JOINT CHANNEL ESTIMATION AND DATA DETECTION USING A BLIND BAYESIAN DECISION FEEDBACK EQUALISER

*S. Chen *, S. McLaughlin, P.M. Grant and B. Mulgrew*
Department of Electrical Engineering, University of Edinburgh,
King's Buildings, Edinburgh EH9 3JL, UK
* Dept of Electrical Engineering, University of Portsmouth

1. Introduction

Since the pioneering work of Sato [1], three main types of blind equalisers for nonminimum phase channels have ben developed. The first family of blind adaptive algorithms, often referred to as Bussgang algorithms, constructs a transversal equaliser by optimising some nonconvex criterion function using a gradient algorithm [1]-[7]. A Bussgang-type blind equaliser typically has very low computational complexity but suffers from the drawback of slow convergence. The second family of blind adaptive algorithms identifies the channel impulse response using techniques based on higher order cumulants (HOCs) [8]-[14] and uses the resulting channel model to design an equaliser. This second class of blind equalisers, although very general and powerful, requires a large number of data samples and extensive computation to estimate HOCs. Recently, blind equalisers based on joint channel and data estimation have been proposed [15]-[20]. This third family of blind adaptive algorithms typically employs some blind approximation of maximum likelihood sequence estimation or its variants. The resulting blind equalisers are therefore computationally very expensive. However, a major advantage of this third approach is that relatively few signal samples are required.

In this paper a blind implementation of the Bayesian symbol-by-symbol DFE [21]-[23] for joint data and channel estimation is derived. A DFE consists of a feedforward section and a feedback section. If the feedforward section contains m channel output samples and the size of the transmitted symbol constellation is M, there are M^m symbol combinations for the length of the feedforward section. At each sample instant, each of these M^m symbol sequences can be used to produce an LMS/RLS channel estimate. Each "conditional" channel estimate is employed to design a Bayesian DFE for symbol detection. The best Bayesian DFE in terms of a posterior probability density function (p.d.f.) is then chosen from the M^m "conditional" DFEs, and its detected symbol is fed back to the equaliser feedback section and used to update an "unconditional" channel estimate. These operations form a basic unit of the proposed blind equaliser. This blind equaliser can be expanded to include several such units, each covering an estimated initial condition. The performance of each unit is monitored and those units which perform poorly can then be switched off. The proposed blind equaliser is conceptually very simple and its total computational load is naturally decomposed into many simple and identical components, leading to an efficient parallel implementation. Simulation results are included to demonstrate its fast convergence property.

Throughout this study, the channel and symbols are assumed to be real-valued. This corresponds to the use of multilevel pulse amplitude modulation scheme (*M-ary* PAM). For the complex-valued channel and modulation schemes such as quadrature amplitude modulation, the derivation of the proposed blind equaliser is similar to the current real case.

Specifically, the channel is modelled as a finite impulse response filter with a transfer function

$$A(z) = \sum_{i=0}^{n_a-1} a_i z^{-i}, \tag{1}$$

where n_a is the length of the channel impulse response and a_i are the channel tap weights. The symbol sequence $\{s(k)\}$ is independently identically distributed (i.i.d.) and has an *M-ary* PAM constellation defined by the set

$$s_i = 2i - M - 1, \ \ 1 \le i \le M. \tag{2}$$

The received signal is given by

$$r(k) = \hat{r}(k) + e(t) = \sum_{i=0}^{n_a-1} a_i s(k-i) + e(k), \tag{3}$$

where $\hat{r}(k)$ is the noiseless channel output and $e(k)$ is an i.i.d. Gaussian noise with zero mean and variance $E[e^2(k)] = \sigma_e^2$.

2. The Bayesian decision feedback equaliser

The structure of a generic DFE is depicted in Fig.1. The equalisation process defined in Fig.1 uses the information present in the observed channel output vector and the past detected symbol vector to produce a delayed estimate of the transmitted symbol. The three important structural parameters of the equaliser are the decision delay d, the feedforward order m and the feedback order n respectively. The feedforward order is usually related to the decision delay by $m = d + 1$ and the feedback order is given by $n = n_a + m - d - 2 = n_a - 1$. In practice, $d = n_a - 1$ is often chosen to cover the entire channel dispersion.

