On the estimate of the noice variance used in A Posteriori Probability based turbo-equalisers

Sylvie Perreau and Gilles Gorlier
Institute for Telecommunication Research,Uni of South Australia,
5095, Mawson Lakes, Australia
email: sylvie.perreau@unisa.edu.au

Abstract— **In this paper, we revisit some fundamental results on turbo-equalisation and more precisely on the prediction of the turbo-equaliser's performance using the EXIT chart. We provide an explaination as to why the performance of certain equalisers do not match the one predicted by the EXIT chart and we propose a method for choosing a single parameter which allows not only a better performance prediction but also improves the overall convergence properties of Aposteriori Probability (APP) based equalisers.**

I. INTRODUCTION

Turbo-decoding has inspired many researchers in the telecommunications field since its discovery by Claude Berrou in 1993 [1]. Many iterative applications followed the revolutionary discovery of turbo-codes, amongst which came the turbo-equaliser, discovered by Douillard *et al.* in 1995 [2]. It is derived from the serially concatenated turbo-encoder scheme where the outer encoder is RSC, and the finite impulse response (FIR) channel plays the role of the inner encoder. At the receiver end an equaliser combat the intersymbol interferences (ISI) introduced by the channel, using the a priori information fed back from the decoder. As for turbo-decoding, the same information enhancement phenomenon occurs along the iterations. The first turbo-equalisers proposed used a posteriori probability (APP) based techniques for both equalisation and decoding. They are penalised by a high computational load. An alternative was found in the interference canceler [3]. Several hybrid methods followed, based on a modified linear equaliser (LE) and a decision feedback equaliser (DFE) [4]. In [5], a turbo equaliser based on a sub-optimum APP based equaliser was proposed, allowing a good trade off between performance and complexity. However, the difficulty laid in obtaining a reliable way of predicting its performance. A popular tool to analyse the convergence of turbo equalisers and turbo codes has been found in the so called EXIT chart. The EXIT chart allows to predict how the mutual information between transmitted symbols and Log Likelihood ratios (LLRs) is exchanged between the equaliser and the decoder. This requires the assumption that the LLRs at the input of one device (i.e decoder or equaliser) are not only Gaussian but in fact, stem from the Maximum Likelihood detection of symbols corrupted by Additive White Gaussian Noise (AWGN). In reality, these properties on the LLRs are rarely met and as a result, the performance of turbo-equalisers are usually degraded compared to the EXIT chart predictions.

In [6] a new technique was proposed in order to increase the mutual information exchanged between a sub-optimal APP equaliser and the APP decoder, by carefully choosing the channel noise variance estimate used in the calculation of the APPs. It was shown that the performance in terms of Bit Error Rate and convergence was improved. The reasons behind this performance improvment however was not clear. In this paper, we extend this method to A Posteriori Probability (APP) based equalisers and we explain why, by tuning the noise variance estimate in the same way as [6], we in fact ensure that the performance of the turbo-equaliser matches the performance predicted by the EXIT chart. In other words, by careful choice of this parameter, we will show that the system formed by the ISI channel, the AWGN and the sub-optimal equaliser is equivalent to a AWGN channel with a known noise variance. We also show that in the case of a maximum a posteriori (MAP) equaliser, the way the noise variance estimate is usually chosen (i.e, equal to the true noise variance) does not guarantee that the LLRs at the input of the decoder are Gaussian as it is often assumed, which explains the mismatch between EXIT chart predictions and simulation results that have been reported in the literature. The paper is organised as follows: section 2 proposes a problem formulation and revisits some fundamental results namely, the calculation of the LLRs and the mutual information in the ideal case (AWGN channel) and in the case of the ISI channel coupled with optimal and sub-optimal APP based equalisers. We also recall the method proposed in [6] and analytically describe the meaning of the resulting noise variance estimate. We finally present some simulation results which illustrate some aspects of the analysis proposed in section 2.

