



LUND UNIVERSITY

New reduced state space BCJR algorithms for the ISI channel

Anderson, John B; Prlja, Adnan; Rusek, Fredrik

Published in:
Proceedings, International Symp. on Information Theory

2009

[Link to publication](#)

Citation for published version (APA):

Anderson, J. B., Prlja, A., & Rusek, F. (2009). New reduced state space BCJR algorithms for the ISI channel. In *Proceedings, International Symp. on Information Theory* (pp. 889-893). IEEE - Institute of Electrical and Electronics Engineers Inc..

Total number of authors:

3

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

New Reduced State Space BCJR Algorithms for the ISI Channel

John B. Anderson, Adnan Prlja and Fredrik Rusek

Electrical and Information Tech. Dept. and Strategic Center for High Speed Wireless Communication

Lund University

Box 118, SE-221 00 Lund SWEDEN

Email: anderson@eit.lth.se, adnan.prlja@eit.lth.se, fredrikr@eit.lth.se

Abstract—A critical component in detection under intersymbol interference (ISI) and in turbo equalization is the BCJR algorithm. We combine three approaches to reducing its computation. First, energy seen by the receiver is focused by a phase-maximizing all pass filter; an improvement on this older idea is proposed. Then the state used by the BCJR is broken into an offset state and a main state. Finally, some reduced-state BCJR procedures are evaluated. These receivers are tested by ISI detection and turbo equalization over strongly bandlimited channels.

I. INTRODUCTION

We investigate the design and complexity of the BCJR calculation when a coded transmission is strongly band limited. The transmission is linear modulation with a baseband pulse $h(t)$ according to

$$s(t) = \sqrt{E_b/T} \sum_n a_n h(t - n\tau T), \quad \tau \leq 1 \quad (1)$$

where $\{a_n\}$ are binary data, E_b is the symbol energy, $h(t)$ has unit energy and τT is the symbol time ($\tau < 1$). The pulse $h(t)$ is much narrower band than $1/2\tau T$, and consequently there is strong intersymbol interference (ISI). An additive white Gaussian noise (AWGN) channel follows. The receiver consists of a matched filter of some sort, a sampler and a post filter, which together reduce the channel model to a discrete-time convolution with $\mathbf{v} = v_0, v_1, \dots$, to whose outputs are added zero-mean IID Gaussians with variance $N_0/2$. We will employ this Linear Modulation—AWGN Channel—BCJR Algorithm system in two ways, as the inner coder/decoder in turbo equalization [9] and as a narrowband communication system in its own right.

Some receiver possibilities for this scenario are sketched in Section II. It has been known for some time (see e.g. [2]) that the samples should be all-pass filtered to produce a maximum phase output; reversing the output frame then produces a minimum phase block that feeds a Viterbi Algorithm (VA) or reduced-search decoder such as the M-algorithm or the M*-algorithm [11]. For finite ISI with total length m_T and taps v_0, \dots, v_{m_T} , minimum phase will not help a full 2^{m_T} state VA, but it strongly aids schemes like the M-algorithm that work better with ISI whose energy is “up front”. For the *reduced state* VA or BCJR, which has 2^m states, $m < m_T$, recent work [3], [4] shows that some sort of minimum phase

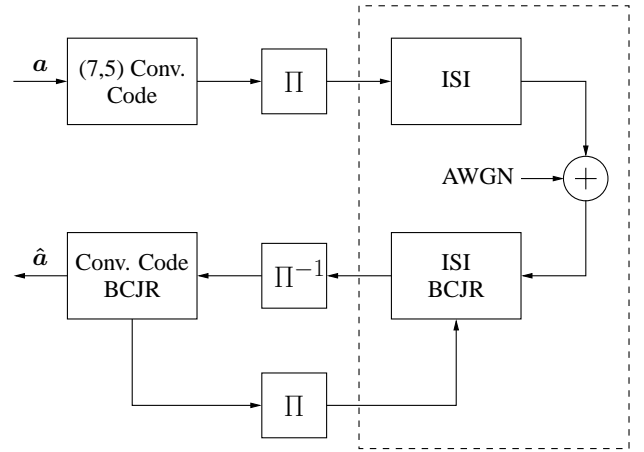


Fig. 1. Turbo equalization with a simple detection inner coder (dashed box).

