Nonlinear Adaptive Filtering
for
Echo Cancellation and Decision Feedback Equalization

by

John Frixou Michaelides
B.Eng. (Electrical)

A thesis submitted to the Faculty of Graduate Studies and
Research in partial fulfillment of the requirements
for the degree of Master of Engineering

Abstract

This thesis studies the problem of nonlinear adaptive filtering for Echo Cancellation (EC) and Decision Feedback Equalization (DFE). The high speed requirements in digital subscriber loops and voiceband data modems have put more constraints on the design of adaptive filters used for EC and DFE due to the significance of nonlinearities present in the echo and communication channels. Different configurations for nonlinear adaptive filters are developed for both echo cancellation and for combined echo cancellation and decision feedback equalization. Most of the configurations are based on the "Table Look-up" structure although the use of nonlinear filters based on the Volterra series and nonlinear compensators (involving separate adaptive linear and nonlinear parts) are also considered. Computer simulations are performed for all the configurations and most of the results are supported by theoretical derivations.

Sommaire

Cette thèse étudue le problme de filtrage adaptative nonlinèaire pour l'Annulation d'Echo (AE) et Decision Egalisation à Reaction (DER). Les exigences de grande vitesse pour le branchement d'abonnè numerique et donnèe modems de bande de fréquences vocales a mis plus de contraintes sur le dessein de filtres adaptative utilizer pour AE et DER à cause de le significaton de nonlinèairites èvident dans les canals d'èchoet de communication. Des configurations diffèrants pourfoltres adaptative nonlinèaire sont développés pour lesdeux, AE et pour AE combiner et DER. La plupart des configurations sont basés d'après le structure de la "Table de Recherche" bien que l'utilization de filtres nonlinéaire basès d'après le série de Volterra et compensateurs nonlinéaire (impliquant les pieces séparé linéaire et nonlinéaire adaptative) sont considérés en plus. Simulations par ordinateur sont exécutés pour tous les configurations et la plupart des resultats sont soutenus par dérivations théorique.

Acknowledgements

I would like to sincerely thank my supervisor Dr. Peter Kabal for his invaluable guidance, encouragement and understanding throughout this study. I am grateful to him for his financial support.

I wish to extend my gratitude to my parents for their continuous support and encouragement throughout my studies. I would also like to thank all my friends for their support and especially Anne O'Donnell for her work in translating the abstract into French as well as for her encouragement throughout this study.

The laboratory facilities provided by INRS is greatly appreciated.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

The trend towards digital networks in the telephone system and the need to provide digital transmission capability for the subscriber has prompted the investigation of the use of existing subscriber loops for transmitting digital data. For economic reasons it is desirable to transmit in both directions over the same pair of wires. Full-duplex (simultaneous two-way) digital transmission over a common medium has arisen in two important applications: (1) digital transmission on the subscriber loop, and (2) voiceband data modems.

In the first case the basic voice service as well as enhanced data services are provided over the two-wire subscriber loop. The total bit rate that has been proposed is 144 kbits/sec in each direction and it includes provision for two voice/data channels at 64 kbits/sec each plus a data channel at 16 kbits/sec. This application is an important element of the emerging integrated services digital network (ISDN) in which integrated voice and data services will be provided to the customer over a common facility. At the subscriber end, voice to be transmitted is converted from analog to digital (received voice is converted from digital to analog) through coding and filtering operations and together with the full-duplex data stream is sent over

the two-wire subscriber loop through the subscriber modem (Fig 1.1). At the central office end, the central office modem connects to the digital central office switch for voice or circuit switched data transmission, and to data networks for packet switched data transport capability.

