

# Noise-Predictive BCJR Equalization for Suppression of Intrachannel Nonlinearities

Ivan B. Djordjevic, *Member, IEEE*, and Bane Vasic, *Senior Member, IEEE*

**Abstract**—An iterative equalization scheme for suppressing intrachannel nonlinearities in high-speed optical transmission systems operating at 40 Gb/s is presented. The proposed method employs a noise-predictive filter to mitigate the colored noise due to amplified spontaneous emission noise and optical filtering. The maximum *a posteriori* probability symbol detection implemented as the Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm is employed to suppress the intrachannel nonlinear effects. Soft decision outputs provided by the noise-predictive BCJR equalizer are processed by an iterative low-density parity-check decoder.

**Index Terms**—Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm, low-density parity-check (LDPC) codes, noise-predictive filter, optical communications.

## I. INTRODUCTION

IT IS widely believed that high-speed optical transmission systems operating at 40 Gb/s or higher are severely limited by intrachannel nonlinearities such as intrachannel four-wave mixing (IFWM) and intrachannel cross-phase modulation [1], [2]. Methods to deal with intrachannel nonlinearities include 1) modulation formats, 2) constrained (modulation or line) coding, and 3) equalization techniques [3]. The modulation formats (e.g., [1]) try to remove the phase short-term coherence of the pulses emitted in a given neighborhood. The constrained encoder [2] translates the information sequence into a sequence not containing the waveforms that are most likely to be received incorrectly. Other techniques proposed recently, in context of chromatic dispersion and polarization mode dispersion, include maximum likelihood sequence detection (MLSD) based on Viterbi algorithm [4]. However, the Viterbi decoder provides hard decisions only preventing the soft iterative decoding.

In this letter, we propose a nonlinear equalizer, referred to as noise-predictive Bahl–Cocke–Jelinek–Raviv (NP-BCJR) equalizer, composed of two components. The first component is a noise-predictive filter that deals with colored noise due to the amplified spontaneous emission (ASE) noise, colored by optical filter in every erbium-doped fiber amplifier (EDFA) stage (concept initially proposed in magnetic recording [7], [8]). The second component is the BCJR equalizer whose role is to suppress intrachannel nonlinearities (the concept introduced by authors in [5]). The iterative low-density parity-check (LDPC) decoder processes further the soft outputs (bit reliabilities) provided by the BCJR equalizer. The noise-predictive filter is implemented as a finite impulse response (FIR) filter. The noise-predictive filter

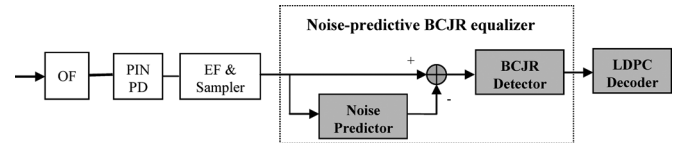


Fig. 1. NP-BCJR equalization. OF: optical filter. PD: photodetector. EF: electrical filter.

estimates the colored noise, which is removed from the sampled signal after the photodetection and electrical filtering and passed to the second nonlinear equalizer ingredient—the BCJR equalizer. The BCJR equalizer removes the intrachannel nonlinearities and requires reduced number of states in trellis compared to the case where no noise-predictive filter is being employed. The main feature of the proposed scheme is that it can operate in the regime of very strong intrachannel nonlinearities.

## II. NP-BCJR EQUALIZATION

The block scheme of the proposed NP-BCJR equalization scheme is shown in Fig. 1. It is based on the BCJR equalization scheme proposed by authors in [5]. The novelty with respect to the scheme from [5] is the introduction of the noise-predictive FIR filter. Its role is to deal with colored noise due to ASE noise and partially to deal with linear intersymbol interference (ISI) introduced by residual dispersion and (optical/electrical) filtering. The key idea is to reduce the trellis complexity of BCJR equalization scheme from [5] by estimating the colored noise by the linear noise-predictive FIR filter. In [5], the BCJR algorithm [6] is used to deal with both the linear and nonlinear impairments. Here we use the BCJR algorithm to deal with intrachannel nonlinearities only. If  $p(D) = p_1D + p_2D^2 + \dots + p_ND^N$  represents the transfer polynomial of the FIR noise prediction filter of length  $N$ , then  $1 - p(D)$  represents the transfer polynomial of the prediction error filter, and the noise prediction error, defined as the difference of the exact noise sample  $w_n$  and the noise predicted sample  $\tilde{w}_n$  can be written as

$$e_n = w_n - \sum_{i=1}^N p_i w_{n-i}, \quad \tilde{w}_n = \sum_{i=1}^N p_i w_{n-i}. \quad (1)$$