II. PROBLEM FORMULATION AND FUNDAMENTAL RESULTS

A. Transmitter

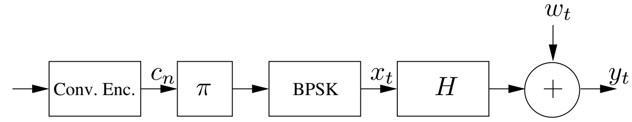

Fig. 1. Transmitter structure: bit interleaved coded modulation

We consider the transmitter and channel as shown in Figure 1. An iid sequence of non coded bits is passed through the channel encoder to produce the length K sequence of coded symbols 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 FIR channel with transfer function $H(z) = \sum_{i=0}^{N-1} h_i z^{-i}$ where the h_i are complex-valued coefficients. We assume AWGN with zero mean and variance σ_w^2. The received signal is thus modeled by

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

where H and X_t are so defined:

$$H^T = [h_0 \; h_1 \; \ldots \; h_{N-1}]$$
$$X_t = [x_t \; x_{t-1} \; \ldots \; x_{t-N+1}]^T$$

Here, the operator $()^T$ denotes the transposition operation.

B. Receiver

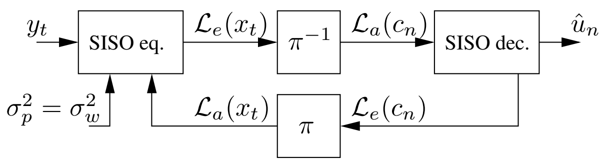

Fig. 2. Receiver structure: turbo-equaliser

Figure 2 illustrates the structure of a turbo-equaliser where the Log Likelihood ratio (LLR) operator $\mathcal{L}(x)$ is applied to coded symbols x and is given by:

$$\mathcal{L}(x) = \log \frac{P(x = +1)}{P(x = -1)}$$

The 2 components of the receiver, namely equaliser and decoder, exchange extrinsic log likelihood ratios (LLRs). The soft input soft output (SISO) equaliser computes the a posteriori LLRs of the symbols x_t, $\mathcal{L}_p(x_t)$, from the channel observations y_t and the a priori LLRs $\mathcal{L}_a(x_t)$. The a priori LLRs come from the decoder previous iteration, except at the first iteration since this information is not available. The extrinsic LLR is the difference between the a posteriori LLR and the a priori LLR:

$$\mathcal{L}_e(x_t) = \mathcal{L}_p(x_t) - \mathcal{L}_a(x_t)$$

and on the SISO decoder side:

$$\mathcal{L}_e(c_n) = \mathcal{L}_p(c_n) - \mathcal{L}_a(c_n)$$

The interleaver shuffles the LLRs so to separate the errors when error burst occurs. Although the interleaver is invisible to the equaliser and the decoder, its presence is essential for the efficiency of the turbo-process [7].

C. LLR and mutual information derivations in the AWGN channel case

In this section, we briefly recall some fundamental derivations of the LLRs and the mutual information which are used to obtain the so called EXIT chart. The first important point to note is that the EXIT chart consists of 2 curves which show how the mutual information is exchanged between the encoder and equaliser. One key assumptions made on the LLRs in order to obtain the EXIT chart is that they are LLRs of the output of an AWGN channel with noise power σ^2 and with input the symbols x, in other words:

$$y = x + w$$

Proposition 1: *In the AWGN case, the mutual information between the LLRs and the symbols x does not depend on the parameter σ_p^2 used in the calculation of the probabilities. However, the mean of the LLRs is equal to half of their variance only when $\sigma_p^2 = \sigma^2$, which corresponds to an optimum the performance of the decoder following the equaliser*
Proof:
It has been shown [8] that in the case of the AWGN channel, the variance of LLRs is $4/\sigma^2$ and the mean $2/\sigma^2$. It is interesting to note that the derivation of the above mean and variance values of LLRs were obtained with the noise estimate parameter used in the Gaussian probability computation equal to the true noise variance. When a different parameter is used in the computation of the probabilities, i.e, $P(x = 1|y) = \exp -\frac{1}{2\sigma_p^2}(y-1)^2$, with $\sigma_p^2 \neq \sigma^2$, the LLRs are written as:

$$LLR = \frac{2}{\sigma_p^2}(1 + w) \tag{2}$$

Obviously, now, the mean of the LLR is $\frac{2}{\sigma_p^2}$ while the variance is $\frac{4\sigma_w^2}{\sigma_p^4}$. It is interesting to see that in this case, when the a posteriori probabilities are computed with a noise variance estimate different from the true one, the mean of the LLRs is no longer equal to half of the variance. The probability density function of the LLR can then be expressed as

$$f_L(l|x) = \frac{\sigma_p^2}{2\sqrt{2\pi}\sigma} e^{-\frac{\sigma_p^4}{8\sigma^2}(l - \frac{2}{\sigma_p^2})^2}$$