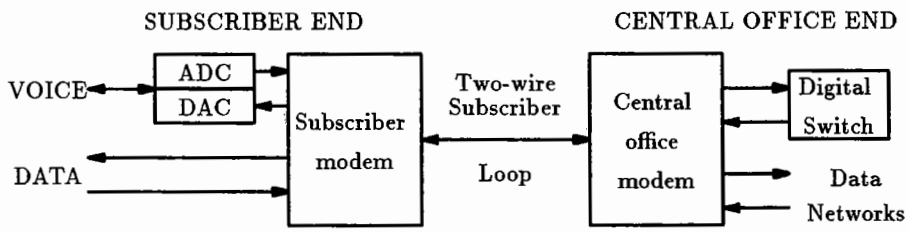


Fig. 1.1 Digital transmission on the subscriber loop (ADC=Analog to Digital Conversion, DAC=Digital to Analog Conversion) [1]

Three methods of realizing two-wire full-duplex transmission have been proposed: (1) Frequency division multiplexing (FDM) (uses separate frequency bands for the two directions), (2) Time compression multiplexing (TCM) (the two transmission directions are time-separated), and (3) Echo cancellation (EC) (uses a hybrid transformer for directional separation and an echo canceller to improve the channel separation). Echo cancellation provides greater range (distance between subscriber and central office) than the other methods. This greater range is due to the approximate halving of the transmitted signal bandwidth relative to the alternatives. During the writing of this thesis the T1D1 subcommittee of the American National Standards Institute (ANSI) has adopted EC with a zero redundancy Quaternary Line Code (2B1Q—two binary bits coded to one quarternary symbol) as a standard.

In the second application of full-duplex transmission, voiceband data modems are

used to transmit data over the telephone lines whose usable bandwidth is restricted to the frequency range of 300 Hz to 3400 Hz. In voiceband data transmission the basic customer interface to the network is often the same two-wire subscriber loop but the overall transmission link may include four-wire trunk facilities as in Fig. 1.2.

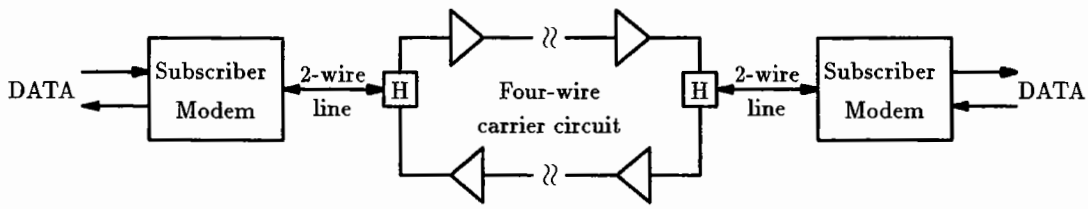


Fig. 1.2 Voiceband data transmission (H=Hybrid)

Using end-to-end analog facilities, full-duplex transmission has been provided at 2.4 kbits/sec and below using FDM. However at 4.8 kbits/sec and above, frequency separation becomes impractical due to an inadequate total bandwidth, and the use of echo cancellation must be considered. Echo cancellation makes available the full voiceband for each direction of transmission. Consequently, data transmission rates of up to 9600 bits/sec (full-duplex) can be reached over ordinary dial-up telephone lines. In March 1984 the CCITT (Consultive Committee International Telegraph and Telephone) has adopted a recommendation, V.32, for a two wire full-duplex modem family that provides for rates of 4800 and 9600 bits/sec using echo cancellation.

1.1 Echo Cancellation

Echoes are generated in most telephone connections and are the result of impedance mismatches in the communication circuit. The first form of echo con-

trol was introduced in the late 1920's in the form of a voice-activated switch called an echo suppressor [2]. An echo suppressor was installed near a four-wire to two-wire interface in order to eliminate the echoes present in the long distance telephone connections. Although a quite satisfactory solution at the time, an echo suppressor has the tendency to "chop" speech signals and it may clip out a noticeable segment at the beginning of an utterance. The problems presented by echo suppressors lead to the proposition of a new technique called echo cancellation. The idea of an echo canceller was first proposed by Sondhi [3]. Although echo cancellation was first introduced for speech transmission, in a short time it found widespread interest for application in data transmission.

1.1.1 Echo Cancellation in Data transmission

When data signals are transmitted through the telephone network the same echoes are encountered as in speech transmission. In full-duplex transmission, echoes from the data signal transmitted in one direction can interfere with the data signal flowing in the opposite direction unless these two data signals are in non-overlapping frequency bands.

