ADAPTIVE ESTIMATION AND EQUALISATION OF THE HIGH FREQUENCY COMMUNICATIONS CHANNEL

Stephen McLaughlin B.Sc. M.I.E.E.E. A.M.I.E.E.

Thesis submitted for the degree of Doctor of Philosophy to the

Faculty of Science in the University of Edinburgh

1989

DECLARATION OF ORIGINALITY

This thesis was composed entirely by myself. The work reported herein was conducted exclusively by myself in the Department of Electrical Engineering at the University of Edinburgh.

ACKNOWLEDGEMENTS

As in many walks of life no man is an island and herein I acknowledge those people who have made the production of this tome easier through their support and help. I would like to thank my supervisor Dr. Colin Cowan for first of all being brave enough to take me on and then for his enlightened guidance. Dr Bernie Mulgrew for numerous educational discussions and Dr. Gavin Gibson for broadening my mathematical education. Professor Peter Grant for his constant advice and encouragement, Professor John Mavor for presiding over a happy department where it is easy to work. Dr Jimmy Dripps for his help early on about HF communication systems. Bill Templeman who persuaded MEL initially to fund me and then took me to the pub when it got difficult! All of my fellow postgraduates in the Signal Processing Group, past and present and too numerous to mention, for all that advice about my presentations.

Professor Dave Falconer of Carleton University for making my visit there an enjoyable one and Bob, Dave, Nessa, Wilf and George for showing me the sights. Professor Neil Bershad of University of California for improving my english and beefing up my analytical skills. My climbing buddies Iain and Stuart who made me relax, on the hills and in the Diggers.

Finally and most importantly I owe a debt of gratitude to my wife, Dorothy, for all her help, support, patience and understanding when I needed it most.

DEDICATION

This thesis is dedicated to my mother and late father for all their help and encouragement over the years, it was they who made it possible.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AWGN	Additive White Gaussian Noise
BER	Bit-Error-Rate
CCIR	International Radio Consultative Committee
DFE	Decision Feedback Equaliser
EKF	Extended Kalman Filter
EVR	Eigenvalue ratio
FIR	Finite Impulse Response
HF	High Frequency
IIR	Infinite Impulse Response
ISI	Intersymbol Interference
LMS	Least Mean Squares
ML	Maximum Likelihood
MLSE	Maximum Likelihood Sequence Estimation
MSD	Mean Square Deviation
MSE	Mean Squared Error
MUF	Maximum Usable Frequency
MVK	Minimum Variance Kalman
RLS	Recursive Least Squares
SNR	Signal-to-Noise Ratio
VLSI	Very Large Scale Integration

LIST OF PRINCIPAL SYMBOLS

Variables and Constants

d	estimation lag of Kalman equalisers
$[x_n]$	input data sequence $x_0, x_1, \ldots, x_n$
$[y_n]$	output data sequence $y_0, y_1, \ldots, y_n$
$[e(n)]$	error sequence $e_0, e_1, \ldots, e_n$
$[h_n]$	impulse response sequence
$\underline{\Phi}_{xx}$	autocorrelation matrix
$\underline{\Phi}_{xy}$	crosscorrelation matrix
$\hat{R}(k)$	estimate of autocorrelation matrix
$\underline{r}_{xx}(k)$	least squares autocorrelation matrix at k_{th} iteration
$\underline{r}_{xy}$	least squares crosscorrelation matrix at k_{th} iteration
α_i	i_{th} eigenvalue of autocorrelation matrix
N	additive white gaussian noise power
K	number of taps in channel (or order of transversal filter)
$\underline{h}_k$	impulse response vector of FIR filter at the k_{th} iteration
$\underline{x}_k$	data vector at the k_{th} iteration
n_k	additive white gaussian noise at time k
$\underline{d}_k$	tap weight incremental change vector at time k
D	variance of $\underline{d}_k$
$\underline{q}_k$	tap weight misadjustment vector at time k
$\underline{q}_k^g$	additive noise term of tap weight misadjustment vector at time k
$\underline{q}_k^l$	lag term of tap weight misadjustment vector at time k
$U_{j,k}$	state space transition matrix (chapter 3)
ζ	mean square error cost function
$\underline{K}_k$	Kalman gain matrix at the k_{th} iteration

A_k	state transition matrix of EKF at the k_{th} iteration
$F(k/k-1)$	state space model state transition matrix from time $k-1$ to time k
$\underline{s}_k$	state vector at time k
$\hat{\underline{s}}_k^k$	estimate of state vector at time k
$\hat{\underline{s}}_k^{k-1}$	prediciton of state vector at time k
$\underline{z}_k$	measurement vector (chapter 4) at time k
$\underline{w}_k$	plant noise vector
Q	plant noise covariance
$\underline{v}_k$	measurement noise vector at time k
R	measurement noise covariance
C_i	Butterworth filter coefficients
$\underline{c}_k$	impulse response at time k of a linear equaliser
$\underline{c}_k^{ff}$	feedforward tap vector of a decision feedback equaliser
$\underline{c}_k^{fb}$	feedback tap vector of a decision feedback equaliser
λ	memory factor in exponentially windowed RLS algorithm
μ	step size used in LMS algorithm
∇	gradient of mean square error cost function
$\hat{\underline{\nabla}}(k)$	estimate of gradient of mean squared error cost function at time k

Operators

$E[\,.\,]$	statistical expectation operator
$tr[\,.\,]$	trace of a matrix
z^{-1}	unit sample delay
Σ	summation
Π	product
$\hat{}$	denotes an estimate

| | denotes the norm of a matrix

Vectors and Matrices

All vectors are specified as column vectors. The matrix transpose operation is denoted by the superscript T.

Chapter 1

INTRODUCTION

This thesis is, as the title suggests, concerned with the development of both algorithms and structures for adaptive signal processing in a communication system operating in the high frequency (HF) band, i.e. 3-30MHz. The motivation is to develop techniques which will enable serial data transmission at data rates considerably higher than are presently achievable. HF communications has long been a neglected area of the spectrum, the advent of satellite communications and the difficult nature of the medium itself apparently having made it redundant. However, the high cost and questionable physical security of satellite links coupled with the advent of relatively cheap very large scale integration (VLSI) have reawakened interest in HF communications[1-7]. Although it must be stated that within the United Kingdom there has been a considerable body of ongoing work in HF communications.

Although receivers based on parallel structures [8] would appear to offer a better performance[9] the resultant increased complexity at the receiver coupled with the limitation imposed on transmitter power make serial structures more attractive. The time-variant nature of the HF communications channel make it ideally suited as an application area for adaptive signal processing techniques.

Adaptive signal processing is a relatively youthful research area, the first pioneering work [10-12] only having appeared thirty years ago, although it is true to say that the groundwork was laid considerably earlier in the work of Gauss [13] and Legendre[14]. The recent plethora of texbooks on the subject [15-20] would appear to suggest that this area is now approaching maturity. However, little attention has been paid to the study of such techniques when applied to environments such as the non-stationary HF channel, with the notable exception of[21-26]; this thesis adds to this work by considering the tracking performance of adaptive algorithms in time-variant environments. The thesis also studies existing methods and develops new techniques for application in the HF communication scenario.

The purpose of this chapter is to provide the necessary definitions required to aid the understanding of this thesis and also to detail the organisation of the material presented in it. Consequently the first section defines precisely what is intended by the term adaptive signal processing within the context of this thesis. The second section then deals with the application of such techniques to the HF communications channel and finally the organisation of the thesis is discussed.

1.1 ADAPTIVE FILTERING

As was indicated in the previous paragraphs, the aim of this chapter is to provide the necessary definitions required to aid understanding of this thesis. In this section a simple introduction to the concept of adaptive filtering is presented. In order to define adaptive filtering it is necessary to first describe what is intended by the term filtering.

One of the primary aims of filtering is to enable the extraction of a signal from one which has been contaminated by noise. In this thesis a filter is considered to be both linear and discrete time in nature; Figure 1.1 represents the structure of such a filter. The input and output signal sequences, x_n and y_n respectively are related to each other by the impulse response of the filter, h_n. Explicitly the output sequence, y_n, is the convolution of the input sequence, x_n, with the the impulse response of the filter, h_n.

Clearly if models for the generation of both the signal and noise processes exist it is possible, in principle, to generate a filter which will optimally enhance the desired signal with respect to the noise. However, in the real world only partial a-priori knowledge of such processes will exist, (at best), and so it is not possible to explicitly derive such a filter. However, it is perfectly reasonable to assume that the necessary information could be obtained through analysis of the real data. That is the optimal filter could perhaps be *learned* from the data. As a consequence then some form of on-line parameter adjustment is required, the adjustments required being derived from analysis of the received data, as illustrated in Figure 1.2.

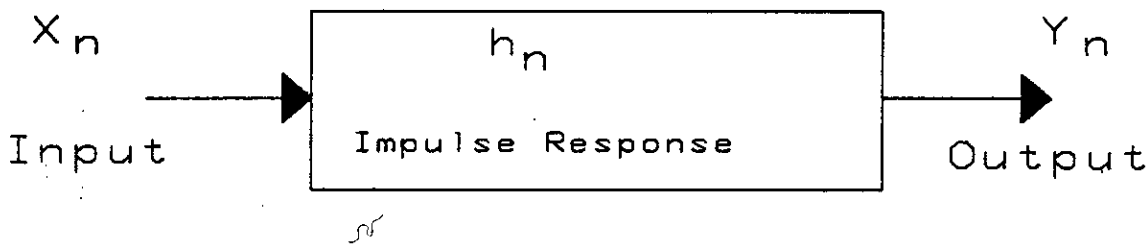


Figure 1.1 - Linear discrete time filter

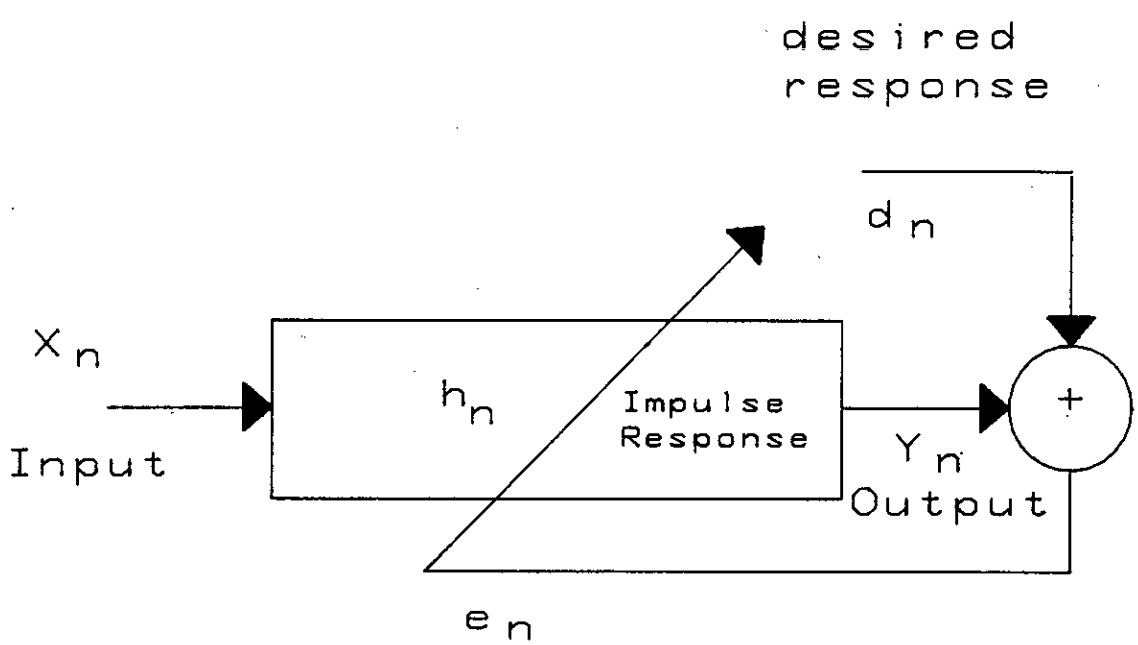


Figure 1.2 - An adaptive filter

Hence, it may be observed that adaptive, in this context at least, can be interpreted as indicating some form of self learning which enables, or at least approximates, optimal behaviour of the filter. The parameter adjustment is normally achieved via adjustment of the filter impulse response based on an algorithm, i.e. a set of rules, which minimise the error between the desired output and the actual output of the filter. This statement raises the question of where the desired response is obtained from since only analysis of actual data is being used. This is normally achieved by transmission of a data sequence, termed the training sequence, which is known a-priori at the destination. There are algorithms which attempt to operate without a training sequence, i.e. *blind* as in[27,28]. These are in general nonlinear in nature and are not considered in this thesis. When an optimal, or as is more realistic near-optimal, solution is reached the algorithm is said to have converged. This concept of convergence also applies when time varying environments are considered, the distinction being that the optimal solution is varying with time, and the convergence behaviour may be measured by how well the adaptive filter tracks the behaviour of the desired output.

In summary, an adaptive filter is a filter with a time varying impulse response, the time variations being selected on-line by an algorithm. The selection is aimed at achieving optimal performance in the sense of minimising a function of the measured error between the desired and actual responses.

1.2 THE APPLICATION OF ADAPTIVE FILTERING TECHNIQUES TO THE HF COMMUNICATIONS CHANNEL

There are many ways in which adaptive filters may be configured for real time applications, [15-20] , however within the context of this thesis there are only two particular structures which are of direct interest. These are channel estimation and channel equalisation, their key differences being clearly illustrated in Figures 1.3 and 1.4. These particular structures are of interest here because, as will be discussed in more detail later, the HF channel may be viewed as a finite impulse response

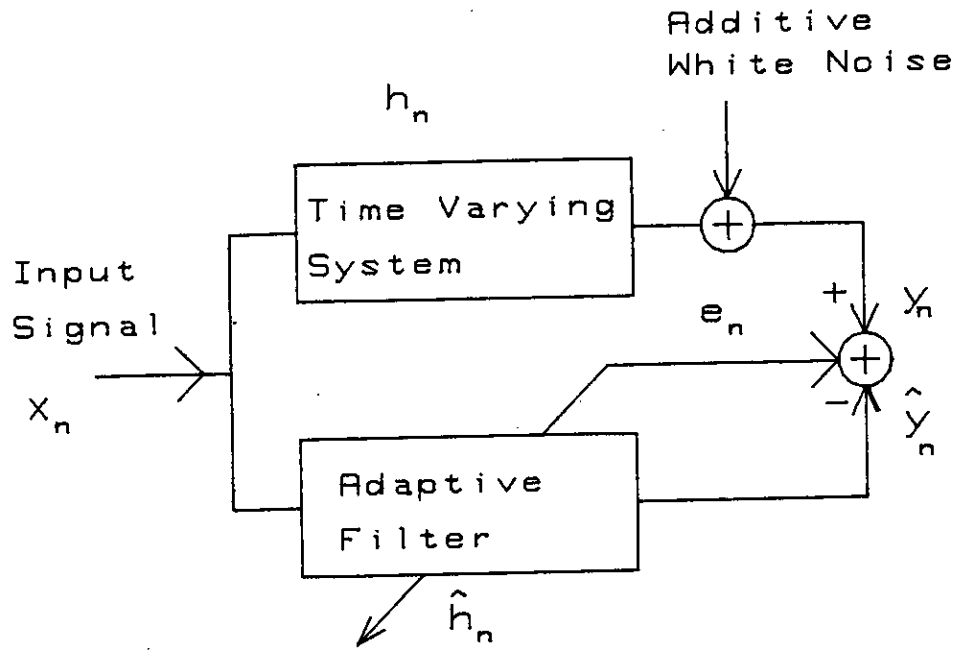


Figure 1.3 - Adaptive estimation of the HF channel

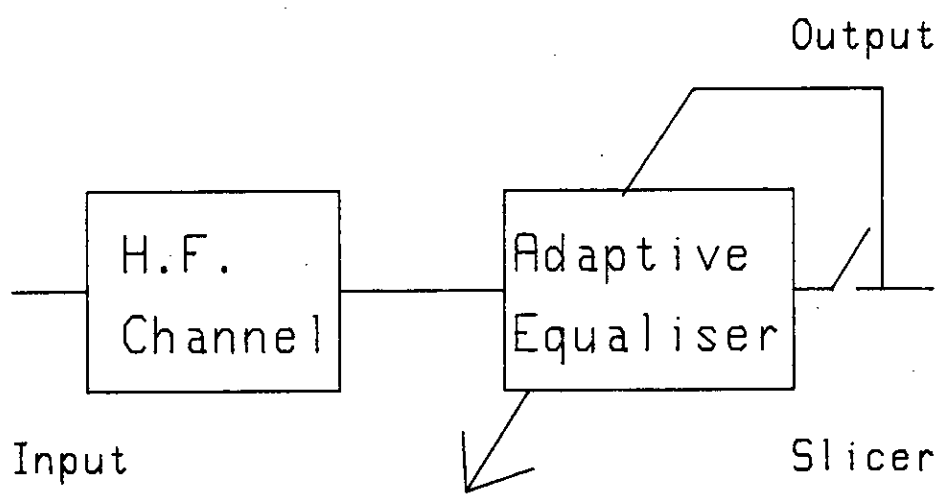


Figure 1.4 - Adaptive equalisation of the HF channel

(FIR) filter whose impulse response is time varying in nature. This fact then allows us to consider applying linear discrete time adaptive filters to estimate directly the impulse response of the channel, i.e. channel estimation, or to equalise distortion introduced into a data sequence transmitted over the channel, i.e. channel equalisation.

In channel estimation, Figure 1.3, the adaptive filter is configured such that the input sequence x_n is also the input to the channel. The output of the filter, $\hat{y}_n$, is then used in conjunction with the channel output, y_n, to generate an error signal, e_n to drive the adjustment of the adaptive filter's impulse response via the algorithm. The aim being to track the time variations of the channel's impulse response.

In channel equalisation, Figure 1.4, the adaptive filter operates in an inverse system modelling approach in contrast to the direct system modelling mode adopted in channel estimation. A transmitted data sequence, y_n' is distorted in passing through the HF channel, the resultant sequence output from the channel, then forms the input to the adaptive filter. The adaptive filter then attempts to reconstruct , $\hat{y}_n$, the original transmitted data sequence.

The key point regarding the application of adaptive filters to HF communications is that due to the time variant nature of the HF channel adaptive filters offer a method by which on-line parameter estimation and adjustment, in both channel equalisation and channel estimation, may be used to improve data rates for serial data communication with no subsequent degradation in performance.

1.3 ORGANISATION OF THESIS

As the title of this thesis suggests it's primary aim is to study existing and develop new adaptive algorithms and structures for channel estimation and equalisation of the HF communications channel. The preceding two sections have provided the definitions of 'adaptive filtering' and the manner in which it will be applied to HF communications systems required for the remainder of this thesis. In chapter 2 necessary background information regarding the nature of HF communications is

presented as well as the structure of the channel model used throughout this thesis for computer simulations. The themes indicated in the title are developed along three separate but inter-related paths in chapters 3,4 and 5.

The performance of the two most common adaptive algorithms, the LMS and RLS [15] as channel estimators in the HF communications environment is first considered in chapter 3. Their tracking performance is analysed to determine which if any would be the most suitable for this application. It is demonstrated that contrary to popular belief the RLS is not particularly suited to this type of environment, offering a performance no better than the LMS for a considerable penalty in terms of computational complexity.

Chapter 4 moves on from these conclusions and attempts to develop novel adaptive algorithms for HF channel estimation utilising a-priori knowledge of the channel structure and incoporating it into the algorithm. A channel estimator, termed the minimum variance Kalman (MVK) is presented which utilises full a-priori knowledge of the channel. The performance of this algorithm is optimal. The next two algorithms aim to overcome the full a-priori knowledge of the MVK. Initially this is done within the context of an extended Kalman filter (EKF), and then by a computationally simpler technique where the LMS is modified to include a prediction filter, effectively increasing the order of the recursion in the LMS. The EKF structure is shown to give excellent performance but suffers from severe numerical instability due the high degree of computational complexity. The modifed LMS is considerably simpler, however it's performance is disapointing although it offers scope for improvement.

In chapter 5 the equalisation problem is considered and a novel equaliser structure is presented based on work presented in [29] , where the channel estimation and sequence estimation processes are separated. The performance of the new structures and their complexity is compared with that of two existing structures, a Godard-Kalman decision feedback equaliser [30] and the adaptive Kalman equaliser

of[29].

Finally in chapter 6 the conclusions put forward in the preceding chapters are summarised and areas worthy of further study based on the work presented here are suggested.

Chapter 2

THE HIGH FREQUENCY COMMUNICATIONS ENVIRONMENT

This chapter deals with the basic nature of and principles by which communication within the HF medium is achieved. The manner by which the HF channel is quantified is discussed and the technique used in the computer simulations in the remainder of this thesis presented.

2.1 INTRODUCTION

Within the historical development of electrical communications systems HF must rank high as a prime mover in the acceleration from the turn of the century to today which has no corner of the earth safe from one form or another of electrical communication. When Marconi first demonstrated transatlantic communication by radio in 1901 [31,32] many people envisaged HF communications providing a world wide communication facility for general use. However the highly time-variant and unpredictable nature of the HF radio prevented this and it's use has, in general, been restricted to the military and amateurs. Only in the United Kingdom was there a continuing research effort. In recent years interest has been growing on a worldwide basis in HF communications because of the increasing availability of cheap digital signal processors coupled with the lack of physical security and high cost of many satellite communications systems.

Initially in this chapter the physics involved in propagating radio waves over the HF medium is discussed and also the nature of said medium. A brief description of the mathematical and physical justification of the manner in which the HF channel is modelled is then presented. Finally a summary of the simple channel model used in this thesis is discussed.

2.2 PROPAGATION AT HF

This section summarises the physics involved in the propagation of HF radio

waves within the atmosphere.

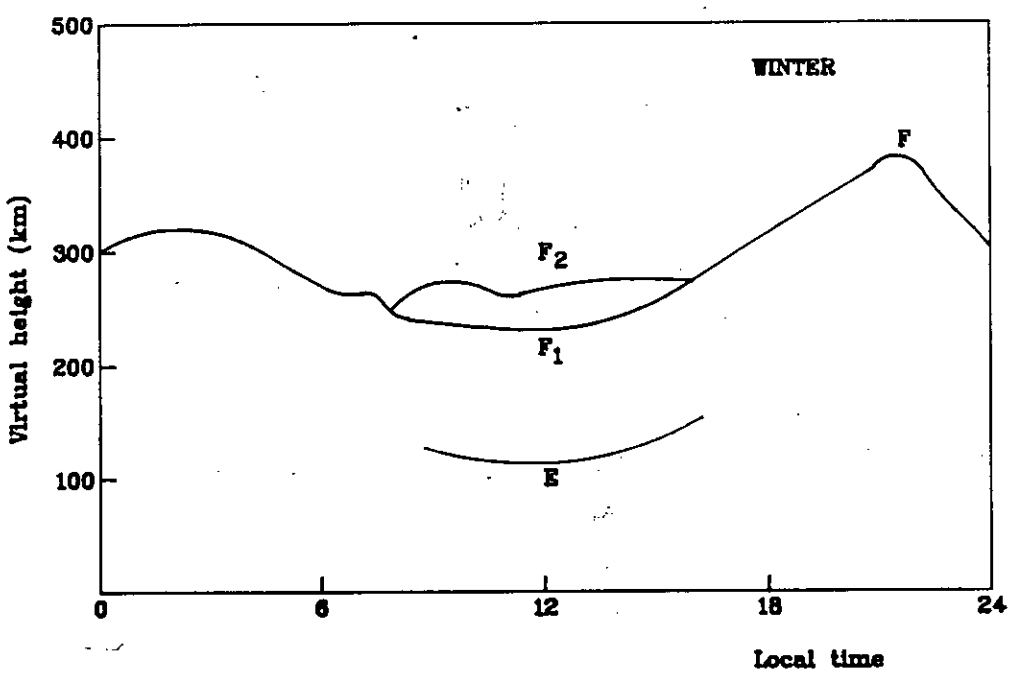
2.2.1 The Ionosphere

The ionosphere is the region of the earth's atmosphere at an altitude of approximately 50-350km. It is formed as a result of the ionisation of atoms and molecules of oxygen and nitrogen in the atmosphere by the sun's radiation. The structure of the ionosphere has been determined by vertical and oblique radio sounding[33]. It is generally divided into three regions labelled D, E and F. Signal propagation in the HF band (3-30MHz) can be thought of as reflection from these layers. It should be emphasised however that these macroscopic layers are in fact peaks of ionisation intensity which vary in position, altitude and mean-square reflectivity as a result of microscopic ionospheric turbulence. The characteristics of each of these regions is discussed below.

The D region lies at an altitude of 60-90km and is a daytime phenomenon, since it disappears at night because molecular recombination is no longer counteracted by the radiation of the sun. The D region does not normally support propagation as a result of the high level of absorption which it exhibits.

The band between 90 and 170km is known as the E region, the maximum ionisation intensity occurring at about 110km. This region does support propagation, but, because there is still a significant concentration of heavy particles, it does not do so particularly well at night.

The region above 170km is termed the F region; in daytime it is divided into the F1 and F2 layers. The F1 layer is centred at around 200km and is not generally considered as a vehicle for transmission on its own. The F2 layer is concentrated about 300-320km and plays the dominant role in long-range (>2000km) communication. At night the F1 and F2 layers merge at around 300km and provide the main propagation mechanism at night. Figure 2.1 [33] illustrates typical variations in height of the E and F regions over a twenty-four hour period for both summer and



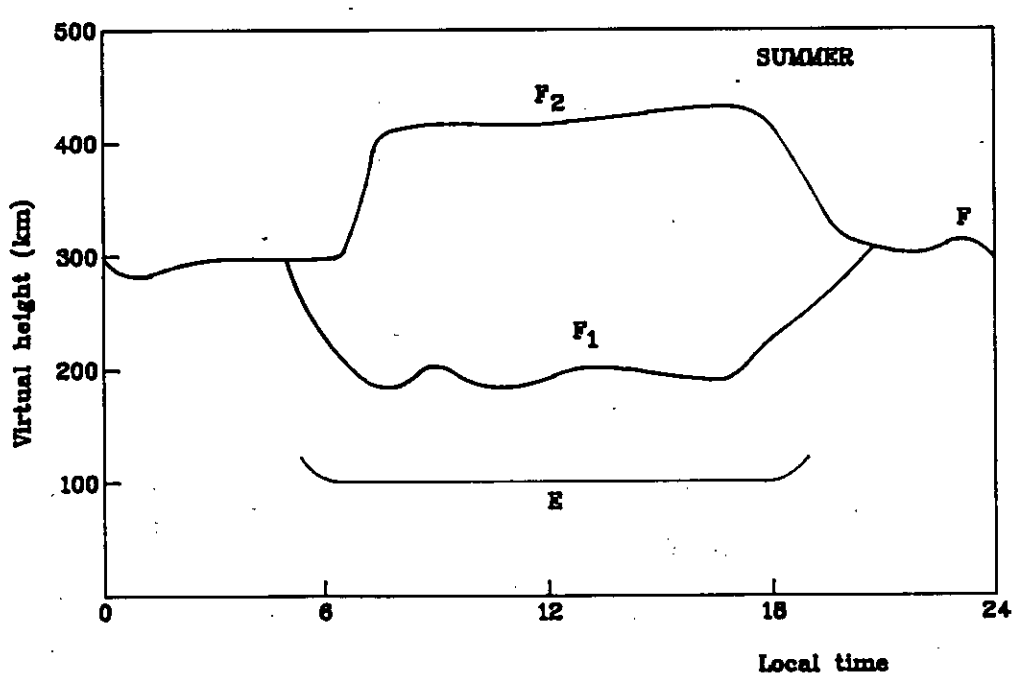


Figure 2.1 - Diurnal variation of ionospheric layers [31]

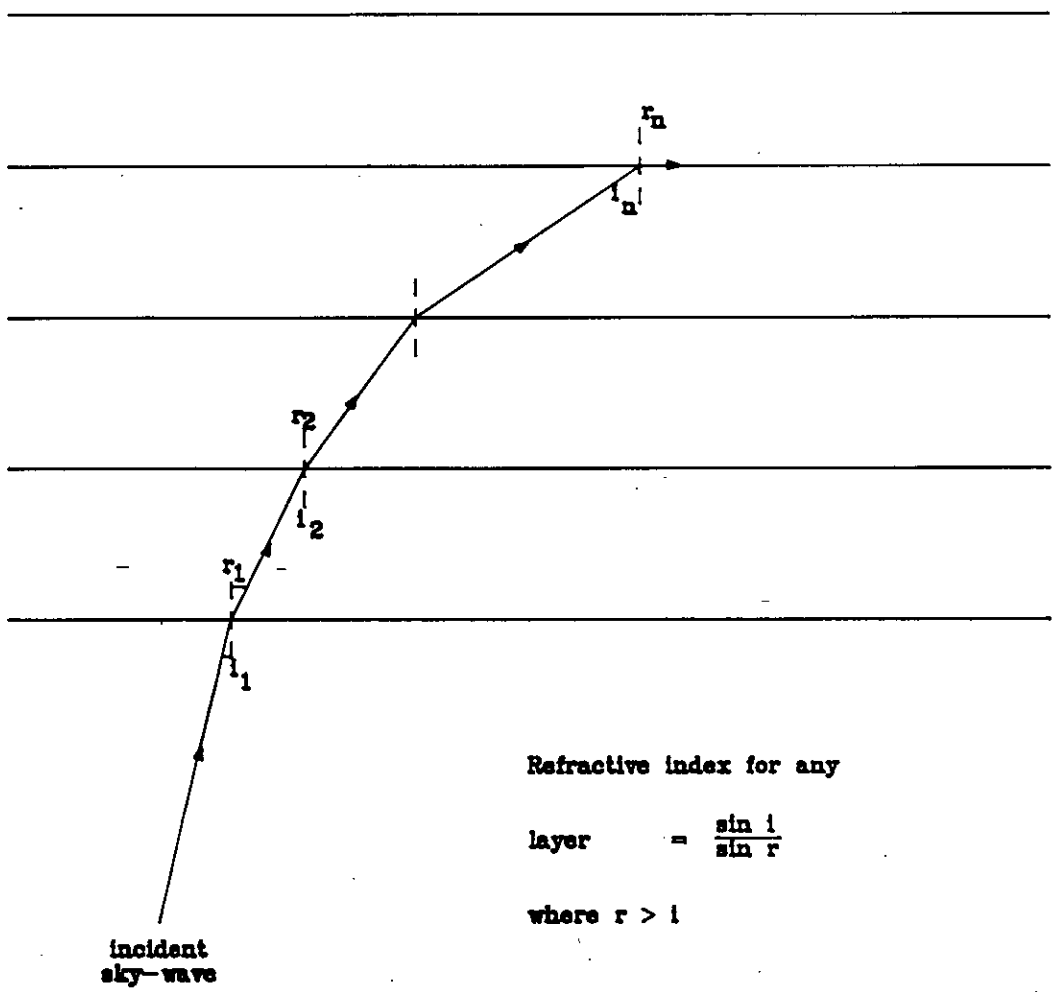


Figure 2.2 - Refraction by ionospheric layers of increasing electron density

winter.

2.2.2 The Method of Wave Propagation

As was stated in the previous section, signal progagation at HF is popularly considered as reflection of the radio waves from the ionosphere. It is in fact a process of refraction which arises from the continuous change in refractive index with altitude of a particular layer as illustrated in Figure 2.2.

Clearly for a wave to be refracted back to the earth's surface, its trajectory must become horizontal at some point. By applying Snell's law of refraction it is possible to obtain an expression for the carrier frequency required to meet these limitations. From this, the maximum frequency at which a vertically incident ray is refracted back to the earth can be calculated and this is known as the critical frequency. If the selected operating frequency is higher than the critical frequency then the ray passes through the layer, and is termed an escape ray, albeit bent by the refraction process, as in Figure 2.3.

Since the behaviour of the layers has been recorded over many years, tables have been produced which indicate the maximum usable frequency (MUF) for a given time of day, geographical location and time of year. The operating frequency of most systems is normally selected to be approximately 10% below the MUF to try to ensure communication since these predictions are essentially long term averages of observation and prone to error as a result of any fluctuations. It should be emphasised that these tables are merely a guide and communication may only be possible at a frequency higher (or lower) than the specified MUF.

One important feature of the refraction process is that it can be shown to be equivalent to mirror type reflection at a particular height, as is indicated in Figure 2.4. This height commonly being referred to as the virtual height with Figure 2.5 illustrating its variation as a function of frequency.

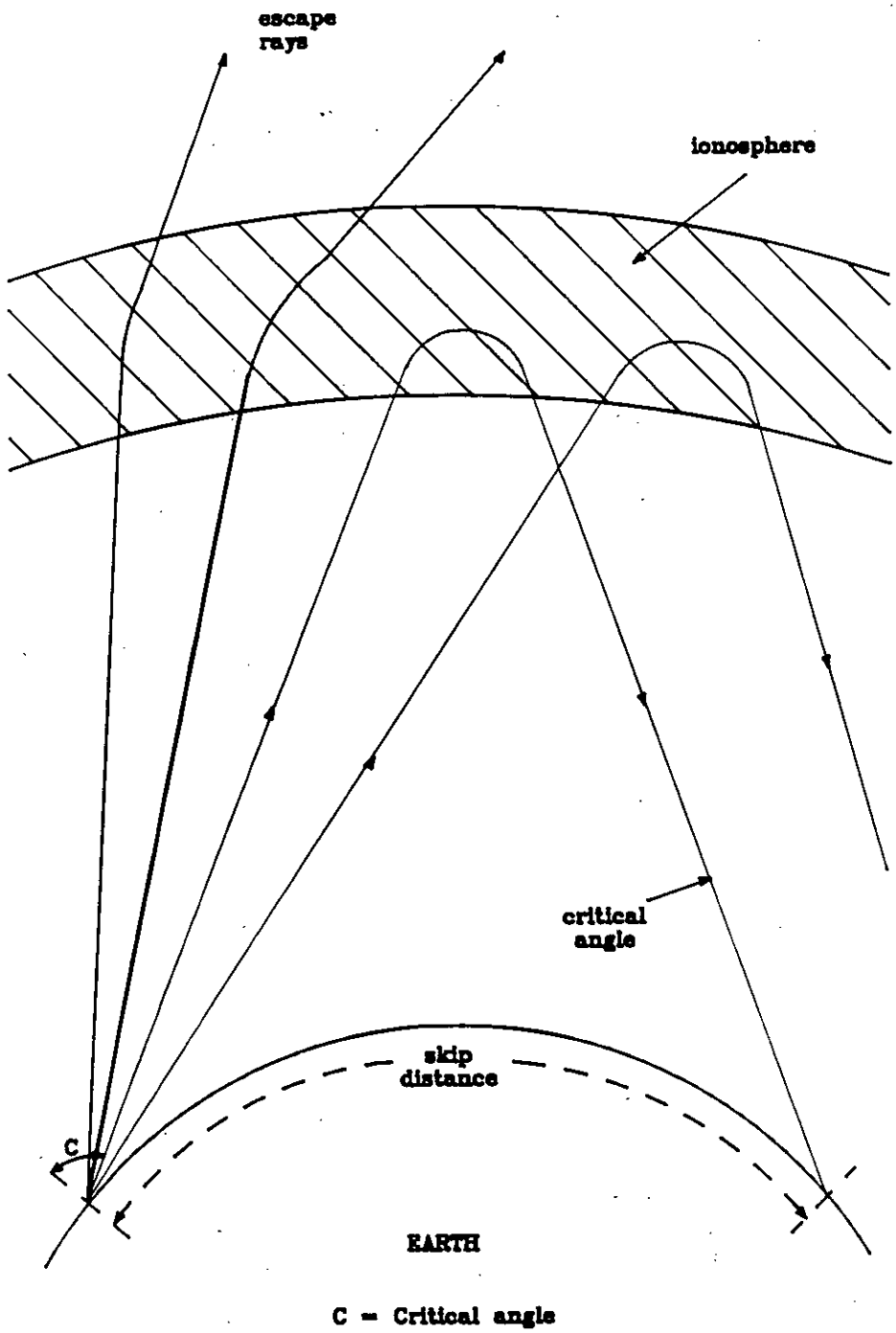


Figure 2.3 - Illustration of the critical angle and critical ray in ionospheric propagation

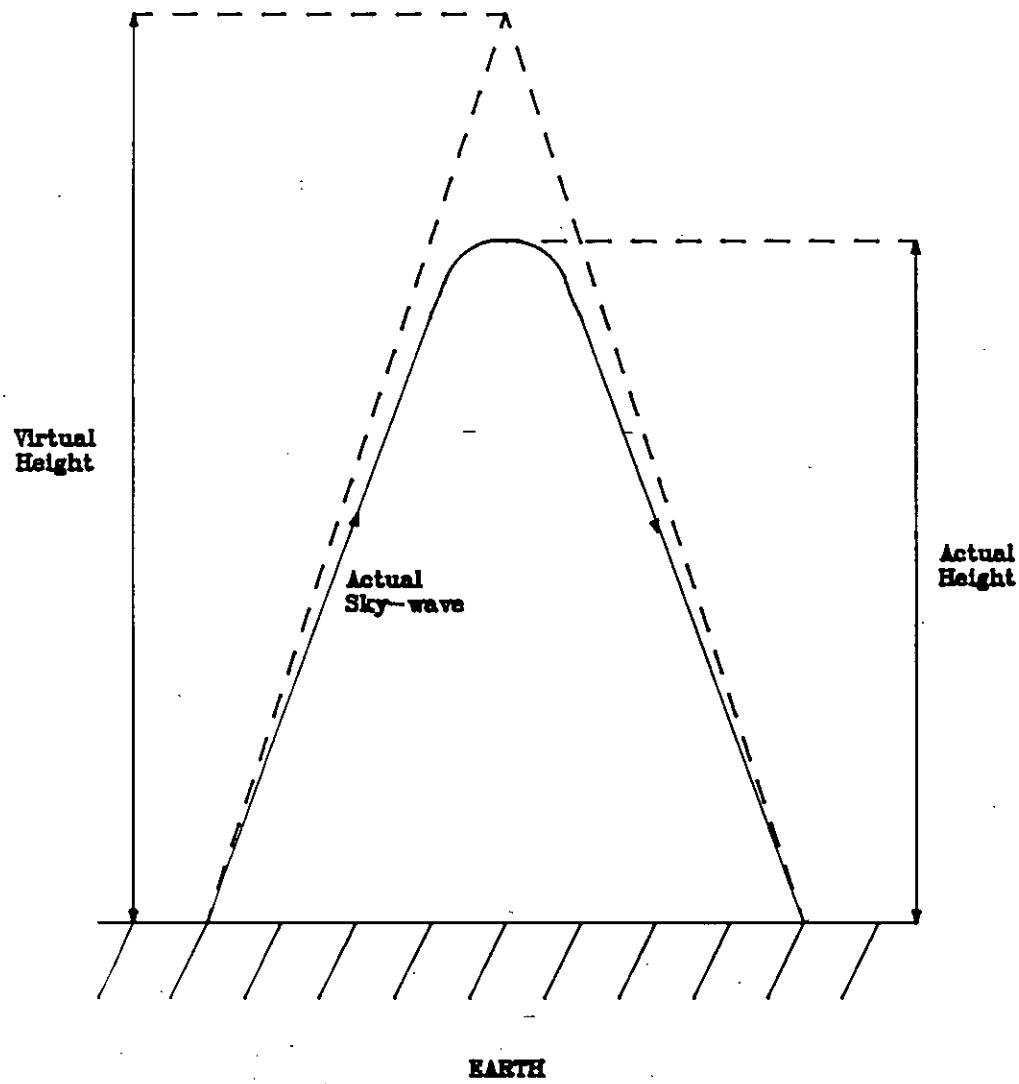


Figure 2.4 - Difference between virtual and actual height in sky-wave propagation

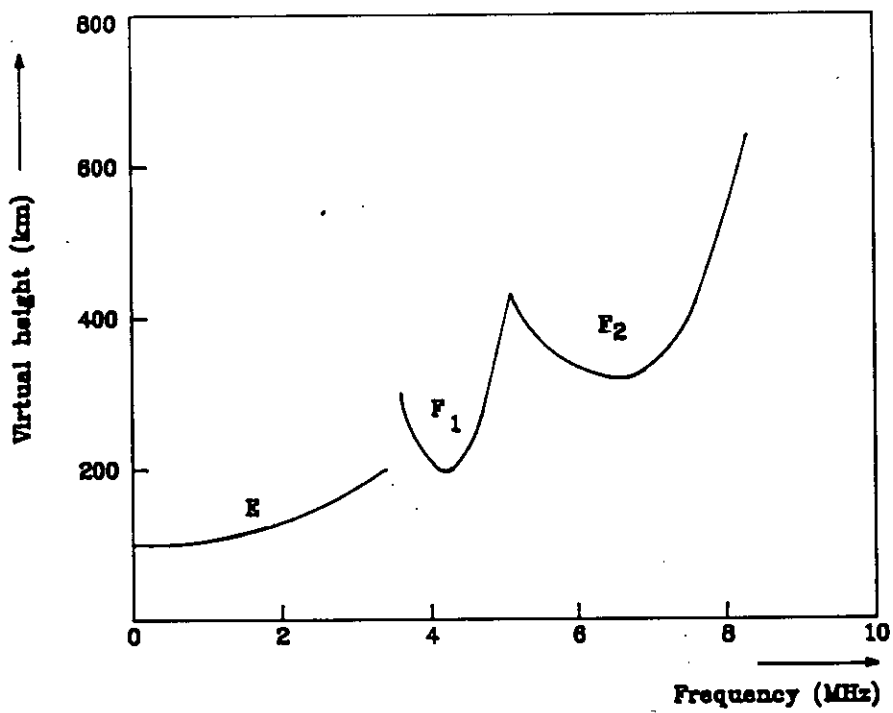


Figure 2.5 - Virtual height as a function of frequency for vertical incidence [31]

2.2.3 Multiple Mode Propagation

Multiple progagation modes are possible in HF communications for three main reasons, the first is as a result of radio waves being electromagnetic in nature. It is well known [34] that on entering a magneto-ionic medium, (such as the ionosphere), any electromagnetic wave will split into two waves, termed the ordinary and extraordinary waves. Each of these waves will traverse different paths in the medium because the expression for their refractive indices will be different. This of course means there are two possible MUFs for each transmission mode and this phenomenon can be observed on ionospheric soundings, as illustrated in Figure 2.6. In addition, since most HF radio systems operate some 10-15% below the MUF, it is possible to have a high and low angle ray, as illustrated in Figure 2.7. This phenomenon occurs because the summit of the ray path at the MUF does not traverse the portion of maximum electron density. Hence, if transmission is lower than the MUF, as is normal, two ray paths are possible. The one corresponding to the higher virtual height is termed the high angle ray the other being the low angle ray.

Finally it is clear that because of the existence of two reflecting layers, (E and F), it is possible for ducting of the wave to occur between the layers and also between the layers and the earth's surface as in Figure 2.8.

2.2.4 Signal Fading

It should be apparent from the preceding section that the multiple rays which arise when communicating between two points, (as in Figure 2.9), cannot all arrive at the receiver simultaneously. The continuously changing nature of the ionosphere ensures that time variations in time of arrival and in the magnitude of each wave will occur. These effects when combined result in constructive and destructive interference at the receiver which causes fading of the received signal.

2.3 OTHER FACTORS INFLUENCING HF COMMUNICATIONS

It is clear that as well as the propagation of multiple paths within HF

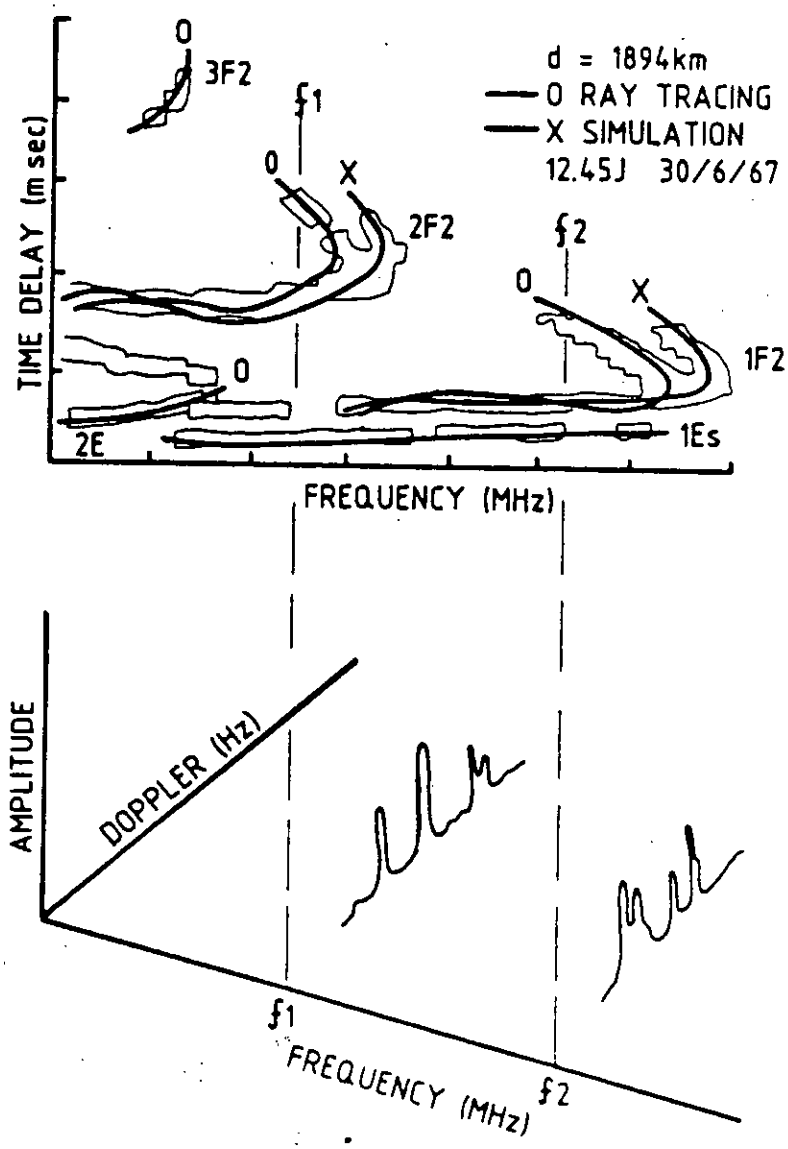


Figure 2.6 - Time and frequency domain displays of oblique sounding data [137]

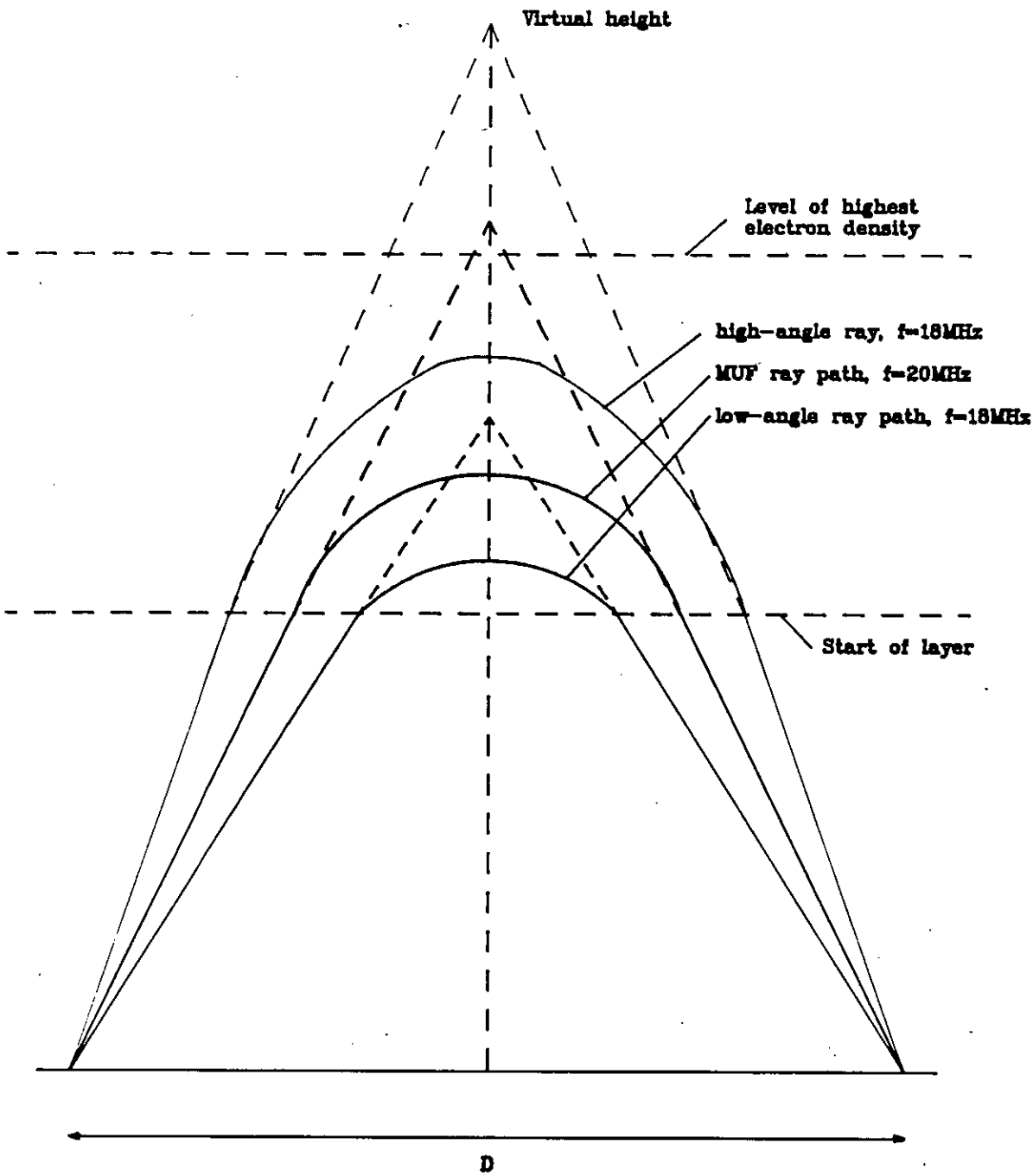


Figure 2.7 - High and low-angle ray paths

communication channels and the resultant signal fading that there are other factors that wi influence the quality of communications at HF. The two most important of these are dea within this section.

2.3.1 Atmospheric Noise

Noise is one of the fundamental limitations of any electrical communication system and atmospheric noise influences radio communication. It may be separated into two main categories, man-made noise and noise associated with natural phenomena in the earth's atmosphere. Man-made noise is generally associated with urban areas where there is a greater concentration of electrical equipment, all generating interference. Natural atmospheric interference is most commonly caused by lightning strikes which radiate large amounts of noise over great distances and a wide range of frequencies. In all simulations within this thesis the noise is modelled as a single additive white Gaussian noise source. No account is taken of impulsive noise [35] which is suggested will have a log-normal distribution. Such noise is normally overcome by means of some form of forward error correcting code[36] .

2.3.2 Equipment Limitations

Clearly the equipment will play a major role in the performance of any HF communication system. It is only in the recent past with the advent of relatively cheap multi-purpose DSP chips that such technology has been available for use in HF radio systems. The growth of this area is demonstrated in [123,124] where DSP technology is being applied to all aspects of HF radio design from speech codecs [125,126] to the filtering operations necessary within the radio receiver [127] and also to the control of adaptive arrays for suppression of jammers as in [128]. Clearly as complex DSP chips become more widely available and new algorithms are developed then adaptive techniques will have an even larger role to play in the HF communications scenario.