presented to the VA is key to improving the error performance: The reduced state space built upon v_0, \dots, v_m should include the high-energy part of the ISI, and this is promoted by a minimum phase input. Our paper confirms this. But we find that the minimum phase notion can be extended, in ways that lead to 2–4 fold further reduction in the VA. The novel work in the paper centers on the marrying this extension to some new reduced BCJR algorithms, and testing them in very narrow band ISI.

Earlier reduced state work primarily treats the VA. The BCJR is a rather different algorithm that consists of two linear recursions, instead of the add-compare-select of the VA. It is reasonable that energy focusing by minimum phase is also important to BCJR state reduction, and we confirm this, but there is little reason to suspect that the inner workings of the BCJR can be similar to the VA's. We evaluate a number of modifications to the BCJR recursions and report in detail on one that is a good compromise between simplicity and accurate log-likelihood ratio (LLR) outputs. It is evaluated as a detector in its own right, (“simple detection” in the sequel) and as an element in turbo equalization. These are shown schematically in Fig. 1. An earlier study of reduced-state BCJR algorithms for non-minimum phase channels is [10]. A factor graph based approach for reduced complexity receivers has also been presented in [5].

The pulses $h(t)$ in this paper are chosen within the so-called faster than Nyquist, or FTN, framework. In this way of thinking, $h(t)$ with $\tau = 1$ is a familiar T -orthogonal pulse. In this paper h is a root raised-cosine (RC) pulse with 30% excess bandwidth. Its error probability in AWGN is $Q(\sqrt{2E_b/N_0})$. As τ drops below 1, pulses come “faster”, but the transmitted average power spectral density (PSD) shape remains the same, namely, a raised cosine. The bit density is $2/\tau$ bits/Hz-s, taking 3 dB bandwidth. The asymptotic error rate remains $Q(\sqrt{2E_b/N_0})$ for $\tau \geq .703$, the “Mazo limit” [1]. Thereafter, it declines with τ according to $Q(\sqrt{d_{\min}^2 E_b/N_0})$, $d_{\min}^2 < 2$, in which d_{\min} is the signals’ Euclidean minimum distance. The methods in this paper apply to any ISI, but there are good reasons for the FTN framework. First, the transmissions generated as τ declines have proven to be an effective way to design coded systems that minimize both energy and bandwidth. Second, they yield increasingly high bit density systems that have identical PSD shapes; this is important because above 3–4 bits/Hz-s both practical schemes and the Shannon capacity are sensitive to the PSD rolloff and hard to compare unless the PSD shape is the same. Finally, the Shannon capacity of FTN signals exceeds that of orthogonal signals with the same PSD [6].

The paper is organized as follows. Section II sets up a white noise receiver structure that facilitates a reduced state space. One element is an all-pass filter, and Section III proposes a generalization on earlier designs that further reduces the state space. Section IV proposes some BCJR alternatives. Section V evaluates the best of these in simple detection and turbo equalization.

II. A SUITABLE DISCRETE-TIME RECEIVER

We require a receiver that creates discrete-time outputs that are suitable for conversion to maximum phase form and subsequent VA or BCJR processing. In addition, we will require that the noise at the VA/BCJR input be white. There are colored noise BCJR algorithms and detectors that appear to be successful with certain ISI patterns, but whether colored or white noise is best is a complex question that depends on the ISI, and in this paper we focus on the white noise case. The transmit/receive chain thus comprises the following: Data each $n\tau T \rightarrow$ Transmit Filter $h(t) \rightarrow$ AWGN \rightarrow Matched Filter \rightarrow Sample at $n\tau T \rightarrow$ Post Filter $B(z) \rightarrow$ Reverse Frame \rightarrow VA/BCJR \rightarrow Data Out.