This work was supported by the UK Science and Engineering Research Council under award GR/G/72380. SM gratefully acknowledges financial support from the UK Royal Society.

Printed and published by the IEE, Savoy Place, London WC2R 0BL, UK

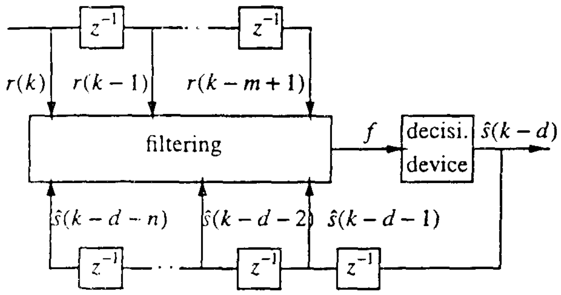

Fig.1 Schematic of a generic decision feedback equaliser.

Given the channel in (1), the value of the noiseless channel output vector

$$\hat{\mathbf{r}}(k) = [\hat{r}(k) \cdots \hat{r}(k-m+1)]^T \quad (4)$$

is specified by the symbol sequence $\mathbf{s}(k) = [\mathbf{s}_f^T(k)\ \mathbf{s}_b^T(k)]^T$, where

$$\mathbf{s}_f(k) = [s(k) \cdots s(k-d)]^T \quad (5)$$

and

$$\mathbf{s}_b(k) = [s(k-d-1) \cdots s(k-d-n)]^T. \quad (6)$$

Under the assumption that the given feedback vector is correct, that is, $\hat{\mathbf{s}}_b(k) = \mathbf{s}_b(k)$, the state of $\hat{\mathbf{r}}(k)$ is determined by $\mathbf{s}_f(k)$. For the *M-ary* PAM constellation, $\mathbf{s}_f(k)$ has $N_s = M^{d+1} = M^m$ combinations and, therefore, $\hat{\mathbf{r}}(k)$ has N_s states. The states of $\hat{\mathbf{r}}(k)$ can be grouped into M sets according to the value of $s(k-d)$:

$$R_f^{(i)} = \{\hat{\mathbf{r}}(k) = \mathbf{r}_j^{(i)} | s(k-d) = s_i\}, \quad 1 \le i \le M. \quad (7)$$

Each $R_f^{(i)}$ contains $N_s^{(i)} = N_s/M = M^d$ states.

The Bayesian DFE [21],[23] can now be summarised. The p.d.f. of $\mathbf{r}(k) = [r(k) \cdots r(k-m+1)]^T$ conditioned on $s(k-d) = s_i$ is

$$p_r(\mathbf{r}(k)|s(k-d) = s_i) = \sum_{j=1}^{N_s^{(i)}} \alpha_j^{(i)} p_e(\mathbf{r}(k) - \mathbf{r}_j^{(i)}), \quad 1 \le i \le M, \quad (8)$$

where $\mathbf{r}_j^{(i)} \in R_f^{(i)}$, $\alpha_j^{(i)}$ are the a-priori probabilities of $\mathbf{r}_j^{(i)}$ and $p_e(\cdot)$ is the p.d.f. of the noise vector $\mathbf{e}(k) = [e(k) \cdots e(k-m+1)]^T$. Since all the channel states can be assumed to be equiprobable and the noise p.d.f. is Gaussian, (8) leads to the M Bayesian decision variables

$$\eta_i(k, \mathbf{a}) = \sum_{j=1}^{N_s^{(i)}} \exp(-\| \mathbf{r}(k) - \mathbf{r}_j^{(i)} \|^2 / 2\sigma_e^2), \quad 1 \le i \le M. \quad (9)$$

Here $\mathbf{a} = [a_0\ a_1 \cdots a_{n_a-1}]^T$ is included in the expression of Bayesian decision variables to emphasise that the channel states are computed based on the given channel $\mathbf{a}$. The minimum-error-probability decision is defined by

$$\hat{s}(k-d) = s_{i^*} \ \text{ if } \ \eta_{i^*}(k, \mathbf{a}) = \max\{\eta_i(k, \mathbf{a}),\ 1 \le i \le M\}, \quad (10)$$

which provides the optimal solution for the equalisation structure of Fig.1.