2.3.3 Co-channel Interference

One of the major limitations in HF communications, especially within the

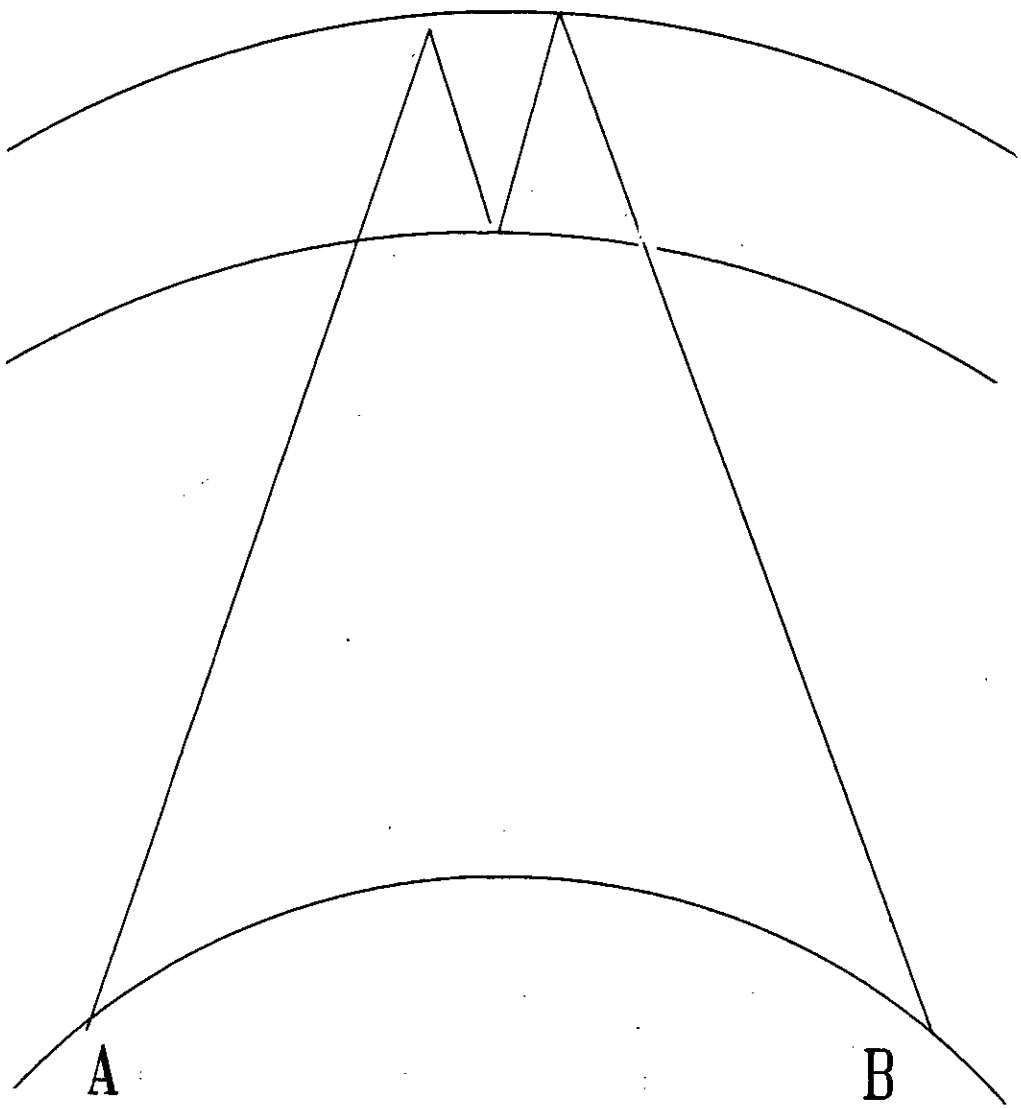


Figure 2.8 - Illustration of ducting between layers in the ionosphere

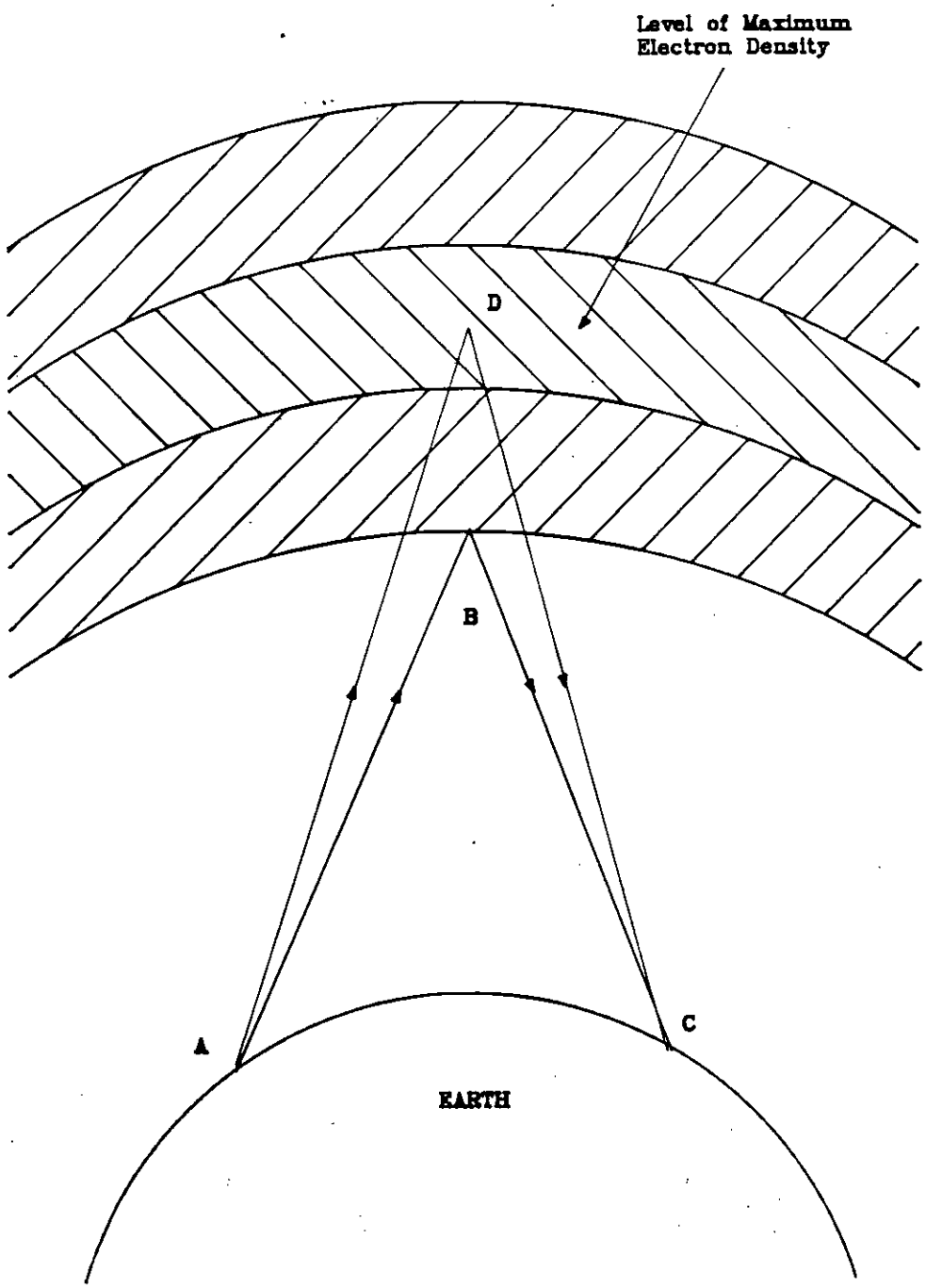


Figure 2.9 - Explanation of fading in HF communications

European context, is co-channel interference [45]. The ionosphere, as stated previously, provides long range sky wave propagation paths; as a consequence users of the HF medium often find that the channel they wish to use is occupied by an interference signal arriving from a source at some considerable distance. These interference signals tend to be narrowband in nature, persist for three to four minutes and are in general non-Gaussian in nature [129,130]. Over the past decade, a great deal of work [131-136] has been carried out in trying to measure and characterise this interference with the aim of generating accurate statistical models of the nature and form of co-channel interference.

One possible approach where adaptive signal processing techniques could be utilised to overcome this problem is within the context of real-time channel evaluation techniques (RTCE). Adaptive techniques are used to monitor the quality of any particular channel via measurements of SNR, fade-rate, etc. In addition these systems can determine if a channel is occupied or not so that an operator can make the most efficient use of the allocated frequencies.

Co-channel interference is not considered within the context of this thesis. This approach is taken because of the nature of the interference, in particular the non-Gaussian aspect, suggest that conventional adaptive filtering will not be enough to eliminate it. Consequently some combination of techniques, such as RTCE and adaptive equalisation, would be required to deal with both the time-variant nature of the channel and the non-Gaussian nature of the co-channel interference. The work reported in this thesis concentrates on the problem of adapting to the frequency selective fading nature of the HF channel.

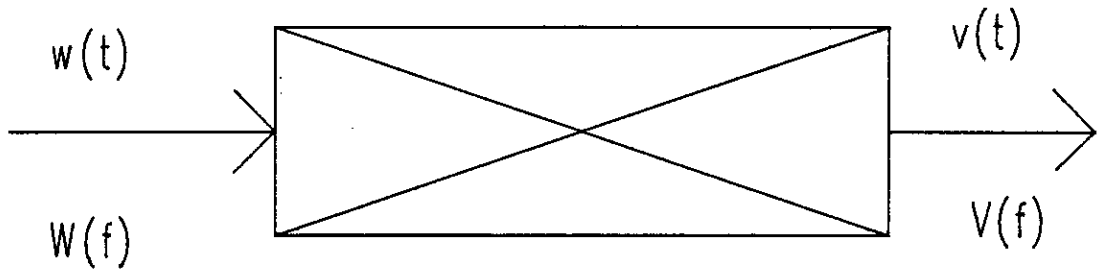


Figure 2.10 - Duality of linear time-variant channel

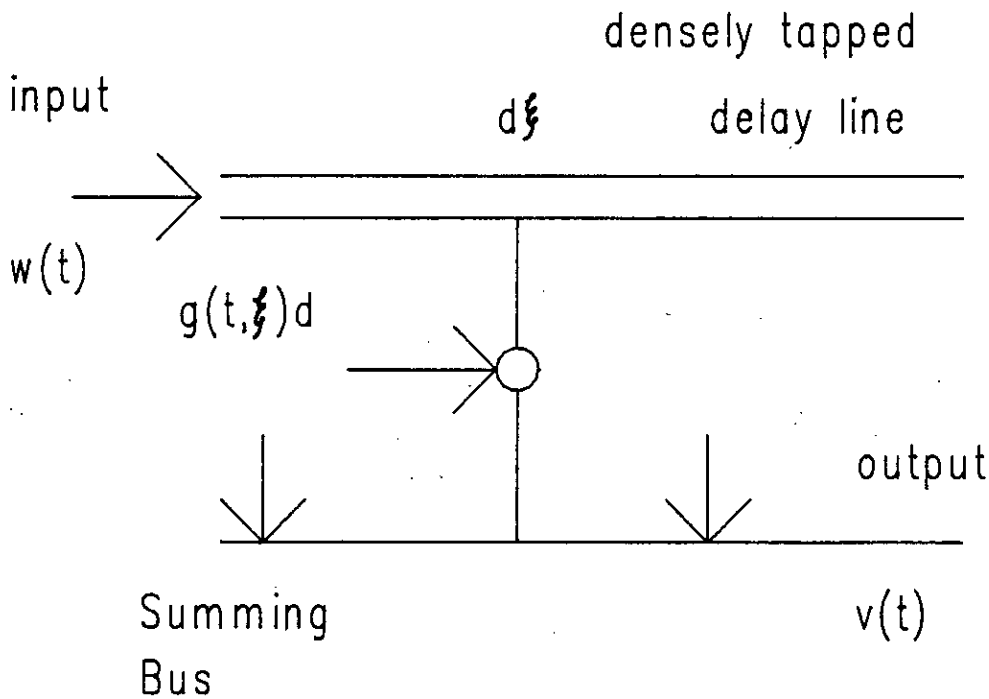


Figure 2.11 - Simplistic channel model

2.4 THE HF CHANNEL MODEL

2.4.1 Mathematical Basis for the Channel Model

In the late 1950's and early 1960's, a considerable amount of work was carried out into the characterisation and behaviour of randomly time-variant linear channels[37-40] , the motivation for this work coming from both communication systems and radar astronomy research. This section provides a summary of the work presented in[40] with particular reference to the HF communications channel.

System Functions

If a randomly time-variant linear communications channel is represented as in Figure 2.10, then it is clear that the concept of time and frequency duality may be used to obtain four possible operators based on input and output representations in both the frequency and time domain, as illustrated below in equations 2.1a-d,

$$v(t) = O_{tt}\ [\ w(t)\ ], \tag{2.1a}$$
$$V(f) = O_{ff}\ [\ W(f)\ ], \tag{2.1b}$$
$$v(t) = O_{tf}\ [\ W(f)\ ], \tag{2.1c}$$
$$V(f) = O_{ft}\ [w(t)\ ]. \tag{2.1d}$$

Where clearly the operators O_{tt}, O_{tf}, O_{ff}, and O_{ft} individually consist of dual operators. Since it is assumed that it is a linear channel which is being dealt with, then these equations may now be more formally expressed as linear integral operators with associated kernels, as demonstrated below.

$$v(t) = \int w(s)\ K_1(t,s)\ ds\ , \tag{2.2a}$$
$$v(t) = \int W(f)\ K_2(t,f)\ df\ , \tag{2.2b}$$
$$V(f) = \int W(l)\ K_3(f,l)\ dl\ , \tag{2.2c}$$
$$V(f) = \int w(t)\ K_4(f,t)\ dt. \tag{2.2d}$$

The variables s and l representing dummy variables for time and frequency respectively.

Although from a mathematical point of view the above expressions are sufficient to describe the system, they do not easily allow a physical interpretation of the input-

output relationship of the system. In order to achieve this it is necessary to derive a series of kernel system functions. Obviously a series of such functions could be obtained and this was clearly illustrated by Bello in[40]. However, for the purposes of this work it is sufficient to consider only two of these functions, the Input Delay Spread function and the Time-Variant Transfer function.

The derivation of the Input Delay Spread function proceeds as follows, the substitution $s = t - \epsilon$ is made in equation (2.2a), thus:-

$$v(t) = \int w(t-\epsilon)\, g(t,\epsilon) d\epsilon, \tag{2.3}$$

where,

$$g(t,\epsilon) = K_1(t,t-\epsilon). \tag{2.4}$$

This form of the equation allows the input-output relationship of the channel to be interpreted as a continuum of stationary scintillating scatterers where $g(t,\epsilon)$ represents the complex modulation produced by the hypothetical elemental scatterers in the range $(\epsilon, \epsilon + d\epsilon)$. The Input Delay Spread function as defined above may be viewed as the channel impulse response at the delay ϵ. This then allows, using equation (2.3), the input signal to be interpreted as first delayed and then multiplied by a differential scattering gain. This is illustrated in Figure 2.11.

The Time-Variant Transfer function, $T(f,t)$, is simply the Fourier transform of the Input Delay Spread function, that is

$$T(f,t) = \int \exp(j\epsilon)\, g(t,\epsilon) d\epsilon \tag{2.5}$$

These functions provide the basis for the channel model used in this study as will be illustrated in the proceeding sections.

Channel Correlation Functions

It is clear that given the time-variant nature of the communication channels considered in this analysis that in order to characterise them completely, it will be necessary to define the associated auto-correlation functions. A full treatment of all the auto-correlation functions which can be developed is provided in Bello [40] but

again only those of interest are illustrated here.

Proceeding with the assumptions that the scattering of the channel at different delays is uncorrelated and that the channel may be considered wide-sense stationary, then the auto-correlation function of the time-variant impulse response, $g(t,\epsilon)$, is,

$$\Phi(\tau_1, \tau_2 ; \Delta t) = \frac{1}{2}E\ [\ g^*\ (\tau_1 ; t)\ g(\tau_2 ; t + \Delta t)\ ]. \tag{2.6}$$

The uncorrelated scattering assumption allows this to be written as,

$$\Phi(\tau_1, \tau_2 ; \Delta t) = \Phi(\tau_1 ; \Delta t)\delta(\tau_1 - \tau_2), \tag{2.7}$$

and if $\Delta t = 0$ then $\Phi(\tau)$ is simply a power spectrum and provides a measure of the average power output as a function of delay. This may be interpreted as the multi-path spread of the channel.

Adopting a similar approach with the Time-Variant Transfer function, $T(f,t)$, a similar auto-correlation function can be defined as below:

$$\Phi(f_1 f_2 ; \Delta t) = \frac{1}{2}E\ [\ T^*(f_1 ; t)\ T(f_2 ; t + \Delta t)\ ], \tag{2.8}$$

and this is known as the spaced-frequency spaced-time correlation function. If Δt is again set to zero and $\Delta f = f_2 - f_1$ substituted then this becomes,

$$\Phi(\ \Delta f\ ) = \int_{-\infty}^{\infty} \Phi(\ \Delta\tau\ )\ \exp(-j2\pi\Delta f\,\tau)\ d\tau \tag{2.9}$$

i.e. the Fourier transform of the multi-path intensity profile. This function provides a means of determining the frequency coherence of the channel and

$$(\ \Delta f\ )_c = \frac{1}{T_m}, \tag{2.10}$$

where T_m is the multipath spread. If $(\ \Delta f\ )_c$ is small in comparison to the signal bandwidth then the channel is termed frequency selective since the signal will be severely distorted. If Δf is set to zero rather than Δt in the expression for the spaced-frequency spaced-time correlation function then the time variations of the channel are demonstrated as Doppler spreading and possibly a Doppler shift of a spectral line.

If the Fourier transform of $\Phi\ (\ \Delta f\ ;\ \Delta t\ )$ is defined as,

$$S(\Delta f;\lambda) = \int_{-\infty}^{\infty} \Phi(\Delta f;\Delta t)\exp(-j2\pi\lambda\Delta t)\, d\Delta t \tag{2.11}$$

and Δf is set to zero, then

$$S(\lambda) = \int_{-\infty}^{\infty} \Phi(\Delta t)\exp(-j2\pi\lambda\Delta t)\, d\Delta t \tag{2.12}$$

This power spectral density function relates the signal intensity as a function of Doppler frequency; hence $S(\lambda)$ is termed the Doppler power spectrum of the channel. The range of values over which $S(\lambda)$ is non-zero is termed the Doppler spread of the channel and consequently via the Fourier transform relationship with $\Phi(\Delta t)$, the coherence time of the channel is

$$(\Delta t)_c = \frac{1}{B_d}, \tag{2.13}$$

where B_d is the Doppler spread of the channel.

Delay Line Model

From the previous sections and the discrete nature of the multi-path phenomenon in the HF channel as discussed in section 2.2, it is possible to propose a general model as illustrated in Figure 2.12. The time-varying frequency response may then be written as,

$$T(f,t) = \sum_{i=1}^{N} G_i(t)\exp(-j2\pi\tau_i), \tag{2.14}$$

where i is the tap or path number, τ_i the time delay associated with the i_{th} path, N represents the total number of paths with $G_i(t)$ the time-varying gain of the i_{th} path.

2.4.2 Physical Basis for Channel Model

The model which has been proposed has been shown to be an accurate representation of the HF channel by Watterstone et al. in[41, 42]. This is because, as measurements have confirmed [43-45] , for 80-85% of the time HF channels exhibit Rayleigh fading characteristics. This implies that independent zero-mean complex Gaussian characteristics would be appropriate to describe the tap gain functions. In addition the discrete nature of the rays leads to the tapped delay line model.

In some situations HF channels have exhibited Ricean fading characteristics (i.e. a specular component exists), this means that the function may no longer be considered to have a zero-mean. In addition, the channel model assumes no dispersion and this is in general true for the bandwidths involved in HF radio systems, however under certain conditions the ionosphere can be very turbulent and in these situations, dispersion may occur. For accurate modelling of such situations, it would be necessary to incorporate all-pass dispersion filters in each path proceeding the multipliers. Table 2.1 illustrates the range of conditions which the CCIR recommend channel simulators be capable of demonstrating.

Parameter	**Range**
Fading depths	2 to 40dB
(*)Fade duration	0.05 to 1.5s
(*)Fade rate	5 to 40 per minute
(*)Delay time	0 to 5ms
(*)Spectral width	0.1 to 1.2kHz
(*)Rate at which fade	0.5 to 2kHz/s
Frequency drifts	0 to 7Hz

Table 2.1 - CCIR recommended range of parameters for HF channel simulators

* - note not all of these parameters are independent of each other.

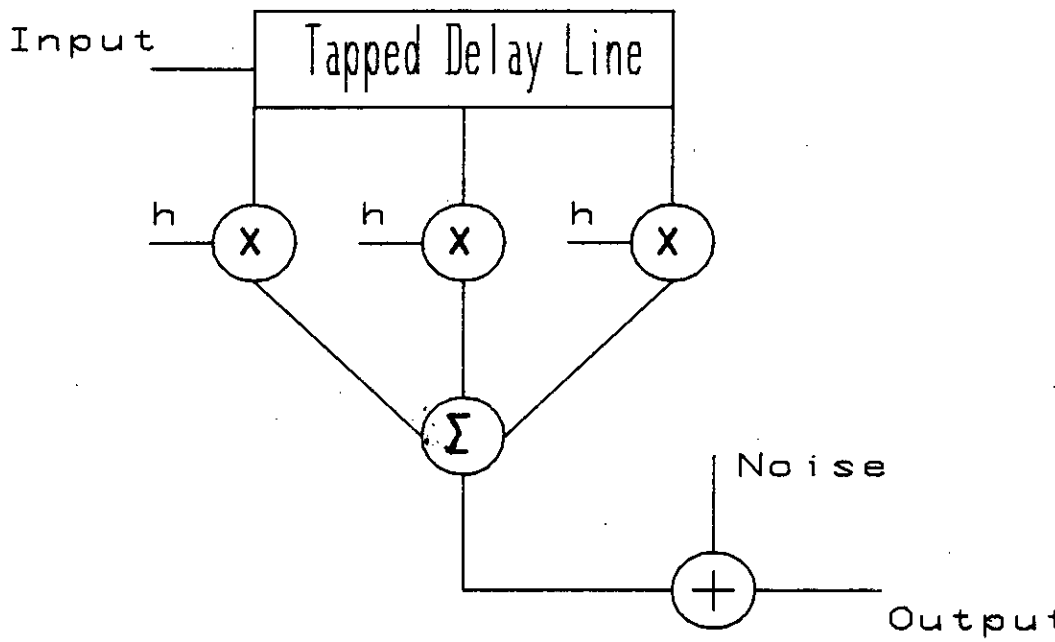


Figure 2.12 - HF Channel model

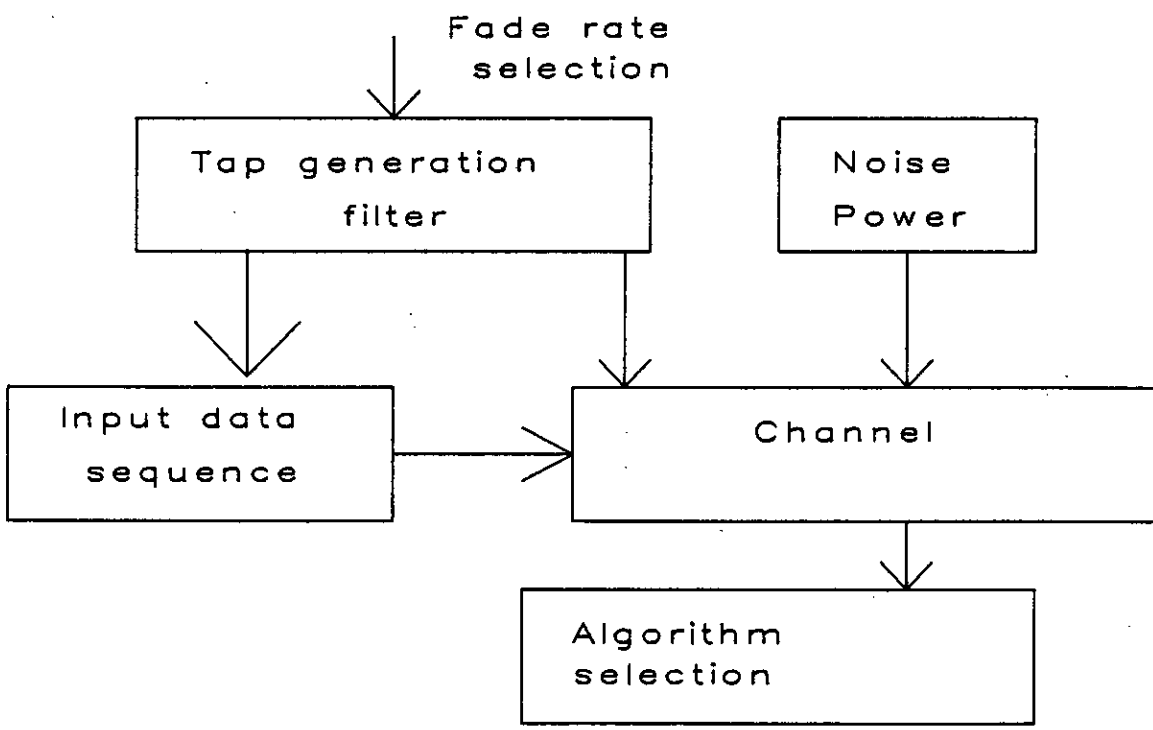


Figure 2.13 - Simulation model structure

2.5 IMPLEMENTATION OF THE CHANNEL MODEL

This section details the manner in which the channel model was implemented. All of the simulations performed for this thesis assumed a three path channel with a symbol rate of 2400 symbols per second. The only extraneous noise source was additive white Gaussian noise, impulse noise not being considered since it is generally overcome via coding and data interleaving which are outwith the scope of this thesis. Figure 2.13 illustrates the overall structure of the model.

The main difficulty in implementing the model lay in generating the tap gain functions in a manner that ensured ease of repeatability; this was achieved by generating a random number sequence with Gaussian characteristics with zero-mean and unit variance. This sequence was then filtered via a digital second order low pass Butterworth filter, the bandwidth of the filter being of the order of the fade rate introduced on to the signal and the filter being generated by the bilinear transformation. The characteristics of the filter are shown in Figures 2.14 and 2.15. This is a recognised method for generating Rayleigh fading characteristics and has been used by many researchers in the past[41, 42, 46-48]. Although not the perfect filter response, the Butterworth is convenient because it is easy to implement and has an approximate linear phase relationship; the structure of the filter is illustrated in Figure 2.16. A separate filter was used for each path, with a different input sequence, to ensure that each tap gain function was statistically independent. In order to scale each tap weight and ensure an overall channel gain of unity, the steady-state gain of the filter was calculated by means of calculus of residues. The behaviour of the tap gain functions for a filter bandwidth of 10Hz is illustrated in Figure 2.17.

Although a complete channel model would be complex in nature, for the simulations demonstrated in this thesis only a real channel was used. The reasoning for this being that the primary concern in this work was the performance of adaptive algorithms and all the information required is provided by a real channel simulation without unnecessarily increasing the level of complexity required to that of a complex

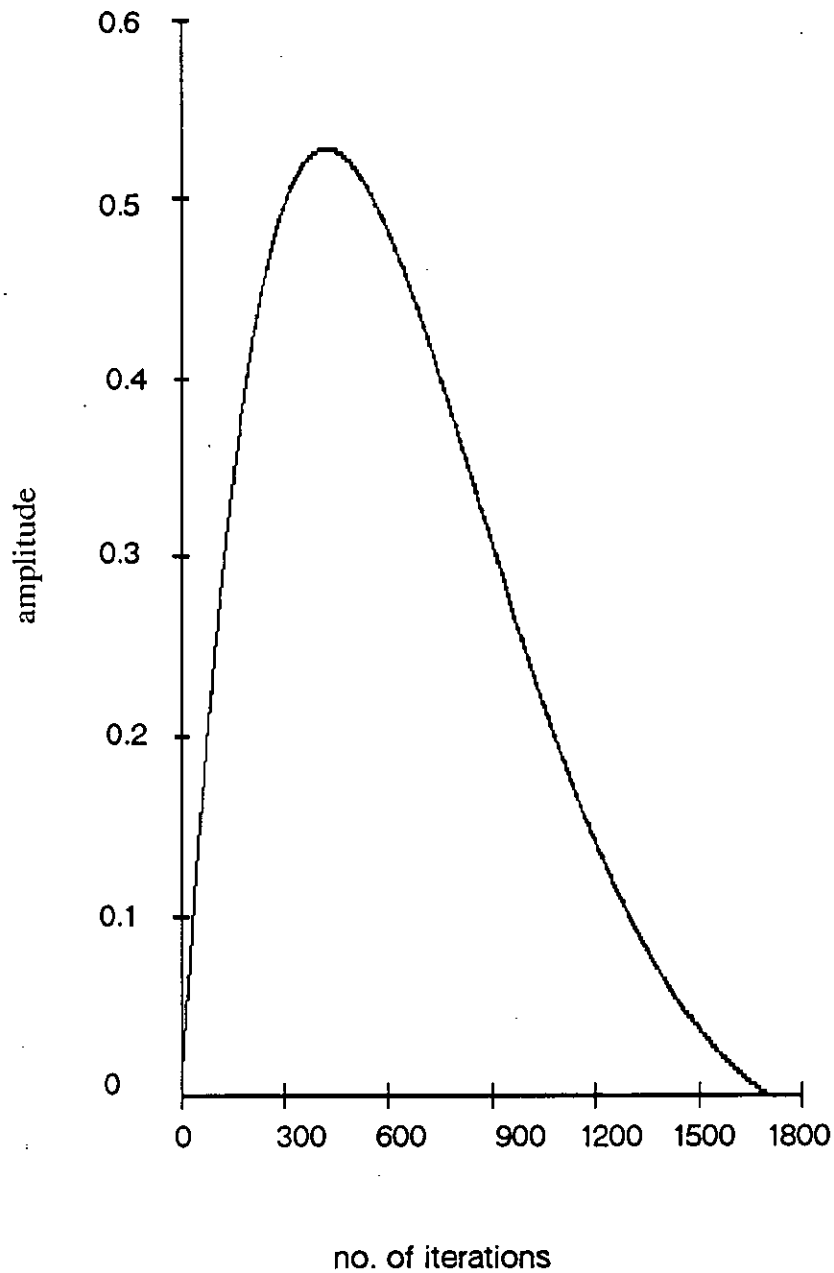


Fig. 2.14 – Impulse response of tap generation filter, bandwidth=1.0Hz, sampling frequency=2400Hz.

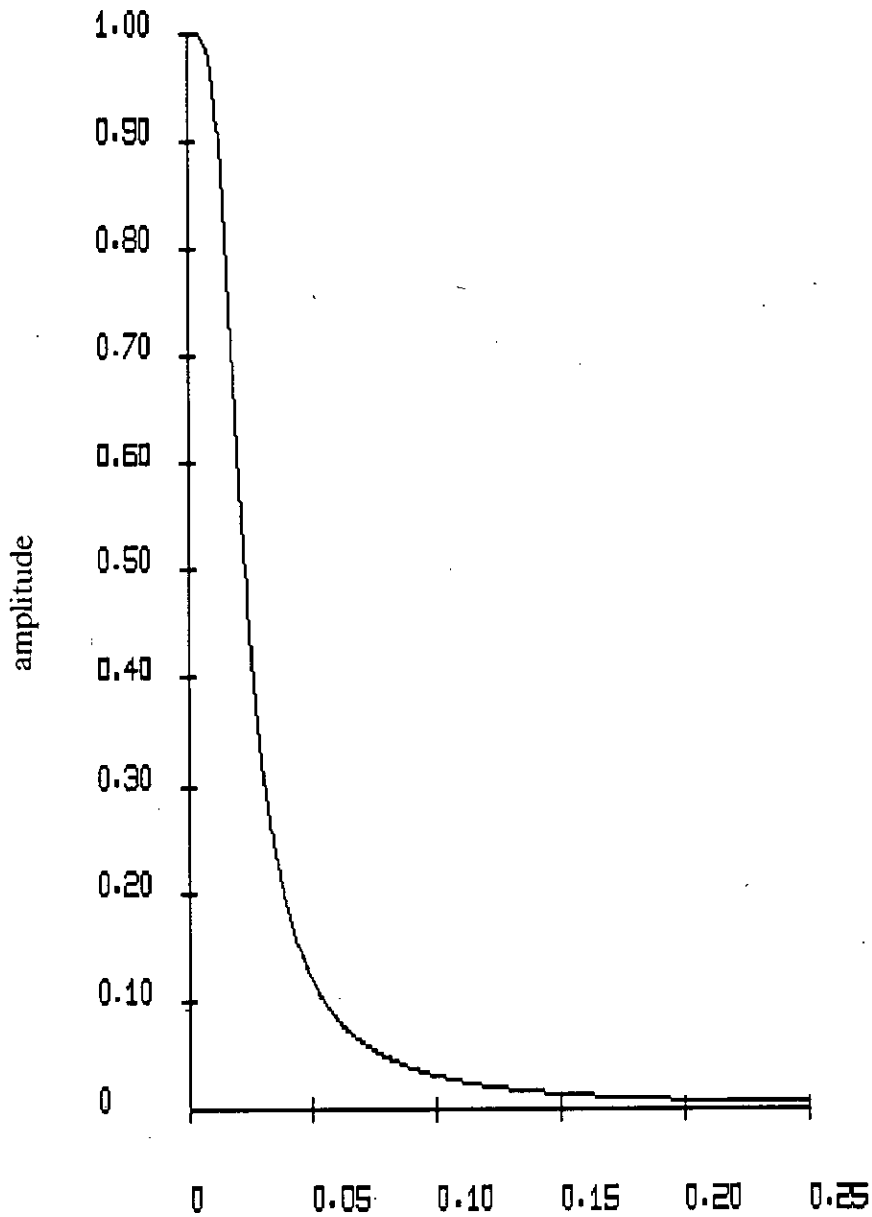


Figure 2.15 - Normalised frequency response of tap generation filter, bandwidth = 1.0Hz, sampling frequency = 2400Hz.

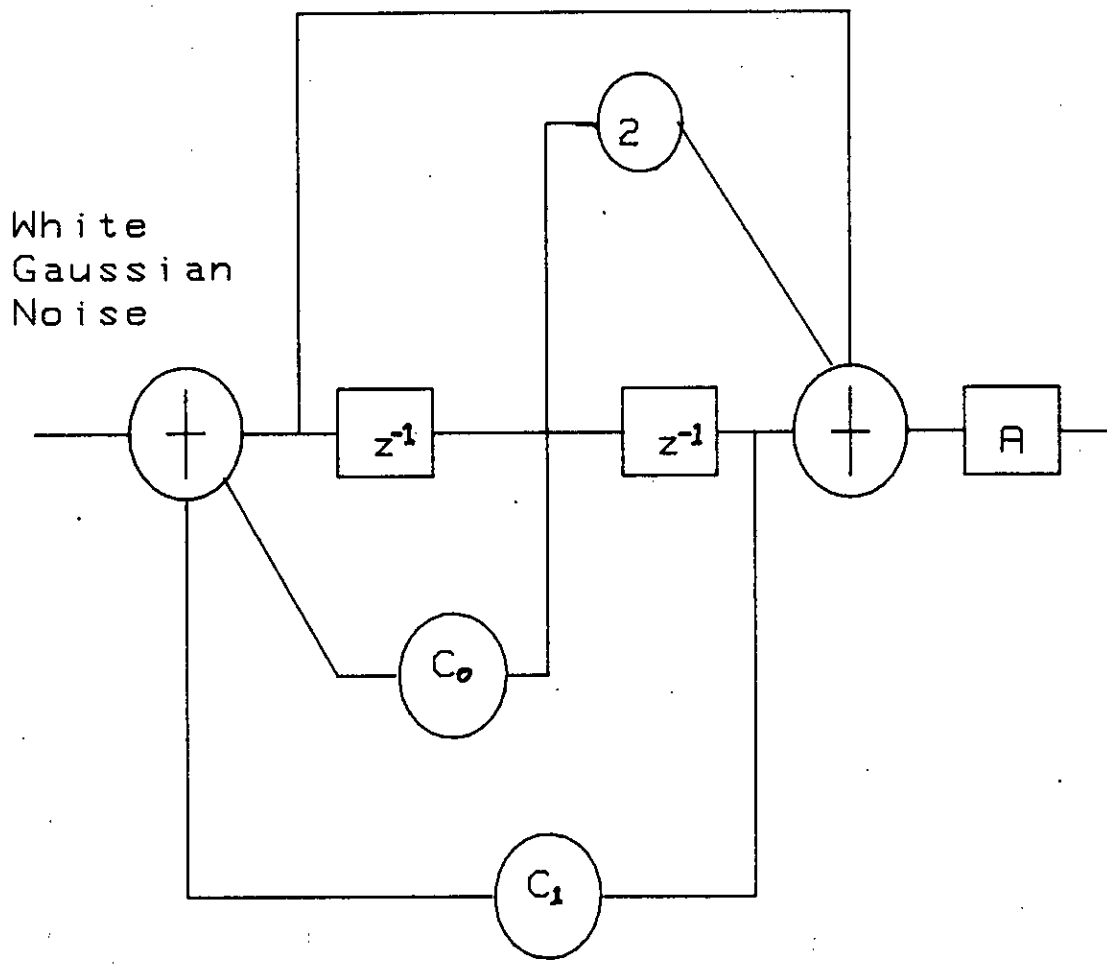


Figure 2.16 - Tap generation filter

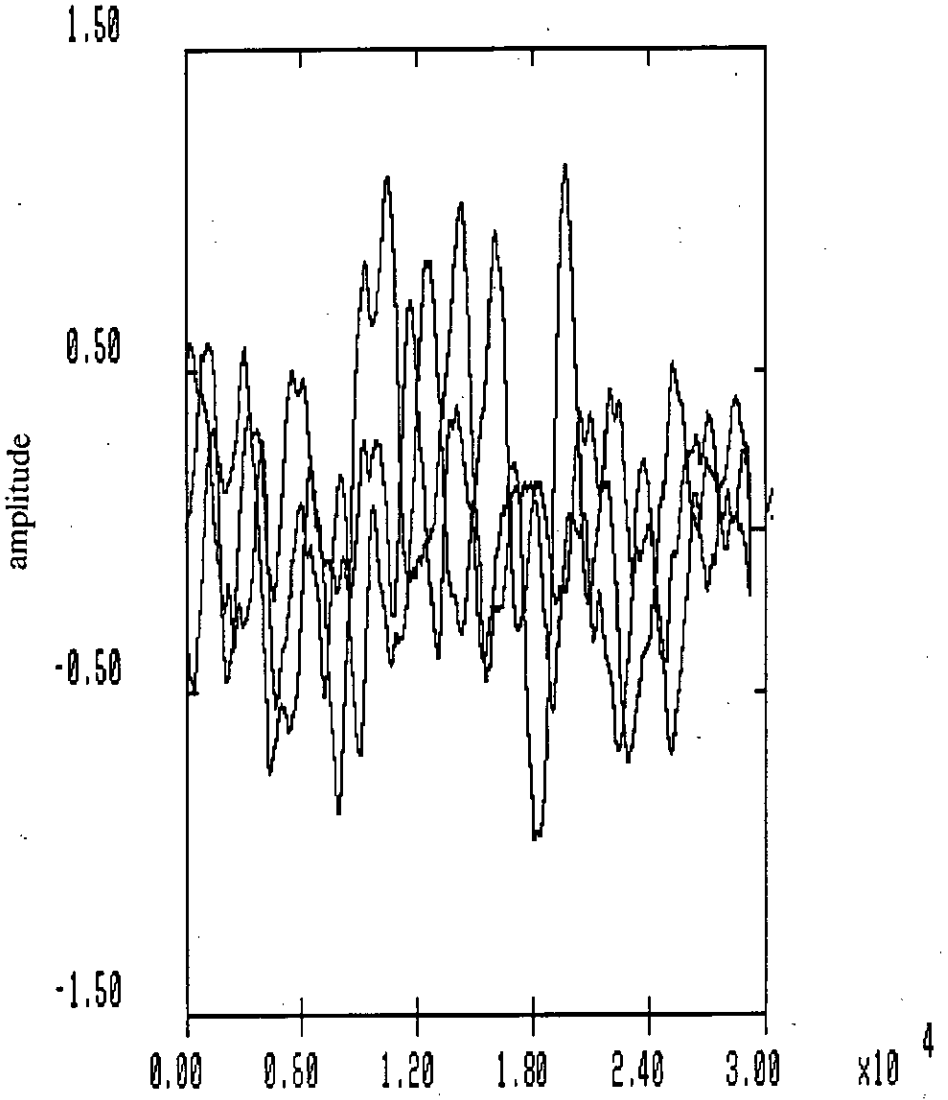


Fig. 2.17 – Trajectory of 3 taps for simulated HF channel with fade rate=1Hz.

channel model. This approach has been adopted by several other researchers in the area, namely Eleftheriou and Falconer [50] and Ling and Proakis[23], a full analysis of the performance would require a more general model.

2.6 SUMMARY

In summary, this chapter has introduced the HF channel and illustrated its nature and the manner in which radio communication is achieved in it. A mathematical and physical justification has been presented which relates the frequency selective fading channel model used for simulation purposes in this thesis to the physical reality. In conclusion the manner in which the channel model was implemented was presented.

Chapter 3

PERFORMANCE STUDY OF ADAPTIVE FIR FILTERING ALGORITHMS AS HF CHANNEL ESTIMATORS

3.1 INTRODUCTION

This chapter is primarily concerned with an analysis of the performance of two adaptive FIR filtering algorithms as HF channel estimators, i.e. the LMS and the RLS. At the present time considerable research effort is being expended to develop adaptive equalisers for use in communication systems where the channel is time-varying in nature, as in the HF channel. The overall aim is to allow data communication at speeds greater than are currently possible.

The chapter is structured such that an initial brief outline of least squares theory is presented with it's relationship to the LMS and RLS adaptive algorithms discussed as well as the related topics of data windowing, numerical robustness and so called fast algorithms. Then the analysis of both algorithms is presented, that of the LMS being merely a summary of the work of Macchi [51, 52] and Eweda and Macchi [53, 54] and is presented for comparative purposes. Finally simulations are presented which illustrate the performance of the algorithms and the accuracy of the theoretical predictions.

It has generally been assumed that the RLS algorithm would be suitable for use in time-varying environments because of it's fast convergence properties in a time-invariant environment. However, recently published work [55-58] would appear to suggest that in both time-varying and high noise environments the RLS suffers a considerable degradation in performance as demonstrated by a slower rate of convergence and higher minimum MSE.

In this work the direct modelling (channel estimator) approach was chosen for analysis as opposed to indirect modelling, i.e. equalisation, because the only unknown time variation considered is that of the channel coefficients, the input signal being stationary. This makes the analysis more tractable and isolates the tracking

performance of the algorithms, since this is part of the algorithm's steady-state behaviour, as opposed to transient behaviour which is related to the initial convergence behaviour of the algorithm. New theoretical results reinforced by simulation are presented here which illustrate the optimal performance bounds for the RLS operating in such an environment. The performance of the RLS is characterised for various values of

i)the level of colouration of the input signal,

ii)the level of additive noise in the system,

iii)the level of time-variations.

The decoupling of the overall error achieved by the RLS in estimating the system into a measurement term and a lag term as in [21] , is used to illustrate the degradation in performance due to the high additive noise and/or time variations in the system. The relative effect of these errors is shown theoretically and reinforced with simulations. The selection of λ, the exponential windowing factor, to give optimal performance is considered and the trade-off required in it's selection is discussed and illustrated.

To provide a comparison for the results obtained in this work use was made of the considerable body of literature [51-54] on the performance of the LMS adaptive algorithm in a time-variant environment, in particular the expression obtained by Macchi in [51] for the residual steady-state mean squared error (MSE) in a nonstationary environment.

3.2 ADAPTIVE FIR FILTERING ALGORITHMS AND LEAST SQUARES ESTIMATION THEORY

Prior to considering the particular adaptive FIR filtering algorithms studied here it is worthwhile to consider a brief review of least-squares estimation theory and it's relationship to adaptive FIR filtering algorithms. A brief outline of the numerical

aspects of least squares estimation theory and the role of data windowing is also considered.

3.2.1 Least Squares Estimation Theory

The need to estimate the state of a stochastic dynamic system from noisy measurements is important in many aspects of engineering and natural science. It was Gauss who first considered this problem in 1805 when studying the movement of celestial bodies in[13] as did Legendre in[14] independently. The basic idea of least squares, as applied to parameter estimation problems is that estimates of the parameter of interest are selected such that the output of the model approximates the data as accurately as possible, as measured via the sum of the squares of the differences. This statement may be expressed more formally by considering the following simple estimation problem.

Consider a stochastic process whose mean value is a linear function of some parameter vector $\underline{\beta}$. Thus the least squares estimate attempts to minimise the error between the desired signal and the estimated, i.e.,

$$\epsilon_i = y_i - \underline{x}_i^T \underline{\beta}. \tag{3.1}$$

To minimise this in the least squares sense it is necessary to minimise,

$$J = \sum_{i=0}^{k} \epsilon_i^2 = \sum_{i=0}^{k} (y_i - \underline{x}_i^T \underline{\beta})^2. \tag{3.2}$$

Consequently, by differentiating expression (3.2) and setting the result to zero the so called Normal equation is obtained[59].

The early work in the application of least squares theory did not consider the estimation problem in a probabilistic sense; rather it was viewed as a deterministic problem in terms of the error minimisation. It was not until Wiener in [60] that this was achieved for stationary continuous time systems. Wiener reduced the continuous filtering problem to the solution of an integral equation, the so called Wiener-Hopf equation. The general linear nonstationary problem was resolved in the pioneering

work of Kalman [61] and Kalman and Bucy[62]. This chapter is primarily concerned with the performance of adaptive FIR filtering algorithms; therefore a summary of the relevant areas only will be presented; more detailed and general accounts of least-squares estimation theory may be found in any of [59, 63-65].

The structure of the estimation problem considered here is illustrated in Figure 3.1, where a random sequence $\{x(n)\}$ is input to a time varying system and the resultant noise contaminated output, y_k, is generated. The problem is to estimate the impulse response vector, $\underline{h}_k$, of the system given that $y_k = \underline{h}_k^T \underline{x}_k + n_k$, and to do so in some optimal manner. The error generated is $e_k = y_k - \hat{y}_k$, where $\hat{y}_k = \hat{\underline{h}}_k^T \underline{x}_k$.

The widely accepted approach has been to minimise the mean-squared value of the error, e_k, the mean-squared value representing a cost or loss function, [59, 63-65] , so called because it indicates the penalty associated with an incorrect estimate. It may be readily shown[59, 63-65] , that given this criterion the optimal estimate of $\underline{h}_k$, denoted by $\underline{h}_{opt}$, is obtained from the so called Wiener solution;

$$\underline{h}_{opt} = \Phi_{xx}^{-1} \Phi_{xy} \tag{3.3}$$

where $\Phi_{xx} = E\ [\ \underline{x}_k \underline{x}_k^T\ ]$ and $\Phi_{xy} = E\ [\ \underline{x}_k y_k\ ]$, and $E[.]$ denoting the expectation operator.

Clearly the presence of the expectation operator in equation (3.3) precludes any practical application. It is the purpose of adaptive FIR filtering algorithms to determine $\underline{h}_{opt}$ given only access to the data sequences. Only two approaches to this problem are considered here, the so called LMS algorithm and the RLS algorithm and these are dealt with in the following sections.

3.2.2 The Least Mean Squares Adaptive Algorithm

The simplest approach adopted to achieve the minimisation of the MSE criterion was by means of a stochastic gradient search technique[66]. The LMS algorithm, first suggested by Widrow in 1960 [67] , is the best known of these techniques. The LMS utilises a weighted estimate of the gradient to recursively estimate in time the optimal

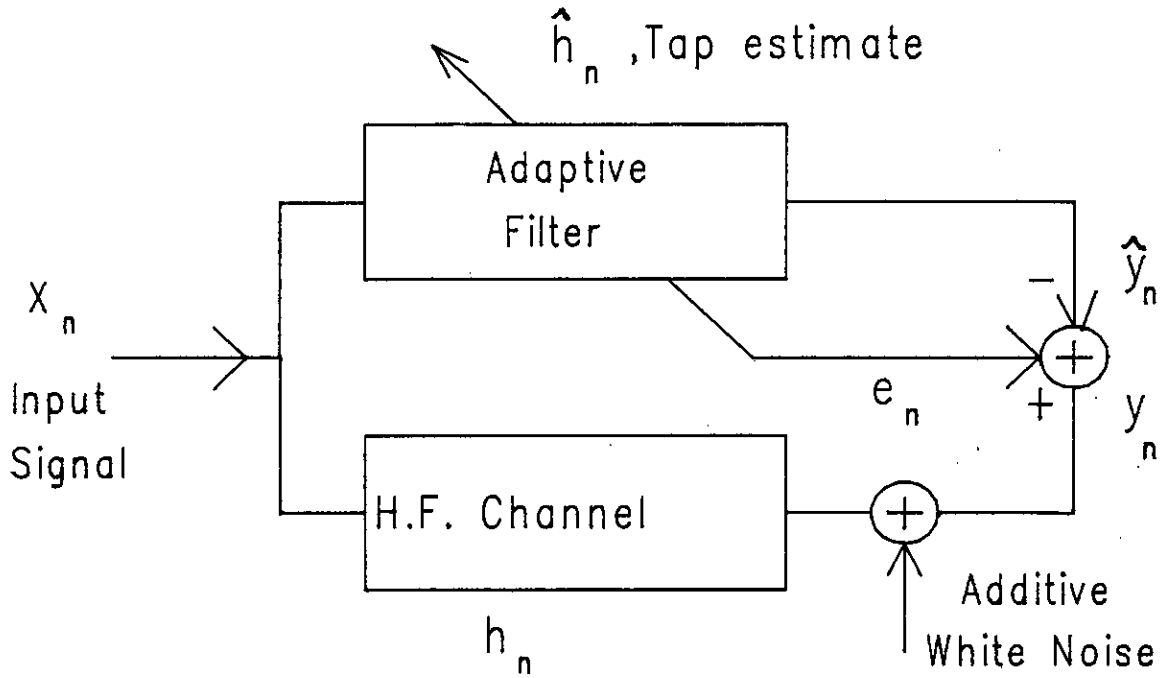


Figure 3.1 - Adaptive channel estimation

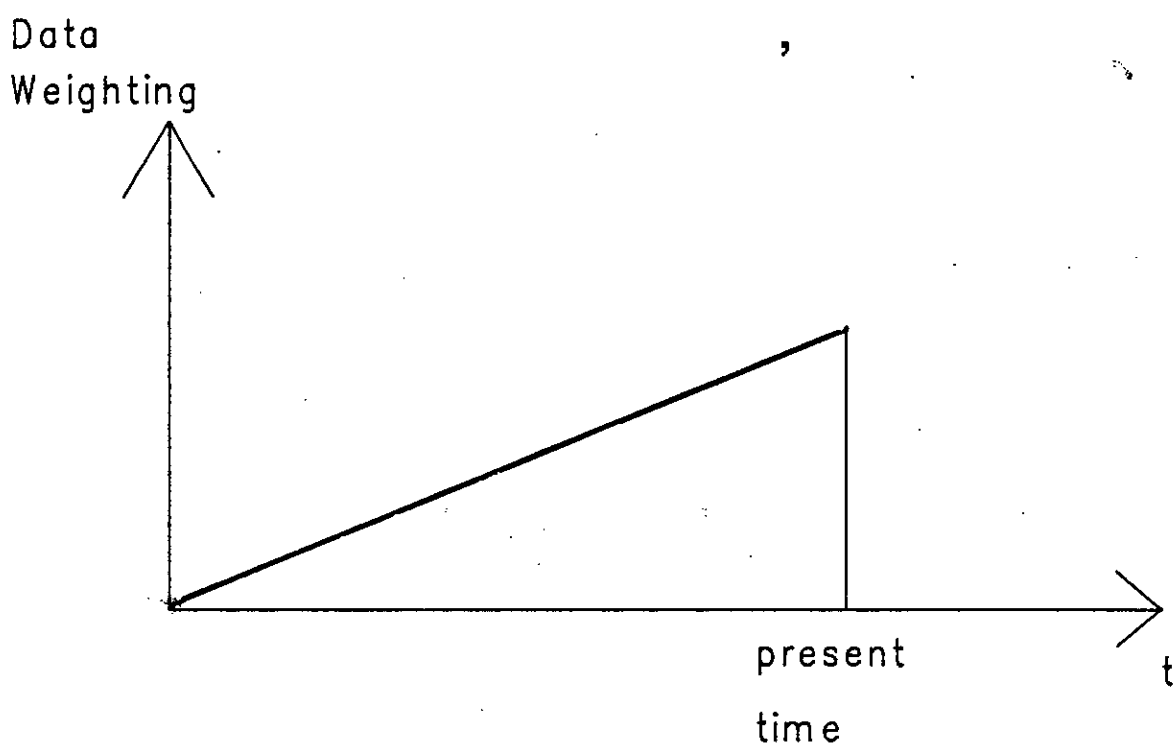


Figure 3.2 - Data windowing concept

tap weights, that is,

$$\hat{\underline{h}}_{k+1} = \hat{\underline{h}}_k - \mu \hat{\underline{\nabla}}_k \tag{3.4}$$

where $\hat{\underline{h}}_k$ represents the estimate of $\underline{h}_{opt}$ at sample instant k and the vector $\hat{\underline{\nabla}}_k$ is the estimate of the gradient of the MSE cost function given the particular estimate $\hat{\underline{h}}_k$. The scalar parameter μ is a convergence factor which determines stability and rate of adaptation of the algorithm. The exact gradient,

$$\underline{\nabla}_k = \frac{\delta \xi}{\delta \underline{h}}(\underline{h}_k) \tag{3.5}$$

which is simply equal to the following,

$$\frac{\delta \xi}{\delta \underline{h}} (E\ [\ (y_k - \underline{h}_k^T \underline{x}_k\ )^2\ ]) \tag{3.6}$$

and thus becomes,

$$\underline{\nabla}_k = -2\ E\ [\ \underline{x}_k\ (y_k - \underline{h}_k^T \underline{x}_k\ )]. \tag{3.7}$$

The estimate, $\hat{\underline{\nabla}}_k$, could be obtained by utilising a time average as opposed to an ensemble average, (assuming ergodicity), but clearly the time average may only be obtained from a single instant since $\hat{\underline{h}}_k$ changes at each sample. Consequently when the time average is used the estimate of the gradient becomes,

$$\hat{\underline{\nabla}}_k = -2 \underline{x}_k e_{k+1}, \tag{3.8}$$

which when substituted in expression (3.6) results in,

$$\hat{\underline{h}}_{k+1} = \hat{\underline{h}}_k - 2\mu \underline{x}_{k+1} e_{k+1}. \tag{3.9}$$

The selection of the parameter μ is crucial to both the performance and stability of the algorithm. The work of Feuer and Weinstein [68] provided a criterion for the stability of the LMS given the length of the filter and the eigenstructure of the input signal. As regards the performance of the algorithm there have been many publications which have looked at the algorithm in many situations, [15] and references therein. The limitations of the LMS are summarised below.

The LMS is by far the simplest (and oldest) adaptive algorithm; however it suffers from a relatively slow initial convergence rate which is affected quite severely by the

eigenstructure of the input signal. This is explained fully in [66] , but intuitively may be explained by the fact that the gain μ can adjust to only one mode of the system at a time. A highly coloured signal will have a large eigenvalue spread and as a result many modes. A fuller description of the performance and characteristics of the algorithm may be found in any of the adaptive filter textbooks referenced in chapter 1.

As a concluding comment, the LMS is the oldest adaptive algorithm but it is also by far the simplest to implement, requiring only $2K$ operations per iteration, K being the order of the filter. Consequently the LMS is the most widely applied of all adaptive algorithms in spite of it's many drawbacks and limitations.

3.2.3 Recursive Least Squares

As was stated previously the exact measures of Φ_{xx} and Φ_{xy} are not readily available. The RLS algorithm utilises the data sequences which are available to construct estimates of these measures. These estimates are as follows,

$$\underline{r}_{xx}(k) = \sum_{n=0}^{k} \underline{x}_n \underline{x}_n^T, \tag{3.10}$$

and,

$$\underline{r}_{xy}(k) = \sum_{n=0}^{k} \underline{x}_k y_k, \tag{3.11}$$

which are of course the auto and cross-correlation estimates for the sequences $\underline{x}_n$ and y_n. With these an estimate of the optimum tap weight vector, at time instant k, may be constructed from,

$$\underline{r}_{xx}(k)\hat{\underline{h}}_k = \underline{r}_{xy}(k). \tag{3.12}$$

In a practical situation a recursive formulation of this expression is obviously required and this achieved via the following substitutions,

$$\underline{r}_{xx}(k) = \underline{r}_{xx}(k-1) + \underline{x}_k \underline{x}_k^T \tag{3.13}$$

and

$$\underline{r}_{xy}(k) = \underline{r}_{xy}(k-1) + \underline{x}_k y_k, \tag{3.14}$$

in conjunction with the expression,

$$\underline{r}_{xx}(k-1)\underline{h}_{k-1} = \underline{r}_{xy}(k-1). \tag{3.15}$$

The recursive formulation may then be written as,

$$\hat{\underline{h}}_k = \underline{h}_{k-1} + \underline{r}_{xx}^{-1}(k)\underline{x}_k e_k. \tag{3.16}$$

The inverse , $\underline{r}_{xx}^{-1}(k)$, is obtained using the Sherman-Morrison Identity [69] ,

$$\underline{r}_{xx}^{-1}(k) = \underline{r}_{xx}^{-1}(k-1) - \frac{\underline{r}_{xx}^{-1}(k-1)\underline{x}_k \underline{x}_k^T \underline{r}_{xx}^{-1}(k-1)}{1 + \underline{x}_k^T \underline{r}_{xx}^{-1}(k-1)\underline{x}_k}. \tag{3.17}$$

The above recursion requires K^2 products per iteration and indicates the penalty involved in implementing the RLS algorithm. Algorithms based on the RLS may be obtained with complexity of order K, but like the RLS they suffer from a high degree of numerical instability and invariably require some form of 'numerical rescue', i.e. periodic initialisation even when implemented on 32 bit floating point processors[70-72]. In it's favour the RLS is guaranteed to converge within $2K$ iterations, K being the order of the filter, provided that the system under consideration is stationary.

