Iterative Channel Estimation for GPRS

Nikolai Nefedov[1,2], Markku Pukkila[1]

[1]Nokia Research Center, P.O.Box 407, Nokia Group FIN-00045, Finland

[2]Helsinki University of Technology, Comm.Lab./IDC, Finland

email: nikolai.nefedov@nokia.com; markku.pukkila@nokia.com

ABSTRACT – In this paper we consider iterative estimation and equalization techniques and present a simple method of updating channel estimates that includes decoder outputs into the iteration process. To clarify performance-complexity trade-off we evaluated iterative estimation and turbo equalization techniques in General Packet Radio System (GPRS). It is found that turbo estimation is more beneficial for GPRS, showing 1dB gain after one channel estimate update, while turbo equalization rounds provide only a slight improvement on the top of that.

I. INTRODUCTION

One of important problems in reliable data communications over frequency selective fading channels is the mitigation of intersymbol interference (ISI). Recently suggested turbo equalization [1] is an iterative equalization technique for coded data that has gained a lot of interest [2,3,4]. This technique performs an iterative ISI removal relaying on channel estimates that usually are obtained based on a known training sequence. In practice channel estimates may have rather poor quality that in turn deteriorates the efficiency of equalization. On the other hand, the problem of joint data and channel estimation has been addressed in many publications with a number of different iterative techniques proposed, e.g. [5][6].

In this paper we present a method of updating channel estimates that includes decoder outputs into the iteration process similar to the turbo equalization. In particular, the decoded symbols are fed back into the channel estimator and the estimate is updated assuming that data now are known at receiver. Based on that we considered different iterative estimation and equalization scenarios for General Packet Radio System (GPRS).

The paper is organized as follows. Section II gives a formal problem description and an overview of conventional receivers. In Section III we briefly describe the turbo equalization technique and apply this method for iterative channel estimation. In the same section we present a lower bound on the variance of iterative estimate. Trade-off between the performance gain and receiver complexity in GPRS under different scenarios is addressed in Section IV, with conclusions following in Section V.

II. PRELIMINARIES

A. Problem formulation

Let's consider binary data transmitted in blocks $\mathbf{u}=(u_0, u_1, ...u_{K-1})^T$, $\mathbf{u} \in \mathbb{Z}^K$ over a channel with memory L in presence of Gaussian noise $\mathbf{w}$, i.e. $w_n = \mathcal{N}(0, \mathbf{R})$, $\mathbf{R}$ is noise covariance matrix. For channels with additive white Gaussian noise (AWGN) $\mathbf{R}=\sigma^2\mathbf{I}$, where noise variance $\sigma^2 = N_0/2$. We assume that channel impulse response (CIR) is unknown and characterized by complex channel taps $\mathbf{h}_L=(h_0, h_1, ..., h_L)^T$. The binary data $\mathbf{u}$ are encoded by a code, such that $\mathbf{c} = \Xi\mathbf{u}$, $\mathbf{c} \in \mathbb{Z}^N$. Additionally, some interleaving scheme may be used to destroy possible error clusters at receiver side. To assist channel estimation a known training sequence $\mathbf{m}$ is inserted in each coded block $\mathbf{c}$. The training sequence $\mathbf{m}$ consists of $L + P$ symbols, with L preamble and P midamble symbols; $\mathbf{m}=(m_0, m_1, ..., m_{(P+L-1)})^T$. The resulting data block $\mathbf{a}$ of length N_b is arranged as follows: $\mathbf{a} = \begin{bmatrix} \mathbf{c}_1^T \mathbf{m}^T \mathbf{c}_2^T \end{bmatrix}^T$, where $\mathbf{c}_i = (c_0^{(i)}, c_1^{(i)}, ...c_{N_d-1}^{(i)})^T$ are coded data symbols separated into sub-blocks; $N_b = (2N_d + L + P)$. Finally the data block $\mathbf{a}$ is mapped into M-ary symbols and transmitted with the normalized symbol energy $E_s=1$.

