DEPENDENCY BETWEEN ERROR VARIANCE OF THE A PRIORI INFORMATION AND A MODIFIED CHANNEL NOISE VARIANCE IN TURBO-EQUALISATION

Gilles Gorlier, Sylvie Perreau

Institute for Telecommunications Research
University of South Australia
Mawson Lakes SA 5095 Australia
Email: Gilles.Gorlier@postgrads.unisa.edu.au
Email: Sylvie.Perreau@unisa.edu.au

Inbar Fijalkow

Equipes Traitement des Images et du Signal
ENSEA - University of Cergy-Pontoise - CNRS
95014 Cergy France
Email: Inbar.Fijalkow@ensea.fr

ABSTRACT

In this paper, we propose a method based on the modification of the channel noise variance at the equaliser input in order to improve the performance of a turbo-equaliser. We will show a relation between the modified channel noise variance and the a priori information statistics. The simulation are done for 2 types of turbo-equalisers.

Keywords: EXIT chart, turbo-equalisation, APP equalisation, interference canceler, SNR estimation.

1. INTRODUCTION

In turbo-equalisation, a decoder is combined with an equaliser to combat intersymbol interference (ISI). The equaliser collects information from the decoder output through the interleaver. They are processed together with the channel observations. That way, the decoder output provides the equaliser with a priori knowledge about the message. As a consequence the symbol estimation improves. The turbo-equaliser is an iterative process introduced in 1995 [1], inspired by the turbo-decoder [2].

Three classes of turbo-equalisers have been deeply investigated. They make use of one of the following 3 types of soft input soft output (SISO) equaliser: a posteriori probability (APP) [3], interference canceler (IC) [4], or reduced complexity probability based equalisers; in association with a SISO APP decoder. The problem with the APP equaliser, although it provides the best results, is its very high calculation complexity. Often one prefers using a low complexity alternative algorithm. The IC is a low complexity, filter based equaliser which shows excellent results at high signal to noise ratio (SNR). It uses the same principle as the decision feedback equaliser (DFE): the symbols estimated from previous iteration are fed back through a filter in order to cancel the interferences. It makes this type of equaliser particularly sensitive to error propagation. Finally, the reduced complexity probability based equalisers can be viewed as a good trade off between complexity and performance. One of them is the DFE incorporating fixed lag smoothing [5]. It uses both the estimated symbols throughout its memory, the channel observations and the a priori information to estimate a symbol. The estimated symbols are updated along the channel memory and the complexity is not as high as the APP equaliser.

In turbo-decoding, it was proved that the knowledge of the exact SNR is not mandatory to achieve sufficiently good performances. For instance in [6] the authors showed that an SNR overestimation does not degrade much the performances of the system in terms of bit error rate (BER). For their experiments, they used the log implementation of the APP algorithm and tested an AWGN channel as well as a Rayleigh fading channel.

In this work, we use a modified channel noise variance at the equaliser input in order to improve the turbo-equaliser performances. The turbo-equaliser is composed of a DFE incorporating fixed lag smoothing and an APP decoder.

This article is structured as follows: Section 2 describes the turbo-equaliser using a DFE incorporating fixed lag smoothing. Next section highlights the method used to chose the modified channel noise variance according to the a priori information. Section 4 shows the effect of the error variance of the feedback loop on the channel noise variance choice.

2. TURBO-EQUALISER

2.1. Transmission model

An independently and identically distributed (iid) sequence of non coded bits is passed through the channel encoder to produce the sequence c_n which is interleaved and mapped using a BPSK modulation. We assume that the coded and modulated symbols x_t are transmitted through a finite impulse response (FIR) channel with transfer function $H(z) = \sum_{i=0}^{N-1} h_i z^{-1}$ where the h_i are complex-valued coefficients. We assume additive Gaussian white noise with zero mean and variance σ_w^2. The received signal is thus modeled by

$$y_t = H^T X_t + w_t \tag{1}$$

$H^T = [h_0 \ h_1 \ \dots \ h_{N-1}]$ and $X_t = [x_t \ x_{t-1} \ \dots \ x_{t-N+1}]^T$. Here, the operator $()^T$ denotes the transposition operation. For sake of simplicity, we also assume throughout this paper a BPSK modulation.

