

BAYESIAN BCJR FOR CHANNEL EQUALIZATION AND DECODING

Luis Salamanca, Juan José Murillo-Fuentes

Teoría de la Señal y Comunicaciones,
Universidad de Sevilla
Camino de los Descubrimientos s/n,
41092, Seville, Spain.

E-mail: {salamanca, murillo}@us.es

Fernando Pérez-Cruz

Teoría de la Señal y Comunicaciones,
Universidad Carlos III de Madrid
Avda. de la Universidad 30,
28911, Leganés (Madrid), Spain.

E-mail: fernando@tsc.uc3m.es

ABSTRACT

In this paper we focus on the probabilistic channel equalization in digital communications. We face the single input single output (SISO) model to show how the statistical information about the multipath channel can be exploited to further improve our estimation of the a posteriori probabilities (APP) during the equalization process. We consider not only the uncertainty due to the noise in the channel, but also in the estimate of the channel estate information (CSI). Thus, we resort to a Bayesian approach for the computation of the APP. This novel algorithm has the same complexity as the BCJR, exhibiting lower bit error rate at the output of the channel decoder than the standard BCJR that considers maximum likelihood (ML) to estimate the CSI.

1. INTRODUCTION

Single input single output (SISO) communication channels can be characterized by a linear finite impulsive response that either represents the dispersive nature of a physical media or the multiple paths of wireless communications [1]. This representation causes inter-symbol interference (ISI) at the receiver end that can impair the digital communication. Given the channel estate information (CSI), the maximum likelihood sequence detector (MLSD) [2] (i.e. the Viterbi algorithm) provides the optimal decision for the whole transmitted sequence at the receiver end. Alternatively, the BCJR [3] algorithm computes the a posteriori probabilities (APP) for each transmitted symbol providing bitwise optimal decisions. These APP measure the uncertainty in the received symbols due to noise in the communication channel. After the equalizer a channel decoder purges the channel errors in the received sequence. This allows communication systems to yield maximal achievable rates when using reliable error correcting codes [4].

The CSI is typically acquired using pilots (a preamble) [1] and a maximum likelihood (ML) estimator. These preambles are typically short to reduce the transmission of non-

informative symbols, yielding inaccurate CSI estimates. In the following, we will refer to the BCJR equalizer with ML estimation of the channel as ML-BCJR. The ML-BCJR only delivers an approximation to the APP for each symbol because it does not include the uncertainty in the estimate. Inaccuracies in the APP estimates degrade the performance of modern channel decoders, such as turbo or low-parity density-check (LDPC) codes [5, 6]. The channel decoder may fail to deliver the correct transmitted codeword or may even fail to converge at all.

In this paper, we explore the use of a Bayesian equalizer that integrates the uncertainty in the CSI estimate to produce more accurate APP estimates. We found that when incorporating the uncertainty in the estimation of the CSI, the Markov property is lost. Hence, the BCJR cannot be run to compute the APP for each transmitted symbol. We propose a simple yet accurate approximation to the Bayesian solution that allows to recover the Markov property. In this novel solution, the Bayesian framework is embedded into the BCJR algorithm. The complexity of the resulting algorithm, hereafter referred to as Bayesian BCJR, is identical to the original BCJR algorithm.

The difference between the bit error rate (BER) of the ML-BCJR and the Bayesian BCJR equalizers is not significant, although it slightly favors the Bayesian equalizer. However, this is not an accurate measure of the quality of the APP estimates for each equalizer due to it only considers hard decisions, in contrast to the soft inputs needed by modern channel decoders. Thus, assuming LDPC coding in our communication system, we can compare the quality of the APP estimates for each equalizer. We experimentally show that at the output of the LDPC decoder the Bayesian BCJR equalizer considerably improves the performance of the ML-BCJR equalizer, when we measure the probability of error. These gains are more significant for high signal to noise ratios, channels with long impulsive responses and/or short training sequences.

We find in the literature related works that consider the

uncertainties in the estimation of the CSI. In the framework of turbo-receivers [7], this uncertainty in the estimation is considered in the iterative process of equalization and decoding, as in [8], where the authors use a MMSE to estimate the channel, and in [9, 10], where they do not focus on the optimal estimation of the APP. The inaccuracies in the estimation of different parameters of an OFDM system can be introduced in the computation of the APP, as proposed in [11]. In [12], the authors consider the channel estimation inaccuracies during the decoding process by means of a practical decoding metric. For nonlinear channels we have shown in [13] that accurate APP estimates increase the performance of LDPC decoders.