According to maximum likelihood (ML) criteria the optimal receiver is to find

$$\hat{\mathbf{u}}= \arg\max_{\mathbf{h},\mathbf{u}} p(\mathbf{r}|\mathbf{h}, \mathbf{a})=\arg\max_{\mathbf{h},\mathbf{u}} p(\mathbf{r}|\mathbf{h}, \mathbf{m},\Xi\mathbf{u}) \qquad (1)$$

The optimal solution of (1) is prohibitively complex, and in practice the general problem (1) is split into several ones which are then considered separately. Separating channel equalization (detection) and decoding, and taking into account that the training sequence is known, a suboptimal solution for (1) may be presented as

$$\hat{\mathbf{a}}= \arg\max_{\mathbf{h},\mathbf{a}} p(\mathbf{r}|\mathbf{h}, \mathbf{a}) \Longrightarrow \hat{\mathbf{c}}=\arg\max_{\mathbf{h},\mathbf{c}} p(\mathbf{r}|\mathbf{h}, \mathbf{c}) \qquad (2)$$

$$\hat{\mathbf{u}} = \arg\max_{\mathbf{u}} p(\hat{\mathbf{c}}|\mathbf{u}) \qquad (3)$$

The optimal solution requires a search over all possible $\mathbf{c}$ and $\mathbf{h}$ that is impractical for realistic values of L, N_d. A typical suboptimal solution of (2) is to separate channel estimation and equalization that leads to

$$\hat{\mathbf{h}} = \arg\max_{\mathbf{h}} p(\mathbf{r} \mid \mathbf{m}, \mathbf{h}) \qquad (4)$$

$$\hat{\mathbf{c}} = \arg\max_{\mathbf{c}} p(\mathbf{r} \mid \hat{\mathbf{h}}, \mathbf{c}) \qquad (5)$$

For example, the first the channel estimate $\hat{\mathbf{h}}$ in GSM is calculated based on the known training sequence

m. Then one of the equalization algorithms is applied to remove ISI and obtain $\hat{c}$. Finally, in case of coded data a decoder recovers transmitted information $\hat{u}$. Below we consider the estimation (4) and data detection (5) in more details.

B. Channel estimation and sequence detection

Assuming a linear channel with time-invariant CIR during the transmitted block, the received block can be presented as

$$\mathbf{r} = \mathbf{Ah} + \mathbf{w} = \begin{bmatrix} \mathbf{r_{c_1}} \\ \mathbf{r_m} \\ \mathbf{r_{c_2}} \end{bmatrix} \text{ where } \mathbf{A} = \begin{bmatrix} \mathbf{A_1} \\ \mathbf{M} \\ \mathbf{A_2} \end{bmatrix} \quad (6)$$

$\mathbf{A}$ a block matrix of size $N_b \times (L+1)$; $\mathbf{M}$ is a $P \times (L+1)$ matrix formed by the training sequence $\mathbf{m}$ as follows

$$\mathbf{M} = \begin{bmatrix} m_L & m_{L-1} & \cdots & m_1 & m_0 \\ m_{L+1} & m_L & \cdots & m_2 & m_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ m_{P+L-1} & m_{P+L-2} & \cdots & m_P & m_{P-1} \end{bmatrix}$$

$$\mathbf{A_1} = \begin{bmatrix} c_L^{(1)} & c_{L-1}^{(1)} & \cdots & c_2^{(1)} & c_1^{(1)} & c_0^{(1)} \\ c_{L+1}^{(1)} & c_L^{(1)} & \cdots & c_3^{(1)} & c_2^{(1)} & c_1^{(1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ m_{L-2} & m_{L-3} & \cdots & m_0 & c_{N_d-1}^{(1)} & c_{N_d-2}^{(1)} \\ m_{L-1} & m_{L-2} & \cdots & m_1 & m_0 & c_{N_d-1}^{(1)} \end{bmatrix}$$

$$\mathbf{A_2} = \begin{bmatrix} c_0^{(2)} & m_{P+L-1} & \cdots & m_{P+1} & m_P \\ c_1^{(2)} & c_0^{(2)} & \cdots & m_{P+2} & m_{P+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{N_d-2}^{(2)} & c_{N_d-3}^{(2)} & \cdots & c_{N_d-L}^{(2)} & c_{N_d-L-2}^{(2)} \\ c_{N_d-1}^{(2)} & c_{N_d-2}^{(2)} & \cdots & c_{N_d-L-2}^{(2)} & c_{N_d-L-1}^{(2)} \end{bmatrix} \quad (7)$$

