A New SISO Algorithm with Application to Turbo Equalization

Marcin Sikora and Daniel J. Costello, Jr.
Department of Electrical Engineering, University of Notre Dame
275 Fitzpatrick Hall, Notre Dame, IN 46556
Email: {msikora|dcostel1}@nd.edu

Submitted to the ISIT 2005 on December 13th, 2004

Abstract—In this paper we propose a new soft-input soft-output equalization algorithm, offering very good performance/complexity tradeoffs. It follows the structure of the BCJR algorithm, but dynamically constructs a simplified trellis during the forward recursion. In each trellis section, only the \( M \) states with the strongest forward metric are preserved, similar to the M-BCJR algorithm. Unlike the M-BCJR, however, the remaining states are not deleted, but rather merged into the surviving states. The new algorithm compares favorably with the reduced-state BCJR algorithm, offering better performance and more flexibility, particularly for systems with higher order modulations.

I. INTRODUCTION

Efficient communication over channels introducing inter-symbol interference (ISI) often requires the receiver to perform channel equalization. Turbo equalization [1] is a technique in which decoding and equalization are performed iteratively, similar to turbo-decoding of serially-concatenated convolutional codes [2]. As depicted in Figure 1, the key element of the receiver employing this method is a soft-input soft-output (SISO) demodulator/equalizer (from now on referred to as just an equalizer), accepting a priori likelihoods of coded bits from the SISO decoder, and producing their a posteriori likelihoods based on the noisy received signal.

The SISO algorithm that computes the exact values of the a posteriori likelihoods is the BCJR algorithm [3]. The complexity of a BCJR equalizer is proportional to the number of states in the trellis representing the modulation alphabet and the ISI, and thus it is exponential in both the length of the channel impulse response (CIR) and in the number of bits per symbol in the modulator. This can be a serious drawback in some scenarios, e.g., transmission at a high data rate over a radio channel, where a large signal bandwidth translates to a long CIR, and a high spectral efficiency translates to a large modulation alphabet. Needed in such cases are alternative SISO equalizers with the ability to achieve large complexity savings at a cost of small performance degradation.

There have been two main trends in the design of such SISOs. The first one relies on reducing the effective length of the channel impulse response, either by linear processing (see, e.g., [4]), or interference cancellation via decision feedback. A particularly good algorithm is this category is the reduced-state BCJR (RS-BCJR) [5], which performs the cancellation of the final channel taps on a per-survivor basis. Iterative decoding with RS-BCJR is very stable, thanks to the high quality of the soft outputs, but the receiver cannot use the signal power contained in the cancelled part of the CIR. Another trend is to adapt “hard-output” sequential algorithms [6] to produce soft outputs [7]. Examples in this category are the M-BCJR and T-BCJR algorithms [8], based on the M- and T-algorithms, and the LISS algorithm [9] based on list sequential decoding. These algorithms have no problem using the signal energy from the whole CIR, and offer much more flexibility in choosing the desired complexity. However, their reliance on ignoring unpromising paths in the trellis or tree causes a bias in the soft output (there are more explored paths with one value of a particular input bit than another), which negatively affects the convergence of iterative decoding.

In this paper we present a new SISO equalization algorithm, inspired by both the M-BCJR and RS-BCJR, which shares many of their advantages, but few of their weaknesses. We call this algorithm the M*-BCJR algorithm, since it resembles the M-BCJR in preserving only a fixed number of trellis states with the largest forward metric. Instead of deleting the excess states, however, the M*-BCJR dynamically merges them with the surviving states — a process that shares some similarity to the static state merging done on a per-survivor basis by the RS-BCJR. For the sake of simpler notation, we present the operation of all BCJR-based algorithms, including the M*BCJR, in the probability domain. Each of them, however, can be implemented in the log domain for better numerical stability.

The rest of the paper is structured as follows. Section 2 describes the communication system and the task of the SISO equalizer and introduces the notation. Section 3 reviews the structure of the BCJR, M-BCJR, and RS-BCJR algorithms, helping us to introduce the M*-BCJR in Section 4. Section 5 presents simulation results, and conclusions are given in Section 6.