The tracking performance of the algorithm in the above form is very poor as a result of the growing memory form of the correlation estimates in expressions (3.10) and (3.11). This problem is normally overcome by the use of data windows [73, 74] which give greatest weight to most recent data. The most commonly implemented windowing function being in the form of an exponential function [74] . Windowing functions will be discussed in slightly more detail in the proceeding section.

To summarise the RLS, in low noise and stationary environments, will converge within $2K$ iterations, K being the order of the filter; it is relatively complex to implement and displays a high degree of numerical sensitivity although recent formulations suggest that this problem may have been overcome[75, 76].

3.2.4 Other Aspects of Least Squares Estimation

Data Windowing

As has been stated the tracking performance of the algorithm in the above form

is very poor. This problem is normally overcome by the use of data windows which give greatest weight to most recent data, as in Figure 3.2, the most commonly implemented windowing function being in the form of an exponential function. This results in a modified cost function from the original sum of errors squared to,

$$\sum_{n=0}^{k} (y_n - \hat{y}_n)^2 \lambda^{k-n} \quad 0<\lambda<1. \tag{3.18}$$

There are clearly many possible window functions but in general only the exponential window [74] and sliding window [73] are ever considered. The exponential window is generally used as less computation is required, also the work of Porat [77] would suggest that the sliding window would offer no significant performance advantage.

The value of λ, the exponential windowing factor, used is normally chosen to lie in the range, $0.9 < \lambda < 1.0$ because of constraints on the accuracy of the correlation function estimates it imposes. This raises a question as to the selection of λ in time-variant environments where it is required that the variations in the parameters to be estimated be small within the window length. The result of Porat in [77] provides a means of equating λ to an equivalent window length M, and this shows up the conflicting requirements. That is as the time-variations increase then the window length must become shorter, hence λ becomes smaller (possibly below 0.9), but the correlation estimate requires that λ remain as close to unity as possible.

Finite Precision Effects

The finite nature of the digital machines on which any least-squares algorithm is to be implemented must be taken into account in any assessment of performance. The inherently complex nature of many least-squares estimation algorithms ensures that they will suffer from errors as a result of digital word truncation and round off in matrix multiplication. This idea of finite word length and its effects is illustrated in Figure 3.3.

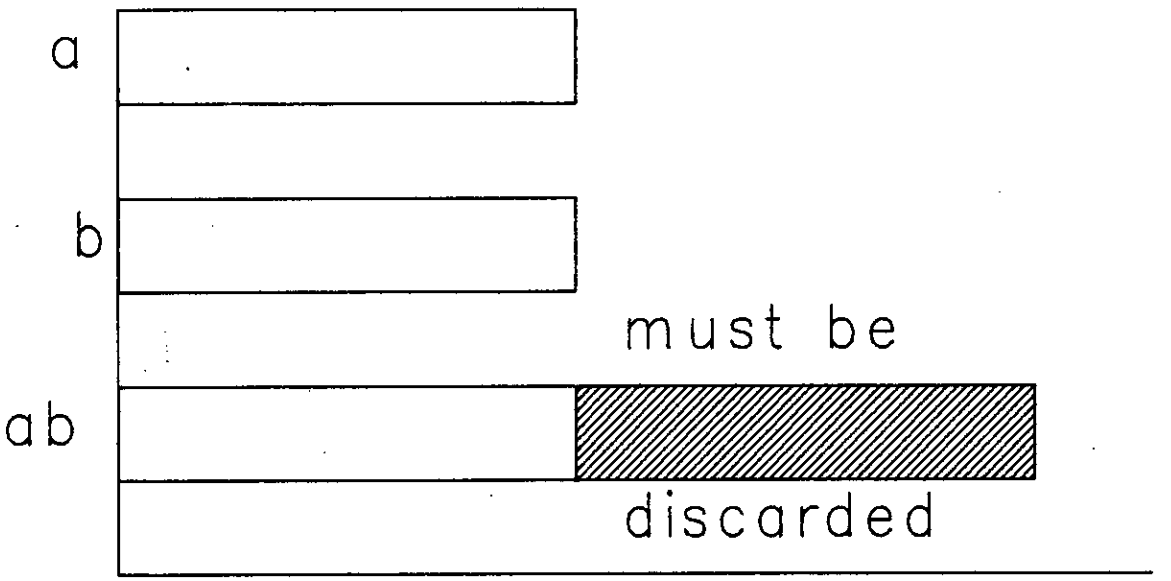


Figure 3.3 - Errors in multiplication due to finite nature of digital machines

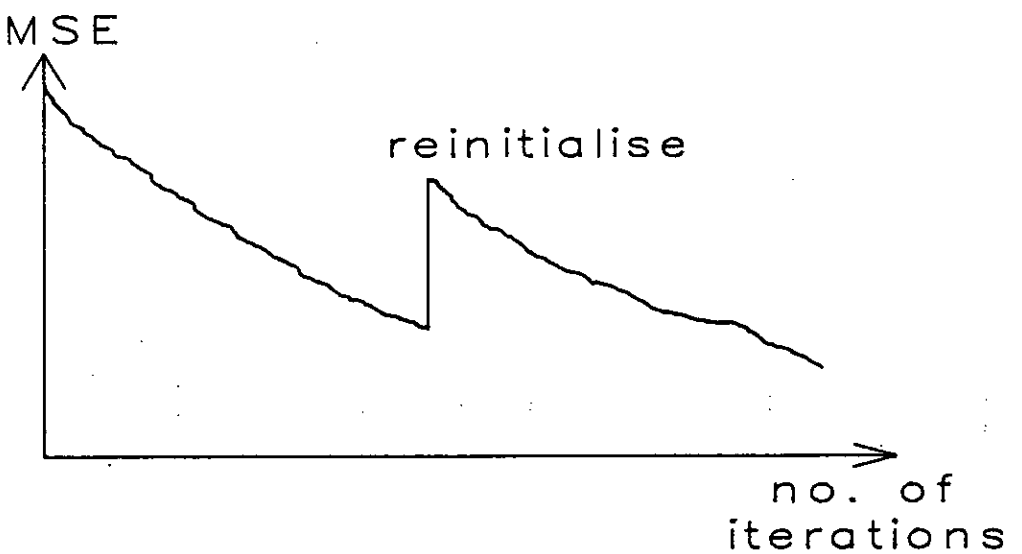


Figure 3.4 - Numerical rescue via reinitilisation

There are many techniques which have been adopted to try and overcome the numerical instability that results from finite precision and these may be divided into two broad categories which can be termed algorithmic manipulation as in [75, 76, 78] and numerical rescue[79-81]. Examples of the former are square root formulations as in [78] or decompositions as in[75, 76]. Numerical rescue is simpler to implement since the technique is simply to detect the onset of numerical instability, i.e. divergence, and then reinitialise the algorithm to some preset values as illustrated in Figure 3.4. It is worth noting that numerical rescue is always likely to be necessary since irrespective of the algorithm manipulations adopted, they will only delay the onset of numerical instability due to the inherent finite nature of digital machines[71].

Fast Algorithms

As a result of the inherently complex nature of algorithms such as the RLS, (where of order K^2 operations are required every iteration), many so called *fast* algorithms have been developed [73, 82, 83] to reduce the computational load when implementing such algorithms. All of these so called *fast* algorithms reduce the complexity to order K operations per iteration by utilising various properties of matrices. It must be emphasised however that these algorithms only exhibit the same level of performance associated with the conventional RLS. The term *fast* referring only to the computational load.

3.3 ANALYSIS OF ADAPTIVE FIR FILTER ALGORITHMS AS HF CHANNEL ESTIMATORS

3.3.1 Introduction

The problem considered is that of direct modelling of a time-varying system which is characterised by the tapped delay line model of chapter 2, where the time-varying taps are generated by filtering random white Gaussian noise through a filter; in this case is a 2nd order digital Butterworth filter is used with bandwidth very much narrower than the symbol rate. The construction used to carry out the system

identification is as discussed in chapter 2. That is, the input to the system is also the input to the adaptive algorithm so that only the time-variations of the system under investigation are being tracked. In the inverse modelling situation decision errors, inputs which would be both nonstationary and coloured would lead to degradation in the performance of the algorithms for reasons not associated with their tracking performance.

Using the representation of Figure 3.1 allows the system output at iteration k to be written as,

$$y_k = \underline{x}_k^T \underline{h}_k + n_k \tag{3.19}$$

where, $\underline{x}_k$ is the input signal vector with the superscript T representing the transpose operator, $\underline{H}_k$ is the tap weight vector of the time varying system and n_k represents the unobservable measurement noise in the system, which in this case is additive white Gaussian noise (AWGN).

The noise, n_k, and input signal vector $\underline{x}_k$ are assumed to satisfy the following assumptions.

A1:-The sequence $\underline{x}_n$ is stationary and Gaussian in nature with finite moments.

A2:-The sequence n_k is identically distributed and independent of $\underline{x}_n$.

A3:-The time variations of $\underline{H}_k$ are random and independent of $\underline{x}_n$ and n_k.

A4:- The estimate of the autocorrelation matrix, $\hat{R}_k = \sum_{j=1}^{k} \underline{x}_j \lambda^{k-j} \underline{x}_j^T$, of the input signal vector $\underline{x}_j$ can, in the limit, be represented as $\hat{R}_k = \Phi(1-\lambda)^{-1}$, where $\Phi = E[\underline{x}_k \underline{x}_k^T]$ and is the exact autocorrelation matrix of the input signal vector $\underline{x}_k$ and initially in the case considered here since the input is white $\Phi = I$, where I is the identity matrix. That is,

$$\lim_{k \to \infty} \hat{R}_k = I(1-\lambda)^{-1}$$

This assumes that λ lies close to 1 (normally > 0.9), if this condition is not satisfied then this assumption cannot be considered valid. It should be noted that assumptions

A1 to A3 are identical to those used by Macchi [52] in her analysis of the LMS in a nonstationary environment.

It may be argued that these assumptions are not representative of the scenario in which the algorithms have to operate, however they provide a means of analysing the performance of the algorithm and as will be shown later the theoretical predictions obtained from the analysis are in close agreement with simulation results and this in itself is sufficient justification for them.

Proceeding with the definition of the following variables,

$$\underline{d}_k = \underline{h}_{k+1} - \underline{h}_k \tag{3.20}$$

and, $\underline{d}_k$ is a measure of the nonstationarity of the channel. Also,

$$\underline{q}_k = \hat{\underline{h}}_k - \underline{h}_k \tag{3.21}$$

which is the tap weight error vector (or misadjustment).

The error e_k would normally be written as,

$$e_k = y_k - \underline{x}_k^T \hat{\underline{h}}_k \tag{3.22}$$

Where $\hat{\underline{h}}_k$ represents the tap weight vector estimate. By using the expressions (3.20) to (3.22) it is possible to rewrite (3.23) as below,

$$e_k = n_k - \underline{x}_k^T \underline{q}_k . \tag{3.23}$$

In this way it is expressed in terms of the unobservable noise in the system and the tap weight error vector.

3.3.2 LMS Mathematical Analysis

This analysis is a brief summary of the work on the LMS by Eweda and Macchi in [52] and[53].

The LMS algorithm is normally written as was indicated previously as,

$$\hat{\underline{h}}_{k+1} = \hat{\underline{h}}_k + 2\mu e_k \underline{x}_k \tag{3.9}$$

where μ is a small positive constant. From this a recursion for the error in the tap estimate is obtained,

$$\underline{q}_{k+1} = (I - 2\mu \underline{x}_k \underline{x}_k^T)\underline{q}_k + 2\mu n_k \underline{x}_k - \underline{d}_k \tag{3.24}$$

This can be split into two identifiable terms,

$$\underline{q}_{k+1}^g = (I - 2\mu \underline{x}_k \underline{x}_k^T)\underline{q}_k^g + 2\mu n_k \underline{x}_k \tag{3.25a}$$

and,

$$\underline{q}_{k+1}^l = (I - 2\mu \underline{x}_k \underline{x}_k^T)\underline{q}_k^l - \underline{d}_k \tag{3.25b}$$

where (3.25a) is associated with the gradient noise and is present even in stationary systems and (3.25b) is the lag error and demonstrates the error contributed by the time variation of the system. Using a standard result of linear algebra [84] it is possible to make the following substitution,

$$U_{j,k} = (I - 2\mu \underline{x}_k \underline{x}_k^T)(I - 2\mu \underline{x}_{k-1} \underline{x}_{k-1}^T) \ldots\ldots (I - 2\mu \underline{x}_{j+1} \underline{x}_{j+1}^T) \tag{3.26}$$

and thus by solving in the usual manner,

$$\underline{q}_{k+1}^g = 2\mu \sum_{j=1}^{j=k} U_{j,k} n_j \underline{x}_j \tag{3.27a}$$

and,

$$\underline{q}_{k+1}^l = -\sum_{j=1}^{j=k} U_{j,k} \underline{d}_j \tag{3.27b}$$

In the limit the Steady State MSD is clearly,

$$\lim_{k\to\infty} E\,|\underline{x}_k^T \underline{q}_k|^2 = \lim_{k\to\infty} E\,|\underline{x}_k^T \underline{q}_k^g|^2 + \lim_{k\to\infty} E\,|\underline{x}_k^T \underline{q}_k^l|^2 \tag{3.28}$$

that is,

$$\textit{Total MSD} = \textit{Stationary MSD} + \textit{Lag MSD} \tag{3.29}$$

Clearly the limits of each term are required in order to obtain the steady state MSD. Thus,

$$\lim_{k\to\infty} E\,|\underline{q}_k^g|^2 = \lim_{k\to\infty} E\,|2\mu \sum_{j=1}^{j=k-1} U_{j,k-1} n_j \underline{x}_j|^2 \tag{3.30a}$$

and also,

$$\lim_{k\to\infty} E\,|\underline{q}_k^l|^2 = \lim_{k\to\infty} E\,|\sum_{j=1}^{j=k-1} U_{j,k-1}\underline{d}_j|^2 \tag{3.30b}$$

It is demonstrated in [52] that these limits are finite by means of standard statistical theory. This leads to the following,

$$MSD = 2N(1-\mu KS)^{-1}(2\mu KS + D/2N\mu) \tag{3.31}$$

where, N is the noise power in the system, K is the order of the system with,

$$|\underline{x}_k|^2 = KS$$

and,

$$D = E[\,|\underline{d}_k|^2]$$

3.3.2 RLS Mathematical Analysis

If the RLS algorithm is written below as previously indicated,

$$\hat{\underline{h}}_{k+1} = \hat{\underline{h}}_k + \hat{R}_k^{-1} e_k \underline{x}_k \tag{3.17}$$

Where $\hat{R}_k$ is the estimated autocorrelation of the input signal vector $\underline{x}_k$.

As has been noted earlier this representation of the RLS is similar to the structure of the LMS with 2μ replaced by the inverse of the estimated autocorrelation of the input signal vector.

Using these expressions and by employing some algebraic manipulation a recursive expression for the parameter error vector $\underline{q}_{k+1}$ can be obtained.

$$\underline{q}_{k+1} = (I - \hat{R}_k^{-1}\underline{x}_k\underline{x}_k^T)\underline{q}_k + \hat{R}_k^{-1} n_k \underline{x}_k - \underline{d}_k \tag{3.32}$$

It is clear that expression (3.32) can be split into two clearly identifiable terms as before and these are shown below,

$$\underline{q}_{k+1}^g = (I - \hat{R}_k^{-1}\underline{x}_k\underline{x}_k^T)\underline{q}_k^g + \hat{R}_k^{-1} n_k \underline{x}_k, \tag{3.33a}$$

and,

$$\underline{q}_{k+1}^l = (I - \hat{R}_k^{-1}\underline{x}_k\underline{x}_k^T)\underline{q}_k^l - \underline{d}_k \tag{3.33b}$$

The first term, $\underline{q}_k^s$, can be viewed as a measurement noise term and is present even in the stationary situation. The second term, $\underline{q}_k^l$ can be considered as a lag term and is associated with the time variations of the system. As a result of assumptions A2 & A3 it may be assumed that the two contributions to the error in expression (5) are independent of each other. In the limit the excess steady state mean square deviation (MSD) is,

$$\lim_{k\to\infty} E\,|\underline{x}_k^T \underline{q}_k\,|^2 = \lim_{k\to\infty} E\,|\underline{x}_k^T \underline{q}_k^s\,|^2 + \lim_{k\to\infty} E\,|\underline{x}_k^T \underline{q}_k^l\,|^2 \tag{3.34}$$

that is,

Total MSD = Stationary MSD + Lag MSD.

The limits of each term are required to be finite in order to obtain the steady state excess MSE. Since the the input signal vector and tap weight error vector are independent and the input signal is stationary it is sufficient to show that the squared norm of the tap weight error vector, q_k, is finite in the limit. The analysis for the general case i.e. not utilising assumption (A4) is possible [85] but extremely complex. For the situation considered here it is unnecessary since by utilising assumption (A4) it is possible to substitute $(1-\lambda)I$ for $\hat{R}^{-1}$ and the proof that the tap weight error vector is finite in the limit then follows as in [52] . Once the limits are shown to be finite an expression for the excess steady state MSD can be obtained as follows.

Using the recursions of (3.33a) and (3.33b) and taking the measurement noise term first,

$$E[\,|\underline{q}_{k+1}^s\,|^2] = E[\,|(I - \hat{R}_k^{-1}\underline{x}_k\underline{x}_k^T)\underline{q}_k^s + \hat{R}_k^{-1} n_k \underline{x}_k\,|^2] \tag{3.35}$$

Now if $E[\,|n_k\,|^2] = N$ and we replace the inverse of the estimated autocorrelation matrix , $\hat{R}_k^{-1}$ by $(1-\lambda)I$ using assumption (A4), where I is the identity matrix, we obtain

$$E[\,|\underline{q}_{k+1}^s\,|^2] = E[\,|\,(\,I - (\,1-\lambda\,)\,\underline{x}_k\underline{x}_k^T\,)\,\underline{q}_k^s + (\,1-\lambda\,)n_k\underline{x}_k\,|^2\,] \tag{3.36}$$

Expanding the expressions and using assumption A2 this becomes,

$$E[\,|\underline{q}_{k+1}^{g}|^2] = E[\,|\,(I-(1-\lambda)\,\underline{x}_k\underline{x}_k^T)\underline{q}_k^{g}|^2] + (1-\lambda)^2 E[\,|\,n_k\underline{x}_k\,|^2\,] \quad (3.37)$$

so that,

$$E[\,|\underline{q}_{k+1}^{g}|^2] = E[\,\underline{q}_k^{g\,T}(I-(1-\lambda)\underline{x}_k\underline{x}_k^T)(I-(1-\lambda)\underline{x}_k\underline{x}_k^T)\underline{q}_k^{g}\,)] + (1-\lambda)^2 NK \quad (3.38)$$

where K is the order of the system. It now follows from the distributive law that,

$$E[\,|\underline{q}_{k+1}^{g}|^2] = E[\underline{q}_k^{g\,T}\underline{q}_k^{g}] - 2(1-\lambda)E[\underline{q}_k^{g\,T}\underline{x}_k\underline{x}_k^T\underline{q}_k^{g}] + (1-\lambda)^2 E[\underline{q}_k^{g\,T}\underline{x}_k\underline{x}_k^T\underline{x}_k\underline{x}_k^T\underline{q}_k^{g}] + (1-\lambda)^2 NK \quad (3.39)$$

At this point we can make use of assumptions A1 and A3 to obtain,

$$E[\,|\underline{q}_{k+1}^{g}|^2] = E[\,|\underline{q}_k^{g}|^2](1-2(1-\lambda)+(1-\lambda)^2\beta) + (1-\lambda)^2 NK \quad (3.40)$$

where $\beta = K - 1 + E[\underline{x}_i^4]$ [86] .

Thus in the limit the steady state MSD associated with the measurement noise is,

$$Measurement\ MSD = KN(1-\lambda)/(2-(1-\lambda)\beta) \quad (3.41a)$$

Similarly for the lag term,

$$Lag\ MSD = D\ /\ ((1-\lambda)\,(2-(1-\lambda)\beta)) \quad (3.41b)$$

Where $D = E\,[\,|\,d_k\,|^2\,]$ and is the variance of the time varying tap increment (assuming zero mean). Therefore the steady state MSE achieved by the RLS algorithm is as below;

$$MSE = N + KN(1-\lambda)/\,(2-\beta(1-\lambda)) + D/((1-\lambda)(2-\beta\,(1-\lambda)) \quad (3.42)$$

When the input is no longer white then it is necessary to proceed as follows. Clearly equation (4) which is the RLS tap-weight update equation can be written as,

$$\hat{\underline{h}}_{k+1} = \hat{\underline{h}}_k + (1-\lambda)\Phi^{-1} e_k\underline{x}_k \quad (3.43)$$

by use of assumption (A4). As a result of the independence assumptions it should be noted that Φ is a diagonal matrix. Although this assumption is clearly untrue for the time sequence considered here, a simple unitary rotation would guarantee that Φ would be diagonal and thus allow the analysis to proceed.

If equation (3.32) is rewritten as,

$$\underline{q}_{k+1} = \underline{q}_k - (1-\lambda)(\underline{q}_k \underline{x}_k + n_k)\,\Phi^{-1}\underline{x}_k - \underline{d}_k . \tag{3.44}$$

In order to obtain the MSD we require $E[\,|\,|q_\infty|\,|^2]$, so if equation (3.44) is squared and the expectation taken at the limit then we may proceed as below.

$$\begin{aligned} E[\,|\underline{q}_{k+1}|^2] = {} & E[\,|\underline{q}_k|^2] - 2(1-\lambda)E[\,|(\underline{q}_k.\underline{x}_k + n_k)(\Phi^{-1}\underline{x}_k)\underline{q}_k|\,] \\ & + (1-\lambda)^2 E[\,|(\underline{q}_k\underline{x}_k + n_k)^2(\Phi^{-1}\underline{x}_k)^2|\,] + E[\,|\underline{d}_k|^2], \end{aligned} \tag{3.45}$$

and if this is expanded then,

$$\begin{aligned} E[\,|\underline{q}_{k+1}|^2] = {} & E[\,|\underline{q}_k|^2] - 2(1-\lambda)E[\,|(\underline{q}_k.\underline{x}_k)(\Phi^{-1}\underline{x}_k).\underline{q}_k|\,] \\ & + (1-\lambda)^2 E[\,|(\underline{q}_k.\underline{x}_k)^2(\Phi^{-1}.\underline{x}_k)^2|\,] + (1-\lambda)^2 N E[\,|(\Phi^{-1}.\underline{x}_k)^2|\,] + D. \end{aligned} \tag{3.46}$$

As a consequence of Φ being a diagonal matrix then,

$$\begin{aligned} E[\,|\underline{q}_{k+1}|^2] = {} & E[\,|\underline{q}_k|^2] - 2(1-\lambda)E[\,|(\sum_{i=1}^{i=K} q_k^i x_k^i)(\sum_{i=1}^{i=K}\frac{1}{\alpha_i}(q_k^i x_k^i)|\,] \\ & + (1-\lambda)^2 E[\,|(\sum_{i=1}^{i=k}(q_k^i x_k^i)((\frac{x_k^i}{\alpha_i})^2)\,|\,] + (\sum_{i=1}^{i=K}\frac{1}{\alpha_i})(1-\lambda)N + D, \end{aligned} \tag{3.47}$$

where x_k^i and q_k^i represent the constituent components in the vectors $\underline{x}_k$ and $\underline{q}_k$ respectively. The terms α_i represent the eigenvalues of the input signal autocorrelation matrix. If this expression is then taken to the limit then the total MSD is expressed in equation (3.48) shown below,

$$Total\ MSD = (\sum_{i=1}^{i=k}\frac{1}{\alpha_i})\,(1-\lambda)\,\frac{N}{(2-\beta(1-\lambda))} + \frac{D}{(1-\lambda)(2-\beta(1-\lambda))} \tag{3.48}$$

It is interesting to note that both expressions for the MSD, i.e. in white and coloured input signal conditions allows the effect of system time-variations and input signal colouration on the rate of convergence to be assessed, as in[86].

3.4 PERFORMANCE COMPARISONS

It has previously been assumed that RLS algorithms would always track time variations of a system faster than the LMS algorithm, and Honig [18] demonstrated that the RLS algorithm will always converge faster than the LMS in a stationary environment even when the input is white. This misconception has arisen as a result of failing to distinguish between the spectral robustness and fast initial rate-of-

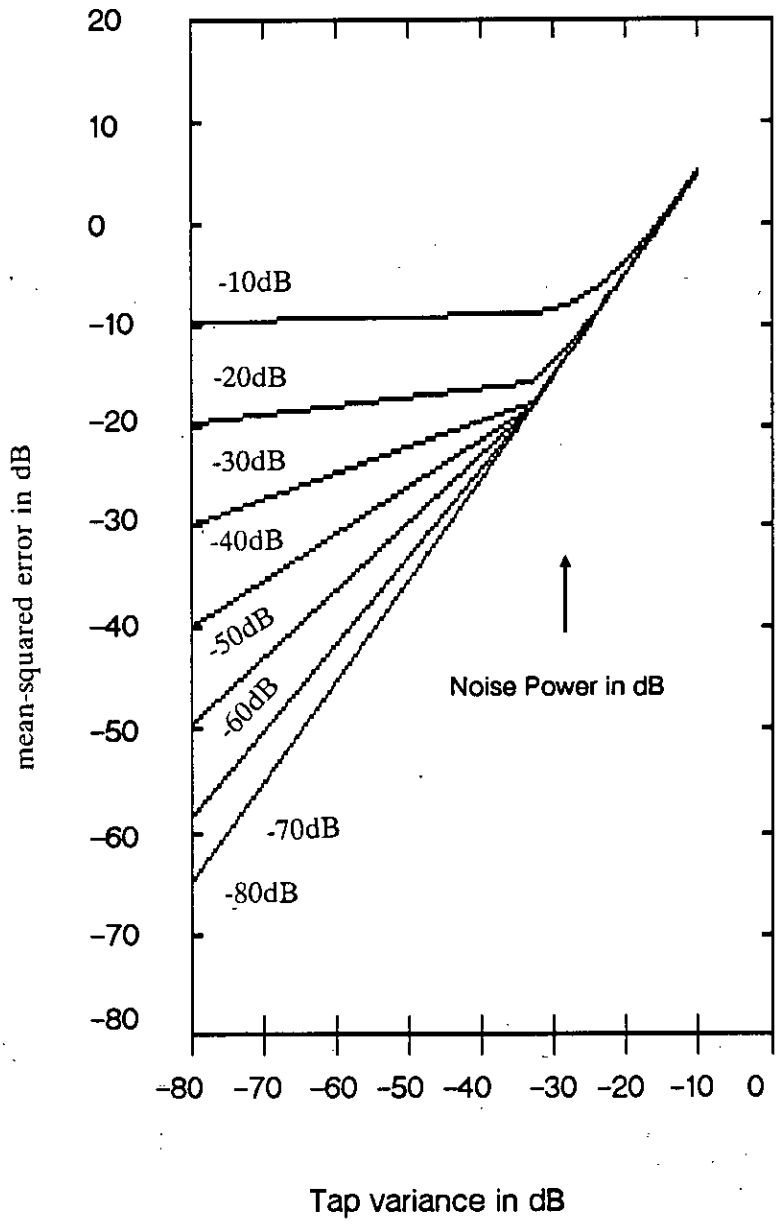


Fig. 3.5a – Theoretical achievable MSE for RLS algorithm in dB plotted against Tap variance in dB, for a fixed Noise power ranging from -80dB to -10dB, exponential windowing factor set at 0.95.

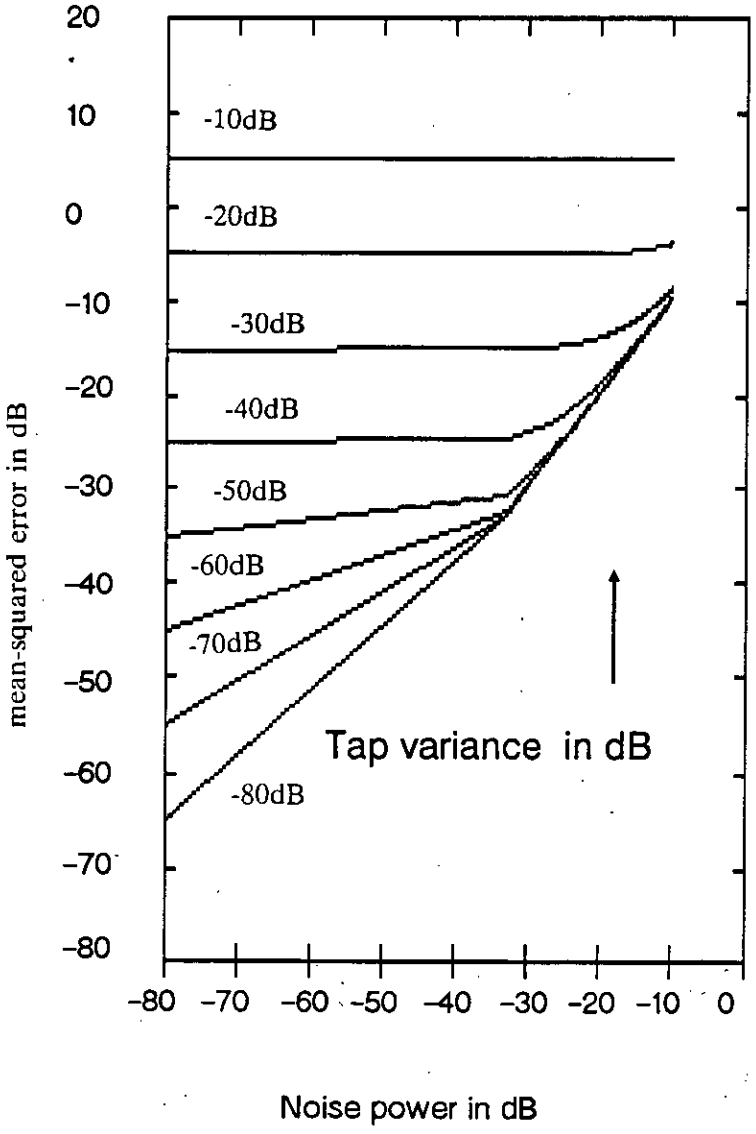


Fig. 3.5b – Theoretical achievable MSE for RLS algorithm plotted against Noise power in dB, with tap variance fixed and ranging from –80dB to –10dB, exponential windowing factor set at 0.95.

convergence, i.e. transient behaviour of the RLS algorithm from it's steady-state, i.e. tracking behaviour. The term spectral robustness may be considered to describe the lack of sensitivity of the RLS to the eigenstructure of the input sequence. The expression obtained for the steady state MSE of the RLS may now be used to evaluate and compare the theoretical performance of the algorithm in a time-varying environment.

Figure 3.5a illustrate the theoretical steady state MSE for the RLS for constant noise in steps of 10dB (-10dB to -80dB) and for tap variance ranging from -10dB to -80dB. Figure 3.5b illustrate the performance of the algorithm for constant tap variance in 10dB steps and noise ranging from -10dB to -80dB. In both situations the input is assumed to be white and $\lambda = 0.95$ for the RLS.

As can clearly be seen the algorithm achieves an asymptotic error floor which it cannot improve upon. It is also interesting to note that if the expressions for the predicted steady state MSE, obtained in [52] , for the LMS are utilised then the predicted MSE is lower than that for the RLS when the nonstationarity is high (tap variance >-40dB).

Figures 3.6a to 3.6h illustrate the simulated performance of the LMS and RLS algorithms as channel estimators for nonstationarity levels of 25dB and 45dB respectively, signal to noise ratios of 30dB and 50dB and for both white and coloured input signal conditions. The level of colouration being determined by an eigenvalue ratio (EVR) of 16.5. The plots show tap vector norms (or mean squared deviation MSD) against number of iterations. The tap vector norm plots were chosen rather than MSE plots because they illustrate the tracking behaviour of the algorithms more accurately. Table 3.1 indicates the appropriate values for each figure with MSD_p representing the theoretically predicted MSD and MSD_m representing the measured values from the simulations.

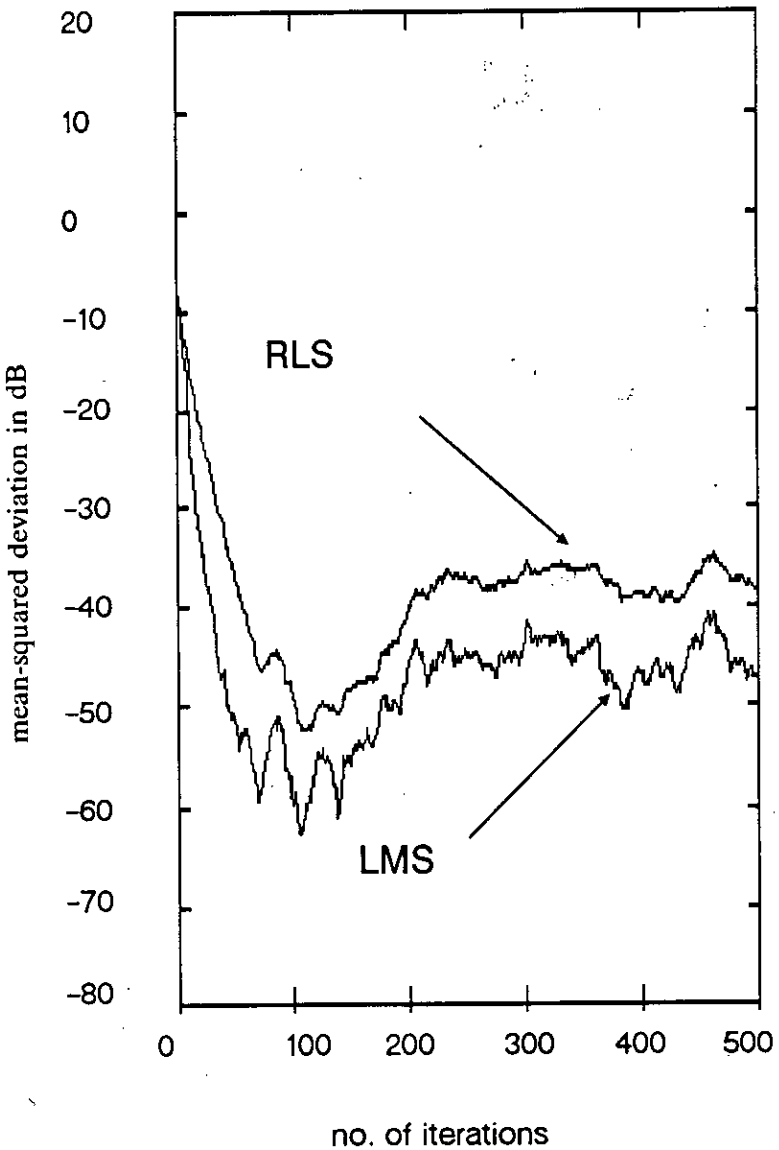


Fig. 3.6a – Performance comparison of LMS and RLS algorithms as HF Channel estimators with white input.

Fade rate=1Hz (D=–45dB) and Noise power ,N=–50dB,
exponential windowing factor for RLS=0.95,
LMS step size =0.16666.

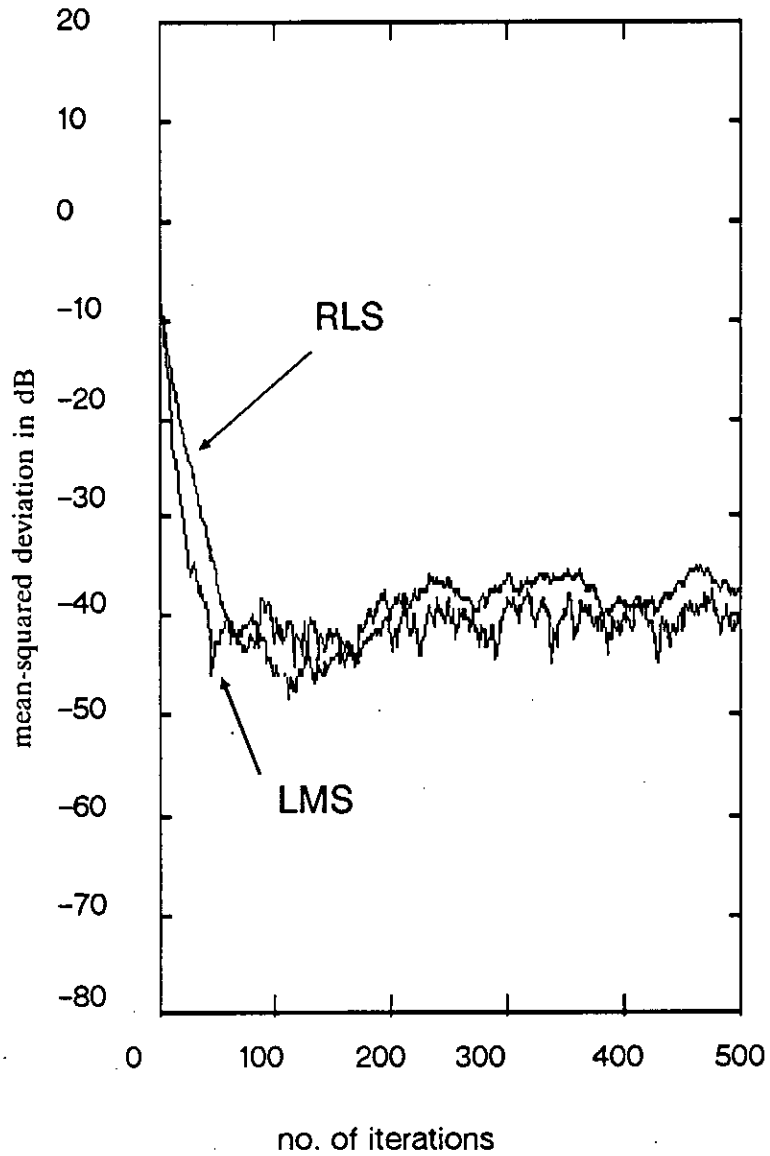


Fig. 3.6b – Performance comparison of LMS and RLS algorithms as HF Channel estimators with white input.

Fade rate=1Hz (D=-45dB) and Noise power, N=-30dB,
exponential windowing factor for RLS=0.95,
LMS step size=0.1666.

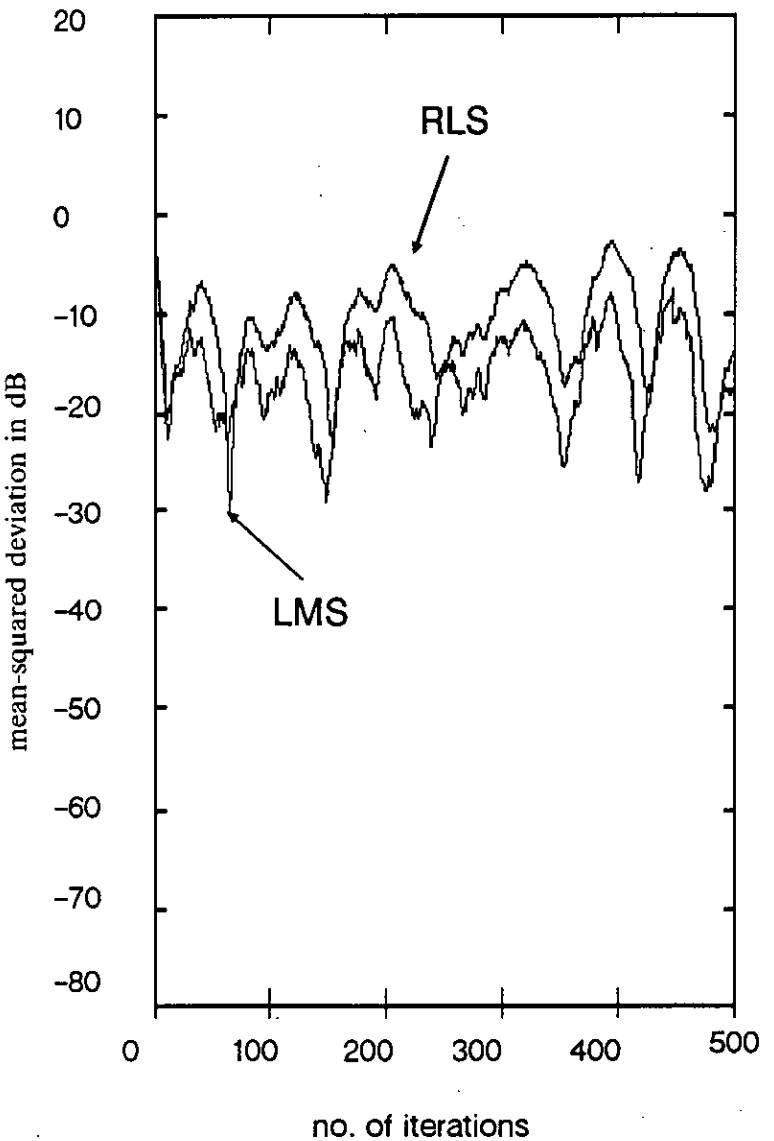


Fig. 3.6c – Performance comparison of LMS and RLS algorithms as HF Channel estimators with white input.

Fade rate=10Hz (D=-25dB) and Noise power, N=-50dB, exponential windowing factor for RLS=0.95, LMS step size=0.16666.

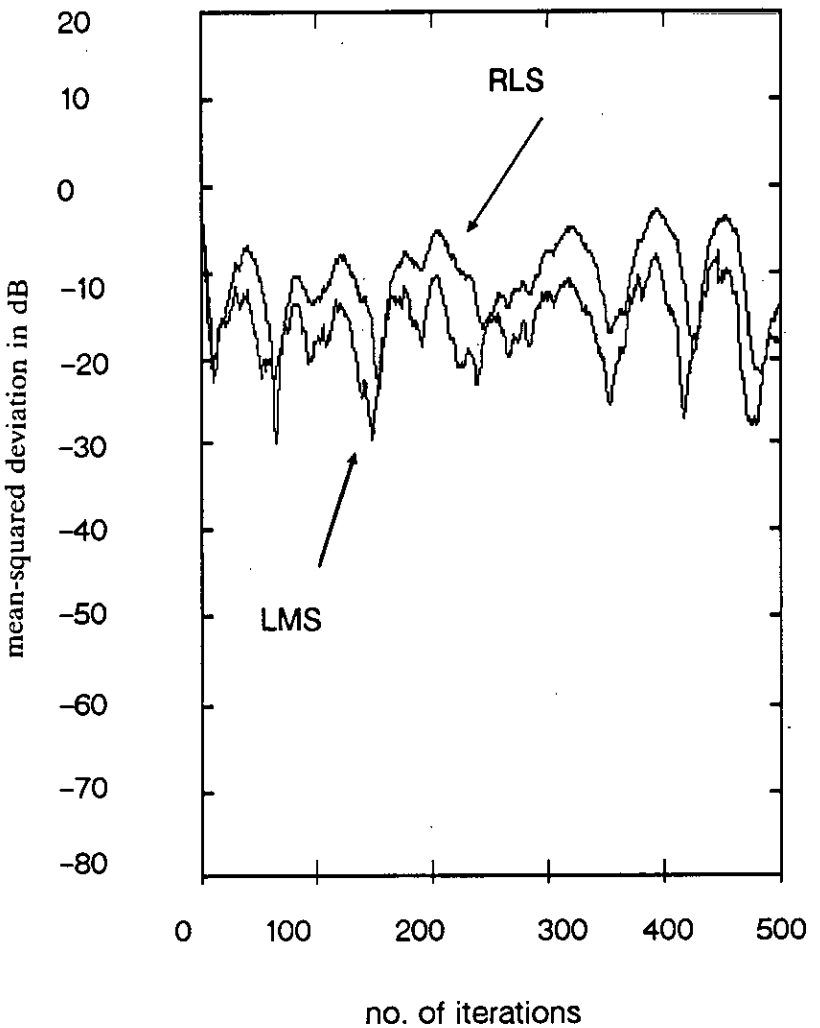


Fig. 3.6d - Performance comparison of LMS and RLS algorithms as HF Channel estimators with white input.

Fade rate=10Hz (D=-25dB) and Noise power, N=-30dB,
exponential windowing factor for RLS=0.95,
LMS step size=0.16666.

λ	D	N	EVR	MSD_P	MSD_M
0.95	-45db	-50dB	1	-34.0dB	-35dB
0.95	-45dB	-30dB	1	-33.2dB	-35dB
0.95	-25dB	-50dB	1	-14.6dB	-13.5dB
0.95	-25dB	-30dB	1	-14.6dB	-13.5dB
0.95	-45db	-50dB	16.5	-34.4dB	-35dB
0.95	-45dB	-30dB	16.5	-32.0dB	-35dB
0.95	-25dB	-30dB	16.5	-14.4dB	-13.5dB
0.95	-25db	-50dB	16.5	-14.5dB	-13.5dB

Table 3.1 - Comparison of simulated and theoretical results

The value of μ selected for the LMS used the stability criterion suggested by Feuer and Weinstein in [68] , $0 \leq \mu < 1/3\ tr[R]$, for this work μ was chosen to be at the proposed optimal value, i.e. $\mu = 1/(6\ tr[R])$ which for the white input conditions is $\mu = 1/6K$ where K is the order of the system. This value of μ was chosen to guarantee stability of the LMS but still ensure reasonable tracking performance by the algorithm for the simulations presented. Therefore, for the three tap channel used in the simulations $\mu = 0.05556$.

The simulations represented in Figures 3.6a-h consider 3 situations.

i) tap variance >> additive noise power,

ii) additive noise power >> tap variance,

iii) additive noise power = tap variance.

In all cases when the input is white the LMS performs as well if not better than the RLS. The predicted values of the RLS are all within 1-2dB of the measured values. In the situations when the input signal is coloured the lack of spectral robustness of the LMS is demonstrated while the RLS is , as expected, relatively unaffected by the signal colouration. The best performance of the RLS (and the theoretical prediction) results

when the window length is short i.e. λ becoming smaller. This is as predicted by the theoretical expression of the RLS. If the expression is differentiated and the measurement and lag terms considered separately it is clear that the contribution of the measurement term becomes smaller as λ approaches 1 and that of the lag term becomes larger as λ approaches 1. This illustrates the trade off required in window selection, i.e. the longer the window length (λ closer to 1) then the better the estimate of the autocorrelation matrix and the shorter the window length (λ getting smaller) then the better the tracking speed and thus the smaller the lag error contribution.

The expression $2\mu = (1 - \lambda)$[15] can be used with the stability criterion shown previously to illustrate the effect of the order of the system. If λ is constrained to lie within the region 0.92-0.999 then clearly if $\lambda=0.92$ (i.e. $\mu = 0.04$) the LMS could not be guaranteed stable for systems of greater than order 8. However, if $\lambda=0.98$ was chosen then the LMS would not be guaranteed stable for systems of order greater than 33. Eleftheriou and Falconer [50] applied a similar technique in their work but utilised a less conservative stability criterion for the LMS and considered a channel model represented by a first order Markov process.

It is clear from the results presented in this chapter that there is a reasonable agreement between the simulation results and theoretical predictions. Work recently published by Clark & Harun [26] and previously by Tront [56] reinforces the results demonstrated here.

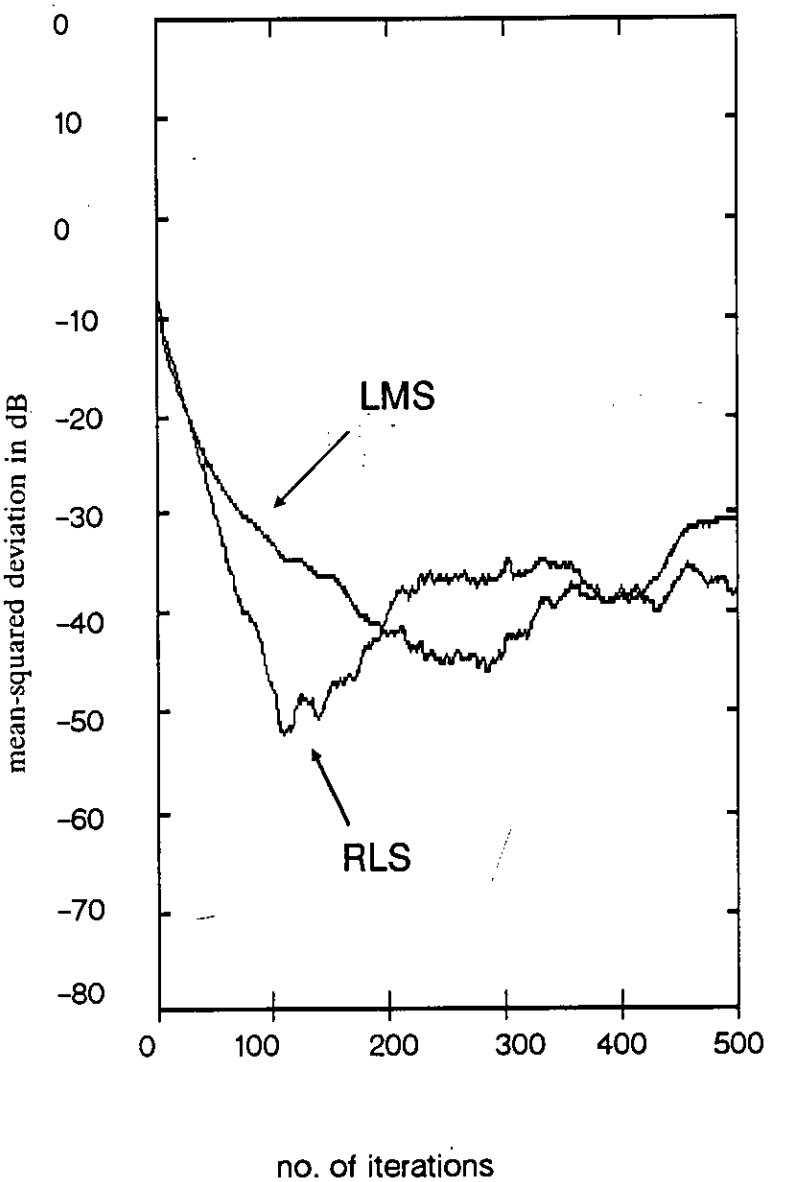


Fig. 3.6e - Performance comparison of LMS and RLS algorithms as HF Channel estimators for coloured input.

Fade rate=1Hz (D=-45dB) and Noise power, N=-50dB,
exponential windowing factor for RLS=0.95,
LMS step size=0.1666.

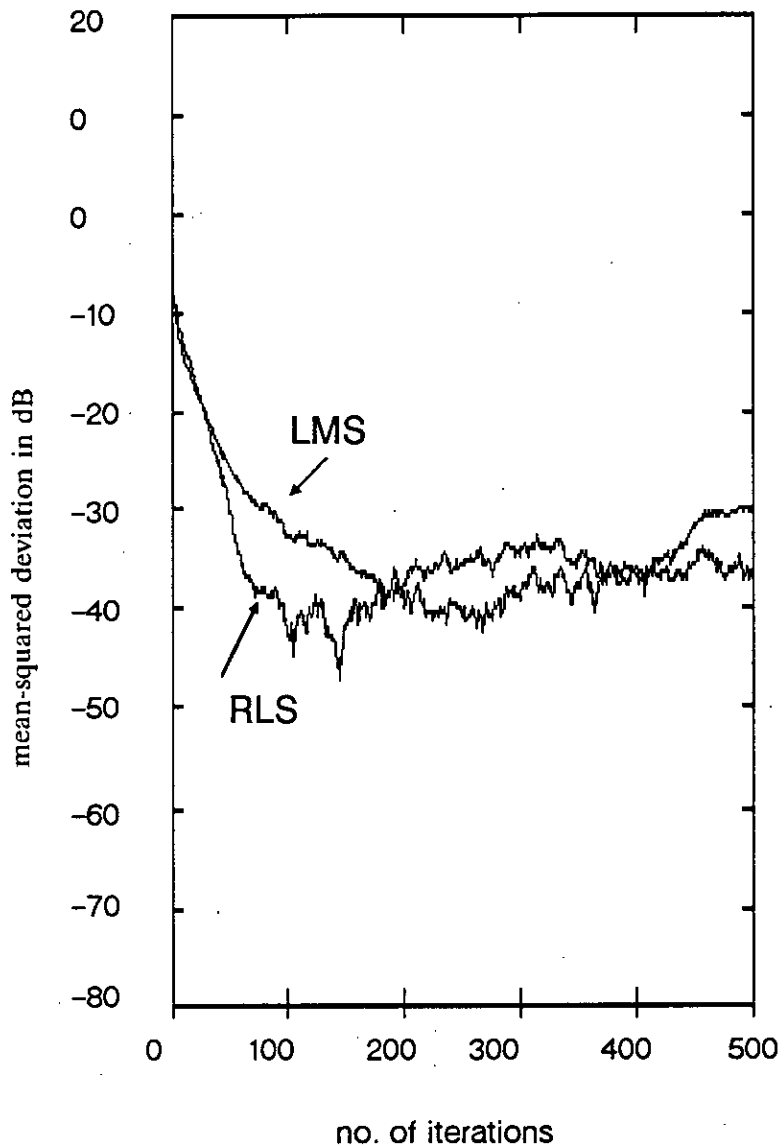


Fig. 3.6f – Performance Comparison of LMS and RLS algorithms as HF Channel estimators with coloured input.

Fade rate=1Hz (D=–45dB) and Noise power, N=–50dB,
exponential windowing factor for RLS=0.95,
LMS step size=0.16666.

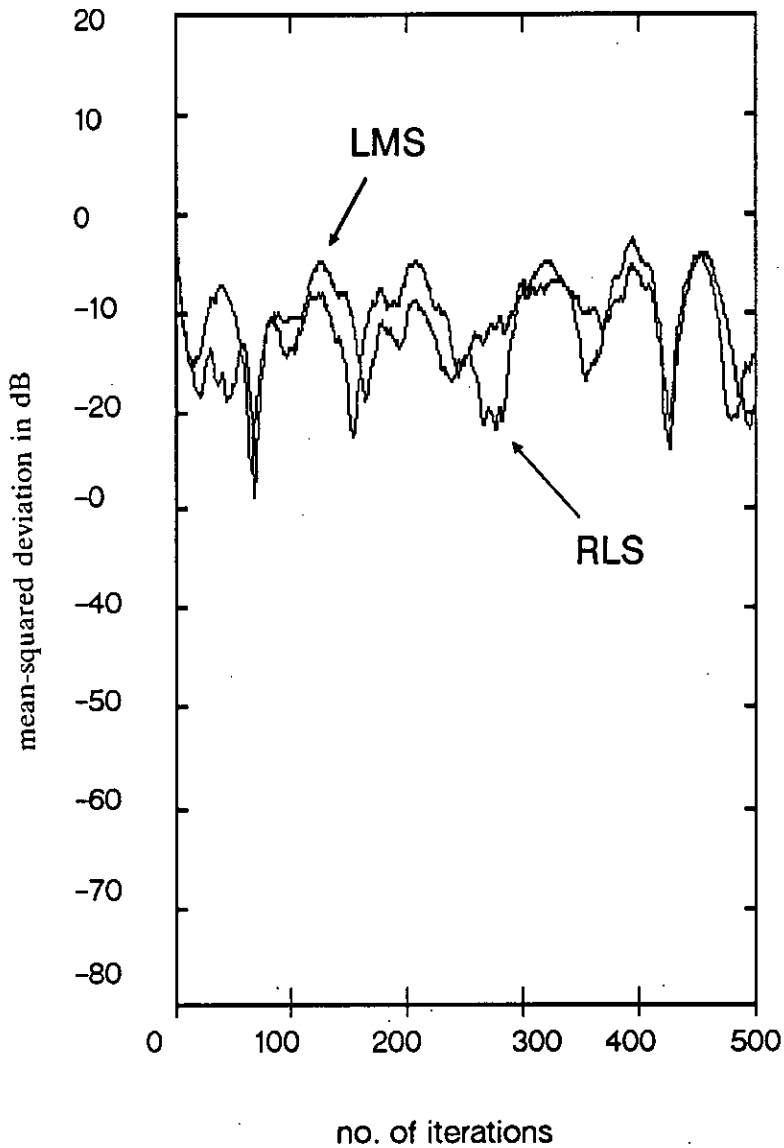


Fig 3.6g - Performance comparison of LMS and RLS algorithms as HF Channel estimators with coloured input.