Given the autocorrelation function  $R_w$ , the coefficients of noise predictor filter are determined by solving the system of well-known normal (Yule–Walker) equations [7]

$$R_w(i) = \sum_{j=1}^N p_j R_w(i-j), \quad i = 1, 2, \dots, N. \quad (2)$$

To solve the system of equations (2), an efficient method due to Burg [9] is used. It involves a recursive procedure of increasing the order of the filter one unit at a time, and re-estimating the

Manuscript received February 28, 2006; revised March 23, 2006. This work was supported by the National Science Foundation (NSF) under Grant ITR 0325979.

The authors are with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721 USA (e-mail: ivan@ece.arizona.edu).  
Digital Object Identifier 10.1109/LPT.2006.876757

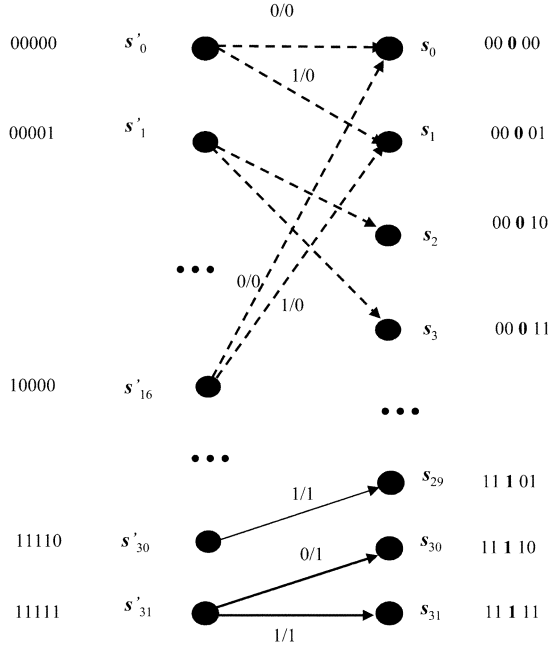


Fig. 2. Trellis used by BCJR (for  $2m + 1 = 5$ ).

filter coefficients in such a way to minimize the mean-square error defined as [7]

$$E[e_n^2] = R_w(0) - \sum_{i=1}^N p_i R_w(i). \quad (3)$$

Notice that the system of equations (2) has to be solved only once, before the start of a transmission, for example by using a short training sequence and collecting noise samples in the receiver. The length of FIR noise predictive filter needs to be chosen carefully in order to avoid error propagation. We have found that the length of 16 is sufficient for nonlinear fiber-optics channels. To reduce the number of coefficients employed in the noise predictive filter, an infinite impulse response (IIR) filter may be used instead [7].

The second component of the nonlinear equalizer proposed here is the BCJR equalizer. The trellis, a dynamic model of the channel shown in Fig. 2, is uniquely determined by the set of triples: previous state, channel output, next state. The state is determined by  $m$  previous and  $m$  next bits influencing the observed bit  $u_j$ ,  $\mathbf{s} = (u_{j-m}, u_{j-m+1}, \dots, u_j, u_{j+1}, \dots, u_{j+m})$ ,  $u_i \in \{0, 1\}$ . As an illustration, a trellis for memory  $2m+1 = 5$ , shown in Fig. 2, has 32 states ( $\mathbf{s}_0, \mathbf{s}_1, \dots, \mathbf{s}_{31}$ ), and each state is given by a different 5-bit pattern. From every state in the trellis there are two possible transitions. A labeled edge is assigned to each allowed transition, and a received sample corresponds to the output symbol of the branch (the central bit of the terminal state). An important difference with respect to magnetic-recording equalization schemes [7], [8] is that the BCJR algorithm operates on a trellis that includes both precursor and postcursor nonlinear ISI, similar to [4]. A significant benefit of using the BCJR algorithm, compared to Viterbi algorithm, is that in addition to detected bits it also provides bit reliabilities, i.e., soft decisions. For the complete description of the finite state machine describing the channel, the transition probability density functions (pdfs)  $p(y_j|\mathbf{s})$ ,  $\mathbf{s} \in \mathcal{S}$  are to be determined ( $y_j$  represent the sample at the input of the BCJR de-