There are two main candidates for this chain.

(i) The *whitened matched filter* (WMF) receiver consists of a receive filter matched to $h(t)$, sampled at $n\tau T$, and a Post Filter designed to whiten the noise. Because $h(t)$ is not τT -orthogonal, the samples have colored noise. According to a standard argument, a stable post filter whitens the noise but creates a maximum phase channel model as seen by the VA/BCJR. With reduced-state VA/BCJRs, this output must be frame-reversed to produce a minimum phase model. The WMF design is useful for τ near 1. Details appear in [4].

(ii) The *orthogonal basis model* (OBM) receiver

expresses $h(t)$ as a sequence of wider band orthonormal pulses. That is,

$$h(t) = \sum_{k=-K}^K c_k \phi(t - k\tau T) \quad (2)$$

where $c_k = \int h(t) \phi(t - k\tau T) dt$. The basis function $\phi(t)$ is τT -orthogonal and the sampling rate $1/\tau T$ must satisfy the sampling theorem for $h(t)$. The Matched Filter is matched to ϕ , not h , and consequently the noise in its outputs is white. The Post Filter is an all-pass that converts the sample stream to maximum phase, while leaving the noise white. As before, Frame Reverse converts the sequence to minimum phase, and the VA/BCJR sees a minimum phase white noise ISI model.

Essentially, (2) expresses $h(t)$ by $2K + 1$ of its samples, with the remaining ones truncated (those $< 0.01/T$ can be safely ignored). Further, the sampling need not be at the symbol rate $1/\tau T$, but can be faster or slower, so long as the sampling theorem is satisfied (we found that neither offers an advantage). Because of its simplicity and versatility, we will use the OBM receiver. The τ will be $1/2$ and $1/3$, a doubling and tripling of bandwidth efficiency. For $\phi(t)$ we simply take the 30% root RC pulse whose bandwidth is scaled by τ , $\sqrt{1/\tau} h(t/\tau)$. This satisfies the sampling criterion for $\tau \leq 1/1.3 \approx .77$.

The $h(t)$ —AWGN—Matched Filter—Sampler—Post Filter chain creates the channel model v_0, \dots, v_{m_T} seen by the VA/BCJR. Without the Post Filter, the model for $\tau = 1/3, 1/2$ is simply the samples h_{-K}, \dots, h_K , with a delay of K inserted for causality. The autocorrelation of $\{v_k\}$ are samples at $n\tau T$ of the autocorrelation of the original $h(t)$. The addition of an all-pass filter does not change this; it only produces a v with a new phase. In every case the (untruncated) v and the true continuous-time signal set have the same d_{\min} , since this depends only on the autocorrelation. Consequently, the different receivers should have about the same error event rate $\approx Q(\sqrt{d_{\min}^2 E_b/N_0})$.

The key to reducing the state space of VA/BCJR that follows is to base that space on the most energetic parts of the ISI model. This is the verdict of several decades of research. The receivers above present the VA/BCJR with an ISI that is minimum phase, i.e., its energy is up front as much as possible for the model autocorrelation. Minimum phase ISI models that stem from strong band limitation have taps in the pattern [low energy precursor] + [high energy response] + [long decaying tail]. These facts motivate the following *offset receiver*. Consider branch labels s_n at trellis stage n each generated from some ± 1 data sequence \mathbf{a} by

$$s_n = \sum_{k=0}^m a_{n-k} v_k + \sum_{k=m+1}^{m_T} a_{n-k} v_k \quad (3)$$

Symbols a_{n-1}, \dots, a_{n-m} comprise the size- m reduced VA/BCJR state. The first term stems from the high energy state symbols, while the second is an offset created by earlier symbol history. An offset is associated with each state in the reduced VA/BCJR but is itself not part of the algorithm’s state.

