A proof of convergence of the MAP turbo-detector to the AWGN case *

Noura Sellami[1*], Aline Roumy[2], Inbar Fijalkow [3]

[1] ISECS, Route Menzel Chaker km 0.5, B.P 868, 3018 Sfax, Tunisia
Email: noura.sellami@isecs.rnu.tn, Tel: (216)97059318, Fax: (216)74274088
[2] IRISA-INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France
Email: aline.roumy@irisa.fr, Tel: (33)299847394, Fax: (33)299847171
[3] ETIS, ENSEA-UCP-CNRS, 95014 Cergy-Pontoise, France
Email: Inbar.Fijalkow@ensea.fr, Tel: (33)130736610, Fax: (33)130736627

Abstract

In this paper, we consider a coded transmission over a frequency selective channel. We propose to study analytically the convergence of the turbo-detector using a maximum *a posteriori* (MAP) equalizer and a MAP decoder. We show that the densities of the extrinsic Log Likelihood Ratios (LLRs) exchanged during the iterations are e-symmetric and output-symmetric. Under the Gaussian approximation, this property allows to perform a one-dimensional (1-D) analysis of the turbo-detector. By deriving the analytical expressions of the extrinsic LLR distributions under the Gaussian approximation, we prove that the bit error rate (BER) performance of the turbo-detector converges to the BER performance of the coded Additive White Gaussian Noise (AWGN) channel at high signal to noise ratio (SNR), for any frequency selective channel.

EDICS:

- PERF Performance analysis and bounds

- CEST Channel estimation and equalization

- MODL Modulation and encoding

1 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. To combat the effects of ISI, an equalizer has to be used. The optimal equalizer, in the sense of minimum sequence error rate (SER) or bit error rate (BER) is based on maximum *a posteriori* (MAP) detection. We distinguish two criteria, MAP sequence detection and MAP symbol detection. When no *a priori* information on the transmitted data is available, MAP detection turns into maximum likelihood (ML) detection. Efficient algorithms exist for MAP sequence detection, for example the SER optimizing Viterbi algorithm [8], and MAP symbol detection, for example the BER optimizing BCJR algorithm [1]. These algorithms are interesting since their complexity grows linearly rather than exponentially with the sequence size. The performance of both algorithms is almost the same [2] and it is limited by the frequency selectivity of the channel. At high signal to noise ratio (SNR), a gap between the Additive White Gaussian Noise (AWGN) channel performance and the equalizer performance is observed and depends on the minimum distance of the frequency selective channel (the minimum Euclidean distance between two noise-free channel output sequences).

In order to improve the quality of the transmission, an error correction code is generally used, together with an equalizer. At the receiver, a solution achieving a good complexity/performance trade-off is to use an iterative receiver consisting of a soft-input soft-output (SISO) equalizer and a SISO decoder [6], following the idea of turbo-codes [3]. The basic idea behind iterative processing is to exchange extrinsic information among the equalizer and the decoder in order to achieve successively refined performance. In this paper, we consider a turbo-detector composed of an equalizer and a decoder using the symbol MAP criterion [1]. A natural question concerns the achievable performance of such a turbo-detector. It is believed (due to simulations) that the scheme converges to the performance without ISI. The aim of this work is to show that the BER performance of the turbo-detector converges to the BER performance of the coded AWGN channel at high SNR.

To address this question, one needs to characterize analytically the convergence point of an iterative receiver as the number of iterations goes to infinity and specify the probability of error achieved at this convergence point. Iterative schemes have been extensively studied. Exact analyses for finite sequence length exist [23, 24] and show the convergence to a constrained form of sequence MAP receiver. However, the convergence point cannot be specified analytically. Iterative receivers have also been studied by noticing that they are a special instance of the message-passing algorithm. The iterative receiver can therefore be analyzed by computing the

2

Evolution of the Densities (DE) of the messages exchanged by the message-passing receiver. This is performed in [16] under the assumption of infinite sequence length. The DE technique is therefore an infinite-dimensional dynamical system, for which the closed form expression of the recursion exists when the size of all variables is two, such that the messages are real numbers (often the Log Likelihood Ratio (LLR) is taken). However, to determine the asymptotic (in terms of the number of iterations) density, one needs to resort to simulations.

In order to reduce the computational burden of the DE, faster techniques have been proposed that approximate the DE by a one-dimensional (1-D) dynamical system. The approximate techniques differ in the scalar to be tracked and in the way to project the density on this scalar. The parameter can be the SNR of the extrinsic LLRs [5, 7]. It can also be mutual information between the transmitted symbol and the associated LLR [21, 22] which leads to the extrinsic information transfer (EXIT) charts. All these methods therefore require two more assumptions:

- the messages are assumed to be independent and identically distributed (i.i.d) Gaussian random variables (conditioned on the transmitted symbols). Wiberg shows in [25] that the Gaussian approximation is efficient for the extrinsic LLRs. In [10], it was proved that it is a good approximation for turbo-decoding with infinite block length. This approximation, with output-symmetry, reduces the dimension of the dynamical system to two.

- the messages are further assumed to be e-symmetric. A density p is e-symmetric if for $x \in \{+1, -1\}$ and $y \in \mathcal{R}$, $p_{Y|X}(y|x) = \exp(xy)p_{Y|X}(-y|x)$. The e-symmetry in our paper is referred to as symmetry by Richardson and Urbanke [17]. This reduces the dimension of the dynamical system to one.

The major disadvantage of these methods is that they use Monte Carlo simulations and provide only numerical results. When the trellis has only two states, closed form expressions of the EXIT functions have been derived [18]. However, in the case of turbo-decoders or turbo-detectors using the MAP criterion, it is difficult to study analytically the performance of a MAP equalizer or a MAP decoder having *a priori* information and a large number of states.

In our work, we seek a closed form expression of the recursion, analyzing the turbo-detector for large trellis sizes (more than two states). Note that such expression exists for two concatenated codes under one more assumption (high SNR) [13]. In the case of linear turbo equalization, analytical studies of the equalizer performance were performed in [12, 15]. To the best of our knowledge, no closed form expression of the 1-D

recursion of the turbo-detector using the MAP equalizer exists.

Main contribution of the paper: We prove the convergence of the turbo-detector BER performance to the coded AWGN channel BER performance for any frequency selective channel. To do this, we derive analytical expressions of the distributions of the extrinsic LLRs at the output of the equalizer and the decoder. Our closed form analysis holds under some commonly used assumptions in the 1-D analyses of iterative schemes, i.e., Gaussianity, e-symmetry and output-symmetry of the extrinsic LLRs, independence of the noise modeling the extrinsic LLRs (true for infinite block length) and high SNR (to insure the tightness of the bounds on the error probability). We prove in this paper that the densities of the extrinsic LLRs are e-symmetric and output-symmetric and that they approach a Gaussian density at high SNR and for infinite block length.

This paper is organized as follows. In Section 2, we describe the system model and the iterative receiver. In Section 3, we give a proof of the preservation of e-symmetry and output-symmetry under MAP equalization. In Section 4, we derive analytically asymptotic approximations of the BER achieved by the equalizer and give the distribution of the extrinsic LLRs at its output. In Section 5, we derive analytical expressions of the distribution of the extrinsic LLRs at the output of the MAP decoder with *a priori* information. In Section 6, we perform the convergence analysis of the turbo-detector and give the proof of convergence to the AWGN case. Throughout this paper, deterministic quantities and random variables are lower and upper case, respectively. Vectors and matrices are underlined and double underlined, respectively. The operator $(.)^T$ denotes the transposition and $\underline{\underline{I}}_N$ is the $N \times N$ identity matrix.

2   General framework

We consider a coded data transmission system over a frequency selective channel as depicted in Figure 1. The input information bit sequence is first encoded with a nonrecursive nonsystematic convolutional encoder with rate r_c. The output of the encoder is interleaved and mapped to symbol alphabet $\mathcal{A}$. For simplicity, we will consider only BPSK modulation ($\mathcal{A} = \{+1, -1\}$). We assume that transmissions are organized into bursts of T symbols. The channel is assumed to be invariant during the transmission. 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 $L \geq 1$ is the channel constraint length and s_k, for $1 - L \leq k \leq T - 1$, are the transmitted symbols. The channel memory is $L - 1$ with 2^{L-1} possible states. In this expression, n_k are modeled as independent

samples from a random variable with normal probability density function (pdf) $\mathcal{N}(0, \sigma^2)$ where $\sigma^2 = N_0/2$ and $\mathcal{N}(\alpha, \sigma^2)$ denotes a Gaussian distribution with mean α and variance σ^2. The term h_l is the lth tap gain of the channel and is assumed to be real valued with $\sum_{l=0}^{L-1} h_l^2 = 1$.

The optimal receiver for this 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 consisting of a soft-input soft-output (SISO) equalizer and a SISO decoder [6]. As shown in Figure 2, we consider the iterative receiver consisting of SISO processors, the equalizer
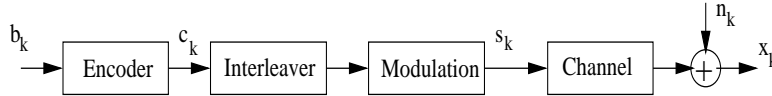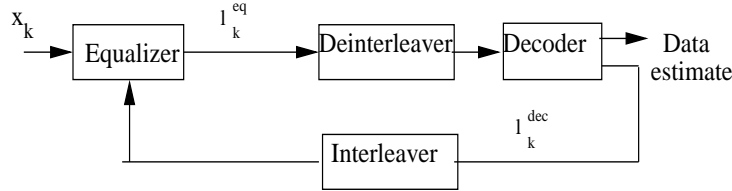

Figure 1: Transmitter structure

Figure 2: Receiver structure

and the decoder. We consider only the MAP approach for both equalization and decoding, using the BCJR algorithm [1]. We assume that the channel is perfectly known at the receiver. The MAP equalizer computes the *a posteriori* probabilities (APPs) on the coded bits, $P_{S_k|\underline{X}}(s|\underline{x})$, $s \in \mathcal{A}$, $1 - L \leq k \leq T - 1$, $\underline{x} = (x_{T-1}, ..., x_0)^T$, and outputs the extrinsic Log Likelihood Ratios (LLRs) [22]:

$$l_k^{eq} = \log \frac{P_{S_k|\underline{X}}(+1|\underline{x})}{P_{S_k|\underline{X}}(-1|\underline{x})} - \log \frac{P_{S_k}(+1)}{P_{S_k}(-1)} \tag{2}$$

which are the *a posteriori* LLRs minus the *a priori* LLRs provided by the decoder. At the first receiver iteration, the *a priori* LLRs, $\log \frac{P_{S_k}(+1)}{P_{S_k}(-1)}$, are equal to zero since no *a priori* information is available. The extrinsic LLRs l_k^{eq} are then deinterleaved and provided to the decoder as input *a priori* information in order to refine its calculations. The MAP decoder computes the APPs $P_{S_k|\underline{E}}(s|\underline{e})$, $\underline{e} = (l_{T-1}^{eq}, \cdots, l_{1-L}^{eq})^T$, and outputs the extrinsic LLRs

$$l_k^{dec} = \log \frac{P_{S_k|\underline{E}}(+1|\underline{e})}{P_{S_k|\underline{E}}(-1|\underline{e})} - \log \frac{P_{S_k}(+1)}{P_{S_k}(-1)}.$$

These LLRs are then interleaved and provided to the equalizer as *a priori* LLRs at the next iteration. After

some iterations, hard decisions are taken on the information bits by the decoder.

In the rest of the paper, we want to prove the convergence of the MAP turbo-detector BER performance to the coded AWGN channel BER performance.

3   Properties of the extrinsic LLRs

We assume that the extrinsic LLRs are independent. This assumption holds for large block length with good and randomly chosen interleavers. Wiberg shows in [25] that the distribution of the extrinsic LLRs in iterative decoding approaches Gaussian like distribution as the number of iterations increases. In [10], it was proved that this is a good approximation for turbo-decoding with infinite block length. We will prove in Section 4 that this approximation holds for MAP equalization. This approximation reduces the dimension of the dynamical system to four (the means and variances of the distributions of the extrinsic LLRs for symbols equal to +1 and to -1). In [17], Richardson *et.al* show that output-symmetry and e-symmetry are preserved under MAP decoding. In this section, we propose to generalize this result to MAP equalization. The output-symmetry reduces the dimension of the dynamical system to two (the distributions of the extrinsic LLRs for symbols equal to +1 and to -1 have the same variance and opposite means). The e-symmetry reduces the dimension of the dynamical system to one (the variance is twice the absolute value of the mean).

3.1   Symmetry

We first recall some definitions of the symmetry.