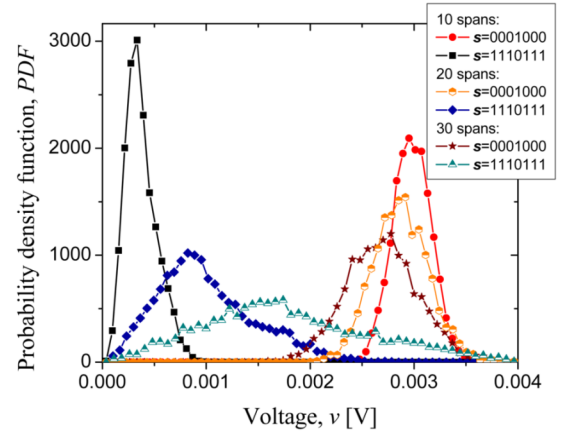


Fig. 3. PDFs for states  $\mathbf{s} = 1110111$  and  $\mathbf{s} = 0001000$  estimated from simulator. (Color version available online at <http://ieeexplore.ieee.org>.)

tor that corresponds to the transmitted bit  $u_j$ , and  $\mathcal{S}$  is the set of states in the trellis). We estimate the transition pdfs from the simulator [2] that takes the most important channel impairments into account. As an illustration, Fig. 3 shows the conditional pdf of the received sample  $y$  given a state  $\mathbf{s}$  for the following two states  $\mathbf{s} = 1110111$  and  $\mathbf{s} = 0001000$  and different number of spans, for dispersion map from [2]. The span length is set to  $L = 120$  km, and each span consists of  $2L/3$  km of  $D_+$  fiber followed by  $L/3$  km of  $D_-$  fiber. Precompensation of  $-800$  ps/nm and corresponding postcompensation of  $800$  ps/nm are also applied. In order to demonstrate the efficiency of the proposed method, a dispersion map is chosen in such a way that strong IFWM is present. Return-to-zero ON-OFF keying of a duty cycle of 33% is observed, the launched power is set to 0 dBm, and the extinction ratio to 14 dB. EDFAs with noise figure of 5 dB are deployed after every fiber section, the bandwidth of optical filter is set to  $3R_b$  and the bandwidth of electrical filter to  $0.65R_b$ , with  $R_b$  being the bit rate (40 Gb/s). As expected, by increasing the number of spans, the ghost pulse at the central bit position for state  $\mathbf{s} = 1110111$  grows, causing the mean of the pdf to shift to the right.

Given  $\mathbf{s}'$ , the previous state,  $\mathbf{s} = (u_{j-m}, u_{j-m+1}, \dots, u_j, u_{j+1}, \dots, u_{j+m})$ , the present state,  $\mathbf{u} = (u_1, u_2, \dots, u_n)$ , the transmitted codeword, and  $\mathbf{y} = (y_1, y_2, \dots, y_n)$ , the BCJR detector input sequence, the log-likelihood ratio (LLR) (denoting the bit reliability) of  $u_j$  ( $j = 1, 2, \dots, n$ ), is calculated as

$$L(u_j) = \max_{(\mathbf{s}', \mathbf{s}): u_j=0}^* [\alpha_{j-1}(\mathbf{s}') + \gamma_j(\mathbf{s}', \mathbf{s}) + \beta_j(\mathbf{s})] - \max_{(\mathbf{s}', \mathbf{s}): u_j=1}^* [\alpha_{j-1}(\mathbf{s}') + \gamma_j(\mathbf{s}', \mathbf{s}) + \beta_j(\mathbf{s})] \quad (4)$$