Fade rate=10Hz (D=-25dB) and Noise power, N=-50dB,
exponential windowing factor for RLS=0.95,
LMS step size=0.1666.

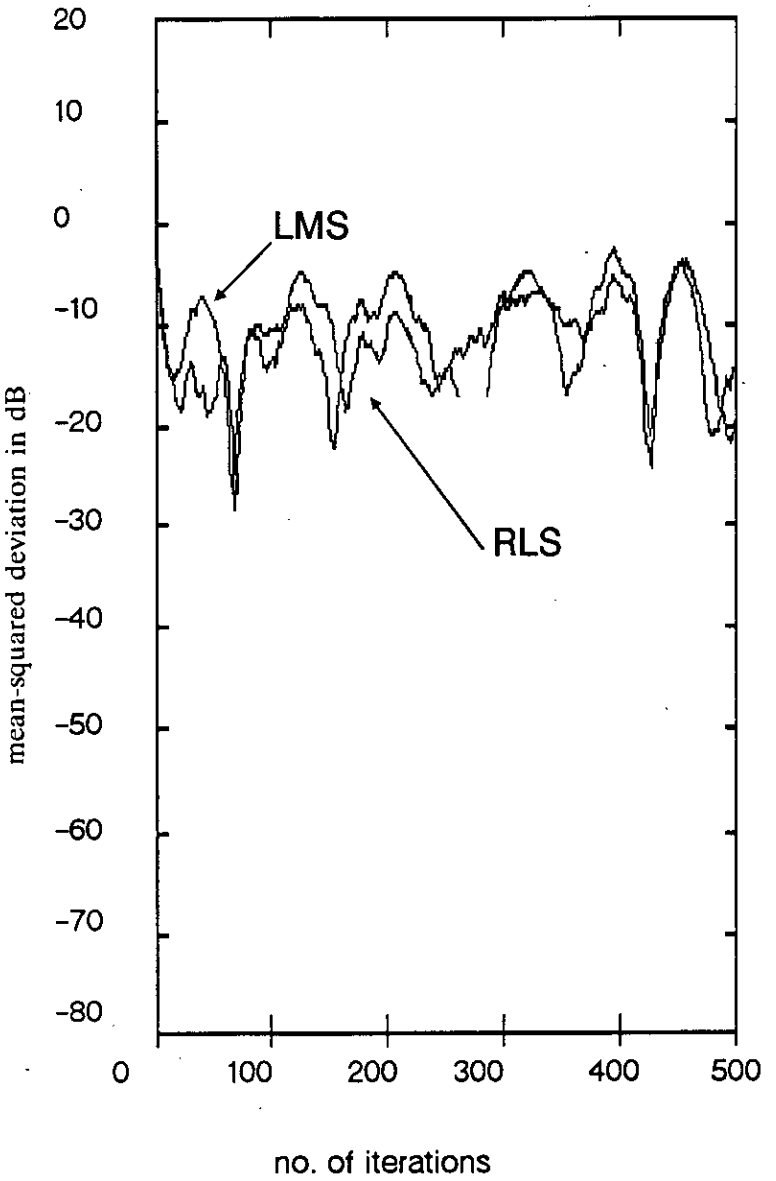


Fig. 3.6h – Performance Comparison of LMS and RLS algorithms as HF Channel estimators for coloured input.

Fade rate=10Hz (D=-25dB) and Noise power, N=-30dB,
exponential windowing factor for RLS =0.95,
LMS step size=0.1666.

3.5 CONCLUSIONS

It is clear from the results presented that the tracking performance expected of RLS algorithms in time varying environments is not achieved. Although this work has only looked at the algorithm's performance in a direct modelling situation it is clear that the tracking performance of the algorithm is an important characteristic of the algorithm which must be clearly separated from other effects such as spectrally robust convergence behaviour.

It would appear that RLS algorithms are not necessarily suitable for use in highly time-variant environments, such as the HF communication channel. The slow rate of convergence and subsequent degradation in performance of the LMS algorithm with coloured input also makes it unsuited for such applications. Consequently it would appear that some form of nonlinear techniques [87] or new algorithms which estimate the parameters generating the nonstationarity may have to be considered to produce adaptation algorithms which can function in such hostile environments, as will now be discussed in the next chapter.

Chapter 4

THE USE OF A-PRIORI KNOWLEDGE IN HF CHANNEL ESTIMATION

The results which were presented in chapter 3 indicate that there is considerable room for improvement in the performance of conventional adaptive algorithms as HF channel estimators. This chapter looks at the development of three algorithms which utilise varying degrees of a-priori knowledge to improve their performance. The term a-priori knowledge refers here to knowledge of the parameters which define the model in a state space sense; this information allowing the use of a Kalman filter approach in the exact sense.

4.1 INTRODUCTION

The publication of [61, 62] by Kalman and Bucy proposed an extremely powerful recursive estimation technique commonly described as Kalman filters. The use of Kalman filter equations pre-supposes that the system under consideration can be described by a set of linear difference equations, for discrete time systems of the type considered here the system is normally described as,

$$\underline{s}_k = F\ (k/k-1)\ \underline{x}_{k-1} + G_k \underline{w}_k, \tag{4.1}$$

and

$$\underline{z}_k = H_k^T\ \underline{s}_k + \underline{v}_k, \tag{4.2}$$

where $F\ (k/k-1)$ represents a $K \times K$ state-transition matrix. The M-dimensional vector $\underline{z}_k$ is normally termed the measurement vector and H_k is the $M \times K$ observation or measurement matrix. The terms w_k and v_k are respectively the K and M-dimensional vectors of the zero-mean white noise processes, which are assumed to be statistically independent with covariances denoted by Q and W respectively.

Given this description optimal estimates of the K-dimensional state vector $\hat{\underline{s}}_k$ are obtained from the noisy observations, $\underline{z}_k$, in a recursive manner by the following equations. The estimation equation,

$$\hat{\underline{s}}_k^k = \hat{\underline{s}}_k^{k-1} + K_k \, [\, \underline{z}_k - H_k \cdot \hat{\underline{s}}_k^{k-1} \,], \tag{4.3}$$

the prediction equation,

$$\hat{\underline{s}}_k^{k-1} = F(k/k-1)\hat{\underline{s}}_{k-1}^{k-1}. \tag{4.4}$$

The Kalman gain, K_k, being described by,

$$K_k = V_k^{k-1} H_k^T \, [\, H_k V_k^{k-1} H_k^T + W_k \,]^{-1}. \tag{4.5}$$

The error covariance, V_k^{k-1}, being defined by $E \, [\, (\, \underline{x}_k - \hat{\underline{x}}_k^{k-1} \,) \, (\, \underline{x}_k - \hat{\underline{x}}_k^{k-1} \,)^T \,]$ and obtained from,

$$V_k^{k-1} = F(k/k-1)V_{k-1}^{k-1}F(k/k-1) + Q_k, \tag{4.6}$$

and the other covariance term, V_{k-1}^{k-1}, is defined by $E \, [\, (\, \underline{x}_k - \hat{\underline{x}}_k^k \,) \, (\, \underline{x}_k - \underline{x}_k^k \,) \,]$, and obtained via,

$$V_{k-1}^{k-1} = V_{k-1}^{k-2} - K_{k-1}H_{k-1}V_{k-1}^{k-2}. \tag{4.7}$$

In order to proceed with the development with the algorithms in this chapter it is necessary to return to the channel model and tap generation filter described in chapter 2 and put it into a suitable format. algorithms in this chapter.

4.2 MINIMUM VARIANCE KALMAN ESTIMATOR

As has been stated the Kalman-Bucy filter is the optimal (i.e. minimum variance) filter for the discrete linear system described by equations (4.1)and (4.2). Unfortunately in most applications not all the parameters which define the system are known a-priori and as a result modelling errors [88] occur in practical applications of the filter. However, by utilising the state space definition of the channel model and incorporating the information provided into the filter it is clear that these modelling errors could be eliminated, and by definition this filter would be the true minimum variance estimator for this system. This fact can then be used to allow the filter to be used to set a bound on the minimum achievable MSE for any HF channel estimator, since it has full a-priori knowledge of the channel.

The widely accepted method for modelling the HF channel views the channel as a FIR filter with time-varying tap weights each of which is statistically independent, as described in chapter 2. The channel may then be described by the equations presented

in the previous section and the state transition matrix used directly in the Kalman filter equations. Thus, the MVK estimator may be defined as the optimal estimator for the representation of the HF channel used in this study. It is optimal because the algorithm has full a-priori knowledge of the system inbuilt, as illustrated in Figure 4.1.

The tap generation filter can be represented by the equations below,

$$X_0(k) = X_1(k-1), \tag{4.8}$$

and,

$$X_1(k) = V(k) - C_0X_0(k-1) - C_1X_1(k-1). \tag{4.9}$$

The values of C_0 and C_1 being dependent on the bandwidth of the filter. These equations may be used in conjunction with the observation equation of the channel to describe the system in terms of the equations below,

$$\underline{s}_{k+1} = F\underline{s}_k + G_k w_k, \tag{4.10}$$

and,

$$y_k = \underline{H}_k^T \underline{s}_k + v_k \tag{4.11}$$

This is clearly of the form of equations (4.1) and (4.2) with the state vector $\underline{s}_k$ defined as, $\underline{s}_k = [\, X_0\, X_1 \,]_k^T$ and the time-invariant state transition matrix being, $F = \begin{bmatrix} 0 & 1 \\ -C_1 & -C_0 \end{bmatrix}$. H_k, is the observation matrix, constructed from the input signal and $C = [1\ -C_1\ 2\ -C_0\,]$, thus $\underline{H}_k = [\,(1 - C_1)\, x_0\ \ (1 - C_0)\, x_1\,]$ where the x_i represent the inputs to the filter.

In order to expand this to the three tap model considered here the matrices are simply augmented appropriately and consequently, the overall state is described by, $[\, X_0\, X_1\, X_2\, X_3\, X_4\, X_5\,]_k^T$, with the state transition matrix being,

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -C_1 & -C_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -C_1 & -C_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -C_1 & -C_0 \end{bmatrix}$$

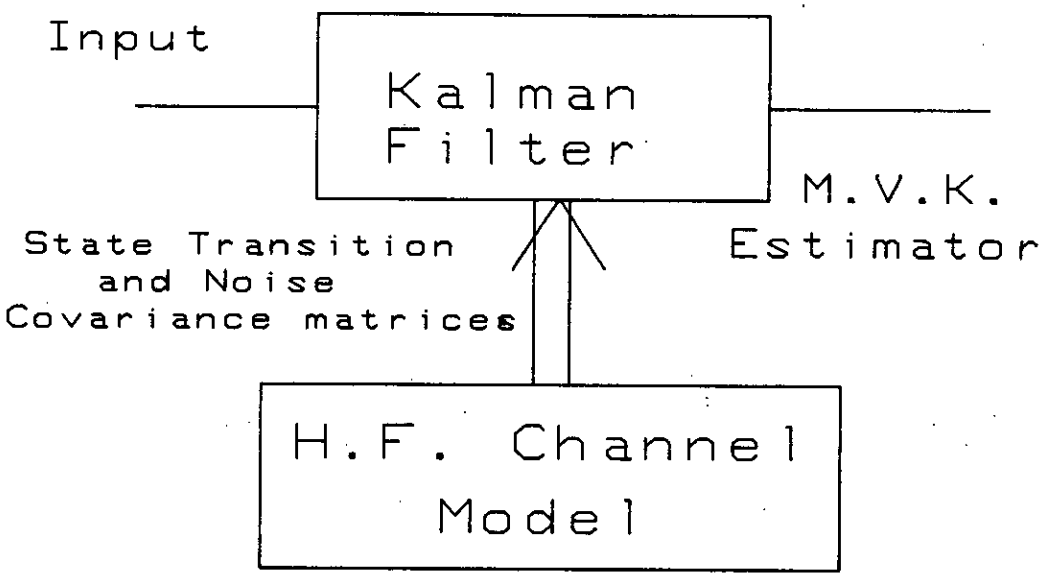


Figure 4.1 - Relationship of M.V.K. estimator to H.F. channel model

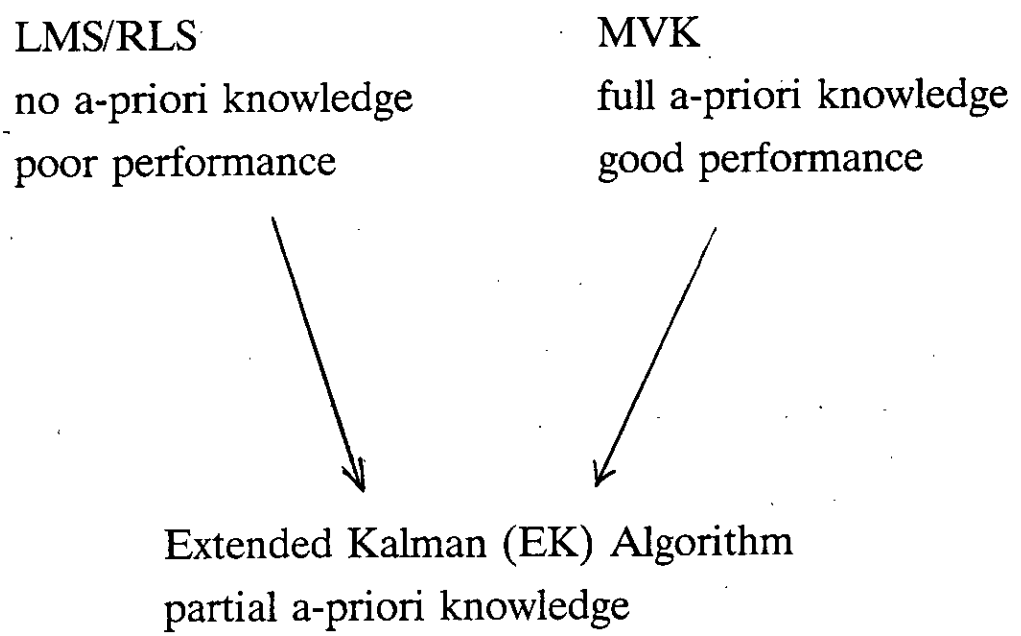


Figure 4.2 - Motivation for using extended Kalman algorithm

This description of the system provides the means of defining exactly what a-priori information is required in the development of the three adaptive algorithms. Obviously the MVK estimator is not implementable in a practical situation since such information is not available, however it is still useful for the insight which it provides and illustrates the motivation for the next stage of algorithm development.

4.3 THE EXTENDED KALMAN ALGORITHM AS A HF CHANNEL ESTIMATOR

Since it is clear that full a-priori knowledge of the channel will not be available in practical situations it would seem logical to develop an algorithm which can approach the performance of the MVK with the full a-priori knowledge constraint removed. The hierarchial approach adopted here is illustrated in Figure 4.2. Essentially, an Extended Kalman filter (EKF) algorithm [64] is utilised to develop the estimator in which only partial a-priori knowledge is used, that being the structure of the model as opposed to particular parameter values. The EKF algorithm is an application of linear Kalman filter theory to nonlinear systems, the nonlinear system is linearised about the current state estimate and the standard Kalman filter algorithm applied to the resultant time-varying linear system.

Before proceeding with the derivation of the EKF algorithm for HF channel estimation it is possible to illustrate, by a simple example, its relationship to the MVK estimator of section 4.2. Figure 4.3 illustrates a simple communication system, the channel being of order two, each of the taps is generated by passing white noise through a simple first order autoregressive filter. If the MVK estimator was applied to this channel then the state vector would be $\underline{s}_k = [\, a_1\, a_2 \,]^T$ but in the case of the extended Kalman this would become, $\underline{s}_k = [\, a_1\, a_2\, \hat{g} \,]^T$ where $\hat{g}$ represents an estimate of the gain of the filter in the tap generation process. Clearly then the algorithm is no longer simply estimating the taps of the channel but also the parameters responsible for the variations in the tap weights.

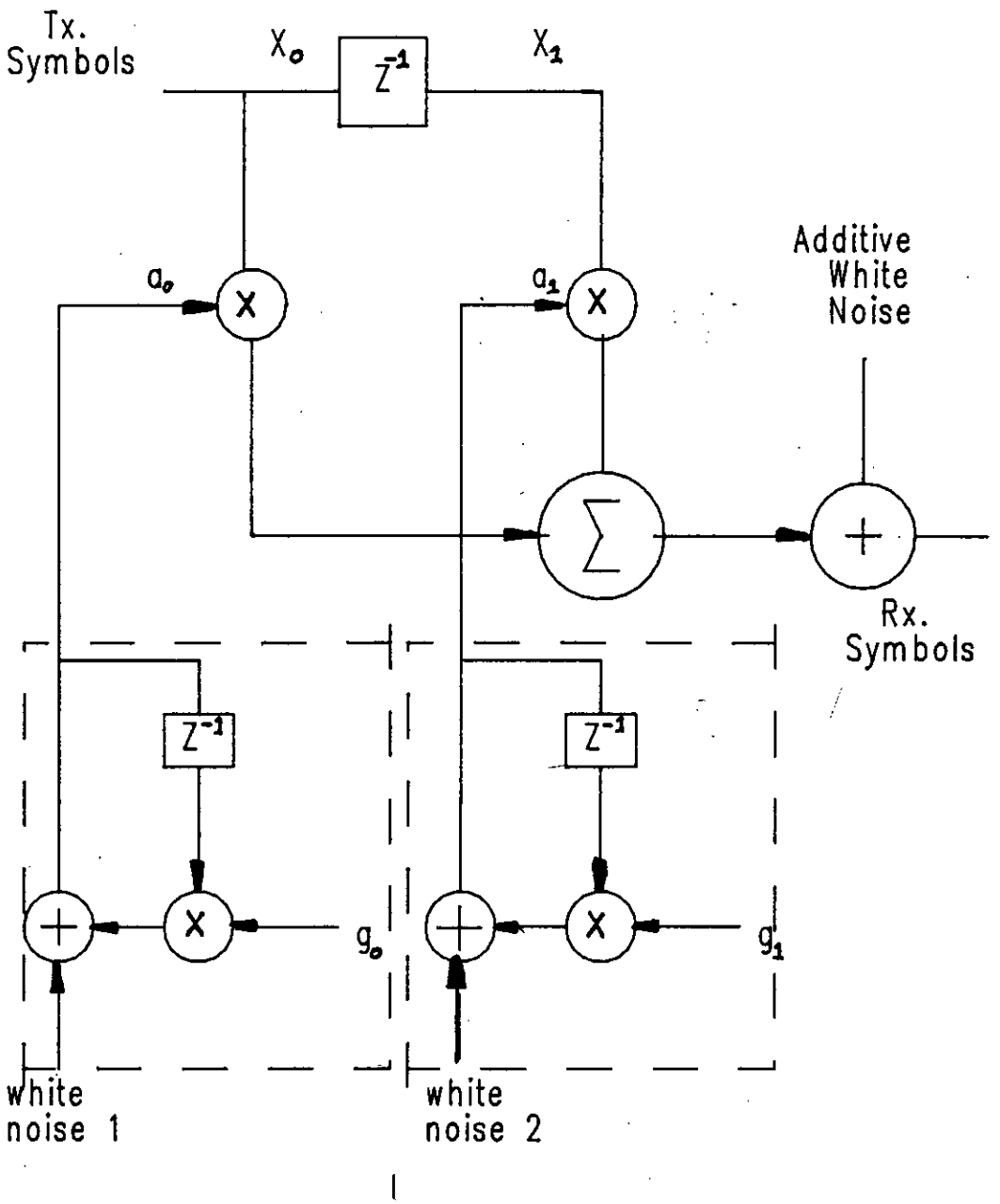


Figure 4.3 - Simple conceptual model of application of extended Kalman algorithm to channel estimation

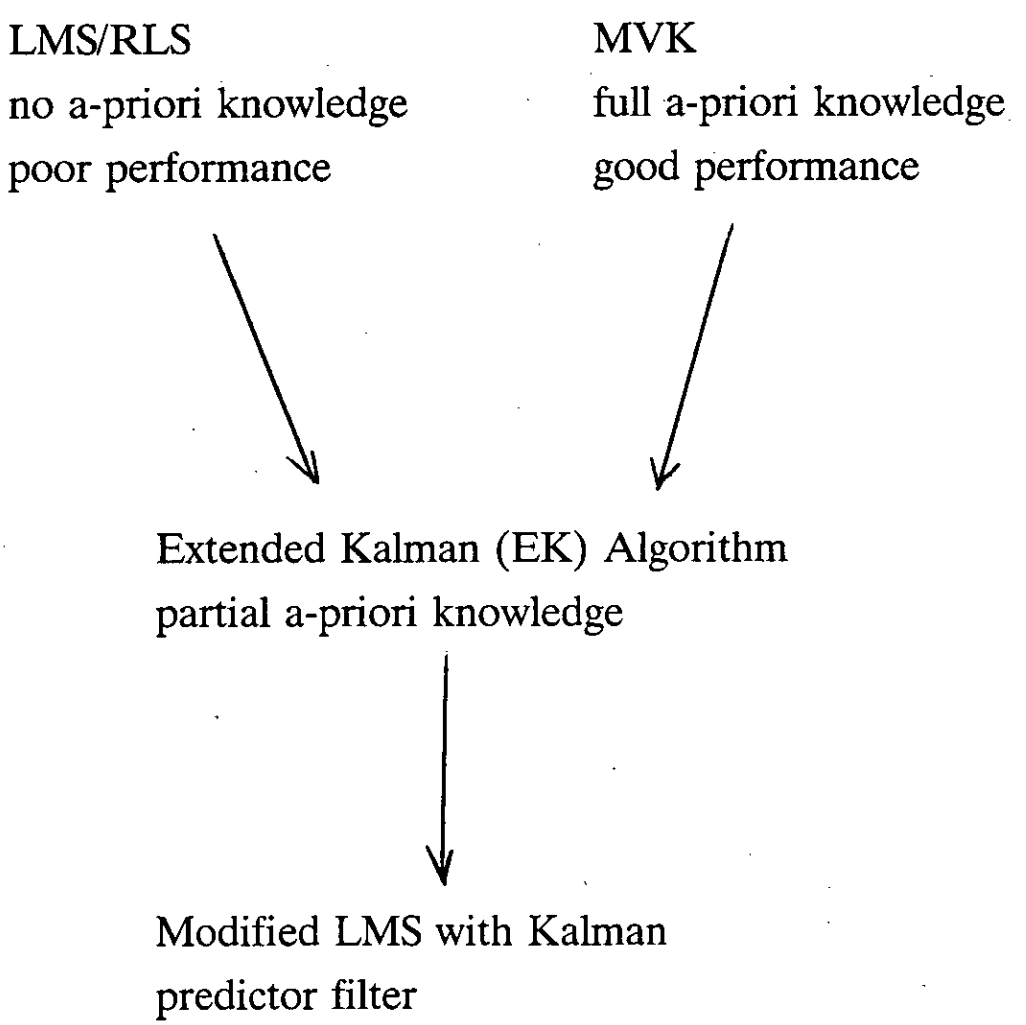


Figure 4.4 - Motivation for Modified LMS algorithm derivation

The state-space representation of the channel model, as described by (4.1) and (4.2), is converted to a nonlinear system by augmenting the state vector of the linear system with the stationary parameters that make up the tap generation model. The key to the performance of the EKF algorithm lies in the accuracy of the initial linearising approximation made to the nonlinear system. Therefore, in order to ensure that the EKF algorithm has a good initial estimate, (thus ensuring good convergence behaviour), an LMS algorithm was used to carry out the initial training of the algorithm.

4.3.1 State Space Formulation

As has been stated previously the Kalman filter is not optimal for nonlinear systems, the problem of optimal filtering for nonlinear scenarios being considerably more complex than in linear system theory. Normally an exact solution via recursive methods is not possible, the conventional approach has been to adapt standard linear algorithms and determine their performance. The EKF algorithm is simply an extension of the conventional linear Kalman filter algorithm to a first order nonlinear system which has undergone a first-order linearisation.

By using the state-space description of the channel model as described in equations previously, the state vector $\underline{s}_k = [\, X_0\, X_1 \,]_k^T$ is augmented with the filter coefficients such that the augmented state $\underline{s}_k'$ equals $[\, X_0\, X_1\, C_0\, C_1 \,]_k^T$. The system is now a nonlinear system and in order to apply the EKF it is necessary to obtain the state transition and measurement (or observation) matrices.

This is achieved by the following substitutions in the state equations of the system, $X_n(k) = X_n'(k) + \delta X_n(k)$ and $C_n = C_n' + \delta C_n$ where X_n' and C_n' may be considered as reference states. When these substitutions are made in equations (4.1) and (4.2) then,

$$X_0'(k) + \delta X_0(k) = X_1'(k) + \delta X_1(k) \tag{4.12}$$

and,

$$X_1'(k) + \delta X_1(k) = V(k) - [\, C_1'(k-1) + \delta C_1(k-1) \,]\,[\, X_0'(k-1) + \delta X_0(k) \,]$$
$$- [\, C_0'(k-1) + \delta C_0(k-1) \,]\,[\, X_1'(k-1) + \delta X_1(k-1) \,] \tag{4.13}$$

result.

Once all the terms have been multiplied out equations (4.12) and (4.13) become

$$\delta X_0(k) = \delta X_1(k-1), \tag{4.14}$$

and

$$\delta X_1(k) = V(k) - X_1'(k-1)\delta C_0(k-1) - C_0'(k-1)\delta X_1(k-1)$$
$$- \delta X_0(k-1)C_1'(k-1) - X_0'(k-1)\delta C_1(k-1) \tag{4.15}$$

If this process is carried out for all six states then the resulting time variant state transition matrix, now termed A_k may be written as,

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -C_1 & -C_0' & 0 & 0 & 0 & 0 & -X_1' & -X_0' \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -C_1' & -C_0' & 0 & 0 & -X_3' & -X_2' \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -C_1' & -C_0' & -X_5' & -X_4' \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

similar arguments can be applied to the measurement matrix resulting in,

$$H_k = [\, x_0\, 2x_0\, x_1\, 2x_1\, x_2\, 2x_2 - (\, x_0 X_1' + x_1 X_3' + x_2 X_5' \,)$$
$$-(\, x_0 X_0' + x_1 X_2' + x_2 X_4' \,)\,]$$

where x_i represents constituents of the input signal vector. Table 4.1 details the algorithm in full making use of the natural block structure.

It should be noted at this point that this realisation of the EKF algorithm assumes only the form of the channel model, it does not force a Butterworth form onto the tap generation filter, merely a second order section structure and as a result the stability of the filter is not guaranteed. This problem can be overcome by ensuring that the initial estimate was reasonably accurate via an LMS algorithm and also if neccessary by monitoring the poles of the filter. If the poles of the estimated filter are outside the unit circle, then by reflecting them inwards along their radii the filter's stability can be maintained. Ovbiously if the EKF algorithm converges that the filter is being

Description	Equation
State estimate	$\hat{x}(k+1) = A_k\,\hat{x}(k) + K(k)[y(k) - C_k\hat{x}(k)]$ $\hat{x}(0) = \hat{x}_0$
State estimate	$\hat{\theta}(k+1) = \hat{\theta}(k) + L(k)[y(k) - C_k\hat{x}(k)]$ $\hat{\theta}(k) = \hat{\theta}_0$
Kalman gain estimate	$K(k) = [\,A_k\,\Sigma_{11}(k)\,C_k^T + M_k\,\Sigma_{12}^T(k)C_k^T$ $+ A_k\Sigma_{12}(k)D_k^T$ $+ M_k\Sigma_{12}(k)D_k^T + S\,]\,P_k^{-1}$
Kalman gain estimate	$L(k) = [\,\Sigma_{12}(k)^T\,C_k^T + \Sigma_{22}(k)\,D_k^T\,]\,P_k^{-1}$
Error covariance estimate	$P_k = C_k\Sigma_{11}(k)C_k^T + C_k\Sigma_{12}(k)D_k^T$ $+ D_k\Sigma_{12}(k)C_k^T + D_k\Sigma_{22}(k)D_k^T + R$
Covariance estimate	$\Sigma_{11}(k+1) = A_k\Sigma_{11}A_k^T + A_k\Sigma_{12}M_k^T$ $+ M_k\Sigma_{12}^T(k)\,A_k^T + M_k\Sigma_{12}(k)M_k^T$ $- K(k)P_kK^T(k) + Q\,;\ \Sigma_{11}(0) = \Sigma_{11}^0$
Covariance estimate	$\Sigma_{12}(k+1) = A_k\Sigma_{12}(k) + M_k\Sigma_{22}(k)$ $- K(k)P_kK^T(k);\ \ \Sigma_{12}(0) = \Sigma_{12}^0$
Covariance estimate	$\Sigma_{22}(k+1) = \Sigma_{22}(k) - L(k)P_kL^T(k);\ \Sigma_{22}(0) = \Sigma_{22}^0$

This form assumes the state space description of equations (4.1) and (4.2), and utilises the natural block structure shown below where the augmented state is

$$s_{aug}(k) = \left[\frac{s(k)}{\theta(k)}\right]$$

and the Kalman gain and Covariance matrix being similarly sectioned as below,

$$K_{aug}(k) = \left[\frac{K(k)}{L(k)}\right]\ ;\quad \Sigma(k) = \left[\frac{\Sigma_{11}(k)\ \Sigma_{12}(k)}{\Sigma_{12}(k)\ \Sigma_{22}(k)}\right]$$

Table 4.1 - EXTENDED KALMAN FILTER ALGORITHM

approximated accurately. to a MSE value approaching the noise floor it would be reasonable to assume

4.3.2 Innovations Based Representation of the EKF Algorithm

In past applications of the EKF method [89-91] has exhibited an alarming tendency to give biased estimates or to diverge if the initial linearisation is inaccurate. Detailed analysis of the convergence behaviour of the EKF algorithm is a difficult problem and it was not until the publication by Ljung of [92] that the reasons for the convergence difficulties were shown to be related to a combination of factors, such as incorrect specification of the system noise covariances and the lack of coupling between the Kalman gain and the parameter estimates. Ljung also demonstrated that convergence, at least to a local minima, was guaranteed if the algorithm was modified to incoporate some coupling between the Kalman gain and the parameter being estimated. Unfortunately this increased the complexity of the algorithm considerably, however as Ljung demonstrated if an innovations representation of the original system is used for the initial linearisation procedure rather than the conventional state space then a less complex algorithm results. Consequently in this section an innovations based representation of the algorithm is derived. For detailed discussions on the applications of the innovations approach to linear least-squares estimation Kaliath et al in [93-99] have explored the matter in great depth and Anderson and Moore in [63] provide useful discussions.

What is meant by an innovations representation? As has been stated previously the Kalman filter is optimal in the mean square error sense for systems described by a set of linear difference equations as described by equations (4.1) and (4.2). It is also true to say that one Kalman filter may be optimal for many different signal models, as discussed in[64], and [63] although it's performance may alter between the models. That is to say that although the filter gain may remain constant for various models the error covariance will be different. These signal models, which have the same Kalman filter have in common however the same output covariance. That is the Kalman filter

is determined by the covariance associated with the particular observation (or measurement) process as opposed to the particular details of the signal models. In this mapping of many signal models to one Kalman filter there is one model which is important, this being the *innovations* process, so called as a result of the white noise driving process being identical to the innovations process of the associated filter. There are many properties associated with the innovations model the most important being that it is causally invertible, i.e. the input noise process may be obtained from the output process in a causal fashion.

In order to obtain the innovations form of the EKF for the problem posed here it is necessary to obtain an innovations representation of the channel model. If the usual model of the HF channel is assumed, as defined in the previous section by equations (4.10) and (4.11), then the innovations form of the channel model may be written down as[63],

$$\bar{\underline{s}}_{k+1} = F\bar{\underline{s}}_k + K_k \epsilon_k, \tag{4.16}$$

and,

$$\bar{y}_k = \underline{H}_k^T \bar{\underline{s}}_k + \epsilon_k. \tag{4.17}$$

Where the innovations sequence, ϵ_k, is defined as;

$$\epsilon_k = y_k - H_k \hat{\underline{s}}_k \tag{4.18}$$

where y_k is the output and H_k the measurement matrix, (as before), and $\hat{\underline{s}}_k$ the current estimate of the state.

In order to obtain the Kalman gain, K_k, it is necessary to define the quantities T_k, M_k, P_k, L_k and Ω_k as in[63], so that the Kalman gain may be obtained recursively from,

$$K_k = -(FT_k\underline{H}_k - M_k)\Omega_k^{-1}. \tag{4.19}$$

The covariance associated with the innovations sequence, ϵ_k, is termed Ω_k being defined as $E[\epsilon_k \epsilon_k^T]$ and obtained recursively from,

$$\Omega_k = L_k - \underline{H}_k^T T_k \underline{H}_k. \tag{4.20}$$

Where L_k represents $E[y_k y_k^T]$ and is the covariance of the output sequence y_k being

obtained recursively from,

$$L_k = \underline{H}_k^T P_k \underline{H}_k + W_k. \tag{4.21}$$

The covariance of the state $\underline{s}_k$ being P_k, that is $P_k = E[\underline{s}_k \underline{s}_k^T]$ and obtained from the recursion,

$$P_{k+1} = FP_kF^T + G_kQ_kG_k^T. \tag{4.22}$$

The quantity M_k representing the cross-covariance of the state and the observation sequence being obtained via,

$$M_k = FP_k\underline{H}_k + G_kS_k \tag{4.23}$$

with the sequence T_k representing the covariance of $\bar{\underline{s}}_k$, that is $T_k = E[\bar{\underline{s}}_k \bar{\underline{s}}_k^T]$ and given that $T_0 = 0$ then,

$$T_{k+1} = FT_kF^T + (FT_k\underline{H}_k - M_k)(L_k - \underline{H}_k^T T_k \underline{H}_k)^{-1} (FT_k\underline{H}_k - M_k)^T \tag{4.24}$$

The innovations form of the extended Kalman filter may then be written down directly as before, Table 4.2 details the algorithm in full.

4.3.3 FINITE PRECISION CONSIDERATIONS

As was discussed in section 4.2 in most applications of Kalman filtering modelling errors arise due to the imperfect knowledge of the system being observed. This concept may also be applied when the finite nature of the digital machines on which the algorithm is implemented are taken into account as was discussed previously in chapter 3. The inherently complex nature of the EKF algorithm ensures that it will suffer from such errors as a result of digital word truncation and round off errors in matrix multiplication.

There are two possible approaches to implementing the solution to this numerical stability, the whole identification process could be restarted and the algorithm reinitialised to the original preset values. However, this would require complete retraining of the algorithm and is clearly unattractive. Alternatively detection of the onset of numerically inspired divergence could be monitored by some means and some

Description	Equation
State estimate	$\hat{s}(k+1) = A_k \hat{s}(k) + K_k \epsilon(k)$
State estimate	$\hat{\theta}(k+1) = \hat{\theta}(k) + L(k)\,\epsilon(k)$
Kalman gain	$L(k) = [\,\Sigma_{12}(k)^T\, C_k^T + \Sigma_{22}(k) D_k^T\,]\,\hat{\Lambda}_k^{-1}.$
Covariance matrix	$\Sigma_{11}(k+1) = A_k \Sigma_{12} + M_k \Sigma_{22}^T(k) - K(k)\hat{\Lambda}_k L^T(k)$
Covariance matrix	$\Sigma_{22}(k+1) = \Sigma_{22}(k) - L(k)\hat{\Lambda}(k) L^T(k) - \delta\Sigma_{22}(k)\Sigma_{22}(k)$
Log likelihood	$\hat{\Lambda}(k) = \hat{\Lambda}(k-1) + \frac{1}{k}[\,\epsilon(k)\epsilon(k) - \hat{\Lambda}_{(k-1)}\,]$

This formulation assumes the state space model previously described in equations (4.1) and (4.2) and uses the natural block structure as in table 4.1

Table 4.2 - INNOVATIONS FORM OF EXTENDED KALMAN FILTER ALGORITHM

form of partial reinintialisation invoked as in Lin[81]. A more fundamental approach to improving the numerical stability of the algorithm could be achieved by the use of orthogonal decompositions such as rotations or reflections as in[75, 76]. In this work it is considered sufficient to demonstrate the performance possible by the use of a-priori knowledge and the development of numerically stable techniques is left until a later date.

4.4 A-PRIORI KNOWLEDGE AND THE LMS ALGORITHM

4.4.1 Introduction

The performance of various adaptive HF channel estimators have been studied so far, the results may be summarised as follows:-

1) Minimum Variance Kalman (MVK) estimator:- this estimator requires full a-priori knowledge of the channel and as a result is not implementable. However it provides the lowest achievable MSE bound on any HF channel estimator.

2) EKF algorithm:- this technique relies on partial a-priori knowledge of the channel and its performance approaches that of the MVK, but it is very computationally complex and is liable to suffer from numerical instability consequently making implementation difficult.

3) Adaptive FIR filters (LMS/RLS algorithms):- these techniques have no a-priori knowledge of the channel and consequently have the poorest performance although the LMS is the least computationally complex.

Clearly the ideal estimator would be one which had a performance approaching that of the MVK but with a level of computational complexity comparable with the LMS algorithm, as illustrated in Figure 4.4. This section considers one approach in attempting to achieve this.

Essentially the LMS recursion is increased from a first order to a second order recursion by utilising a prediction filter which incorporates partial a-priori knowledge

of the tap generation process. The algorithm was implemented on computer as a set of parallel prediction filters and LMS algorithms (as illustrated in Figure 4.5), each prediction filter catering for a possible fade rate.

The illustration in Figure 4.5 bears a striking resemblance to the time sequenced adaptive filter suggested by Ferrara and Widrow in[97], the difference lying in the presence of a prediction filter. In their application of enhancement of electrocardiogram traces, Ferrara and Widrow adopted a similar approach to here in that they utilised a-priori knowledge of the signal being analysed to improve the performance of their system.

4.4.2 Algorithm Development

Again assuming the state space description of the tap generation model of the channel, a prediction filter for each of the taps may be constructed. The ideal input for these prediction filters would be the actual taps, which are of course not directly observable. However, the estimate of the taps from the LMS algorithm is available and this may be used, although it is in fact a noisy observation. The predicted value of the taps obtained may then be used in the LMS algorithm by increasing the order of the recursion from a first to a second order.

In order to derive the required prediction filters directly, it would be necessary to carry out a minimum phase spectral factorisation on the tap generation process, this would be both difficult and computationally intensive since the input of the filters is a contaminated tap estimate. As a result a Kalman filter approach is adopted since this is equivalent for white noise processes as is the case here.

If the representation of the channel specified previously is used once again then it is possible to develop a one stage predictor algorithm[101], thus,

$$\hat{\underline{s}}_k^{k-1} = F\hat{\underline{s}}_{k-1}^{k-2} + K_k^{k-1} v_k \tag{4.25}$$

where $v_k = \underline{h}_k - H_k \hat{\underline{s}}_{k-1}^{k-2}$, i.e. the error in the tap estimate and the Kalman gain,

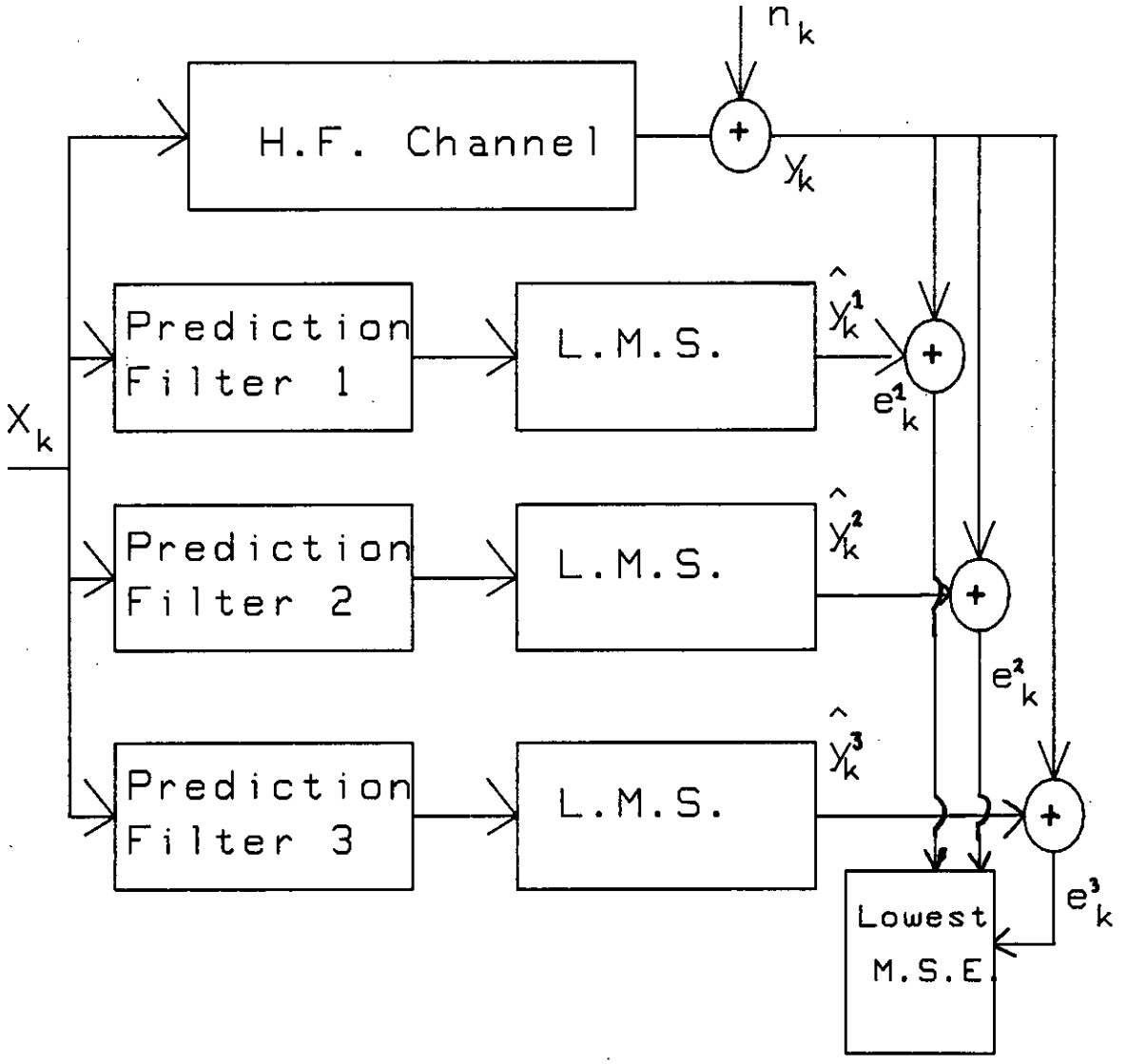


Figure 4.5 - Structure of modified LMS algorithm

K_k^{k-1} is obtained from,

$$K_k^{k-1} = F_k \; V_{\tilde{s}}(k/k-1) \; H_k^T \; V_v^{-1}(k). \tag{4.26}$$

The covariances $V_{\tilde{s}}(k/k-1)$ and $V_v(k)$ are defined by,

$$V_{\tilde{s}}(k/k-1) = E \; [\; (\underline{s}_k - \hat{\underline{s}}_k^{k)} \; (\underline{s}_k - \hat{\underline{s}}_k^{k)}{}^T \;]$$

and,

$$V_v(k) = E \; [\; v_k \; v_l \;]$$

respectively and may be obtained from the following equations,

$$V_{\tilde{s}}(k+1/k) = F \; V_{\tilde{s}}(k/k-1) \; F^T + G_k \; V_w(k) \; G_k^T - K_{k+1}^k \; H_k \; V_{\tilde{s}}(k/k-1) \; F^T, \tag{4.27}$$

and,

$$V_v(k) = H_k^T V_{\tilde{s}}(k/k-1) + V_u(k). \tag{4.28}$$

If the variance and gain equations are taken to the steady state values by computer simulation, then the steady state gain, $\underline{K}_{ss}$, will be obtained and may then be used for prediction. This mechanism is then used to increase the recursion of the LMS. If the variance and gain equations (eqns. 4.25-4.27) are taken to the steady state by computer simulation then the steady state gain, $\underline{K}_{ss}$ is obtained and this can be used in the algorithm for prediction as detailed in the equations shown below.

$$\hat{h}_k^{k-1} = H_k \; \hat{\underline{s}}_k^{k-1} \tag{4.29}$$

that is the predicted value of the tap based on an estimate of the state obtained from,

$$\hat{\underline{s}}_k^{k-1} = F \; \hat{\underline{s}}_{k-1}^{k-1} \tag{4.30}$$

and,

$$\hat{\underline{s}}_{k-1}^{k-1} = \hat{\underline{s}}_{k-1}^{k-2} + \underline{K}_{ss} \; [\; \hat{h}_{k-1} - \hat{h}_{k-1}^{k-2} \;] \tag{4.31}$$

That is, since the actual taps are not observable then the estimate obtained from the LMS, $\hat{h}_{k-1}$, as illustrated below, is used to aid the prediction of the state.

$$\hat{h}_{k-1} = \hat{h}_{k-1}^{k-2} + 2\mu x_{k-1} e_{k-1} \tag{4.32}$$

where x_{k-1} represents the input signal vector and the error e_{k-1} is obtained from,

$$e_{k-1} = y_{k-1} - \hat{y}_{k-1} \tag{4.33}$$

that is the output of the channel less the estimated output where the output

$y_{k-1} = h_{k-1}^T x_{k-1} + n_k$ where n_k is additive white gaussian noise. The estimated output is,

$$\hat{y}_{k-1} = \hat{h}_{k-1}^{k-2} x_{k-1} \tag{4.34}$$

Table 4.3 below summarises the modified LMS algorithm.

Description of operation	Equation
predicted value of tap based on state estimate	$h_k^{k-1} = C\hat{s}_k^{k-1}$
predicted value of state based on an estimate of the state	$\hat{s}_k^{k-1} = F\ \hat{s}_{k-1}^{k-1}$
estimate of state	$\hat{s}_{k-1}^{k-1} = \hat{s}_{k-1}^{k-2} + K_{ss}\ [\ \hat{h}_{k-1} - \hat{h}_{k-1}^{k-2}\ ]$
estimate of tap coefficient	$\hat{h}_{k-1} = \hat{h}_{k-1}^{k-2} + 2\mu x_{k-1} e_{k-1}$
the error	$e_{k-1} = y_{k-1} - \hat{y}_{k-1}$
estimated output	$\hat{y}_{k-1} = \hat{h}_{k-1}^{k-2} x_{k-1}$

(Where x_{k-1} represents the input signal vector.)

Table 4.3 - LMS Channel Estimator with a-priori Modifications

4.5 PERFORMANCE COMPARISONS

The simulations carried out which are reported here were performed for a channel of fade rate 10.0Hz and with and additive noise power set at -50dB.

The performance of the LMS, MVK and EKF algorithms are illustrated in Figures 4.6 to 4.10. Figures 4.6-4.8 show the performance for white input signal conditions and Figures 4.9-4.10 for coloured input signal, (eigenvalue ratio of 11.8) conditions. Table 4.4 below summarises the performance of each of the algorithms as channel estimators for white and coloured inputs.

Figure	Algorithm	EVR	Steady-State MSE
4.6	MVK	1	-45dB
4.7	Modified LMS	1	-25dB
4.8	EK	1	-45dB (*)
4.9	MVK	11.8	-43dB
4.10	EK	11.8	-28dB (!)

* - achieved after initialisation with LMS

! - algorithm diverges due to numerical instability very quickly

Table 4.4 - Summary of performance of algorithms in simulations

Comparing the performance of the LMS with priori knowledge that of the MVK it is clear that the LMS is some 20dB from the noise floor. This is as a result of the contribution to the error by the time variations in the system as discussed by Macchi in[52] and described in chapter 3. It can be seen from Figure 4.8 that the EKF provides an improvement in perfromance of some 5-20dB. The EKF utilised the LMS to provide an initial estimate and as can be seen clearly improves upon it. Figures 4.9-4.10 illustrate the algorithms performance under similar conditions to those above except that the input signal is now coloured. The EKF approaches that of the MVK

which is independent of input signal colouration since its performance is based on the use of a-priori knowledge about the channel model. Unfortunately the performance of the EKF degrades with time, this is due to the inherent numerical instability of the algorithm as discussed in section 4.3.3.

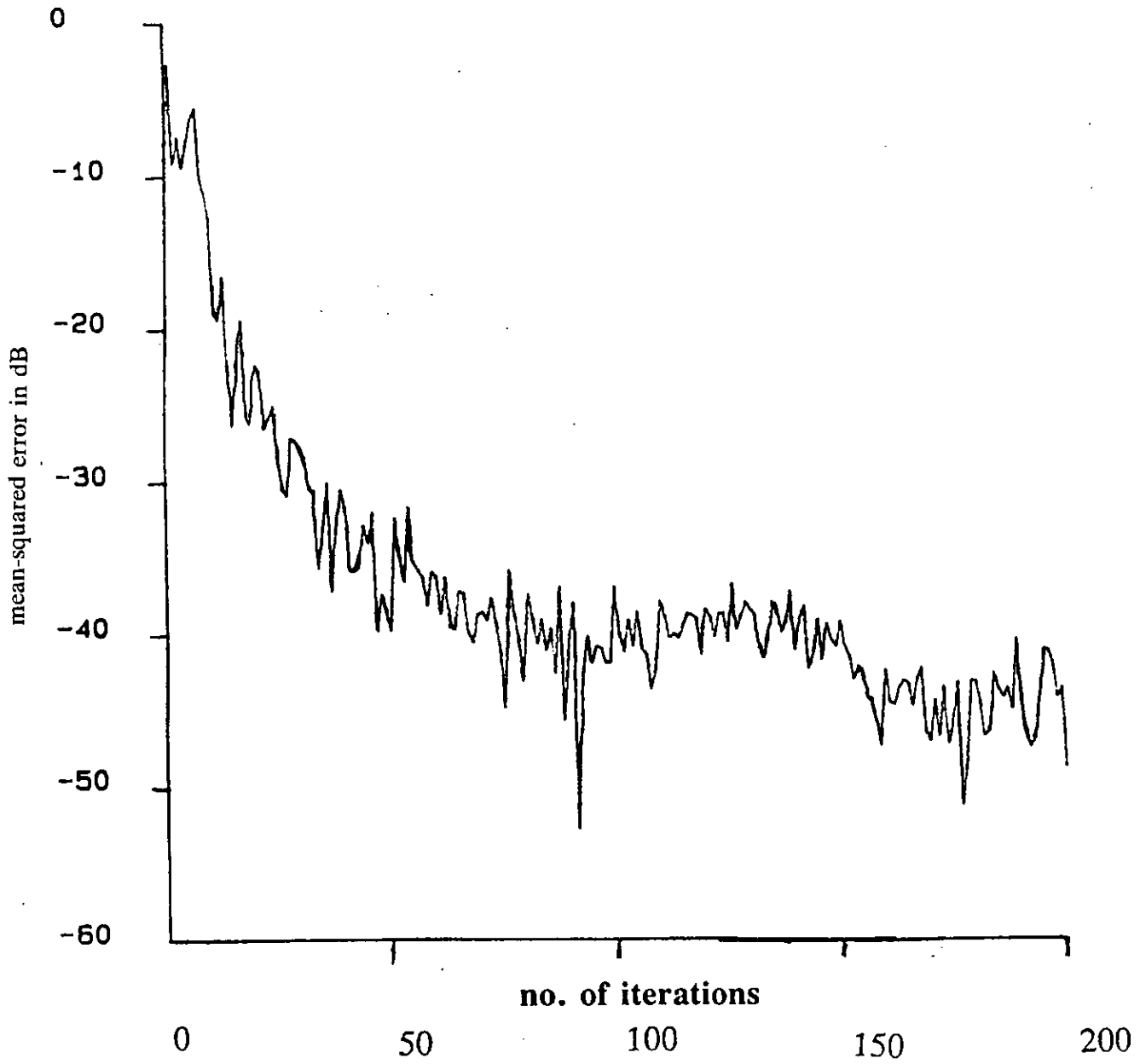


Figure 4.6 - Performance of MVK for white input signal on 10Hz fading channel with additive noise power = -50dB.

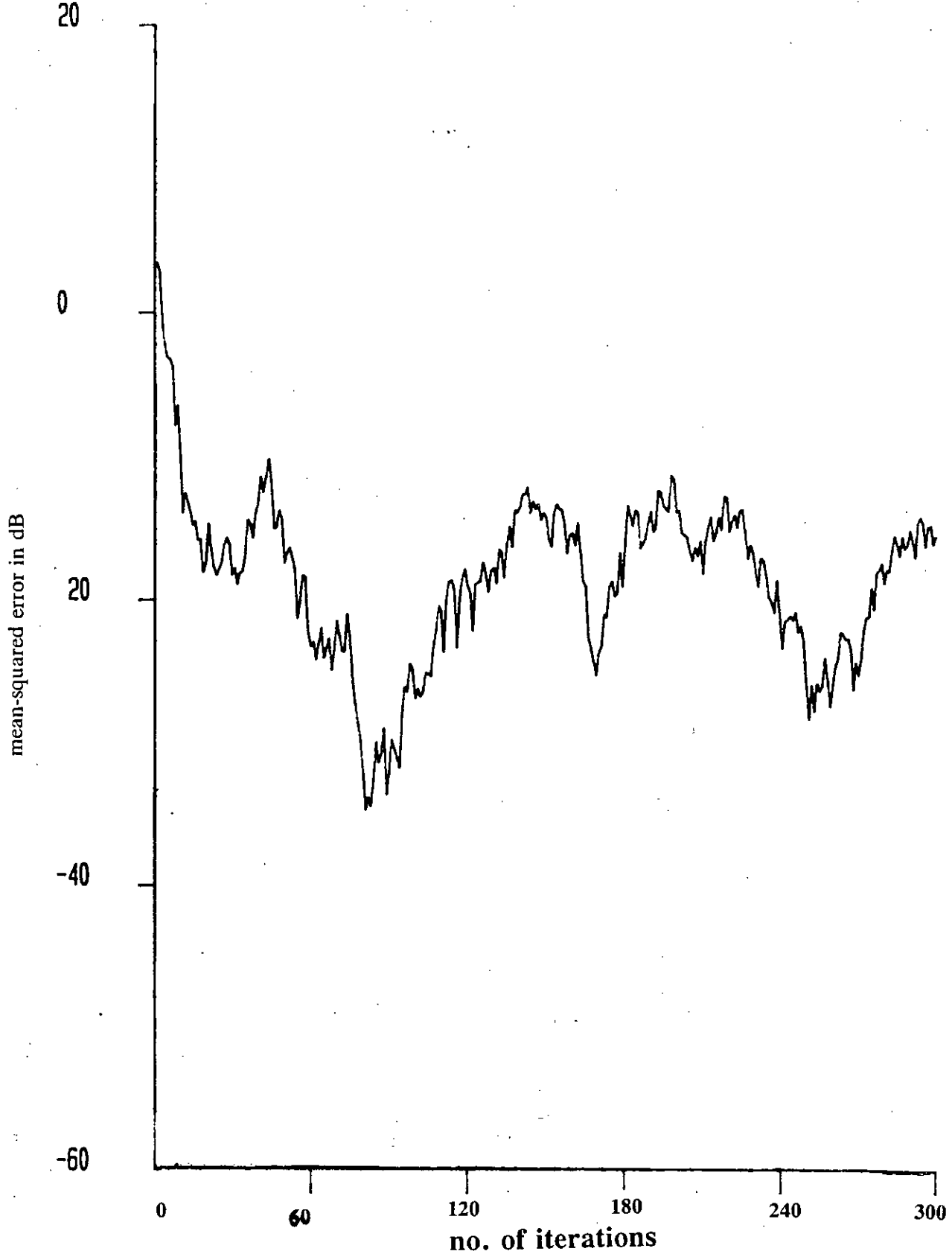


Figure 4.7 - Performance of modified LMS algorithm for white input signal on 10Hz fading channel with additive noise power = -50dB.

