On the analysis of the MAP equalizer performance within an iterative receiver

Noura Sellami
ISECS
Route Menzel Chaker km 0.5
B.P 868, 3018 Sfax, Tunisia

Aline Roumy
IRISA-INRIA
Campus de Beaulieu
35042 Rennes Cedex, France

Inbar Fijalkow
Laboratoire ETIS, UMR 8051
ENSEA-UCP-CNRS
6 av. du Ponceau
95014 Cergy-Pontoise, France

Abstract — **In this paper, we consider an iterative receiver composed of a Maximum *A Posteriori* (MAP) equalizer and a MAP decoder. During the iterations, the equalizer and the decoder exchange extrinsic information and use them as *a priori* in order to improve their performance. We propose here to study analytically the impact of the *a priori* information provided by the channel decoder on the equalizer performance. We show that it is equivalent to a gain in terms of signal-to-noise ratio (SNR) and we provide an analytical expression. Simulation results show that this expession approximates quite well the SNR gain.**

I. Introduction

An important source of degradation in high data rate communication systems is the presence of intersymbol interference (ISI) between consecutive data symbols which is due to the frequency selectivity of mobile radio channels. In order to improve the quality of the transmission, an error correction code is generally used on top of an equalizer. The optimal receiver for a coded system performs joint equalization and decoding treating the concatenation of the encoder and the ISI channel as one code. However, the complexity of this receiver is in general prohibitive especially when an interleaver is used. A solution achieving a good complexity/performance trade-off is to use an iterative receiver constituted of a soft-input soft-output (SISO) equalizer and a SISO decoder [1], following the idea of turbo-codes [2]. The basic idea behind iterative processing is to exchange extrinsic information among the equalizer and the decoder in order to achieve successively refined performance.

The optimal SISO algorithm, in the sense of minimum bit error rate (BER), to be used for equalization and decoding is the symbol MAP algorithm [3]. Hence, the context of this paper considers an iterative receiver composed of a MAP equalizer and a MAP decoder. We propose to study analytically the impact of the *a priori* information provided by the channel decoder on the equalizer performance. To do that, we follow the approach of [4] and [5] which studied the impact of channel estimation errors on the equalizer performance. In [4], Gorokhov studied the impact of channel estimation errors on the performance of the Viterbi equalizer and showed that it is equivalent to a loss in SNR and evaluated this loss. In [5], we have extended the study to a List-type MAP equalizer prefiltered by the whitened matched filter, in the case of multiple-input multiple-output (MIMO) systems. In this paper, we will show that the use of the *a priori* information by the equalizer is equivalent to a gain in SNR and we will give an approximation of this gain.
This study is the first step in the convergence analysis of iterative receivers that we will present in future works. Our aim is to make the analysis in an analytical way on the contrary to the analysis based on extrinsic information transfer (EXIT) charts which uses simulations and becomes very difficult if it avoids them, for trellises with more than two states [6].
The paper is organized as follows. In section 2, we describe the system model. Section 3 recalls the principle of the iterative receiver based on MAP equalization and decoding. In section 4, we study the impact of the *a priori* information on the equalizer performance. In section 5, we give simulation results.

Throughout this paper scalars and matrices are lower and upper case respectively and vectors are underlined lower case. $(.)^T$ denotes the transposition.

II. System model

We consider a coded data transmission system over a frequency selective channel depicted in Figure 1. The input information bit sequence is first encoded with a convolutional encoder. The output of the encoder is interleaved, mapped to the symbol alphabet $\mathcal{A}$. For simplicity, we will consider only the BPSK modulation ($\mathcal{A}=\{+1,-1\}$). We assume that transmissions are organized into bursts of T symbols. The channel is supposed to be invariant during one burst. The received baseband signal sampled at the symbol rate at time k is