One can easily show that the corresponding mutual information between the LLRs and the symbols x, defined by

$$I(L_i, X) = \frac{1}{2} \sum_x \int_{-infty}^{+\infty} f_L(l|x) \log \frac{2 f_L(l|x)}{f_L(l|+1) + f_L(l|-1)}$$

remains the same. In this case, one can show that:

$$I(L_i, X) = J(\sigma) = 1 - \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} exp\left(-\frac{(l - \frac{\sigma^2}{2})^2}{2\sigma^2}\right) \cdot log_2[1 + e^{-l}] dl \tag{3}$$

D. LLR and mutual information derivation for a MAP equaliser

Proposition 2:
For a certain noise variance estimate σ_p^2 used in the computation of the APP of a MAP turbo-equaliser, the mutual information between the LLRs at the output of the equaliser and the symbols at the input of the channel is equal to that of an AWGN channel with variance σ_p^2, where the LLR mean is equal to half of the variance. **Proof**:
In [9], [10], it was shown that the performance of a MAP

equaliser following a Finite Impulse Response (FIR) channel with AWGN of variance σ^2 is equivalent to the performance achieved for an AWGN channel with variance $\sigma_d^2 = \frac{4\sigma^2}{d_{min}^2}$ where d_{min} is the channel mimimum distance. Therefore, using results obtained in the previous section, we can conclude that if and only if the variance estimate used in the computation of the APPs of the MAP is equal to σ_d^2, the distribution of the LLRs at the output of the MAP will be Gaussian with mean $\frac{4}{\sigma_d^2}$ and variance $\frac{2}{\sigma_d^2}$. This result is significant because it indicates that in this case, we can expect a perfect match between the performance prediction obtained using the EXIT chart and the actual simulation results. It has to be pointed out that previous works simply assume that the variance estimate used in the computation of the APPs for a MAP equaliser should be equal to the actual noise variance. In this case, from the results obtained in the previous section, the LLRs would be such as their mean and variance are respectively $\mu = \frac{d_{min}^2}{2\sigma^2}$ and $\sigma_{LLR}^2 = \frac{d_{min}^4}{4\sigma^2}$. Unless the minimum distance $d_{min}^2 = 4$, the LLRs do not satisfy $\sigma_{LLR}^2 = 2\mu$ and the simulation results do not match the EXIT chart predictions. Another important point is that since the decoder is optimum when its input is equivalent to an AWGN channel, sub optimum results are to be expected at the output of the decoder when the equaliser fails to provide APPs corresponding to such AWGN channel.

We have therefore shown that the choice of the noise variance estimate used in the APPs computation is crucial, to ensure not only that the EXIT chart predictions are accurate but also that the decoder is able to provide optimum results.

E. The case of a sub-optimum APP based equaliser

In this section, we concentrate on the decision feedback equaliser incorporating fixed lag smoothing (DFEifls) presented in [5] and we derive the expression for the LLRs, highligting why the mutual information between the LLRs and coded symbols is dependent on the choice of the noise variance estimate (which was not the case for the AWGN channel as shown in section 2.1). We also provide a method for optimising this parameter and the simulation results presented in the next section will highlight the dramatic improvments brought out by our proposed method.
The full derivation of the DFEifls can be found in [5]. In this paper, for sake of simplicity for the LLR derivation (but without losing generality), we restrict our attention to a length 3 FIR channel. The DFEifls produces at time t the marginal aposteriori probability $\alpha_{t|t}^{(n)}$, i.e $Pr(x_{t-n+1}|Y_0^t)$ for the symbol x_{t-n+1} taking into account the vector of observations $Y_0^t = \{y_0, y_1, \ \dots \ y_t\}$ up to time t and apriori information in the context of a turbo-equaliser. The equaliser delivers statistical information to the decoder on symbols when they are "seen" by the channel for the last time. At time t, the equaliser will therefore provide

$$\alpha_{t|t}^{(2)} = Pr(x_{t-2} = 1|Y_0^t, L_a(x_t) \ \dots \ L_a(x_0)) \tag{4}$$