The first symbols are the *main state*, while the second form the auxilliary *offset state*. In the add-compare-select step of the VA, the offset states of the survivors become the new offsets for each main state.

This sort of trellis search dates back at least to the 1970s [7] and was popularized by several authors in the 1980s, the best known paper being Duel-Hallen and Heegard [8], who calculated the minimum distance that applies to the receiver. Studies then and now [4] show that under narrowband ISI a large truncation of the VA is possible without significant loss of distance and that the VA needs only encompass the high energy taps. The long tail taps cannot be ignored because they significantly offset the trellis branch labels, but only the offset needs to be remembered, one per main state.

The offset receiver needs to ignore any low energy precursor. It incorporates the decaying tail without much complexity, but there seems to be no reasonable way to incorporate the precursor, other than setting its taps to zero. Unfortunately, important minimum phase models can have significant precursor length. We have found that for a unit-normalized channel model, precursor taps below 0.02 and tail taps below 0.03 in absolute value can be truncated without significant effect.

In summary, the following are necessary, in priority order, to reduce the states of the VA or BCJR:

- First, small precursor taps should be set to zero.
- The VA/BCJR should see a minimum phase channel.
- An offset VA/BCJR should be used, focused on the energy concentration.

III. NEW MINIMUM PHASE MODELS

There are actually many OBM receiver possibilities because any all-pass Post Filter block satisfies the white noise and no d_{\min} loss requirements. Is there a better one than the phase maximizing all-pass? The answer is yes, and the reason is a subtlety in the idea of maximum phase: The precise maximum phase solution depends on the support of the input. It depends on the pulse time frame, here $[-K\tau T, K\tau T]$, the sampling rate, and whether the domain is continuous or discrete. All rates that meet the sampling theorem on $h(t)$ are allowed in principle, which for the $\tau = 1/3$ case allows samples at less than half the symbol rate.

A fruitful design procedure for $B(z)$ is as follows: Design $B(z)$ as the all-pass that creates the maximum phase version of $2K + 1$ centered samples of $h(t)$, sampled at $1/\tau T$; all such all-passes are legal and produce a channel model with the full d_{\min} . Then truncate outer small taps such that d_{\min} is not significantly reduced. Search over K for the (reversed) solution that best concentrates the minimum distance into the later taps.

For the root RC base pulse in this paper, one finds that reducing either K or the sampling rate leads to more compaction of the energy in the maximum phase outcome $\{h_{-K}, \dots, h_K\} * \mathbf{b}$. Very low postcursor taps (precursor in the reversed model) develop, but these can be truncated directly. Some truncation of the tail taps leads to models (4) at $\tau = 1/3$

and (5) at $1/2$; (4) derives from $2K + 1 = 13$ samples of $h(k\tau T)$, while (5) derives from 9. The continuous-time reference d_{\min}^2 is 0.58 for the first and 1.02 for the second, and both discrete-time models achieve 98% of these. Four precursor taps are deleted in the first and 8 in the second.

$$\mathbf{v} = \{.191, .464, .623, .506, .176, -.123, -.197, -.075, .060, .080, .013, -.035, -.022\} \quad (4)$$

$$\mathbf{v} = \{.375, .741, .499, -.070, -.214, .019, .087, -.020, -.028\} \quad (5)$$

$$\mathbf{v} = \{.130, .484, .706, .368, -.178, -.228, .083, .125, -.057, -.056, .043\} \quad (6)$$

The last model (6) is for $\tau = 1/2$ and is provided for comparison to [4]. It is derived by a less effective minimum phase method and leads to a 2–4 times larger state space.

IV. SOME TRUNCATED BCJR ALGORITHMS

This section reviews the BCJR algorithm and presents a scenario for its state reduction. We then present a number of reduced algorithms. All are based on the main+offset state idea.