Definition 1 *A density p is e-symmetric if $p_{Y|X}(y|x) = \exp(xy)p_{Y|X}(-y|x)$ for all $x \in \{+1, -1\}$ and $y \in \mathcal{R}$.*

Definition 2 *A density p is said to be output-symmetric if $p_{Y|X}(y|x) = p_{Y|X}(-y| -x)$. A channel with input x and output y is said output-symmetric if its transition probability is output-symmetric.*
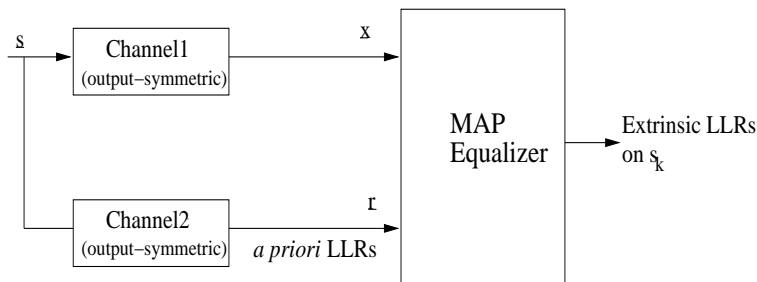

Figure 3: Transmission scheme

Proposition 3 *Suppose that the information sequence $\underline{s}$ of BPSK symbols is transmitted over an output-symmetric channel, Channel 1. The received sequence is denoted $\underline{x}$. Assume that a sequence of a priori LLRs is available at the output of an output-symmetric channel with input $\underline{s}$, Channel 2, and also provided to the*

MAP equalizer (see Figure 3). Then, the density of the a posteriori LLRs at the output of the MAP equalizer, having at its input $\underline{x}$ and the a priori LLRs, is both e-symmetric and output-symmetric.

We give the proof of Proposition 3 in Appendix I.

Remark: Since Channel 2 is output-symmetric, the *a priori* LLRs have an e-symmetric and output-symmetric density [17] (see Figure 3).

Corollary 4 *Assuming that Channel 1 and Channel 2 are independent and Channel 2 is memoryless, the density of the extrinsic LLRs at the output of the MAP equalizer, having at its input $\underline{x}$ and the a priori LLRs, is e-symmetric and output-symmetric.*

We give the proof of Corollary 4 in Appendix II.

Remark: Proposition 3 holds for any output-symmetric channels. There is no need to further assume that the channels are independent and memoryless. However, the assumptions in Corollary 4 are needed for the e-symmetry of the extrinsic LLRs to hold.

Corollary 5 *Assuming that the equivalent channel at the output of the decoder (Channel 2) is an AWGN channel and that the interleaver has infinite length, the densities of the extrinsic LLRs exchanged between the MAP equalizer and the MAP decoder in the turbo-detector of Figure 2 are e-symmetric and output-symmetric.*

Proof of Corollary 5:

We use in the following notations of Proposition 3. As shown in Figure 1, the information sequence is transmitted over a frequency selective channel with AWGN (Channel 1) which is an output-symmetric channel. At the first iteration of the receiver, no *a priori* information is available at the input of the equalizer. It is a particular case where we can apply Proposition 3. Hence, the densities of the extrinsic LLRs at the output of the equalizer are e-symmetric and output-symmetric. Since e-symmetry and output-symmetry are preserved under MAP decoding [17], the densities of the extrinsic LLRs at the output of the decoder are e-symmetric and output-symmetric at the first iteration. Moreover, since the interleaver has infinite length, Channel 1 and Channel 2 are independent. Since Channel 2 is an AWGN channel, it is also memoryless. Then, the equalizer in the turbo-detector is equivalent to the MAP equalizer of Corollary 4. Notice that the MAP equalizer and the MAP decoder in the turbo-detector use the independence assumption which is not verified for short length interleavers.

7

Given the property of e-symmetry and output-symmetry preservation under MAP decoding [17] and MAP equalization (Corollary 4), we can prove by induction that the densities of the extrinsic LLRs exchanged during the next iterations between the equalizer and the decoder are also e-symmetric and output-symmetric.

3.2 Gaussian approximation

In [10], it was shown that for a convolutional code (in a turbo-code) with an infinite block size, if the received signal is subject to AWGN and the *a priori* LLRs at the input of the decoder have a Gaussian density, then the density of the extrinsic LLRs at its output can be well approximated by a Gaussian density. The proof given in [10] cannot be generalized to the case of MAP equalization since the received signal is not only subject to AWGN but also to ISI. In Section 4, we will prove the convergence of the distribution of the extrinsic LLRs at the output of the equalizer to a Gaussian distribution at high SNR and for infinite block size.

The LLRs l_k^{eq} and l_k^{dec} are samples of random variables L_k^{eq} and L_k^{dec}, respectively. In the following, we want to derive the distributions of the extrinsic LLRs L_k^{eq} and L_k^{dec} at the output of the equalizer and the decoder. We have shown that the densities of the extrinsic LLRs are e-symmetric and output-symmetric. Since they are well approximated by a Gaussian density, the variance is the double of the absolute value of the mean [5, 7]. This allows us to perform a 1-D analysis of the turbo-detector.

4 Analysis of the equalizer

The performance of the Viterbi equalizer in the presence of AWGN was studied by Forney in [8]. This study assumed that no *a priori* information is provided to the equalizer. In the following, we propose to study analytically the impact of the *a priori* information on the MAP equalizer. The study will be done here for the equalizer using the MAP sequence criterion. We assume that this study holds for the MAP symbol equalizer using the BCJR algorithm [1] since the two equalizers have almost the same performance at high SNR as observed in [2, page 814]. We first derive analytically asymptotic approximations of P_e, the BER based on the *a posteriori* LLRs at the output of the equalizer fed with *a priori* information. Then, we derive analytical expressions of the distribution of the extrinsic LLRs L_k^{eq}.

4.1 Upper bound and asymptotic approximations of the BER

In [8, 2], an upper bound was derived on the symbol error probability achieved by the Viterbi equalizer for an ISI channel when no *a priori* information is available. Here, we propose to follow the reasoning of [8, 2] to derive an upper bound on P_e when the MAP sequence equalizer is fed with *a priori* LLRs. Based on this upper

bound, we derive asymptotic approximations of P_e.

We assume that the *a priori* observations at the input of the MAP equalizer are modeled as the outputs of an AWGN channel with zero mean and variance $\sigma^2_{eq,in}$. The *a priori* LLRs are then modeled as independent and identically distributed (i.i.d) samples from a random variable with the conditional pdf $\mathcal{N}(\frac{2s_k\mu^2_{eq,in}}{\sigma^2}, \frac{4\mu^2_{eq,in}}{\sigma^2})$, where $\mu_{eq,in} \triangleq \frac{\sigma}{\sigma_{eq,in}}$. We consider an error event as it is defined in [8, 2]. We define a state at time k as $\sigma_k = (s_{k-1,}, ..., s_{k-L+1})$. An error event is said to extend from time k_1 to k_2 if the estimated states and the correct states are equal at times k_1 and k_2 and are different for $k_1 < k < k_2$. Let $\underline{s}_e$ be the vector of transmitted symbols and $\underline{\hat{s}}_e$ be the estimated vector corresponding to the error event. Let $\underline{e} = \underline{\hat{s}}_e - \underline{s}_e$ be the vector of symbol errors. Let E be the set of all nonzero error events, $m(\underline{e})$ the number of decision errors in $\underline{e}$ and $P(\underline{e})$ the probability of $\underline{e}$ to occur. The overall probability of error is given by

$$P_e = \sum_{\underline{e} \in E} m(\underline{e})P(\underline{e}).$$

Let $A(\underline{e})$ be the event that the transmitted sequence $\underline{s}_e$ of data symbols is compatible with the occurrence of $\underline{e}$ ($\underline{\hat{s}}_e = \underline{s}_e + \underline{e}$ is an allowable input sequence). Then, $\underline{e}$ occurs if $A(\underline{e})$ occurs and $\underline{\hat{s}}_e = \underline{s}_e + \underline{e}$ has a MAP sequence metric better than any other possible sequence of symbols, including $\underline{s}_e$. Thus, the probability of $\underline{e}$ can be upper bounded as follows:

$$P(\underline{e}) \leq P_{\underline{s}_e, \underline{\hat{s}}_e} P(A(\underline{e}))$$

where $P_{\underline{s}_e, \underline{\hat{s}}_e}$ is the probability that $\underline{\hat{s}}_e = \underline{s}_e + \underline{e}$ has a MAP sequence metric better than $\underline{s}_e$.

We showed in [19] that the pairwise error probability that the MAP sequence equalizer with *a priori* information chooses $\underline{\hat{s}}_e$ instead of $\underline{s}_e$ is given by

$$P_{\underline{s}_e, \underline{\hat{s}}_e} = Q\left(\frac{\sqrt{\|\underline{d}(\underline{e})\|^2 + 4m(\underline{e})\mu^2_{eq,in}}}{2\sigma}\right),$$

where $Q(\alpha) = \frac{1}{\sqrt{2\pi}} \int_\alpha^\infty \exp(-\frac{y^2}{2})dy$, $\underline{d}(\underline{e})$ is the convolution of $\underline{e}$ with the channel and $\mu_{eq,in} \triangleq \frac{\sigma}{\sigma_{eq,in}}$.

Then,

$$P_e \leq \sum_{\underline{e} \in E} m(\underline{e})P(A(\underline{e}))Q\left(\frac{\sqrt{\|\underline{d}(\underline{e})\|^2 + 4m(\underline{e})\mu^2_{eq,in}}}{2\sigma}\right). \tag{3}$$

We define α_{min} as

$$\alpha_{min} = \min_{\underline{e} \in E} \sqrt{\|\underline{d}(\underline{e})\|^2 + 4m(\underline{e})\mu^2_{eq,in}}.$$

Let $E_{\min}$ be the set of all $\underline{e}$ achieving this minimum value α_{min}. When the SNR is high, because of the exponential decrease of the Gaussian distribution function, P_e can be approximated by the term involving the minimum value α_{min} as:

$$P_e \simeq \psi_{\min} Q\left(\frac{\alpha_{min}}{2\sigma}\right)$$

where $\psi_{\min} = \sum_{\underline{e} \in E_{\min}} m(\underline{e}) P\left(A(\underline{e})\right).$

Proposition 6 *Suppose we are given a frequency selective channel with L taps and AWGN with variance σ^2. Assume that the a priori observations at the input of the MAP equalizer are modeled as the outputs of an AWGN channel with zero mean and variance $\sigma^2_{eq,in}$. Then, at high SNR, the overall probability of error P_e at the output of the equalizer is approximated as:*

$$\begin{cases} P_e \simeq \psi_{\min} Q\left(\frac{\sqrt{\tilde{d}^2_{\min} + 4m_{\min}\mu^2_{eq,in}}}{2\sigma}\right), & \text{if } \tilde{d}^2_{\min} < 4 \text{ and } \mu_{eq,in} < \mu_{eq,\lim} \\ P_e \simeq Q\left(\frac{\sqrt{1+\mu^2_{eq,in}}}{\sigma}\right), & \text{if } \tilde{d}^2_{\min} \geq 4 \text{ or } (\tilde{d}^2_{\min} < 4 \text{ and } \mu_{eq,in} \geq \mu_{eq,\lim}) \end{cases} \tag{4}$$

where $\mu_{eq,in} \triangleq \frac{\sigma}{\sigma_{eq,in}}$, $\left(m_{\min}, \tilde{d}_{\min}\right) = \arg\min_{m(\underline{e}) \geq 2, \|\underline{d}(\underline{e})\|} \sqrt{\|\underline{d}(\underline{e})\|^2 + 4m(\underline{e})\mu^2_{eq,in}}$ and $\psi_{\min} = \sum_{\underline{e} \in E_{\min}} m(\underline{e}) P\left(A(\underline{e})\right).$ When $\tilde{d}^2_{\min} < 4$, $\mu_{eq,\lim} = \sqrt{\frac{4-\tilde{d}^2_{\min}}{4(m_{\min}-1)}}$.

Proof of Proposition 6:

In order to derive the asymptotic approximate expressions of the BER, we have to give a precise expression of α_{min} depending on the channel and on the *a priori* information characteristics. We distinguish two cases: the case of unreliable *a priori* information ($\mu_{eq,in}$ is low) and the case of reliable *a priori* information ($\mu_{eq,in}$ is high).

We first consider the case of unreliable *a priori* information. Generally, in the MAP equalizer, errors occur in bursts. This is still true here since the *a priori* information is not very reliable. Thus, we do not consider isolated errors since they occur rarely and we assume that $m(\underline{e}) \geq 2$. Let

$$\left(m_{\min}, \tilde{d}_{\min}\right) = \arg\min_{m(\underline{e}) \geq 2, \|\underline{d}(\underline{e})\|} \sqrt{\|\underline{d}(\underline{e})\|^2 + 4m(\underline{e})\mu^2_{eq,in}}.$$

Thus,

$$\alpha_{min} = \sqrt{\tilde{d}^2_{\min} + 4m_{\min}\mu^2_{eq,in}}. \tag{5}$$

When $\mu_{eq,in}$ is high, most of the *a priori* observations are very reliable and have more influence on the detection than the channel observations. Since the *a priori* information are independent, errors will not occur in bursts.