2.2. Receiver structure

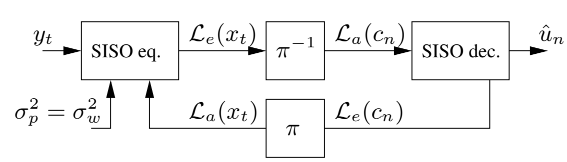


Fig. 1. Receiver structure: turbo-equaliser

Fig. 1 depicts the general structure for turbo equalisation. We will see that the equaliser we envisage in this paper produces at time $t+N-1$ the marginal APP for the symbol x_t taking into account observations and a priori information, up to time $t+N-1$. In other words, the equaliser delivers statistical information on symbols when they are "seen" by the channel for the last time. We denote this APP by :
$\alpha_{t|t+N-1}(\xi) = P(x_t = \xi | Y_{t+N-1}, \mathcal{L}_a(x_{t+N-1}) \ \ldots \ \mathcal{L}_a(x_0))$
for $\xi = +1$ or $\xi = -1$; where $Y_{t+N-1} = \{y_0, y_1, \ \ldots \ y_{t+N-1}\}$ is the sequence of observations up to time t and $\mathcal{L}_a(x_{t+N-1})$ the a priori LLRs. Taking into account the interleaver, each symbol x_t corresponds to a coded bit c_n.

The APP equaliser takes as prior, soft information on the coded bits (provided by the decoder) in the form of log likelihood ratios (LLRs) and outputs the a posteriori LLR minus the a priori LLR:

$$\mathcal{L}_e(x_t) = \underbrace{log(\frac{\alpha_{t|t+N-1}(+1)}{\alpha_{t|t+N-1}(-1)})}_{\mathcal{L}_p(x_t)} - \underbrace{log(\frac{P_a(x_t=+1)}{P_a(x_t=-1)})}_{\mathcal{L}_a(x_t)} \quad (2)$$

Note that the term $\mathcal{L}_a(x_t)$ represents prior information on the occurrence probability of x_t and is provided by the decoder at the previous iteration. Therefore, $\mathcal{L}_e(x_t)$ is independent of this prior information.

Using $\mathcal{L}_a(c_n)$ (i.e $\mathcal{L}_e(x_t)$ through the interleaver) as prior information, the decoder then generates extrinsic LLR on the code bit:

$$\mathcal{L}_e(c_n) = \underbrace{log(\frac{P(c_n=1|\mathcal{L}_a(c_1,\ldots c_K))}{P(c_n=0|\mathcal{L}_a(c_1,\ldots c_K))})}_{\mathcal{L}_p(c_n)} - \underbrace{log(\frac{P_a(c_n=1)}{P_a(c_n=0)})}_{\mathcal{L}_a(c_n)} \quad (3)$$

Note that at the first iteration, the equaliser computes $\alpha_{t|t+N-1}(\xi)$ without using $\mathcal{L}_a(x_t)$ since this information is not available.

2.3. The proposed equaliser

The vector X_t can be seen as the state vector of the Markov process described by the following state equation :

$$X_{t+1} = AX_t + x_{t+1} * [1 \ 0 \ \ldots \ 0]^T \quad (4)$$

where A is a shift matrix with $A_{i,j} = 1 \leftrightarrow i = j+1$. This Markov process is only observable through the observation equation (1). Suppose that the channel H is available. We define the so-called forward variable, expressing the probability that the state X_t be equal to some realisation $[\xi_0 \ \ldots \ \xi_{N-1}]^T$ according to the channel H, and the set of measurements Y_t by:

$$\alpha_{t|t}(\xi_0, \xi_1, \ldots, \xi_{N-1}) = P(X_t^T = [\xi_0, \xi_1, \ldots, \xi_{N-1}]|H, Y_t) \quad (5)$$

The exact computation of this probability involves the so-called forward recursion. We refer the reader to [7] for more details. This recursion requires the calculation of the above probability for every possible realisation of the stochastic process X_t. Such an evaluation obviously requires the computation of M^N probabilities at each step. Thus, it is desirable to seek for a simplified algorithm which permits state revisiting but does not have an exponential complexity in N. Such an algorithm was presented in [5]. This algorithm uses the marginal posterior probabilities of the symbols in the channel and has linear computational complexity in the channel duration. We now briefly recall the forward recursion for this algorithm. We assume in the following that the channel is known. Assume the following quantities be available at time t :