The paper is organized as follows. In Section 2 we describe the structure of a SISO communication system and the ML-BCJR solution. The proposed Bayesian approach and the algorithm to equalize are presented in Section 3. Experimental results in Section 4 help to illustrate the benefits of our method. Finally, Section 5 ends with conclusions and some proposals for future work.

2. ML-BCJR EQUALIZATION

2.1. System Model

We consider the discrete-time dispersive communication system depicted in Fig. 1. The channel $H(z)$ is completely specified by the CSI, i.e., $\mathbf{h} = [h_1, h_2, \dots, h_L]^\top$, where L is the length of the channel. We model \mathbf{h} as independent Gaussians with zero-mean and variance equal to $1/L$ (Rayleigh fading). A block of K message symbols, $\mathbf{m} = [m_1, m_2, \dots, m_K]^\top$, is encoded with a rate code $R = K/N$ to obtain the codeword $\mathbf{b} = [b_1, b_2, \dots, b_N]^\top$, which is transmitted over the channel using a BPSK modulation:

$$x_i = \mathbf{b}_i^\top \mathbf{h} + w_i, \quad (1)$$

where $\mathbf{b}_i = [b_i, b_{i-1}, \dots, b_{i-L+1}]^\top$ and w_i is additive white Gaussian noise (AWGN) with variance σ_w^2 . Thus, the received sequence is $\mathbf{x} = [x_1, x_2, \dots, x_N]^\top$.

At the beginning of every block we transmit a preamble with n known bits ($b_1^\circ, \dots, b_n^\circ$) and the receiver uses $\mathcal{D} = \{x_i^\circ, b_i^\circ\}_{i=1}^n$, the training sequence, to estimate the channel. The ML criterion maximizes:

$$\hat{\mathbf{h}}_{ML} = \arg \max_{\mathbf{h}} p(\mathbf{x}^\circ | \mathbf{b}^\circ, \mathbf{h}). \quad (2)$$

Second, we apply the BCJR algorithm to obtain an approximation to the APP for each transmitted bit:

$$p(b_i = b | \mathbf{x}, \hat{\mathbf{h}}_{ML}) \quad i = 1 \dots N, \quad (3)$$

where $b = \pm 1$ ¹. With these APP we decode the received

¹For ease of presentation, we consider that the transmitted symbols are bits, but the results in this paper can be readily generalized for non-binary alphabets.

word using the LDPC decoder to obtain a maximum a posteriori estimate for m_i .

2.2. BCJR Algorithm

In this section we review the BCJR algorithm [3] as it is the basis of the practical implementation of our algorithm. Given the CSI, the BCJR provides the APP estimates. These probabilities are computed as [14]:

$$p(b_i = b | \mathbf{x}, \mathbf{h}) = \sum_{(p,q) \in \mathcal{S}_b} \frac{p(s_i = p, s_{i+1} = q, \mathbf{x} | \mathbf{h})}{p(\mathbf{x} | \mathbf{h})}, \quad (4)$$

where s_i and s_{i+1} refer to the states in the equivalent Trellis at time i and $i+1$, and \mathcal{S}_b is the set of all possible transitions from $s_i = p$ to $s_{i+1} = q$ caused by the input $b_i = b$. The numerator in (4) can be expressed as:

$$p(s_i = p, s_{i+1} = q, \mathbf{x}_1^{i-1}, x_i, \mathbf{x}_{i+1}^N | \mathbf{h}) = \underbrace{p(s_i = p, \mathbf{x}_1^{i-1} | \mathbf{h})}_{\text{I}} \underbrace{p(s_{i+1} = q, x_i | s_i = p, \mathbf{h})}_{\text{II}} \underbrace{p(\mathbf{x}_{i+1}^N | s_{i+1} = q, \mathbf{h})}_{\text{III}}, \quad (5)$$

where the vector \mathbf{x} is divided in three different sets: received samples before instant i , \mathbf{x}_1^{i-1} ; after instant i , \mathbf{x}_{i+1}^N ; and the received sample at instant i , x_i . Some variables are deleted from the conditional probabilities due to the Markov property. In term II, \mathbf{x}_1^{i-1} does not provide any information given $s_i = p$, and in term III all the information about the received samples and the evolution in the Trellis is contained in the last state $s_{i+1} = q$. These probabilities are the three terms that the BCJR algorithm computes during its forward and backward recursions. We next include them as a function of the states before and after the symbol i has been transmitted:

