A NEW METHOD FOR CHANNEL ESTIMATION AND DATA DETECTION IN THE CONTEXT OF TURBO EQUALISATION

Sylvie Perreau and Gilles Gorlier*

* Institute for Telecommunications Research
University of South Australia
Mawson Lakes SA 5095 Australia
E-Mail: sylvie@spr.levels.unisa.edu.au

Abstract: In this paper, we propose a new turbo equaliser which outputs marginal a-posteriori probabilities of each symbol. These marginal a-posteriori probabilities can be used directly by the decoder in the turbo process. We show how this equaliser is particularly interesting in the context of turbo decoding and also allows to perform blind channel estimation. We will finally show that this new method brings out significant improvements over the classical Interference Canceller (lower SNR threshold or trigger point) with an acceptable computational cost.

1. INTRODUCTION

In recent years, there has been an growing interest in the so called turbo equalisation. The basic idea is that in systems using channel coding, the sequence produced by the equaliser is not equally likely. Therefore, there is much to be gained during the equalisation process, by using the extra a-priori information provided by the encoder. It is however obvious that an optimal formulation to this problem is not feasible when an interleaver is present. Therefore, instead of considering a one step optimal system which would fully map the allowed sequences at the input of the channel, turbo equalisation proceeds by iteratively passing soft information between the decoder and the equaliser. With each iteration, the Bit Error Rate usually decreases as in turbo decoding. However, it has been shown that there exists an SNR threshold under which iterating between equalisation and decoding in no longer beneficial. Several studies have provided some insight into this trigger point effect. In [1], an analysis tool, based on the extrinsic information transfer (EXIT) chart, was used to compare the performance of several turbo equalisers on a single channel. However, the study was only performed on a single channel realisation and did not allow to predict the performance of turbo equalisation for any given channel. In [2], an interesting study also using the EXIT chart provided an insight into predicting the behavior of a particular equaliser, namely the Interference Canceller (IC), depending on the channel dispersion. This study was very useful since it allowed to predict whether turbo equalisation would be useful and it also provided an estimation of the number of iterations necessary to reach convergence. The other interesting point raised in this paper was that the performance of the IC canceller over “tough” channels depends on the type of equaliser used for the first iteration: if a DFE is used, turbo equalisation fails, although when a MAP equaliser is used, turbo equalisation is successful. This observation motivates the work presented in this paper. We propose to perform turbo equalisation using the DFE incorporating fixed lag smoothing presented in [3], which can be viewed as a very good trade off (in terms of computational complexity and performance) between an MAP and a DFE equaliser. Simulations show a significant improvement in terms of trigger point
2. TURBO-EQUALISATION PRINCIPLES

2.1 Problem formulation

We consider the transmitter and channel shown in figure 1. An i.i.d sequence of non coded bits is passed through the channel encoder to produce the sequence $c_m$ which is interleaved and mapped using a BPSK modulation. We assume that the coded and modulated symbols $x_t$ are transmitted through an 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 additive Gaussian white noise with zero mean and variance $\sigma^2$. The received signal is thus modeled by

\[ y_t = H^T X_t + n_t, \]

where $H$ and $X_t$ are defined by

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

Here, the operator $(\cdot)^T$ denotes the transposition operation. For sake of simplicity, we also assume throughout this paper a BPSK modulation.

2.2 Decoding and equalisation

Figure 2 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 a-posteriori probability for the symbol $x_t$ taking into account observations 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 a-posteriori probability by

\[ \alpha_{t|t+N-1}(i) = Pr(x_t = i | Y_{t+N-1}^t) \]

for $i = 1$ or $i = -1$, where $Y_{t+N-1}^t = \{y_1, y_2, \ldots, y_{t+N-1}\}$ denotes the sequence of observations up to time $t + N - 1$.

Taking into account the interleaver, each symbol $x_t$ corresponds to a coded bit $c_m$.

The MAP equalizer 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:

\[ L(x_t) = \log(\frac{\alpha_{t|t+N-1}(1)}{\alpha_{t|t+N-1}(-1)}) - \log(\frac{P(c_m = 1)}{P(c_m = 0)}) \]

Note that the term $\log(\frac{P(c_m = 1)}{P(c_m = 0)})$ represents prior information the occurrence probability of $c_m$ and is provided by the decoder at the previous iteration. Therefore, $L(x_t)$ is independent of this prior information.

