to the Soft-Output Viterbi Algorithm (SOVA). By expanding the conventional VA with a “soft-output unit”, it is possible to deliver the ML-path together with reliability information individually for each symbol [11, 12, 13]. In the binary case, the degradation compared with the optimum symbol-by-symbol MAP algorithm is small. However, in the nonbinary case we observe a significant degradation when realizing the algorithm derived in [13], as shown later. Therefore, we introduce a modified \( M \)-ary algorithm as follows.

Suppose we are given the ISI-trellis of Fig. 2. For each state \( \mu_k, 0 \leq \mu_k \leq S-1, \) the conventional VA extends the \( M \) previous states by adding the new metric
The novelty of the algorithm introduced above is that updating is not restricted to the symbols of the survivor path [13]. Hence, it is possible to output an M-ary reliability vector instead of a scalar. In summary, the modified M-ary Soft-Output VA (M-ary SOVA) is as follows:

- For all states, determine survivor \( \mathcal{A} \)
- For all paths \( \mathcal{A}, \ldots, \Omega \) for the “relevant length”
- For the “relevant symbols”
- Update stored reliability
- Output final decision together with the M-ary reliability vector.

Unfortunately, especially in the nonbinary case, the algorithm is much too complex (the symbol-by-symbol MAP algorithm behaves better in this sense). Complexity reduction at the cost of suboptimality is possible as follows: the number of update operations can be reduced significantly when the update is done for only a limited number of paths and or a limited number of time slots [11]. One special case is to update only the globally best survivor path at a given time \( k \) for the relevant length (“one-dimensional horizontal update”). Another special case is where updating is done for each state, but only one slot in depth (“one-dimensional vertical update”) [11]. Since the ISI-trellis is of the shift-register type, the information symbols always differ at time \( k - (K + 1) \). Updating only the symbols in the time-slot from \( k - (K + 1) \) to \( k - K \) has computational advantages (update is only once, load is constant, pairs to be compared can be precomputed), and analysis is easier. Another strategy would be to update only the symbols where the paths diverge, motivated by the above mentioned distance gain. But this procedure does not insure that all symbols of the finally selected survivor are updated at least once. In Section 4 we give results for the one-dimensional vertical update.

### 4 Simulation Results

The different equalizers will be compared now on the basis of the channel given in Table 1.

This channel is the discrete-time substitute of a terrestrial mobile channel in a bad urban area as specified by CEPT COST-207, where delay is up to 10 ms, and bit duration is \( T_b = 3.7 \) ms, which corresponds to the pan-European cellular mobile system GSM. Pulse shaping is a square-root raised-cosine filter with rolloff \( r = 0.4 \), instead of Gaussian MSK as proposed by

---

**Figure 2:** 4-state ISI-trellis with memory \( K = 1 \) and \( M = 4 \) transitions per state.

- Increments \( \lambda(\xi_{k-1}) \) to the stored accumulated metrics \( \Gamma(\mu_{k-1}) \), compares these sums, and selects the path with the lowest metric:

\[
\Gamma_A(\mu_k) = \min_Y [\Gamma(\mu_{k-1}) + \lambda(\xi_{k-1})],
\]

where \( Y \in \{A, \ldots, \Omega\} \) denote the \( M \) paths and \( A \) is the survivor. The probability of the \( Y \)-th path is

\[
P_Y := P(e_k^{\mu_{k-1}}| \text{path } Y) \sim e^{-\Gamma(\mu_k)}.
\]

By applying Bayes' theorem, we get the probability \( P_X \), of the \( j \)-th symbol \( X_j \in \{A, \ldots, \Omega\} \) of path \( Y \in \{A, \ldots, \Omega\} \)

\[
P_{X_j} = \frac{\sum_{Y \in \mathcal{Y}_j} P_Y}{\sum_{Y \in \{A, \ldots, \Omega\}} P_Y},
\]

where \( \mathcal{Y}_j \) is the set of all paths for which the \( j \)-th symbol \( X_j \) corresponds to the \( j \)-th symbol of path \( Y \), \( j \leq k \). For the survivor path \( \mathcal{A} \) we have \( 1/M \leq P_{X_j} \leq 1 \), for the second best path we have \( 0 \leq P_{X_j} \leq 1/2 \), and so on until \( 0 \leq P_{X_j} \leq 1/M \) for the worst path.

Let \( P_{X_{j, \text{hist}}} \) be the stored probability for the symbol \( X_j \), initialized by one. Since a symbol \( X_j \) is correct, if it was previously correct and the path is right, we have the recursive update [13]

\[
P_{X_{j, \text{hist}}} \leftarrow P_{X_{j, \text{hist}}} \cdot P_{X_j}.
\]

This updating procedure has to be performed whenever at least two symbols differ. These are, what we call, the “relevant symbols”. The “relevant range” where symbols may differ is from \( k - \delta_m \) to \( k - K \), compare Fig. 2.

Finally, at time \( k \) we output the M-ary vector \((-\log P_{X_{j, \text{hist}}} ; 1 \leq X \leq M)\), where the decision delay \( \delta \) is such that all paths have converged.