The BCJR algorithm computes the probabilities of states and paths in a signal trellis, given the channel outputs $\mathbf{y} = y_1, \dots, y_N$ and the apriori data probabilities. A compact statement of the algorithm is given in terms of two matrix recursions¹ that calculate trellis working variables α_n and β_n at stage n . These vectors have components

$$\begin{aligned} \alpha_n[j] &\triangleq P[\text{Observe } y_1, \dots, y_n \cap \mathfrak{S}_n = j] \\ \beta_n[i] &\triangleq P[\text{Observe } y_{n+1}, \dots, y_N \mid \mathfrak{S}_n = i] \end{aligned} \quad (7)$$

where \mathfrak{S}_n is the encoder state at time n . The following *forward* and *backward recursions* hold:

$$\begin{aligned} \alpha_n &= \alpha_{n-1} \mathbf{\Gamma}_n, \quad n = 1, \dots, N \\ \beta_n &= \mathbf{\Gamma}_{n+1} \beta_{n+1}, \quad n = N - 1, \dots, 1 \end{aligned} \quad (8)$$

Here $\mathbf{\Gamma}_n$ is the matrix with $[i, j]$ element

$$\begin{aligned} \Gamma_n[i, j] &\triangleq P[y_n \cap \mathfrak{S}_n = j \mid \mathfrak{S}_{n-1} = i] \\ &= [P(a') / \sqrt{\pi N_0 / E_s}] \exp[-(N_0 / E_s)(y_n - \ell_{i,j})^2] \end{aligned} \quad (9)$$

where $\ell_{i,j}$ is the label (3) on the branch from state i to j and a' is the value of data symbol a_n that causes the transition. Our data frames terminate at both ends at the all-0 state, so $\alpha_0 = (1, 0, \dots, 0)$ and $\beta_N = (1, 0, \dots, 0)'$. From the $\{\alpha\}$ and $\{\beta\}$ come many quantities, but we are interested only in the LLR of a_n , which is

$$\text{LLR}(a_n) \triangleq \ln \frac{P[a_n = +1]}{P[a_n = -1]} = \ln \frac{\sum_{j \in \mathcal{L}_{+1}} \alpha_n[j] \beta_n[j]}{\sum_{j \in \mathcal{L}_{-1}} \alpha_n[j] \beta_n[j]} \quad (10)$$

where \mathcal{L}_{+1} is the set of states reached by $a_n = +1$.

¹The matrix is sparse. The recursions execute $2 \cdot 2^m$ multiplications per stage, not 2^{2m} . We use this in our implementations. The breakdown into main+offset states makes precomputation and storage of all labels practical.

The heart of the BCJR is (9). Its elements contribute whenever an $\ell_{i,j}$ is close to y_n . Whereas the VA “picks winners”, continually dropping path segments that fall short, the BCJR counts every contributing region of the trellis. Hopefully, “hot regions” can be specified—and an accurate LLR computed—by a reduced main state. With narrow band ISI, the labels $\ell_{i,j}$ in both the VA and BCJR depend strongly on both the main and offset states. Only a reasonable approximation of the latter is needed, but it must be present.

The application scenario for ISI is as follows. The α recursion is taken to be the one proceeding left to right in the direction of time, the direction in which the sequence phase is minimum. At the extension to trellis stage $n+1$, the alignment of ISI model taps, state symbols and alphas is

$$\begin{array}{ccccccc} v_{m_T}, \dots, v_1, & v_0 \\ \dots, a'_{n-m_T}, a'_{n-m_T+1}, \dots, a'_n, & a'_{n+1} \\ & \dots, \alpha_n, & \alpha_{n+1} & \longrightarrow \end{array}$$

with the main state and main v energy to the right. The main state is defined by the symbols at stages $n-m+1, \dots, n$. The extension to stage $n+1$ computes α_{n+1} ; all such decided α are stored. Once the state symbols a' become fixed they become the *tentative data*; this can occur at position $n-m$ when a symbol enters the offset state, or it can occur at an earlier n . In any case, the decision is made from α information and $a'_{n-m_T+1}, \dots, a'_{n+1}$ determine the label on their respective branch at stage $n+1$.