Using $L(x_t)$ (i.e. $L(c_m)$), with the coded bit $c_m$ corresponding to the modulated symbol $x_t$ as prior information, the decoder then generates posterior log likelihood on the information bits:

\[ L(c_m) = \log(\frac{P(c_m = 1 | L(c_1), \ldots, L(c_K))}{P(c_m = 0 | L(c_1), \ldots, L(c_K))}) \]

\[ - \log(\frac{P(c_m = 1)}{P(c_m = 0)}) \]

Note that at the next iteration, the equaliser re-computes $\alpha_{t|t+N-1}(i)$ using $L(c_m)$ as explained in the following section.

3. THE PROPOSED EQUALISER

In [6], an optimal formulation for maximum a posteriori probability estimation of the transmitted symbols was presented. This formulation is based on an HMM formulation. Indeed, the vector $X_t$ can be seen as the state vector of the Markov process described by the following state equation:

\[ X_{t+1} = AX_t + x_{t+1} * [1 \ldots 0]^T \]

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 $H_t$
the current estimate of the channel, is available at time $t$. As in [6], define the so-called forward variable, expressing the probability that the state $X_t$ be equal to some realization $[i_0,...,i_{N-1}]^T$ according to the current channel estimate $H_t$, and the set of measurement $Y_t$ by:

$$
\alpha_{t|t}(i_0, i_1,..., i_{N-1}) = P(X_t = [i_0,...,i_{N-1}]|H_t, Y_t)
$$

The exact computation of this probability involves the so-called forward recursion. We refer the reader to [4] for more details. This recursion requires the calculation of the above probability for every possible realization 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 the exponential complexity in $N$ of the approaches of [5] and [6]. Such an algorithm was presented in [3]. 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. (a method for jointly performing turbo equalisation and channel estimation is presented in the next section).

Assume the following quantities be available at time $t$:

- the approximate filtered probabilities,

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

$$
\alpha_{t-1|t-1}^{(n)}(i) = P(X_t = [i]Y_{t-1}, X_{t-1|n}^m = X_{t-1|t-1}^{(m)}| Y_{t-1}) \forall m \neq n, \forall n = 0, ..., N-1, \forall i = 1 or -1 ;
$$

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

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

Then by substituting $\hat{X}_{t|t-1}^{(n)}$ for $X_t^{(n)} \forall n = 1 : N-1$, we obtain the approximate filtered probability at time $t$ of the only component of the state vector on which Eq. (7) does not provide information:

$$
\alpha_{t|t}^{(0)}(i) = P(c_m = i)P(X_t^{(n)} = [i]Y_t, X_t^{(n)} = \hat{X}_{t|t-1}^{(n)}).
$$

where $P(c_m = i)$ is evaluated using $L(c_m)$ (i.e the output of the decoder) obtained from the previous iteration of the turbo process.

Substituting from (1) yields

$$
\alpha_{t|t}^{(0)}(i) = a_t^{(0)} P(c_m = i)
$$

$$
N(\hat{y}_t - \hat{H}_t^{(1)}i, \hat{X}_{t|t-1}^{(1)}, ..., \hat{X}_{t|t-1}^{(N-1)})(8)
$$

where $a_t^{(0)}$ is a normalizing constant and $N(,) is$ a zero mean Gaussian function with variance $\sigma^2$.

In the forthcoming, $L^{(n)}_t(i)$ denotes the quantity $N(\hat{y}_t - \hat{H}_t^{(1)}i, \hat{X}_{t|t-1}^{(1)}, ..., i, ..., \hat{X}_{t|t-1}^{(N-1)})(8)$ where $\hat{X}_{t|t-1}^{(n)}$ has been replaced by $i$.

The remaining updated probabilities $\alpha_{t|t}^{(n)}(i)$ are also approximated by applying the classical forward recursion of the HMM formulation on conditional instead of joint probabilities. The quantities $\alpha_{t|t}^{(n)}(i)$ recorded as smoothed probabilities are thus obtained as

$$
\alpha_{t|t}^{(n)}(i) = a_t^{(n)}\alpha_{t-1|t-1}^{(n-1)}(i) L_t^{(n)}(i).
$$

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

Its formulation makes it very easy to use in the context of turbo-equalisation since the prior information provided in the decoder can be used in the calculation of the marginal APPs of a symbol which is seen by the channel for the first time (see equation (8)). This prior information subsequently naturally propagates with the forward recursion (equation (9)). 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.