II. COMMUNICATION SYSTEM

A communication system with turbo equalization is depicted in Figure 1. The information bits are first arranged into blocks and encoded with a convolutional code. The blocks of coded bits are permuted using an interleaver and mapped onto a sequence of complex symbols by the modulator. (In general, the modulator can have memory, but for simplicity we will assume a memoryless mapper.) The channel acts as a
discrete-time finite impulse response (FIR) filter introducing ISI, the output of which is further corrupted by additive white Gaussian noise (AWGN). We assume the receiver knows the ISI channel coefficients and the noise variance, and it attempts to recover the information bits by iteratively performing SISO equalization and decoding.

The part of the system significant from the point of view of the equalizer is shown in Figure 2. Let \( \mathbf{a} = (a_1, a_2, \ldots, a_L) \) denote a sequence of \( LK \) bits entering the modulator, arranged into \( L \) groups \( a_i = (a_{i1}, a_{i2}, \ldots, a_{iK}) \) of \( K \) bits. Each \( K \)-tuple \( a_i \) selects a complex-valued output symbol \( x_i \) from a constellation of size \( 2^K \) to be transmitted. The sequence of symbols \( y = (y_1, y_2, \ldots, y_{L+S}) \) obtained at the receiver is modeled as

\[
y_i = \sum_{j=0}^{S} h_j x_{i-j} + n_i,
\]

(1)

where \( S \) is the memory of the channel, \( h_j, j = 0, 1, \ldots, S \), are the channel coefficients, and \( n_i, i = 1, 2, \ldots, L + S \), are i.i.d. zero-mean complex-valued Gaussian random variables with variance \( \sigma^2 \) per complex dimension. Equation (1) assumes that \( x_i \) is zero outside \( i = 1, 2, \ldots, L \).

The SISO equalizer for the above channel takes the received symbols \( y \) and the a priori log-likelihood ratios \( \Lambda_a(a^k_i) \) for each bit \( a^k_i \), defined as

\[
\Lambda_a(a^k_i) = \log \frac{P(a^k_i = +1)}{P(a^k_i = -1)},
\]

(2)

and outputs the a posteriori L-values \( L(a^k_i) \)

\[
L(a^k_i) = \log \frac{P(a^k_i = +1 | y)}{P(a^k_i = -1 | y)}.
\]

(3)

The values actually fed to the SISO decoder are extrinsic L-values, computed as \( L_x(a^k_i) = L(a^k_i) - \Lambda_a(a^k_i) \).

Let \( \Lambda(a) \) denote the joint probability that \( a \) was transmitted and \( y \) was received. Then \( \Lambda(a) \) can be expressed as

\[
\Lambda(a) = \sum_{\mathbf{a}, \mathbf{b}} \Lambda(a \mid \mathbf{b}) \gamma(s_i, a_i) \beta(s_{i+1}),
\]

(10)

where the summations are performed over all \( a \) consistent with \( a^k_i = \pm 1 \). Furthermore,

\[
\Lambda(a) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{1}{\sigma^2} \sum_{j=0}^{S} h_j x_{i-j} \right),
\]

(5)

where \( h_j, j = 0, 1, \ldots, S \), and \( \sigma^2 \) are assumed known at the receiver and \( P(a) \) is obtained from \( L_a \) as

\[
P(a) = \prod_{i=1}^{L} \prod_{k=1}^{K} P(a^k_i),
\]

(6)

with

\[
P(a^k_i = \pm 1) = \exp \left( \pm L_a(a^k_i) \right) \frac{1}{1 + \exp(\pm L_a(a^k_i))}.
\]

(7)

Since the number of paths involved in the summations of \( \Lambda(a) \) is extremely large for realistic values of \( K \) and \( L \), a practical algorithm seeks to simplify or approximate this calculation.