In the reverse recursion, labels are formed from m main state symbols and $m_T - m$ offset symbols that lie in the α -decided tentative data. The reverse recursion has not yet reached the area of the offset symbols. The situation is aligned

$$\begin{array}{ccccccc} v_{m_T}, \dots, v_1, & v_0 \\ a'_{n-m_T+1}, \dots, a'_n, & a'_{n+1} \\ \longleftarrow \beta_n, & \beta_{n+1} & \beta_{n+2}, & \dots \end{array}$$

The β vectors are not stored. β_n is immediately used with α_n to find the LLR and from it an update to the n th tentative path symbol. Note that the updated path is much better than the α -only path; it typically has one tenth the bit error rate.

Here are some of the approaches to a reduced state BCJR that we have investigated within this scenario.

(i) *True offset algorithm*: A different offset state is associated with each main state. Some method is needed to choose which offset survives after the move forward to new main states at stage $n+1$. In analogy to the VA, it could be the offset+main path that makes the largest contribution to the new $\alpha_{n+1}[j]$ (there are two at each new state j), but this strategy is poorer than others because the contributions do not correlate tightly enough with hot trellis regions. A simple but effective alternative has proven difficult to find.

(ii) *Single offset algorithm*: Associate the same offset with all main states. Take the new offset bit as the oldest main state bit that leads to the larger set of α_{n+1} contributions (2^m of these stem from each bit value). In performance and complexity, this is in fact the best strategy we have found.

This oldest main state bit is a very good indicator of a hot region.

(iii) *VA aided BCJR*: While the BCJR calculates state probabilities and not bit decisions, the VA decision path is nonetheless a good pointer to hot regions, even if the VA is strongly reduced. An accurate reduced BCJR of this type needs about one unit less m than a BCJR of type (ii), but it adds the VA complexity, and is therefore no overall improvement.

(iv) *Worst case offset each state*: Since the BCJR looks for hot spots wherever they occur, they can be located by finding the α -maximizing offset for each new state at $n+1$. This strategy is also not as effective as (ii).

(v) *Iterating*: Multiple passes of any of the preceding can be executed. This improves the α, β estimate because each new forward recursion can use for its tentative path the LLR-determined path from the previous reverse recursion. The tentative paths thus constantly improve. We find that the α, β accuracy so obtained is not worth the complexity of the additional iterations.

V. SIMPLE DETECTION AND TURBO EQUALIZATION WITH A REDUCED BCJR

We have constructed the ISI—AWGN Channel—BCJR blocks at the right side of Fig. 1 in discrete time and tested them with simulated AWGN. The signals are (1) with τ either 1/2 or 1/3, a bandwidth efficiency 2 and 3 times that of orthogonal-pulse root RC. The ISI model² is that of (5) and (4), respectively, and the reduced BCJR is of type (ii) with 2^m main states. Bit decisions follow from the LLR (10).

Figure 2 plots observed BCJR and offset VA [4] error event rates (EERs) for both τ and a number of state memories m (note that the BCJR optimizes BER, not EER). A distance study shows that the d_{\min} -causing difference sequence is 2, -2, 2 for $\tau = 1/2$; the multiplicity for this is 1/4 (see [4]), so that the full-state VA or BCJR EER is $\approx .25Q(\sqrt{1.02E_b/N_0})$. The $\tau = 1/3$ case is more complex and is $\approx .35Q(\sqrt{.58E_b/N_0})$. These appear as heavy curves. The bit error rate (BER) is about 3 times the EER at higher SNR and 4–5 times at lower SNR. The BCJR needs only 8–16 states at $\tau = 1/2$ and about 64 at $\tau = 1/3$.