$$x_k = \sum_{l=0}^{L-1} h_l s_{k-l} + n_k \tag{1}$$

where L is the channel memory. In this expression, n_k are modeled as independent samples of a real white Gaussian noise with normal probability density function (pdf) $\mathcal{N}(0,\sigma^2)$ where $\mathcal{N}(\mu,\sigma^2)$ denotes a Gaussian distribution with mean μ and variance σ^2. The term h_l is the l^{th} tap gain of the channel, which is assumed to be real valued. Let $\underline{s} = (s_{T-1},...,s_{1-L})^T$ be the $(L+T-1)$-long vector of coded symbols and $\underline{n} = (n_{T-1},...,n_0)^T$ be the T-long noise vector. The output of the channel is the T-long vector $\underline{x} = (x_{T-1},...,x_0)^T$ defined as

$$\underline{x} = \tau(h)\underline{s} + \underline{n} \tag{2}$$

where $\tau(h)$ is a $T \times (T+L-1)$ Toeplitz matrix with its first row equal to $(h_0, h_1, ..., h_{L-1}, 0, ..., 0)$ and its first column $(h_0, 0, ..., 0)^T$.

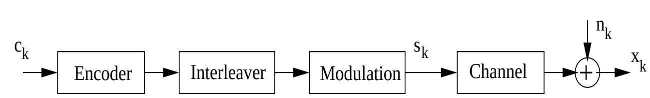

Figure 1: Transmitter structure

When the channel is known and no *a priori* information is provided to the equalizer, the data estimate according to the sequence MAP criterion (equivalently the maximum likelihood (ML) criterion since there is no *a priori*) is given by

$$\underline{\hat{s}}_{MAP} = \arg\min_{\underline{u}} \left(\|\underline{x} - \tau(h)\underline{u}\| : \underline{u} \in \mathcal{A}^{T+L-1} \right). \tag{3}$$

We now consider a particular error event characterized by its length m [7]. Thus, we suppose that there exists an interval of size m such that all the symbols of $\underline{\hat{s}}$ are different from the corresponding symbols of $\underline{s}$ while the preceding symbol and the following one are the same for $\underline{s}$ and $\underline{\hat{s}}$. Define $\underline{s}_m$ and $\underline{\hat{s}}_m$ to be the vectors of symbols corresponding to this interval and the vector of errors $\underline{e}_m = \underline{\hat{s}}_m - \underline{s}_m$. A subevent $\mathcal{E}_m$ of the error event is that $\underline{\hat{s}}$ is better than $\underline{s}$ in the sense of the ML metric

$$\mathcal{E}_m : \|\underline{x}_m - \tau_m(h)\underline{\hat{s}}_m\| \le \|\underline{x}_m - \tau_m(h)\underline{s}_m\| \tag{4}$$

where $\underline{x}_m$ is the subvector of $\underline{x}$ and $\tau_m(h)$ is the block of $\tau(h)$ corresponding to the error interval. The probability $P(\mathcal{E}_m)$ of $\mathcal{E}_m$ is given by [7]:

$$P(\mathcal{E}_m) = Q\left(\frac{\|\underline{\epsilon}_m\|}{2\sigma}\right) \tag{5}$$

where $\underline{\epsilon}_m = \tau_m(h)\underline{e}_m$ and $Q(\alpha) = \frac{1}{\sqrt{\pi}}\int_\alpha^\infty \exp(-y^2)dy$. Let Σ_m be the set of all possible error events of length m. Then, the probability, $P(\Sigma_m)$, that any error event is of length m is bounded by the sum of the probabilities of the subevents $\mathcal{E}_m$

$$P(\Sigma_m) \le \sum_{\mathcal{E}_m} P(\mathcal{E}_m). \tag{6}$$

Let d_{min} be the channel minimum distance [7]. Because of the exponential decrease of the Gaussian distribution function, the overall probability of error $P(\Sigma) \le \sum_m P(\Sigma_m)$ will be dominated at high SNR by the term involving the minimum value $d_{\min}$ of $\|\underline{\epsilon}_m\|$. Thus

$$P(\Sigma) \simeq Q\left(\frac{d_{\min}}{2\sigma}\right) \tag{7}$$

Our goal is to find an approximation of $P(\Sigma)$ when the equalizer is integrated into an iterative receiver. In this case, at each iteration, *a priori* information are provided to the equalizer by the channel decoder.