3. Joint channel estimation and symbol detection

When the channel is unknown and no training period is provided, joint channel estimation and symbol detection can be performed based on a blind implementation of the Bayesian DFE. The basic idea is to identify the N_s "conditional" channel estimates using the N_s sequences of $\mathbf{s}_f(k)$ and to design the N_s corresponding Bayesian DFEs. The detected symbol is chosen to be the best solution of these N_s "conditional" DFEs. Specifically, at sample k, given the feedback vector $\hat{\mathbf{s}}_b(k)$, an "unconditional" channel estimate $\hat{\mathbf{a}}(k-1-d)$ and an estimated MSE $\sigma_\varepsilon^2(k-1-d)$, the operations of the blind equaliser are as follows:

(i) N_s "conditional" normalised LMS (NLMS) channel estimators update N_s "conditional" channel estimates. Given $\hat{\mathbf{a}}_l(k-1-d) = \hat{\mathbf{a}}(k-1-d)$, the lth estimator forms $\hat{\mathbf{a}}_l(k)$ from $[\mathbf{s}_{f,l}^T(k)\ \hat{\mathbf{s}}_b^T(k)]^T$ and $\mathbf{r}(k)$, where $\mathbf{s}_{f,l}(k)$ is the lth sequence of $\mathbf{s}_f(k)$.

(ii) For each $\hat{\mathbf{a}}_l(k)$, a Bayesian DFE is designed with the required noise variance σ_e^2 being substituted by $\sigma_\varepsilon^2(k-1-d)$. The lth "conditional" Bayesian DFE provides a tentative decision $\hat{s}^{(l)}(k-d) = s_{i_l}$.

(iii) The detected symbol $\hat{s}(k-d)$ is then chosen to be the best solution of the N_s tentative decisions $\hat{s}^{(l)}(k-d)$, $1 \le l \le N_s$.

(iv) Given $r(k-d)$ and $[\hat{s}(k-d) \cdots \hat{s}(k-d-n_a+1)]^T$, $\hat{\mathbf{a}}(k-1-d)$ is updated to $\hat{\mathbf{a}}(k-d)$ using the NLMS algorithm. The estimated MSE is adjusted according to

$$\left.\begin{aligned} \varepsilon(k-d) &= r(k-d) - \sum_{i=0}^{n_a-1} \hat{a}_i(k-d)\hat{s}(k-d-i), \\ \sigma_\varepsilon^2(k-d) &= (1-\mu_\varepsilon)\sigma_\varepsilon^2(k-1-d) + \mu_\varepsilon \varepsilon^2(k-d), \end{aligned}\right\} \quad (11)$$

where $0 < \mu_\varepsilon < 1$ is an adaptive gain.

The complexity of the above blind adaptive equaliser depends on $N_s = M^m$, and the steps (i) and (ii) involve extensive computation. However, these two steps consist of N_s identical components and are suitable for parallel implementation. Although increasing the length m of $\mathbf{r}(k)$ generally speeds up the convergence of the adaptive algorithm, it is advisable to choose a modest value of m such as is normally used in the DFE.

3.1 Sign ambiguity of the channel estimate

Before discussing how to initialise the above blind equaliser, it is necessary to discuss the phenomenon known as sign ambiguity. When the adaptive algorithm converges, the channel estimate $\hat{\mathbf{a}}$ can converge either to $\mathbf{a}$ or $-\mathbf{a}$. This phenomenon is essentially due to the symmetry of the signal constellation and because there is insufficient information for the blind channel estimator to distinguish between $\mathbf{a}$ and $-\mathbf{a}$. This ambiguity problem is not unique to the present blind equaliser. All other existing blind adaptive algorithms based on joint channel and data estimation suffer the same problem.