Actually, in this case, isolated errors corresponding to the few non reliable *a priori* observations will occur and will dominate the overall probability of error. Thus, α_{min} is achieved for $m(\underline{e}) = 1$ and $\|\underline{d}(\underline{e})\|^2 = 4$,

$$\alpha_{min} = 2\sqrt{1 + \mu_{eq,in}^2}. \tag{6}$$

We first assume that $\tilde{d}_{\min}^2 < 4$. Let $\mu_{eq,\lim}$ be the value of $\mu_{eq,in}$ giving equality between (5) and (6):

$$\mu_{eq,\lim} = \sqrt{\frac{4 - \tilde{d}_{\min}^2}{4(m_{\min} - 1)}}. \tag{7}$$

We can easily show that when $\mu_{eq,in} < \mu_{eq,\lim}$ (unreliable *a priori* information), the minimum value of $\sqrt{\|\underline{d}(\underline{e})\|^2 + 4m(\underline{e})\mu_{eq,in}^2}$ is achieved for $m(\underline{e}) = m_{\min}$ and $\|\underline{d}(\underline{e})\|^2 = \tilde{d}_{\min}^2$ and when $\mu_{eq,in} > \mu_{eq,\lim}$ (reliable *a priori* information), it is achieved for $m(\underline{e}) = 1$ and $\|\underline{d}(\underline{e})\|^2 = 4$.

When $\mu_{eq,in} < \mu_{eq,\lim}$, the overall probability of error is approximated at high SNR by

$$P_e \simeq \psi_{\min} Q\left(\frac{\sqrt{\tilde{d}_{\min}^2 + 4m_{\min}\mu_{eq,in}^2}}{2\sigma}\right), \tag{8}$$

where $\psi_{\min} = \sum_{\underline{e} \in E_{\min}} m(\underline{e}) P(A(\underline{e}))$.

When $\mu_{eq,in} > \mu_{eq,\lim}$ (reliable *a priori* information), since the modulation used is BPSK, $E_{\min} = (\pm 2)$. We have $A(+2) = -1$ and $A(-2) = +1$, and then $\psi_{\min} = \frac{1}{2} + \frac{1}{2} = 1$. Hence, the overall probability of error is approximated at high SNR by:

$$P_e \simeq Q\left(\frac{\sqrt{1 + \mu_{eq,in}^2}}{\sigma}\right). \tag{9}$$

If $\tilde{d}_{\min}^2 \geq 4$, α_{min} is always achieved for $m(\underline{e}) = 1$ and $\|\underline{d}(\underline{e})\|^2 = 4$. Thus, P_e is approximated as in (9) for all $\mu_{eq,in}$.

4.1.1 Simulation results

We propose to verify by simulations the tightness of the asymptotic expressions given in (4). In the simulations, we do not use channel coding and the turbo-detector yet. We consider Channel 5 and Channel 7 with impulse responses $(0.289; 0.499; 0.579; 0.499; 0.289)$ and $(0.179; 0.318; 0.477; 0.527; 0.477; 0.318; 0.179)$, respectively. In the case of unreliable *a priori* information, for Channel 5, α_{min} is achieved for $m_{\min} = 2$ and $\tilde{d}_{\min} = 1.035$ [20], giving $\mu_{eq,\lim} = 0.855$. Notice that in this case, $\tilde{d}_{min}$ is equal to the channel minimum distance [20]. Therefore, the same sequences will achieve the minimum value α_{min} and the channel minimum distance. By contrast, for Channel 7, α_{min} and the channel minimum distance are achieved by different sequences. More precisely, α_{min} is achieved for $m_{\min} = 2$ and $\tilde{d}_{\min} = 0.796$ giving $\mu_{eq,\lim} = 0.917$, whereas the channel minimum distance is

0.724, obtained by error sequences of length 6 [20]. For both channels, when $\mu_{eq,in} < \mu_{eq,\lim}$, the set of error sequences achieving α_{min} is $E_{\min} = \{(2,-2),(-2,2)\}$. Since $A(+2,-2) = (-1,+1)$ and $A(-2,+2) = (+1,-1)$, $\psi_{\min} = 2.\frac{1}{4} + 2.\frac{1}{4} = 1$. We provide the equalizer with Gaussian *a priori* LLRs with the conditional pdf $\mathcal{N}(\frac{2s_k\mu^2_{eq,in}}{\sigma^2}, \frac{4\mu^2_{eq,in}}{\sigma^2})$, for a given $\mu_{eq,in} = \frac{\sigma}{\sigma_{eq,in}}$. In Figures 4 and 5, we plot the BER curves with respect to the SNR, for different values of the ratio $\mu_{eq,in}$, respectively for Channel 5 and Channel 7. Here, SNR$= E_s/N_0$ where E_s is the energy per transmitted symbol and $N_0 = 2\sigma^2$. Each curve is obtained while the ratio $\mu_{eq,in}$ is kept constant. The solid lines indicate the BER performance based on the *a posteriori* LLRs obtained by simulations. The dotted lines are obtained by considering the analytical expressions given in Proposition 6. We notice that the analytical curves approximate well the curves obtained by simulations when the BER is less than 10^{-2}. The approximation is less accurate when $\mu_{eq,in}$ is close to $\mu_{eq,\lim}$. Indeed, in this case, isolated errors and double errors occur and have almost equal probabilities. The error probability then becomes a combination of the probabilities of these two types of error events and the approximate expressions become less accurate. However, as will be shown in the following, the approximation is still satisfactory to predict the distribution of the extrinsic LLRs at the output of the equalizer.
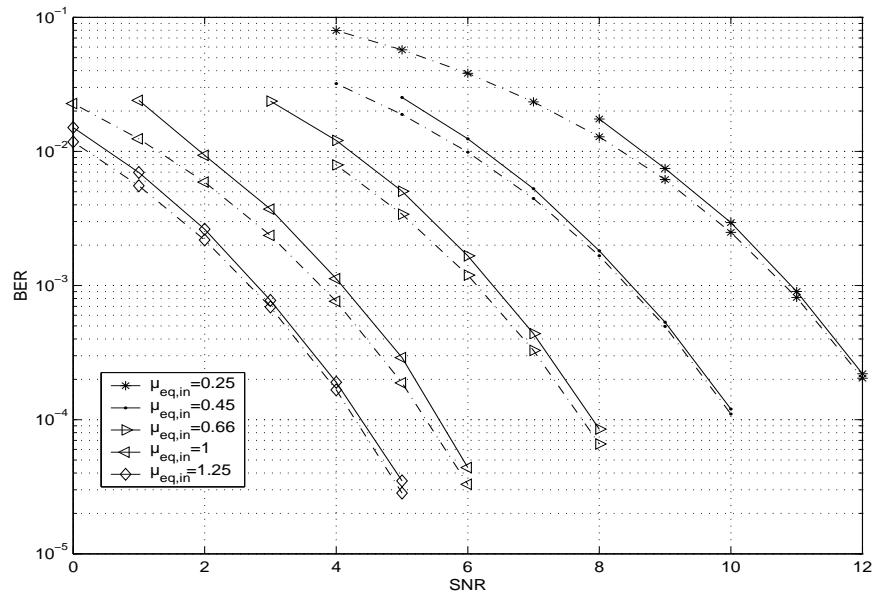

Figure 4: BER versus SNR: comparison of the equalizer BER performance (solid curves) and the asymptotic approximation (dotted curves), when the detection is based on the *a posteriori* LLRs, for Channel 5.

4.2 Gaussian approximation at the output of the equalizer

In [9, 2], a lower bound was derived on the symbol error probability achieved by the Viterbi equalizer for an ISI channel when no *a priori* information is available. By following the reasoning of [9, 2], we can derive expressions of tight lower bounds on the BER of the *a posteriori* LLRs at the output of the MAP sequence equalizer when
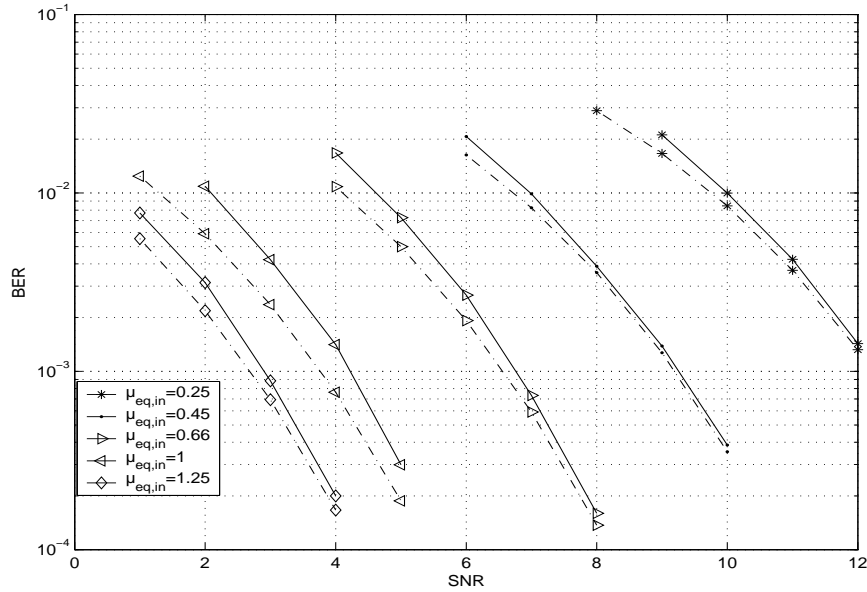
Figure 5: BER versus SNR: comparison of the equalizer BER performance (solid curves) and the asymptotic approximation (dotted curves), when the detection is based on the *a posteriori* LLRs, for Channel 7.

it is fed with *a priori* LLRs. We can show that

$$
\begin{cases}
P_e \geq \varphi_{\min} Q\left(\dfrac{\sqrt{\tilde{d}_{\min}^2 + 4m_{\min}\mu_{eq,in}^2}}{2\sigma}\right), & \text{if } \tilde{d}_{\min}^2 < 4 \text{ and } \mu_{eq,in} < \mu_{eq,\lim} \\[3mm]
P_e \geq Q\left(\dfrac{\sqrt{1+\mu_{eq,in}^2}}{\sigma}\right), & \text{if } \tilde{d}_{\min}^2 \geq 4 \text{ or } (\tilde{d}_{\min}^2 < 4 \text{ and } \mu_{eq,in} \geq \mu_{eq,\lim}),
\end{cases}
\tag{10}
$$

where $\varphi_{\min} = \sum_{\underline{e} \in E_{\min}} P\left(A(\underline{e})\right)$.

The error probability is upper bounded by (3) and lower bounded by (10). The upper bound is dominated by the term involving the minimum value α_{min}, at high SNR. Thus, at high SNR, there exists constants A, B and C such that $AQ(\frac{\alpha_{min}}{2\sigma}) \leq P_e \leq BQ(\frac{\alpha_{min}}{2\sigma}) + C$ where A and B are defined in (10) and (4) and C is the term of the upper bound negligible at high SNR. Notice that the influence of A and B becomes negligible at high SNR. Since the error probability achieved for LLRs with a Gaussian density is given by the function $Q(.)$, we conclude that at high SNR and for infinite block length, the density of the *a posteriori* LLRs at the output of the MAP equalizer with *a priori* information converges to a Gaussian density. Notice that when $\tilde{d}_{\min}^2 \geq 4$ or $\mu_{eq,in} \geq \mu_{eq,\lim}$, $A = B = 1$ and P_e is well approximated by $Q(\frac{\alpha_{min}}{2\sigma})$. Then, the Gaussian approximation becomes more accurate in this case.