4. CHANNEL ESTIMATION

Channel and noise variance estimation can be easily implemented using the Expectation-Maximisation (EM) algorithm. In this paper, we only explain how to obtain the channel estimate. However, we believe that future work should concentrate on the noise variance estimate as well since it is involved in every probability computed by the equaliser and therefore, is a crucial parameter for convergence issues.
Since we are operating in the turbo equalisation context, we iteratively process the same block of data. Therefore, channel estimation can be applied off-line. In this case, it can be shown that the EM algorithm leads to the following expression for the channel estimate at iteration $p$:

$$\hat{H}(p) = R^{-1}u$$  \hspace{1cm} (10)

Where $R$ is the conditional expectation of the auto-correlation of vector $X_t$ and $u$ is the conditional expectation of the cross-correlation between vector $X_t$ and the observation $y_t$. These quantities are calculated at iteration $p$, using the marginal a-posteriori probabilities $\alpha_{H}(n)(i)$ to approximate the a-posteriori probability of vector $X_t$.

More precisely, $R$ and $u$ are computed according to:

$$R = \sum_{t=1}^{T} \hat{X}(t|t)\hat{X}(t|t)^T$$  \hspace{1cm} (11)

$$u = \sum_{t=1}^{T} \hat{X}(t|t)y_t$$  \hspace{1cm} (12)

with $\hat{X}(t|t) = [x(t|t) x(t - 1|t) \cdots x(t - N + 1|t)]^T$, where the $(x)(t-n|t)$ is computed as the conditional expectation of symbol $x_{t-n}$ using $\alpha_{H}(n)(i)$.

It has to be pointed out that the EM algorithm may be subject to local minima problems. However, the introduction of a-priori information on the sequence $x_t$ via the decoder output significantly reduces the likelihood of local minima.

5. SIMULATIONS

Simulations have been performed on a type B Proakis channel, i.e., $H = [0.407 \ 0.814 \ 0.407]$. The main aim of our simulations is to compare our proposed scheme to a turbo equalizer using the IC (DFE is used for the first iteration).

The performance of the proposed turbo equalizer in terms of bit error rate is shown in figures 4 (equaliser output) and 3 (decoder output). The performance of the IC based turbo equaliser can be seen in figure 5 and 6. One can note that the turbo effect is observed for lower SNRs when using our proposed scheme, compared to the IC equaliser. One important observation to be made is that the IC canceller seems to outperform our proposed equaliser seems at the first iteration. However, it is not true for subsequent iterations which seems to indicate that the lower performance of our proposed equaliser at the first iteration is compensated by the fact that the exchange

![Decoder output (proposed method)](image)

Fig. 3. Decoder output-proposed method

![Output of proposed equaliser](image)

Fig. 4. Output of proposed equaliser

of information between the equaliser and decoder seems much more efficient when using an MAP based equaliser. It would be interested to check this particular point using an EXIT chart.

Another important feature of our proposed method is the channel estimation scheme. Most existing turbo equalisers, such as the IC or DFE assume the knowledge of the channel. Our proposed method allows to relax this requirement as shown in figure 7. The parameters of the channel have been initialised to $H = [0.330 \ 0.330 \ 0.330]$ and $E_b/N_0$ is set to 9dB. One can see that the channel estimate converges after only 4 iterations approximately to the true values of the channel.

6. CONCLUSION AND FURTHER COMMENTS

In this paper, we have proposed a new turbo equaliser which is based on a reduced complexity MAP equaliser. When compared to the IC based turbo equaliser, a significant improvement has been observed both in terms of robustness
channel is unknown. This proposed method offers many directions for future research: indeed, the EM algorithm here used to estimate the channel coefficients could also be used to estimate the noise variance which is an important parameter in the probability computations used in the turbo process. It would be interesting to see how using a noise variance estimate instead of the true one affects the turbo process.

REFERENCES