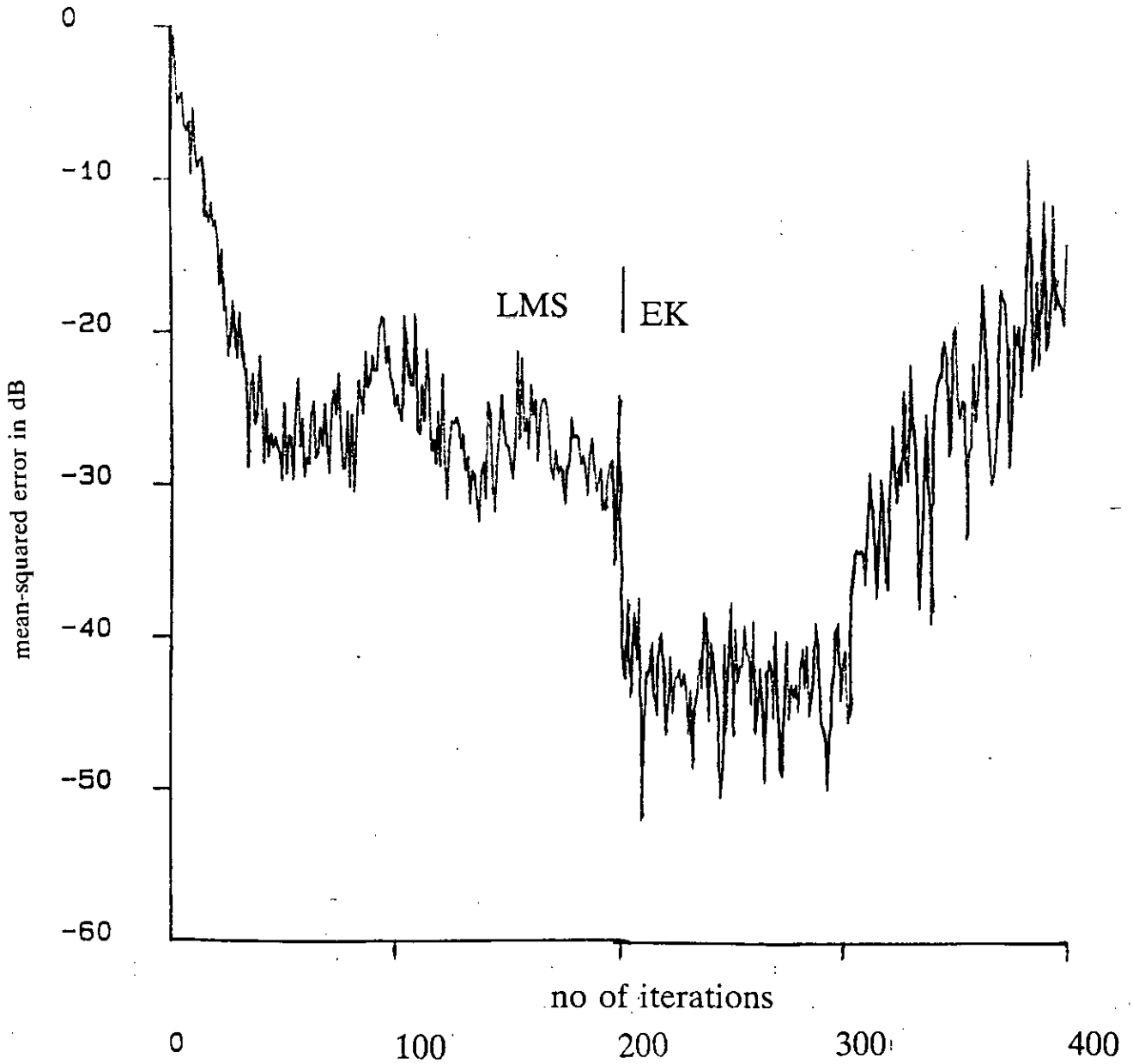


Figure 4.8 - Performance of Extended Kalman algorithm for white input signal on 10Hz fading channel with additive noise power = -50dB.

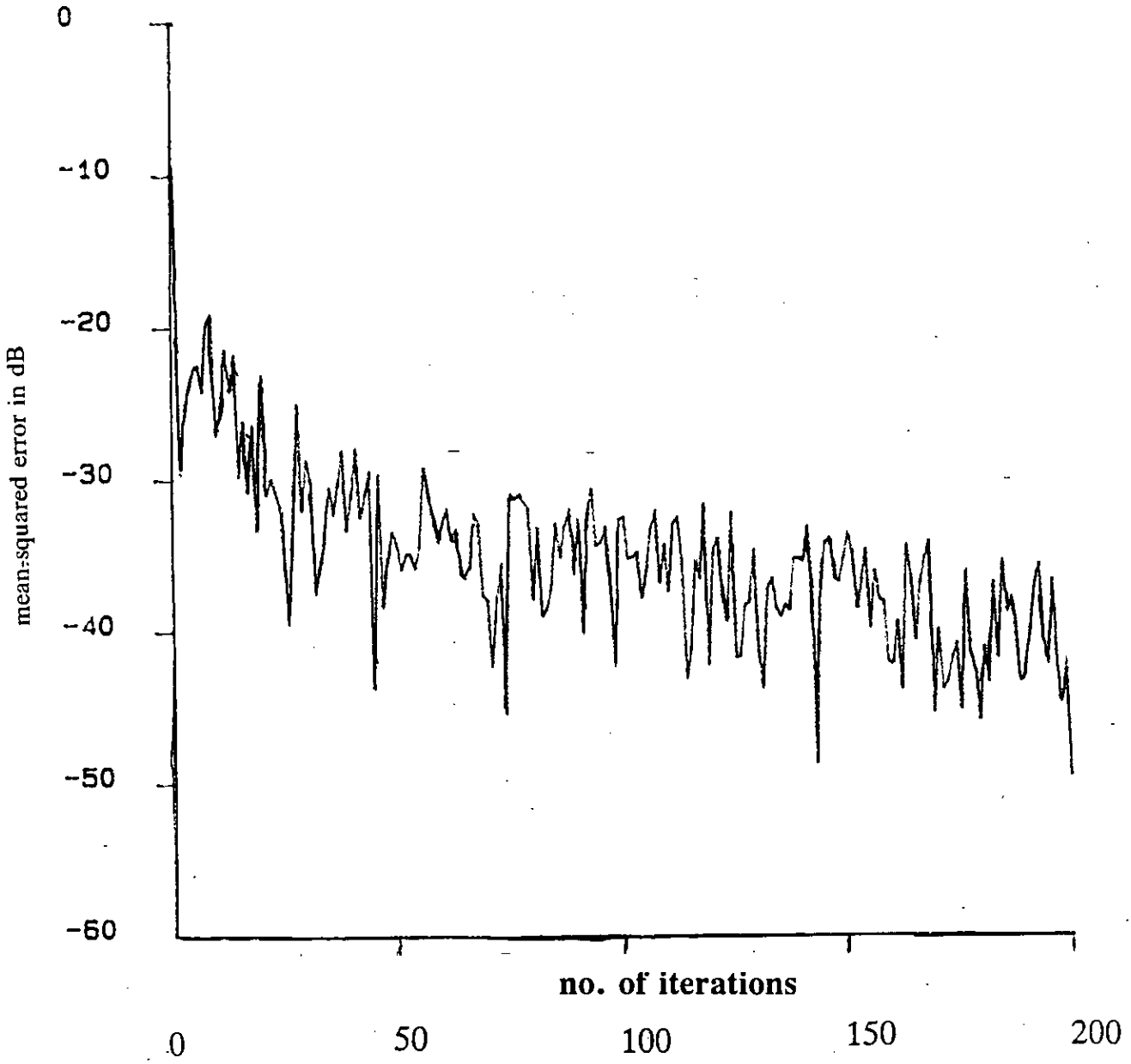


Figure 4.9 - Performance of MVK algorithm for coloured input signal (evr=11.8) on 10Hz fading channel with additive noise power = -50dB.

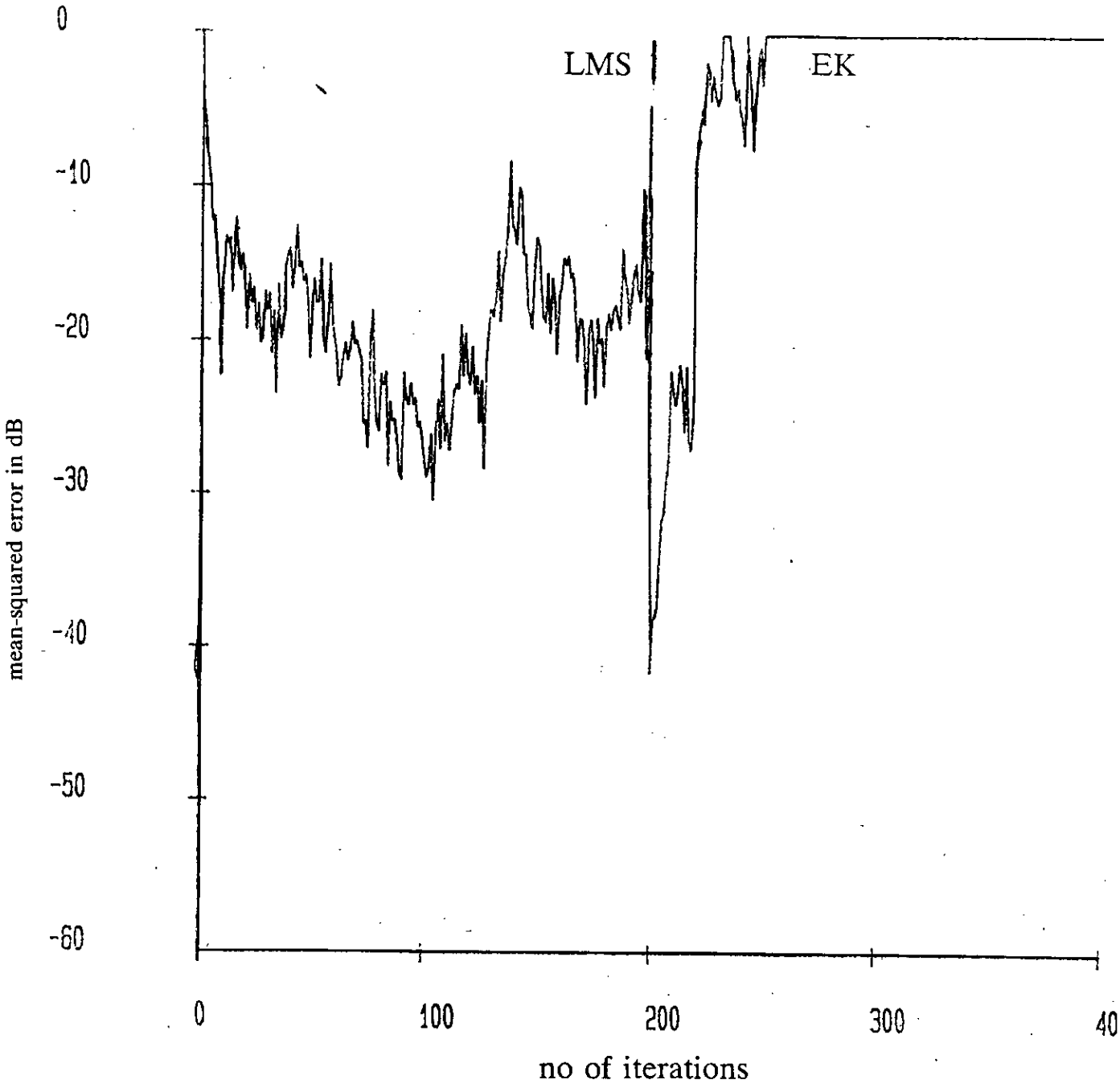


Figure 4.10 - Performance of Extended Kalman algorithm for coloured input signal (evr=11.8) on 10Hz fading channel with additive noise power = -50dB.

4.6 CONCLUSIONS

In this chapter three new algorithms for the HF channel have been developed each performing with various degrees of success. The MVK requires full a-priori knowledge, a constraint which precludes implementation. The EKF although providing a good performance in the steady-state MSE sense is complex and numerically unstable. Finally the PLMS is the least complex but performed least well, however it has not fully been explored and the work of Clark et al in [102-104] offer some hope that this type of technique may still be useful if a suitable predictor filter was utilised.

In summary this work has demonstrated the possibility of utilising a-priori knowledge of the channel being identified to improve the performance. Ultimately such an estimator would be incoporated in an equaliser where the channel estimation and decision process are seperated, as reported in, Mulgrew paper 1987 the following chapter now considers such structures.

Chapter 5

ADAPTIVE EQUALISATION OF THE HF CHANNEL

5.1 INTRODUCTION

All the chapters preceeding this one have considered channel estimation only, in this chapter adaptive equalisation is considered. Equalisation is used in communication systems to compensate for the distortion introduced into the transmitted data sequence by the communications channel, Quershi in [105] illustrates several applications. When an equaliser is termed adaptive, it is assumed capable of some form of self-adjustment to deal with variations which arise in the channel impulse response consequently causing distortion. This self-adjustment is normally achieved by incorporating an adaptive algorithm such as the LMS or RLS into the structure to set the tap weights of the equaliser based on some criterion, such as MSE, to a value which minimises the distortion in the system.

This chapter considers three possible structures, one novel, and their MSE performance in the HF communication scenario. The three structures are:

1) A conventional decision feedback equaliser (DFE) which utilises a Godard-Kalman adaptive algorithm to carry out the tap weight update.

2) A linear Kalman equaliser which utilises an LMS channel estimator, reported in[29], separating the channel and sequence estimation processes.

3) A Kalman based equaliser which like the above uses an LMS algorithm to carry out the channel estimation but which incorporates an element of decision feedback.

The structures studied in this chapter are detailed in the following sections, however their relationship to each other may be shown in a qualitative manner by Figures 5.1a-c. Figure 5.1a illustrates, in block diagram form, a conventional DFE in which the adaptive estimation and data equalisation functions are performed at the same time, Figure 5.1b represents the adaptive Kalman equaliser reported in [29] and

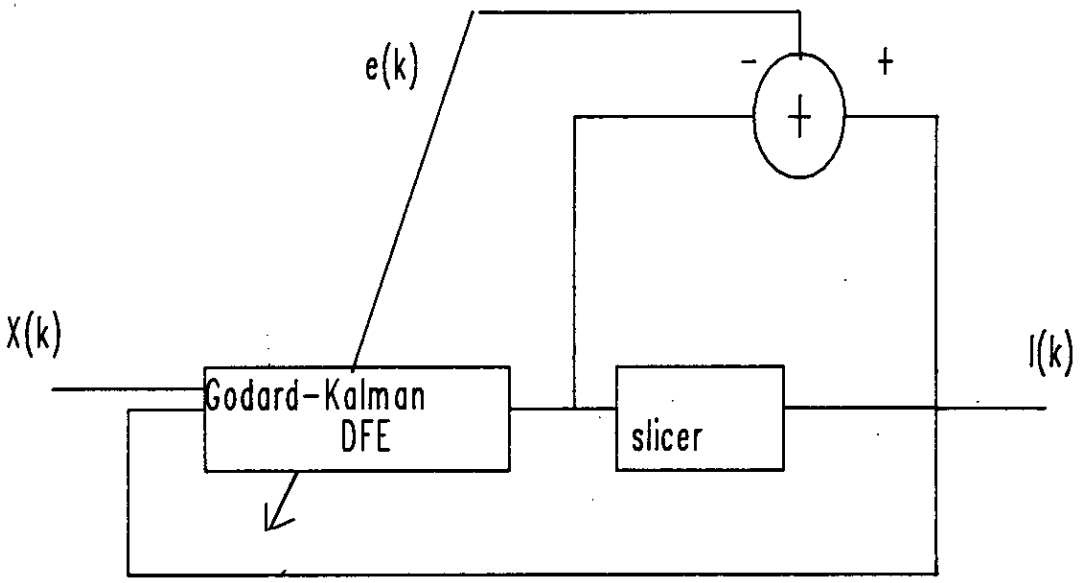


Figure 5.1a - Godard-Kalman decision feedback equaliser

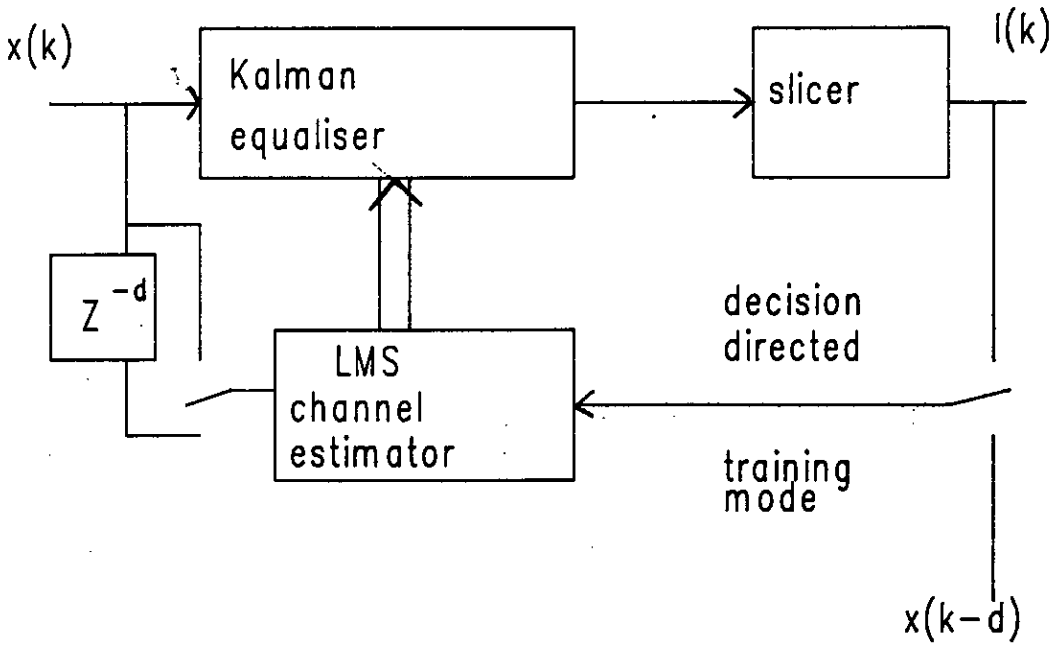


Figure 5.1b - Adaptive Kalman equaliser

as can be seen the channel estimation and data equalisation processes have been separated. The final illustration, Figure 5.1c, depicts the novel adaptive Kalman decision feedback equaliser, where as in Figure 5.1b, the estimation and equalisation processes have been separated but now decisions are fedback into the structure.

This chapter is structured as follows, initially the equalisation problem is highlighted and existing structures which have been reported elsewhere summarised with their advantages and disadvantages discussed. Following on from this, the development of the adaptive Kalman equaliser of [29] is briefly illustrated and used to develop the novel Kalman decision feedback structure. Finally comparisons of the performance of the three structures considered here are presented and conclusions drawn based on the results, the relative performances considered in terms of their steady state MSE.

5.2 EQUALISER STRUCTURES

In this section a brief resume of existing equaliser structures is presented before illustrating the development of the Kalman decision feedback structure.

5.2.1 Introduction

As has been intimated previously adaptive equalisation of radio and telephone communication channels is used to compensate for the time dispersion introduced to the transmitted data sequence. This time dispersion introduces intersymbol-interference (ISI) into the transmitted data sequence. The nature of this ISI in the HF channel can be appreciated by consideration of the multipath nature of the channel which results in the energy associated with the transmission of one symbol being smeared across several symbol periods as illustrated in Figure 5.2.

Research over the last twenty years has produced a large body of literature [105] and references therein, there are many types of equaliser structure, and they may be summarised as follows:

a) Linear transversal equalisers which in general suffer from an inability to represent

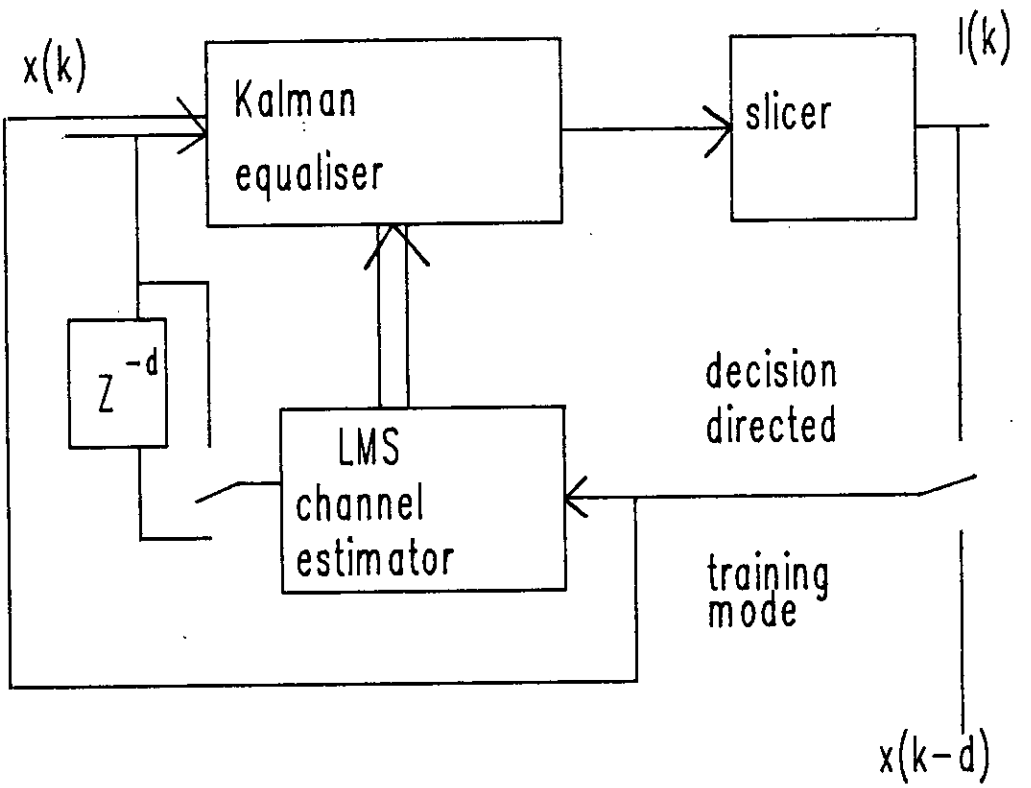


Figure 5.1c - Adaptive Kalman DFE

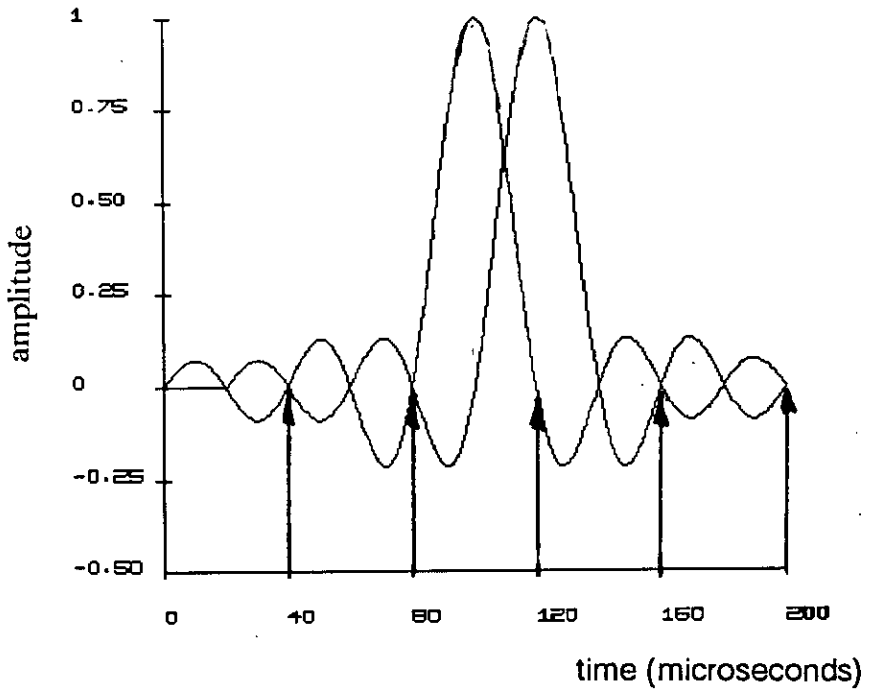


Fig. 5.2a – No Inter–Symbol Interference (ISI) present at sample points.

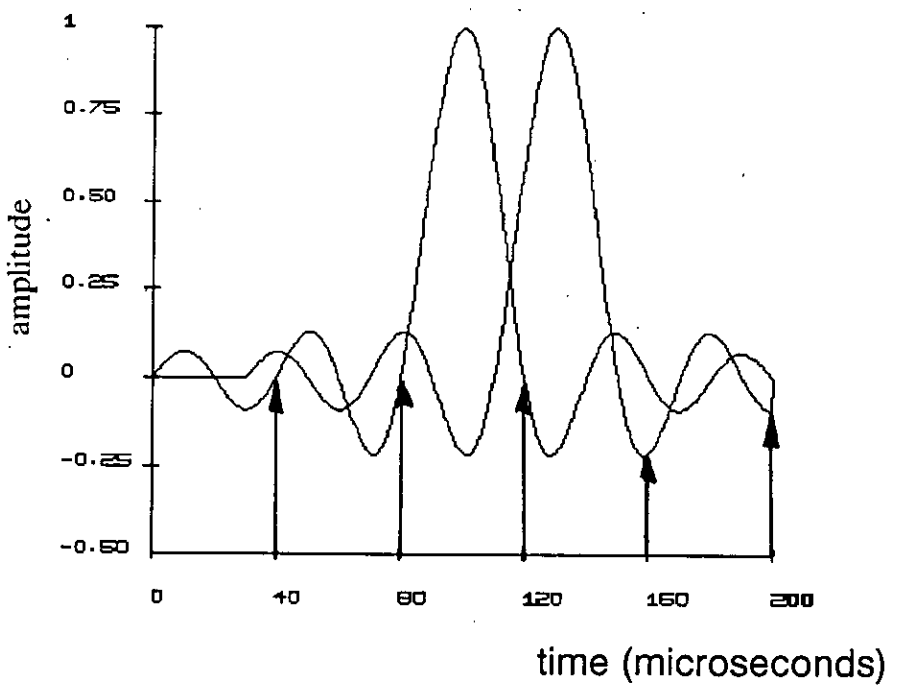


Fig. 5.2b – Inter–Symbol Interference (ISI) present at sample points.

the inverse of the channel transfer function adequately.

b) Conventional decision feedback equalisers which, although providing a better performance than linear equalisers, suffer from a degraded performance due to error propagation in the feedback section.

c) Maximum likelihood sequence estimation is a technique which is not considered in this thesis but whose main disadvantage would appear to be its computational complexity, however considerable effort is being expended in developing more efficient implementations

In general most equalisers have two modes of operation, training and decision-directed. In training mode the transmitted data sequence is known a-priori and is termed the training sequence. In this mode the a-priori knowledge of the training sequence is utilised to ensure the coefficients of the equaliser achieve the appropriate values to mitigate the ISI. On completion of the training sequence the equaliser switches into decision-directed mode, i.e. the data sequence is not known a-priori and the equaliser must assume that all decisions it takes are correct. Clearly this may not always be the case and the equaliser will clearly suffer a degraded performance in such situations. There are some applications, such as microwave line-of-sight communication systems[106, 107], where a training sequence is not present. This is termed *blind* equalisation [108] since the equaliser is required to bootstrap into decision-directed mode. It will be apparent that in such situations the level of complexity required in the equaliser is much greater than the situation being considered here. This is because the system has no prior knowledge of the channel or of the transmitted data sequence as would normally be provided by the training sequence.

The selection of coefficients for the equaliser to minmise the effect of ISI may be based on many criteria[109-111], the most suitable would be the probability of error. However this is a highly nonlinear function and generally not practicable, hence the most common criteria is the MSE, which is the sum of squares of all the ISI terms and the noise power at the output of the equaliser. Many equalisers operate by generating

an estimate of the inverse filter which when convolved with the channel response allows the transmitted data sequence to be reconstructed accurately, as in the linear equaliser illustrated in Figure 5.3. For conventional DFE's, a feedback filter is inserted after the decision device, (as in Figure 5.4), and is used to cancel out any trailing intersymbol interference (ISI) by using previously detected symbols, which are assumed to be correct.

In all the situations which will be considered here the equaliser tap weights are symbol spaced, however it is perfectly feasible to have fractionally-spaced coefficients. In general the spacing is chosen to be T/2, where T represents the symbol period. The motivation for using a fractionally-spaced equaliser, (FSE), are it's relative insensitivity to timing phase and ability to deal with more severe delay distortion than a symbol-spaced equaliser. However, as in all engineering applications there is a penalty to be paid and in this case the complexity is increased since a FSE requires twice as many coefficients as a symbol-spaced equaliser.

A range of adaptive algorithms are used in adaptive equalisation, the two most common being the LMS and RLS algorithms which have been discussed in detail previously. The LMS offers an easily implementable algorithm but lacks the spectral robustness and initial fast convergence of the RLS, which unfortunately is relatively complex to implement. The conclusions presented in chapter three suggest that when operated as a channel estimator, as opposed to an equaliser, the LMS offers a similar if not improved performance, in terms of the steady state MSE achieved, than the RLS on channels which are time-varying and this is reinforced by the results reported elsewhere[26, 56-58]. It is these results which have provided some of the motivation for the novel structure considered here, in that the data sequence estimation and channel impulse response estimation processes are separated.

5.2.2 Linear Equaliser Structures

The simplest structure used for equalisation is the transversal equaliser, Fig. 5.3.

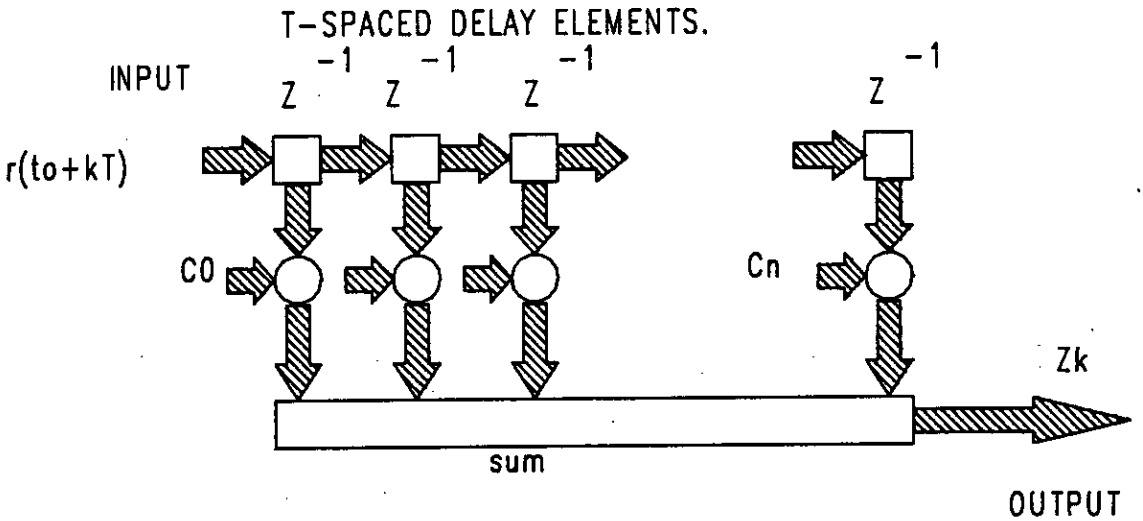


Figure 5.3 - Linear transversal equaliser structure

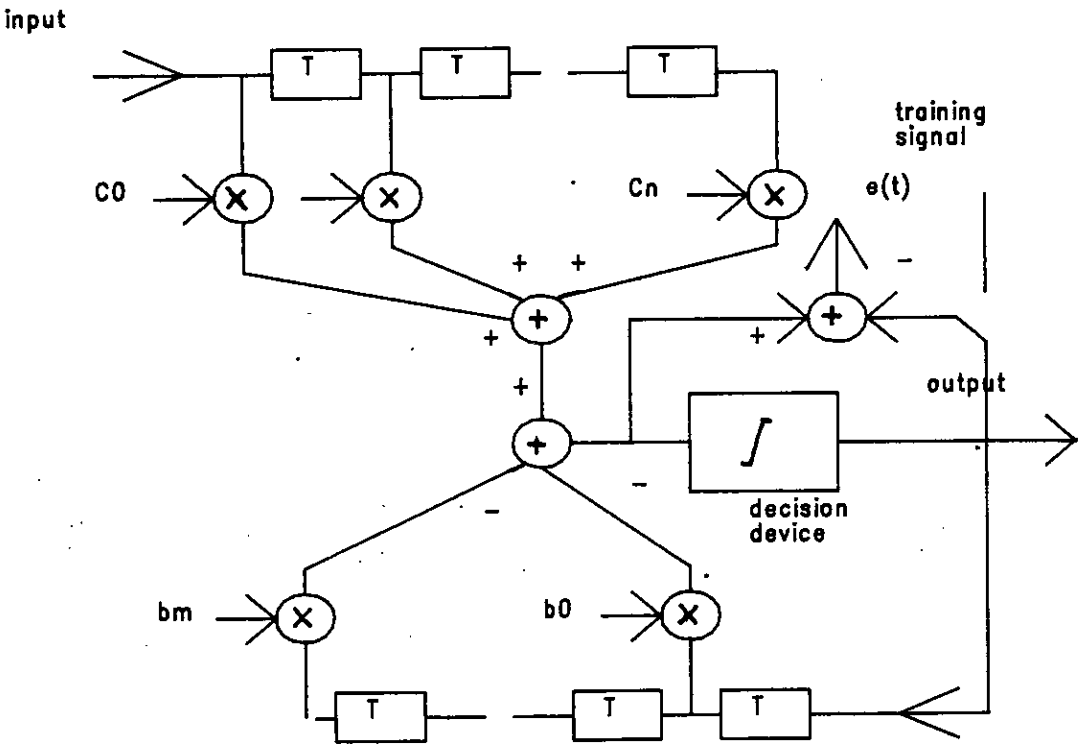


Figure 5.4 - Decision feedback equaliser structure

If a digital communications scenario is assumed, then the channel may be modelled as a discrete time transversal filter with additive white Gaussian noise, the output of the channel can be written down as,

$$y_k = \underline{h}_k^T \underline{x}_k + n_k \tag{5.1}$$

where $\underline{x}_k$ represents the channel input vector with $x_{i,k}$ representing it's constituent components and $\underline{h}_k$ is the M-point impulse response vector, which may or may not be time-variant, with components represented by $h_{i,k}$. In a similar way the equaliser output may be written down as,

$$\hat{x}_{i,k} = \underline{c}_k^T \underline{Z}_k \tag{5.2}$$

where $\underline{c}_k$ represents the N-point impulse response vector of the equaliser and

$$\underline{Z}_k = [\ y_k\ y_{k-1}\ y_{k-2} \cdots\cdots\cdots y_{k-N}\ ] \tag{5.3}$$

is the vector which contains the N previous channel outputs. It is clear that the coefficients of such an equaliser are essentially being selected to force the combined channel and equaliser impulse response to approximate a unit pulse. That is, the equaliser must approximate the inverse filter of the channel. This requirement results in the equaliser suffering from excessive noise enhancement and sensitivity to sampler timing phase[105].

5.2.3 Decision Feedback Equalisers.

In attempt to overcome the performance limitations of linear equalisers, as discussed previously, a simple non-linear equaliser was developed, the decision feedback equaliser as illustrated in Figure 5.4. Essentially it is a transversal equaliser with a feedback section which uses past decisions to cancel out the ISI associated with these detected symbols.

$$\hat{x}_{i,k} = \underline{c}_k^{FF} \underline{Z}_k + \underline{c}_k^{FB} \underline{\hat{x}}_k \tag{5.4}$$

where,

$\underline{c}_k^{FF}$ represents the A coefficients in the feedforward section of the equaliser,

$\underline{c}_k^{FB}$ represents the B coefficients in the feedback section of the equaliser, and

$\underline{\hat{x}}_k$ is the vector containing the B previous decisons.

The DFE as a result of the feedback section's ability to cancel the ISI associated with past symbols removes the constraint from the feedforward section that it approximate the inverse filter of the channel. This means that excessive noise enhancement and sensitivity to sampler timing phase is reduced. However it is clear that if an incorrect decision is made that it will propagate through the feedback section thus increasing the likelihood of more incorrect decisions, i.e. error propagation will occur. The effect of error propagation has received little attention[112], since for many applications the error bursts which would result in error propagation are relatively short. In the time-varying HF channel it is clear from previous results [56, 113] that this is not the case, consequently a degraded performance of the DFE results.

5.2.4 Maximum Likelihood Sequence Estimation

As was discussed in section 5.2.1 conventional adaptive equalisers utilise the MSE criteria in general in order to minimise the ISI in the received data sequence. This is not an ideal criterion the probability of error being more suitable but unfortunately highly nonlinear. This has motivated many researchers to investigate the use of other nonlinear criteria, as in[114]. Such receivers usually use the maximum a-posteriori probabability rule[115] to maximise the probability of correctly detecting each symbol as in[116], or the entire transmitted sequence. These receivers are termed maximum likelihood receivers (ML), the classical ML receiver [114] may be viewed as a bank of m^k matched filters, where k is the length of the sequence of symbols which come from a discrete alphabet of size m. Unfortunately the computational complexity of such receivers increases exponentially as the sequence length increases although the Viterbi algorithm [117, 118] partially overcomes this problem.

In general MLSE receievers require knowledge of the channel and it is necessary to utilise an adaptive channel estimator [119] as illustrated in Figure 5.5. This estimator will of course only provide an estimate of the channel and this introduces a possible source of error under severe non-stationary or high noise conditions. Also,

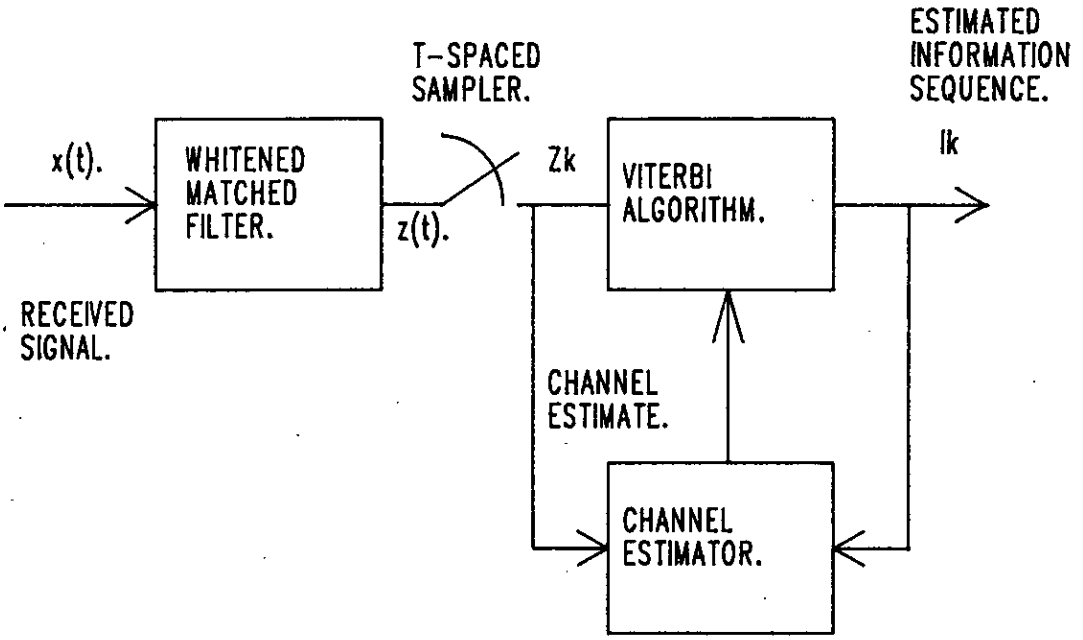


Figure 5.5 - Adaptive maximum likelihood structure

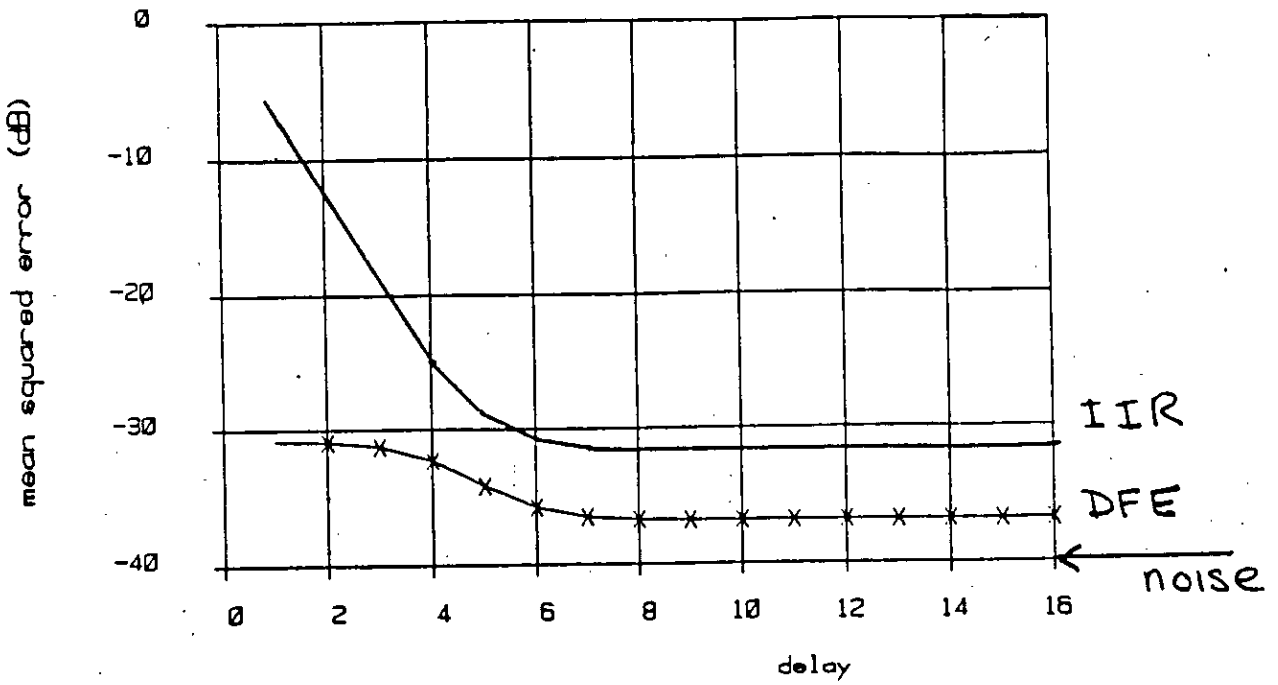


Figure 5.6 - Optimum MSE performance of IIR and DFE structures

although the Viterbi stops the exponential growth in complexity it still places a considerable load on any processor, as a result of this much work has been devoted in recent years to developing reduced state MLSE detectors as in the following[120-122]. This coupled with the growth in capability of VLSI techniques ensure that MLSE receivers will receive more and more attention in the future.

5.3 GODARD KALMAN DECISION FEEDBACK EQUALISER

The conventional DFE structure differs, as has been stated earlier, from the linear structure by the addition of a feedback section which is used to cancel out the ISI associated with these symbols. The feedback section allows a greater freedom for the linear section in selecting tap weight coefficients. Conventional DFE's of this type have been found to operate very well over wire line channels but in rapidly time-varying environments the performance appears to be degraded by error propagation in the feedback section.

In this chapter the algorithm which was used to adjust the tap weight co-efficients of the equaliser was the algorithm postulated by Godard [30] in 1974 in which he chose not to replace the equaliser with a conventional Kalman filter, but rather adopted a transversal equaliser structure and used the Wiener solution for the optium tap weights as a starting point. The algorithm offers very fast initial convergence and is spectrally robust but suffers from relatively high level of complexity. To apply it to the DFE, the observation vector contains both the feedforward and feedback coefficients, the algorithm for the DFE being summarised in Table 5.1. The solution obtained by Godard was for a stationary channel and its application was extended to slowly time-varying channels by means of exponential data windowing.

In terms of the conventional Kalman filter equations, Godard assumed that the state transition matrix was the Identity matrix and the state vector chosen to be the tap weights of the equaliser. The selection of states for the application of Kalman filter theory to data equalisation is crucial and is discussed later.

Description of operation	Equation
tap weight vector estimate	$\hat{\underline{H}}(k) = \hat{\underline{H}}(k-1) + \underline{K}(k)[x(k) - \hat{\underline{H}}(k)\hat{\underline{L}}(k-1)]$
Kalman gain vector	$\underline{K}(k) = \underline{V}(k-1)\hat{\underline{H}}^T(k)[\hat{\underline{H}}(k)\underline{V}(k-1)\hat{\underline{H}}^T(k) + \lambda]^{-1}$
Covariance matrix estimate	$\underline{V}(k) = 1.0/\lambda[I - \underline{K}(k)\hat{\underline{H}}(k)]\underline{V}(k-1)$

Where the estimated tap weight vector is $\hat{\underline{H}}(k) = [\hat{a}_0(k)\,\hat{a}_1(k) \ldots\ldots \hat{a}_f(k)\,\hat{b}_0(k)\,\hat{b}_1(k) \cdots \hat{b}_{fb}(k)]$, the Kalman gain vector is $\underline{K}^T(k) = [K_0(k)\,K_1(k) \ldots\ldots K_n(k)]$ with f the number of feedforward taps and fb the number of feedback taps.

Table 5.1 - Summary of Godard-Kalman DFE Equations

5.4 AN ADAPTIVE KALMAN EQUALISER

In [29] Mulgrew and Cowan presented a novel equaliser structure, the derivation of which may be summarised as follows. Initially, a channel model based on a FIR filter was postulated and the constraint that the optimum transversal equaliser for such a channel, which requires minimisation of its MSE subject to the impulse response being finite, causal and stable is relaxed. The new relaxed constraint requires only that the filter be causal and stable, this results in the solution to the minimisation problem being provided by a Wiener infinite impulse response (IIR) filter.

The motivation for considering an IIR structure was to try and overcome the limitations of conventional FIR equaliser structures. In many applications FIR solutions have been found to be perfectly adequate, and they are generally preferable since they are unconditionally stable. However these FIR filters suffer from indeterminate order when required to model transfer function poles, especially poles

close to the unit circle. The obvious alternative has been IIR filters, but, IIR filters are not unconditionally stable The IIR equaliser has received little attention recently with the development of the DFE. The DFE which although it has a superior MSE performance compared to an IIR equaliser when the number of signal levels are low or the noise is high, suffers from error propagation unlike the IIR equaliser. The development of the DFE has restricted the wider application of IIR structures because of the improvement it offers over FIR structures and the lack of guaranteed stability associated with IIR filters. However, DFE structures suffer from error propagation as a result of the feedback of previously detected symbols, the IIR based structure offers a means of overcoming this problem.

The Wiener IIR filter offers advantages over the conventional linear (FIR) equaliser in terms of the order required for the same level of performance for minimum phase channels. However, the realisation of such a filter would require a minimum phase spectral factorisation, this would present a major difficulty. The solution adopted was to use a Kalman filter to realise the solution since, if the processes are stationary and the observation noise white, then the steady state Kalman and Wiener IIR filters are identical. The FIR filter model of the communications channel in common usage although readily adapted to a state space representation, and hence to a Kalman filter, requires care in the selection of states which will constitute the state vector.

The care in the selection of states for the state vector is because the FIR filter model of the channel assumes M taps, that is to say it may be completely described by $M-1$ states. If this approach was adopted however it would result in the plant and observation noise terms being correlated. The Kalman filter for such a situation, although it exists, would only be conditionally stable, the stability being dependent on the channel impulse response. If the channel is described by a state vector with M states then the appropriate Kalman filter, as in [123] and[124] is unconditionally stable.

In order to deal with non-minimum phase channels, a fixed lag smoothing form of the Kalman filter was used. Fixed-lag smoothing [63] is concerned with the on-line smoothing of data with a fixed delay d between the signal being received and an estimate being made. As was intimated in section 4.1 in Kalman filters an estimate at iteration k is based on a set of noisy observations, clearly there need be no delay between the last observation and the next estimate. However should a delay be permitted then it is clear that more observations are available on which the estimate may be made. Unfortunately as in all engineering applications although a smoother would be expected to provide a better estimate than a filter because it has more observations on which to make estimates, a penalty of increased complexity is incurred. A fixed lag smoother would normally imply that for a fixed lag, d, a state vector augmented to,

$$[\, \underline{s}^T(k)\ \underline{s}^T(k-1)\ \ldots\ldots\ \underline{s}^T(k-d)\,]. \tag{5.5}$$

However, this is unnecessarily complex because the state transition matrix is simply a shift matrix, thus the state vector is augmented to contain $d+1$ elements,

$$\underline{s}^T(k) = [\, s(k)\ s(k-1)\ \ldots\ldots\ s(k-M+1)\ \cdots\ s(k-d)\,], \tag{5.6}$$

where d is the fixed lag and M is the number of taps in the channel. The state transition equation then becomes,

$$\underline{s}(k) = \underline{a}\underline{s}(k-1) + \underline{b}s(k), \tag{5.7}$$

where $\underline{a}$ is a $(d+1) \times (d+1)$ shift matrix and $\underline{b}$ is a vector with $(d+1)$ elements,

$$\underline{b}^T = [\, 1\ 0\ 0\ \ldots.\ 0\,]. \tag{5.8}$$

The observation equation is clearly,

$$x(k) = \underline{h}^T\underline{s}(k) + n(k), \tag{5.9}$$

where $\underline{h}$ is a column vector with $(d+1)$ elements,

$$\underline{h}^T = [\, h_0\ h_1\ \cdots\ h_{M-1}\ 0\ 0\ \cdots\ 0\,], \tag{5.10}$$

that is the channel tap weight vector augmented to d + 1 elements by the addition of zeros.

The problem still remains however of making the equaliser adaptive, since knowledge of the channel impulse response is required and this is achieved by using an LMS algorithm as a channel estimator to provide the required estimate. Some measure of the observation noise in the system is also required for the Kalman structure and this is achieved by the recursion shown below, the derivation of which is detailed in[29],

$$\hat{\sigma}_e^2(k+1) = (1 - 1/M)\hat{\sigma}_e^2(k) + e^2(k+1)/M, \tag{5.11}$$

where $e(k+1)$ is the error obtained from the LMS. This recursion also provides a measure of the model uncertainty in the system and therefore offers a means for model order reduction by using the residuals to provide some information on any paths not modelled by the LMS. The structure of the equaliser is illustrated in Figure 5.7. The main point to note about the general structure of this equaliser is the separation of the state and channel estimation processes, and this approach is extended in the next section. The Kalman filter can then be written down directly from these definitions and is detailed in Table 5.2.

5.5 AN ADAPTIVE KALMAN DECISION FEEDBACK EQUALISER

The motivation in the previous section for adopting an IIR structure was it's freedom from error propagation. However, the DFE offers an inherent MSE performance advantage, as illustrated in Figure 5.7. Figure 5.7 illustrates the *theoretical* MSE performance of both a DFE and an IIR equaliser on a stationary non-minimum phase channel, whose impulse response was described by $0.2602z^{-1} + 0.928z^{-2} + 0.2602z^{-3}$, where the additive noise power was set at -40dB. It can clearly be seen that for a range of estimation lags the performance advantage of the DFE is constant at approximately 5dB. This inherent *theoretical* performance advantage provides the motivation for the development of the Adaptive Kalman DFE now discussed.