In a conventional receiver the channel estimation is based on midamble P symbols from the received sequence $\mathbf{r_m} = \mathbf{Mh} + \mathbf{w}$. For a channel with Gaussian noise the likelihood function

$$p(\mathbf{r_m}|\mathbf{m}, \mathbf{h}) \sim \exp(-\tfrac{1}{2}(\mathbf{r_m} - \mathbf{Mh})^H \mathbf{R}^{-1}(\mathbf{r_m} - \mathbf{Mh}))$$

that gives ML channel estimate [7] as

$$\hat{\mathbf{h}}_{ML} = \arg \max_{\mathbf{h}} p(\mathbf{r_m}|\mathbf{M}, \mathbf{h}) = \mathbf{C}(\hat{\mathbf{h}}_{ML})\mathbf{M}^H \mathbf{R}^{-1}\mathbf{r_m} \quad (8)$$

where $\mathbf{C}(\hat{\mathbf{h}}_{ML}) = (\mathbf{M}^H \mathbf{R}^{-1}\mathbf{M})^{-1}$ is a covariance matrix of the estimate. For the given case the ML channel estimate (8) is the minimum variance unbiased (MVU) channel estimate based on the data $\mathbf{r_m}$.

Given $\hat{\mathbf{h}}$, the ML sequence detection is made by maximizing the likelihood function with respect to $\mathbf{c}$

$$\hat{\mathbf{c}} = \arg \max_{\mathbf{c}} p(\mathbf{r}|\hat{\mathbf{h}}, \mathbf{c})$$

that can be implemented via M^L-state Viterbi algorithm.

III. ITERATIVE (TURBO) ESTIMATION

In many cases even after splitting (1) into (2) and (3), the complexity/performance trade-off of (2) remains unacceptable, and that stimulates the use of different forms of decision feed back equalizers (DFE). One of the solutions that is closely related to DFE class is so-called turbo equalization method [1]. Instead of solving (2), this method tries to find iteratively

a solution for $\hat{u}$ over a combined trellis formed by channel and encoder, i.e.

$$\hat{u} = \arg \max_{\mathbf{u}} p(\mathbf{r}|\hat{\mathbf{h}}, \mathbf{m}, \Xi(\mathbf{u})) = \arg \max_{\mathbf{u}} p(\mathbf{r}|\hat{\mathbf{h}}, \mathbf{a}). \quad (9)$$

In the turbo equalization scheme (9) the iteration proceeds only between the signal detector and channel decoder assuming a known channel state $\hat{\mathbf{h}}$ during iterations. Given a known training sequence $\mathbf{m}$, the channel estimate may be obtained by (8) based on the data $\mathbf{r_m}$.

However, in many cases the accuracy of channel estimate, which is based only on a relatively short training sequence $\mathbf{m}$, may be rather low. That in turn may cause a significant performance degradation at the receiver that cannot be fully compensated by the turbo equalization. This fact motivated us to use a decision-directed adaptive channel estimation method during the iteration process similar to [5]. The idea is to feed back the decoded symbols to the channel estimator and update previous channel estimates assuming that the whole burst is now known by the receiver (Fig.1). Hence, the receiver relies on the decoded data symbols $\hat{u}$ and the known training sequence $\mathbf{m}$ and forms a new channel estimate. In other words, the receiver iteratively updates the channel estimate based on the "extended" training sequence. In particular, after decoding procedure the data $\hat{u}$ are re-encoded as $\check{c} = \Xi\hat{u}$ and then combined with the training sequence $\mathbf{m}$, forming a new "extended" training sequence $\check{a}$ of length $N_b = P + L + 2N_d$. If we would use all available data $\mathbf{a}$ as the known training sequence, then in AWGN channel the ML channel estimate is

$$\hat{\mathbf{h}}^{extend} = \mathbf{C}(\hat{\mathbf{h}}^{extend})\mathbf{A}^H \mathbf{r} \quad (10)$$

where the covariance matrix of the new "extended" estimate is

$$\mathbf{C}(\hat{\mathbf{h}}^{extend}) = (\mathbf{A}^H \mathbf{A})^{-1}$$
$$= (\mathbf{A_1}^H \mathbf{A_1} + \mathbf{M}^H \mathbf{M} + \mathbf{A_2}^H \mathbf{A_2})^{-1} \quad (11)$$

and matrix $\mathbf{A}$ is defined by (6),(7).