- the approximate filtered probabilities, $\alpha^{(n)}_{t-1|t-1}(\xi)$ denoting the probability that the $(n+1)^{th}$ symbol in the channel memory at time $t-1$ (i.e x_{t-n-1}) be equal to ξ, knowing the observations up to time $t-1$ and the prediction of the other symbols stored in the channel memory at time $t-1$ ($X^{(m)}_{t-1}$, $\forall m \neq n$, $X^{(m)}_{t-1}$ denoting the $(m+1)^{th}$ component of vector X_{t-1})

$$\begin{array}{lcl} \alpha^{(n)}_{t-1|t-1}(\xi) & = & P(X^{(n)}_{t-1} = \xi | Y_{t-1}, X^{(m)}_{t-1} = \hat{X}^{(m)}_{t-1|t-2}) \\ \quad \forall m \neq n & & \forall n = 0, \ldots \ N-1, \xi = 1 \text{ or } \xi = -1; \end{array}$$

- the current estimate $\hat{X}_{t-1}$ of the vector X_{t-1} as given by the previous recursion. A prediction $\hat{X}_{t|t-1}$ of vector X_t is easily obtained by taking advantage of the shift structure of the process X_t. Clearly we have, for $n = 1 \ \ldots \ N-1$

$$\hat{X}^{(n)}_{t|t-1} = \hat{X}^{(n-1)}_{t-1|t-1} \quad (6)$$

Then by substituting $\hat{X}^{(n)}_{t|t-1}$ for $X^{(n)}_t$ $\forall n = 1 \ : \ N-1$ we obtain the approximate filtered probability at time t of the only component of the state vector on which (6) does not provide information:

$$\alpha^{(0)}_{t|t}(\xi) = P_a(x_t = \xi) P(X^{(0)}_t = \xi | Y_t, X^{(n)}_t = X^{(n)}_{t|t-1}) \quad (7)$$

where $P_a(x_t = \xi)$ is evaluated using $\mathcal{L}_a(x_t)$ (i.e the output of the decoder) obtained from the previous iteration of the turbo process. Substituting from (1) yields

$$\alpha^{(0)}_{t|t}(\xi) = a^{(0)}_t P_a(x_t = \xi) N_{\sigma^2_w}(y - \hat{m}_{(\xi)}) \quad (8)$$

where $a^{(0)}_t$ is a normalising constants, $\hat{m}_{(\xi)}$ stands for the vectorial product $H^T[\xi, \hat{X}^{(1)}_{t|t-1}, \ \ldots \ \hat{X}^{(N-1)}_{t|t-1}]^T$, and $N_{\sigma^2_w}(.)$ is the zero mean Gaussian function with variance σ^2_w:

$$N_{\sigma^2_w}(y - m_{(\xi)}) = \frac{1}{\sqrt{2\pi}\sigma_w} exp(-\frac{(y - m_{(\xi)})^2}{2\sigma^2_w}) \quad (9)$$

In the forthcoming, $L^{(n)}_t(\xi)$ denotes the quantity

$$N_{\sigma^2_w}(y_t - H^T[\hat{X}^{(0)}_{t|t}, \hat{X}^{(1)}_{t|t-1}, \ \ldots \ , \xi, \ \ldots \ , \hat{X}^{(N-1)}_{t|t-1}]^T) \quad (10)$$

where $\hat{X}^{(n)}_{t|t-1}$ has been replaced by ξ. The remaining updated probabilities $\alpha^{(n)}_{t|t}(\xi)$ are also approximated by applying the classical forward recursion of the HMM formulation on conditional instead of joint probabilities.The quantities $\alpha^{(n)}_{t|t}(\xi)$ recorded as smoothed probabilities are thus obtained as

$$\alpha^{(n)}_{t|t}(\xi) = a^{(n)}_t \alpha^{(n-1)}_{t-1|t-1}(\xi) L^{(n)}_t(\xi) \quad (11)$$

The equaliser presented in this section has been shown to provide a much better performance than that of a DFE [5]. Indeed, based on a sub-optimal formulation of HMM theory, it allows to revisit the symbol probabilities as long as the symbol is seen by the channel memory.