$$\alpha_i(p) = p(s_i = p, \mathbf{x}_1^{i-1} | \mathbf{h}), \quad (6)$$

$$\gamma_i(p, q) = p(s_{i+1} = q, x_i | s_i = p, \mathbf{h}), \quad (7)$$

$$\beta_{i+1}(q) = p(\mathbf{x}_{i+1}^N | s_{i+1} = q, \mathbf{h}). \quad (8)$$

In the BCJR algorithm, (7) is computed at each stage as:

$$\gamma_i(p, q) = p(x_i | s_{i+1} = q, s_i = p, \mathbf{h}) p(s_{i+1} = q | s_i = p), \quad (9)$$

where by Markovity some variables are extracted from both probabilities. The second term is equal to a half for a binary alphabet. In the first term, the transition from $s_i = p$ to $s_{i+1} = q$ can be rewritten in terms of the set \mathbf{b}_i of transmitted bits, and assuming a channel with AWGN it follows that:

$$p(x_i | s_{i+1} = q, s_i = p, \mathbf{h}) = p(x_i | \mathbf{b}_i, \mathbf{h}) \sim \mathcal{N}(\mathbf{b}_i^\top \mathbf{h}, \sigma_w^2 \mathbf{I}). \quad (10)$$

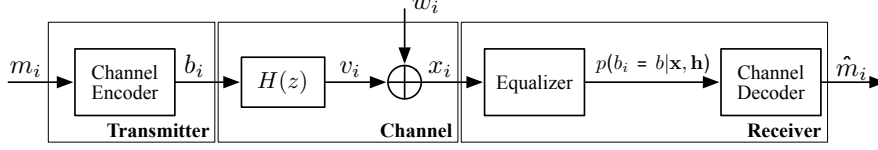


Fig. 1. System model.

The probabilities (6) and (8) are computed at each stage through the forward recursion,

$$\alpha_{i+1}(q) = \sum_{p=0}^{Q-1} \gamma_i(p, q) \alpha_i(p) \quad (11)$$

$$= p(s_{i+1} = q, \mathbf{x}_i^1 | \mathbf{h}), \quad (12)$$

and backward recursion,

$$\beta_i(p) = \sum_{q=0}^{Q-1} \gamma_i(p, q) \beta_{i+1}(q) \quad (13)$$

$$= p(\mathbf{x}_i^N | s_i = p, \mathbf{h}), \quad (14)$$

where Q are the states p (or q) at time i (or $i + 1$) that, for any value of b_i , end at time $i + 1$ (or i) in state q (or p). We assume that both recursions start from known states.

Through the forward and backward recursions [14] and by exploiting the Markov property the BCJR algorithm efficiently computes the summation in (4), providing optimal APP estimates for AWGN channels [3].

3. BAYESIAN EQUALIZATION

3.1. Introduction

If we provide the BCJR with the true CSI we have a perfect APP estimation [3]. However, in practice we estimate the CSI from a training sequence. The ML in (2) is a quite extended CSI estimator. And the APP estimates of the BCJR degrade with errors in the CSI. In this paper, we resort to a Bayesian approach to incorporate the uncertainty in the CSI through its posterior probability. Therefore, when we run the BCJR algorithm we take into account the uncertainty due to the noise over each received symbol and the uncertainty in the CSI estimate. We estimate the APP for each symbol as follows:

$$p(b_i = b | \mathbf{x}, \mathcal{D}) = \int p(b_i = b | \mathbf{x}, \mathbf{h}) p(\mathbf{h} | \mathcal{D}) d\mathbf{h}, \quad (15)$$

where $p(b_i = b | \mathbf{x}, \mathbf{h})$ is the APP computed by the BCJR algorithm for a given \mathbf{h} and $p(\mathbf{h} | \mathcal{D})$ is the CSI's posterior:

$$\begin{aligned} p(\mathbf{h} | \mathcal{D}) &= \frac{p(\mathbf{h}) p(\mathbf{x}^\circ | \mathbf{b}^\circ, \mathbf{h})}{p(\mathbf{x}^\circ | \mathbf{b}^\circ)} \\ &= \frac{p(\mathbf{h}) \prod_{i=1}^n p(x_i^\circ | \mathbf{b}_i^\circ, \mathbf{h})}{p(x_1^\circ, \dots, x_n^\circ | b_1^\circ, \dots, b_n^\circ)}, \end{aligned} \quad (16)$$