Full-duplex transmission over two wire lines can be achieved by the application of the hybrid (H) which functions as four-wire to two-wire interface. A typical configuration of a full-duplex transmission system using echo cancellation is shown in Fig. 1.3 (the two-wire line may include a four-wire trunk facility in the middle of the connection as in Fig. 1.2). The hybrid gives rise to imperfect isolation between transmitter and receiver at either end ("near" and "far") of a connection. A part of

the transmitted signal directly leaks into the local receiver as a result of an imbalance within the hybrid circuit (near-end echo). Another part of the transmitted signal is reflected from impedance irregularities in the two-wire line, and ends up as part of the received signal (far-end echo). This is schematically indicated for the signal transmitted by the near-end transmitter by means of dashed arrows in Fig. 1.3.

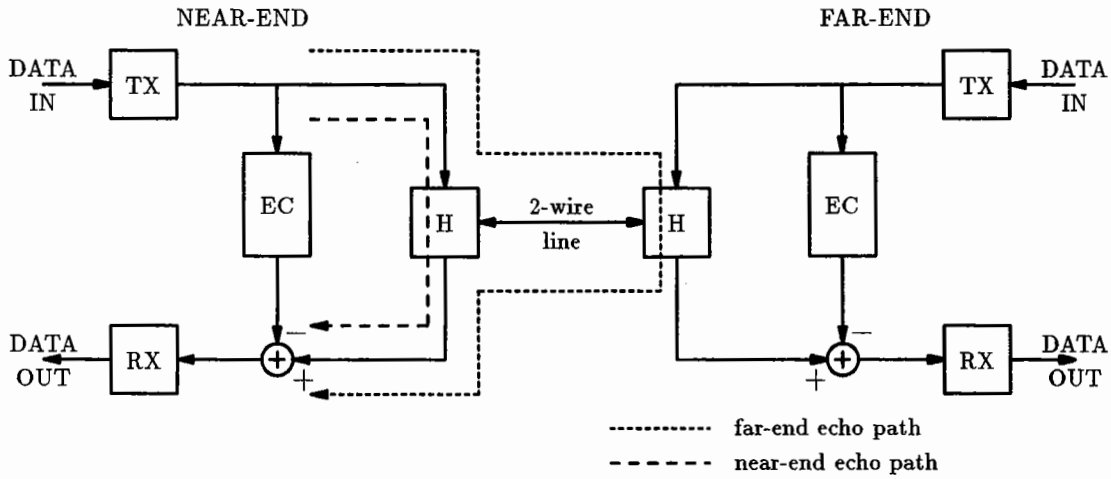


Fig. 1.3 Echo cancellation in full duplex data transmission (TX=Transmitter, RX=Receiver)

The echo signal is modeled as the result of an unintended transmission path (the "echo channel") between the transmitter and the receiver. An echo canceller is an adaptive filter that synthesizes a replica of the echo by gradually adapting its impulse response to the impulse response of the echo channel. The replica of the echo is subtracted from the echo until a very small residual echo remains.

The degree of attenuation of the leakage from transmitter to receiver of the hybrid depends on how well its balancing network matches the line impedance, which differs from line to line. In practice no better than 10 dB or so of attenuation through the hybrid can be guaranteed with a single compromise termination. In both sub-

scriber loop and voiceband data transmission the data signal coming from the distant transmitter (far-end signal) may be highly attenuated (40–50 dB). Thus the required attenuation of the near-end echo signal is of the order of 50–60 dB in order to achieve an acceptable far-end signal-to-echo interference ratio (of the order of 20 dB) at the receiver input for the maximum expected line attenuation. The far-end echo could be, in the worst case, 48 dB below the near-end transmitted signal and an attenuation (of the far-end echo) of the order of 12–22 dB is required to achieve the 20 dB far-end signal to echo interference ratio.

The canceller configurations described so far neglect the effect of nonlinear distortion in the echo path or in the echo replica. Because of the degree of cancellation required for data transmission the compensation of nonlinearities becomes an important consideration. The primary sources of nonlinearity are the data converters, the transmitted pulse asymmetry and the saturation of transformers. While these mechanisms result in a small degree of nonlinearity, they are nevertheless of importance when the objective is 50–60 dB of cancellation.