4.3 Extrinsic LLR modeling

Proposition 7 *Suppose we are given a frequency selective channel with L taps and AWGN with variance σ^2. Assume that the a priori observations at the input of the MAP equalizer are modeled as the outputs of*

an AWGN channel with zero mean and variance $\sigma^2_{eq,in}$. Then, at high SNR, the extrinsic LLRs L^{eq}_k on the transmitted symbols s_k at the output of the MAP equalizer are modeled as random variables with the conditional pdf $\mathcal{N}(\frac{2s_k \mu^2_{eq,out}}{\sigma^2}, \frac{4\mu^2_{eq,out}}{\sigma^2})$ where

$$
\begin{cases}
\mu^2_{eq,out} = f_{eq}(\mu^2_{eq,in}) \overset{\Delta}{=} \frac{\tilde{d}^2_{\min} + 4(m_{\min}-1)\mu^2_{eq,in} + 4\gamma\sigma^2}{4} \ , & \text{if } \tilde{d}^2_{\min} < 4 \ and \ \mu_{eq,in} < \mu_{eq,\lim} \\
\mu^2_{eq,out} = f_{eq}(\mu^2_{eq,in}) \overset{\Delta}{=} 1, & \text{if } \tilde{d}^2_{\min} \geq 4 \ or \ (\tilde{d}^2_{\min} < 4 \ and \ \mu_{eq,in} \geq \mu_{eq,\lim}),
\end{cases}
\tag{11}
$$

where $\mu_{eq,in} = \frac{\sigma}{\sigma_{eq,in}}$, $\mu_{eq,\lim} = \sqrt{\frac{4-\tilde{d}^2_{\min}}{4(m_{\min}-1)}}$, $\gamma = -2\log(\psi_{\min})$ and $\psi_{\min} = \sum_{\underline{e} \in E_{\min}} m(\underline{e}) P(A(\underline{e}))$.

Remark: Notice that when $\mu_{eq,in} \geq \mu_{eq,\lim}$ (good *a priori* information) or $\tilde{d}^2_{\min} \geq 4$, the extrinsic LLRs at the equalizer output are modeled as random variables with the conditional pdf $\mathcal{N}(\frac{2s_k}{\sigma^2}, \frac{4}{\sigma^2})$ and are then equivalent to the LLRs corresponding to the AWGN channel with zero mean and variance σ^2, as though the effects of the ISI were eliminated. In a turbo-detector, the equalizer provides the decoder with the extrinsic LLRs l^{eq}_k at each iteration. Hence, if the *a priori* information becomes reliable or $\tilde{d}^2_{\min} \geq 4$, the BER of the decoder on the information bits (which is also the BER of the turbo-detector) is equivalent to the BER of the coded AWGN channel. In Section 6, we will give the proof of the convergence of the BER performance of the turbo-detector to the BER performance of the coded AWGN channel for any frequency selective channel.

Proof of Proposition 7:

We first consider the case where $\tilde{d}^2_{\min} < 4$. When the *a priori* information are unreliable ($\mu_{eq,in} < \mu_{eq,\lim}$), the BER at the output of the MAP equalizer is approximated by

$$
P_e \simeq \psi_{\min} Q\left(\frac{\sqrt{\tilde{d}^2_{\min} + 4m_{\min}\mu^2_{eq,in}}}{2\sigma}\right).
\tag{12}
$$

Since $Q(\sqrt{z})e^{-\frac{y}{2}} \geq Q(\sqrt{z+y})$ [2], we obtain

$$
P_e \gtrsim Q\left(\frac{\sqrt{\tilde{d}^2_{\min} + 4m_{\min}\mu^2_{eq,in} + 4\gamma\sigma^2}}{2\sigma}\right)
\tag{13}
$$

where $\gamma = -2\log(\psi_{\min})$.

At high SNR, the BER can be approximated by:

$$
P_e \simeq Q\left(\frac{\sqrt{\tilde{d}^2_{\min} + 4m_{\min}\mu^2_{eq,in} + 4\gamma\sigma^2}}{2\sigma}\right).
\tag{14}
$$

14

We know that the BER for BPSK modulation, for an AWGN channel with zero mean and variance σ_1^2 is

$$P_e = Q\left(\frac{1}{\sigma_1}\right). \tag{15}$$

Comparing (14) and (15), we can conclude that the BER performance of the MAP equalizer when it is provided with *a priori* information is equivalent to the BER performance achieved for an AWGN channel with zero mean noise and variance $\sigma_1^2 = \frac{4\sigma^2}{\tilde{d}_{\min}^2 + 4m_{\min}\mu_{eq,in}^2 + 4\gamma\sigma^2}$. Hence, the *a posteriori* LLRs at the output of the MAP equalizer can be modeled as i.i.d samples from a random variable with pdf $\mathcal{N}(\frac{2s_k}{\sigma_1^2}, \frac{4}{\sigma_1^2})$.

Using (2) and assuming that the *a priori* LLRs and the extrinsic LLRs are independent thanks to the interleaver [11], we obtain

$$L_k^{eq} \sim \mathcal{N}\left(s_k\left(\frac{2}{\sigma_1^2} - \frac{2\mu_{eq,in}^2}{\sigma^2}\right), \left(\frac{4}{\sigma_1^2} - \frac{4\mu_{eq,in}^2}{\sigma^2}\right)\right).$$

Therefore,

$$L_k^{eq} \sim \mathcal{N}\left(s_k\left(\frac{\tilde{d}_{\min}^2 + 4\left(m_{\min} - 1\right)\mu_{eq,in}^2 + 4\gamma\sigma^2}{2\sigma^2}\right), \left(\frac{\tilde{d}_{\min}^2 + 4\left(m_{\min} - 1\right)\mu_{eq,in}^2 + 4\gamma\sigma^2}{\sigma^2}\right)\right). \tag{16}$$

When the *a priori* LLRs are reliable ($\mu_{eq,in} > \mu_{eq,\lim}$) or $\tilde{d}_{\min}^2 \geq 4$, the BER can be approximated at high SNR by

$$P_e \simeq Q\left(\frac{\sqrt{1 + \mu_{eq,in}^2}}{\sigma}\right). \tag{17}$$

In this case, the BER of the MAP equalizer is equivalent to the BER of an AWGN channel with zero mean and variance $\sigma_2^2 = \frac{\sigma^2}{\left(1 + \mu_{eq,in}^2\right)}$ (note that the BER is independent from the channel). Hence, the *a posteriori* LLRs at the output of the MAP equalizer can be modeled as i.i.d samples from a random variable with pdf $\mathcal{N}(\frac{2s_k}{\sigma_2^2}, \frac{4}{\sigma_2^2})$.

Assuming that the *a priori* LLRs and the extrinsic LLRs are independent [11], we obtain

$$L_k^{eq} \sim \mathcal{N}\left(s_k\left(\frac{2}{\sigma_2^2} - \frac{2\mu_{eq,in}^2}{\sigma^2}\right), \left(\frac{4}{\sigma_2^2} - \frac{4\mu_{eq,in}^2}{\sigma^2}\right)\right).$$

Hence,

$$L_k^{eq} \sim \mathcal{N}\left(s_k\left(\frac{2}{\sigma^2}\right), \frac{4}{\sigma^2}\right). \tag{18}$$

By analogy with the model of *a priori* LLRs, we define $\mu_{eq,out}$ such that the extrinsic LLRs L_k^{eq} are random variables with the conditional pdf $\mathcal{N}(\frac{2s_k\mu_{eq,out}^2}{\sigma^2}, \frac{4\mu_{eq,out}^2}{\sigma^2})$. Using (16) and (18), we can write the quantity $\mu_{eq,out}^2$ as a function of $\mu_{eq,in}^2$:

$$\begin{cases} \mu_{eq,out}^2 = f_{eq}(\mu_{eq,in}^2) \triangleq \frac{\tilde{d}_{\min}^2 + 4(m_{\min}-1)\mu_{eq,in}^2 + 4\gamma\sigma^2}{4}, & \text{if } \tilde{d}_{\min}^2 < 4 \text{ and } \mu_{eq,in} < \mu_{eq,\lim} \\ \mu_{eq,out}^2 = f_{eq}(\mu_{eq,in}^2) \triangleq 1, & \text{if } \tilde{d}_{\min}^2 \geq 4 \text{ or } (\tilde{d}_{\min}^2 < 4 \text{ and } \mu_{eq,in} \geq \mu_{eq,\lim}). \end{cases} \tag{19}$$

4.4 Simulation results

We consider the same simulation conditions as in Section 4.1.1. We do not use channel coding. We consider two channels: Channel 3 and Channel 5 respectively with impulse responses $(0.499; 0.708; 0.499)$ and $(0.289; 0.499; 0.579; 0.499; 0.289)$. For Channel 3, $m_{\min} = 2$ and $\tilde{d}_{\min} = 1.5307$. By using (7), we obtain that for Channel 3, $\mu_{eq,\lim} = 0.644$. For both channels, when $\mu_{eq,in} < \mu_{eq,\lim}$, the set of error sequences achieving α_{min} is $E_{\min} = \{(2, -2), (-2, 2)\}$. Thus, $\psi_{\min} = 2.\frac{1}{4} + 2.\frac{1}{4} = 1$. We provide the equalizer with Gaussian *a priori* LLRs with the conditional pdf $\mathcal{N}(\frac{2 s_k \mu_{eq,in}^2}{\sigma^2}, \frac{4 \mu_{eq,in}^2}{\sigma^2})$, for a given $\mu_{eq,in} = \frac{\sigma}{\sigma_{eq,in}}$. Figures 6 and 7 show $\mu_{eq,out}^2$ versus $\mu_{eq,in}^2$ respectively, for Channels 3 and 5. The dotted curves are obtained by simulations for different values of the SNR= E_s/N_0. We indicate on the figure the corresponding values of BER at the output of the equalizer when no *a priori* information is available at its input. The solid curves are the analytical curves obtained using (11). Notice that the analytical curves do not depend on the SNR since $\psi_{\min} = 1$. We can see that the analytical curves approximate well the curves given by simulations as the SNR increases and for relatively high BER (BER=0.0023 for Channel 3 and BER=0.03 for Channel 5). We notice that for $\mu_{eq,in} \gg \mu_{eq,\lim}$ and $\mu_{eq,in} \ll \mu_{eq,\lim}$, the analytical curves approximate well the curves obtained by simulations. As in Section 4.1.1 and for the same reasons, around the limit value $\mu_{eq,\lim}$, the approximation is less accurate. We also notice that the approximation is better for Channel 3 which is less difficult to equalize than Channel 5.
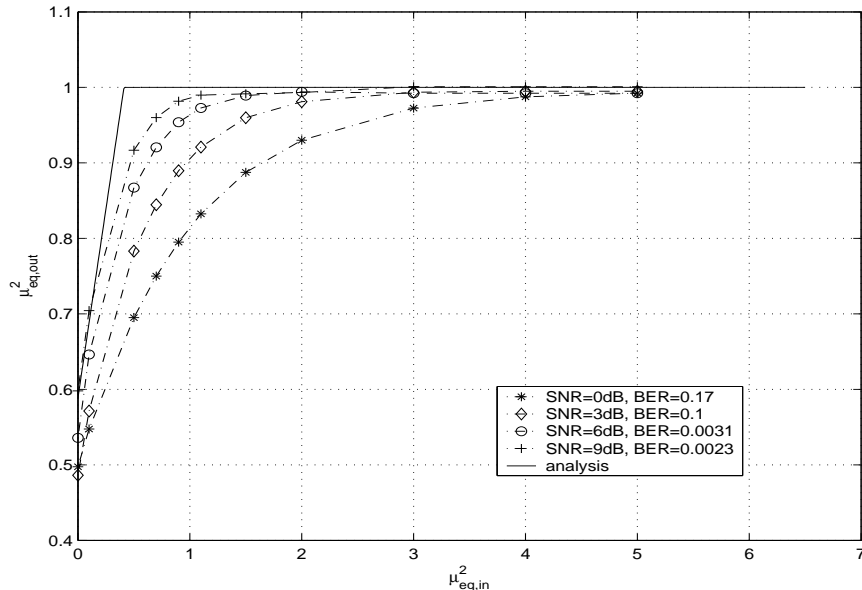

Figure 6: Equalizer analysis for Channel 3: $\mu_{eq,out}^2$ versus $\mu_{eq,in}^2$.

5 Analysis of the decoder

In order to perform the whole analysis of the turbo-detector, we have to find a closed form expression for the distribution of the extrinsic LLRs at the output of the decoder.
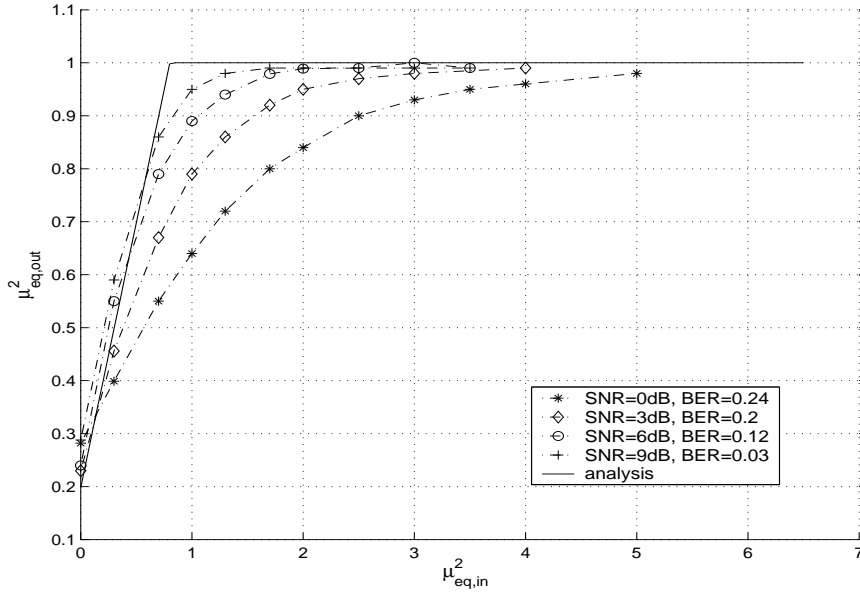
Figure 7: Equalizer analysis for Channel 5: $\mu^2_{eq,out}$ versus $\mu^2_{eq,in}$.