### III. SISO EQUALIZATION

#### A. The BCJR algorithm

The classical algorithm for efficiently computing \( \Lambda(a) \) by exploiting the trellis structure of the set of all paths is the BCJR algorithm [3]. By defining the state \( s_i \) at time \( i \) as the past \( S \) input symbol \( K \)-tuples \( \mathbf{a}_i, s_i = (a_{i-1}, \ldots, a_{i-S}) \), and a branch metric \( \gamma(s_i, a_i) \) as

\[
\gamma(s_i, a_i) = P(a_i) \sqrt{2\pi\sigma^2} \exp \left( -\frac{1}{2\sigma^2} \sum_{j=0}^{S} h_j x_{i-j} \right),
\]

(8)

the path metric can be factored into

\[
\Lambda(a) = \prod_{i=1}^{L+S} \gamma(s_i, a_i).
\]

(9)

For indices outside the range \( i = 1, \ldots, L \), the variables \( a_i \) are regarded as empty sequences \( \phi \) with \( P(a_i = \phi) = 1 \).

For every trellis branch \( b_i = (s_i, a_i, s_{i+1}) \) starting in state \( s_i \), labeled by input bits \( a_i \), and ending in state \( s_{i+1} \), the BCJR algorithm computes the sum of the path metrics \( \Lambda(a) \) over all paths passing through this branch as

\[
\sum_{\mathbf{a}, \mathbf{b}_i} \Lambda(a) = \alpha(s_i) \gamma(s_i, a_i) \beta(s_{i+1}).
\]

(10)

The computation of the forward state metrics \( \alpha(s_i) \) is performed in the forward recursion for \( i = 1, 2, \ldots, L + S - 1 \):

\[
\alpha(s_{i+1}) = \sum_{b_i = (s_i, a_i, s_{i+1})} \alpha(s_i) \gamma(s_i, a_i),
\]

(11)

with the initial state value \( \alpha(s_1) = 1 \). Similarly, the backward recursion computes the backward state metrics \( \beta(s_i) \) for \( i = L + S, L + S - 1, \ldots, 2 \):

\[
\beta(s_i) = \sum_{b_i = (s_i, a_i, s_{i+1})} \gamma(s_i, a_i) \beta(s_{i+1}),
\]

(12)
with the terminal state value \( \beta(s_{L+S+1}) = 1 \). With all \( \alpha \)'s, \( \beta \)'s, and \( \gamma \)'s computed, the summations over paths in (12) can be replaced by the summations over branches,

\[
L(a_k^i) = \log \frac{\sum b_i^s a_{k+1}^s = +1 \alpha(s_i) \gamma(s_i, a_i) \beta(s_{i+1})}{\sum b_i^s a_{k+1}^s = -1 \alpha(s_i) \gamma(s_i, a_i) \beta(s_{i+1})}.
\]

The completion phase, in which (15) is evaluated for every \( a_k^i \), concludes the algorithm.

The complexity of the BCJR equalizer is proportional to the number of trellis states, \( 2^{KS} \). The following subsections describe the operation of the RS-BCJR [5] and M-BCJR [8] algorithms, which preserve the general structure of the BCJR, but instead operate on dynamically built simplified trellises with a number of states controlled via a parameter. In the original form of both algorithms, the construction of this simplified trellis occurs during the forward recursion and is based on the values of the forward state metrics, while the backward recursion and the completion phase just reuse the same trellis.

### B. The RS-BCJR algorithm

The way we will describe the operation of the RS-BCJR algorithm is slightly different from the presentation in [5], but is in fact equivalent.

Let us consider two states in the trellis,

\[
s_i = (a_{i-1}, ..., a_{i-S'}, a_{i-S'-1}, ..., a_{i-S}),
\]

\[
s'_i = (a_{i-1}, ..., a_{i-S'}, a'_{i-S'-1}, ..., a'_{i-S}),
\]

differing only in the last \( S' \) binary \( K \)-tuples. Furthermore, consider two partial paths beginning in states \( s_i \) and \( s'_i \) and corresponding to the same partial input sequence \( a_{i[L]} = (a_i, ..., a_L) \). Both paths are guaranteed to merge after \( S' \) time indices, and hence their partial path metrics are

\[
\lambda(s_i, a_{i[L]}) = \prod_{j=i}^{i+S'-1} \gamma(s_j, a_j) \prod_{j=i+S'-S'}^{L} \gamma(s_j, a_j),
\]

\[
\lambda(s'_i, a_{i[L]}) = \prod_{j=i}^{i+S'-1} \gamma(s'_j, a_j) \prod_{j=i+S'-S'}^{L} \gamma(s_j, a_j).
\]

Additionally, close examination of (8) reveals that the difference between \( \gamma(s_j, a_j) \) and \( \gamma(s'_j, a_j) \) for \( j = i, ..., i+S'-S'-1 \) is not large. Hence, the difference between \( \lambda(s_i, a) \) and \( \lambda(s'_i, a) \), for \( a_{i[L]} \), is also not large.