III. Iterative receiver

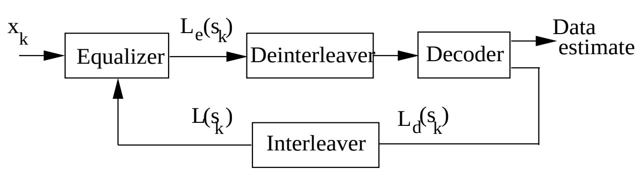

Figure 2: Transmitter structure

As shown in Figure 2, the receiver consists of two soft-input soft-output (SISO) processors, the equalizer and the decoder. We consider only the MAP approach for both equalization and decoding, using the BCJR algorithm [3]. The MAP equalizer computes the *a posteriori* probabilities (APPs) on the coded bits, $P(s_k = s|\underline{x})$, $s \in \mathcal{A}$, $1 - L \le k \le T - 1$, and outputs the log-likelihood ratios (LLRs) [6]:

$$\begin{aligned} L_e(s_k) &= L(s_k|\underline{x}) - L(s_k) \\ &= \log\frac{P(s_k = +1|\underline{x})}{P(s_k = -1|\underline{x})} - \log\frac{P(s_k = +1)}{P(s_k = -1)} \end{aligned} \tag{8}$$

which are the *a posteriori* LLRs $L(s_k|\underline{x})$ minus the *a priori* LLRs $L(s_k)$. These *a priori* LLRs are provided by the decoder. At the first receiver iteration, $L(s_k) = 0$ since no *a priori* information are available. The LLRs $L_e(s_k)$ are then deinterleaved and provided to the decoder as input information, in order to refine its calculations. The MAP decoder computes the APPs $P(s_k = s|\underline{r})$, $\underline{r} = (L_e(s_{1-L}), \cdots, L_e(s_{T-1}))^T$, and outputs the LLRs

$$L_d(s_k) = \log\frac{P(s_k = +1|\underline{r})}{P(s_k = -1|\underline{r})} - \log\frac{P(s_k = +1)}{P(s_k = -1)}.$$

These LLRs are then interleaved and provided to the equalizer as *a priori*, $L(s_k)$, at the next iteration. After some iterations, hard decisions are taken on the information bits by the decoder.

IV. Performance analysis

Now, we want to evaluate the gain in performance due to the use of *a priori* information by the MAP equalizer. The study will be done here for the equalizer using the sequence MAP criterion. It holds for the symbol MAP equalizer using the BCJR algorithm [3] since the two equalizers have almost the same performance as shown in [8, page 814].

Proposition: Suppose we are given a frequency selective channel with additive white gaussian noise (AWGN) and noise variance σ^2. Assume that the outputs of an AWGN channel with noise variance σ_a^2 are also available as observations (corresponding here to the *a priori* observations). Then, at high SNR, the MAP equalizer using the *a priori* information is equivalent to the MAP equalizer having no *a priori* information but with an equivalent signal-to-noise ratio

$$\hat{SNR} = SNR(1 + \frac{8\mu^2}{d_{\min}^2}) \tag{9}$$

where SNR is the true signal-to-noise ratio, $\mu = \frac{\sigma}{\sigma_a}$ and d_{min} is the channel minimum distance as defined in [7].

Remark: The representation of the *a priori* information as the outputs of an AWGN channel is an accurate representation of the decoder outputs. Actually, it was shown in [9][10] that it is equivalent to have at the equalizer input a set of observations

$$z_k = s_k + w_k \tag{10}$$

where $w_k \sim \mathcal{N}(0, \sigma_a^2)$. Thus, the LLRs $L(s_k)$ fed back from the decoder can be modeled as independent and identically distributed (i.i.d) samples from a random variable with the conditional pdf $\mathcal{N}(\pm\frac{2}{\sigma_a^2}, \frac{4}{\sigma_a^2})$ for some σ_a^2 [9], where the polarity of the mean is equal to s_k.