The prior information provided by the decoder is used in the calculation of the marginal probability of a symbol which is seen by the channel for the first time (8). This prior information naturally propagates with the forward recursion (11). Therefore, when the equaliser delivers the LLR on symbol x_t, it has taken into account information provided by the soft decoder and it has also exploited the time redundancy introduced by the ISI. In a sense, this equaliser takes advantage of the ISI while the IC simply tries to compensate for it.

3. MODIFIED CHANNEL NOISE VARIANCE

3.1. Proposed Method

A basic principal of the turbo systems is the exchange of soft information which allows to quantify the occurrence probability of a symbol. It is important to find a judicious way to determine σ_p^2 so that the equaliser output probability matches the confidence we can actually grant it with. To do so, we use the EXIT chart principle, which is a simulation based analysis of a turbo system, that separates the two components forming the system, and studies each of them independently of the other one [8] [9]. One component is viewed as a device mapping a quantity of input mutual information (I_i) into a quantity of output mutual information (I_o). An important assumption of the EXIT charts is that the LLRs fed back, for instance from the equaliser to the decoder, can be modeled as iid, with a Gaussian conditional pdf. The variance σ_i^2 of the LLRs at one component input is linked to the mutual information at the same component input I_i by the reversible function J. $I_i = J(\sigma_i)$:

$$I_i = 1-\frac{1}{\sqrt{2\pi}\sigma_i}\int_{-\infty}^{+\infty} exp\left(-\frac{(l-\frac{\sigma_i^2}{2})^2}{2\sigma_i^2}\right)\cdot log_2[1+e^{-l}]dl \quad (12)$$

Although the 2 components are analysed independently we have to bear in mind that the equaliser output is also the decoder input, through the interleaver. Thus the statistics equality $\sigma_{o\{eq\}}^2 = \sigma_{i\{dec\}}^2$ and $I_{o\{eq\}} = I_{i\{dec\}}$ should be assumed. Here, we added the subscript $\{eq\}$ and $\{dec\}$ in order to clearly differentiate the equaliser statistics from the decoder statistics. In fact the equaliser behavior depends on the channel observation (with noise σ_w^2), σ_p^2 and $I_{i\{eq\}}$ coming from the decoder output. We want to tune σ_p^2 so that the two representations of σ^2 versus I match.

The original receiver (Fig. 1) is modified into a system including a new function which determines σ_p^2 from the equaliser input mutual information $I_{i\{eq\}}$: $\sigma_p^2 = f_{\sigma_w^2}(I_{i\{eq\}})$. $I_{i\{eq\}}$ is calculated from the decoder extrinsic LLRs (using the same measurement method as for the EXIT chart, see [8] [9] for exact equation) after each iteration. $f_{\sigma_w^2}(.)$ is such that the equaliser extrinsic LLR statistics follow the function $\sigma_{o\{eq\}}^2 = J^{-1}(I_{o\{eq\}})$.

In order to obtain the function $f_{\sigma_w^2}$, we run EXIT simulations. For each value of the channel noise variance σ_w^2, we draw a graph superimposing the grid of points obtained, over the graphical representation of the theoretical function J^{-1} (12). Results are shown in Fig. 2. Within the whole set of points of $\sigma_{o\{eq\}}$ versus $I_{o\{eq\}}$, we select the ones in the neighborhood of the theoretical curve. We suppose these points to be ideal for optimising the information exchange between equaliser and decoder. Horizontally we read the input mutual information $I_{i\{eq\}}$, vertically we read σ_p^2 parameter. The desired function $f_{\sigma_w^2}(.)$ is then obtained by intervals. An example is shown in the following section.

3.2. Simulation Results

Simulations have been performed on a type B Proakis channel, i.e, $H = [0.407 \quad 0.814 \quad 0.407]$. The channel is assumed to be constant and known by the receiver. The aim of the simulations is to observe the effect of the proposed modification on the BER performances of the turbo-equaliser combining a DFE incorporating fixed lag smoothing with an APP decoder.