The DFE incorporating fixed lag smoothing [11] takes advantage of a shifting property of the successive states to calculate the marginal probability of a symbol. In other words,

$$\alpha_{t|t}^{(2)} = a_t^{(2)} \alpha_{t-1|t-1}^{(1)} \exp\{-\frac{1}{2\sigma_p^2}(y_t - H^T[\hat{x}(t|t), \hat{x}(t-1|t-1), 1]^T)^2\} \tag{5}$$

where $a_t^{(2)}$ are normalising constants, σ_p^2 is the noise variance estimate, $\hat{x}(t|t)$ and $\hat{x}(t-1|t-1)$ are smoothed estimates of symbols x_t and x_{t-1} respectively. Note that these smoothed estimates are computed as conditional expectations. For instance,

$$\hat{x}(t|t) = 2\alpha_{t|t}^{(0)} - 1$$

It is worth noting that these estimates depend on σ_p^2. By applying the recursion on $\alpha_{t-1|t-1}^{(1)}$, we can show that:

$$\alpha_{t|t}^{(2)} = \prod_{i=0}^{2} a_{t-1}^{(i)} \exp\{-\frac{1}{2\sigma_p^2} \sum_{i=0}^{2} (y_{t-2+i} - h_i - \sum_{j \neq i} h_j \hat{x}_{t-j})^2\} \tag{6}$$

After few manipulations which we cannot provide in this paper by lack space but can be found in [**?**], we can show that the LLRs can be expressed by:

$$\mathcal{L}_e(x_t) = \frac{2}{\sigma_p^2}(1 + \mathcal{W}(\sigma_p^2)) \tag{7}$$

where $\mathcal{W}(\sigma_p^2) = \sum_{i=0}^{2} w_{t-i} + \sum_{j \neq i} h_j h_i (x_{t-i} - \hat{x}(t-i|t-j))$. It is interesting to note that this expression resembles equation (2) which is the expression of the LLR in the case of the Gaussian channel. Unfortunately, in the case of the DFEifls, the noise term $\mathcal{W}(\sigma_p^2)$ is not Gaussian (due to the presence of the error terms $e(t-i) = x_{t-i} - \hat{x}(t-i|t-j)$). Moreover, this noise term strongly depends on the value of the parameter σ_p^2 which indicates that the corresponding mutual information between LLRs and symbols depends on σ_p^2. In [6], it is suggested that in order to optimise the performance of the turbo equaliser, one has to choose σ_p^2 such that the corresponding mutual information is equal to the mutual information of LLRs derived from data corrupted by an AWGN channel. As shown in figure 3, this operation is realised by superimposing the output mutual information I_o and LLR variance σ_{LLRo}^2 obtained for the equaliser over the graphical representation of the function $J(\sigma)$ given by equation (3). It is worth mentioning that no real justification was providedin [6] as to why such a choice of σ_p^2 resulted in such significant improvments as compared with a usual way of estimating the channel noise variance. Using the argument similar to that of the previous section, we can now justify this choice by the fact that when the mutual information $\mathcal{L}_e(x_t)$ is equal to a mutual informationwhich can be expressed using (**??**), then the performance of the system formed by the FIR channel and equaliser is equivalent to that of an AWGN channel with noise variance equal to $\sigma_w^2 = \frac{4}{\sigma_{LLRo}^2}$ where σ_{LLRo} is found where the curves of mutual information at the output of the equaliser and the curve related to (**??**) intersect. Again, since the decoder is optimum when fed by LLRs satisfying the property guaranteed by (3), we can explain why the turbo-equaliser provides the best performance when σ_p^2 is chosen using this method. The other important conclusion that can be drawn is that the predictions provided by the EXIT chart will be more accurate using this choice of parameter.

III. Simulations

The simulations presented in this section have been obtained using the Proakis B channel, i.e $H = [0.4080.8170.407]$.

Adapting the noise variance estimate for the MAP: Figure 3 illustrates the findings of section 2.4. In this figure, two curves are superimposed: the function $\sigma^2_{LLR} = J^{(-1)}(I_o)$ obtained theoretically when the input LLRs are Gaussians with a mean equal to half the variance. The other curve is the result of simulations of an MAP based turbo equaliser at the first iteration (when no a-priori information is available). We plot the variance of the LLRs in the y-axis and the corresponding mutual information in the x-axis. Simulations are repeated for various values of σ^2_p until the theoretical and simulation curves intersect which occurs as indicated by the dashed line for $\sigma^2_p = 0.6$. According to our analysis provided in section 2.4, for this value of σ^2_p obtained at the intersection of the 2 curves, the LLRs are Gaussian and in fact σ^2_p can be found as $\sigma^2_p = \frac{4\sigma^2}{d^2_{min}}$. For $H = [0.4070.8170.408]$, $d^2_{min} = 2.6$. Therefore, $\frac{4\sigma^2}{d^2_{min}} = 0.61$ which matches the value σ^2_p obtained on the curve.