Proof:
The proof is divided into three parts. First, the probability of an error subevent of length m, $P(\mathcal{E}_m)$, is derived and then upper bounded. Finally, the overall probability of error, $P(\Sigma)$, is calculated in order to find an approximation of the equivalent SNR.

Proof-part1: $P(\mathcal{E}_m)$
Taking into account the *a priori* information, the *a posteriori* probability of the sequence $\underline{s}$ is given by

$$p(\underline{s}|\underline{x},\underline{z}) \propto \exp\left(-\frac{\|\underline{x}-\tau(h)\underline{s}\|^2}{2\sigma^2}\right)\exp\left(-\frac{\|\underline{z}-\underline{s}\|^2}{2\sigma_a^2}\right) \quad (11)$$

where $\underline{z}= (z_{T-1},...,z_{1-L})^T$. The data estimate according to the sequence MAP criterion is then given by

$$\hat{\underline{s}}_{MAP} = \arg\min_{\underline{u}}\left(\|\underline{x}-\tau(h)\underline{u}\|^2+\frac{\sigma^2}{\sigma_a^2}\|\underline{z}-\underline{u}\|^2 : \underline{u}\in\mathcal{A}^{T+L-1}\right).$$

A subevent $\mathcal{E}_m$ of the error event of length m is that $\hat{\underline{s}}_m$ is better than $\underline{s}_m$ in the sense of the sequence MAP metric

$$\mathcal{E}_m : \|\underline{x}_m-\tau_m(h)\hat{\underline{s}}_m\|^2+\frac{\sigma^2}{\sigma_a^2}\|\underline{z}_m-\hat{\underline{s}}_m\|^2 \leq$$
$$\|\underline{x}_m-\tau_m(h)\underline{s}_m\|^2+\frac{\sigma^2}{\sigma_a^2}\|\underline{z}_m-\underline{s}_m\|^2. \quad (12)$$

Let $\mu=\frac{\sigma}{\sigma_a}$, $\underline{y}= (x_{T-1},x_{T-2},\cdots,x_0,\mu z_{T-1},\cdots,\mu z_{1-L})^T$, $M=\left((\tau(h))^T,\mu I_{T+L-1}\right)^T$ a $(L-1)\times(2T+L-1)$ matrix and $\underline{n}_w=(n_{T-1},n_{T-2},\cdots,n_0,\mu w_{T-1},\cdots,\mu w_{1-L})^T$. Using (1) and (10), we can write

$$\underline{y}=M\underline{s}+\underline{n}_w. \quad (13)$$

The data estimate according to the sequence MAP criterion is then given by

$$\hat{\underline{s}}_{MAP}=\arg\min_{\underline{u}}\left(\|\underline{y}-M\underline{u}\|^2 : \underline{u}\in\mathcal{A}^{T+L-1}\right).$$

Hence, (12) is equivalent to

$$\mathcal{E}_m : \left\|\underline{y}_m-M_m\hat{\underline{s}}_m\right\|^2 \leq \left\|\underline{y}_m-M_m\underline{s}_m\right\|^2 \quad (14)$$

where $\underline{y}_m$ is the $(2m+L-1)\times 1$ subvector of $\underline{y}$ corresponding to the error interval and $M_m=\left((\tau_m(h))^T,\mu I_m\right)^T$.

Since the components of the vector $\underline{n}_w$ are independent samples of a real white Gaussian noise with pdf $\mathcal{N}(0,\sigma^2)$ and using the result given in (5), the probability of the error event $P(\mathcal{E}_m)$ is

$$P(\mathcal{E}_m)=Q\left(\frac{\|\underline{E}_m\|}{2\sigma}\right) \quad (15)$$

where $\underline{E}_m=\underline{M}_m(\hat{\underline{s}}_m-\underline{s}_m)=\underline{M}_m\underline{\epsilon}_m$. Since the modulation used is the BPSK and then $(\hat{\underline{s}}_m-\underline{s}_m)$ is a vector with m components equal to ± 2, we obtain,

$$\begin{aligned}\|\underline{E}_m\| &= \|\underline{M}_m(\hat{\underline{s}}_m-\underline{s}_m)\| \\ &= \left\|\left((\tau_m(h))^T,\mu I_m\right)^T(\hat{\underline{s}}_m-\underline{s}_m)\right\| \\ &= \sqrt{\|\epsilon_m\|^2+4m\mu^2}. \quad (16)\end{aligned}$$