Proposition 8 *Assume that a priori observations at the input of the MAP decoder are available and modeled as the outputs of an AWGN channel with zero mean and variance $\sigma^2_{dec,in}$. Then, at high SNR, the extrinsic LLRs L^{dec}_k on the coded symbols s_k at the output of the decoder, are random variables with the conditional pdf $\mathcal{N}(\frac{2s_k\mu^2_{dec,out}}{\sigma^2}, \frac{4\mu^2_{dec,out}}{\sigma^2})$ where*

$$\mu^2_{dec,out} = f_{dec}(\mu^2_{dec,in}) \stackrel{\Delta}{=} \mu^2_{dec,in}(d_{free} - 1) + \beta\sigma^2, \tag{20}$$

with $\mu_{dec,in} \stackrel{\Delta}{=} \frac{\sigma}{\sigma_{dec,in}}$, d_{free} is the code minimum distance, $\beta = -2\log(m_c d_{free})$, $m_c = k^0_{d_{free}}r_c$, $k^0_{d_{free}}$ being the number of codewords of weight d_{free} caused by information sequences whose first one occurs at time 0, and r_c is the code rate.

The proof of Proposition 8 is given in Appendix III.

Simulation results

The information data are encoded using the rate $r_c = 1/2$ convolutional code having 4 states and generator polynomials $(7,5)$ in octal. For this code, $m_c = r_c = 1/2$. We provide the decoder with Gaussian *a priori* LLRs with the conditional pdf $\mathcal{N}(\frac{2s_k\mu^2_{dec,in}}{\sigma^2}, \frac{4\mu^2_{dec,in}}{\sigma^2})$, for a given $\mu_{dec,in} = \frac{\sigma}{\sigma_{dec,in}}$. Figure 8 shows $\mu^2_{dec,out}$ versus $\mu^2_{dec,in}$. The dotted curves are obtained by simulations for different values of the SNR. Here SNR=E_b/N_0 where E_b is the energy per information bit and $N_0 = 2\sigma^2$. The solid curves are the theoretical curves obtained using (20). Simulations show that the analysis of the decoder holds for high SNR values and is less accurate for low SNR values. Hence, we will use (20) in the asymptotic convergence analysis in the next section, for high

SNR values. This will allow us to prove, without relying on simulations, the convergence to the coded AWGN channel performance at high SNR. For low SNR values, we will perform a simulation of the decoder.
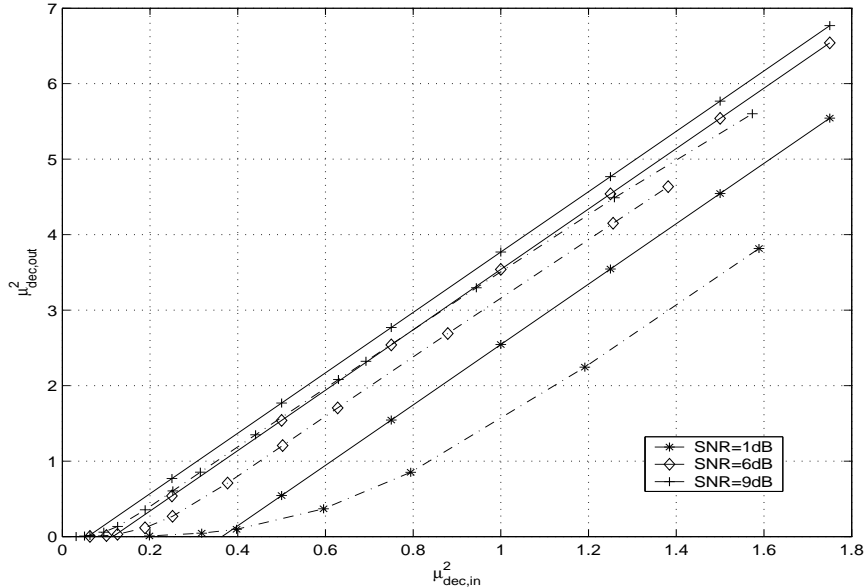

Figure 8: Decoder analysis: $\mu^2_{dec,out}$ versus $\mu^2_{dec,in}$, for the convolutional code with generator polynomials $(7, 5)$.

6 Convergence analysis

In our turbo-detector, equalization and decoding steps are iterated by passing the extrinsic LLRs l^{eq}_k and l^{dec}_k between the equalizer and the decoder. The mechanism of turbo-detection can be described by the evolution of the density of the extrinsic LLRs. Under Gaussian and independence approximations (and thanks to e-symmetry and output-symmetry preservation), the density of the extrinsic LLRs can be described by a single parameter, $\mu^2_{eq,out}$ at the output of the equalizer and $\mu^2_{dec,out}$ at the output of the decoder. Hence, the density evolution can be approximated by the changes of $\mu^2_{eq,in}$ to $\mu^2_{eq,out} = f_{eq}(\mu^2_{eq,in})$ and $\mu^2_{dec,in}$ to $\mu^2_{dec,out} = f_{dec}(\mu^2_{dec,in})$. At the first iteration, there is no *a priori* information at the input of the equalizer, thus $\mu^2_{eq,in} = 0$. Then, the extrinsic output LLRs l^{eq}_k described by $\mu^2_{eq,out} = \mu^2_{dec,in}$ are fed into the decoder yielding extrinsic LLRs l^{dec}_k described by $\mu^2_{dec,out} = \mu^2_{eq,in}$ which are fed back to the equalizer and so forth.

In this paper, our analysis differs from the conventional EXIT charts in two aspects. Firstly, conventional EXIT charts [12, 21, 22] are based on the evolution of the mutual information whereas our analysis is based on the evolution of the parameters $\mu^2_{eq,out} = \mu^2_{dec,in}$ and $\mu^2_{dec,out} = \mu^2_{eq,in}$. We choose this representation because it follows naturally from our analysis and allows us to use it easily to prove the convergence to the AWGN case. Secondly, we give a closed form expression for the evolution of these parameters at high SNR (there is no need for simulations). When the MAP equalizer is used, classical analyses generating EXIT functions [22] rely

on extensive computer simulations. To the best of our knowledge, analytical studies exist only when a linear equalizer is used [12, 15]. Notice that the study is more complicated in the case of the MAP equalizer since it is not a linear processor.

In the following, we consider two cases: the case of low SNR and the case of high SNR.

6.1 Semi analytical analysis

We consider here the case of low SNR. In this case, as shown by simulations in Section 5, the analytical expression (20) in Proposition 8 is not very accurate. Thus, $\mu^2_{dec,out}$ is obtained here by performing only one simulation while varying $\sigma^2_{dec,in}$, as performed in conventional EXIT charts [21, 22]. The analytical expression (20) will be used in the asymptotic analysis (high SNR) in Section 6.2 to prove analytically the convergence to the coded AWGN channel performance.

Simulation results

We propose to test for the validity of the convergence analysis of the turbo-detector. We consider the whole system with the channel coding at the transmitter and the turbo-detector at the receiver. In the simulations, the modulation used is BPSK. The information data are encoded using the rate $r_c = 1/2$ convolutional code with generator polynomials $(7, 5)$ or $(23, 35)$ in octal. The interleaver is randomly chosen and its size is 2048. Figures 9 and 10 represent $\mu^2_{eq,out} = \mu^2_{dec,in}$ versus $\mu^2_{dec,out} = \mu^2_{eq,in}$ at SNR=6dB respectively for Channels 3 and 5. We plot the iterative trajectory obtained by using simulations when the turbo-detector of Figure 2 is considered (we simulate a whole turbo-detector without using any artificial *a priori* LLRs). The solid curves for the equalizer (Equalizer analysis) are obtained by using the theoretical analysis (Proposition 7). The curves for the decoder (Decoder conventional method) are obtained as in conventional methods [5, 7, 21, 22] by simulations using artificial Gaussian *a priori* LLRs with the pdf $\mathcal{N}(\frac{2s_k\mu^2_{dec,in}}{\sigma^2}, \frac{4\mu^2_{dec,in}}{\sigma^2})$ at its input, when the codes used have generator polynomials $(7, 5)$ or $(23, 35)$. We notice that the convergence points are well predicted by our analysis. These points are on the curve of the coded AWGN channel corresponding to $\mu^2_{eq,out} = \mu^2_{dec,in} = 1$. As in Figures 6 and 7, we notice that the analytical curves for the equalizer give an accurate approximation of the curves obtained by simulations for $\mu_{eq,in} >> \mu_{eq,\lim}$ and $\mu_{eq,in} << \mu_{eq,\lim}$ and less accurate around the limit value $\mu_{eq,\lim}$. We also notice that the approximation is better when Channel 3 is considered. This can be explained by the fact that Channel 3 is easier to equalize than Channel 5 (the minimum distance of Channel 3 is greater than the one of Channel 5). The dashed curves for the equalizer (Equalizer conventional method) are obtained as in conventional methods by simulations using artificial Gaussian *a priori* LLRs with the pdf

$\mathcal{N}(\frac{2s_k\mu_{eq,in}^2}{\sigma^2}, \frac{4\mu_{eq,in}^2}{\sigma^2})$ at its input. These curves give a better approximation than the curves obtained by the analysis. However, they require a simulation of the equalizer for each value of the couple $\left(\sigma^2, \sigma_{eq,in}^2\right)$, whereas our analysis does not need any simulation of the equalizer.
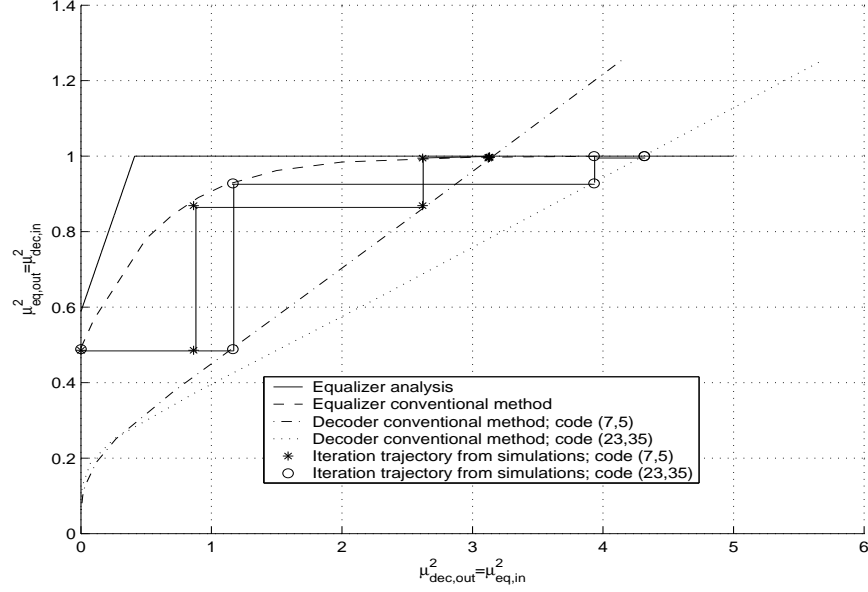

Figure 9: Convergence analysis at SNR=6dB for Channel 3: $\mu_{eq,out}^2 = \mu_{dec,in}^2$ versus $\mu_{dec,out}^2 = \mu_{eq,in}^2$.
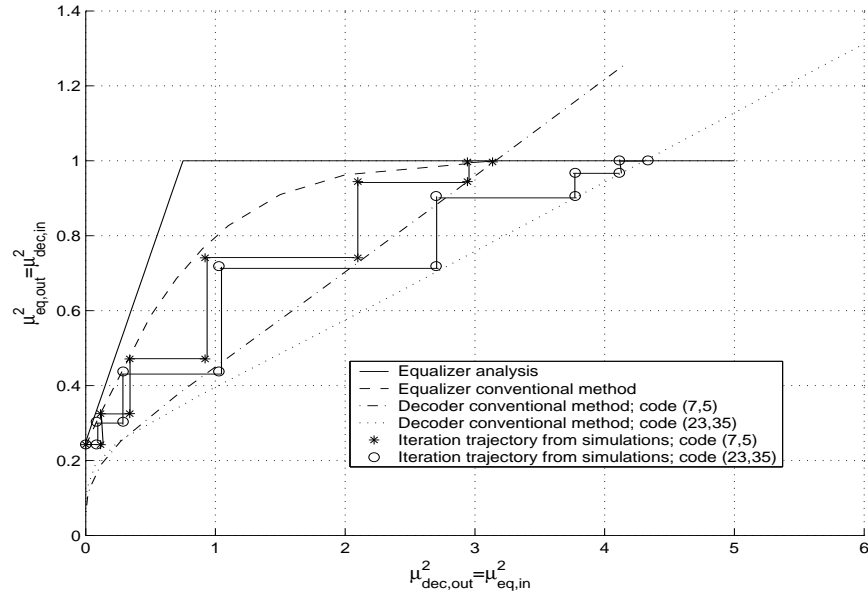

Figure 10: Convergence analysis at SNR=6dB for Channel 5: $\mu_{eq,out}^2 = \mu_{dec,in}^2$ versus $\mu_{dec,out}^2 = \mu_{eq,in}^2$.