To avoid heavy computation of the matrix inverse in (11) we can use some adaptive algorithm to update the estimate. In this paper we applied stochastic adaptation of the estimate based on the LMS algorithm [7]

$$\hat{\mathbf{h}}^{(k+1)} = \hat{\mathbf{h}}^{(k)} - \mu(\check{\mathbf{A}}^{(k)})^H(\check{\mathbf{A}}^{(k)}\hat{\mathbf{h}}^{(k)} - \mathbf{r}) \quad (12)$$

where $\hat{\mathbf{h}}^{(k)}$ is the estimate from kth iteration, $\check{\mathbf{A}}^{(k)}$ is an estimated data matrix containing all (data+training) symbols known at kth iteration, $\mathbf{r}$ is the received vector and μ is a step size of the iterative algorithm.

At the initial round the channel estimate could be based on some conventional method, e.g. one-shot ML estimate (8), which exploits only the known training sequence.

It is useful to evaluate the quality of the new "extended" estimate. For that purpose let's consider AWGN channel with noise samples $w_n = \mathcal{N}(0, \sigma_n^2)$ and a constant variance during the received block $\mathbf{r}$, i.e. $\sigma_n^2 = \sigma^2, n = 1...N_b$. The variance of the estimate $\hat{\mathbf{h}}$ based only on the training sequence $\mathbf{m}$ may

be bounded by Cramer-Rao lower bound (CRLB) [7] as $var(\hat{h}_i) \geqslant [\mathbf{I}(\mathbf{h})]_{ii}^{-1}$, where $\mathbf{I}(\mathbf{h}) = \{I_{ij}\}$ is the $(L+1) \times (L+1)$ Fisher information matrix with elements

$$I_{ij} = \frac{\partial^2 p(\mathbf{r_m}|\mathbf{m}, \mathbf{h})}{\partial h_i \partial h_j}.$$

In our case CRLB may be written as

$$var(\hat{h}_i) \geqslant [\mathbf{M}^H \mathbf{R}^{-1} \mathbf{M}]_{ii}^{-1} \approx \frac{1}{\sum_{n=1}^{P} \frac{1}{\sigma_n^2}} = \frac{\sigma^2}{P} \quad (13)$$

where $i=0...L$.

For the "extended" training sequence $\mathbf{r}$ it gives

$$var(\hat{h}_i^{extend}) \geqslant [\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A}]_{ii}^{-1}$$
$$\approx \frac{1}{\sum_{n=1}^{N_b} \frac{1}{\sigma_n^2}} = \frac{1}{\frac{P}{\sigma^2} + \frac{2N_d}{\sigma_d^2}} = \frac{\sigma^2}{P + 2N_d \frac{\sigma^2}{\sigma_d^2}} \quad (14)$$

where σ_d^2 is the variance associated with the "extended" data provided by the decoder.

To evaluate (14) let's denote p_n as an error probability for nth symbol $\check{c}_n$ after decoding/re-encoding operations. The probabilities p_n are calculated by averaging over a number of received blocks. Now we can treat data $\check{c}$ forming the extension of training sequence as a set of random variables with the mean and the variance as follows

$$E[\check{c}_n] = (1 - p_n)c_n + p_n \bar{c}_n$$
$$var[\check{c}_n] = E[\check{c}_n^2] - E[\check{c}_n]^2$$
$$= c_n^2 - p_n(c_n^2 - \bar{c}_n^2) - (c_n - p_n(c_n - \tilde{c}_n))^2,$$

where $\bar{c}_k$ stands for $\bar{c}_n \neq c_n$.

Assuming a time-invariant channel during the transmitted block (i.e. $p_n = p$, $n=0,...,N_b-1$) and the antipodal signalling ($c_n = 1$, $\bar{c}_n = -1$) it results in

$$var[\check{c}_n] = 4p(1 - p). \quad (15)$$

The variance of the "extended" data is

$$\sigma_d^2 = var[\check{c}_n + w_n] = var[\check{c}_n] + var[w_n] + 2cov[\check{c}_n w_n] \quad (16)$$

Taking into account that more errors appear after decoding/re-encoding operations at high noise levels (i.e. larger σ^2 results in larger p, and hence, in large $var[\check{c}_n]$), hence the correlation term $cov[\check{c}_n w_n] \geqslant 0$ and $\sigma_d^2 \geqslant var[\check{c}_n] + var[w_n]$. As can be seen from (16) the CRLB (14) is a function of $cov[\check{c}_n w_n]$, and it achieves its minimum $var(\underline{\hat{h}}_i^{extend})$ if $cov[\check{c}_n w_n] =0$. Based on (14) and (15) this minimum may be presented as

$$var(\underline{\hat{h}}_i^{extend}) \geqslant \frac{\sigma^2}{P + \frac{2N_d \sigma^2}{4p - 4p^2 + \sigma^2}} \quad (17)$$