Thus, we can write,

$$P(\mathcal{E}_m)=Q\left(\frac{\sqrt{\|\underline{\epsilon}_m\|^2+4m\mu^2}}{2\sigma}\right). \quad (17)$$

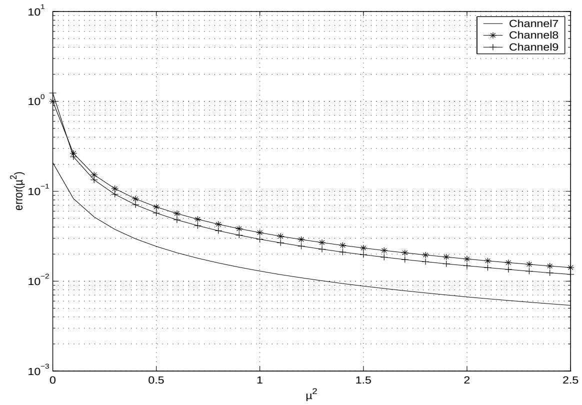

Figure 3: The error (19) made on the bound $\left(d_{\min}^2+8\mu^2\right)$ with respect to μ^2.

Note that when no *a priori* information is available, $\mu=0$ and then (5) is equal to (17).

Proof-part2: lower bound for $\left(\|\underline{\epsilon}_m\|^2+4m\mu^2\right)$ **:**
In order to find an approximation of $P(\Sigma)$, the overall probability of error, we want now to find a lower bound for the quantity $\left(\|\underline{\epsilon}_m\|^2+4m\mu^2\right)$. Actually, at high SNR, this term will dominate the sum of the probabilities of the error events (because of the exponential decrease of the Gaussian distribution function). By definition, $\|\underline{\epsilon}_m\|^2\geq d_{\min}^2$. Moreover, we have $m\geq 2$. Thus, a lower bound for the quantity $\left(\|\underline{\epsilon}_m\|^2+4m\mu^2\right)$ is given by

$$bound(\mu^2)=\left(d_{\min}^2+8\mu^2\right). \quad (18)$$

This bound is reached generally for channels with memory L less than 6, since for these channels the error sequence allowing to attain the minimum distance is of length $m=2$ (see examples of channels in [11]). For channels with memory greater than 6, this bound is still a tight bound. We propose to show this for three channels given here by their impulse responses [11]:

- Channel7: (0.18; 0.32; 0.48; 0.53; 0.48; 0.32; 0.18)
- Channel8: (0.16; 0.24; 0.43; 0.49; 0.49; 0.43; 0.24; 0.16)
- Channel9: (0.11; 0.21; 0.35; 0.46; 0.48; 0.46; 0.35; 0.21; 0.11).

For these channels, the minimum distance error sequence is of length $m=6$ [11] and thus it does not allow to approach the bound (18). However, the input sequence error $(+2,-2)$ allows to approach closely the bound and that's why we consider it in the following. On figure 3, we plot the normalized error between the bound $bound(\mu^2)=\left(d_{\min}^2+8\mu^2\right)$ and the quantity $\left(\|\underline{\epsilon}_m\|^2+4m\mu^2\right)$ computed for the input sequence error $(+2,-2)$, for the three channels. This error is given by

$$error(\mu^2)=\frac{\|\underline{\epsilon}_m\|^2-d_{\min}^2}{bound(\mu^2)}. \quad (19)$$

Figure 3 shows that this error is low. Actually, for $\mu>0.3$, i.e., $\sigma>0.3\sigma_a$, the error is less than 10%. Thus the considered bound is a tight bound, especially when μ is high.