6.2 Asymptotic analysis

Proposition 9 *For any frequency selective channel, under the Gaussian approximation, the BER performance of the MAP turbo-detector converges at high SNR to the BER performance of the coded AWGN channel.*

Proof of Proposition 9:

We consider the asymptotic case of high SNR. We use here the closed form expression (20) in Proposition 8

20

describing the extrinsic LLRs at the output of the decoder, which becomes accurate. We want to prove that the BER performance of the turbo-detector converges to the BER performance of the coded AWGN channel. The asymptotic convergence point of the turbo-detector is represented by the fixed point of $f_{eq} \circ f_{dec}$. To have the convergence of the turbo-detector performance to the AWGN channel BER performance, the fixed point needs to be in the region of reliable *a priori* information at the equalizer input such that (18) is valid. This is always the case when $\tilde{d}^2_{\min} \geq 4$. When $\tilde{d}^2_{\min} < 4$, this means that the value of $\mu^2_{dec,out}$ in (20) such that $\mu^2_{dec,in} = 1$ must be greater than $\mu^2_{eq,\lim} = \frac{4-\tilde{d}^2_{\min}}{4(m_{\min}-1)}$. The condition can be rewritten as:

$$\mu^2_{dec,out} = \mu^2_{dec,in}(d_{free} - 1) + \beta\sigma^2|_{\mu^2_{dec,in}=1} = d_{free} - 1 + \beta\sigma^2 \geq \frac{4 - \tilde{d}^2_{\min}}{4(m_{\min} - 1)}.$$

This leads to

$$d_{free} \geq 1 + \frac{1}{(m_{\min} - 1)} - \beta\sigma^2 - \frac{\tilde{d}^2_{min}}{4(m_{\min} - 1)}. \tag{21}$$

Since σ is low (high SNR), we can neglect the term $\beta\sigma^2$ and we obtain:

$$d_{free} \geq 1 + \frac{1}{(m_{\min} - 1)} - \frac{\tilde{d}^2_{min}}{4(m_{\min} - 1)}. \tag{22}$$

Since $m_{\min} \geq 2$, we have $1 + \frac{1}{(m_{\min}-1)} \leq 2$. Since d_{free} is greater than 2 for a convolutional code, except the identity code, the condition (22) is always verified.

In the following, we validate by simulation this asymptotic analysis by comparing it with the convergence of a turbo-detector for finite block length.

Simulation results

We use here the same simulation conditions as in Section 6.1. The SNR is now equal to 9dB (high SNR). Figures 11 and 12 represent $\mu^2_{eq,out} = \mu^2_{dec,in}$ versus $\mu^2_{dec,out} = \mu^2_{eq,in}$ respectively for Channels 3 and 5. We notice that the analysis of the equalizer becomes accurate here. As in Section 6.1, the curves of the decoder (Decoder conventional method) obtained by using artificial Gaussian *a priori* LLRs give an accurate approximation of the decoder behavior. We also plot in Figures 11 and 12 the curves obtained by using (20) in Proposition 8 (Decoder analysis) which predict well the decoder performance since the SNR is high. Thus, the behavior of the turbo-detector is accurately predicted by our analysis at high SNR. The simulations show that the performance of the turbo-detector converges to the BER performance of the coded AWGN channel as proved by the analysis.

To confirm this result, we plot in Figure 13 the BER performance curves on the information bits at the output of the turbo-detector for one to four iterations, for Channel 3 and for one to six iterations, for Channel

5. The convolutional code has generator polynomials $(7,5)$. Figure 13 also shows the BER performance of the coded AWGN channel. We notice that for moderate to high SNR (SNR$\geq 4dB$ for Channel 3 and SNR$\geq 6dB$ for Channel 5), the BER performance at the output of the turbo-detector reaches the BER performance obtained for the coded AWGN channel as predicted by the analysis. For lower SNR values, the BER performance of the turbo-detector does not converge to the BER performance of the coded AWGN channel. In this case, our analysis does not predict the convergence point of the turbo-detector since the expressions of the extrinsic LLRs distributions we derived are not close enough in this case (as explained in Sections 4 and 5) and the Gaussian approximation is not very accurate.
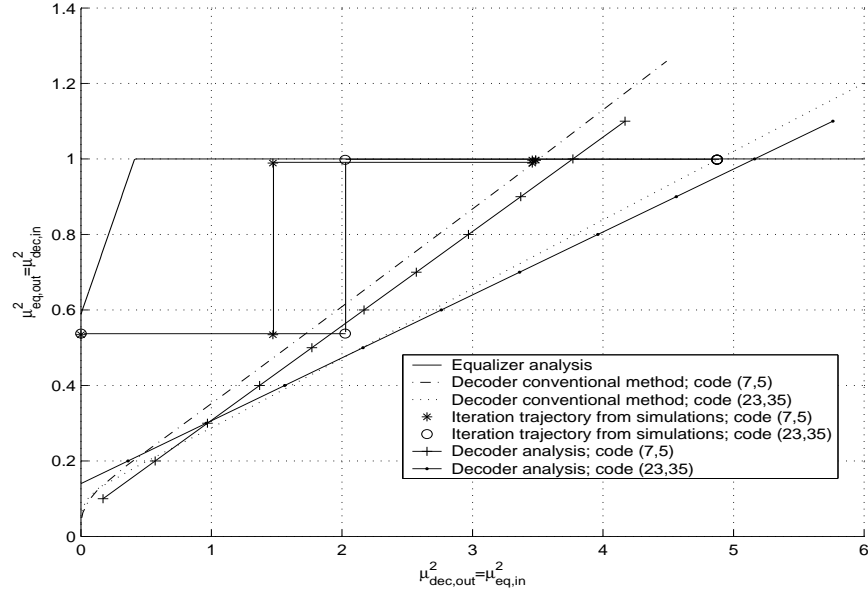

Figure 11: Convergence analysis at SNR=9dB, $\mu_{eq,out}^2 = \mu_{dec,in}^2$ versus $\mu_{dec,out}^2 = \mu_{eq,in}^2$, for Channel 3
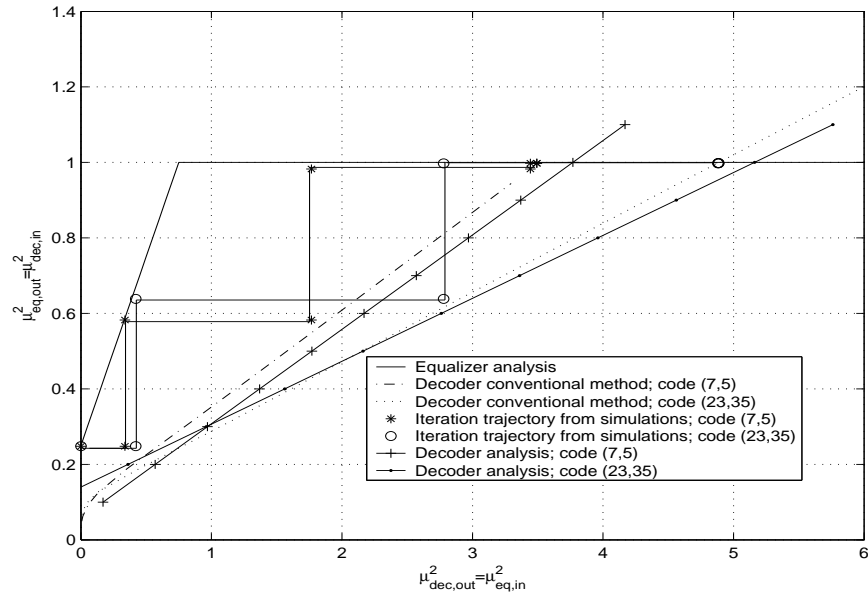

Figure 12: Convergence analysis at SNR=9dB, $\mu_{eq,out}^2 = \mu_{dec,in}^2$ versus $\mu_{dec,out}^2 = \mu_{eq,in}^2$, for Channel 5
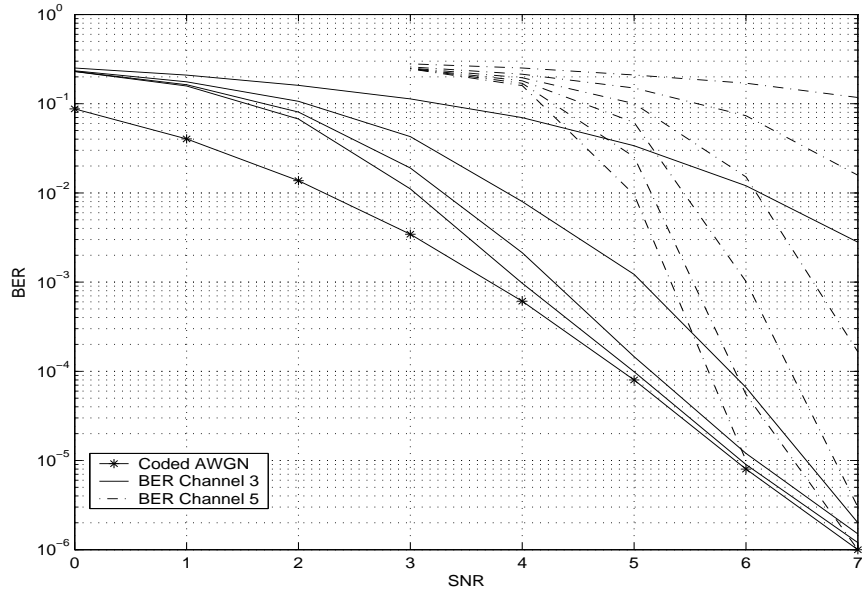
22

Figure 13: BER performance of the turbo-detector for Channel 3 (solid curves) and Channel 5 (dotted curves).

6.3 Validity of the Gaussian approximation

In [10], it was shown that for a convolutional code with an infinite block size, the density of the extrinsic LLRs at the decoder output can be well approximated by a Gaussian density. In Section 4, we proved the convergence of the distribution of the extrinsic LLRs at the output of the MAP equalizer to a Gaussian distribution at high SNR and for infinite block size. We want now to test for the validity of the Gaussian approximation by simulation for a turbo-detector with finite block length. Figures 14 and 15 show the Kullback Leibler distance [4] of the distribution of the extrinsic LLRs, when the transmitted symbol is equal to $+1$, to the Gaussian distribution with the same mean and variance, respectively at the output of the equalizer and the decoder versus the iteration number of the iterative receiver for SNR$\in \{3, 6, 9\}$. We consider here Channels 3 and 5. We use the convolutional code with generator polynomials $(7, 5)$. The solid curves are obtained when Channel 3 is used and the dotted curves are obtained when Channel 5 is used. Figures 14 and 15 show that as the number of iterations and the SNR increase, the Kullback Leibler distance decreases. Thus, the pdf of the extrinsic LLRs approaches a Gaussian pdf when the SNR is moderately high and as the number of iterations increases (the *a priori* information becomes reliable). We notice that on average the values of the Kullback Leibler distance obtained for Channel 5 are higher than those obtained for Channel 3. Thus, the Gaussian approximation is better verified for Channel 3. This can be explained by the fact that Channel 3 is easier to equalize than Channel 5. We also notice that the distribution of the extrinsic LLRs at the output of the decoder is less close to a Gaussian distribution than the distribution of the extrinsic LLRs at the output of the equalizer (the Kullback Leibler distance is higher).
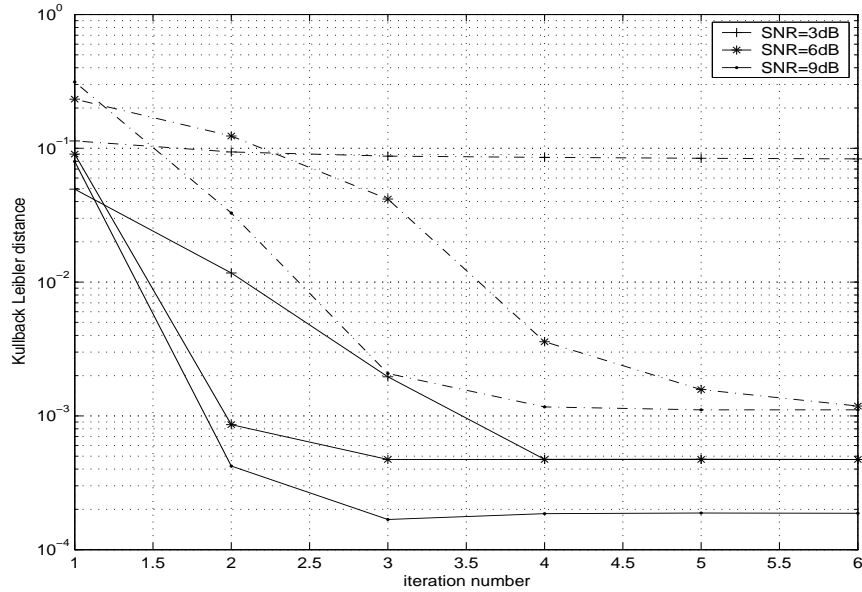
Figure 14: Kullback Leibler distance of the distribution of the extrinsic LLRs at the output of the equalizer when the transmitted symbol is equal to $+1$ to the Gaussian distribution with the same mean and variance, versus the iteration number of the iterative receiver for SNR$\in \{3, 6, 9\}$, for Channel 3 (solid curves) and for Channel 5 (dotted curves).