Adapting the noise variance estimate for the DFEifls

Figure 4 presents results similar to the ones in figure 3 in the case of a DFEifls. The first observation is that the mutual information varies depending on the chosen σ^2_p. Another interesting observation is that for $\sigma^2_p = 0.5$ (which is the optimum value for the turbo equaliser), the corresponding LLR variance is $\sigma^2_{LLR} = 6$ and according to our analysis, the performance of the system comprising the FIR channel, the AWGN with noise variance $\sigma^2 = 0.1$ and the DFEifls is equivalent to that of a AWGN channel with variance $\sigma^2_d = \frac{4}{\sigma^2_{LLR}} = 0.666$. If we now turn our attention to figure 8 which shows the bit Error Rate (BER) of the DFEifls as well as that of the AWGN channel, we can see that at the first iteration (which is the one considered in the paper), the BER obtained by the DFEifls for an SNR of 10dB (corresponding to $\sigma^2 = 0.1$ in figure 4, is 10^{-1} which is the same BER obtained for an AWGN channel with SNR of approximately $1.7dB = 10\log_{10}(0.666)$ which matches the value $\sigma^2_d = 0.666$ obtained from the LLR variance. This illustrates the proposition put forward in the previous section.

One of our claims in this paper regarding the benefits obtained by the apropriate choice of the noise variance estimate, is that because we make our system equivalent to a AWGN channel, the performance of the turbo equaliser can be more accurately predicted by the EXIT chart. This claim is validated by figures 5 and 6. In figure 5, the noise variance estimate is equal to the true noise variance. One can see that the exchange of mutual information as plotted on the graph does not follow what the EXIT would have predicted. In the case of appropriate noise variance estimate (in figure 6), not only the EXIT chart predicts very accurately the trajectory of the mutual exchange information but also, the number of iterations necessary to reach convergence is reduced as compared to the non adapted case.

Finally, figure 7 highlights the benefits of adapting the channel noise variance estimate in terms of BER. One can see that the BER is significantly reduced at each iteration of the turbo equaliser when the noise variance is adapted.

IV. Conclusion

In this paper, we have hilighted the importance of properly choosing the noise variance estimate used in the calculation of probabilities for APP based turbo equalisers. We have shown how to determine such estimate in the case of a MAP equaliser and a sub-optimal version of the MAP called DFEifls so that not only better BER results are obtained, but also very accurate prediction of the the turbo equaliser performance is feasible using the EXIT chart. It has to be pointed out that although the proposed choice of variance estimate was implemented at each iteration of the turbo-equaliser, the analysis presented here concentrated on the first iteration only. Future work will include a similar theoretical analysis applicable to subsequent iterations.

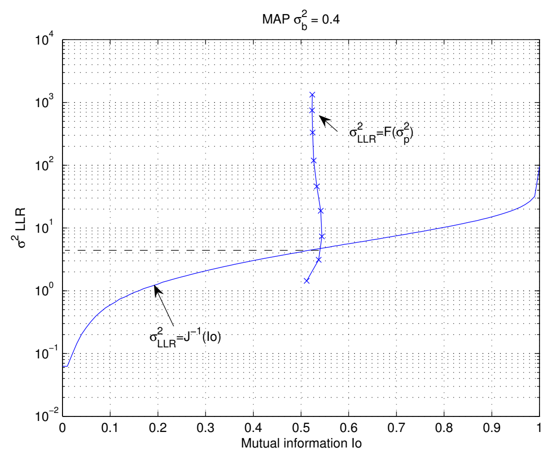

Fig. 3. Optimum σ^2_p for MAP equaliser

References

[1] 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.

[2] 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.

[3] 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.

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

[5] S. Perreau and G. Gorlier, “A new method for channel estimation and data detection on the context of turbo equalization,” in *Proceedings of the 13th IFAC Symposium on System Identification*, 2003, pp. 590–594.