3.2. Initialisation

Initial choices of $\sigma_\varepsilon^2(k-1-d)$, $\hat{\mathbf{s}}_b(k)$ and $\hat{\mathbf{a}}(k-1-d)$ at $k=0$ are required to start the blind adaptive process. The initial value of the estimated MSE does not have any serious effect on the performance of the blind equaliser, and $\sigma_\varepsilon^2(-1-d)$ can simply be set to a small positive value. A previous simulation study [23],[25],[26] has suggested that performance of the Bayesian equaliser is relatively insensitive to error in the estimated noise variance. A common choice for the initial feedback vector is $\hat{\mathbf{s}}_b(0)=[0\cdots0]^T$ with the initial channel estimate usually set to $\hat{\mathbf{a}}(-1-d)=[0\cdots0]^T$. However, given this zero initial estimate, the blind channel estimator tends to converge falsely to an "equivalent" minimum phase channel. To overcome this difficulty, one element of $\hat{\mathbf{a}}$ can be initialised to a non-zero value.

If the *i*th channel tap a_i is known to have the largest amplitude, setting $\hat{a}_i(-1-d)$ to 1.0 and the rest of $\hat{\mathbf{a}}(-1-d)$ to zeros is obviously a better initialisation strategy. Since the channel tap which has the largest amplitude is unknown, several possibilities must be tested. This suggests an expansion of the blind equaliser to include several units of the basic algorithm (steps (i) to (iv)). In theory, n_a units are needed to cover the n_a initial estimates, and the *i*th unit, $0 \le i \le n_a-1$, is given the initial channel estimate

$$\hat{a}_i(-1-d)=1,\ \hat{a}_j(-1-d)=0,\ 0\le j\le n_a-1 \text{ and } j\ne i. \quad (12)$$

In practice, not all of these candidates need to be examined. For example, statistically speaking, it is unlikely that the last channel tap a_{n_a-1} will have the largest amplitude. There are another n_a candidates, each having an initial channel estimate

$$\hat{a}_i(-1-d)=-1,\ \hat{a}_j(-1-d)=0,\ 0\le j\le n_a-1 \text{ and } j\ne i. \quad (13)$$

However, these n_a units need not be tested since the blind equaliser is incapable of distinguishing between $\mathbf{a}$ and $-\mathbf{a}$. If a unit converges, its channel estimate converges either to $\mathbf{a}$ or to $-\mathbf{a}$, and its estimated MSE σ_ε^2 will be significantly smaller than those of the units which do not achieve convergence. Thus those units which perform poorly in terms of estimated MSE can then be switched off. Similar expansion can be applied to the initial feedback vector. This enhances the reliability of the blind equaliser at the cost of increased complexity. A reduced constellation approach may be adopted to assign the initial choices of $\hat{\mathbf{s}}_b(0)$ by assuming that $\hat{s}(-d),\cdots,\hat{s}(-d-n_a+1)$ are binary. The number of initial choices can further be reduced by only assigning $\hat{s}(-d),\cdots,\hat{s}(-d-i)$, $i<n_a-1$, to nonzero values.

3.3. Simulation study

A simulated channel involving the 4-*ary* PAM symbol constellation is used to illustrate the behaviour of the proposed blind equaliser. In practice, the performance of the blind equaliser can only be observed through the estimated MSE (11). In simulation, the true performance of the blind equaliser can be assessed by the channel estimation error, which is defined as the following mean tap error (MTE)

$$\sigma_{\mathbf{a}}^2(k-d)=\|\pm\hat{\mathbf{a}}(k-d)-\mathbf{a}\|^2=\sum_{i=0}^{n_a-1}(\pm\hat{a}_i(k-d)-a_i)^2. \quad (14)$$