7 Conclusion

In this paper, we considered a coded transmission over a frequency selective channel. We proposed to study analytically the convergence of the turbo-detector using a MAP equalizer and a MAP decoder. We showed that as for MAP decoding [17], e-symmetry and output-symmetry are preserved under MAP equalization. Using the i.i.d Gaussian approximation, we showed that the BER performance of the turbo-detector converges to the BER performance of the coded AWGN channel at high SNR. Simulations showed that our analysis allows one to accurately predict the turbo-detector behavior at high SNR and the performance of the convergence stationary points.

Appendix I: Proof of Proposition 3

As shown in Figure 3, we consider the transmission of a sequence $\underline{s} = (s_{T-1}, ..., s_{1-L})^T$ of BPSK symbols over an output-symmetric channel, Channel 1. The output sequence is called $\underline{x}$. We assume that a set of *a priori* LLRs with an output-symmetric density is also available. These *a priori* LLRs are associated with an output-symmetric channel, Channel 2, with input $\underline{s}$ [17] (see Figure 3). Let $\underline{r} = (r_{T-1}, ..., r_{1-L})^T$ be the sequence of *a priori* observations at the output of Channel 2. Let $\underline{y} = (\underline{x}^T, \underline{r}^T)^T$. Since Channel 1 and Channel
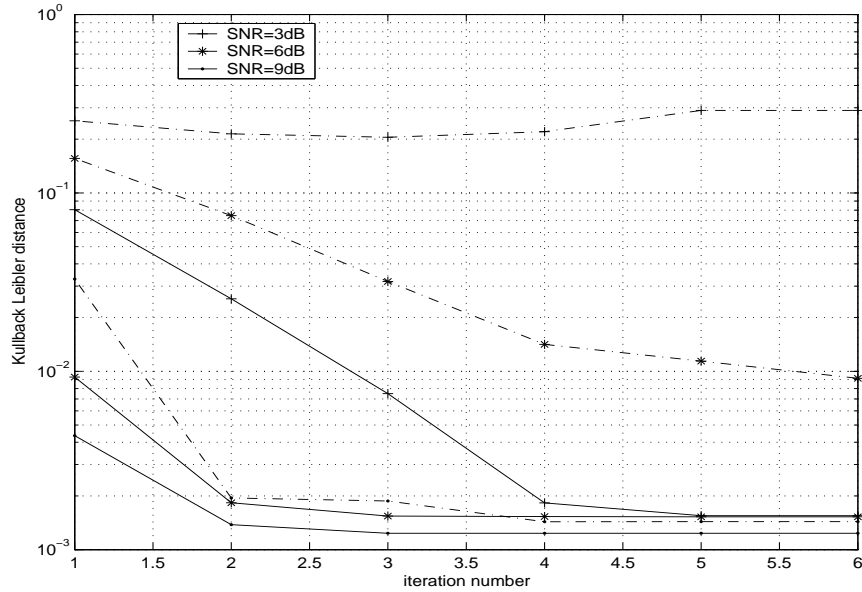
Figure 15: Kullback Leibler distance of the distribution of the extrinsic LLRs at the output of the decoder when the transmitted symbol is equal to $+1$ to the Gaussian distribution with the same mean and variance, versus the iteration number of the iterative receiver for SNR$\in \{3, 6, 9\}$, for Channel 3 (solid curves) and for Channel 5 (dotted curves).

2 are output-symmetric, $p_{\underline{Y}|\underline{S}}\left(\underline{y}|\underline{s}\right) = p_{\underline{Y}|\underline{S}}\left(-\underline{y}|-\underline{s}\right)$.

The received sequence $\underline{y}$ is then fed to the MAP equalizer (see Figure 3). The MAP equalizer computes the *a posteriori* probabilities on the transmitted symbols s_i, $P_{S_i|\underline{Y}}(s|\underline{y})$, for $s \in \{+1, -1\}$. Using Bayes rule, we obtain

$$P_{S_i|\underline{Y}}(s|\underline{y}) = \frac{p_{\underline{Y}|S_i}(\underline{y}|s)P_{S_i}(s)}{p_{\underline{Y}}(\underline{y})}.$$

Thus,

$$\begin{aligned}
\log\left(\frac{P_{S_i|\underline{Y}}(+1|\underline{y})}{P_{S_i|\underline{Y}}(-1|\underline{y})}\right) &= \log\left(\frac{p_{\underline{Y}|S_i}(\underline{y}|+1)}{p_{\underline{Y}|S_i}(\underline{y}|-1)}\right) + \log\left(\frac{P_{S_i}(+1)}{P_{S_i}(-1)}\right) \\
&= \log\left(\frac{p_{\underline{Y}|S_i}(\underline{y}|+1)}{p_{\underline{Y}|S_i}(\underline{y}|-1)}\right)
\end{aligned}$$

since $P_{S_i}(+1) = P_{S_i}(-1) = 1/2$.

Let $p_{L_i|S_i}(z|s)$ be the density of the *a posteriori* LLRs $L_i = \log\left(\frac{p_{\underline{Y}|S_i}(\underline{Y}|+1)}{p_{\underline{Y}|S_i}(\underline{Y}|-1)}\right)$ conditioned on $S_i = s \in \{+1, -1\}$. We want to show that the density of these LLRs is e-symmetric and output-symmetric, i.e., $p_{L_i|S_i}(z|s) = p_{L_i|S_i}(-z|s)\exp(sz)$ and $p_{L_i|S_i}(z|s) = p_{L_i|S_i}(-z|-s)$.

First, notice that

$$
\begin{aligned}
\frac{p_{\underline{Y}|S_i}(\underline{y}|s)}{p_{\underline{Y}|S_i}(\underline{y}|-s)} &= \frac{\sum_{\underline{s}:s_i=s} p_{\underline{Y}|\underline{S}}(\underline{y}|\underline{s})}{\sum_{\underline{s}:s_i=-s} p_{\underline{Y}|\underline{S}}(\underline{y}|\underline{s})} \\
&= \frac{\sum_{\underline{s}:s_i=s} p_{\underline{Y}|\underline{S}}(-\underline{y}|-\underline{s})}{\sum_{\underline{s}:s_i=-s} p_{\underline{Y}|\underline{S}}(-\underline{y}|-\underline{s})} \\
&= \frac{\sum_{\underline{s}:s_i=-s} p_{\underline{Y}|\underline{S}}(-\underline{y}|\underline{s})}{\sum_{\underline{s}:s_i=s} p_{\underline{Y}|\underline{S}}(-\underline{y}|\underline{s})} \\
&= \frac{p_{\underline{Y}|S_i}(-\underline{y}|-s)}{p_{\underline{Y}|S_i}(-\underline{y}|s)}.
\end{aligned}
\tag{23}
$$

Define N as the length of $\underline{y}$ and $l_i^{-1}(z)$ as the set of all $\underline{y} \in \mathcal{R}^N$ such that $\log\left(\frac{p_{\underline{Y}|S_i}(\underline{y}|+1)}{p_{\underline{Y}|S_i}(\underline{y}|-1)}\right) = z$. From (23), we have $l_i\left(\underline{y}\right) = \log\left(\frac{p_{\underline{Y}|S_i}(\underline{y}|+1)}{p_{\underline{Y}|S_i}(\underline{y}|-1)}\right) = \log\left(\frac{p_{\underline{Y}|S_i}(-\underline{y}|-1)}{p_{\underline{Y}|S_i}(-\underline{y}|+1)}\right) = -l_i\left(-\underline{y}\right)$, thus

$$
\underline{y} \in l_i^{-1}(z) \Leftrightarrow -\underline{y} \in l_i^{-1}(-z).
$$

Let $\underline{y} = (\underline{x}^T, \underline{r}^T)^T = (y_0, y_1, \cdots, y_{N-1})^T$. In order to express the density $p_{L_i|S_i}(z|s)$, we want to perform the change of variable $\underline{y}$ towards z. Thus, we introduce $\underline{g} = (y_0, y_1, \cdots, y_{N-2}, z)^T = (g_0, g_1, \cdots, g_{N-1})^T$. Let $\underline{d}$ be the $N \times N$ matrix with the element on the kth row and j^{th} column $d_{k,j} = \frac{\partial g_k}{\partial y_j}$, for $0 \le k \le N-1$ and $0 \le j \le N-1$, defined as

$$
\begin{cases}
d_{k,j} = 1, \text{ for } 0 \le k \le N-2, \ 0 \le j \le N-2 \text{ and } k = j \\[2mm]
d_{k,j} = 0, \text{ for } 0 \le k \le N-2, \ 0 \le j \le N-1 \text{ and } k \ne j \\[2mm]
d_{k,j} = \frac{\partial z}{\partial y_j} = \frac{\left(\frac{\partial p_{\underline{Y}|S_i}(\underline{y}|+1)}{\partial y_j}\right)}{p_{\underline{Y}|S_i}(\underline{y}|+1)} - \frac{\left(\frac{\partial p_{\underline{Y}|S_i}(\underline{y}|-1)}{\partial y_j}\right)}{p_{\underline{Y}|S_i}(\underline{y}|-1)}, \text{ for } k = N-1 \text{ and } 0 \le j \le N-1
\end{cases}
\tag{24}
$$

Let $\alpha\left(\underline{y}\right)$ be the inverse of the determinant of the matrix $\underline{d}$. Since $l_i\left(\underline{y}\right) = -l_i\left(-\underline{y}\right)$, we obtain, for all $0 \le j \le N-1$, $\frac{\partial l_i(\underline{y})}{\partial y_j} = -\frac{\partial l_i(-\underline{y})}{\partial y_j}$ and then $\alpha\left(\underline{y}\right) = -\alpha\left(-\underline{y}\right)$.