The RS-BCJR equalizer relies on the above observation and, for some predefined \( S' \), declares states differing only in the last \( S' \) binary \( K \)-tuples indistinguishable. Every such set of states is subsequently reduced to a single state, by selecting the state with the highest forward metric and merging all remaining states into it. Here, we define merging of the state \( s'_i \) into \( s_i \) as updating the forward metric \( \alpha(s_i) := \alpha(s_i) + \alpha(s'_i) \), redirecting all trellis branches ending at \( s'_i \) into \( s_i \), and deleting \( s'_i \) from the trellis. This reduction is performed during the forward recursion, and the \( \gamma \)'s for the paths that originate from removed states need never be computed. The trellis that results has only \( 2^{KS} \) states, compared to \( 2^{KS'} \) in the original trellis. The same trellis is then reused in the backward recursion and the completion stage.

The RS-BCJR equalizer is particularly effective when the final coefficients of the ISI channel are small in magnitude. Furthermore, the reduced-state trellis retains the same branch-to-state ratio (branch density) and has the same number of branches with \( a_k^i = +1 \) and \( a_k^i = -1 \) for any \( i \) and \( k \) — properties that ensure a high quality for the soft outputs and good convergence of iterative decoding. Unfortunately, the RS-BCJR algorithm cannot use the signal power in the final \( S' \) channel taps, effectively reducing the minimum Euclidean distance between paths. Moreover, the number of surviving states can only be set to a power of \( 2^K \), which could be a problem for large \( K \) (e.g., for a system with 16QAM modulation, equalization using 16 states could result in poor performance, while 256 states could exceed acceptable complexity).

### C. The M-BCJR algorithm

The M-BCJR algorithm is based on the M-algorithm [6], originally designed for the problem of maximum likelihood sequence estimation. The M-algorithm keeps track only of the \( M \) most likely paths at the same depth, throwing away any excess paths. In the M-BCJR equalizer this idea is applied to the trellis states during the forward recursion. At every level \( i \), when all \( \alpha(s_i) \) have been computed, the \( M \) states with the largest forward metrics are retained, and all remaining states are deleted from the trellis (together with all the branches that lead to or depart from them). The same trellis is then reused in the backward recursion and completion phase.

In [8] it was shown that the M-BCJR algorithm performs well when the state reduction ratio \( 2^{KS}/M \) is not very large. Also, unlike the RS-BCJR algorithm, it can use the power from all the channel taps. For small \( M \), however, the reduced trellis is very sparse, i.e., the branch-to-state ratio is much smaller than in the full trellis and there is often a disproportion between the number of branches labeled with \( a_k^i = +1 \) and \( a_k^i = -1 \) for any \( i \) and \( k \). These factors reduce the quality of the soft outputs and the convergence performance and may require an alternative way of computing the a posteriori likelihoods (like the Bayesian estimation approach presented in [10]). Finally, the M-BCJR algorithm requires performing a partial sort (finding the \( M \) largest elements out of \( M2^K \)) at every trellis section, which increases the complexity per state.

### IV. THE M*-BCJR ALGORITHM

In this section we demonstrate how the concept of state merging present in the RS-BCJR equalizer can be used to enhance the performance of the M-BCJR algorithm. We call the resulting algorithm the M*-BCJR algorithm.