In the expression (14), $-\hat{\mathbf{a}}(k-d)$ is used if $\hat{\mathbf{a}}$ converges to $-\mathbf{a}$. Otherwise, $\hat{\mathbf{a}}(k-d)$ is used. The example used was a five-tap channel with the transfer function

$$A(z)=-0.205-0.513z^{-1}+0.719z^{-2}+0.369z^{-3}+0.205z^{-4}. \quad (15)$$

The three structure parameters of the Bayesian DFE were chosen to be $d=n_a-1=4$, $m=d+1=5$ and $n=n_a-1=4$. The noise variance was chosen as $\sigma_e^2=0.005$, giving rise to a SNR=30dB for the 4-*ary* PAM constellation. The NLMS algorithm had an adaptive gain of 0.1 while the adaptive gain for estimating the MSE was 0.02. Fig.2 depicts the estimated MSEs and the MTEs of the blind adaptive algorithm starting from different initial channel estimates (12), where the label I_i indicates that the nonzero element of the initial estimate is $\hat{a}_i(-1-d)$. From Fig.2, it can be seen that the blind adaptive unit I_2 achieved convergence. The channel estimate of this converged unit is plotted in Fig.3. Average performance of the blind adaptive unit I_2 over 10 different runs with $\hat{\mathbf{s}}_b(0)=[s(-5)\ s(-6)\ 0\ 0]^T$ is illustrated in Figs. 4 and 5.

The simulation results clearly demonstrate the fast convergence of the blind Bayesian DFE for joint channel estimation and symbol detection. For the 4-level symbol constellation, convergence was achieved in a few hundred symbols. In the case of the channel (15) with 4-*ary* PAM symbols, the convergence rate was observed to be less consistent in the different runs in comparison to 2-*ary* PAM simulation results not presented here. In some runs, the algorithm achieved convergence in less than 500 samples, while in other runs it needed 600 to 700 symbols to achieve convergence. Increasing the decision delay to $d=5$ and, consequently, $m=6$ will result in faster and more consistent convergence performance. However, this would result in a dramatic increase in computational complexity. For multiple signal levels, the estimation error fluctuates more violently compared with the binary case. Consequently, care must be exercised in the selection of the two adaptive gains.

4. Conclusions

A blind Bayesian decision feedback equaliser has been developed for joint channel estimation and symbol detection. It has been shown how the complete blind equaliser is built up with many identical adaptive units. Each of these units consists of a bank of simple least mean square channel estimators and Bayesian decision feedback equalisers. An efficient parallel implementation can therefore be realised readily. Simulation results have demonstrated fast convergence of this blind equaliser. Convergence can generally be achieved in less than 100 symbols when binary symbol constellation is used and within a few hundred symbols when a 4-level symbol constellation is used.

5. References

[1] Y. Sato, "A method of self-recovering equalisation for multilevel amplitude-modulation systems," *IEEE Trans. Communications*, Vol.COM-23, pp.679-682, 1975.

[2] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Communications*, Vol.COM-28, pp.1867-1875, 1980.

[3] J.R. Treichler and B.G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.ASSP-31, No.2, pp.459-472, 1983.

[4] G. Picchi and G. Prati, "Blind equalization and carrier recovering using a stop-and-go decision-directed algorithm," *IEEE Trans. Communications*, Vol.COM-35, pp.877-887, 1987.

[5] J.R. Treichler, "Application of blind equalization techniques to voiceband and RF modems," in *Preprints 4th IFAC Int. Symposium Adaptive Systems in Control and Signal Processing* (France), 1992, pp.705-713.

[6] J. Karaoguz and S.H. Ardalan, "A soft decision-directed blind equalization algorithm applied to equalization of mobile communication channels," in *Proc. ICC'92* (Chicago), 1992, Vol.3, pp.343.4.1-343.4.5.

[7] S. Chen, S. McLaughlin, P.M. Grant and B. Mulgrew, "Reduced-complexity multi-stage blind clustering equaliser," in *Proc. ICC'93* (Geneva), 1993, Vol.2, pp.1149-1153.

[8] K.S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes," *Ann. Statist.*, Vol.10, pp.1195-1208, 1982.

[9] G.B. Giannakis and J.M. Mendel, "Identification of nonminimum phase system using higher order statistics," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.ASSP-37, pp.360-377, 1989.

[10] H.-H. Chiang and C.L. Nikias, "Adaptive deconvolution and identification of nonminimum phase FIR systems based on cumulants," *IEEE Trans. Automatic Control*, Vol.AC-35, pp.36-47, 1990.