As in development of the previous equaliser, the structure of the channel model is

Operation	Equation
state estimate	$\hat{\underline{s}}\ (k/k) = \hat{\underline{s}}\ (k/k-1) + \underline{K}(k)\ [\ x(k) - \underline{H}\ \hat{\underline{s}}\ (k/k-1)\ ]$
state prediction	$\hat{\underline{s}}\ (k/k-1) = \underline{a}\hat{\underline{s}}\ (k-1/k-1)$
Kalman gain	$\underline{K}(k) = \underline{V}(k/k-1)\underline{H}^T\ [\ \underline{HV}(k/k-1)\underline{H}^T + \hat{\sigma}_n^2\ ]^{-1}$
Covariance prediction	$\underline{V}(k/k-1) = \underline{aV}(k-1/k-1)\underline{a}^T + \underline{bb}^T\sigma_s^2$
Covariance estimate	$\underline{V}(k/k) = [\ I - \underline{K}(k)\underline{H}\ ]\underline{V}(k/k-1)$
Tap estimate	$\hat{\underline{h}}\ (k+1) = \hat{\underline{h}}\ (k) + 2\mu\ \hat{\underline{s}}\ (k+1)e(k+1)$
Channel estimation error	$e(k+1) = x(k+1) - \hat{\underline{h}}^{\ T}\ (k)\ \hat{\underline{s}}\ (k+1)$
observation noise estimate	$\hat{\sigma}_n^2\ (k+1) = (1 - 1/M\ )\ \hat{\sigma}_n^2\ (k) + e^2\ (k+1)/M$

The Kalman gain vector is as below,

$$\underline{K}^T\ (k) = [\ K_0\ (k)\ K_1\ (k)\ .....\ K_d(k)\ ].$$

with the estimated tap weight vector $\hat{\underline{h}}\ (k)$ being used in the d element vector $\underline{H}$ as below,

$$\underline{H} = [\ \hat{h}_0\ \hat{h}_1 .....\ \hat{h}_{M-1}\ 0\ 0\ \cdots\ 0\ ],$$

Table 5.2 - Summary of Equations for an Adaptive Kalman Equaliser

again the starting point. The channel is considered as a FIR filter which can be represented in the state space form as shown below,

$$\underline{s}(k+1) = F(k+1/k)\underline{s}(k) + n(k). \tag{5.12}$$

where $F(k+1/k)$ represents the state space transition matrix associated with the channel and $\underline{s}(k)$ is the input data vector. The following observation equations may then be associated with this model,

$$x_1(k) = \underline{h}^T\underline{s}(k) + n(k) \tag{5.13a}$$

and,

$$x_2(k) = \underline{s}(k-d). \tag{5.13b}$$

The first equation represents the channel output and the second represents a decision feedback term. It is worth noting at this point that it would be possible to have many similar observation equations to the second one, that is feedback many decisions. However preliminary investigations[125], suggested that nothing would be gained in terms of performance for a considerable increase in complexity. The two observation equations may now be combined as below,

$$\underline{x}(k) = \underline{H}\underline{s}(k) + \underline{N}(k). \tag{5.14}$$

$\underline{H}$ being a $(d+1) \times 2$ matrix as shown below,

$$\begin{bmatrix} h_0 & h_1 & \cdots & h_{M-1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

It is clear that the problem presented here is identical to that of the previous section, the difference lying in the observation equations, it is necessary then to generate the Kalman filter which provides the solution given these particular observation equations. With the definitions above, this may be done directly and the equations are detailed in Table 5.3. It should now be noted that the inversion of the innovations term,

$$[\,\underline{H}\underline{V}(k/k-1)\underline{H}^T + \underline{\sigma}_n^2\,]^{-1}, \tag{5.15}$$

is no longer a scalar and consequently the inversion of a two by two matrix is required. The LMS channel estimator is used in exactly the same way as in the previous structure

to provide the tap weight vector estimate and a measure of the observation noise. The detailed structure of the equaliser is shown in Figure 5.8.

The final question which remains is to determine how many states are required for the equaliser to operate efficiently, and clearly $d+1$, that is the estimation with lag delay is d and there is one feedback term.

Operation	Equation
state estimate	$\hat{\underline{s}}\ (k/k) = \hat{\underline{s}}\ (k/k-1) + \underline{K}(k)\ [\ x(k) - \underline{H}\ \hat{\underline{s}}\ (k/k-1)\ ]$
state prediction	$\hat{\underline{s}}\ (k/k-1) = \underline{a}\ \hat{\underline{s}}\ (k-1/k-1)$
Kalman gain	$\underline{K}(k) = \underline{V}(k/k-1)\underline{H}^T\ [\ \underline{H}\underline{V}(k/k-1)\underline{H}^T + \underline{\sigma}_n^2\ ]^{-1}$
Covariance prediction	$\underline{V}(k/k-1) = \underline{a}\ \underline{V}(k-1/k-1)\ \underline{a}^T + \underline{b}\underline{b}^T\sigma_s^2$
Covariance estimate	$\underline{V}(k/k) = [\ I - \underline{K}(k)\underline{H}\ ]\underline{V}(k/k-1)$
Tap estimate	$\hat{\underline{h}}\ (k+1) = \hat{\underline{h}}\ (k) + 2\mu\ \hat{\underline{s}}\ (k+1)e(k+1)$
Channel estimation error	$e(k+1) = x(k+1) - \hat{\underline{h}}^T\ (k)\ \hat{\underline{s}}\ (k+1)$
observation noise estimate	$\hat{\sigma}_n^2\ (k+1) = (1 - 1/M\ )\ \hat{\sigma}_n^2\ (k) + e^2\ (k+1)/M$

where the Kalman gain vector is as below,

$$\begin{bmatrix} K_{0,0} & K_{0,1} & . & . & . & K_{0,d+1} \\ K_{1,0} & K_{1,1} & . & . & . & K_{1,d+1} \end{bmatrix}$$

The estimated tap weight vector $\hat{\underline{h}}\ (k)$ being used in the d element vector $\underline{H}$ as below,

$$\begin{bmatrix} h_0 & h_1 & . & . & . & h_{M-1} & 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 & 0 & 0 & . & . & . & 1 \end{bmatrix}$$

Table 5.3 - Summary of Adaptive Kalman DFE Equations

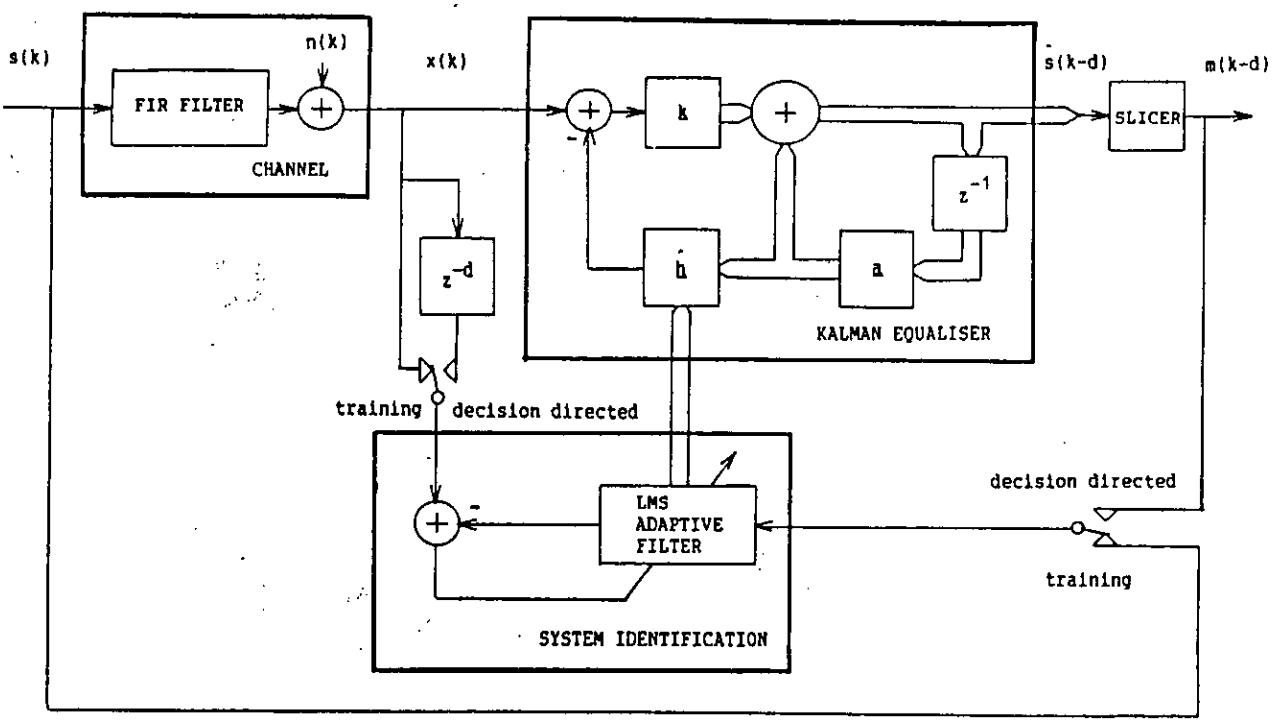


Figure 5.7 - Adaptive Kalman equaliser structure

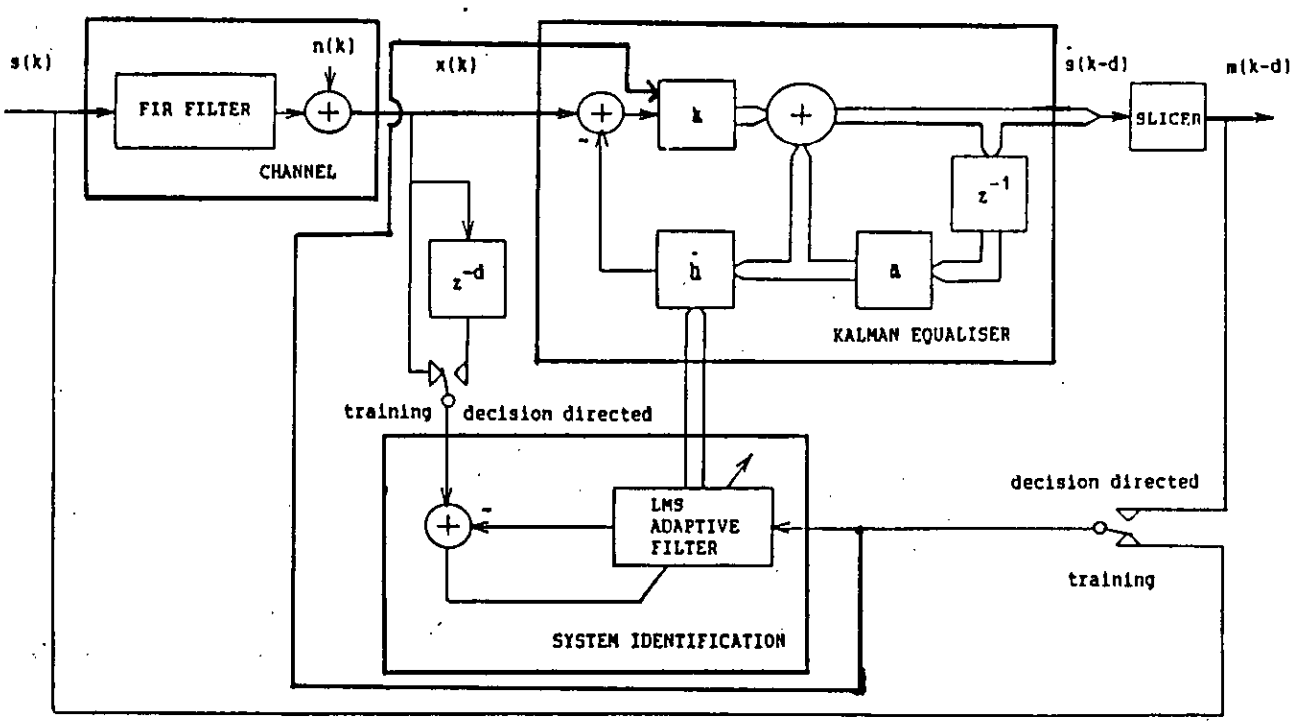


Figure 5.8 - Adaptive Kalman DFE structure

5.6 ALGORITHM COMPLEXITY CONSIDERATIONS

A breakdown of the computation required to process each of the three algorithms considered here is presented in this section. Tables 5.4 and 5.5 present the computation required for each process in the algorithm given particular values of the lag d and the number of taps in the channel.

This results in the adaptive Kalman DFE requiring 123 multiplications and 99 additions/subtractions per iteration to carry out the tests performed in this report. The adaptive Kalman equaliser required 47 multiplications and 37 additions/subtractions and is clearly less complex, the conventional DFE using the Godard Kalman required 133 operations per iteration which is comparable with the adaptive Kalman equaliser. It is also worth remembering that more computationally efficient implementations are possible by utilising the standard matrix algebra techniques as has been demonstrated in[75, 76, 78].

operation	mult.	add/sub.
$\underline{V}(k/k-1)\,\underline{H}(k)$	$d(M-1)+1$	$d(M-2)$
$\underline{H}^T(k)\,\underline{V}(k/k-1)\,\underline{H}(k) + \sigma_n^2$	M	M
$x(k) - \underline{H}^T(k)\,\hat{\underline{s}}(k/k-1)$	$M-1$	$M-1$
$\underline{V}(k/k-1)\,\underline{H}(k)\,[\,\underline{H}^T(k)\,\underline{V}(k/k-1)\,\underline{H}(k) + \sigma_n^2\,]^{-1}$	$d+1$	
$\hat{\underline{s}}(k/k-1) + \underline{K}(k)\,[\,x(k) - \underline{H}^T(k)\,\hat{\underline{s}}(k/k-1)\,]$	$d+1$	d
$\underline{K}(k)\,\underline{H}^T(k)\,V(k/k-1)$	$\frac{d^2}{2}+\frac{d}{2}$	
$\underline{V}(k/k-1) - \underline{K}(k)\,\underline{H}^T(k)\,\underline{V}(k/k-1)$		$\frac{d^2}{2}+\frac{d}{2}$
$y(k) - \underline{h}^T(k-1)\,\underline{x}(k)$	M	M
$2\,\mu\,\underline{x}(k)\,(\,y(k) - \underline{h}^T(k-1)\,\underline{x}(k)\,)$	$M+1$	
$\underline{h}(k-1) + 2\,\mu\,\underline{x}(k)\,(\,y(k) - \underline{h}^T(k-1)\,\underline{x}(k)\,)$		M
$(\,1 - \frac{1}{M}\,)\,\hat{\sigma}_e^2(k-1) + \frac{e^2(k)}{M}$	3	1

no. of states = d+1

no of channel taps = M

Table 5.4 - Adaptive Kalman Equaliser Complexity

operation	mult.	add/sub.
$\underline{V}(k/k-1)\,\underline{H}(k)$	$2(d+1)M$	$2(d+1)(M-2)$
$\underline{H}^T(k)\,\underline{V}(k/k-1)\,\underline{H}(k) + \sigma_n^2$	$4M$	$4M-4$
$x(k) - \underline{H}^T(k)\,\hat{\underline{s}}(k/k-1)$	$2(M-1)+2(d+1)$	$2M-2$
$\underline{V}(k/k-1)\,\underline{H}(k)\,[\,\underline{H}^T(k)\,\underline{V}(k/k-1)\,\underline{H}(k) + \sigma_n^2\,]^{-1}$	$4(d+1)+4$	$2(d+1)+1$
$\hat{\underline{s}}(k/k-1) + \underline{K}(k)\,[\,x(k) - \underline{H}^T(k)\,\hat{\underline{s}}(k/k-1)\,]$	$2(d+1)$	$2(d+1)$
$\underline{K}(k)\,\underline{H}^T(k)\,V(k/k-1)$	$\frac{2(d+1)M+(d+1)^2}{2}$	$\frac{2(d+2)(M-2)+(d+1)^2}{2}$
$\underline{V}(k/k-1) - \underline{K}(k)\,\underline{H}^T(k)\,\underline{V}(k/k-1)$		$\frac{(d+1)^2}{2}$
$y(k) - \underline{h}^T(k-1)\,\underline{x}(k)$	M	M
$2\,\mu\,\underline{x}(k)\,(\,y(k) - \underline{h}^T(k-1)\,\underline{x}(k)\,)$	$M+1$	
$\underline{h}(k-1) + 2\,\mu\,\underline{x}(k)\,(\,y(k) - \underline{h}^T(k-1)\,\underline{x}(k)\,)$		M
$(\,1 - \frac{1}{M}\,)\,\hat{\sigma}_e^2(k-1) + \frac{e^2(k)}{M}$	3	1

no. of states = d+2

no. of channel taps = M

Table 5.5 - Adaptive Kalman DFE Complexity

5.7 PERFORMANCE RESULTS

The results presented in this section detail the performance of the three equaliser structures under a variety of channel scenarios. Initially simulations are presented for a stationary channel situation. These tests were performed to determine if the inherent performance advantage, in terms of the steady state MSE achieved, of decision feedback structures was actually achieved. Subsequently tests were performed on the channel simulator discussed in chapter 2 to determine the level of performance achieved under the time varying conditions of the HF channel.

Channel No.	Impulse Response	Classification
1	$0.6082 + 0.7603\, z^{-1} + 0.228\, z^{-2}$	minimum phase
2	$0.2602 + 0.9298\, z^{-1} + 0.2602\, z^{-2}$	non-minimum phase

Table 5.6 - CHANNEL IMPULSE RESPONSES

All of the simulations were performed on a Sun 3/60 workstation using the 'C' computer language. The two stationary channels used in these simulations are detailed in table 5.6 above. In the HF channel simulations the Doppler spread was 1Hz, as indicated on the plots. In all cases an additive noise power level of -50dB was adopted. The tests were performed for 100% and 50% periodic retraining on the stationary channels and 50% for the HF channel. The 100% retraining being used to provide a reference. When periodic retraining is 50% only 50 out of every 100 symbols is known a-priori, the rest of the time the equaliser is operating in decision directed mode.

It is clear from the results on the stationary channels, illustrated in Figures 5.9-5.18 and summarised below in table 5.7, that the best performance, in terms of final MSE, is offered by the the decision feedback structures, This is particularly noticeable on the non-minimum phase channel where an improvement of approximately 10dB is available, this is as expected.

Figure	Equaliser	Channel	% training	Steady-state MSE
5.9a	Godard-Kalman DFE	1	100%	-40dB
5.9b	Godard-Kalman DFE	2	100%	-30dB
5.10a	Godard-Kalman DFE	1	50%	-40dB
5.10b	Godard Kalman DFE	2	50%	-30dB
5.11a	Adaptive Kalman	1	100%	-35dB
5.12a	Adaptive Kalman	1	50%	-35dB
5.13a	Adaptive Kalman	2	100%	-22dB
5.14a	Adaptive Kalman	2	50%	-22dB
5.15a	Adaptive Kalman DFE	1	100%	-38dB
5.16a	Adaptive Kalman DFE	1	50%	-38dB
5.17a	Adaptive Kalman DFE	2	100%	-30dB
5.18a	Adaptive Kalman DFE	2	50%	-30dB

Table 5.7 -Steady-state MSE of equalisers on stationary channels

In the HF scenario, as Figures 5.19-5.21 illustrate, the results are markedly different from the stationary case. Table 5.8 below summarises the performance of the various equalisers for these simulations.

Figure	Equaliser	Fade rate	% training	Steady-state MSE
5.19	Godard-Kalman DFE	1Hz	50%	-15dB (*)
5.20a	Adaptive Kalman	1Hz	50%	-23dB
5.21a	Adaptive Kalman DFE	1Hz	50%	-23dB

* - algorithm diverges due to numerical instability.

Table 5.8 -Steady-state MSE of equalisers on HF channel

It is clear that both adaptive Kalman structures offer a lower final MSE than that achieved by the conventional DFE by some 5-10dB. This rather suprising result may be explained in several ways. The inherent robustness of channel estimator to decision errors compared with channel equalisation, the separation of the channel and sequence estimation processes. In addition both of the adaptive Kalman structures offer greater freedom in selection of the number of taps. That is for a conventional equaliser if the length is increased beyond a certain point, for a fixed channel length, the increased algorithm noise this causes outweighs any advantage in improved performance the additional taps offer. The numerical sensitivity of techniques such as the Godard Kalman algorithms is also clearly illustrated in Figure 5.19.

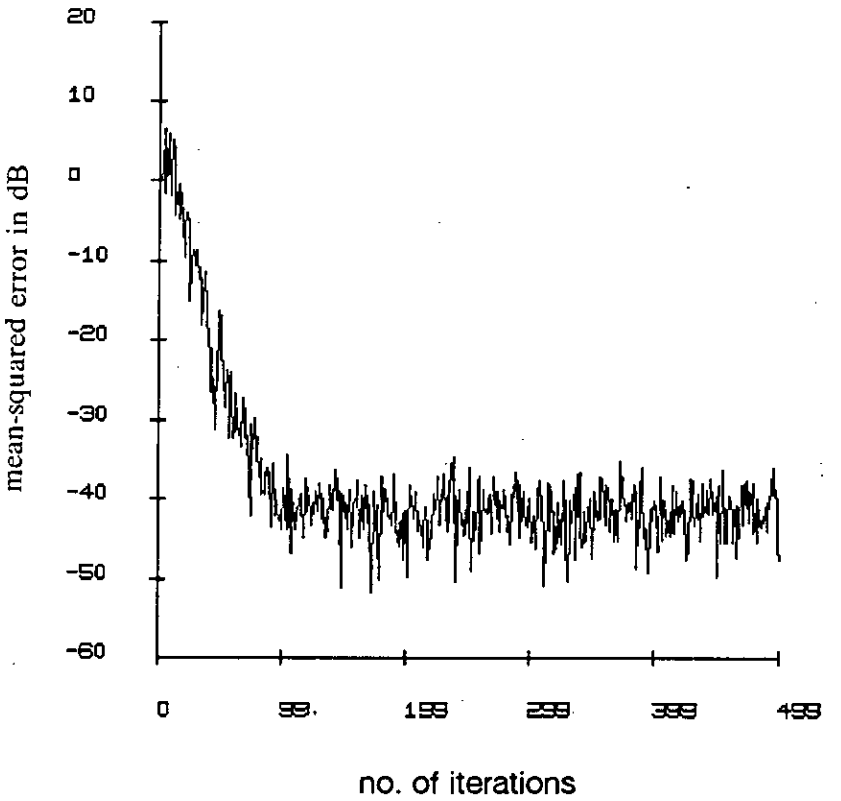


Fig. 5.9a - MSE performance of Godard-Kalman DFE on channel 1, additive noise power=-50dB.
(5 feedforward taps, 2 feedback taps, lag=4,100% training).

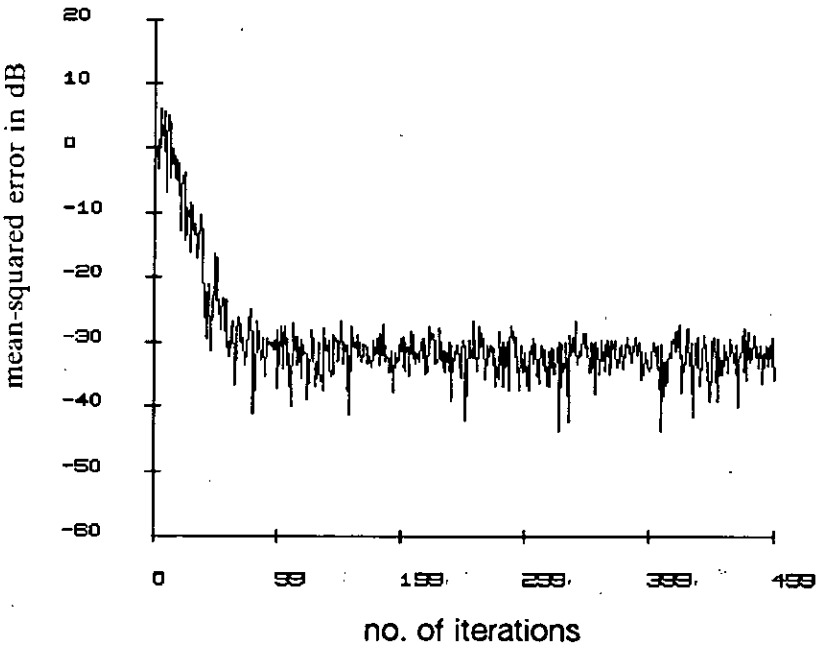


Fig. 5.9b - MSE performance of Godard-Kalman DFE on channel 2, additive noise power=-50dB.
(5 feedforward taps, 2 feedback taps, lag=4, 100% training).

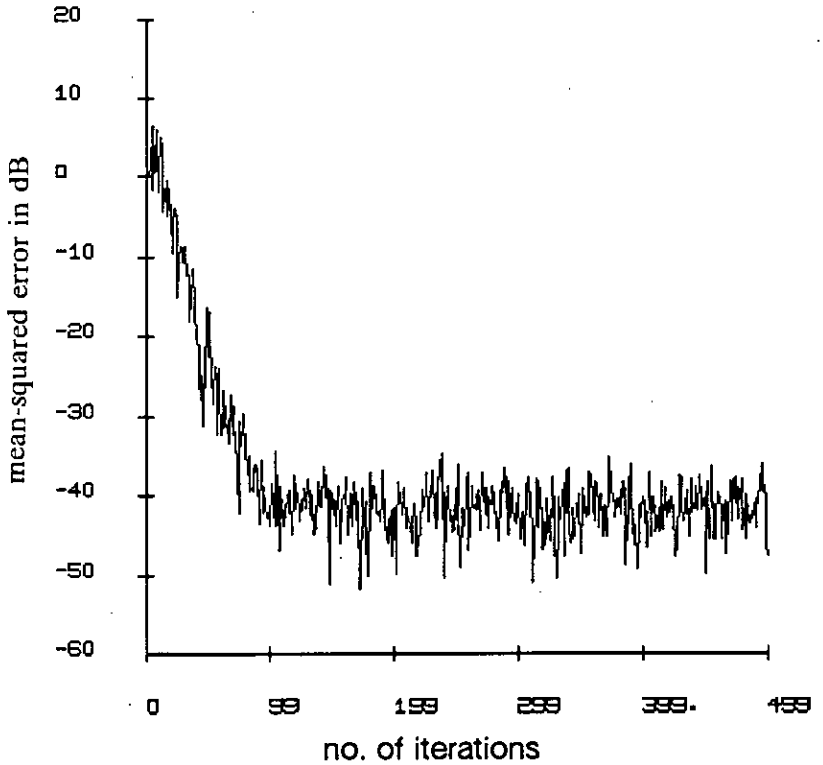


Fig 5.10a - MSE performance of Godard-Kalman DFE on channel 1, additive noise power=-50dB. (5 feedforward taps, 2 feedback taps, lag=4, 50% training)

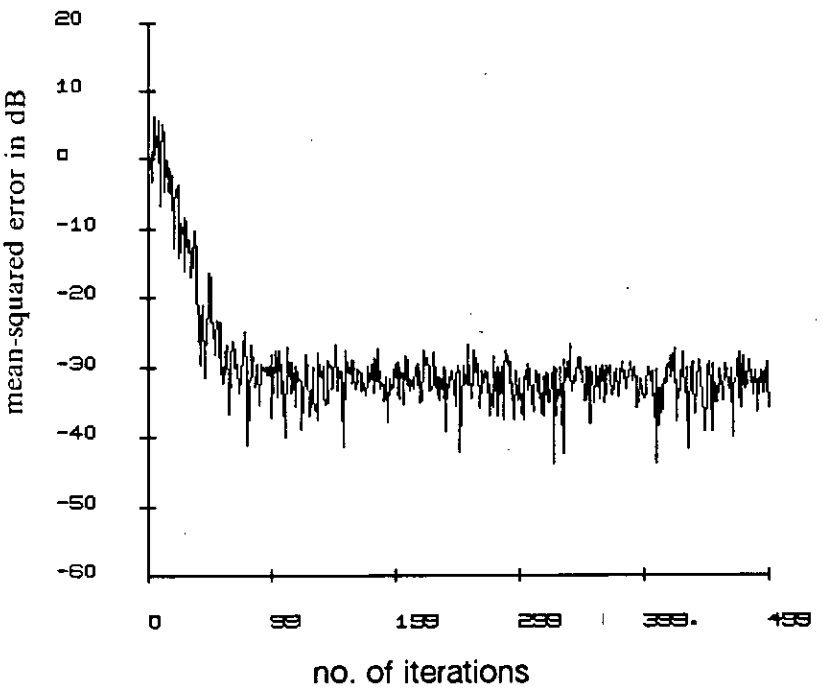


Fig 5.10b - MSE performance of Godard-Kalman DFE on channel 2, additive noise power=-50dB. (5 feedforward taps, 2 feedback taps, lag=4, 50% training)

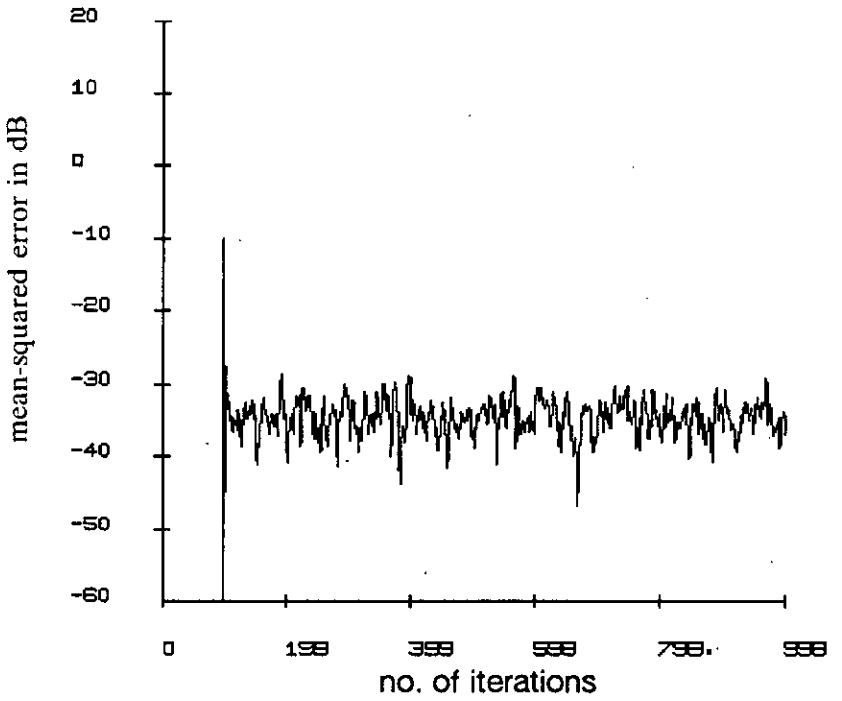


Fig 5.11a - MSE performance of Adaptive Kalman equaliser on channel 1,additive noise power=-50dB, (6 states, 100% training).

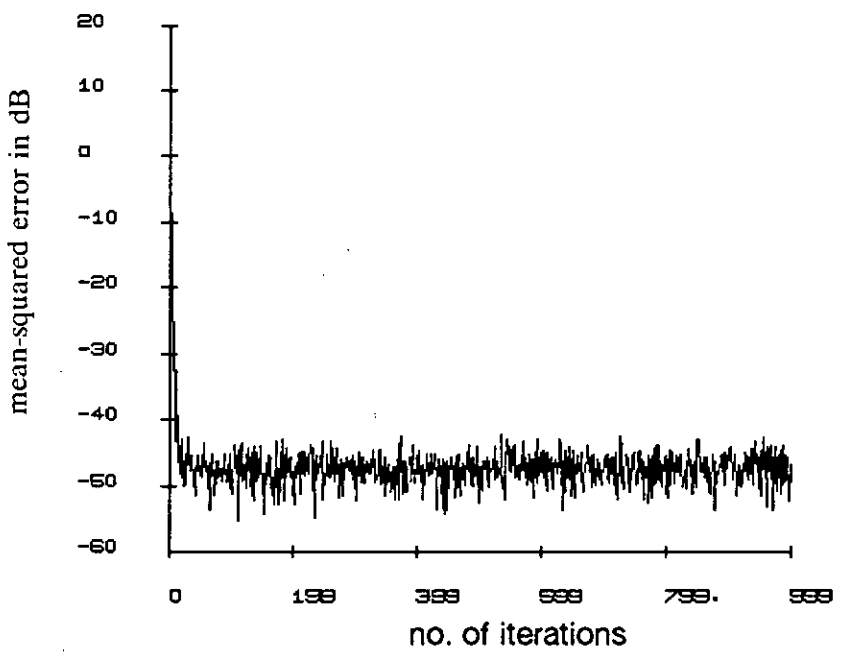


Fig 5.11b - MSE performance of LMS channel estimation for Adaptive Kalman equaliser on channel 1, additive noise power=-50dB, (3 taps, 100% training).

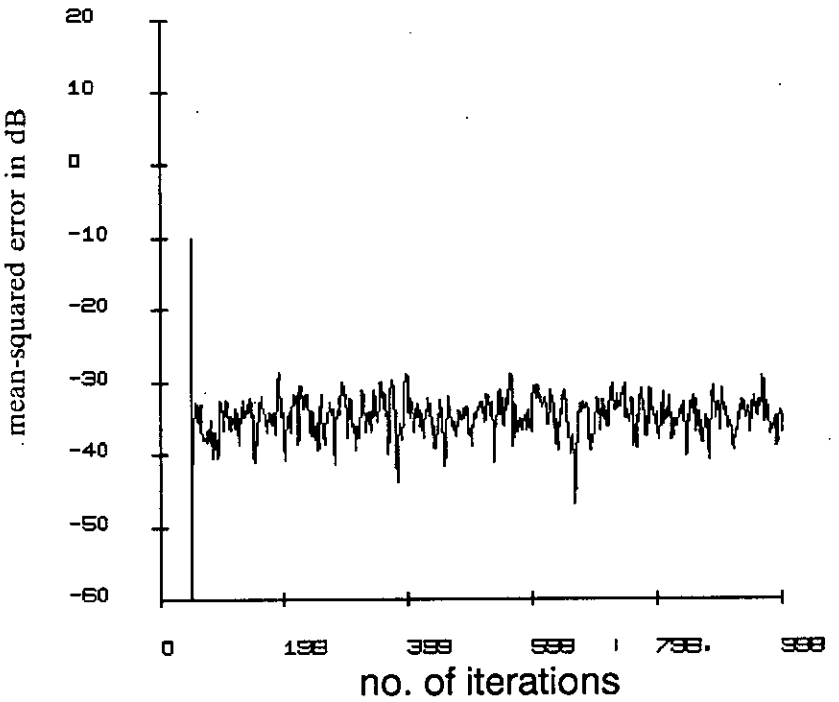


Fig 5.12a - MSE performance of Adaptive Kalman equaliser on channel 1, additive noise power= -50dB, (6 states, 50% training).

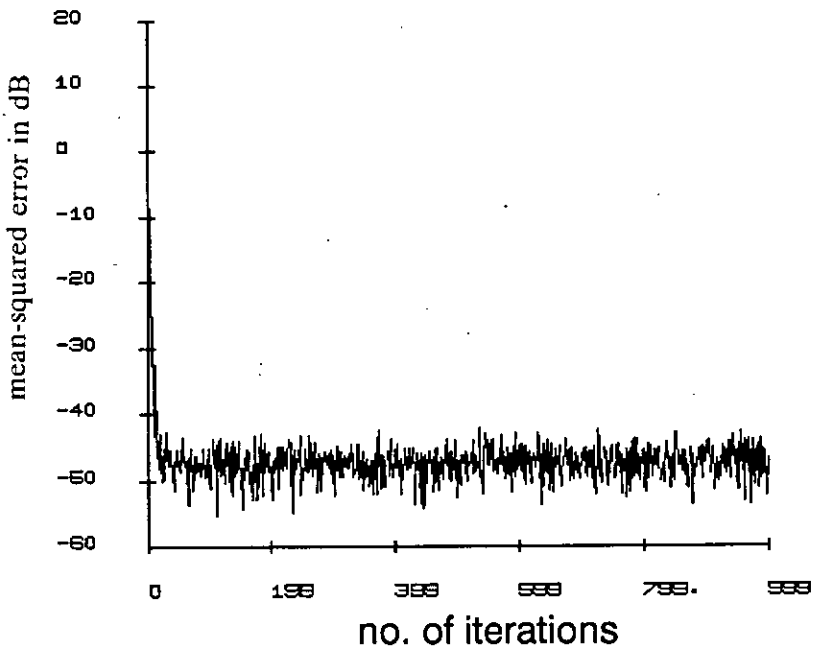


Fig 5.12b - MSE performance of LMS channel estimator for Adaptive Kalman equaliser on channel 1, additive noise power=-50dB, (3 taps, 50% training).

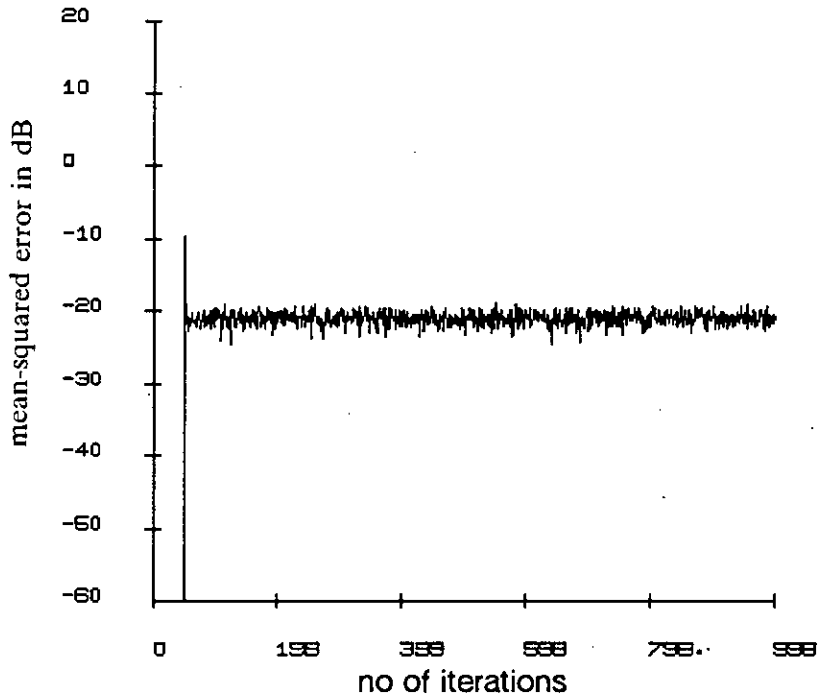


Fig 5.13a - MSE performance of Adaptive Kalman equaliser on channel 2, additive noise power= -50dB, (6 states, 100% training).

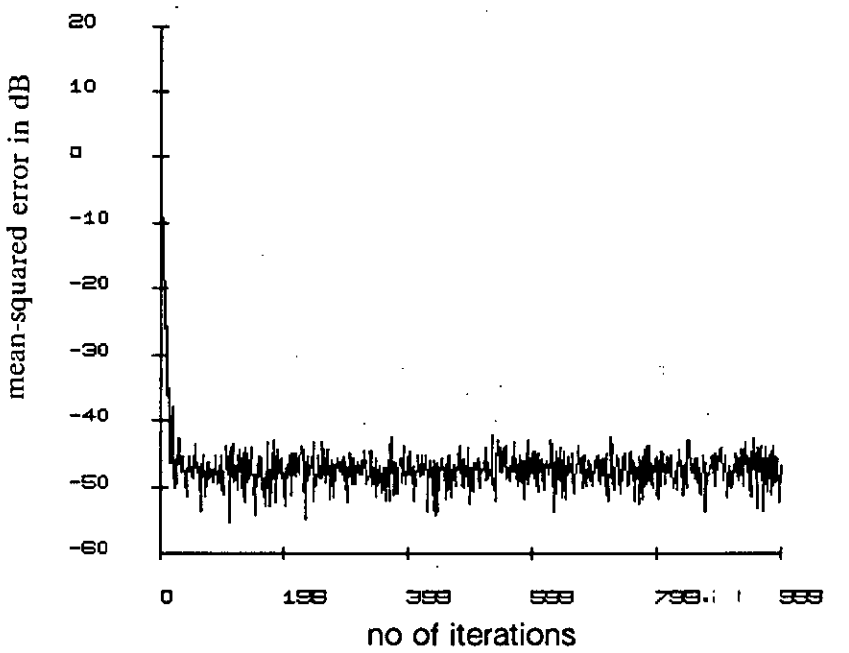


Fig 5.13b - MSE performance of LMS channel estimator for Adaptive Kalman equaliser on channel 2, additive noise power=-50dB, (3 taps, 100% training).

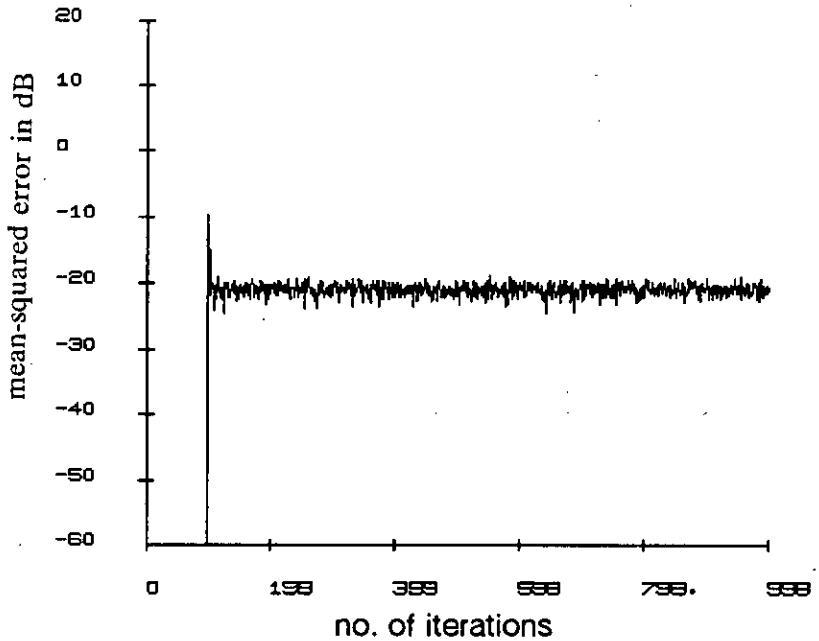


Fig. 5.14a – MSE performance of Adaptive Kalman equaliser on channel 2, additive noise power=–50dB, (6 states, 50% training).

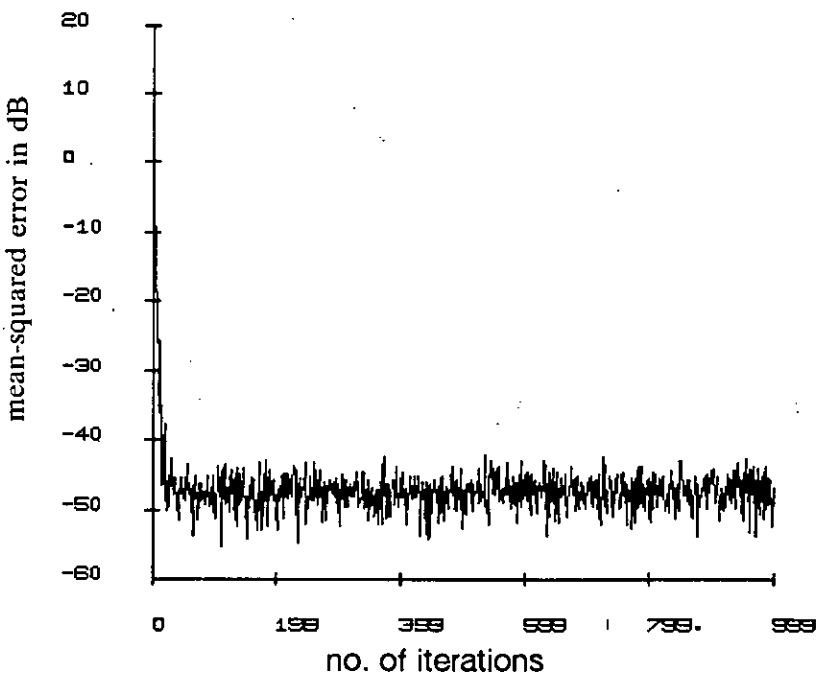


Fig 5.14b – MSE performance of LMS channel estimator for Adaptive Kalman equaliser on channel 2, additive noise power=–50dB, (3 taps, 50% training).

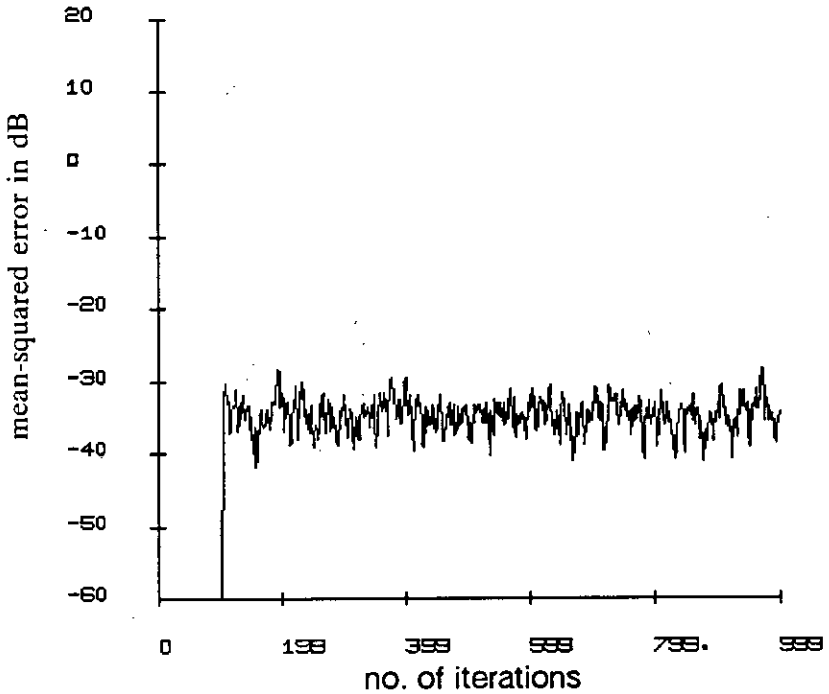


Fig 5.15a - MSE performance of Adaptive Kalman DFE on channel 1, additive noise power= -50dB, (7 states, lag=6, 100% training).

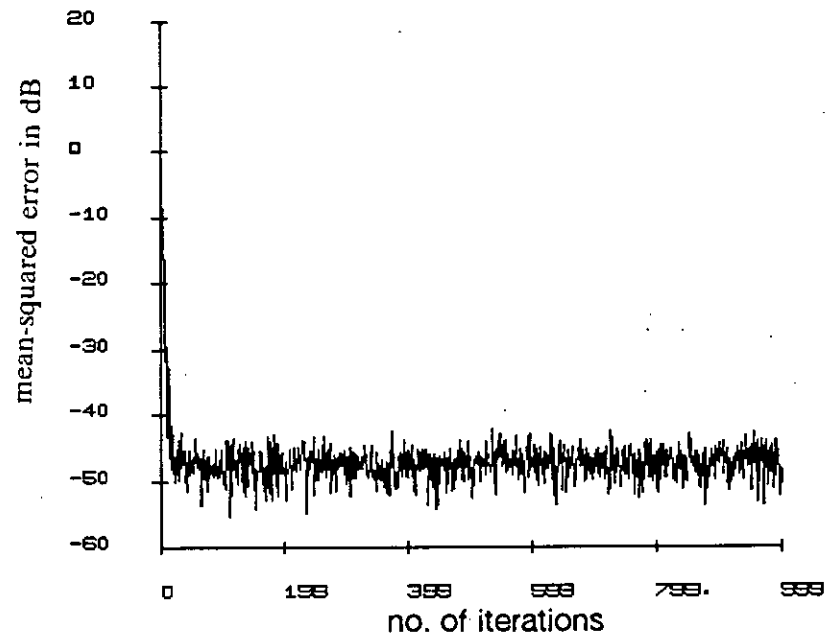


Fig. 5.15b - MSE performance of LMS channel estimator for Adaptive Kalman DFE on channel 1, additive noise power=-50dB, (3 taps, 100% training).

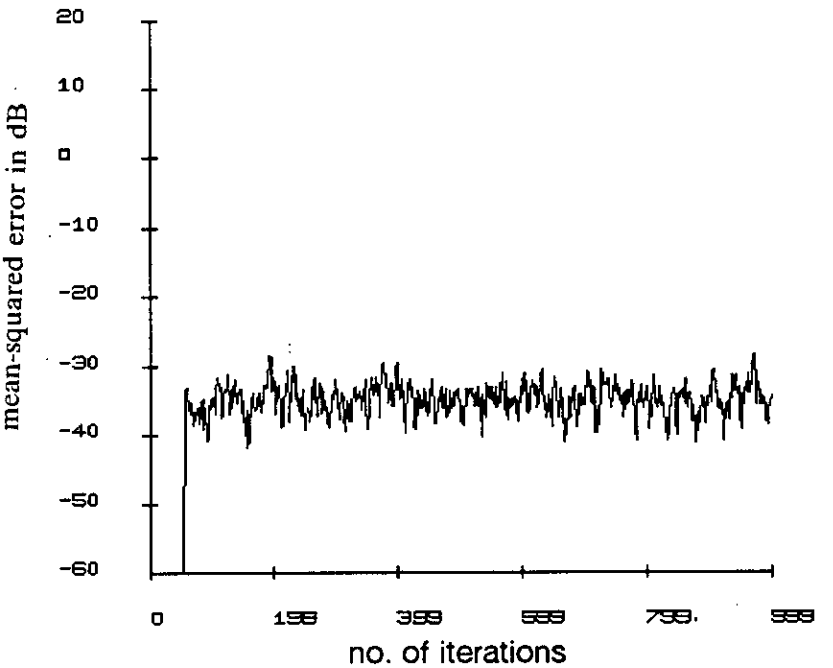


Fig 5.16a - MSE performance of Adaptive Kalman DFE on channel 1, additive noise power= -50dB, (7 states, lag=6, 50% training).

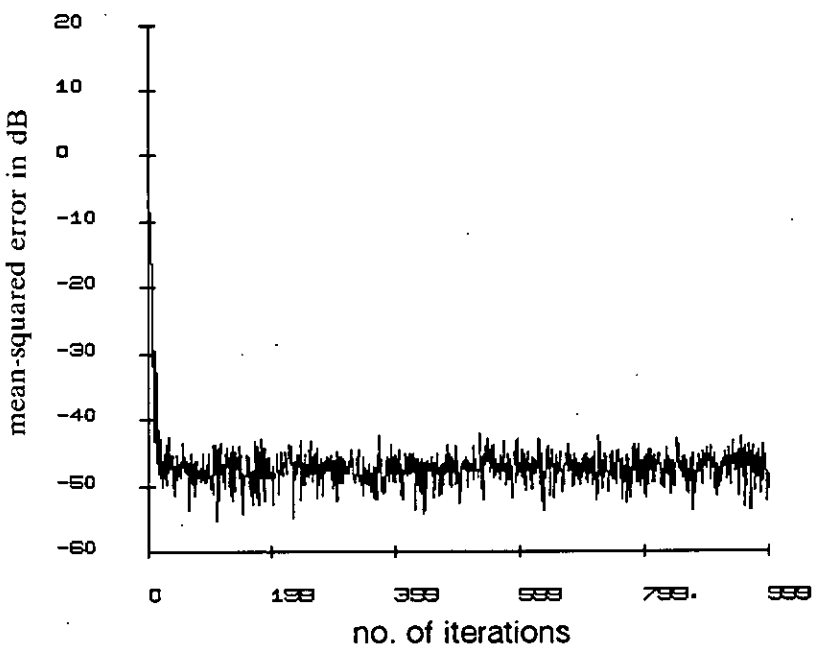


Fig. 5.16b - MSE performance of LMS channel estimator for Adaptive Kalman DFE on channel 1, additive noise power=-50dB, (3 taps, 50% training).

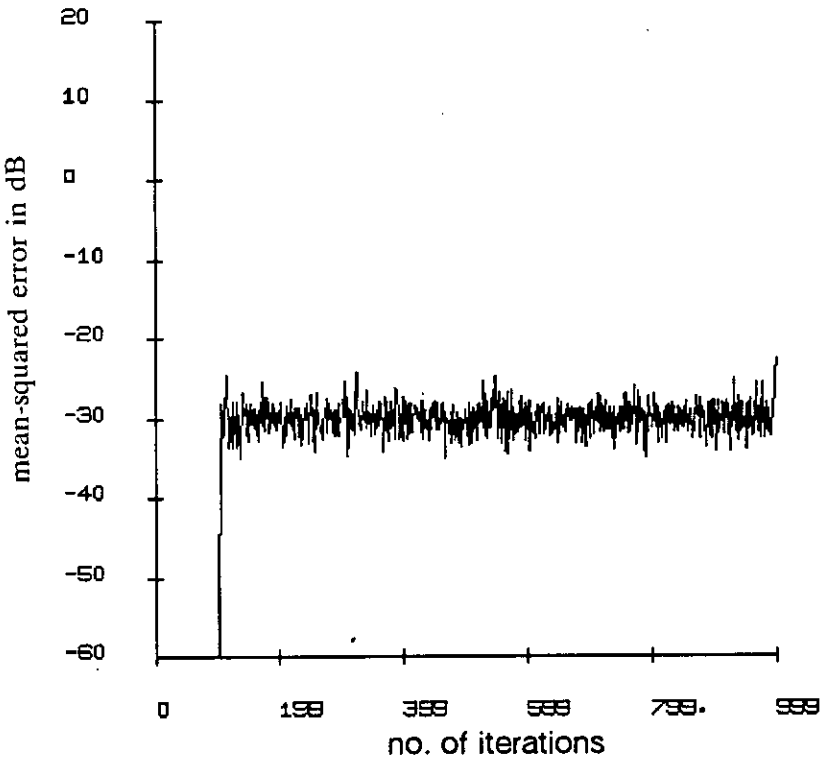


Fig . 5.17a - MSE performance of Adaptive Kalman DFE on channel 2, additive noise power=-50dB, 7 states, lag=6, 100% training)

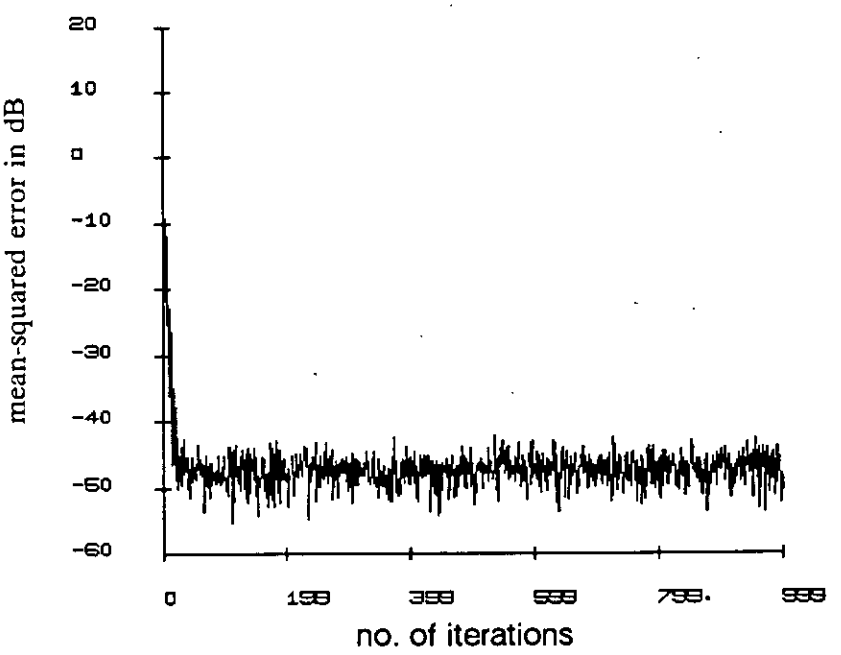


Fig 5.17b -MSE performance of LMS channel estimator for Adaptive Kalman DFE on channel 2, additive noise power=-50dB, (3 taps, 100% training).