Proof-part3: $P(\Sigma)$
As in the case without *a priori,* at high SNR, the probability $P(\Sigma)$ can be approximated by

$$P(\Sigma) \simeq Q\left(\frac{\sqrt{d_{\min}^2 + 8\mu^2}}{2\sigma}\right)$$
$$= Q\left(\frac{d_{\min}}{2\sigma}\sqrt{1 + \frac{8\mu^2}{d_{\min}^2}}\right). \quad (20)$$

Comparing (7) with (20), we can conclude that the effect of the *a priori* is similar to a gain in SNR. Actually, the expression of the error probability given in (20) can be seen as the one given in (7) with an equivalent signal-to-noise ratio

$$S\hat{N}R = SNR(1 + \frac{8\mu^2}{d_{\min}^2}). \quad (21)$$

In the following, we propose to verify this analytical result by simulations.

V. Simulation results

In our simulations, we consider the following channels [11]:

- Channel3: $(0.5; 0.71; 0.5)$
- Channel5: $(0.29; 0.50; 0.58; 0.50; 0.29)$

The modulation used is the BPSK. The transmissions are organized into bursts of 512 symbols. Figures 4 and 5 show the Bit Error Rate (BER) curves with respect to the SNR, for different values of the ratio $\mu = \frac{\sigma}{\sigma_a}$, for Channel3 and Channel5. Each curve is obtained while the ratio μ is kept constant. The solid lines indicate the equalizer performance obtained by simulations. The dotted lines are obtained by shifting the curve corresponding to the case with no *a priori* and with a perfect channel knowledge ($\mu = 0$) by the values of the SNR shift: $10\log_{10}(1 + \frac{8\mu^2}{d_{\min}^2})$. Table.1 shows the values of the minimum error distance $d_{\min}$ and the minimum distance input error sequence for the channels of interest [11].

	Channel3	Channel5
$d_{\min}$	1.5308	1.0532
Error sequence	$(2, -2)$	$(2, -2)$

Table.1

For both channels, the bound given in (18) is reached for the error sequence $(2, -2)$ since it is the input sequence allowing to reach the minimum distance. Figure 4 shows that the theoretical curves (dotted lines) approximate well the BER for $\mu < 1$. When μ increases, the approximation becomes erroneous. Figure 5 shows that the theoretical curves fit better those obtained by simulations for different values of μ for Channel5. Thus, the approximation is better for Channel5. We can also conclude that the approximation holds in general for $\sigma < \sigma_a$. For a given σ_a, the analytical expression holds for high SNR, such as $\sigma < \sigma_a$. Figure 6, compares the performance obtained by simulations and the theoretical performance when σ_a is kept constant equal to 0.5. We notice that the theoretical curve approximates very well the BER when the SNR is high (SNR>5dB). This is coherent with the assumption we made previously in the performance analysis. Actually, in order to obtain (20), we assumed that the SNR is high.

VI. Conclusion

In this paper, we considered an iterative receiver composed of a Maximum *a posteriori* (MAP) equalizer and a MAP decoder. We proposed to study analytically the impact of the *a*

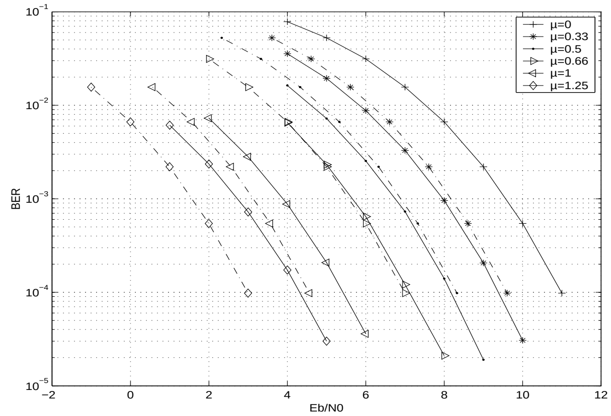

Figure 4: Comparison of the equalizer performance (solid curves) and the theoretical performance (dotted curves) obtained using (21) for Channel3.