where in a channel with AWGN the likelihood is distributed as in (10) and, assuming Rayleigh fading, we can consider a zero-mean and unit-variance Gaussian prior for the computation of the posterior. Hence, the posterior is also Gaussian whose mean and covariance matrix are, respectively, equal to [15]:

$$\begin{aligned} \bar{\mathbf{h}}_{\mathcal{D}} &= (\mathbf{I} + \mathbf{B}^\circ (\mathbf{B}^\circ)^\top \sigma_w^{-2})^{-1} \mathbf{B}^\circ \sigma_w^{-2} \mathbf{x}^\circ, \\ \mathbf{C}_{\mathbf{h} | \mathcal{D}} &= (\mathbf{I} + \mathbf{B}^\circ (\mathbf{B}^\circ)^\top \sigma_w^{-2})^{-1}, \end{aligned} \quad (17)$$

where $\mathbf{B}^\circ = [\mathbf{b}_1^\circ, \mathbf{b}_2^\circ, \dots, \mathbf{b}_n^\circ]$.

Equation (15) is not analytically solvable. Numerical integration techniques, such as Monte Carlo, lead into time-demanding algorithms for the computation of the optimal APP in (15). Therefore, in the following we propose an approximate yet accurate algorithm that allows to incorporate the CSI's posterior into the BCJR algorithm, with no increase in the complexity compared to the ML-BCJR.

3.2. Bayesian BCJR

By Bayes rule, the APP in (4) is obtained as:

$$p(b_i = b | \mathbf{x}, \mathbf{h}) = \sum_{\mathbf{b}/b_i} p(\mathbf{b} | \mathbf{x}, \mathbf{h}) \quad (18)$$

$$= \sum_{\mathbf{b}/b_i} \frac{p(\mathbf{x} | \mathbf{b}, \mathbf{h}) p(\mathbf{b})}{p(\mathbf{x})} \quad (19)$$

$$= \frac{1}{Z} \sum_{\mathbf{b}/b_i} p(\mathbf{x} | \mathbf{b}, \mathbf{h}), \quad (20)$$

where \mathbf{b}/b_i denotes the sum over all possible values of $b_j \forall j \neq i$, i.e., the marginal probability for $b_i = b$. The prior is $p(\mathbf{b}) = 1/2^N$.

In a channel with AWGN, the likelihood in (20) yields:

$$p(\mathbf{x} | \mathbf{b}, \mathbf{h}) \sim \mathcal{N}(\mathbf{B}^\top \mathbf{h}, \sigma_w^2 \mathbf{I}), \quad (21)$$

and we can consider Markovity since x_i only depends on \mathbf{b}_i . It follows that the joint probability for \mathbf{x} in (20) factorizes as a product of independent Gaussians:

$$p(b_i = b|\mathbf{x}, \mathbf{h}) = \frac{1}{Z} \sum_{\mathbf{b}/b_i} \prod_{i=1}^{N+L-1} p(x_i|\mathbf{b}_i, \mathbf{h}). \quad (22)$$

As seen in Subsection 2.2, the BCJR algorithm exploits the Markov property to efficiently compute this summation in (22) through the forward and backward recursions [14], providing optimal APP estimates for AWGN channels [3].

Considering (20) in the Bayesian estimator in (15), it yields:

$$p(b_i = b|\mathbf{x}, \mathcal{D}) = \frac{1}{Z} \sum_{\mathbf{b}/b_i} \int p(\mathbf{x}|\mathbf{b}, \mathbf{h})p(\mathbf{h}|\mathcal{D})d\mathbf{h}, \quad (23)$$

that provides the optimal APP estimates thanks to consider the inaccuracies in the CSI estimate. As stated in (17), the CSI posterior is Gaussian and the marginalization in (23) can be analytically computed leading to:

$$p(b_i = b|\mathbf{x}, \mathcal{D}) = \frac{1}{Z} \sum_{\mathbf{b}/b_i} p(\mathbf{x}|\mathbf{b}, \mathcal{D}), \quad (24)$$

where this Gaussian likelihood is distributed as:

$$p(\mathbf{x}|\mathbf{b}, \mathcal{D}) \sim \mathcal{N}(\mathbf{B}^\top \bar{\mathbf{h}}_{\mathbf{h}|\mathcal{D}}, \mathbf{B}^\top C_{\mathbf{h}|\mathcal{D}} \mathbf{B} + \sigma_w^2 \mathbf{I}). \quad (25)$$