where the *forward metric*  $\alpha_j(\mathbf{s}) = \log p(\mathbf{s}_j = \mathbf{s}, \mathbf{y}_1^j)$  and the *backward metric*  $\beta_j(\mathbf{s}) = \log p(\mathbf{y}_{j+1}^n | \mathbf{s}_j = \mathbf{s})$  are determined as described in [5]. The dashed lines in Fig. 2 correspond to transitions  $(\mathbf{s}', \mathbf{s}): u_j = 0$ , and the solid lines to transitions  $(\mathbf{s}', \mathbf{s}): u_j = 1$ . The *branch metric*  $\gamma_j(\mathbf{s}', \mathbf{s})$  is given by  $\gamma_j(\mathbf{s}', \mathbf{s}) = \log p(y_j | u_j) p(u_j)$ .  $\mathbf{y}_l^m = (y_l, \dots, y_m)$  and  $p(y_j | u_j)$  is determined from simulator. The  $\max^*$ -operator is defined in [2]. Two main differences between the magnetic-recording turbo-equalization schemes (e.g., [8]) and NP-BCJR equalizer implemented here are: 1) LLRs of the *output bits* correspond to the central bit of terminal states, and 2) BCJR algorithm operates on a

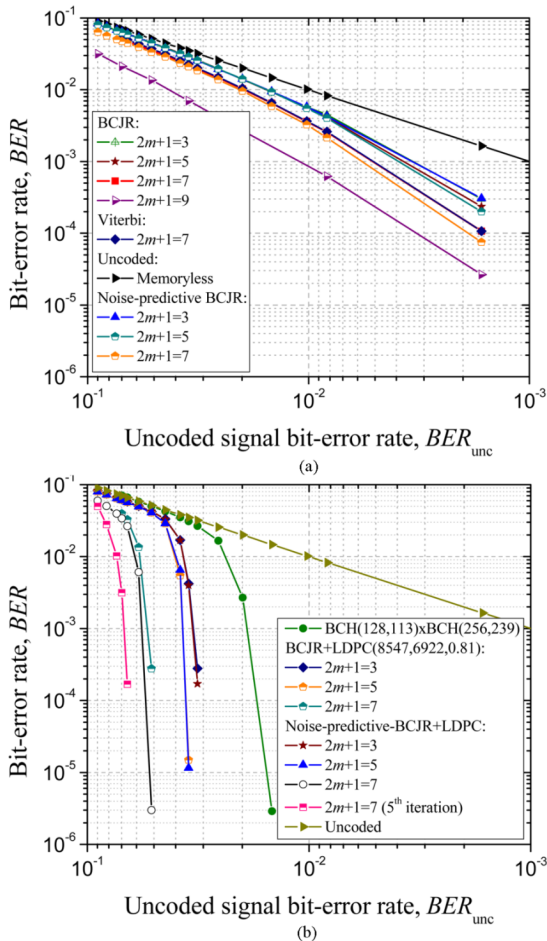


Fig. 4. (a) BER of NP-BCJR equalizer (b) supplemented with LDPC decoder. (Color version available online at <http://ieeexplore.ieee.org>.)

trellis that includes both precursor and postcursor ISI. The BCJR LLRs outputs  $L(u_j)$  ( $j = 1, 2, \dots, n$ ) are fed to an iterative LDPC decoder, implemented using an efficient implementation of message-passing (MP) algorithm. To facilitate the implementation a structured lattice based LDPC code of girth-eight is employed in simulations.

### III. PERFORMANCE ANALYSIS AND CONCLUSION

The bit-error-rate (BER) performance comparison of NP-BCJR equalizer scheme combined with the lattice LDPC (8547,6922,0.81) code of girth eight and column weight four against corresponding scheme without noise prediction is given in Fig. 4 for different memories in the channel trellis. Notice that BER performance improvement due to noise-predictive filtering is negligible for the channel memories up to  $2m+1=5$ . For the memory  $2m+1=7$ , the NP-BCJR equalizer BER approaches (in limit) that of BCJR equalizer with memory  $2m+1=9$ , facilitating therefore the implementation at high speed. Better BER improvement is expected when the EDFAs with higher noise figure are employed.