A simplified version of (13) is to compute

\[
\log P_{X_{j, \text{hist}}} \leftarrow \min \{ \log P_{X_{j, \text{hist}}} ; \log P_{X_j} \},
\]

which simplifies fixed-point arithmetic [11, 12, 13].
| $l$, $-L_{-} \leq l \leq L_{+}$ | $E[|f_k^{(l)}|^2]$ | $\sum_{j=0}^{L_{+}} E[|f_k^{(l)}|^2]$ |
|-----------------|------------------|-----------------|
| -1              | -19.3 dB         | 0.008           |
| 0               | 0 dB             | 0.667           |
| 1               | -3.0 dB          | 0.991           |
| 2               | -19.0 dB         | 0.999           |
| 3               | -29.7 dB         | 1.000           |

Table 1: Average power of the channel tap-gains (GSM "bad urban area (BU)", $T_b = T_s/2 = 3.7 \mu s$, $r = 0.4$).

GSM. The receiving filter is matched to the transmitting filter, and hence fixed. Each tap is a Rayleigh-fading process. Correlation between the tap-gains was taken into account [15]. From Table 1 we recognise that $L = 3$ echoes have to be taken into account when the cutoff is about -25 dB.

We investigated trellis-coded 8PSK (C8PSK) and the uncoded 4PSK reference system. The selected code was the 8-state Ungerboeck code in feedforward realization; this code is the simplest nontrivial code with no parallel transitions, but still the best known one-dimensional 8-state code for both AWGN [1] and flat fading [3]. The length of the shortest path which deviates from the minimum distance path is 2. The decision delay was 24 symbols.

A conventional VE would require $M^L = 8^3 = 512$ states, which is very complex. By using the principles of state truncation [9], we restrict $M^L$ to $8^1 = 8$ states. $K = 1$ ($K = 2$) suffices so that approximately 98.3% (99.1%) of the energy of the tap-gains are taken into account on the average, compare Table 1. The precursor tap ($L_{-} = 1$) was truncated, and the remaining successors where cancelled according to [9] instead of truncated.

Simulation results assuming perfect knowledge of the channel tap-gains are given in Fig. 3. The top curve is for a conventional 8-state VE with hard output. The next curve is for a soft-output 8-state equalizer with scalar output according to [13], but computation is restricted to the above introduced 1-dimensional vertical update. Also, computation is simplified to a comparison between the survivor and the second best path. The gain to the hard-output equalizer is negligible. By using a 1-dimensional vertical update and $M$-ary output, a further 2 dB improvement at $10^{-3}$ BER is possible, which belongs to the next curve.

Simplified versions of the symbol-by-symbol MAP algorithm are more powerful. The lowest curve is for a 64-state MAP algorithm with no further simplifications. Hence, up to 5 dB at $10^{-3}$ can be gained by making use of soft-decisions. The second best curve is for an 8-state MAP algorithm, where backward extension is reduced to one step. Degradation is small, when the maximum rule (10) is used instead of exponentials, which is the next curve. This equalizer is even less complex than the VE of the top curve. Nevertheless, 4 dB are gained by soft-decisions.

For comparison, the uncoded reference system is also plotted. The gain between the coded scheme with an optimum 64-state MAP equalizer and the reference system with a 16-state VE is more than 6 dB at $10^{-3}$. The broken lines are for flat fading and are also plotted for convenience. The multipath diversity gain is seen to be about 1-2 dB at $10^{-3}$ for the coded system. We have to keep in mind that the selected trellis code is a simple one. Also, we have shown earlier that the gains are more significant when going from uncoded 2PSK to trellis-coded 4PSK [17].

Fig. 4 presents simulation results of the most promising algorithms discussed so far including real channel estimation at Doppler speeds of 208 Hz ($f_0=900$ MHz, $v=250$ km/h). Since closed-loop channel estimation is very critical, we computed the tap-gains only during short training segments, which were inserted periodically into the data sequence. Between these probes interpolation of the tap-gains was done according to the sampling theorem, see [15] for details. The redundancy of 20 % training (-1 dB) is not included in Fig. 4. Note that the normalized Doppler frequency exceeds that of [14, 1987] by more than one order.

5 Summary

The key issue of this paper was the investigation of nonbinary soft-output symbol-by-symbol and se-
sequence estimators. Starting from the optimum structures, reduced complexity receivers were derived. The algorithms allow concatenation of multiple "codes". Application was given for equalizing a frequency-selective fading channel, having an outer TCM coding scheme. Channel estimation was included in the investigations. About 4-5 dB gain was shown at $10^{-3}$ by making use of the soft-decisions. Under the constraint of similar complexity, the modified symbol-by-symbol estimator outperforms the sequence estimator when the signal space is nonbinary. Multipath diversity can be achieved at the cost of complexity by extending the backward recursion at a given number of states, irrespectively of the prefilter. This is in contrast to decision-feedback equalization schemes such as [14] or [5, Section 6.5.2], where multipath diversity can only be resolved by adaptive prefiltering (once the equivalent channel is settled, all the energy of the echoes will be cancelled there). This, however, is in contradiction with an efficient channel estimation algorithm with interpolation.

Having soft input and output (compound channel output), the equalizer can be interpreted as a filter. This fills the gap to conventional linear transversal equalizers or nonlinear decision-feedback equalizers. In general, each decoder working on a trellis can now be interpreted as a filter [12].

References