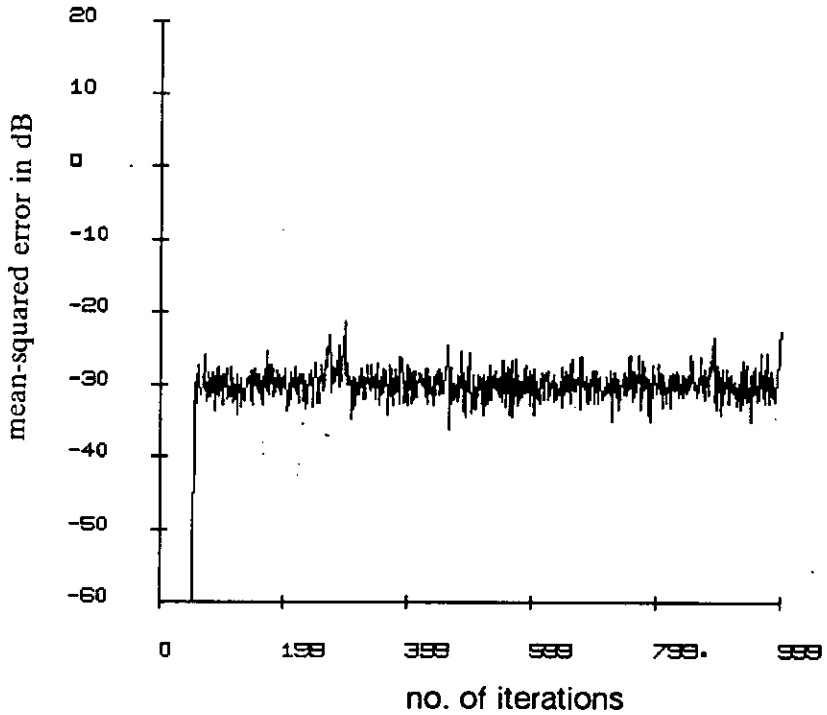


Fig. 5.18a – MSE performance of Adaptive Kalman DFE on channel 2, additive noise power=–50dB, (7 states, lag=6, 50% training).

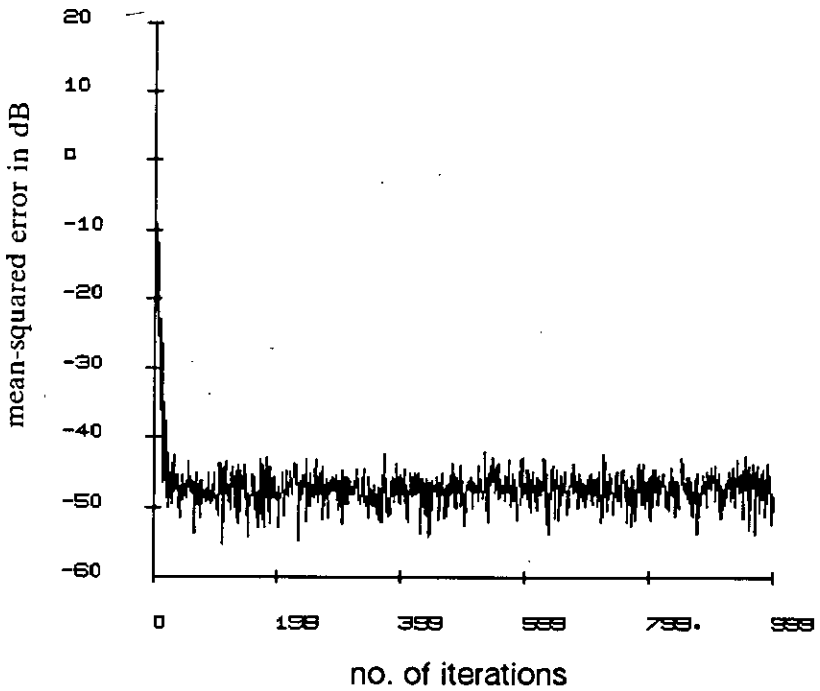


Fig. 5.18b – MSE performance of LMS channel estimator for Adaptive Kalman DFE on channel 2, additive noise power=–50dB, (3 taps, 50% training).

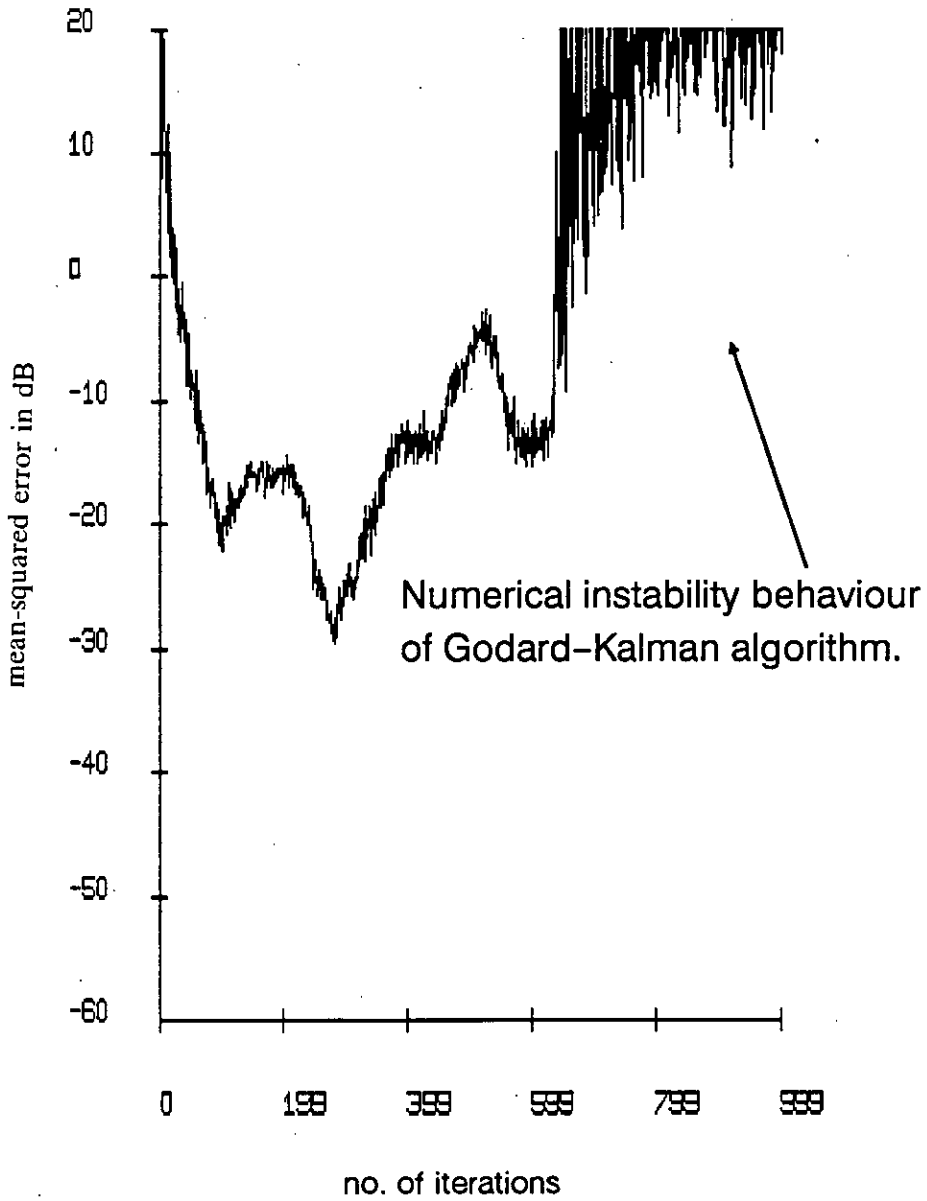


Fig. 5.19 - MSE performance of Godard-Kalman DFE on HF channel, additive noise power= -50dB, (5 feedforward taps, 2 feedback taps, lag=4, fade rate=1Hz, 50% training).

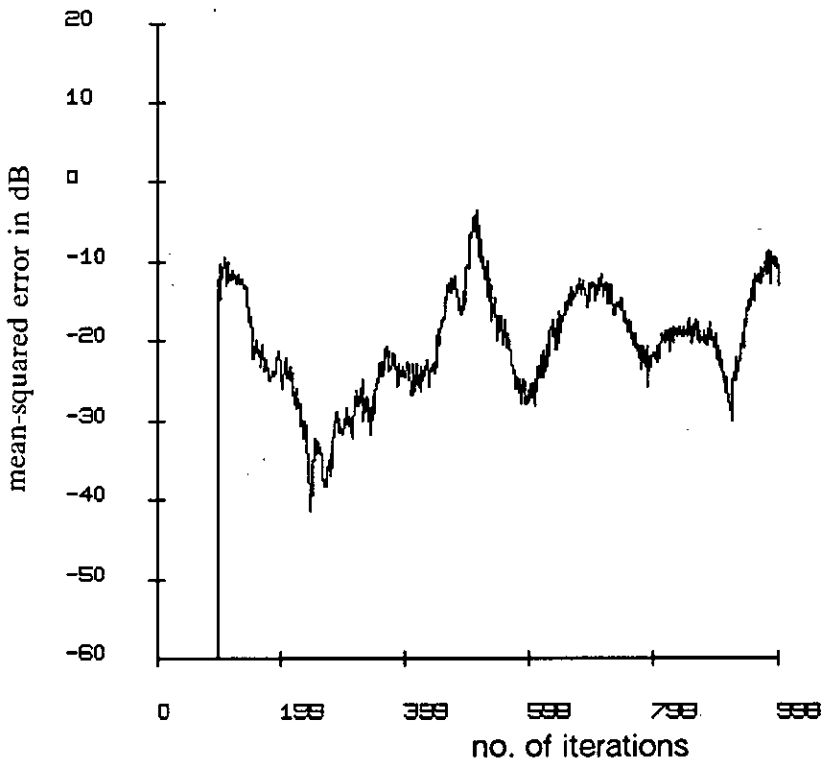


Fig. 5.20a - MSE performance of Adaptive Kalman equaliser on HF channel, additive noise power= -50dB, (6 states, lag=5, fade rate=1Hz, 50% training).

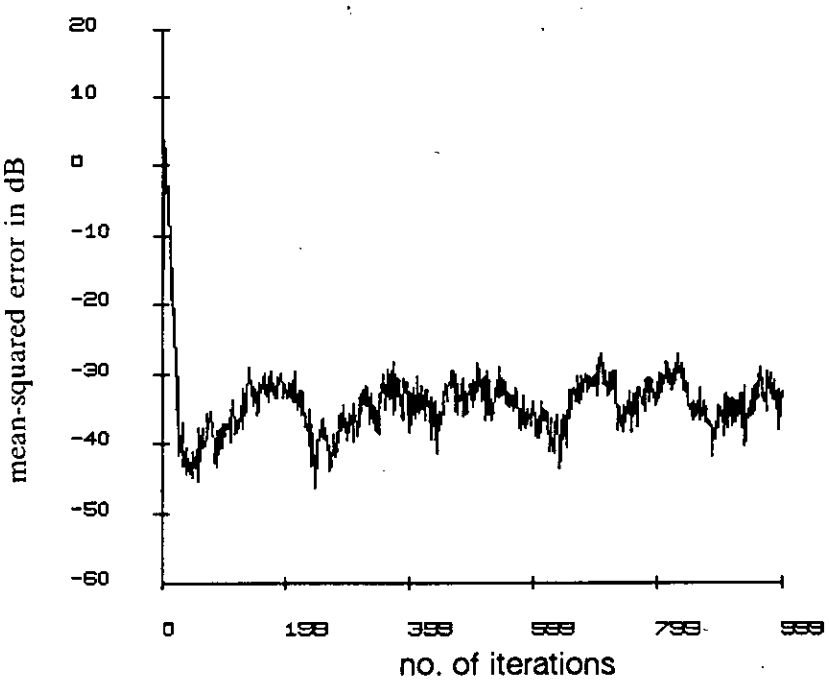


Fig. 5.20b - MSE performance of LMS channel estimator for Adaptive Kalman equaliser on HF channel, additive noise power= -50dB, (3 taps, 50% training).

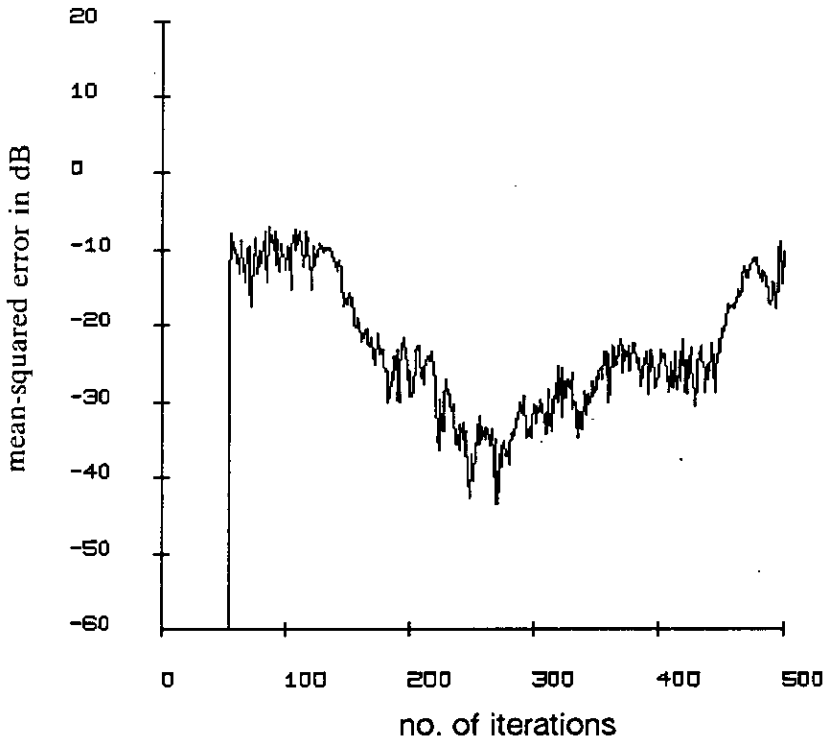


Fig. 5.21a – MSE performance of Adaptive Kalman DFE on HF channel, additive noise power = –50dB, (7 states, lag=6, fade rate=1Hz, 50% training).

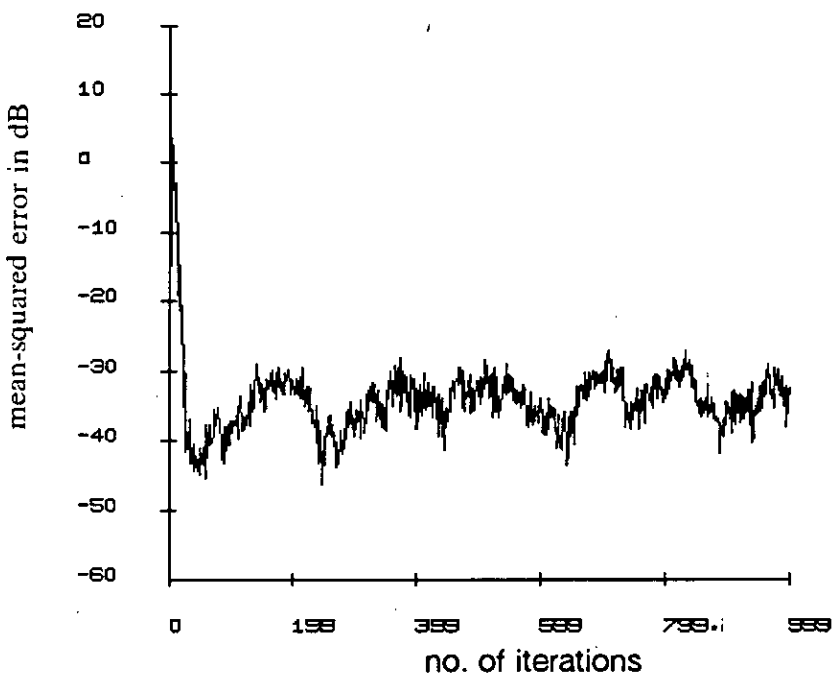


Fig. 5.21b MSE performance of LMS channel estimator for Adaptive Kalman DFE on HF channel, additive noise power = –50dB, (3 taps, 50% training).

5.8 CONCLUSIONS

It is clear from the results presented here that the adaptive Kalman IIR offers the best performance in the HF channel of the three structures tested here. This is based on the final MSE achieved and the computational complexity of the algorithm. The adaptive Kalman DFE would appear to offer a better MSE performance in several channel scenarios. However, it should be noted that it is likely to suffer a degradation in performance due to error propagation as in the conventional DFE.

It is worth noting that if a more accurate channel estimator, of reasonable level of complexity, could be developed then an improvement in performance could be expected from the adaptive Kalman based structures.

Chapter 6

CONCLUSIONS

In this thesis the problem of designing both adaptive algorithms and structures for use as channel estimators and equalisers within the HF communications scenario has been addressed. In this chapter the main conlusions of this work are highlighted with suggestions of possible further research in the area.

In chapter 3, a study of the performance of two existing adaptive algorithms, the LMS and RLS, as HF channel estimators was carried out. The work determined that, contrary to popular opinion, the more complex RLS algorithm offered no performance advantage over the computationally simpler LMS algorithm. A new theoretical expression was derived which allowed the steady state MSE performance of the RLS algorithm to be predicted, given prior knowledge of the levels of both noise and time-variations which would be encountered in the system.

Although the performance of the LMS as a channel estimator was as good if not better than the RLS, its lack of spectral robustness and relatively slow convergence make it less than ideal for application in this environment. Consequently in chapter 4 three new adaptive algorithms were derived for specific use as HF channel estimators. The key to the development of each of the algorithms is in their particular use of a-priori knowledge of the channel structure. Each of the algorithms uses, to a greater or lesser degree, some a-priori knowledge of the state space representation of both the channel structure and the tap generation process. As was described in chapter 2, it is widely accepted that the HF channel is accurately modeled by a tapped delay line structure with time varying taps which are generated by filtering a zero-mean Gaussian sequence. This is readily represented in a state-space form and thus lends itself to the use of Kalman filters.

The first algorithm, the MVK, requires full a-priori knowledge of the channel, (and is consequently not implementable), in the form of both the noise covariance and

state-space transition matrices being incorporated into a standard Kalman filter. This algorithm has optimal performance and provided the motivation for the derivation of the second new algorithm, the EKF channel estimator. In this algorithm, only partial a-priori knowledge of the channel was required in the form of the structure of the channel model. In addition, an LMS was used to *bootstrap* the algorithm on initialisation by providing estimates of the required variables. Although the performance of the EKF is near optimal its computational complexity and numerical sensitivity preclude implementation in the present form.

The results from the first two algorithms provided the impetus for the third novel algorithm. This time, the a-priori knowledge of the structure of the channel model was used to provide a series of prediction filters, configured for various parameters, in conjunction with an LMS algorithm. This algorithm operated by essentially increasing the order of the recursion within the LMS algorithm from first to second order. The performance of the algorithm, although disappointing, still offers room for improvement by selecting the prediction filter on different criteria.

In chapter 5, the problem of channel equalisation was addressed and three equaliser structures, one novel, were considered. The structures were a conventional DFE utilising a Godard-Kalman algorithm for the tap adjustment, an adaptive Kalman structure which utilised an LMS channel estimator and a new adaptive Kalman DFE structure which also used an LMS channel estimator. The performance of each of the structures was studied under a variety of conditions, the MSE criteria being used to assess their performance. It was determined that the structures where the channel and sequence estimation processes were separate offered an enhanced performance.

Although this thesis has considered both channel estimation and equalisation for HF communication systems it is far from a complete study of the application of adaptive techniques to this area. In particular there are specific areas which have not been addressed fully or only touched upon in passing. These particular points provide suggestions for possible areas of further research and are summarised below.

As has been alluded to on several occasions within the thesis, there are several techniques by which the complexity and numerical sensitivity of adaptive algorithms may be reduced. It is possible that, if these techniques were applied, that the EKF algorithm would become more attractive for implementation. The rapid growth in VLSI, as demonstrated by the increasing gate count, has now made it possible for implementation of the more complex algorithms, such as the EKF, to be considered. In fact a description of the implementation of a square root covariance form of the EKF was recently reported in [137]. This now makes it possible to consider implementation of the EKF technique so that further analysis of the algorithm's performance on actual HF channels as opposed to a simulator could be performed. This would allow confirmation of the assumptions made regarding the channel model description in the algorithms derivation to be verified. Consequently, if the EKF offered the enhanced performance suggested by the results in this thesis, then it could be incorporated as the channel estimator for the Kalman equaliser structures described in chapter 5.

Since the tests performed in this thesis utilised a simplified channel model it would be necessary to confirm the performance of the algorithms and structures on a more realistic model or on actual channel data. If this was carried out, then the problems of carrier frequency acquisition and tracking, timing recovery and resistance to co-channel interference would have to be addressed. The channel model used in this work is a relatively simple model; As such, in addition to modelling the HF channel, it is very useful in the representation of a wide range of communication channels, such as troposcatter, meteor burst, line-of sight microwave and mobile communication channels all of which exhibit frequency selective fading. One consequence of this is that the results presented in this thesis are of more general use than just the HF communications scenario.

In terms of the equaliser structures, it would be interesting to develop the predictor LMS algorithm so that it could predict several samples ahead thus removing,

or at least reducing, the lag in the equaliser structures. In addition, some work on the performance of the equalisers when configured in a fractionally spaced format to determine if the increased complexity is acceptable for any significant improvement in performance.

APPENDIX A - RELEVANT PUBLICATIONS

1) S.McLaughlin,C.F.N. Cowan and B. Mulgrew, "Tracking Performance of Least Squares Algorithms as HF Channel Estimators", Presented at the IEE Digital Signal Processing Colloqium, January 1987, Digest No.1987/14 pp 16/1-7.

2)* S.McLaughlin,C.F.N. Cowan and B. Mulgrew "Performance Comparison of Least Squares and Least Mean Squares Algorithms as HF Channel Estimators", Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, U.S.A., April 1987, pp 2105-2108.

3)* S.McLaughlin, B. Mulgrew and C.F.N. Cowan "A Performance Study of the Extended Kalman Algorithm as a HF Channel Estimator", Proceedings of Fourth International Conference on HF Radio Systems and Techniques, London April 1988.

4)* S.McLaughlin, B. Mulgrew and C.F.N. Cowan "The Use of A-Priori Knowledge for HF Channel Estimation" Proceedings of the Fourth European Conference on Signal Processing (EUSIPCO) held in Grenoble September 1988.

5) S.McLaughlin, B.Mulgrew and C.F.N. Cowan "Performance Bounds for Exponentially Windowed RLS Algorithms in a Nonstationary Environment", Proceedings of the 2nd International Conference on Mathematics in Signal Processing, University of Warwick, Dec. 13-15 1988.

6)* S.McLaughlin, B.Mulgrew and C.F.N. Cowan "A Novel Adaptive Equaliser for Nonstationary Communication Channels", Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, May 1989, Glasgow

7) S.McLaughlin, B.Mulgrew and C.F.N. Cowan "A Performance Study of 3 Adaptive Equalisers in the Mobile Communications Environment", Proceedings of the IEEE International Conference on Communications 1989,Boston.

* Reprinted at back of thesis.

8) S. McLaughlin and C.F.N. Cowan. "A Performance Study of the RLS algorithm as a Channel Estimator in a Nonstationary Environment", Proceedings of the second IEE Adaptive Filters Colloquium 22nd of March 1989.

References

1. McBride, W.J., "Adaptive Equalisation for Data Transmission on HF Radio Channels," Ph.D. Thesis, University of Strathclyde, 1976.

2. Nadji, H.Y., "HF Communication," Ph.D. Thesis, University of Technology, Loughborough, 1984.

3. McVerry, F., "HF Communication," Ph.D. Thesis, University of Technology, Loughborough, 1985.

4. Harun, R., "Techniques of Channel Estimation for HF Radio Links," Ph.D. Thesis, University of Technology, Loughborough, 1984.

5. Hsu, F.M., "Square Root Kalman Filtering for High-Speed Data Received over Fading Dispersive HF Channels," *IEEE Trans.*, vol. IT-28, No. 5, pp. 753-763, September 1982.

6. Darnell, M., "Problems of Mobile HF Communication and Techniques for Performance Improvement," *IEE Proc. Part F*, vol. 132, No. 5, pp. 433-440, Aug. 1985.

7. Hodgkiss, W., "Digital Transmission Systems Operating over HF Radio Channels," Ph.D. Thesis, Imperial College, University of London, 1980.

8. Bello, P.A., "Selective Fading Limitations of the Kathryn Modem and Some System Design Considerations," *IEEE Trans.*, vol. CT-13, No. 3, pp. 320-333, Sept. 1965.

9. Dripps, J.H., "An Optimum Modulation Method for Digital Data Transmission on HF Ionospheric Channels," Ph.D. Thesis, University of Strathclyde, 1977.

10. Widrow, B., Hoff Jr., M., "Adaptive Switching Circuits," *IRE WESCON Convention Record* , vol. 4, pp. 96-104, 1960.

11. Gabor, D., Wilby, W.P.L., Woodcock, R., "A Universal Non-linear Filter, Predictor and Simulator Which Optimises itself by a Learning Process," *Proceed-*

ings IEE, Part B., vol. 108, pp. 422-438, 1961.

12. Glaser, E.M., "Signal Detection by Adaptive Filters," *IEEE Trans.*, vol. IT-7, No. 2, pp. 87-98, 1961.

13. Gauss, K.G., *Theory of Motion of Heavenly Bodies*, Dover, New York, 1931.

14. Legendre, A.M., *Elements de Geometrie*, Paris, 1794.

15. Cowan, C.F.N., Grant, P.M., *Adaptive Filters*, Prentice Hall, Englewood Cliffs, NJ, 1985.

16. Haykin, S., *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, NJ, 1985.

17. Goodwin, G.C., Sin, K.S., *Adaptive Filtering Prediction and Control*, Prentice Hall, Englewood Cliffs, NJ, 1984.

18. Honig, M.L., Mescherschmitt, D.G., *Adaptive Filters: Theory and Practice*, Kluwer Academic Publishers, 1984.

19. Widrow, B., Stearns, S., *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1985.

20. Mulgrew, B., Cowan, C.F.N., *Adaptive Equalisers: Algorithms and Structures*, Kluwer Academic Publishers, 1988.

21. Widrow, B. et al, "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter," *IEEE Proc.*, vol. 64, No. 8, pp. 1151-1162, 1976.

22. Widrow, B., Walach, E., "On the Statistical Efficiency of the LMS Algorithm with Nonstationary Inputs," *IEEE Trans.*, vol. IT-30, No. 2, pp. 211-221, 1984.

23. Ling, F., Proakis, J.G., "Nonstationary Learning Characteristics of Least Squares Adaptive Estimation Algorithms," *Proceedings ICASSP*, vol. 2, pp. 3.7.1-3.7.4, 1984.

24. Macchi, O., "Adaptive Equalisation of Time Varying Channels: What is meant by "slow variations"?," *Proceddings of ICC*, vol. 3, pp. 1247-1249, Amsterdam, 1984.

25. Haykin, S., Zhang, Q., "Tracking Characteristics of the Kalman Filter in a Nonstationary Environment for Adaptive Filter Applications," *Proceedings ICASSP*, pp. 671-674, 1983.

26. Clark, A.P., Harun, R., "Assessment of Kalman Filter Channel Estimators for an HF Radio Link," *IEE Proceedings, Part F.*, vol. 133, No. 6, pp. 513-521, 1986.

27. Beneviste, A., Goursat, M., "Blind Equalisers," *IEEE Trans.*, vol. COM-32, No. 8, pp. 871-883, 1986.

28. Picchi, G., Prati, G., "Blind Equalisation using a Stop-and-Go Decision-Directed Algorithm," *IEEE Trans.*, vol. COM-35, No. 9, pp. 877-887, 1987.

29. Mulgrew, B., Cowan, C.F.N., "An Adaptive Kalman Equaliser: Structure and Performance," *IEEE Trans.*, vol. ASSP-35, No. 12, pp. 1727-1735, 1987.

30. Godard, D., "Channel Equalisation Using a Kalman Filter for Fast Data Transmission," *IBM Journal of Res. & Dev.*, pp. 267-273, May 1974.

31. Glazier, E.V.D., Lamont, H.R.L., *The Services Textbook of Radio*, 5, HMSO, 1958.

32. Kraus, J.D., Carver, K.R., *Electromagnetics*, McGraw-Hill, 1973.

33. Furutsu, K., Ishida, T., "On the Theory of the amplitude distribution of impulsive random noise," *Applied Physics*, vol. 32, pp. 1206-1221, 1961.

34. Zadeh, L., "Frequency Analysis of Variable Networks," *Proceedings of IRE*, vol. 38, pp. 291-299, 1950.

35. Gersho, A., "Characterisation of Time Varying Linear Systems," *IEEE Proc.*, vol. 51, pp. 238-248, 1963.

36. Kaliath, T., "Time Variant Dispersive Channels," in *Chapter 6 in Lectures on Communication Systems Theory, Baghdady ed.*, McGraw-Hill, 1961.

37. Bello, P.A., "Characterisation of Randomly Time Variant Linear Channels,"

IEEE Trans., vol. CS 11, pp. 360-393, 1963.

38. Watterston, C.C. et al, "Experimental Confirmation of an HF Channel Model," *IEEE Trans.*, vol. COM-18, pp. 792-803, 1970.

39. Watterston, C.C., "Methods of Improving the Performance of HF Digital Radio Systems," *NTIA-ITS Report 79-29*, 1979.

40. Shaver, et al, "Evaluation of a Gaussian HF Channel Model," *IEEE Trans.*, vol. COM-15, pp. 79-88, 1967.

41. Shepherd, R.A., Lomax, J.B., "Frequency Spread in Ionospheric Radio Propagation," *IEEE Trans.*, vol. COM-15, pp. 268-275, 1967.

42. Farrell, P.G. et al, "Measurement and Verification of an HF Channel Model," *Proceedings of IEE Intnl. Conf. on HF Radio Systems*, pp. 52-57, 1985.

43. Giordano, A.A., Hsu, F.M., *Least Squares Applications to Digital Signal Processing*, Wiley, 1986.

44. Ehrman, L. et al, "Real-Time Software Simulation of the HF Radio Channel," *IEEE Trans.*, vol. COM-30, No. 8, pp. 1809-1817, Aug. 1980.

45. Proakis, J.G., *Digital Communications*, McGraw-Hill, 1984.

46. Friedlander, B., "System Identification Techniques for Adaptive Noise Cancelling," *IEEE Trans.*, vol. ASSP-30, pp. 708-716, 1982.

47. Eleftheriou, E., Falconer, D.D., "Tracking Properties and Steady State Performance of RLS Adaptive Filter Algorithm," *IEEE Trans*, vol. ASSP-34, pp. 1097-1109, 1986.

48. Macchi, O., Eweda, E., "Convergence Analysis of Self Adaptive Equalisers," *IEEE Trans.*, vol. IT-30, pp. 161-176, March 1984.

49. Macchi, O., "Optimisation of Adaptive Identification for Time Varying Filters," *IEEE Trans.*, vol. AC-31, pp. 283-287, 1986.

50. Eweda, E., Macchi, O., "Second Order Convergence Analysis of Stochastic

Adaptive Filtering," *IEEE Trans.*, vol. AC-28, pp. 76-85, 1983.

51. Eweda, E., Macchi, O., "Tracking Error Bounds of Adaptive Nonstationary Filtering," *Automatica*, vol. 21, No. 3, pp. 293-302, 1985.

52. Cowan, C.F.N., "Performance Comparisons of Finite Linear Adaptive Filters," *IEE Proc. Part F*, vol. 134, No. 3, pp. 211-216, 1987.

53. Tront, R.J., "Performance of Kalman Decision Feedback Equalisation in HF Radio Modems," Masters Thesis, University of British Columbia, 1982.

54. Bershad, N.J., Macchi, O., "Comparison of RLS and LMS algorithms for Tracking Chirped Sinusoid," *Proceedings ICASSP*, vol. 2, pp. 896-899, Glasgow, May 1989.

55. Macchi, O., Bershad, N.J., "Superiority of LMS over RLS for Tracking a Chirped Signal," *Proceedings of 2nd IEE Adaptive Filters Colloquium*, pp. 2/1-2/6, March 1989.

56. Wiener. N., *Extratoplation, Interpolation and Smoothing of Stationary Time Series*, Wiley, New York, 1949.

57. Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems," *Trans. ASME, Ser. D., J. Basic Eng.*, vol. 82, pp. 34-45, 1960.

58. Kalman, R.E., Bucy, R.S., "New Results in Linear Filtering and Prediction," *Trans. ASME, Ser. D., J Basic Eng.*, vol. 83, pp. 95-108, 1961.

59. Anderson, B.D.O., Moore, J.B., *Optimal Filtering*, Prentice Hall, Englewood Cliffs, NJ, 1979.

60. Jazwinski, A.H., *Stochastic Processes and Filtering Theory*, Academic Press, 1970.

61. Goodwin, G.C., Payne, R.L., *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press, 1977.

62. Ljung, L., Sodestrom, T., *Theory and Practice of Recursive Identification*, MIT

Press, 1983.

63. Gardner, W.A., "Learning Characteristicsof Stochastic-Gradient Descent Algorithms: A General Study, Analysis and Critique," *Signal Processing*, vol. 6, No. 2, pp. 113-133, 1984.

64. Widrow, B., "Adaptive Filters," in *Aspects of Network and System Theory R.E. Kalman & N. De Claris ed.*, Holt-Rinehart, Winston, NY, 1971.

65. Feuer, A., Weinstein, E., "Convergence Analysis of LMS Filters with Uncorrelated Gaussian Data," *IEEE Trans.*, vol. ASSP-33, No. 1, pp. 222-230, 1985.

66. Ayres, F., *Theory and Problems of Matrices*, Schaums Outline Series, McGraw-Hill, 1974.

67. Gelb, A., ed, *Applied Optimal Estimation*, MIT Press, 1982.

68. Cioffi, J.M., "Limited Precision Effects in Adaptive Filtering," *IEEE Trans.*, vol. CAS-34, No. 7, pp. 821-833, 1987.

69. Bellanger, M.G., *Digital Processing of Signals: Theory and Practice*, John Wiley, 1985.

70. Falconer, D.D., Ljung, L., "Application of Fast Kalman Estimation to Adaptive Equalisation," *IEEE Trans.*, vol. COM-26, pp. 304-337, 1978.

71. Mulgrew, B., "On Adaptive Filter Structure and Performance," Ph.D. Thesis, University of Edinburgh, 1987.

72. Proudler, I.K., McWhirter, J.G., Shepherd. T.J., "Fast QRD-based Algorithms for Least Squares Linear Prediction," *Proc. of 2nd IMA Conf. on Mathematics in Signal Processing*, Warwick, 1988.

73. Cioffi, J.M., "A Fast Adaptive Rotors RLS Algorithm," *Submitted IEEE Trans. ASSP*, Feb. 1989.

74. Porat, B., "Second Order Equivalence of Rectangular and Exponential Windows in Least Squares Estimation of Gaussian Autoregressive Processes," *IEEE Trans.*,

vol. ASSP-33, pp. 1209-1212, 1985.

75. Morf, M., Kaliath, T., "Square Root Algorithms for Least Squares Estimation," *IEEE Trans.*, vol. AC-20, No. 6, pp. 487-497, 1975.

76. Bellanger, M.G., "Coefficient Wordlength Limitation in FLS Adaptive Filters," *Proc. of ICASSP*, vol. 4, pp. 3011-3014, Tokyo, April 1986.

77. Bellanger, M.G., "Engineering Aspects of Fast Least Squares Algorithms in Transversal Adaptive Filters," *Proc. of ICASSP*, vol. 4, pp. 2149-2152, Dallas, USA, April 1987.

78. Lin. D.W., "On Digital Implementation of the Fast Kalman Algorithms," *IEEE Trans.*, vol. ASSP-32, No. 5, pp. 998-1005, 1984.

79. Cioffi, J.M., Kaliath, T., "Fast RLS Transversal Filters for Adaptive Filtering," *IEEE Trans.*, vol. ASSP-32, No. 2, pp. 304-337, April 1984.

80. Carryanis, G., Manolakis, D., Kalouptidis, N., "A Fast Sequential Algorithm for Least Squares Filtering and Prediction," *IEEE Trans.*, vol. ASSP-31, No. 6, pp. 1392-1402, 1983.

81. Kuo, B.C., *Digital Control Systems*, Holt-Saunders International, 1981.

82. Davie, A.M., *Private Communication*, 1988.

83. Gibson, G.J., "A Comparison of Linear Random Search and Least Mean Square Adaptive Algorithms," SPG Internal Report, University of Edinburgh, 1987.

84. Gibson, G.J., Cowan, C.F.N., "Communications Data Equalisation with Neural Networks," *Submitted to IEEE Trans, ASSP*, Nov. 1988.

85. Heffes, H., "The Effect of Erroneous Models on the Kalman Filter Response," *IEEE Trans.*, vol. AC-11, No. 3, pp. 541-543, 1966.

86. Nelson, L.W., Stear, E., "The Simultaneous On-line Estimation of Parameters and States in Linear Systems," *IEEE Trans.*, vol. AC-21, pp. 94-98, 1976.

87. Orlhac, D.G., Athanas, M., Speyer, J. & Houpt, P.K., "Dynamic Stochastic

Control of Freeway Corridor Systems Vol. IV, Estimation of Traffic Variables via an Extended Kalman Filter," *Report no. ESL-R-611*, MIT, Cambridge, MA, Sept 1975.

88. Schlee,F.H., Standish, C.J. & Toda, N.F., "Divergence in the Kalman Filter," *AIAA*, vol. 5, pp. 1114-1120, 1967.

89. Ljung, L., "Asymptotic Behaviour of the Extended Kalman Filter as a Parameter Estimator for Linear Systems," *IEEE Trans.*, vol. AC-24, No. 1, pp. 36-50, Feb. 1979.

90. Kaliath, T., "An Innovations Approach to Least-Squares Estimation Part I: Linear Filtering in Additive White Noise," *IEEE Trans. Automatic Control*, vol. AC-13, No. 6, pp. 646-655, Dec. 1968.

91. Frost, P., Kaliath, T., "An Innovations Approach to Least-Squares Estimation Part III: Nonlinear Estimation in White Gaussian Noise," *IEEE Trans.*, vol. Ac-16, No. 3, pp. 217-226, June 1971.

92. Kaliath, T., Geesey, R.A., "An Innovations Approach to Least Squares Estimation Part IV: Recursive Estimation Given Lumped Covariance Functions," *IEEE Trans.*, vol. AC-16, No. 6, pp. 720-727, Dec. 1971.

93. Kaliath, T., Geesey, R.A., "An Innovations Approach to Least-Squares Estimation Part V: Innovations Representations and Recursive Estimation in Coloured Noise," *IEEE Trans.*, vol. AC-18, No. 5, pp. 435-453, October 1973.

94. Gevers, M.R., Kaliath, T., "An Innovations Approach to Least-Squares Estimation Part VI: Discrete Time Innovations Representations and Recursive Estimation," *IEEE Trans.*, vol. AC-18, No. 6, pp. 588-600, Dec. 1973.

95. Aasnaes, H.B., Kaliath, T., "An Innovations Approach to Least Squares Estimation Part VII: Some Applications of Vector Autoregressive-Moving Average Models," *IEEE Trans.*, vol. AC-18, No. 6, pp. 601-607, Dec. 1973.

96. Kaliath, T., "The Innovations Approach to Detection and Estimation Theory," *Proc. IEEE*, vol. 58, No. 5, pp. 680-695, May 1970.

97. Ferrara Jr., E., Widrow, B., "The Time Sequenced Adaptive Filter," *IEEE Trans.*, vol. ASSP-29, No. 3, pp. 679-683, June 1981.

98. Srinath, M.D., Rajasekaran, P.K., *An Introduction to Statistical Signal Processing with Applications*, John Wiley & Son, New York, 1979.

99. Clark, A.P., Ashgar, S.M., "Dtection of Digital Signals Transmitted over a Known Time Varying Channel," *IEE Proc., Part F*, vol. 128, No. 3, pp. 167-174, June 1981.

100. Clark, A.P., McVerry, F., "Channel Estimation for an HF Radio Link," *IEE Proc., Part F*, vol. 128, No. 1, pp. 33-42, Feb. 1981.

101. Clark, A.P., "Near-Maximum-Likelihood Detectors," *1st IEE Adaptive Filters Colloquium*, vol. IEE Digest No. 1985/76, October 1985.

102. Quershi, S., "Adaptive Equalisation," *IEEE Proc.*, vol. 73, No. 9, pp. 1349-1387, Sept. 1985.

103. Young, M.C.S., "Application of Adaptive Equalisation to Microwave Digital Radio," Ph.D. Thesis, University of Edinburgh, 1989.

104. Wong, W.K., "Adaptive Equalisers for Multipath Compensation in Digital Microwave Communications," Ph.D. Thesis, University of Edinburgh, 1987.

105. Lucky, R.W., "Automatic Equalisation of a Digital Communications Channel," *BSTJ*, vol. 44, pp. 547-588, April 1965.

106. Lucky, R.W., Salz, J. & Weldon Jr., E.J., *Principles of Data Communication*, McGraw-Hill, New York, 1968.

107. Clark, A.P., *Equalisers for Digital Data Modems*, Pentech Press, 1985.

108. Monsen, P., "MMSE Equalisation of Interference on Fading Diversity Channels," *IEEE Trans.*, vol. COM-32, pp. 5-12, Jan. 1984.

109. de Olivera Duarte, A.M. & O'Reilly, J.J., "Simplified Techniques for Bounding Error Statitics of DFB Receivers," *IEE Proc. Part F*, vol. 132, No. 7, pp. 567-575, Dec. 1985.

110. Eleftheriou, E. & Falconer, D.D., "Adaptive Equalisation for HF Channels," *IEEE J. Selected Areas on Communications*, vol. 5, No. 2, pp. 238-247, Feb. 1987.

111. Lucky, R.W., "A Survey of the Communication Theroy Literature," *IEEE Trans.*, vol. IT-19, pp. 725-739, Nov. 1973.

112. Silvey, S.D., *Statistical Inference*, Chapman & Hall, London, 1975.

113. Abend, K., Fritchman, B.D., "Statistical Detection for Communication Channels with Intersymbol Interference," *IEEE Proc.*, vol. 58, pp. 779-785, May 1970.

114. Forney Jr., G.D., "Maximum-likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," *IEEE Trans.*, vol. IT-18, pp. 363-378, May 1972.

115. Forney Jr., G.D., "The Viterbi Algorithm," *IEEE Proc.*, vol. 61, pp. 268-278, May 1973.

116. Magee Jr., F.R., Prokis, J.G., "Adaptive Maximum-likelihood Sequence Estimation for Digital Signalling in the Presence of Intersymbol Interference," *IEEE Trans.*, vol. IT-19, pp. 120-124, Jan. 1973.

117. Russel, S.P., Shahara, A., "A Feasability Study of Reduced-State Viterbi Detection," *EASCON 1976 Record*, p. 89A, Washington, USA, Sept. 1976.

118. Falconer, D.D. et al, "Reduced Complexity Trellis Coding/Decoding," *Proc. of ICC*, vol. 1, Boston, June 1989.

119. Clark, A.P., Abdullah, S.N., "Near-Maximum-Likelihood Detectors for Voiceband Channels," *IEE Proc., Part F*, vol. 134, No. 3, pp. 217-226, June 1987.

120. Lawrence, R.E., Kaufmann, H., "The Kalman Filter for Equalisation of a Digital Communications Channel," *IEEE Trans.*, vol. CT-19, No. 6, pp. 1137-1141, Dec. 1971.

121. Bendetto, S. & Bigileri, E., "On Linear Receivers for Digital Transmission Systems," *IEEE Trans.*, vol. COM-22, No. 9, pp. 1205-1215, Sept. 1974.

122. Mulgrew, B., *Private Communication*, 1988.

123. Bagwell, D.J. & Considine, V., "Digital processing architectures for HF radio receivers," *Proceedings of the 3rd International Conference on HF Radio Systems and Techniques*, pp. 86-88, IEE, London, 1985.

124. Anderson, D.T. & Whikehart J.W., "A digital signal processing receiver," *Proceedings of the 3rd International Conference on HF Radio Systems and Techniques*, pp. 89-93, IEE, London, 1985.

125. Rochette, D. et al, "A Real time 800bit/s linear predicitive vocoder for HF communications," *Proceedings of the 4th International Conference on HF Radio Systems and Techniques*, pp. 285-289, IEE, London, 1988.

126. Miyake, M., Kitajima, M. & Urano A., "HF Digital Voice Terminal," *Proceedings of the 4th International Conference on HF Radio Systems and Techniques*, pp. 290-294, IEE, London, 1988.

127. Pearce, T.H. et al, "The application of DSP techniques in HF radio receiver design," *Proceedings of the 4th International Conference on HF Radio Systems and Techniques*, pp. 205-209, IEE, London, 1988.

128. Dutta, S., "A null steering adaptive array processor for voiceband HF links," *Proceedings of the 4th International Conference on HF Radio Systems and Techniques*, pp. 295-299, IEE, London, 1988.

129. Laycock, P.J. et al, "A model for HF spectral occupancy," *Proceedings of the 4th International Conference on HF Radio Systems and Techniques*, pp. 165-171, IEE, London, 1988.

130. Perry, B.D. & Abraham L.C., "A Wideband HF Interference and Noise Model based on Measured Data," *Proceedings of the 4th International Conference on HF Radio Systems and Techniques*, pp. 172-176, IEE, London, 1988.

131. Moulsely, T.J., "HF data transmission in the presence of interference," *Proceedings of the 3rd International Conference on HF Radio Systems and Techniques*, pp. 67-70, IEE, London, 1985.

132. Gibson, A.J. et al, "HF Spectrum Occupancy Measurements in Southern England," *Proceedings of the 3rd International Conference on HF Radio Systems and Techniques*, pp. 71-75, IEE, London, 1985.

133. Gibson, A.J. & Arnett L., "New HF Spectrum Occupancy Measurements in Southern England," *Proceedings of the 4th International Conference on HF Radio Systems and Techniques*, pp. 159-164, IEE, London, 1988.

134. Pennington, J., "Interference Tests on HF Modems," *Proceedings of the 4th International Conference on HF Radio Systems and Techniques*, pp. 318-320, IEE, London, 1988.

135. Gott, G.F. et al,, "Analysis of HF Interference with Application to Digital Communications," *IEE Proc. part F*, vol. 130, No. 5, 1983.

136. Dutta, S. & Gott, G.F., "Correlation of HF Interference Spectra with Range," *IEE Proc. part F*, vol. 128, No. 4, 1981.

137. Gaston, F.M. & Irwin, G., "VLSI architectures for square root covariance Kalman filtering," *Proc. SPIE - Adavanced Algorithms and Architectures for Signal processing IV*, vol. 1152, 1989.

A PERFORMANCE STUDY OF THE EXTENDED KALMAN ALGORITHM AS A HF CHANNEL ESTIMATOR

S McLaughlin, B Mulgrew and C F N Cowan

University of Edinburgh, UK

ABSTRACT

In this paper a study is made of the tracking and convergence properties of an extended Kalman (EK) algorithm as a high frequency (HF) channel estimator. This study uses the EK algorithm to adaptively estimate not only the time varying taps of the system but the parameters in the system model which generate the nonstationarity. This approach makes use of the a-priori knowledge of the HF channel to construct the extended Kalman algorithm for this system. Simulations are presented of the performance of the EK algorithm as a HF channel estimator for both white and coloured input signal conditions and are compared with the performance of the least mean square (LMS) and minimum variance Kalman (MVK) algorithms.

INTRODUCTION

The application of adaptive equalisation techniques to HF communication systems is necessary if higher data transmission rates than are presently possible are to be obtained. This means that robust adaptive algorithms which operate effectively in nonstationary environments are required. Unfortunately few studies of the convergence and tracking properties of adaptive algorithms in nonstationary environments have been published. The published work which is available has concentrated on the two most common adaptive algorithms, the least mean square (LMS) [1,2] and the recursive least squares (RLS) [3,4].

The poor performance of the LMS algorithm under coloured input signal conditions make it unsuitable for use in an HF channel equaliser. The RLS has been considered a likely candidate for such applications because of it's fast rate of convergence, which is independent of input signal colouration, in stationary environments. However, recently published studies indicate that the RLS has a considerably degraded performance in high noise [5] and nonstationary environments [6].

As a result new techniques will have to be developed to achieve the necessary performance required for HF communication systems. The work reported here is an exploratory study of a technique where estimation of the parameters which generate the nonstationarity are incoporated into the adaptive algorithm. In this paper a study is made of the tracking and convergence properties of an extended Kalman (EK) algorithm as a high frequency (HF) channel estimator. The direct modelling situation allows the properties of the algorithms to be analysed under controllable input signal conditions.

The EK algorithm is an application of linear Kalman filter theory to nonlinear systems, where the system is linearised around the current state estimate and the standard Kalman algorithm applied to the resulting time-varying linear system.

The standard model [7] of the HF channel represents it as a finite impulse response (FIR) filter (figure 1) with time varying complex taps each of which is statistically independent and has Gaussian statistics. A nonlinear system is obtained by representing the generation of the time varying taps in a state space formulation and then augmenting the state vector of the linear system with the stationary parameters that make up the tap generation model. The EK algorithm is then applied to this system. The key to the performance of the extended Kalman algorithm is the accuracy of the initial linearising approximation applied to the nonlinear system. In order that the EK algorithm has a good initial estimate of the system (which ensures convergence) for the EK an LMS algorithm is used to initially train the EK algorithm.

Simulations are presented of the performance of the EK and LMS algorithms as HF channel estimators for both white and coloured input signal conditions. A minimum variance Kalman (MVK) [6] estimator obtained by using a priori knowledge of the linear system and which achieves the lowest possible mean squared error (MSE) of any linear channel estimator is used as comparative measure of the performance of all the algorithms. The simulations illustrate the improved performance of the EK algorithm over the LMS algorithm as an HF channel estimator.

DERIVATION OF EXTENDED KALMAN ALGORITHM FOR HF CHANNEL ESTIMATION

It is well known that the Kalman filter is the optimal filter for the linear system described by equations 1 and 2 below;

$$x_{k+1} = F_k x_k + G_k w_k \tag{1}$$

$$y_k = H_k^T x_k + v_k \tag{2}$$

Clearly however, this filter is not optimal for a nonlinear system. The problem of optimal filtering for nonlinear systems is considerably more complex than is the case in linear systems theory. For nonlinear systems an exact solution via recursive methods is not normally possible, the conventional initial approach has been to adapt standard linear algorithms and determine their performance. The extended Kalman algorithm is simply an application of the linear Kalman filtering algorithm to a nonlinear system which has undergone a first order linearisation. Table 1 summarises the extended Kalman algorithm [10].

In the HF channel model the time varying taps are normally generated by filtering random numbers which have Gaussian statistics through a 2nd order Butterworth (or similar) filter whose bandwidth is dependent on the fade rate of the channel. For the simulations carried out in this paper a digital 2nd order Butterworth filter with filter structure as in figure 2 was used. This filter can be represented by equations (3) to (5) shown below,

$$X_0(k) = X_1(k-1) \tag{3}$$

$$X_1(k) = V(k) - C_1 X_0(k-1) - C_0 X_1(k-1) \tag{4}$$

the values of C_0 and C_1 being dependent on the bandwidth of the filter and consequently the fade rate. The taps for the HF channel model are then obtained using equation (5) below;

$$T_k = Cx_k + v(k) \tag{5}$$

This can be represented in the form of equations (1) and (2) with state $x_k = \begin{bmatrix} X_0(k) \\ X_1(k) \end{bmatrix}$ and state transition matrix $F_k = \begin{bmatrix} 0 & 1 \\ -C_1 & -C_0 \end{bmatrix}$ and $C = [1 \ -C_1 \ 2 \ -C_0]$.

One of these filters is used to generate each of the taps required for the channel model, thus, the overall state for a three tap channel model is $[X_0 X_1 X_2 X_3 X_4 X_5]_k^T$. By augmenting the state vector x_k with the filter coefficients C_0 and C_1 the system is made nonlinear. In order to apply the extended Kalman algorithm the state transition matrix $\Phi(k/k-1)$, and the measurement matrix M(k) are required.

To obtain the following substitutions, $X_n(k) = X_n(k) + \delta X_n(k)$, where X_n represents a reference state and n=0,1,...5 and $C_m(k) = C_m(k) + \delta C_m(k)$ m=0,1 and C_m a reference as before. Inserting these substitutions in equations (3) and (4) results in,

$$\delta X_0(k) = \delta X_1(k-1) \qquad (6)$$

and,

$$\delta X_1(k) = V(k) - X_1'(k-1)\delta C_0(k-1) - C_0'(k-1)\delta X_1(k-1) - \delta X_0(k-1)C_1'(k-1) - X_0'(k-1)\delta C_1(k-1) \qquad (7)$$

If this is applied to all of the individual states then the resulting state transition matrix is,

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -C_1' & -C_0' & 0 & 0 & 0 & 0 & -X_1' & -X_0' \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -C_1' & -C_0' & 0 & 0 & -X_3' & -X_2' \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -C_1' & -C_0' & -X_5' & -X_4' \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

similar arguments can be applied to the measurement matrix and lead to

$$M_k = [\, m_0 \; 2m_0 \; m_1 \; 2m_1 \; m_2 \; 2m_2 \; -(m_0 X_1' + m_1 X_3' + m_2 X_5') \; -(m_0 X_0 + m_1 X_2' + m_2 X_4') \,]$$

where m_i represents constituents of the input signal vector.

In the operation of the algorithm each new state estimate provided by the algorithm is used as the reference state. thus $X_n' = \hat{X}_n$ and $C_m' = \hat{C}_m$, where ^ indicates an estimate.

It should be noted that this realisation of the extended Kalman algorithm does not force a Butterworth form onto the tap generation filter nor does it guarantee that the estimated filter would be stable. This problem is overcome by ensuring that the intial approximation, i.e. the intial estimate used in the extended Kalman filter, is reasonably accurate in the mean squared sense by using the LMS to train it. Stability can be ensured by monitoring the poles of the filter , if they are not within the unit circle then by reflecting them through the unit circle along the same radii the filters stability is maintained. It is clear that if the extended Kalman converges then the filter approximation must be reasonably accurate.

The question of convergence of the extended Kalman is difficult and few results have been published. Ljung in [8] demonstrated that global convergence of the extended Kalman could be guaranteed if an innovations representation was used for the linearisation procedure rather than the normal state space formulation. This is because of the lack of coupling between the Kalman gain and the state being estimated. Recently Joshi in [9] considered the robustness of extended Kalman observers in the control field given certain actuator or sensor nonlinearities and showed convergence would be achieved given a certain limited range of the nonlinearity.

In this paper no theoretical proof of convergence is given but simulations are presented which show that given a reasonable initial estimate in the mean squared sense then convergence occurs.

Simulation Results

The simulations carried out for this paper were performed for a channel of fade rate 10.0Hz and signal-to-noise (SNR) of 50dB. Figures 3-5 show the algorithms performance for white input signal conditions and figures 6-8 for coloured input signal (eigenvalue ratio of 11.8) conditions.

Comparing the performance of the LMS with that of the MVK it is clear that the LMS is some 20dB from the noise floor. This is as a result of the contribution to the error by the time variations in the system as discussed by Macchi in [2]. It can be seen from figure 5 that the EK provides an improvement in perfromance of some 5-10dB. The EK utilised the LMS to provide it's initial estimate and as can be seen clearly improves upon it.

Unfortunately the performance of the EK degrades with time, this is due to the inherent numerical instability of the algorithm, although there are techniques for overcoming this problem they are not within the scope of this paper. Nevertheless the simulation demonstrates that the technique suggested in this paper may provide a means of improving upon the performance of existing algorithms.

Figures 6-8 illustrate the algorithms performance under similar conditions to those above except that the input signal is now coloured. The LMS shows a slight degradation in performance but again the EK approaches that of the MVK which is independent of input signal colouration since it's performance is based on the use of a-priori knowledge about the channel model.