The coding gain of NP-BCJR equalizer (with memory seven) supplemented with LDPC decoder over an uncoded signal is about 9.3 dB (at BER of  $1 \cdot 10^{-6}$  for 25 iterations in MP decoder). For comparison purposes, the turbo product code (TPC) BER curve is inserted as well. TPC of code rate 0.82 based on

BCH(128 113)  $\times$  BCH(256 239), and about four times longer, is 2.5 dB away from proposed nonlinear equalizer, at BER of  $1 \cdot 10^{-6}$ . By iterating between BCJR equalizer block and the LDPC decoder, the BER performance can be further improved (on the expense of an increased decoding delay). At BER of  $10^{-6}$ , the coding gain improvement in fifth iteration (for ten iterations in MP decoder), over uncoded case is around 10 dB, and the improvement over TPC is around 3.2 dB. The expected coding gain at BER of  $10^{-12}$ , obtained by extrapolation, is around 13.4 dB. We have recently shown [11] that the finite geometry codes and lattice codes of high girth do not exhibit an error floor in the region of interest for fiber-optics communications, so that the extrapolation is justifiable once the waterfall region is reached. The results of simulations are obtained by maintaining a double precision of LLRs. We have recently shown in [10] that proper choice in number of quantization bits results in negligible BER performance loss.

In conclusion, we have shown that NP-BCJR equalizer supplemented with iterative decoding performs very well in systems heavily degraded by ISI due to intrachannel fiber nonlinearities, where advanced FEC schemes, such as turbo or LDPC codes, are not designed to operate. We note that other techniques for suppression of intrachannel nonlinearities, such the ones based on Volterra series method [3] or MLSD, do not provide soft outputs required for soft iterative decoding. To reduce the number of states required for BCJR equalization further, the dispersion map has to be designed such that during transmission over  $D_+$  fiber the pulse is spread up to 10 bits. By replacing the  $\max^*$ -operator in (4) by the  $\max$ -operator results in negligible performance loss, and forward and backward recursion of BCJR algorithm become forward and backward Viterbi algorithm, respectively. The MLSD chips are already commercially available at 10 Gb/s, and are in a development phase for 40-Gb/s transmission, suggesting that the nonlinear NP-BCJR equalization scheme proposed in this letter is timely.

### REFERENCES

- [1] R.-J. Essiambre, G. Raybon, and B. Mikkelsen, "Pseudo-linear transmission of high-speed TDM signals 40 and 160 Gb/s," in *Optical Fiber Telecommunications IVB*, I. P. Kaminov and T. Li, Eds. San Diego, CA: Academic, 2002, pp. 233–304.
- [2] I. B. Djordjevic and B. Vasic, "Constrained coding techniques for suppression of intrachannel nonlinear effects in high-speed optical transmission," *J. Lightw. Technol.*, vol. 24, no. 1, pp. 411–419, Jan. 2006.
- [3] C. Xia and W. Rosenkranz, "Performance enhancement for duobinary modulation through nonlinear electrical equalization," in *Proc. ECOC 2005*, vol. 2, Glasgow, U.K., Sep. 25–29, 2005, pp. 257–258.
- [4] N. Alic *et al.*, "Signal statistics and maximum likelihood sequence estimation in intensity modulated fiber optic links containing a single optical pre-amplifier," *Opt. Express*, vol. 13, pp. 4568–4579, Jun. 2005.
- [5] I. B. Djordjevic, B. Vasic, and V. S. Rao, "Suppression of intrachannel nonlinearities using BCJR algorithm and iterative decoding," in *Proc. OFC/NFOEC 2006*, Anaheim, CA, Mar. 5–10, 2006, Paper OTuK5.
- [6] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 2, pp. 284–287, Mar. 1974.
- [7] J. D. Coker *et al.*, "Noise-predictive maximum likelihood (NPML) detection," *IEEE Trans. Magn.*, pt. 1, vol. 34, no. 1, pp. 110–117, Jan. 1998.
- [8] S. Aviran *et al.*, "Noise-predictive turbo equalization for partial-response channels," *IEEE Trans. Magn.*, vol. 41, no. 10, pp. 2959–2961, Oct. 2005.
- [9] J. P. Burg, "A New Analysis Technique for Time Series Data," NATO Adv. Study Inst., Sig. Proc. With Emphasis on Underwater Acoust., Aug. 1968.
- [10] S. Sankaranarayanan, I. B. Djordjevic, and B. Vasic, "Iteratively decodable codes on  $m$ -flats for WDM high-speed long-haul transmission," *J. Lightw. Technol.*, vol. 23, no. 11, pp. 3696–3701, Nov. 2005.
- [11] S. K. Chilappagari *et al.*, "Error floors of LDPC codes on binary symmetric channel," in *Proc. ICC 2006*, Istanbul, Turkey, 2006.