Several methods have been proposed that extend a linear echo canceller into a nonlinear one. One of the alternatives uses the "memory compensation principle" or "table look-up" method that replaces the conventional linear transversal filter with a random access memory (RAM) and can compensate for general nonlinearities [4]. The table look-up structure has received considerable attention for an ISDN environment because it uses simple signal processing and thus allows high bit rates and/or low power consumption. Another method is one that uses the Volterra series expansion to extend the linear transversal filter to a nonlinear one by the addition of extra

"nonlinear" taps [5]. The Volterra series expansion is capable of representing any echo mechanism which is time invariant. That is, the nonlinearity can have memory, but the nature of the nonlinearity cannot change with time. Of course, since the canceller is adaptive, the nonlinearity can change slowly with time, as long as the adaptation mechanism can keep up. Most of the other methods proposed are based on either the table look-up structure or on a nonlinear filter that is based on the Volterra series expansion [6][7][8][9].

1.2 Scope and Organization of the Thesis

This thesis studies the problem of nonlinear adaptive filtering for echo cancellation and decision feedback equalization. It does not deal explicitly with decision feedback equalization but it examines a configuration which combines decision feedback equalization with echo cancellation. We study previously used methods for the compensation of nonlinear channels and suggest possible modifications to them as well as new configurations in order to improve performance.

Most of the models to be presented are based on the table look-up structure although the use of nonlinear filters based on the Volterra series expansion and nonlinear compensators (involving separate linear and nonlinear parts) are also considered. Different configurations for nonlinear adaptive filters are developed for both echo cancellation and for combined echo cancellation and decision feedback equalization.

Results are derived by computer simulations and are supported by theoretical derivations. Comparisons with existing models are also made. In almost all the

simulations the most general echo and communication channels are considered so that the nonlinear adaptive filters can compensate for general nonlinearities in both digital subscriber loops and voiceband data modems applications (these general channels can be represented by the Volterra series expansion). The emphasis will be on data transmission of binary values although many of the results can be generalized to multilevel signalling.

The thesis is organized into seven chapters. After the introduction, the second chapter describes the most important sources of nonlinear distortion and gives a mathematical representation of the resulting nonlinear models. The third chapter is concerned with the table look-up method. It gives theoretical results for the algorithms used and compares them with computer simulations. The fourth chapter introduces a new two-memory structure that is used for adaptive reference echo cancellation. One of the memories is used as an echo canceller and the other is used as a reference-former to form a reference signal which is an estimate of the far-end signal. The reference-former is further exploited as a decision feedback equalizer as well, with the addition of some simple arithmetic operations. Two other alternatives for combined echo cancellation and decision feedback equalization are also suggested. The fifth chapter presents a model that is a combination of a linear FIR (finite impulse response) filter with the table look-up structure. In Chapter six a new type of compensator is introduced. This compensator consists of separate adaptive linear and nonlinear parts and compensates for a specific nonlinear channel model. Finally the seventh chapter summarizes the work to be presented, draws some conclusions, and suggests possible topics for further study on the subject.

Chapter 2 Sources of Nonlinear Distortion

Compensation of nonlinear distortion has become an important factor in echo cancellation due to the high degree of cancellation required for acceptable reception of data. In this chapter the most important sources of nonlinear distortion are described. This study will prove to be very helpful for the design of nonlinear adaptive filters for echo cancellation.

There are two signal paths in which nonlinearities may exist. These are the near-end echo path and the far-end echo path. Compensation for the nonlinearities present in the near-end echo path is more important than for the far-end echo path due to the high degree of cancellation needed for the near echo (especially for the digital subscriber loop application where large signal attenuations are present when high data rates are used). Thus more emphasis will be given to the description of nonlinearities in the near-end echo path of a digital baseband transmission system. At a later section, the nonlinearities present in the communication channel will also be described in order to examine the need for nonlinear equalization of the received signal.

2.1 Transmission System Description

Fig. 2.1 shows the general configuration of a digital baseband transmission system. It can be seen that the transmitter is part of the echo path. If some of the transmitter components are nonlinear then the near-end echo path becomes nonlinear and a nonlinear echo canceller is required. The kind of nonlinearities present in the echo path will depend on how the transmitter is implemented and how the cancellation is to be performed.