Next we investigate the same reduced BCJR as part of the turbo equalization system in Fig. 1. The transmitter consists of the (7,5) rate 1/2 convolutional encoder, an interleaver, and the same two τ in (1). A block of $N = 1000$ information bits is encoded to produce $2N$ bits, which feed a size $2N$ random interleaver; these map to ± 1 via $\{0 \rightarrow +1, 1 \rightarrow -1\}$, terminating +1s are added, and the sequence is passed through the discrete time ISI model.

²The discrete time transmitter uses a longer ISI model with precursors and extra tail taps so that the root RC autocorrelation is accurate to $\approx \pm .0002$. The receiver is thus slightly mismatched, but the effect on BER is small. The test setup is as in our earlier paper [4]: Size 800 frames of random ± 1 data, with enough frames to give 40–100 error events; frames are terminated before and after by m_T ‘+1’ symbols; error events are taken to begin when the receiver output state splits from the transmitter state path and to end after 5 output data are correct.

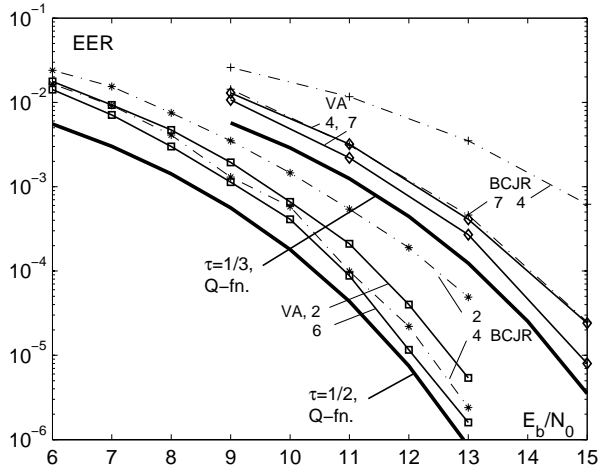


Fig. 2. Error event rates for simple ISI detection vs. E_b/N_0 in dB; VA at memory m (solid), BCJR at m (dash dot). Heavy line is Q-function estimate.

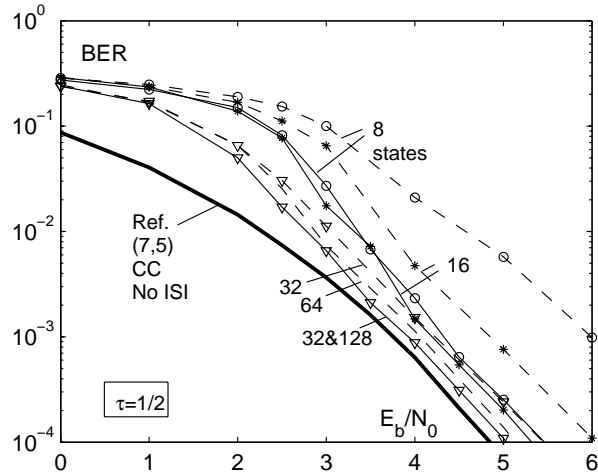


Fig. 3. Turbo equalization for $\tau = 1/2$, comparing simple truncation (dashed) to BCJR offset algorithm (ii) with several m (solid).

Figures 3–4 plot BER outcomes for the offset BCJR and a BCJR that simply truncates its state to the strongest m positions. Enough frames are decoded to produce at least 70 error events. The ISI model is (4) for $\tau = 1/3$ and (6) for $\tau = 1/2$; the last is the one in [4] and comparison can be made to that paper. Were this model (5), performance would be better because the BCJR would see a model with more concentrated energy. In these figures the offset BCJR clearly improves simple truncation. The size needed is 16–32 states at $\tau = 1/2$ and 64–128 at $1/3$.