For each σ_w^2 value which is used to draw the BER curves, a mapping table between mutual information $I_{i\{eq\}}$ and σ_p^2 parameter

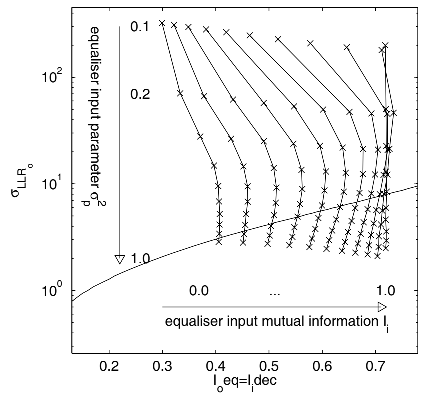


Fig. 2. Equaliser output variance function of the equaliser output mutual information for different values of input mutual information and σ_p^2 parameter. Channel noise variance $\sigma_w^2 = 0.5$. Proakis B channel [10]

needs to be created in accordance with the previously described criterion. As an example, Table 1 presents the mapping between I_i and σ_p^2 for a channel noise variance $\sigma_w^2 = 0.5$.

Table 1. Mapping between the equaliser mutual information I_i and the equaliser parameter $\sigma_p^2 = f_{\sigma_w^2}(I_{i\{eq\}})$, for $\sigma_w^2 = 0.5$

$I_{i\{eq\}} < 0.05$	$\rightarrow$	$\sigma_p^2 = 1.0$
$0.05 \leq I_{i\{eq\}} < 0.15$	$\rightarrow$	$\sigma_p^2 = 0.9$
$0.15 \leq I_{i\{eq\}} < 0.25$	$\rightarrow$	$\sigma_p^2 = 0.8$
$0.25 \leq I_{i\{eq\}} < 0.40$	$\rightarrow$	$\sigma_p^2 = 0.7$
$0.40 \leq I_{i\{eq\}} < 0.65$	$\rightarrow$	$\sigma_p^2 = 0.6$
$0.65 \leq I_{i\{eq\}}$	$\rightarrow$	$\sigma_p^2 = 0.5$

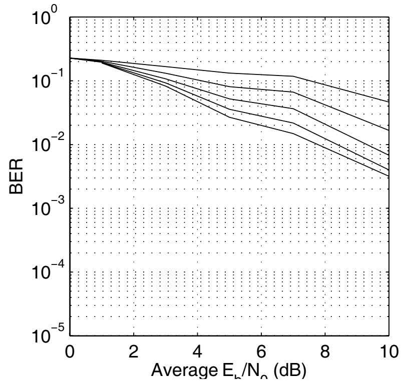
We observe that when $I_{i\{eq\}}$ is small, σ_p^2 takes larger values. Then σ_p^2 gradually decreases as the input mutual information increases. Eventually when $0.65 \leq I_{i\{eq\}}$, σ_p^2 takes the true channel noise variance value 0.5. It is interesting to notice that the lower bound that σ_p^2 reaches is the true channel noise variance σ_w^2.

The performances in terms of BER of the turbo-decoder, are plotted. Fig. 3 shows the decoder output for each of the 5 iterations of the process. Compared to Fig. 3(a), where the true channel noise variance $\sigma_p^2 = \sigma_w^2$ is input, Fig. 3(b) with an adapted σ_b^2 shows smoother curves and reaches a better BER for any SNR.

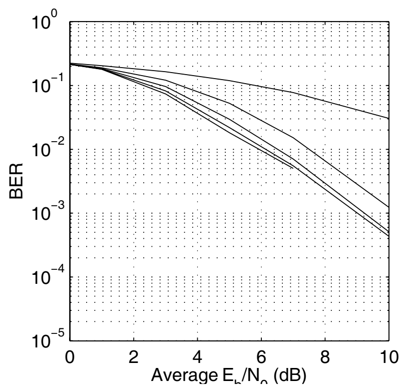
It is well known that the a priori input mutual information of an equaliser included in a converging turbo-equaliser is null at the first iteration (no a priori before the first decoding step) and increases at each iteration. It is a logical consequence that σ_p^2 has larger values at the first iterations than at the last ones.