[11] D. Hatzinakos and C.L. Nikias, "Blind equalization using a tricepstrum-based algorithm," *IEEE Trans. Communications*, Vol.39, No.5, pp.669-682, 1991.

[12] A.G. Bessios and C.L. Nikias, "Blind equalization based on cepstra of power spectrum and tricoherence," in *Proc. SPIE* (San Diego), 1991, Vol.1565, pp.166-177.

[13] F.-C. Zheng, S. McLaughlin and B. Mulgrew, "Cumulant-based deconvolution and identification: several new families of linear equations," *EURASIP Signal Processing J.*, Vol.30, No.2, pp.199-219, 1993.

[14] F.-C. Zheng, S. McLaughlin and B. Mulgrew, "Blind equalization of nonminimum phase channels: higher order cumulant based algorithm," *IEEE Trans. Signal Processing*, Vol.41, No.2, pp.681-691, 1993.

[15] G. Kawas and R. Vallet, "Joint detection and estimation for transmission over unknown channels," in *Proc. Douzienne Colloque GRETSI* (France), 1989.

[16] M. Ghosh and C.L. Weber, "Maximum-likelihood blind equalization," in *Proc. SPIE* (San Diego), *1991*, Vol.1565, pp.188-195.

[17] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," submitted to *IEEE Trans. Communications*, 1991.

[18] K. Giridhar, J.J. Shynk and R.A. Iltis, "A modified Bayesian algorithm with decision feedback for blind adaptive equalization," in *Preprints 4th IFAC Int. Symposium Adaptive Systems in Control and Signal Processing* (France), 1992, pp.737-742.

[19] E. Zervas, J. Proakis and V. Eyuboglu, "A quantized channel approach to blind equalization," in *Proc. ICC'92* (Chicago), 1992, Vol.3, pp.351.8.1-351.8.5.

[20] J.G. Proakis, "Adaptive algorithms for blind channel equalization," in *Proc. 3rd IMA Conf. Mathematics Signal Processing* (University of Warwich, UK), 1992.

[21] S. Chen, B. Mulgrew and S. McLaughlin, "Adaptive Bayesian decision feedback equaliser based on a radial basis function network," in *Proc. ICC'92*, (Chicago), 1992, Vol.3, pp.343.3.1-343.3.5.

[22] D. Williamson, R.A. Kennedy and G.W. Pulford, "Block decision feedback equalization," *IEEE Trans. Communications*, Vol.40, No.2, pp.255-264, 1992.

[23] S. Chen, B. Mulgrew and S. McLaughlin, "Adaptive Bayesian equaliser with decision feedback," *IEEE Trans. Signal Processing*, to appear, September, 1993.

[24] K. Abend and B.D. Fritchman, "Statistical detection for communication channels with intersymbol interference," *Proc. IEEE*, Vol.58, No.5, pp.779-785, 1970.

[25] S. Chen and B. Mulgrew, "Overcoming co-channel interference using an adaptive radial basis function equaliser," *EURASIP Signal Processing J.*, Vol.28, No.1, pp.91-107, 1992.

[26] S. Chen, B. Mulgrew and P.M. Grant, "A clustering technique for digital communications channel equalisation using radial basis function networks," *IEEE Trans. Neural Networks*, to appear, May, 1993.