As an illustration the bounds (13),(17) are visualized at Fig.2 for parameters $P=20$ and $N_d=58$ accepted in GSM. As can be seen from (14), in case of MVU channel estimator the variance of the "extended" estimate is always lower than one calculated only from the training sequence, i.e. $var(\underline{\hat{h}}_i^{extend}) \leqslant var(\hat{h}_i^{extend}) < var(\hat{h}_i)$ for $p > 0$. That can be explained by an observation that by extending training sequence even with unreliable symbols, in average we make covariance matrix (11) more diagonal

dominant, that finally improve the channel estimate. Another point to mention is that the gain from the "extended" training sequence is mainly visible at low signal/noise ratios (SNR) and practically disappears at high SNR where initial estimate is already rather accurate.

IV. SIMULATION RESULTS

To find an efficient way to perform iterations described above it is necessary to study the trade-off between receiver complexity and the performance gain provided by different iteration scenarios. Turbo equalization method is based on soft decisions provided by the decoder and that itself significantly (2-4 times) increases the decoder complexity. On the other hand, the turbo-estimation method in the form presented above does not require soft decisions and hence, modifications of the decoder. A simple channel update based on an adaptation rule allows to avoid complex calculations associated with the matrix inverse. Additionally, it would be useful to evaluate the potential gain that soft values from the decoder could provide, if available, in case of turbo-estimation. Furthermore, to reduce complexity increase due to iterations, a less complex receiver also should be considered, where iterative channel estimate is based only on detected (instead of decoded) symbols.

As a practical testbed we considered performance of different iterative receivers for GPRS in typical mobile radio channels. In this paper we present simulation results for the strongest channel coding scheme offered by GPRS, that is CS1, which employs $\frac{1}{2}$-rate convolutional coding with constraint length 5 without any puncturing [8]. The interleaving used in GPRS is rectangular over 4 bursts. We used ML estimator (8) with the LMS adaptation rule (12) and equalizers with 5 and 6-taps. Quality of service in packet transmission is characterized by block error rate (BLER), but bit error rate (BER) is also shown at performance figures.

The performance figures are presented for a) the conventional receiver, b) after iterative channel estimate update, c) after turbo equalization iterations, where the channel estimate is no more changed. The turbo estimation gain given in the texts below denotes the gain after the first channel estimate update.

First we evaluate performance of turbo-estimation and turbo equalization schemes in a test channel with symmetrical CIR and channel memory $L=4$. Results presented at Fig.3 show that the first channel estimate update gives 1.0 dB, and further turbo iterations with the improved estimate are providing 0.7 dB at BLER 10^{-2}, giving the total gain of about 2 dB. The turbo equalization on the top of that gives only 0.3 dB gain.

For more realistic evaluation we used typical urban channel with speed $v=3$km/h (TU3) and hilly terrain channel, $v=100$km/h (HT100). The performance of turbo estimation receiver over TU3 is depicted at Fig.4. The gain achieved after one channel estimate

update is around 0.9 dB at BLER=10^{-2}, which is already most of the achievable gain compared to a case with the known channel states. In other words, a receiver with the ideal channel information gives only 0.2 dB gain compared to the receiver with one channel estimate update. The turbo equalization is able to provide only 0.3 dB extra gain on top of that.

Fig.5 and 6 present the performance in HT100 channel in case of 5-tap and 6-tap equalizers, respectively. These figures reveal that the turbo estimation scheme and the turbo-equalization on the top of that are less efficient if the whole channel impulse response does not fit into the channel estimator and detector windows. This is due to the long delay spread of the HT channel, which requires at least 6 channel taps to be estimated and detected at GSM symbol resolution. Hence, the gain is moderate (0.5 dB at BLER=10^{-2}) with 5-tap equalizer, and it is increased to 1.0 dB if one extra channel tap is taken into account. Once more, the turbo equalization does not noticeably improve the performance in this case.