During the forward recursion the M*-BCJR algorithm retains a maximum of \( M \) states for any time index \( i \). Unlike the M-BCJR algorithm, however, the excess states are not deleted, but merely merged into some of the surviving states. This means that none of the branches seen so far are deleted from the trellis, but they are just redirected into a more likely state.
The forward recursion of the algorithm can be described as follows:

1. Set $i := 1$. For the initial trellis state $s_1$, set $\alpha(s_1) := 1$. Also, fix the set of states surviving at depth 1 to be $S_1 := s_1$.
2. Initialize the set of surviving states at depth $i + 1$ to an empty set, $S_{i+1} = \emptyset$.
3. For every state $s_i$ in the set $S_i$, and every branch $b = (s_i, a_i, s_{i+1})$ originating from that state, compute the metric $\gamma(s_i, a_i)$, and add $s_{i+1}$ to the set $S_{i+1}$.
4. For every state $s_{i+1}$ in $S_{i+1}$ compute the forward state metric as a sum of $\alpha(s_i)\gamma(s_i, a_i)$ over all branches $b = (s_i, a_i, s_{i+1})$ visited in step 3 that end in $s_{i+1}$.
5. If the number of states in $S_{i+1}$ is no more than $M$, proceed to step 8. Otherwise continue with step 6.
6. Determine the $M$ states in $S_{i+1}$ with the largest value of the forward state metric. Remove all remaining states from $S_{i+1}$ and put them in a temporary set $S'_{i+1}$.
7. Go over all states $s'_{i+1}$ in the set $S'_{i+1}$ and perform the following tasks for each of them:
   - Find a state $s_{i+1}$ in $S_{i+1}$ that differs from $s'_{i+1}$ by the least number of final $K$-tuples $a_i$.
   - Redirect all branches ending in $s'_{i+1}$ to $s_{i+1}$.
   - Add $\alpha(s'_{i+1})$ to the metric $\alpha(s_{i+1})$.
   - Delete $s'_{i+1}$ from the set $S'_{i+1}$.
8. Increment $i$ by 1. If $i \leq L+S−1$, go to step 2. Otherwise the forward recursion is finished.

The merging of $s'_{i}$ into $s_{i}$ in step 7 is also illustrated in Figure 3. The backward recursion and the completion phase are subsequently performed only over states remaining in the sets $S_i$ and only over visited branches (i.e., branches for which the metrics $\gamma$ were calculated in step 3).

Just as for the M-BCJR, the $M^*\text{-BCJR}$ algorithm can use the power from all channel taps and offers full freedom in choosing the number of surviving states $M$. At the same time, the $M^*\text{-BCJR}$ never deletes visited branches, and hence it retains the branch density of the full trellis and avoids a disproportion between the number of branches labeled with $a^k = +1$ and $a^k = −1$. As a result, the soft outputs generated by the $M^*\text{-BCJR}$ equalizer ensure good convergence of the iterative receiver. Complexity-wise, the algorithm requires some additional processing per state (due to step 7) and some additional memory per branch (the ending state must be remembered for each branch). However, if we regard the calculation of the branch metrics $\gamma$ as the dominant operation, the complexities of the M-BCJR, RS-BCJR, and $M^*\text{-BCJR}$ equalizers are the same for fixed $M = 2^KS^*$.

V. Simulation results

To evaluate the performance of the $M^*\text{-BCJR}$ equalizer, we considered two turbo-equalization systems. Both systems used a recursive, memory 5, rate 1/2 terminated convolutional code as an outer code. The first system used BPSK modulation and a 5-tap channel (maximum 16 states), and a block of 507 information bits (size 1024 DRP [11] interleaver). The second system used 16QAM modulation, but only a 3-tap channel (maximum 256 states), and a block of 2043 information bits (size 4096 DRP interleaver). The remaining parameters and the channel impulse responses are summarized in Table I.

Both systems were simulated with the $M^*\text{-BCJR}$ and RS-BCJR equalizers, for several values of $M$ and $S'$. In each case we allowed the receiver to perform 6 iterations. The bit error rates $P_e$ for a range of $E_b/N_0$ (average energy per bit over noise spectral density) are plotted in Figure 4. To better illustrate the complexity-performance tradeoffs achievable with both algorithms, we also plotted the number of states $M$ or $2^KS^*$ against the $E_b/N_0$ needed to achieve certain $P_e$ ($10^{-4}$ for system 1 and $10^{-3}$ for system 2) in Figure 5.