Compared to (21) and since the covariance matrix of the likelihood $p(\mathbf{x}|\mathbf{b}, \mathcal{D})$ is non-diagonal, this term can not be expressed as a product of independent Gaussians, and the Markov property is lost in (24). Therefore, we conclude that the forward and backward recursions approach cannot be used in the Bayesian equalizer in (15). This poses a major problem when computing the solution. In this paper we propose to reestablish the Markov property by using an approximation, which allows solving the Bayesian approach through the forward and backward recursion. In particular, we approximate (25) by:

$$p(\mathbf{x}|\mathbf{b}, \mathcal{D}) \approx p'(\mathbf{x}|\mathbf{b}, \mathcal{D}) \sim \mathcal{N}(\mathbf{B}^\top \bar{\mathbf{h}}_{\mathbf{h}|\mathcal{D}}, \Sigma), \quad (26)$$

where:

$$\Sigma_{ij} = \begin{cases} (\mathbf{B}^\top C_{\mathbf{h}|\mathcal{D}} \mathbf{B} + \sigma_w^2 \mathbf{I})_{ii}, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \quad (27)$$

Thus, (24) yields:

$$p(b_i = b|\mathbf{x}, \mathcal{D}) \approx \frac{1}{Z} \sum_{\mathbf{b}/b_i} \prod_{i=1}^{N+L-1} p'(x_i|\mathbf{b}_i, \mathcal{D}). \quad (28)$$

The APP obtained are a good approximation of the optimal APP depicted in (24), since the diagonal of the covariance matrix in (25) is dominant. Thanks to the reestablishment of the Markov property, (28) can be solved as (22). We have more accurate APP estimates at the same complexity as the ML-BCJR solution.

3.3. Practical Implementation

Based on the results in Section 2.2, we provide some details on the practical implementation of the forward-backward recursions algorithm to solve (28). Instead of computing γ assuming \mathbf{h} is obtained through some estimation criterion such as ML, we marginalize the CSI's posterior in (16) from the Gaussian likelihood in (9):

$$p'(x_i|\mathbf{b}_i, \mathcal{D}) = \int p(x_i|\mathbf{b}_i, \mathbf{h})p(\mathbf{h}|\mathcal{D})d\mathbf{h}. \quad (29)$$

The result of this integral can be analytically calculated since both terms are Gaussians:

$$p'(x_i|\mathbf{b}_i, \mathcal{D}) \sim \mathcal{N}(\mathbf{b}_i^\top \bar{\mathbf{h}}_{\mathbf{h}|\mathcal{D}}, \mathbf{b}_i^\top C_{\mathbf{h}|\mathcal{D}} \mathbf{b}_i + \sigma_w^2). \quad (30)$$

Thus, the γ term is obtained at each stage as:

$$\gamma_i^{\mathcal{D}}(p, q) = p'(x_i|\mathbf{b}_i, \mathcal{D})p(s_{i+1} = q|s_i = p) \quad (31)$$

$$= p(s_{i+1} = q, x_i|s_i = p, \mathcal{D}). \quad (32)$$

The α and β term are computed by means of the forward and backward recursions, as seen in (11) and (13) respectively, assuming they start from known states. Finally, the APP is calculated from the transition probabilities as in (4).

4. SIMULATION RESULTS

To illustrate the performance of the Bayesian BCJR, we compare its bit error rate curves to the ones of the ML-BCJR, before and after the decoder. In all experiments we consider the following scenario:

- Block frames of 500 random bits encoded with a regular LDPC code (3,6) of rate 1/2.
- Up to 10^6 frames of 1000 bits are transmitted over the channel.
- Between frames, a training sequence of n uncoded bits is transmitted to estimate the channel.
- Every frame, and its associated training sequence, is sent over the same Rayleigh fading channel. We consider that the channel coherence time is greater than the duration of the frame, i.e., the channel does not change during this time. All the frames are transmitted over the same channel to compute accurate estimates of the BER.
- We consider for the Bayesian estimation a prior with zero mean and variance equal to 1 for all taps.