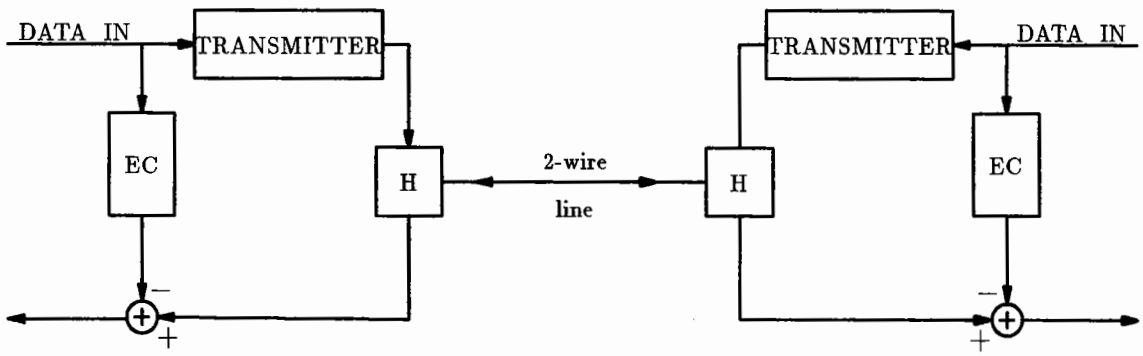


Fig. 2.1 A digital baseband transmission system

There are three alternatives, in terms of analog versus digital representation, as to how the echo and its replica can be processed. The three alternatives are schematically represented in Fig. 2.2 where