The turbo estimation scheme presented above used decoded bits in the channel estimate updating. We also considered a less complex solution, which utilizes hard decisions already from the detector output. Here we avoided interleaving and decoding procedures, but for the price of less reliable decisions in the feedback. Dashed lines at Fig.7 show the performance in TU3 channel for the case, when the detected bits are used in the channel estimator. In that case the gain from the iterative estimation is only 0.3 dB with 0.4 dB extra gain provided by the turbo equalization on the top of that.

V. DISCUSSIONS AND CONCLUSIONS

In this paper we presented a scheme, where similar to turbo equalization technique, we iteratively update channel estimates relying on data after decoder. We studied the trade-off between the receiver complexity and the performance gain provided by different receivers utilizing iterative channel estimation and turbo equalization techniques. It was found that turbo-estimation is beneficial mainly at low SNR, or/and in channels with large delay spread, provided that length of estimation and detection windows fit to the channel response.

We evaluated iterative estimation and equalization techniques in GPRS, where turbo equalization seems more attractive compared to speech services because of used rectangular interleaver. Nevertheless, it is found that in GPRS it is more beneficial to update channel estimate during the iterative process than to use more complex turbo equalization technique. The simulation results show almost 1 dB gain in TU3 and HT100 channels after only one channel estimate update, and only a slight improvement is observed by further turbo equalization rounds. The iterative channel estimation causes only a minor complexity increase in channel estimator, so the conventional detector and channel decoder can be used. The up-

dating of channel estimate may be done by an adaptive procedure that does not require the calculation of the matrix inverse. Also no practical need was found for soft decisions from the decoder unless turbo equalization technique is applied. We also considered a less complex solution, which exploits the detected bits instead of decoded in the channel state updating and thus avoids interleaving and decoding procedures. However this solution seems not to be the best one from performance-complexity point of view.

REFERENCES

[1] C. Douillard et al, "Iterative Correction of Inter-symbol Interference: Turbo-Equalization", *European Trans. on Telecomm.*, v. 6, No. 5, Sep-Nov 1995, pp. 507-511.

[2] A. Picart, P. Didier, A. Glavieux, "Turbo-Detection: A new approach to combat channel frequency selectivity", *ICC'97*, pp.1498-1502.

[3] G. Bauch and V. Franz, "Iterative Equalization and Decoding for the GSM-System," *VTC'98*, May 1998, pp. 2262-2266.

[4] P. Strauch et al, "Turbo Equalization for an 8-PSK Modulation Scheme in a Mobile TDMA Communication System", *VTC'99*, pp. 1605-1609.

[5] K-H. Chang, C.N.Georghiades, "Iterative Joint Sequence and Channel Estimation for Fast Time-Varying Intersymbol Interference Channels", *ICC'95*, pp.357-361, Seattle 1995.

[6] N. Seshardi, "Joint Data and Channel Estimation Using Blind Trellis Search Techniques", *IEEE Trans.Comm.*,v.42,No.2/3/4, 1994, pp.1000-1011

[7] S. Haykin, *Adaptive Filter Theory*, 3d ed., Prentice Hall, 1996.