VI. CONCLUSION

We have investigated a number of reduced state BCJR algorithms, and tested them as simple detectors and as part of a turbo equalizer over two narrow band ISI channels. These BCJR are of an offset type that depends on the fact that the ISI model is minimum phase. We have proposed a generalization of that concept that reduces state size by a further 2–4 times. The outcome is a decoder of reasonable complexity, which

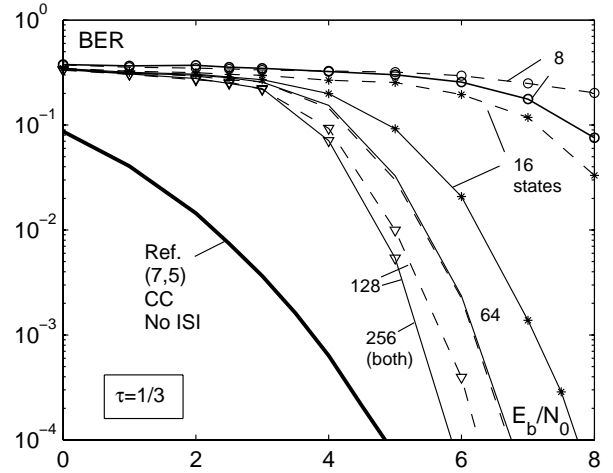


Fig. 4. Turbo equalization as in Fig. 3, but $\tau = 1/3$.

in a turbo configuration can lead simultaneously to an energy saving of 4 dB and a bandwidth reduction of 35%.

ACKNOWLEDGMENTS

This work was supported by the Swedish Research Council (VR) through Grant 621-2003-3210, and by the Swedish Foundation for Strategic Research (SSF) through its Strategic Center for High Speed Wireless Communication at Lund.

REFERENCES

- [1] J.E. Mazo, "Faster-than-Nyquist signaling," *Bell Syst. Tech. J.*, vol. 54, pp. 1451–1462, Oct. 1975.
- [2] K. Balachandran, J.B. Anderson, "Reduced complexity sequence detection for nonminimum phase intersymbol interference channels," *IEEE Trans. Information Theory*, vol. 43, pp. 275–280, Jan. 1997.
- [3] D. Fertonani, A. Barbieri, G. Colavolpe, "Reduced-complexity BCJR algorithm for turbo equalization," *IEEE Trans. Commun.*, vol. 55, pp. 2279–2287, Dec. 2007.
- [4] A. Prlja, J.B. Anderson, F. Rusek, "Receivers for faster-than-Nyquist signaling with and without turbo equalization," *Proc.*, 2008 IEEE Int. Symp. Information Theory, Toronto, July 2008.
- [5] D. Fertonani, A. Barbieri, G. Colavolpe, "Novel graph-based algorithms for soft-output detection over dispersive channels," *Proc.*, IEEE 2008 Global Commun. Conf., Dec. 2008.
- [6] F. Rusek, J.B. Anderson, "Constrained capacities for faster-than-Nyquist signaling," *IEEE Trans. Information Theory*, vol. 55, pp. 764–775, Feb. 2009.
- [7] J.B. Anderson, "Tree encoding of speech," *IEEE Trans. Information Theory*, vol. 21, pp. 379–387, July 1975.
- [8] A. Duel-Hallen, C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, pp. 428–436, May 1989.
- [9] C. Douillard *et al.*, "Iterative correction of intersymbol interference: Turbo equalization," *Eur. Trans. Telecomm.*, vol. 6, pp. 507–511, Sept./Oct. 1995.
- [10] C.M. Vithanage, C. Andrieu, R.J. Piechocki, "Novel reduced-state BCJR algorithms," *IEEE Trans. Commun.*, vol. 55, pp. 1144–1152, June 2007.
- [11] M. Sikora, D.J. Costello, "A new SISO algorithm with application to turbo equalization," *Proc.*, IEEE Int. Symp. Information Theory, Adelaide, pp. 2031–2035, Sept. 2005.