Using change of variables, we can then write,

$$
\begin{aligned}
\frac{p_{L_i|S_i}(z|s)}{p_{L_i|S_i}(-z|s)} &= \frac{\int_{y_i, i \ne N-1} \sum_{\underline{y} \in l_i^{-1}(z)} p_{\underline{Y}|S_i}(\underline{y}|s) \left|\alpha\left(\underline{y}\right)\right| dy_0 dy_1 \cdots dy_{N-2}}{\int_{y_i, i \ne N-1} \sum_{\underline{y} \in l_i^{-1}(-z)} p_{\underline{Y}|S_i}(\underline{y}|s) \left|\alpha\left(\underline{y}\right)\right| dy_0 dy_1 \cdots dy_{N-2}} \\
&= \frac{\int_{y_i, i \ne N-1} \sum_{\underline{y} \in l_i^{-1}(z)} p_{\underline{Y}|S_i}(\underline{y}|s) \left|\alpha\left(\underline{y}\right)\right| dy_0 dy_1 \cdots dy_{N-2}}{\int_{y_i, i \ne N-1} \sum_{\underline{y} \in l_i^{-1}(z)} p_{\underline{Y}|S_i}(-\underline{y}|s) \left|\alpha\left(-\underline{y}\right)\right| dy_0 dy_1 \cdots dy_{N-2}}, \text{ since } l_i^{-1}(z) = -l_i^{-1}(-z) \\
&= \frac{\int_{y_i, i \ne N-1} \sum_{\underline{y} \in l_i^{-1}(z)} p_{\underline{Y}|S_i}(\underline{y}|s) \left|\alpha\left(\underline{y}\right)\right| dy_0 dy_1 \cdots dy_{N-2}}{\int_{y_i, i \ne N-1} \sum_{\underline{y} \in l_i^{-1}(z)} p_{\underline{Y}|S_i}(-\underline{y}|s) \left|\alpha\left(\underline{y}\right)\right| dy_0 dy_1 \cdots dy_{N-2}}, \text{ since } \alpha\left(\underline{y}\right) = -\alpha\left(-\underline{y}\right) \\
&= \exp\left(sz\right)
\end{aligned}
$$

where the last step follows from $\frac{p_{\underline{Y}|S_i}(\underline{y}|s)}{p_{\underline{Y}|S_i}(-\underline{y}|s)} = \frac{p_{\underline{Y}|S_i}(\underline{y}|s)}{p_{\underline{Y}|S_i}(\underline{y}|-s)} = \exp(sz)$ for all $\underline{y} \in l_i^{-1}(z)$. Thus, the density of the a

posteriori LLRs is e-symmetric. We can also write,

$$
\begin{aligned}
p_{L_i|S_i}(-z|-s) &= \int_{y_i, i \neq N-1} \sum_{\underline{y} \in l_i^{-1}(-z)} p_{\underline{Y}|S_i}(\underline{y}|-s) \left| \alpha\left(\underline{y}\right) \right| dy_0 dy_1 \cdots dy_{N-2} \\
&= \int_{y_i, i \neq N-1} \sum_{\underline{y} \in l_i^{-1}(z)} p_{\underline{Y}|S_i}(-\underline{y}|-s) \left| \alpha\left(-\underline{y}\right) \right| dy_0 dy_1 \cdots dy_{N-2} \\
&= \int_{y_i, i \neq N-1} \sum_{\underline{y} \in l_i^{-1}(z)} p_{\underline{Y}|S_i}(\underline{y}|s) \left| \alpha\left(\underline{y}\right) \right| dy_0 dy_1 \cdots dy_{N-2} \\
&= p_{L_i|S_i}(z|s).
\end{aligned}
$$

Thus, the density is also output-symmetric.

Appendix II: Proof of Corollary 4

The extrinsic LLR on the transmitted symbol s_i is given by

$$
\begin{aligned}
l_i^{ext} &= \log\left(\frac{p_{\underline{Y}|S_i}(\underline{y}|+1)}{p_{\underline{Y}|S_i}(\underline{y}|-1)} \right) - \log\left(\frac{p_{R_i|S_i}(r_i|+1)}{p_{R_i|S_i}(r_i|-1)} \right) \\
&= \log\left(\frac{\sum_{\underline{s}:s_i=1} p_{\underline{Y}|\underline{S}}(\underline{y}|\underline{s})}{\sum_{\underline{s}:s_i=-1} p_{\underline{Y}|\underline{S}}(\underline{y}|\underline{s})} \right) - \log\left(\frac{p_{R_i|S_i}(r_i|+1)}{p_{R_i|S_i}(r_i|-1)} \right).
\end{aligned}
$$

Since Channel 1 and Channel 2 are independent and Channel 2 is memoryless, we obtain

$$
\begin{aligned}
l_i^{ext} &= \log\left(\frac{\sum_{\underline{s}:s_i=1} p_{\underline{X}|\underline{S}}(\underline{x}|\underline{s}) \left(\frac{\prod_{j=1-L}^{T-1} p_{R_j|S_j}(r_j|s_j)}{p_{R_i|S_i}(r_i|+1)} \right)}{\sum_{\underline{s}:s_i=-1} p_{\underline{X}|\underline{S}}(\underline{x}|\underline{s}) \left(\frac{\prod_{j=1-L}^{T-1} p_{R_j|S_j}(r_j|s_j)}{p_{R_i|S_i}(r_i|-1)} \right)} \right) \\
&= \log\left(\frac{\sum_{\underline{s}:s_i=1} p_{\tilde{\underline{Y}}|\underline{S}}(\tilde{\underline{y}}|\underline{s})}{\sum_{\underline{s}:s_i=-1} p_{\tilde{\underline{Y}}|\underline{S}}(\tilde{\underline{y}}|\underline{s})} \right), \quad \text{where } \tilde{\underline{y}} = \underline{y}\backslash\{r_i\} \text{ and } \tilde{\underline{Y}} = \underline{Y}\backslash\{R_i\} \\
&= \log\left(\frac{p_{\tilde{\underline{Y}}|S_i}(\tilde{\underline{y}}|+1)}{p_{\tilde{\underline{Y}}|S_i}(\tilde{\underline{y}}|-1)} \right),
\end{aligned}
$$

$\tilde{\underline{y}} = \underline{y}\backslash\{r_i\}$ denotes $\underline{y}$ deprived of r_i.

In order to show that the density of the extrinsic LLRs $L_i^{ext} = \log\left(\frac{p_{\underline{Y}|S_i}(\underline{Y}|+1)}{p_{\underline{Y}|S_i}(\underline{Y}|-1)} \right) - \log\left(\frac{p_{R_i|S_i}(R_i|+1)}{p_{R_i|S_i}(R_i|-1)} \right)$ is

e-symmetric and output-symmetric, we can use the same proof as in Appendix I, by replacing $\underline{y}$ by $\tilde{\underline{y}}$ and $\underline{Y}$ by

$\tilde{\underline{Y}}$.

Appendix III: Proof of Proposition 8

Let $\underline{c} = (c_0, \cdots, c_{T-1})^T$ be the sequence of coded bits that are transmitted. The BER on the coded bits can

be upper bounded using a union bound technique [2, 14] by

$$
P_e \leq \sum_{d=d_{free}}^{\infty} \frac{k_d d}{T} Q\left(\frac{\sqrt{d}}{\sigma_{dec,in}} \right)
$$

where k_d is the number of codewords of size T and weight d and d_{free} is the code minimum distance.

At high SNR, the BER can be approximated by the term involving d_{free} as

$$
\begin{aligned}
P_e &\simeq \frac{k_{d_{free}} d_{free}}{T} Q\left(\frac{\sqrt{d_{free}}}{\sigma_{dec,in}}\right) \\
&= m_c d_{free} Q\left(\frac{\sqrt{d_{free}}}{\sigma_{dec,in}}\right),
\end{aligned}
$$

where $k_{d_{free}}$ is the number of codewords of weight d_{free} and $m_c = \frac{k_{d_{free}}}{T}$.

When T is large enough, as is the case in this paper, if the convolutional code has $k^0_{d_{free}}$ codewords of weight d_{free} caused by information sequences whose first one occurs at time 0, then $\lim_{T \to +\infty} \frac{k_{d_{free}}}{T} = k^0_{d_{free}} r_c$ [14].

Since $Q\left(\sqrt{z}\right) e^{-\frac{y}{2}} \geq Q\left(\sqrt{z+y}\right)$ [2], we obtain

$$
P_e >\sim Q\left(\sqrt{\frac{d_{free}}{\sigma^2_{dec,in}} + \beta}\right) \tag{25}
$$

where $\beta = -2\log\left(m_c d_{free}\right)$.

This bound is tight at high SNR and the overall probability of error can then be approximated by

$$
P_e \simeq Q\left(\sqrt{\frac{d_{free}\mu^2_{dec,in}}{\sigma^2} + \beta}\right). \tag{26}
$$

This is equivalent to the BER performance of an AWGN channel with a noise having a zero mean and variance $\sigma^2_3 \triangleq \frac{\sigma^2}{d_{free}\mu^2_{dec,in} + \beta\sigma^2}$. Hence, the *a posteriori* LLRs at the output of the decoder can be modeled as i.i.d samples from a random variable with pdf $\mathcal{N}(\frac{2s_k}{\sigma^2_3}, \frac{4}{\sigma^2_3})$.

Assuming that the *a priori* and extrinsic LLRs are independent [11], we obtain

$$
L^{dec}_k \sim \mathcal{N}\left(s_k\left(\frac{2}{\sigma^2_3} - \frac{2\mu^2_{dec,in}}{\sigma^2}\right), \left(\frac{4}{\sigma^2_3} - \frac{4\mu^2_{dec,in}}{\sigma^2}\right)\right).
$$

Hence

$$
L^{dec}_k \sim \mathcal{N}\left(s_k\left(\frac{2\left(\mu^2_{dec,in}(d_{free}-1) + \beta\sigma^2\right)}{\sigma^2}\right), \frac{4\left(\mu^2_{dec,in}(d_{free}-1) + \beta\sigma^2\right)}{\sigma^2}\right). \tag{27}
$$

As for the equalizer, we can define $\mu^2_{dec,out}$ such that the extrinsic LLRs L^{dec}_k are random variables with the conditional pdf $\mathcal{N}(\frac{2s_k\mu^2_{dec,out}}{\sigma^2}, \frac{4\mu^2_{dec,out}}{\sigma^2})$. The quantity $\mu^2_{dec,out}$ is a function of $\mu^2_{dec,in}$ and is defined as

$$
\mu^2_{dec,out} = f_{dec}(\mu^2_{dec,in}) \triangleq \mu^2_{dec,in}(d_{free}-1) + \beta\sigma^2. \tag{28}
$$

References

[1] 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-32, pp.284-287, March 1974.

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

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

[4] T.M.Cover and J.A.Thomas, *Elements of information theory,* NewYork: Wiley, 1991.

[5] D.Divsalar, S.Dolinar, and F.Pollara, "Iterative turbo decoder analysis based on density evolution," *IEEE J. Sel. Areas Comm. ,* vol. 19, no. 5, pp. 891-907, May 2001.

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

[7] H.El Gamal and A.R.Hammons, "Analyzing the turbo decoder using the Gaussian approximation," *IEEE Trans. on Inf. Theory*, vol.47, no.2, pp.671-686, Feb.2001.

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

[9] G.D.Forney, Jr., "Lower bounds on error probability in the presence of large intersymbol interference," *IEEE Trans. on Comm.*, pp. 76-77 , Feb. 1972.

[10] M.Fu, "Stochastic analysis of turbo decoding," *IEEE Trans. on Inf. Theory*, vol. 51, no. 1, pp. 81-100, January 2005.

[11] J.Hokfelt, O.Edfors, and T.Maseng, "Interleaver design for turbo codes based on the performance of iterative decoding," *IEEE Int. Conf. on Comm.*, Vancouver, BC, Canada, June 1999.

[12] S.Lee, A.Singer, and N.R.Shanbhag, "Linear turbo equalization analysis via BER transfer and EXIT charts," *IEEE Trans. on Signal Processing*, vol. 53, no. 8, pp. 2883-2897, August 2005.

[13] F.Lehmann and G.M.Maggio "Nonlinear analysis of the iterative decoding of parallel concatenated convolutional codes," *IEEE Trans. on Inf.Theory*, vol. 51, no. 6, pp. 2190-2198, June 2005.

[14] L.C.Perez, J.Seghers, and D.J.Costello, Jr., "A distance spectrum interpretation of turbo codes," *IEEE Trans. on Inf. Theory*, vol. 42, no. 6, pp. 1698-1709, November 1996.

[15] V.Ramon, C.Herzet, and L.Vandendorpe, "A semi-analytical method for predicting the performance and convergence behavior of a multi-user turbo-equalizer/demapper," *IEEE Trans. on Signal Processing*, vol. 55, no. 3, pp. 1104-1117, March 2007.

[16] T.J.Richardson, M.A.Shokrollahi, R.L.Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. on Inf. Theory*, vol.47, no.2, pp.619-637, Feb.2001.

[17] T.J.Richardson and R. Urbanke, "Modern Coding Theory" (draft of book), http://lthcwww.epfl.ch/mct/

[18] A.Roumy, S.Guemghar, G.Caire, and S.Verdu, "Design methods for irregular repeat accumulate codes," *IEEE Trans. on Inf. Theory*, pp. 1711-1727, August 2004.

[19] N.Sellami, A.Roumy, and I.Fijalkow, "The impact of both a priori information and channel estimation errors on the MAP equalizer performance, " *IEEE Trans. on Signal Processing*, vol. 54, no. 7, pp. 2716-2724, July 2006.

[20] 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.

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

[22] 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.

[23] J.M.Walsh, C.R.Johnson, Jr., and P.A.Regalia, "A refined information geometric interpretation of turbo decoding," *IEEE ICASSP'05,* vol. III, pp. 481-484, Philadelphia, PA, March 2005.

[24] J.M.Walsh, P.A.Regalia, and C.R.Johnson, Jr., "Belief propagation is iterative constrained Maximum Likelihood sequence detection," submitted to *IEEE Trans. on Inf. Theory*, Sept.2005.

[25] N.Wiberg, "Codes and decoding on general graphs," Ph.D. dissertation, Dept. Electircal Engineering, Linkoping University, Linkoping, Sweden, 1996.