$a_k \longrightarrow$ transmitted data

$\hat{e}_t \longrightarrow$ echo estimate

$e_t \longrightarrow$ echo

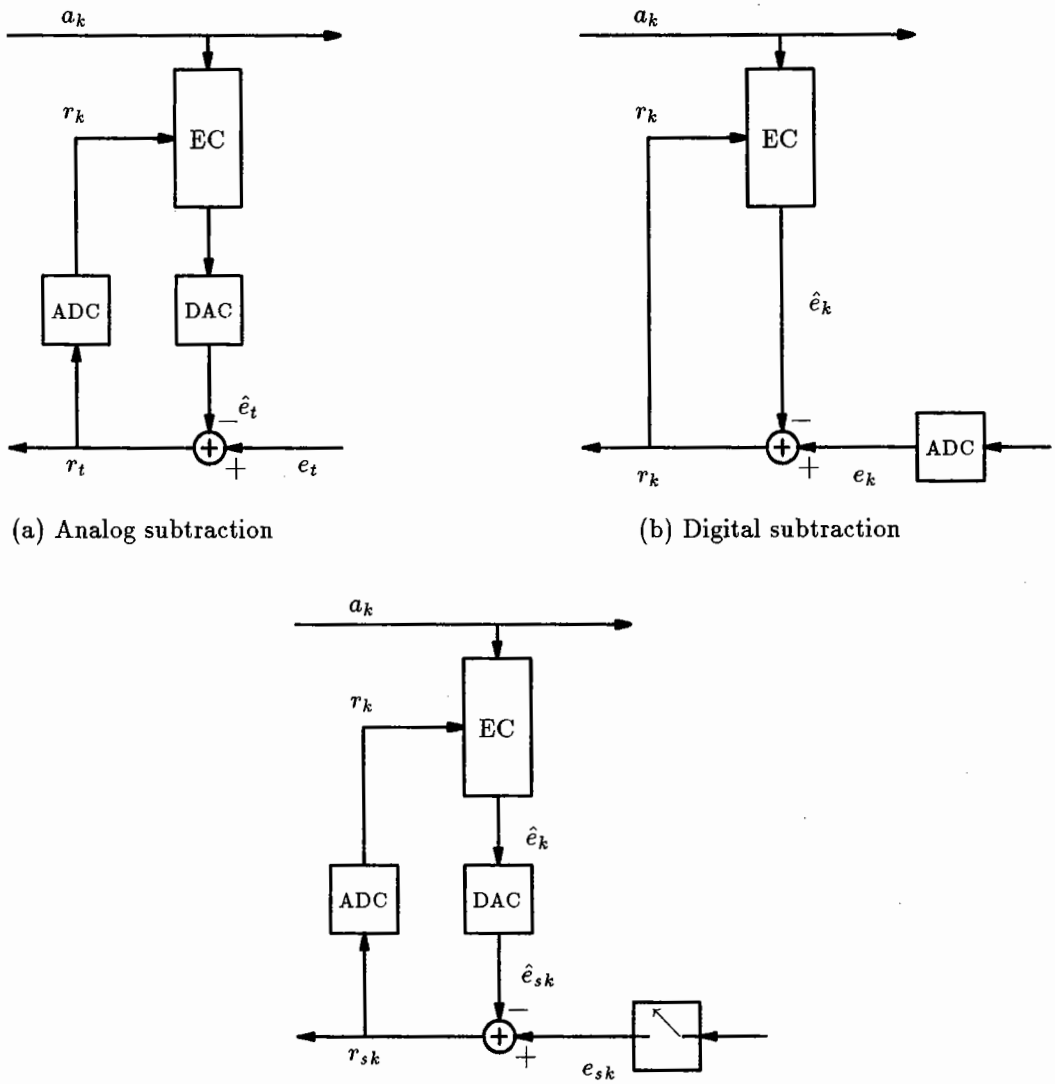
$r_t \longrightarrow$ residual signal

and are

(a) analog subtraction of the signals e_t and $\hat{e}_t$

(b) digital subtraction of the signals e_k and $\hat{e}_k$

(c) sampled data subtraction of the signals e_{sk} and $\hat{e}_{sk}$



(a) Analog subtraction

(b) Digital subtraction

(c) Sampled data subtraction

Fig. 2.2 Method of cancellation

A more detailed description of the transmission system is shown in Fig. 2.3 using digital subtraction (the method of Fig. 2.2(b)). The scrambler serves two purposes. First of all by applying different scramblers at the two directions of transmission, any

statistical dependence between the two input data streams is decreased. This helps the convergence of the echo canceller and equalization in the receiver.

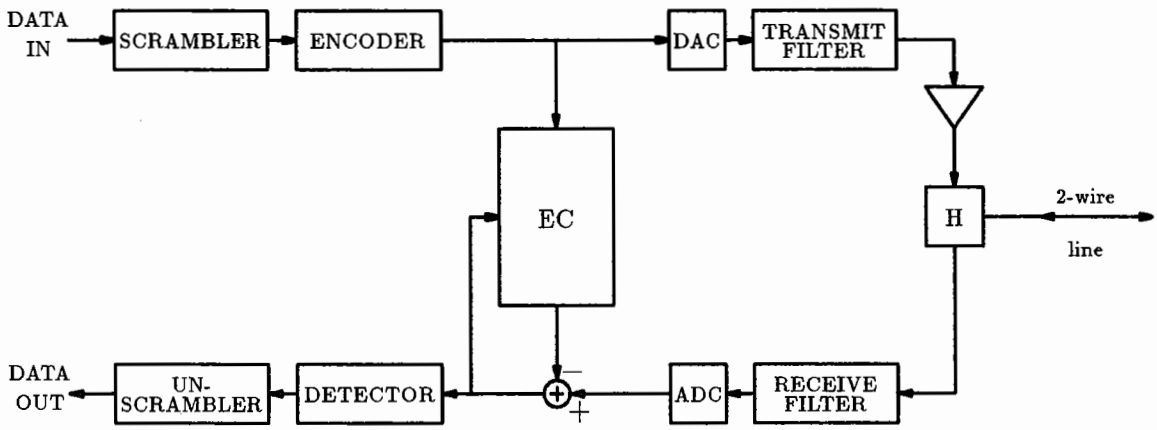


Fig. 2.3 Transmission system block diagram

The encoder may have different forms depending on the choice of the line code. "Bipolar" or "Alternate Mark Inversion" (AMI) coding has been very popular in many design considerations in view of its simplicity and robustness. However 2B1Q has been found to be a strong candidate in an ISDN environment. The transmit filter shapes the pulse placed on the line to minimize high-frequency components that would increase crosstalk and RFI (radio frequency interference). The receive filter removes the out of band noise.