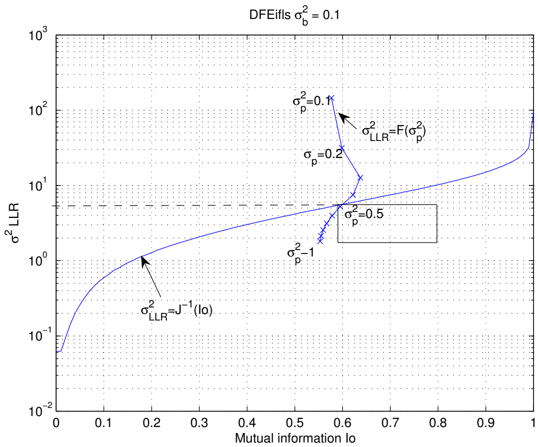

Fig. 4. Optimum σ_p^2 for DFEifls equaliser

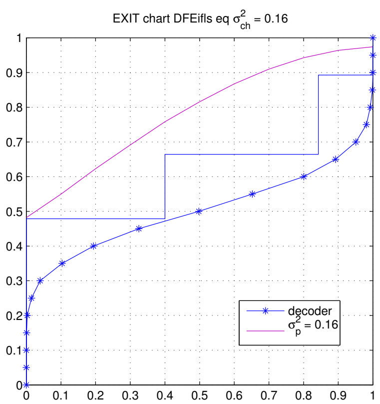

Fig. 5. EXIT chart for DFEifls with $\sigma_p^2 = \sigma^2$

[6] G. Gorlier, S. Perreau, and I. Fijalkow, "Dependency between error variance of the a priori information and a modified channel noise variance in turbo-equalisation," in *to appear in Proceedings IEEE Internationnal Conference on Acoustics, Speech and Sgnal Processing*. IEEE, 2006.

[7] S. Dolinar and D. Divsalar, "Weight Distributions for Turbo Codes Using Random and Nonrandom Permutations," JPL, TDA Progress Report 42-121, Aug. 1995.

[8] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE transactions on communications*, vol. 49, october 2001.

[9] G. D. Forney, "Maximum likelihood sequence estimationof digital sequences in the presence of intersymbol interference," vol. 18, pp. 363–378, May 1972.

[10] N. Sellami, A. Roumy, and I. Fijalkow, "The impact of both a priori information and channel estimation errors on the map equalizer performance," *to appear in IEEE Trans on signal processing*, 2006.

[11] 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.

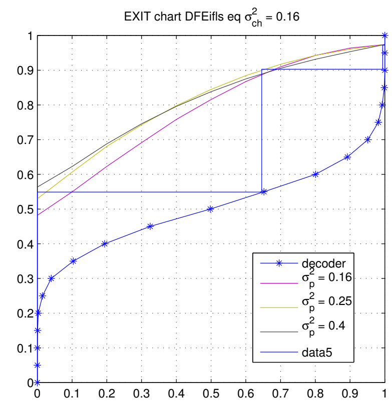

Fig. 6. EXIT chart for DFEifls with optimised σ_p^2

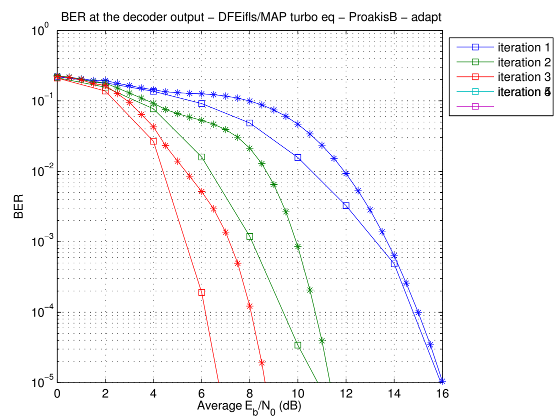

Fig. 7. BER for optimised (squares) and non-optimised (stars) σ_p^2

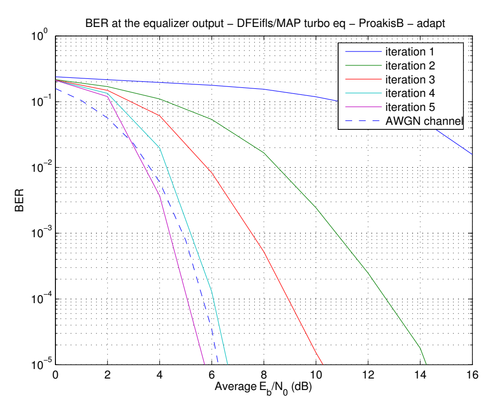

Fig. 8. BER for AWGN channel and DFEifls with noise variance adaptation