CONCLUSIONS

It is clear from the results presented in this paper that the approach adopted has resulted in an improved channel estimator. Although the extended Kalman algorithm is not practical for implementation purposes, because of it's complexity and numerical instability, it has provided an insight into the perfomance of techniques which incoporate estimation of the parameters which generate the nonstationarity. The next stage will be to adapt existing simpler algorithms to operate in a similar manner to the extended Kalman and determine if the improved performance can be maintained with a simpler implementation structure.

In summary this paper has demonstrated the possibility of utilising a-priori knowledge of the system being identified to improve the performance. Ultimately it may even be possible to incoporate such a channel estimator into an equaliser of the form suggested by Clark et al in [11] or Macchi et al in [12] where the channel estimation and decision process are seperated.

References

1)Widrow,B et al.
Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter
IEEE Proceedings,Vol. 64,No.8,August 1976

2)Eweda,E. & Macchi,O.
Tracking Error Bounds of Adaptive Nonstationary Filtering
Automatica,Vol.21,No.3,pp. 293-302 1985

3)Ling,F. & Proakis,J.G.
Nonstationary Learning Characteristics of Least Squares Adaptive Estimation Algorithms
Proceedings ICASSP 1984 Vol.2,pp3.7.1-3.7.4

4)Evangelos,E, & Falconer,D.D.,
Tracking Properties and Steady State Performance of RLS Adaptive Filter Algorithms,
IEEE Trans. ASSP, Vol. ASSP 34,No. 5,October 1986

5)Cowan,C.F.N.,
Performance comparisons of finite linear adaptive filters
IEE Proceedings Part F, Vol. 134, No. 3, June 1987.

6)McLaughlin, S., Mulgrew, B. and Cowan, C.F.N.,
Performance Comparisons of Least Squares algorithms as HF channel estimators
Proceedings of ICASSP 1987, Vol. 4, pp 2105-2108.

7) Watteston, et al.,
Experimental Confirmation of an HF Channel Model,
IEEE Trans. Comms., Vol. COM-18 Sept.,1970

8)Ljung, L.
Asymptotic Behaviour of the Extended Kalman Filter as a Parameter Estimator for Linear Systems
IEEE Trans. on Automatic Control, Vol. AC-24, No. 1, Feb. 1979

9)Joshi, S.M.
Robustness of extended-Kalman-type observers
Int. J. of Control, 1987, Vol. 45, No. 5, pp 1857-1866

10)Jazwinski,A.,
Stochastic Processes and Filtering Theory,
Academic Press,1970

11)Clark,A.P.& McVerry,F.
Channel estimation for an HF radio link
IEE Proc.,Vol.128,Pt.F,No.1,Feb 1981

12)Macchi, O. & Hachicha, A.
Self-Adaptive Equalisation based on a prediction principle
Proceedings of GLOBECOM 1986, Vol. 3, pp 1641-1645.

TABLE 1 - EXTENDED KALMAN FILTER ALGORITHM

The nonlinear system is described by,

$$\frac{dx_t}{dt} = f(x_t, t) + G(t) w_t \ , \ t \geq t_0 \ , \ x_{t_0} \approx N(\hat{x}_{t_0}, P_{t_0})$$

and

$$y_{t_k} = h(x_{t_k}, t_k) + v_k$$

which undergoes a first order linearisation with $\delta x_t \approx x_t - \hat{x}(t)$ and $\delta y_{t_k} \approx y_{t_k} - \hat{y}(t_k)$. This leads to the linearised system described by,

$$\delta x_{t_{k+1}} = \Phi[t_{k+1}, t_k ; \hat{x}(t_k)] \delta x_{t_k} + w_{t_{k+1}}$$

and,

$$\delta y_{t_k} = M[t_k ; \hat{x}(t_k)] \delta x_{t_k} + v_k$$

The standard Kalman filter equations are then applied to this linearised system and the resulting algorithm consists of prediction via :

$$\hat{x}(t_{k+1} | t_k) = \hat{x}(t_k | t_k) + \int_{t_k}^{t_{k+1}} f(\hat{x}(t | t_k), t) \, dt$$

and

$$P(t_{k+1} | t_k) = \Phi[t_{k+1}, t_k ; \hat{x}(t_k | t_k)] \, P(t_k | t_k) \, \Phi^T[t_{k+1}, t_k ; \hat{x}(t_k | t_k)] + Q(t_{k+1})$$

and at an observation,

$$\hat{x}(t_{k+1} | t_{k+1}) = \hat{x}(t_{k+1} | t_k) + K[t_{k+1} ; \hat{x}(t_{k+1} | t_k)] \, [y_{t_{k+1}} - h(\hat{x}(t_{k+1} | t_k), t_{k+1})]$$

and,

$$P(t_{k+1} | t_{k+1}) = [I - K(t_{k+1} ; \hat{x}(t_{k+1} | t_k)) \, M(t_{k+1} ; \hat{x}(t_{k+1} | t_k))] \times P(t_{k+1} | t_k) \times [I - K(\,.\,) M(\,.\,)]^T + K(\,.\,) R(k+1) K^T(\,.\,)$$

where (.) represents what has gone before. Finally the Kalman gain is,

$$K[t_{k+1} ; \hat{x}(t_{k+1} | t_k)] = P(t_{k+1} | t_k) \, M^T(\,.\,) \times [M(\,.\,) P(\,.\,) M^T(\,.\,) + R(k+1)]^{-1}$$

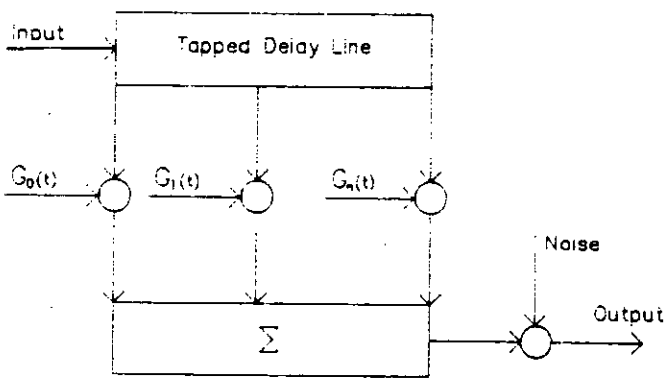


Figure 1 **HF channel model**

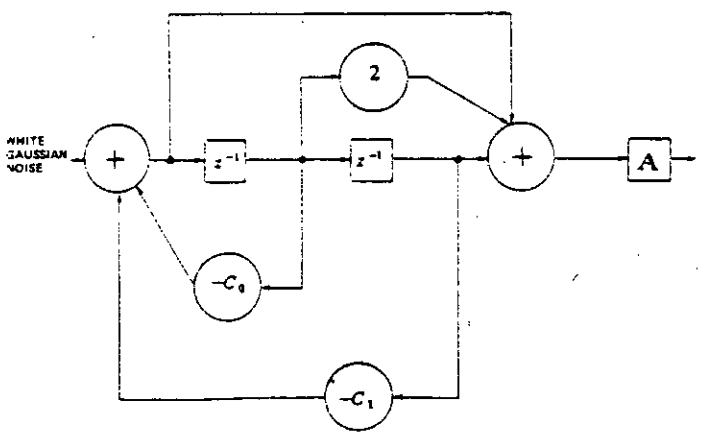


Figure 2 **Tap generation filter**

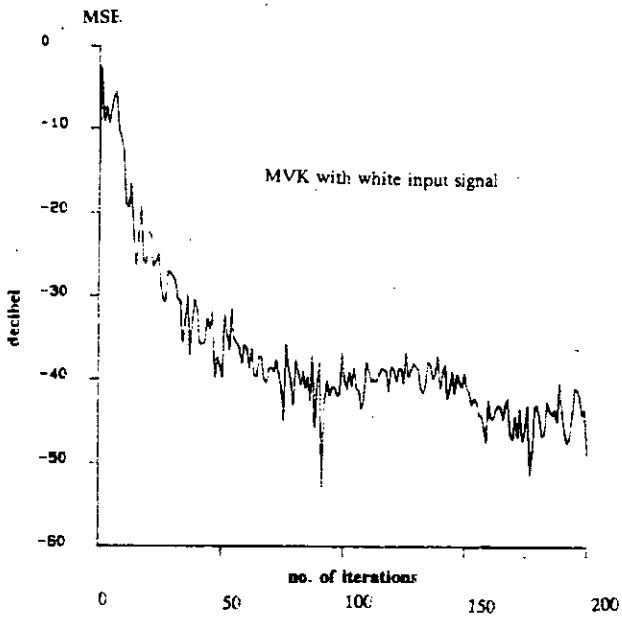


Figure 3

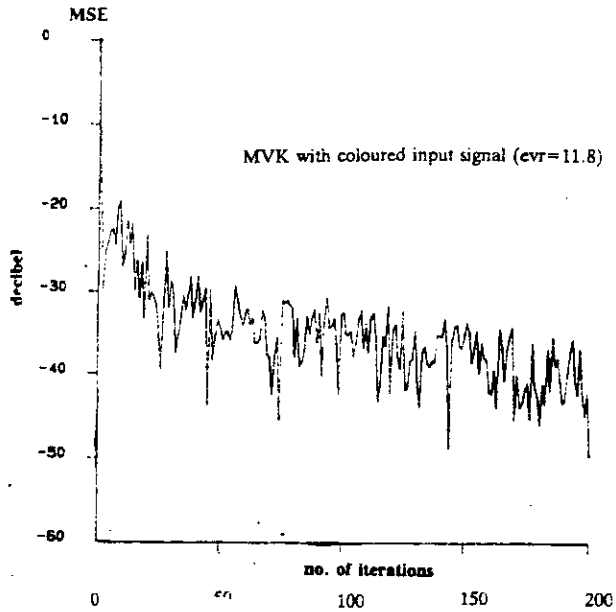
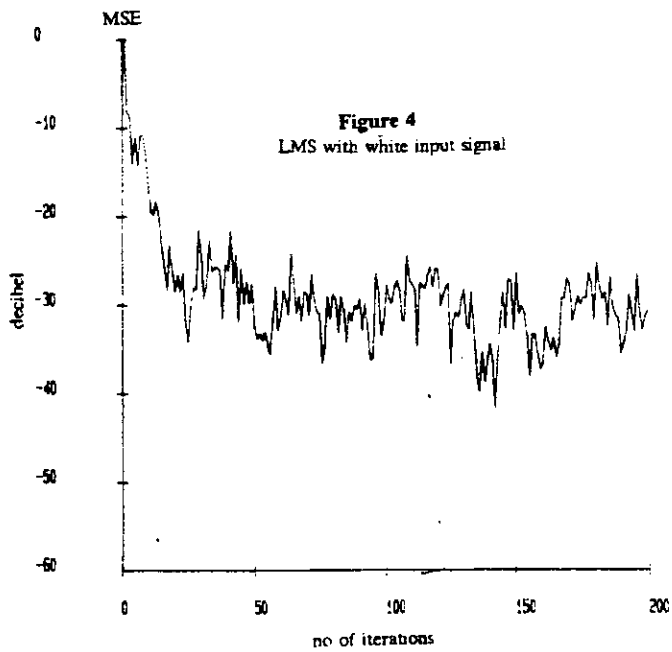
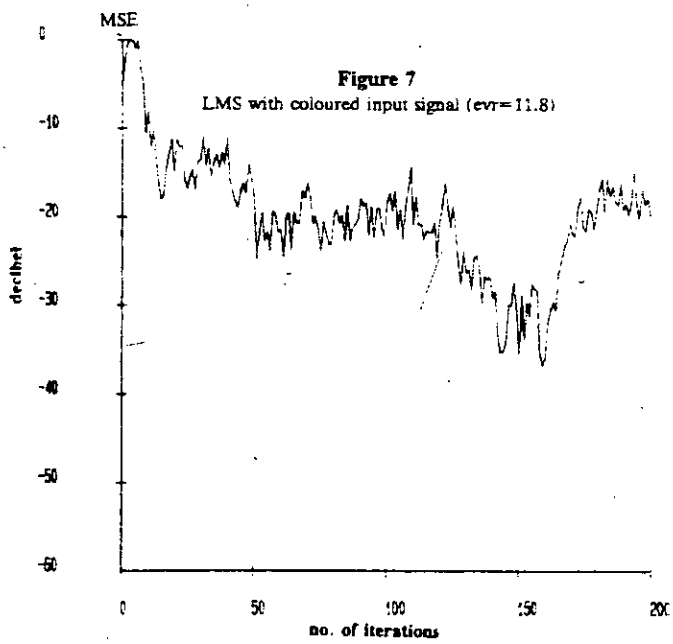
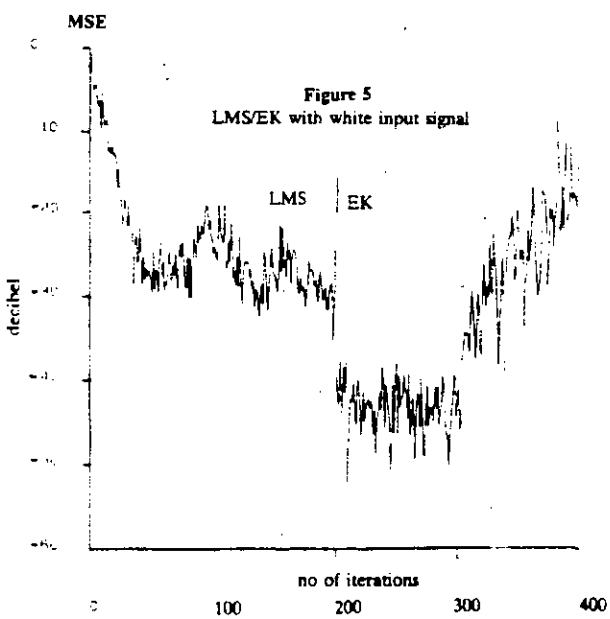
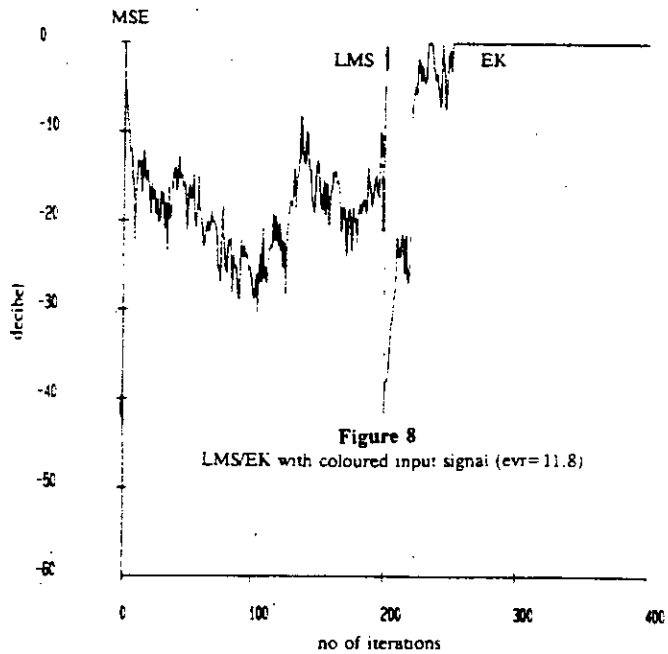


Figure 6









SIGNAL PROCESSING IV: Theories and Applications
J.L. Lacoume, A. Chehikian, N. Martin, and J. Malbos (Editors)
Elsevier Science Publishers B.V. (North-Holland)

THE USE OF A-PRIORI KNOWLEDGE FOR HF CHANNEL ESTIMATION

Stephen McLaughlin, Bernard Mulgrew and Colin F.N. Cowan

Dept. of Electrical Engineering, University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh, EH9 3JL

ABSTRACT

Within the high frequency (HF) communications system scenario there is a requirement for higher rates of data transmission than are presently possible. The use of adaptive equalisers is considered as one possible solution to this problem. However, until recently there have been few published studies which characterise the performance of adaptive algorithms in time-varying environments (such as the HF communications channel). In this paper a new algorithm for use in HF channel estimators is developed which utilises a-priori knowledge of the channel structure. A comparitive study of the algorithm's performance against more conventional adaptive algorithms is presented.

I INTRODUCTION

Until recently there have been few published studies which characterise the performance of adaptive algorithms in time-varying environments (such as the HF communications channel). Recently published studies [1,2,3], would appear to suggest that the algorithms considered most suitable for this application, the least mean squares (LMS) and recursive least squares (RLS) [4], are incapable of tracking the time variations of the channel accurately enough. The aim of this paper is to investigate possibile alternatives to such algorithms by utilising the available a-priori information about the structure of the HF channel.

It is worth noting that several authors have indicated previously [1,5] the importance of accurate channel prediction in adaptive equalisation, this and the fact that the best channel estimators are those with a-priori knowledge of the channel make it logical to incoporate some form of prediction which utilises the a-priori knowledge of the simulated tap generation process for more accurate estimation of the channel.

The performance of various HF channel estimators was studied in [6,7] and the results can be summarised as follows:-

1) Minimum Variance Kalman (MVK) - this estimator requires full a-priori knowledge of the channel and as a result is not implementable but provides the lowest achievable mean squared error (MSE) bound of any HF channel estimator.

2) Extended Kalman Filter - This estimator has partial a-priori knowledge of the channel and it's performance approaches that of the MVK but convergence is not guaranteed for this algorithm and it also suffers from a computational complexity which excludes implementation.

3) Adaptive FIR Filters (LMS/RLS) - These estimators have no a-priori knowledge of the channel and also have the poorest performance although the least mean square (LMS) is the least computationally complex.

Clearly the ideal estimator would be one which has a performance approaching that of the MVK but with a computational complexity comparable with that of the LMS.

The standard model for the HF channel is a FIR filter with statistically independent time varying taps (as in fig. 1), the taps being generated by passing a white random sequence with Gaussian statistics through a filter with a bandwidth determined by the fade rate of the channel and an (approximate) Gaussian frequncy response. A second order Butterworth filter was used for the simulation and analysis presented here. This a-prori knowledge is then used to develop an algorithm which utilises an LMS algorithm and incoporates a prediction filter based on the tap generation model assumed for the simulations presented.

This modified LMS is a first attempt at achieving the performance of the MVK algorithms while reducing the

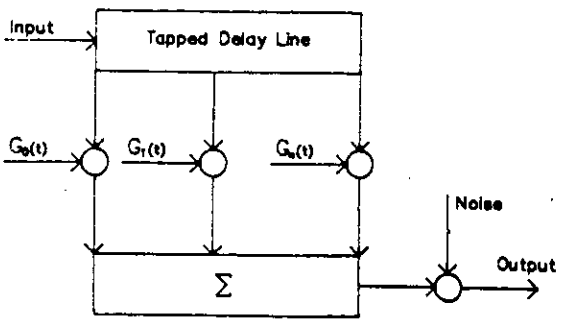


Figure 1 - HF Channel model

computational complexity. The algorithm was implemented on computer as a set of parallel prediction filters and LMS algorithms (as in fig. 2). Each prediction filter catering for a possible fade rate and the algorithm with the lowest MSE was utilised. In this way the system could cope with changing fade rates and always achieved the optimal performance.

Simulation results are presented which demonstrate the performance of this modified LMS as a HF channel estimator, and compared with the conventional LMS and a MVK estimator, which is used to demonstrate the minimum achievable MSE.

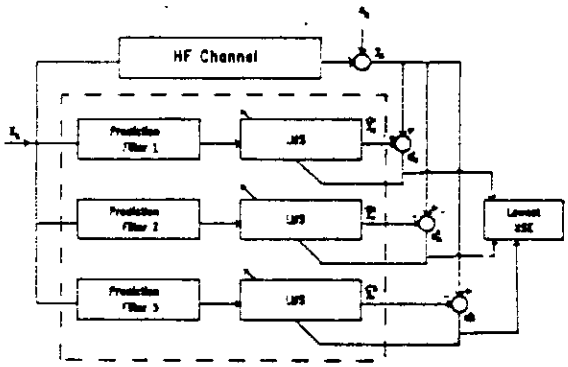


Figure 2 - Algorithm Set up

II ALGORITHM DEVELOPMENT

As has been stated the conventional model for the HF communications channel is a FIR filter with time varying taps. The tap generation filter used is a 2nd order digital butterworth filter as illustrated in figure 3, the coefficients C_0 and C_1 being determined by the bandwidth of the filter. The filter may be considered as a linear system represented by equations 1 and 2 below:

$$\underline{s}_{k+1} = F\underline{s}_k + Gw_k \tag{1}$$

$$h_k = C\underline{s}_k + u_k \tag{2}$$

Where the state $s_k = \begin{bmatrix} S_0(k) \\ S_1(k) \end{bmatrix}$ and the state transition matrix $F = \begin{bmatrix} 0 & 1 \\ -C_1 & -C_0 \end{bmatrix}$ with $C = [\,1 \;\; -C_1 \;\; 2 \;\; -C_0\,]$. Both w_k and u_k are zero mean with variances $V_w(k)$ and $V_u(k)$ respectively.

From this assumed tap-generation model a prediction filter for each of the taps can be constructed. The ideal inputs for these prediction filters would be the actual taps which are however not directly observable, but the estimate of the taps provided by the LMS is a reasonable approximation. The predicted value of the taps obtained from this filter can then be used in the LMS algorithm. This could be viewed as increasing the order of the recursion in the LMS from first to second.

In order to derive the required prediction filters directly it would be necessary to carry out a spectral factorisation on the tap generation process, this would be both difficult and computationally intensive because the input to the prediction filters is a noisy estimate of the actual tap. As a result a Kalman filter is used since it is equivalent to carrying out the spectral factorisation.

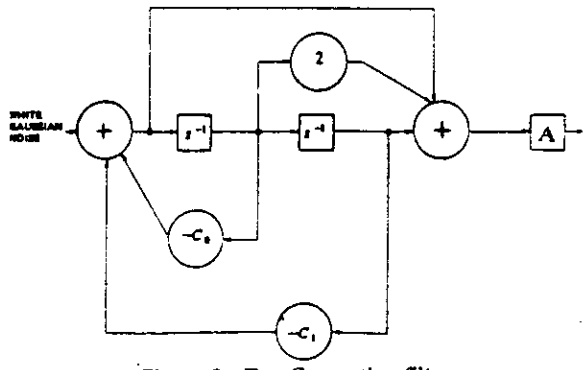


Figure 3 - Tap Generation filter

Using the model of the tap generation process detailed in equations 1 and 2 allows a one stage predictor algorithm, as detailed in [8] to be written as follows

$$\hat{\underline{s}}^{k}_{k+1} = F\,\hat{\underline{s}}^{k-1}_{k} + \underline{K}^{k}_{k-1}\,v_k \tag{3}$$

with $v_k = h_k - C\,\hat{\underline{s}}^{k-1}_{k}$, i.e. the error in the tap estimate, and the Kalman gain is,

$$\underline{K}^{k}_{k-1} = FV_s(k|k-1)C^T V_v^{-1}(k) \tag{4}$$

where the variances are as follows,

$$V_v(k) = C^T V_s(k|k-1) + V_u(k) \tag{5}$$

and,

$$V_s(k+1|k) = FV_s(k|k-1)F^T + GV_w(k)G^T - \underline{K}^{k}_{k+1}CV_s(k|k-1)F^T \tag{6}$$

If the variance and gain equations (eqns. 4-6) are taken to the steady state by computer simulation then the steady state gain, $\underline{K}_{ss}$ is obtained and this can be used in the algorithm for prediction as detailed in the equations shown below.

$$\hat{h}^{k-1}_{k} = C\,\hat{\underline{s}}^{k-1}_{k} \tag{7}$$

that is the predicted value of the tap based on an estimate of the state obtained from,

$$\hat{\underline{s}}^{k-1}_{k} = F\,\hat{\underline{s}}^{k-1}_{k-1} \tag{8}$$

and,

$$\hat{\underline{s}}^{k-1}_{k-1} = \hat{\underline{s}}^{k-2}_{k-1} + \underline{K}_{ss}\,[\,\hat{h}_{k-1} - \hat{h}^{k-2}_{k-1}\,] \tag{9}$$

That is, since the actual taps are not observable then the estimate obtained from the LMS, $\hat{h}_{k-1}$, as illustrated below, is used to aid the prediction of the state.

$$\hat{h}_{k-1} = \hat{h}^{k-2}_{k-1} + 2\mu X_{k-1}e_{k-1} \tag{10}$$

where X_{k-1} represents the input signal and the error e_{k-1} is obtained from,

$$e_{k-1} = y_{k-1} - \hat{y}_{k-1} \tag{11}$$

that is the output of the channel less the estimated output where the output $y_{k-1} = h^T_{k-1}X_{k-1} + n_k$ where n_k is additive white gaussian noise. The estimated output is,

$$\hat{y}_{k-1} = \hat{h}^{k-2}_{k-1}X_{k-1} \tag{12}$$

These equations are shown for one tap only but clearly it is trivial to extend it to multiple taps each with their own prediction filter.

It is worth noting that this form of prediction could readily be extended to predict the channel impulse response M samples ahead, M being some integer, for use in a Viterbi based equaliser as in [5]. Normally as M gets larger the equalisation improves but the system identification degrades (i.e. the tracking performance degrades). But hopefully with this technique the tracking performance could be maintained while still allowing the decision to be delayed M samples.

III SIMULATION RESULTS

The simulations carried out for this paper were performed for a three tap channel with a fade rate of 10.0Hz and signal-to-noise ratio (SNR) of 50dB. All simulations are for a white input signal and are averaged over an ensemble of 30. The simulations were performed on a Sun 350 workstation in 64 bit double precision.

Figure 4 illustrates the performance of the MVK which approaches the noise floor even under the severe fading conditions simulated. Figure 5 shows the predictor filter utilising the actual taps to demonstrate it's optimal performance which approaches that of the MVK. The performance of the LMS, as shown in fig 6, is some 20dB from the noise floor in the mean with a considerable variance. Figure 7 which illustrates the performance of the combined LMS and prediction filter which is no better than that of the LMS.

This performance although dissapointing in the sense that it offers no improvement over the LMS suggests that were it configured for prediction M samples ahead, as suggested earlier would offer an improved performance over a Viterbi equaliuser using a conventional LMS channel estimator.

IV CONCLUSIONS

In this paper a new algorithm for use in the HF communications environment has been developed and it's performance demonstrated. Although the performance of the algorithm was no better than the LMS it offers possible improved performance over a conventional LMS in a Viterbi type equaliser of the type suggested in [5].

Further investigation of the algorithm's performance under such conditions will be carried out and reported at a later date.

ACKNOWLEDGEMENTS

This work was funded by the Science and Engineering Research Council of Great Britain.

REFERENCES

1)Clark,A.P. & Harun,R.,
Assessment of Kalman-filter channel estimators for an HF radio link.
IEE Proc. Part F, Vol.133, No.6, October 1986, pp513-521.

2)Evangelos,E. & Falconer, D.D.,
Tracking Properties and Steady State Performance of RLS Adaptive Filter Algorithms,
IEEE Trans. ASSP, Vol. ASSP-34, No. 5, October 1986, pp1097-1109.

3) Tront,R.J.,
Performance of Kalman Decision Feedback Equalization in HF Radio Modems
Masters Thesis,University of British Columbia,1982.

4)Cowan,C.F.N. & Grant,P.M.,
Adaptive Filters,
Prentice Hall,1985.

5) Clark, A.P. and Ashgar, S.M.,
Detection of digital signals transmitted over a known time-varying channel
IEE Proc. Part F, Vol. 128,No. 3, June 1981

6) McLaughlin, S., Mulgrew, B. and Cowan, C.F.N.
Performance Comparison of Least Squares and Least Mean Squares Algorithms as HF Channel Estimators
Proceedings of ICASSP 1987, Vol.4, pp. 2105-2108

7) McLaughlin, S., Mulgrew, B. and Cowan,C.F.N.,
A Performance Study of the Extended Kalman Algorithm as a HF Channel Estimator
Proceedings of IEE 4th International Conference on 'HF RADIO SYSTEMS AND TECHNIQUES' 11-13 April 1988

8)Srinath, M.D. & Rajasekaran, P.K.,
An Introduction to Statistical Signal Processing with Applications
John Wiley & Sons, 1979.

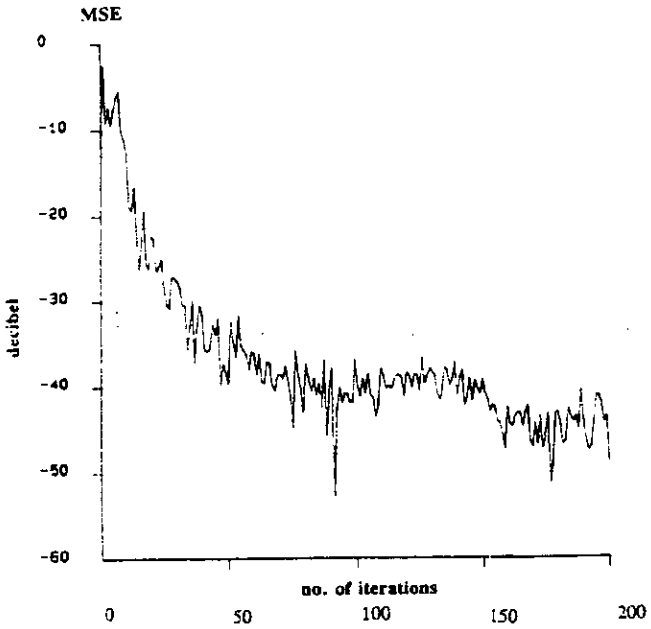


Figure 4 - MVK performance

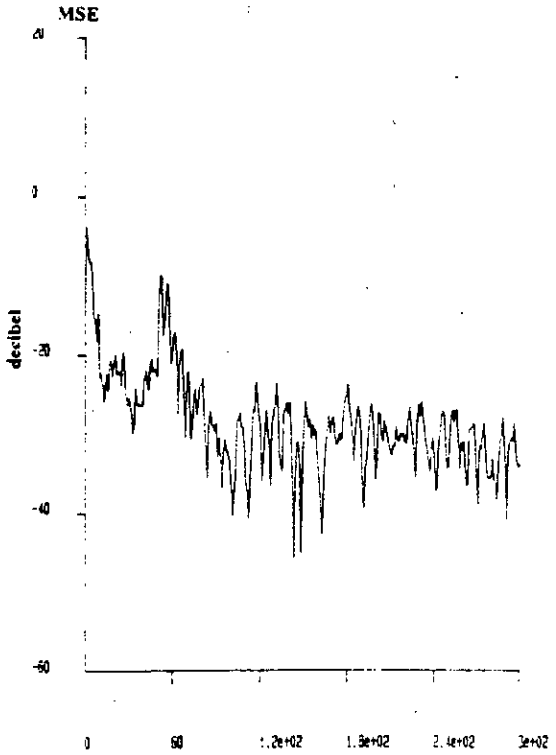


Figure 5 - Optimal predictor performance

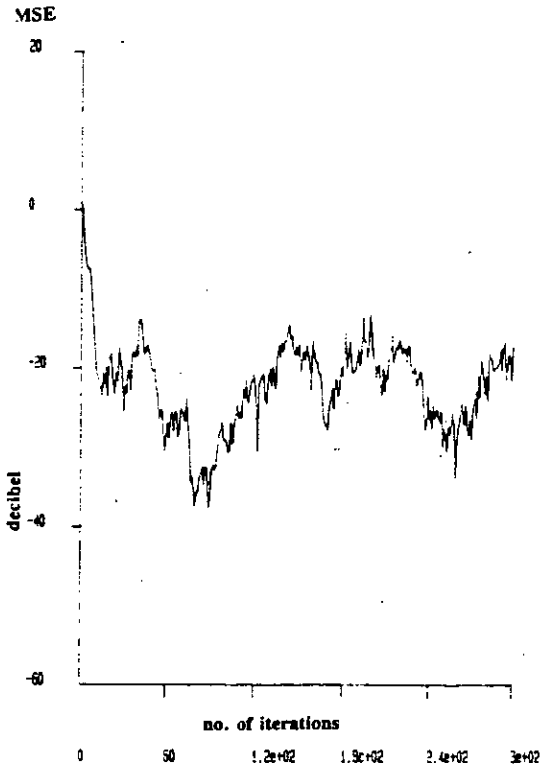


Figure 6 - LMS performance

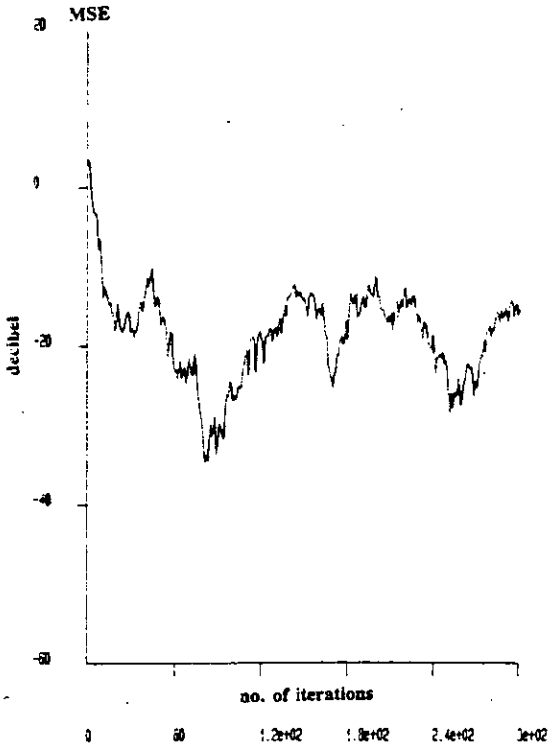


Figure 7 - Combined LMS and Predictor performance

A Novel Adaptive Equaliser for Nonstationary Communication Channels

Stephen McLaughlin, Bernard Mulgrew and Colin F.N. Cowan Member I.E.E.E.

Dept. of Electrical Engineering, University of Edinburgh,

Mayfield Road, Edinburgh, EH9 3JL, Scotland, United Kingdom.

ABSTRACT

Adaptive equalisation of communications channels is used to compensate for time dispersion introduced into the transmitted data sequence. In this paper the performance of a novel Kalman decision feedback equaliser (DFE) which uses a channel estimator via a least mean squares (LMS) algorithm is studied for a variety of stationary and nonstationary communications channels. This structure providing a means of model order reduction by using the residuals of the LMS to provide information on the unmodelled paths in the communication channel, which is then incorporated into the Kalman DFE structure as observation noise. The structure is compared with a conventional DFE which is trained by a Godard-Kalman algorithm with exponential windowing and an adaptive Kalman structure previously reported [2].

INTRODUCTION

In the digital communications scenario many of the media over which communications is attempted are nonstationary, i.e. time varying, in nature, e.g. high frequency and mobile communications channels. As a consequence it is necessary to utilise adaptive equalisation techniques to overcome this problem and ensure robust communication at the required data rate. Existing adaptive equaliser structures [1] although adequate for many applications suffer a degraded performance under the more severe conditions which can occur. This can be attributed to several causes, the adaptive algorithms lack spectral robustness or suffer from numerical instabilities, the equaliser structures fail to approximate the inverse of the channel accurately, or propagation of decision errors in the structure. In this paper an attempt is made to overcome some of these problems with a novel equaliser structure which consists of an equaliser with decision feedback and utilises a least mean square (LMS) algorithm for channel estimation in a manner similar to that discussed in [2].

This equaliser provides several benefits over existing structures, in particular it offers a means of model order reduction. This is achieved by using the residuals of the LMS to provide information on the unmodelled paths in the communication channel and incorporating this information as observation noise in the Kalman structure. Also, the number of taps required by the LMS for channel estimation will normally be less than that required by either a linear equaliser or of a conventional decision feedback equaliser (DFE), thus reducing the number of taps requiring to be adjusted adaptively. The structure is also more robust to decision error propagation than a conventional DFE because the channel estimator has fewer taps and due to the inherently robust nature of the LMS.

Simulation results are presented in the paper showing the equalisers' mean squared error (MSE) performance compared to that of a conventional DFE which uses a Godard-Kalman adaptive algorithm with exponential windowing for a nonstationary channel.

EQUALISER STRUCTURES

Research over the last twenty years has produced a large body of literature, [1] and references therein, and there are many types of equaliser structure, however, they may be summarised as follows:

a) Linear transversal equalisers which in general suffer from an inability to represent the inverse of the channel impulse response adequately.

b) Conventional decision feedback equalisers which, although providing a better performance than linear equalisers, suffer from a degraded performance due to error propagation in the feedback section.

c) Maximum likelihood sequence estimation is a technique which is not considered in this paper but whose main disadvantage would appear to be its computational complexity. However considerable effort is being expended in developing more efficient implementations.

In general most equalisers operate by generating an estimate of the inverse filter which when convolved with the channel response allows the transmitted data sequence to be reconstructed accurately, as in the linear equaliser illustrated in Figure 1. For the case of conventional DFE's, a feedback filter is inserted after the decision device, (as in Figure 2a), and is used to cancel out any trailing intersymbol interference (ISI) by using previously detected symbols, which are assumed to be correct.

A range of adaptive algorithms are used in adaptive equalisation, the two most common being the LMS and RLS algorithms [3]. The LMS offers an easily implementable algorithm but lacks the spectral robustness and fast convergence of the RLS, which unfortunately is relatively complex to implement. The conclusions of [4,5,6] would appear to suggest that when operated as a channel estimator, as opposed to an equaliser, the LMS offers a similar if not better performance than the RLS on channels which are time-varying. These results have provided some of the motivation for the novel structure considered here, in that the data sequence estimation and channel impulse response estimation processes are separated.

The conventional DFE structure differs, as has been stated earlier, from the linear structure by the addition of a feedback section which is used to cancel out the ISI associated with these symbols. The feedback section allows a greater freedom for the linear section in selecting tap weight coefficients. Conventional DFE's of this type have been found to operate very well over wire line channels but in rapidly time-varying environments the performance appears to be degraded by error propagation in the feedback section [7].

In this paper the algorithm which was used to adjust the tap weight coefficients of the equaliser was the algorithm postulated by Godard [8] in 1974 in which he chose not to replace the equaliser with a conventional Kalman filter, but rather adopted a transversal equaliser structure and used the Wiener solution for the optimal tap weights as a starting point.

The algorithm offers very fast initial convergence and is spectrally robust but suffers from a relatively high level of complexity. To apply it to the DFE, the observation vector contains both the feedforward and feedback coefficients. The solution obtained by Godard was for a stationary channel and its application was extended to slowly time-varying channels by means of exponential data windowing.

In [2] Mulgrew and Cowan presented a novel equaliser structure, the derivation of which may be summarised as follows. Initially, a channel model based on a FIR filter was postulated and the constraint that the optimum transversal equaliser for such a channel, which requires minimisation of its MSE subject to the impulse response being finite, casual and stable is relaxed. The new relaxed constraint requires only that the filter be casual and stable, this results in the solution to the minimisation problem being provided by a Wiener infinite impulse response (IIR) filter.

The Wiener IIR filter offers advantages over the conventional linear (FIR) equaliser in terms of the order required for the same level of performance for minimum phase channels. However, the realisation of such a filter would require a minimum phase spectral factorisation, this would present a major difficulty. The solution adopted in [2] was to use a Kalman filter to realise the solution since, if the processes are stationary and the observation noise white, then the steady state Kalman and Wiener IIR filters are identical. The FIR filter model of the communications channel in common usage [1,2] although readily adapted to a state space representation, and hence to a Kalman filter, requires care in the selection of states which will constitute the state vector.

In order to deal with non-minimum phase channels, a fixed lag smoothing [9] form of the Kalman filter was used. Normally this would imply that for a fixed lag, d, a state vector augmented to,

$$[\ \underline{s}^T(k)\ \underline{s}^T(k-1)\ .....\ \underline{s}^T(k-d)\ ].$$

However, this is uneccessarily complex because the state transition matrix $\underline{a}$ is simply a shift matrix, thus the state vector is augmented to contain $d+1$ elements [2],

$$\underline{s}^T(k) = [\ s(k)\ s(k-1)\ .....\ s(k-M+1)\ \cdots\ s(k-d)\ ],$$

where d is the fixed lag and M is the number of taps in the channel. The state transition equation then becomes,

$$\underline{s}(k) = \underline{a}\underline{s}(k-1) + \underline{b}s(k),$$

where $\underline{a}$ is a $(d+1) \times (d+1)$ shift matrix and $\underline{b}$ is a vector with $(d+1)$ elements,

$$\underline{b}^T = [\ 1\ 0\ 0\ ....\ 0\ ].$$

The observation equation is clearly,

$$x(k) = \underline{h}^T\underline{s}(k) + n(k),$$

where $\underline{h}$ is a column vector with $(d+1)$ elements,

$$\underline{h}^T = [\ h_0\ h_1\ \cdots\ h_{M-1}\ 0\ 0\ \cdots\ 0\ ].$$

That is the channel tap weight vector augmented to $(d+1)$ elements by the addition of some zeros.

The problem still remains however of making the equaliser adaptive, since knowledge of the channel impulse response is required and this is achieved by using an LMS algorithm as a channel estimator to provide the required estimate. Some measure of the observation noise in the system is also required for the Kalman structure and this is achieved by the recursion shown below; the derivation of which is detailed in [2].

$$\hat{\sigma}_e^2(k+1) = (\ 1 - 1/M\ )\hat{\sigma}_e^2(k) + e^2(k+1)/M,$$

where $e(k+1)$ is the error obtained from the LMS. The above recursion also effectively provides a measure of the model uncertainty in the system and therefore offers a means for model order reduction by using the residuals to provide some information on any paths not modeled by the LMS. The main point to note about the general structure of this equaliser is the separation of the state and channel estimation processes, and this approach is extended in the next section, the structure of the equaliser is ilustrated in figure 2b.

AN ADAPTIVE KALMAN DECISION FEEDBACK EQUALISER

As in the development of the previous equaliser [3] the structure of the channel model is again the starting point, the channel is considered as a FIR filter which can represented in the state space form shown below,

$$\underline{s}(k+1) = \Phi(k+1,k)\underline{s}(k) + n(k).$$

Where $\Phi(k+1,k)$ represents the state space transition matrix associated with the channel and $\underline{s}(k)$ is the input vector. The following observation equations can be associated with this model,

$$x_1(k) = \underline{h}^T\underline{s}(k) + n(k)$$

and

$$x_2(k) = \underline{s}(k-d).$$

The first equation represents the channel output, as in the conventional DFE structure, and the second represents a decision feedback term. It is worth noting at this point that it would be possible to have many similar observation equations to the second one, that is feed back many decisions, but preliminary investigations suggested that nothing would be gained in terms of performance for a considerable increase in complexity. The two observation equations can be combined as below,

$$\underline{x}(k) = \underline{H}\underline{s}(k) + \underline{N}(k).$$

$\underline{H}$ being a $(d+1) \times 2$ matrix as shown below,

$$\begin{bmatrix} h_0 & h_1 & \cdots & h_{M-1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

It is clear that the problem presented here is identical to that of the previous section, the difference lying in the observation equations. It is necessary then to generate the Kalman filter which provides the solution given these particular observation equations. With the definitions above it is possible to write down the Kalman filter equations which are detailed in table 2. Note that the inversion of the innovations now requires the inversion of a 2×2 matrix. The LMS channel estimator is used in exactly the same way as the previous structure to provide the tap weight vector estimate and a measure of the observation noise. The detailed structure of the equaliser is shown in figure 2c.

The question remains as to how many states are required for the equaliser to operate efficiently, and clearly $n_{min} = d + 1 + 1$. That is, the estimation with lag delay is $d + 1$ and there is one feedback term.

Table 1 - An Adaptive Kalman Decision Feedback Equaliser

$$\hat{\underline{s}}(k,k) = \hat{\underline{s}}(k/k-1) + \underline{K}(k)\ [\ x(k) - \underline{H}\ \hat{\underline{s}}(k/k-1)\ ]$$

$$\hat{\underline{s}}(k/k-1) = \underline{a}\ \hat{\underline{s}}(k-1/k-1)$$

$$\underline{K}(k) = \underline{V}(k/k-1)\underline{H}^T\ [\ \underline{H}\underline{V}(k/k-1)\underline{H}^T + \underline{\sigma}_v^2\ ]^{-1}$$

$$\underline{V}(k/k-1) = \underline{a}\ \underline{V}(k-1/k-1)\ \underline{a}^T + \underline{b}\underline{b}^T\sigma_v^2$$

$$\underline{V}(k/k) = [\ I - \underline{K}(k)\underline{H}\ ]\underline{V}(k/k-1)$$

where the Kalman gain vector is as below,

$$\begin{bmatrix} K_{0,0} & K_{0,1} & \cdots & K_{0,d-1} \\ K_{1,0} & K_{1,1} & \cdots & K_{1,d-1} \end{bmatrix}$$

The LMS system identification being,

$$\hat{\underline{h}}(k+1) = \hat{\underline{h}}(k) + 2\mu\ \hat{\underline{s}}(k+1)e(k+1)$$

$$e(k+1) = x(k+1) - \hat{\underline{h}}^T(k)\ \hat{\underline{s}}(k+1)$$

The estimated tap weight vector $\hat{\underline{h}}(k)$ being used in the d element vector $\underline{H}$ as below.

$$\begin{bmatrix} h_0 & h_1 & \cdots & h_{M-1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

ALGORITHM COMPLEXITY CONSIDERATIONS

A breakdown of the computation required to process each of the three algorithms considered here is presented in this section. Tables 2 and 3 present the computation required for each process in the algorithm given particular values of the lag d and the number of taps in the channel.

For the simulations presented in this paper the number of taps in the channel was $M=3$ with the lag for both the adaptive Kalman structures being $d=5$. The Godard-Kalman driven DFE had 5 feedforward and 2 feedback taps. This results in the adaptive Kalman DFE requiring 123 multiplications and 99 additions/subtractions per iteration to carry out the tests performed in this report. The adaptive Kalman equaliser required 47 multiplications and 37 additions/subtractions and is clearly less complex, the conventional DFE using the Godard Kalman required 133 operations per iteration which is comparable with the adaptive Kalman equaliser.

It is also worth remembering that more computationally efficient implementations are possible by utilising standard matrix algebra techniques as has been demonstrated in [9].

PERFORMANCE RESULTS

The results presented in this section detail the performance of the three equalisers on time varying channels. The channel has 3 time varying taps, generated by filtering random white noise sequence through a 2nd order filter as in [10], the data rate being 100kbit/s and a Doppler spread of 100Hz, in all cases the signal-to-noise ratio was 50dB. The tests were carried out for 50% training, that is only 50 out of every 100 symbols is known a-priori the rest of the time the equaliser is operating in a decision-directed mode, this is more akin to normal operation of the equalisers. All tests were simulated in the 'C' language on a Sun 3/50 workstation.

It is clear from the results presented in figures 3-5 that the best performance, in terms of final MSE, is offered by the adaptive Kalman DFE structure the final MSE being lower than that achieved by the conventional DFE by some 5-10dB.

Adaptive Kalman Equaliser Complexity

no. of states = d+1

no of channel taps = M

operation	mult.	add/sub.
$\underline{V}(k/k-1)\,\underline{H}(k)$	$d(M-1)+1$	$d(M-2)$
$\underline{H}^T(k)\,\underline{V}(k/k-1)\,\underline{H}(k)+\sigma_v^2$	M	M
$x(k)-\underline{H}^T(k)\,\hat{\underline{x}}(k/k-1)$	$M-1$	$M-1$
$\underline{V}(k/k-1)\,\underline{H}(k)\,[\,\underline{H}^T(k)\,\underline{V}(k/k-1)\,\underline{H}(k)+\sigma_v^2\,]^{-1}$	$d+1$	
$\hat{\underline{x}}(k/k-1)+\underline{K}(k)\,[\,x(k)-\underline{H}^T(k)\,\hat{\underline{x}}(k/k-1)\,]$	$d+1$	d
$\underline{K}(k)\,\underline{H}^T(k)\,V(k/k-1)$	$\frac{d^2}{2}+\frac{d}{2}$	
$\underline{V}(k/k-1)-\underline{K}(k)\,\underline{H}^T(k)\,\underline{V}(k/k-1)$		$\frac{d^2}{2}+\frac{d}{2}$
$y(k)-\underline{h}^T(k-1)\,\underline{x}(k)$	M	M
$2\mu\,\underline{x}(k)\,(\,y(k)-\underline{h}^T(k-1)\,\underline{x}(k)\,)$	$M+1$	
$\underline{h}(k-1)+2\mu\,\underline{x}(k)\,(\,y(k)-\underline{h}^T(k-1)\,\underline{x}(k)\,)$		M
$(\,1-\frac{1}{M}\,)\,\hat{\sigma}_v^2(k-1)+\frac{e^2(k)}{M}$	3	1

Table 2

Adaptive Kalman DFE Complexity

no. of states = d+2

no. of channel taps = M

operation	mult.	add/sub.
$\underline{V}(k/k-1)\,\underline{H}(k)$	$2(d+1)M$	$2(d+1)(M-2)$
$\underline{H}^T(k)\,\underline{V}(k/k-1)\,\underline{H}(k)+\sigma_v^2$	$4M$	$4M-4$
$x(k)-\underline{H}^T(k)\,\hat{\underline{x}}(k/k-1)$	$2(M-1)+2(d+1)$	$2M-2$
$\underline{V}(k/k-1)\,\underline{H}(k)\,[\,\underline{H}^T(k)\,\underline{V}(k/k-1)\,\underline{H}(k)+\sigma_v^2\,]^{-1}$	$4(d+1)+4$	$2(d+1)+1$
$\hat{\underline{x}}(k/k-1)+\underline{K}(k)\,[\,x(k)-\underline{H}^T(k)\,\hat{\underline{x}}(k/k-1)\,]$	$2(d+1)$	$2(d+1)$
$\underline{K}(k)\,\underline{H}^T(k)\,V(k/k-1)$	$\frac{2(d+1)M+(d+1)^2}{2}$	$\frac{2(d+2)(M-2)+(d+1)^2}{2}$
$\underline{V}(k/k-1)-\underline{K}(k)\,\underline{H}^T(k)\,\underline{V}(k/k-1)$		$\frac{(d+1)^2}{2}$
$y(k)-\underline{h}^T(k-1)\,\underline{x}(k)$	M	M
$2\mu\,\underline{x}(k)\,(\,y(k)-\underline{h}^T(k-1)\,\underline{x}(k)\,)$	$M+1$	
$\underline{h}(k-1)+2\mu\,\underline{x}(k)\,(\,y(k)-\underline{h}^T(k-1)\,\underline{x}(k)\,)$		M
$(\,1-\frac{1}{M}\,)\,\hat{\sigma}_v^2(k-1)+\frac{e^2(k)}{M}$	3	1

Table 3

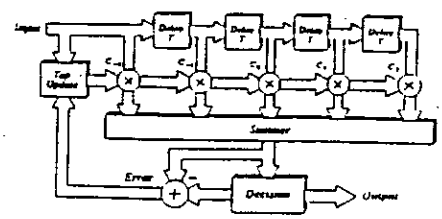


Figure 1

LINEAR EQUALISATION

Figure 2a

Godard-Kalman Decision Feedback Equaliser

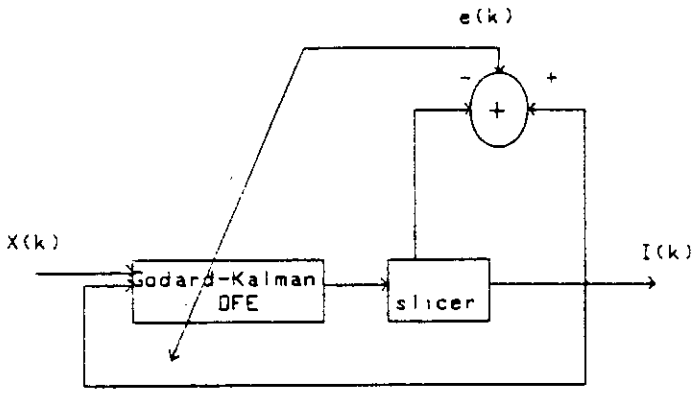


Figure 2b

Adaptive Kalman equaliser

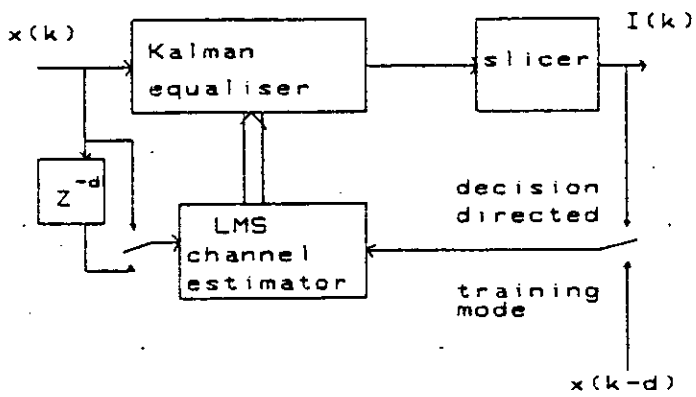
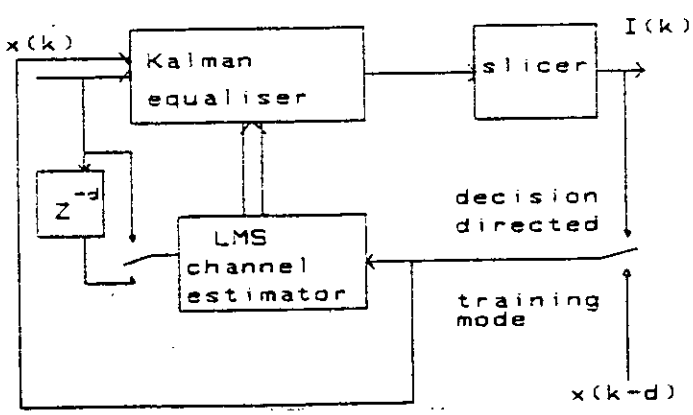


Figure 2c

Adaptive Kalman DFE



REFERENCES

1)Qureshi, S.U.H., Proc. IEEE, Vol. 73, No. 9, pp1349-1387, Sept. 1985.

2)Mulgrew, B. and Cowan, C.F.N., IEEE Trans. on ASSP, Vol. 35, No. 12, pp1727-1735, Dec. 1987.

3)Cowan, C.F.N. and Grant, P.M., Prentice Hall, 1985.

4)Eleftheriou, E. and Falconer, D.D., IEEE Trans. on ASSP, Vol. 34, No. 5, pp1097-1109, Oct. 1986.

5)Clark, A.P. and Harun, R., IEE Proc. Part F, Vol. 133, No. 6, pp513-521, Oct. 1986.

6)McLaughlin, S., Mulgrew, B. and Cowan, C.F.N., Proceedings of ICASSP '87, pp49.2.1-49.2.4, Dallas, U.S.A.

7)de Oliveira Duarte, A.M. and O'Reilly, J.J., IEE Proc. Part F, Vol. 132, No. 7, pp567-575, Dec. 1985.

8)Godard, D., IBM Journal of Research and Development, pp267-273, May 1974.

9) Anderson, B.D.O. & Moore, J.B. Prentice Hall 1979.

10) Hsu, F.M. and Giordano, A. John Wiley, 1986.