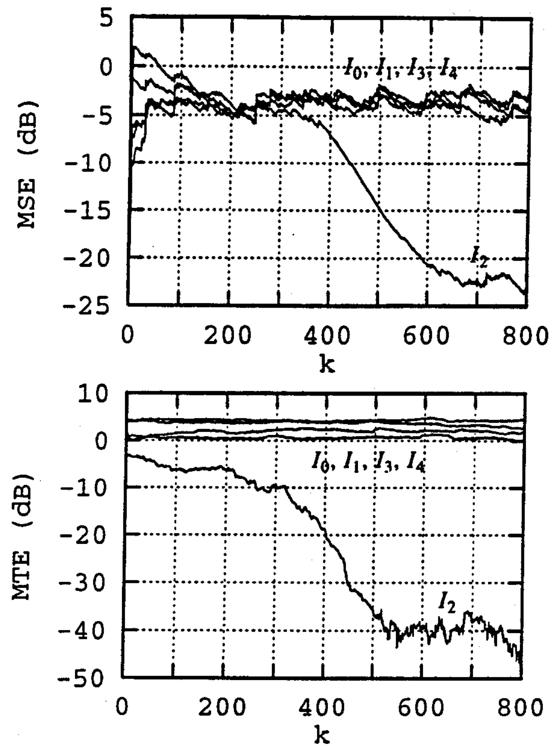

Fig.2 Estimated mean square errors and mean tap errors of adaptive units with different initial estimates. $\hat{\mathbf{s}}_b(0) = [0\ 0\ 0\ 0]^T$ and the label I_i indicates the ith unit.

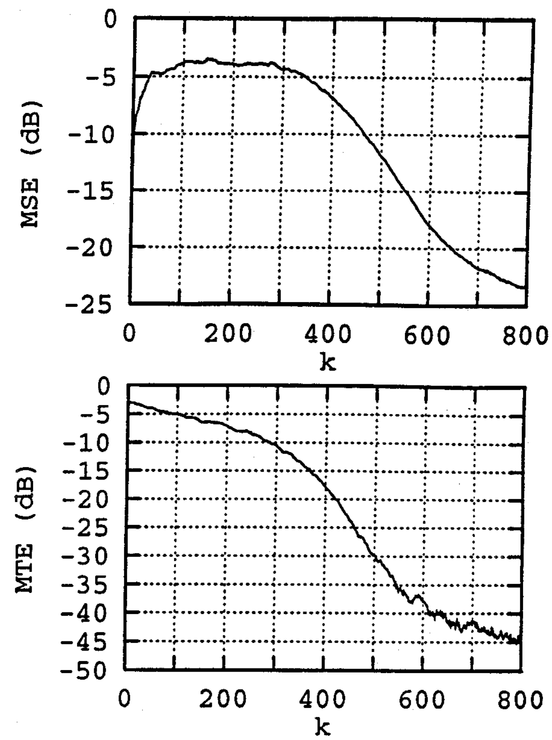

Fig.4 Average estimated mean square error and mean tap error of adaptive unit I_2 over 10 runs. $\hat{\mathbf{s}}_b(0) = [s(-5)\ s(-6)\ 0\ 0]^T$.

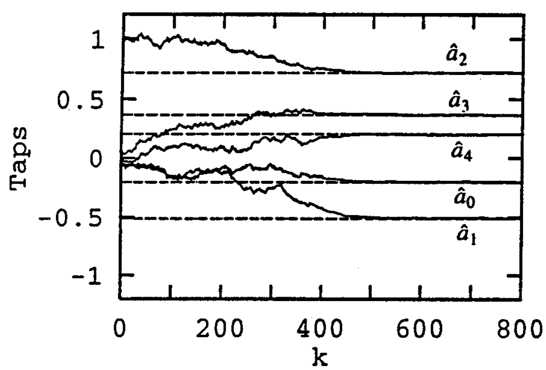

Fig.3 Channel estimate of adaptive unit I_2. $\hat{\mathbf{s}}_b(0) = [0\ 0\ 0\ 0]^T$.

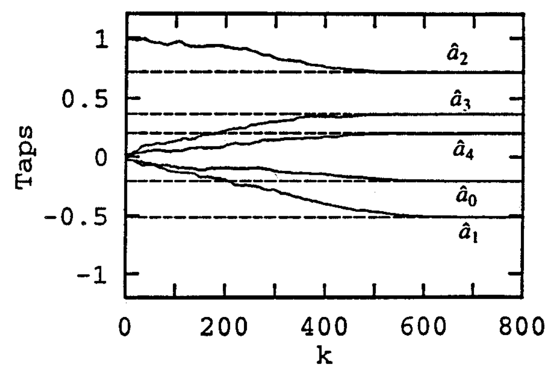

Fig.5 Average channel estimate of adaptive unit I_2 over 10 runs. $\hat{\mathbf{s}}_b(0) = [s(-5)\ s(-6)\ 0\ 0]^T$.