4. INFLUENCE OF THE A PRIORI ERROR VARIANCE ON THE MODIFIED CHANNEL NOISE VARIANCE

We saw in the previous section that it is possible to improve the equalisation by modifying the estimated channel noise variance σ_p^2. The resulting variance is function of the actual channel noise variance σ_w^2 and the equaliser input mutual information $I_{i\{eq\}}$. In a way, the a priori information compensates for the sub-optimality of the equaliser. The difference with σ_w^2 decreases with the number of



(a) Non adapted $\sigma_p^2 = \sigma_w^2$



(b) Adapted σ_p^2

Fig. 3. Decoder output - BER versus SNR: comparison of the “traditional method” using $\sigma_p^2 = \sigma_w^2$ to the adapted method using $\sigma_p^2 = f_{\sigma_w^2}(I_{i\{eq\}})$

iterations as the turbo-decoder converges. In [8] and [9] the authors suppose the a priori information at the input of the equaliser modeled as the outputs of an AWGN channel with zero mean and variance: $\sigma_a^2 = 4/\sigma_{i\{eq\}}^2$. It is possible to observe a linear relation between the logarithm of σ_a^2 that we just defined and the chosen variance σ_p^2. This relation is shown on Fig.3.2. The curves marked with 'x' are the actual curves obtained from simulations. On the top of it is added the first order polynomial fitting, marked with 'o'. The slope is well respected at every iteration.

5. CONCLUSION

In this paper, we consider a turbo-equaliser based on a sub-optimal probability based equaliser. The channel is known by the equaliser. In the case of reduced complexity probability based equaliser coupled with an APP decoder, we say that the influence of each component should not be equal in the process. We demonstrate that we can use the channel noise input of the equaliser to tune the system. The BER simulations improve thanks to this method. Finally, we highlight a relation between the chosen variance σ_p^2 and the error variance σ_a^2 at the equaliser input. The aim of this work is to obtain in the future an analytical way to predict the optimal value of the parameter σ_p^2.

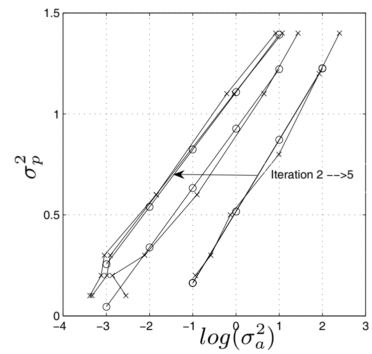


Fig. 4. σ_p^2 vs $log(\sigma_a^2)$

6. REFERENCES

[1] C. Douillard, M. Jézéquel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, “Iterative correction of intersymbol interference: Turbo equalization,” *European Transactions on Telecommunications and Related Technologies*, vol. 6, no. 5, pp. 507–511, Sept. – Oct. 1995.

[2] C. Berrou, A. Glavieux, and P. Thitimajshima, “Near Shannon limit error-correcting coding and decoding: Turbo codes,” in *IEEE Int. Conf. on Communications (ICC-1993)*, vol. 2. Geneva, Switzerland: IEEE, May 1993, pp. 1064–1070.

[3] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.

[4] A. Glavieux, C. Laot, and J. Labat, “Turbo equalization over a frequency selective channel,” in *Internationnal Sympothium on Turbo-codes*. Brest, France: IEEE, 1997, pp. 96–102.

[5] S. Perreau, L. White, and P. Duhamel, “A blind decision feedback equaliser incorporating fixed lag smoothing,” *IEEE trans on signal processing*, vol. 48, no. 5, pp. 1315–1328, May 2000.

[6] A. Worm, P. Hoeher, and N. Wehn, “Turbo-decoding without snr estimation,” *IEEE Communications Letters*, vol. 4, no. 6, June 2000.

[7] L. Rabiner, “A tutorial on hidden markov models and selected applications in speech recognition,” in *Readings in Speech Recognition*, A. Waibel and K.-F. Lee, Eds. San Mateo, CA: Kaufmann, 1990, pp. 267–296.

[8] S. ten Brink, “Convergence of iterative decoding,” *Electron. Lett*, vol. 35, pp. 806–808, May 1999.

[9] M. Tüchler, R. Koetter, and A. Singer, “Turbo Equalization: Principles and new results,” *IEEE Transaction on communication*, vol. 50, pp. 754–767, May 2002.

[10] J. G. Proakis, *Digital Communications*, 3rd ed. New York, NY, USA: McGraw-Hill, 1989.