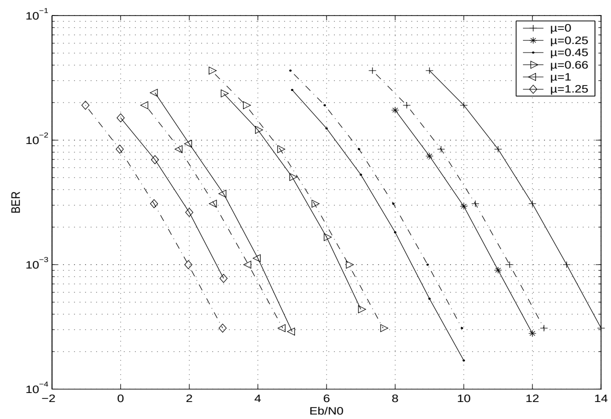

Figure 5: Comparison of the equalizer performance and the theoretical performance obtained using (21) for Channel5.

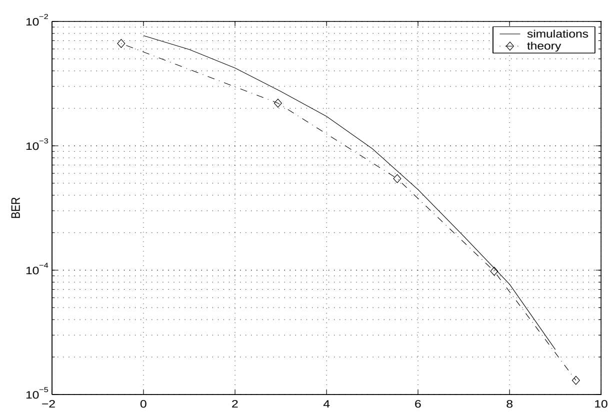

Figure 6: Comparison of the equalizer performance and the theoretical performance obtained using (21) when $\sigma_a = 0.5$, for Channel3.

priori information on the equalizer performance. We gave an approximation of the error probability which allows us to find an expression of the gain in terms of the SNR due to the use of the *a priori* information. Simulation results show that this approximation is quite good especially for long channels. This work is a first step in the study of the convergence analysis of iterative receivers.

References

[1] C.Douillard, M.Jézéquel, C.Berrou, A.Picart, P.Didier, and A.Glavieux, "Iterative correction of intersymbol interference: turbo-equalization," *European Trans. Telecommun.,* vol. 6, no. 5, pp. 507-511, 1995.

[2] C.Berrou, A.Glavieux, and P.Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," *IEEE Int. Conf. Communications*, Geneva, Switzerland, pp.1064-1070, May 1993.

[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-32, pp.284-287, March 1974.

[4] A.Gorokhov, "On the performance of the Viterbi equalizer in the presence of channel estimation errors," *IEEE Signal Process. Letters,* vol. 5, no. 12, pp. 321-324, December 1998.

[5] N.Sellami, I.Fijalkow, and S.Perreau, "Performance analysis of a List-type equalizer over estimated MIMO frequency selective channels, " *EUSIPCO'02*, Toulouse, France, September 2002.

[6] M.Tüchler, R.Koetter, and A.Singer, "Turbo equalization: principles and new results," *IEEE Trans. on Comm.*, vol. 50, no. 5, pp. 754-767, May 2002.

[7] G.D.Forney, Jr., "Maximum-likelihood sequence estimation for digital sequences in the presence of intersymbol interference," *IEEE Trans. Infom. Theory*, vol. 18, pp. 363-378 , May 1972.

[8] S. Benedetto and E. Biglieri, *Principles of Digital Transmission with Wireless Applications,* NewYork: Kluwer/Plenum, 1999.

[9] S.Ten Brink, "Convergence of iterative decoding," *IEEE Electronic Letters,* vol. 35, pp.806-808, May 1999.

[10] J.Hagenauer, E.Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. on Inf. Theory*, vol. 42, Issue 2, pp. 429-445, March 1996.

[11] W.Ser, K.Tan, and K.Ho, "A new method for determining "unknown" worst-case channels for maximum-likelihood sequence estimation," *IEEE Trans. on Comm.*, vol. 46, no. 2, pp. 164-168, February 1998.