The simulations demonstrate the superior performance of the $M^*\text{-BCJR}$ equalizer. In scenario 1, the $M^*\text{-BCJR}$ equalizer with 3 states outperforms the RS-BCJR with 8 states by 0.1 dB for $P_e$ below $10^{-4}$. When both algorithms use 4 states, the $M^*\text{-BCJR}$ equalizer offers a 0.7 dB gain compared to the RS-BCJR. In scenario 2, the $M^*\text{-BCJR}$ with 16 states achieves almost a 3 dB gain over the RS-BCJR with the same number of states.

VI. Summary

We have examined the problem of complexity reduction in turbo equalization for systems with large constellation sizes and/or long channel impulse responses. We have defined the operation of merging one state into another and used it to give an alternative interpretation of the RS-BCJR algorithm. Finally we modified the M-BCJR algorithm, replacing the

---

**TABLE I**

| Simulated turbo-equalization scenarios. |
|----------------------------------------|
| Scenario 1                             | Scenario 2 |
| Outer code                             | CC(2,1,5)  | CC(2,1,5)  |
| Modulation                             | BPSK       | 16QAM      |
| Channel memory $S$                     | 4          | 2          |
| CIR                                    | $\{0.45, 0.25\}$, $\{0.15, 0.71, 0.09\}$ | $\{1, 1, 1\}$ |
| BCJR states                            | 16         | 256        |
| Interleaver size                       | 1024       | 4096       |
| No. of iterations                      | 6          | 6          |
deletion of excess states by the merging of these states into the surviving states. The resulting algorithm, called the M*-BCJR algorithm, was shown to generate reduced-complexity trellises more suitable for SISO equalization than those obtained by the RS-BCJR and M-BCJR algorithms. Simulation results demonstrated very good performance for turbo-equalization systems employing the M*-BCJR, exceeding that of the RS-BCJR even with much smaller complexities.

REFERENCES
[1] R. Koetter, A. C. Singer, and M. Tuchler, “Turbo equalization,” IEEE Signal Proc. Mag., pp. 67–80, Jan. 2004.
[2] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, “Serial concatenation of interleaved codes: performance analysis, design, and iterative decoding,” IEEE Trans. Inform. Theory, vol. 44, pp. 909–926, May 1998.
[3] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” IEEE Trans. Inform. Theory, vol. IT-20, pp. 284–287, Mar. 1974.
[4] M. Tuchler, A. C. Singer, and R. Koetter, “Minimum mean squared error equalization using a priori information,” IEEE Trans. Signal Proc., vol. 50, no. 3, pp. 673–683, Mar. 2002.
[5] G. Colavolpe, G. Ferrari, and R. Raheli, “Reduced-state BCJR-type algorithms,” IEEE J. Select. Areas Commun., vol. 19, pp. 848–859, May 2001.
[6] J. B. Anderson and S. Mohan, “Sequential coding algorithms: a survey and cost analysis,” IEEE Trans. on Communications, vol. COM-32, no. 2, pp. 169–176, Feb. 1984.
[7] J. Hagenauer, “The renaissance of an ancient decoding method: Sequential decoding for turbo schemes,” presented at the IEEE Communication Theory Workshop, Capri, Italy, May 2004.
[8] V. Franz and J. B. Anderson, “Concatenated decoding with reduced-search BCJR algorithm,” IEEE J. Select. Areas Commun., vol. 16, no. 2, pp. 186–195, Feb. 1998.
[9] J. Hagenauer and K. Kuhn, “Turbo equalization for channels with high memory using a list-sequential equalizer,” in 3rd International Symposium on Turbo Codes, ENST de Bretagne, Sept. 2003.
[10] K. R. Narayanan, R. V. Tamma, E. Kurtas, and X. Yang, “Performance limits of turbo equalization using M-BCJR algorithm,” in 3rd International Symposium on Turbo Codes, ENST de Bretagne, Sept. 2003.
[11] S. Crozier and P. Guinand, “High-performance low-memory interleaver banks for turbo-codes,” in IEEE Vehicular Tech. Conf. (VTC’01 Fall), Oct. 2001.