In Fig. 2 we first depict the BER at the output of the equalizer, for a 3 taps channel and different lengths of the training sequence. As predicted, the difference between

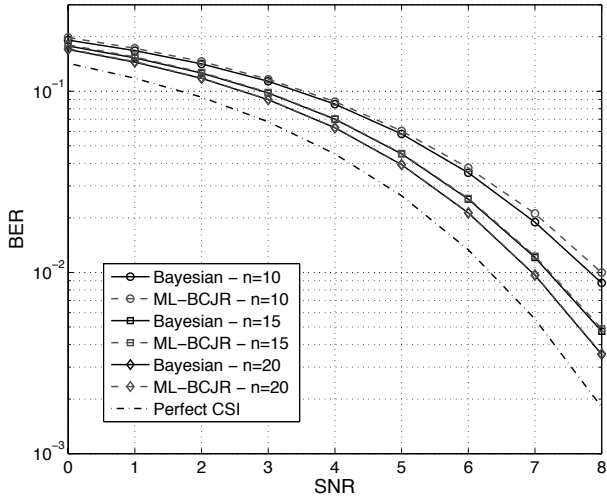


Fig. 2. BER performance for Bayesian equalizer (solid lines) and ML-BCJR (dashed lines) before the decoder, for a channel with 3 taps and different lengths of the training sequence, $n = 10$ (\circ), $n = 15$ (\square) and $n = 20$ (\diamond). Dash-dotted line illustrates performance assuming perfect CSI.

BER curves is negligible and decreases with the length of the training sequence.

It is the decoder the process that takes advantage of these better estimations of the APP provided by the Bayesian equalizer, improving its performance and achieving better results in terms of BER. In Fig. 3 we observe the results for different lengths of the training sequence for a channel with $L = 3$. At higher values of SNR and worst estimations of the channel, due to shorter training sequences, the difference between SNR values of both methods for a certain bit error rate increases. This yields to a gain over 0.5 dB around a SNR of 9 dB for $n = 15$. The BER curve assuming a perfect knowledge of the channel is included too to set a lower bound of performance for the system.

In Fig. 4 we include the BER curves for five lengths of the training sequence and a channel with $L = 6$. For this channel length we achieve a gain around 0.5 dB in the range of 9 – 10 dB and $n = 20$.

As the length of the training sequence becomes greater, the performance of the Bayesian equalizer tends to the ML-BCJR, and both methods achieve similar results in terms of BER, as we observe in Fig. 3 and Fig. 4.

5. CONCLUSIONS AND FUTURE WORK

Channel equalization has been traditionally solved using a discriminative model where the only variable modeled as random is noise. The generative model introduced in this paper, where the posterior probability density function of

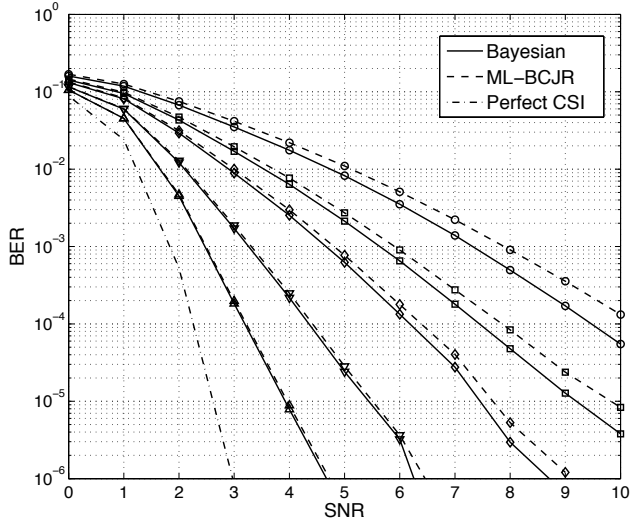


Fig. 3. BER performance for Bayesian equalizer (solid lines) and ML-BCJR (dashed lines) after the LDPC decoder, for a channel with 3 taps and different lengths of the training sequence, $n = 10$ (\circ), $n = 15$ (\square), $n = 20$ (\diamond), $n = 35$ (∇) and $n = 60$ (\triangle). Dash-dotted line illustrates performance assuming perfect CSI.