[8] ETSI GSM05.03 *Digital cellular telecommunications system (Phase 2+).Channel coding, v.7*, release 1999.
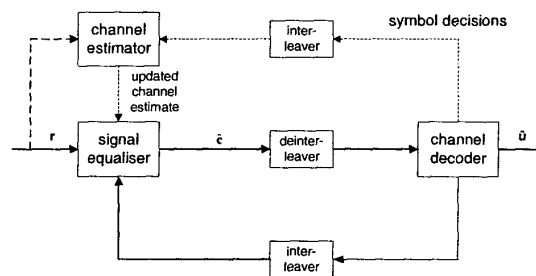
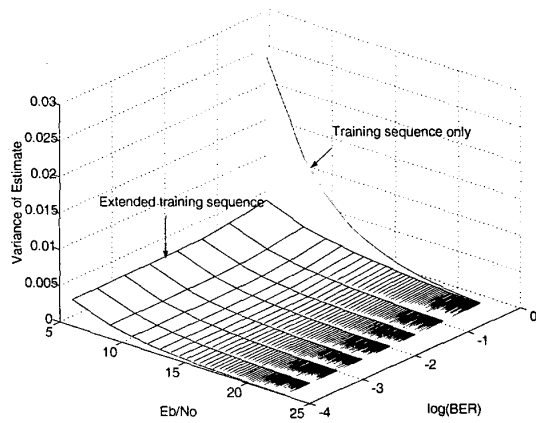Fig.1. Block diagram receiver with iterative estimation-equalization.
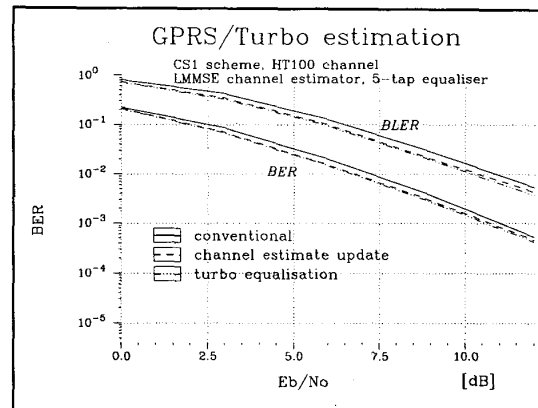
Fig.2. Variance of estimate
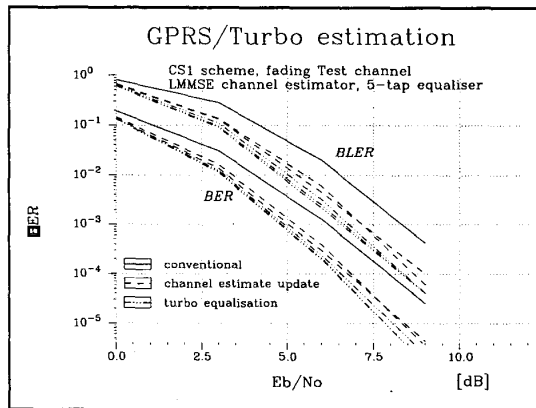

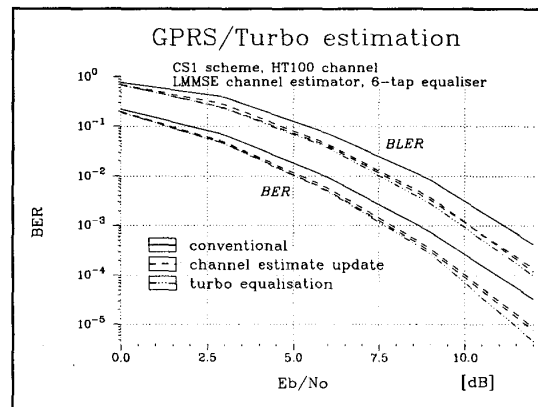Fig.5 HT100 channel, 5-tap equalizer

Fig.3. Test channel

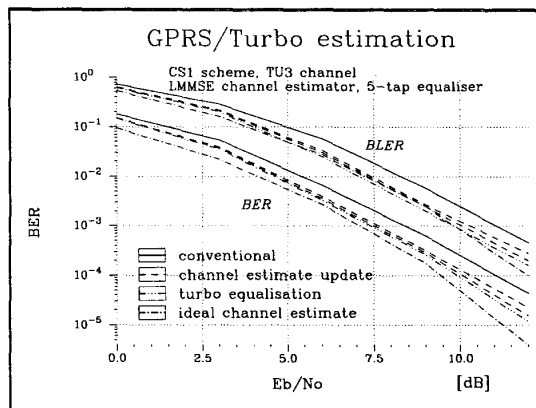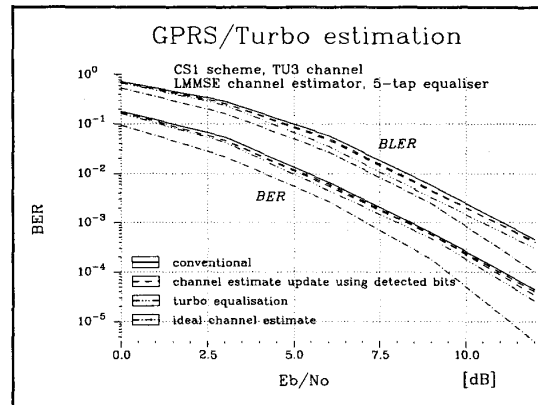Fig.6. HT100 channel, 6-tap equalizer.

Fig.4. TU3 channel, 5-tap equalizer.

Fig.7 TU3 channel. Iterative channel estimate and
turbo-equalization.