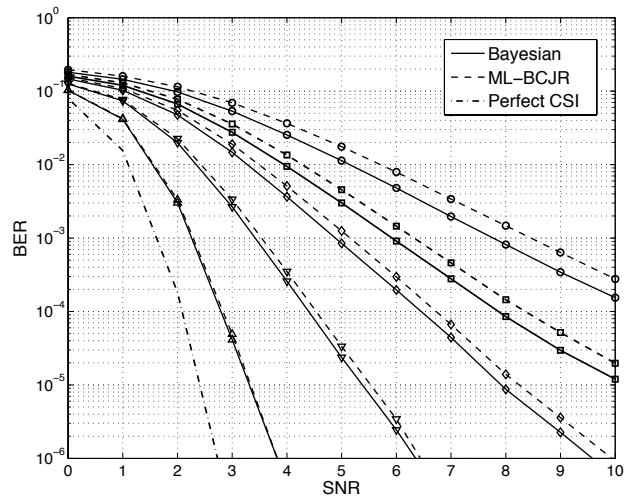


Fig. 4. BER performance for Bayesian equalizer (solid lines) and ML-BCJR (dashed lines) after the decoder, for a channel with 6 taps and different lengths of the training sequence, $n = 15$ (\circ), $n = 20$ (\square), $n = 25$ (\diamond), $n = 40$ (∇) and $n = 90$ (\triangle). Dash-dotted line illustrates performance assuming perfect CSI.

the estimated CSI is included, is a more principled solution. If we are to just estimate the encoded transmitted symbols, the discriminative model is a good choice. However, if the estimation of the APP is needed, i.e., the decoder very much

benefits from this information, the discriminative solution exhibits poor results whenever the CSI is badly estimated. On the contrary, the Bayesian approach exploits the full statistical model to provide better APP estimates. We show in the experimental section that these estimations are useful if a LDPC encoding is used. Other soft-decoders may take advantage of this solution as well.

The proposed algorithm approximates the Bayesian equalizer and we obtain the APP estimates, close to the optimal solution, with no increase in the complexity or the computation time. In case of inaccurate estimations, we obtain a remarkable gain at the decoder's output, while long enough training sequences lead us to the same results in terms of BER as the ML-BCJR.

Acknowledgment

This work was partially funded by Spanish government (Ministerio de Educación y Ciencia TEC2009-14504-C02-01, 02}, Consolider-Ingenio 2010 CSD2008-00010), and the European Union (FEDER).

6. REFERENCES

- [1] J. G. Proakis, *Digital Communications*, 5th ed. New York, NY: McGraw-Hill, 2008.
- [2] D. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 2, pp. 268–278, Mar. 1973.
- [3] 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. 20, no. 2, pp. 284–287, Mar. 1974.
- [4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New Jersey, USA: John Wiley & Sons, 2006.
- [5] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, March 2008.
- [6] S. Chung, D. Forney, T. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Commun. Lett.*, vol. 5, no. 2, pp. 58–60, Feb. 2001.
- [7] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glevieux, "Iterative correction of inter-symbol interference: Turbo-equalization," *European Trans. Telecomm.*, vol. 6, no. 5, pp. 507–511, Sep.-Oct. 1995.
- [8] L. M. Davis, I. B. Collings, and P. Hoeher, "Joint MAP equalization and channel estimation for frequency-selective and frequency-flat fast-fading channels," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2106–2114, Dec. 2001.
- [9] R. Otnes and M. Tuchler, "On iterative equalization, estimation, and decoding," in *Proc. IEEE Int. Conf. on Commun.*, vol. 4, May 2003, pp. 2958–2962.
- [10] X. Wang and R. Chen, "Blind Turbo equalization in Gaussian and impulsive noise," *IEEE Trans. Veh. Technology*, vol. 50, no. 4, pp. 1092–1105, Jul. 2001.
- [11] B. Lu and X. Wang, "Bayesian blind Turbo receiver for coded OFDM systems with frequency offset and frequency-selective fading," in *Proc. IEEE Int. Conf. on Commun.*, vol. 1, Apr.-May 2002, pp. 44–48.
- [12] P. Piantanida, S. Sadough, and P. Duhamel, "On the outage capacity of a practical decoder accounting for channel estimation inaccuracies," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1341–1350, May 2009.
- [13] P. M. Olmos, J. Murillo-Fuentes, and F. Perez-Cruz, "Joint nonlinear channel equalization and soft LDPC decoding with Gaussian processes," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1183–1192, Mar 2010.
- [14] T. K. Moon, *Error Correction Coding: mathematical methods and algorithms*, ser. Information Theory. John Wiley & Sons, 2005.
- [15] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation theory*. New York, NY: Prentice-Hall, 1993, vol. 1.