2.2 Nonlinearities in the Near-End Echo Path

Many sources of nonlinearities exist in the near-end echo path. A description of the most important ones is given in this section.

2.2.1 Nonlinearity due to Transmitted Pulse Asymmetry

When nominally balanced positive and negative pulses are transmitted, a high degree of symmetry between the two pulses is required. Perfectly square pulses as input to the transmit filter are not possible because the output of a logic gate departs from that ideal due to rising edge and falling edge transients that are not equal in general. In practice there will be a slight imbalance which a linear echo canceller cannot compensate for. If a linear echo canceller is to be used then the distortion due to pulse asymmetry should be lower than approximately -60 dB. This is a level which can be achieved with careful circuit design at the cost of an increase in complexity.

It was established in [10] that the components of the transmitted data signal at multiples of the bit rate are a direct and useful measure of transmitted pulse asymmetry. It was shown that a bipolar encoded signal (which has zeros in the power spectrum at all harmonics of the bit rate) in the presence of pulse asymmetry has a line component at the bit frequency.

A representation of the transmitted pulse asymmetry for binary and quarternary signals is next given.

Binary signals

The following assumptions are made

(i) The transmitted data bits assume the values $C_k = +1$ and $C_k = -1$

(ii) A positive pulse has a shape $h_+(t)$ and a negative pulse has shape $h_-(t)$

From the above assumptions the transmitted signal can be represented as

$$x(t) = \sum_k \left[f_+(C_k) h_+(t - kT) + f_-(C_k) h_-(t - kT) \right] \tag{2.1}$$

where $f_+(C_k)$ is 1 when $C_k = 1$, and 0 otherwise, and similarly $f_-(C_k)$ is 1 only when $C_k = -1$. The functions $f_+(C_k)$ and $f_-(C_k)$ are given by

$$\begin{aligned} f_+(C_k) &= \frac{1+C_k}{2} \\ f_-(C_k) &= \frac{1-C_k}{2} \ . \end{aligned} \tag{2.2}$$

Thus Eq. (2.1) becomes

$$\begin{aligned} x(t) &= \frac{1}{2}\sum_k \Big[(1+C_k)h_+(t-kT) + (1-C_k)h_-(t-kT)\Big] \\ &= \frac{1}{2}\sum_k \Big[\Big[h_+(t-kT) + h_-(t-kT)\Big] + C_k\Big[h_+(t-kT) - h_-(t-kT)\Big]\Big] \end{aligned} \tag{2.3}$$

which shows that nonlinear distortion due to pulse asymmetry for binary (± 1) signals is in the form of a dc offset pulse that is transmitted for every data symbol (of course in the case of pulse symmetry where $h_+(\cdot) + h_-(\cdot) = 0$, this dc offset pulse is equal to zero).

Assuming equiprobable signal levels, the average value of the transmitted signal can be found from Eq. (2.3) to be

$$E[x(t)] = \frac{1}{2}\sum_k \Big[(h_+(t-kT) + h_-(t-kT)\Big] \ . \tag{2.4}$$

Then the total signal can be written as a zero mean random component plus the deterministic mean component of the above equation. The random component has a continuous power spectrum while the deterministic component consists of a line spectrum since it is periodic in T seconds. Expanding Eq. (2.4) in a Fourier series,

$$E[x(t)] = \frac{1}{2T}\sum_n \Big[H_+(n\frac{2\pi}{T}) + H_-(n\frac{2\pi}{T})\Big]e^{j\frac{2\pi nt}{T}} \ . \tag{2.5}$$

Thus the transmitted signal contains line components at multiples of the bit rate due to the pulse asymmetry. Taking into account the effect of the transmit filter,

the only significant component will be that at the bit rate, which is proportional to the difference of the Fourier transforms of the positive and negative pulses at that frequency.

Quarternary Signals

In this case the following assumptions are made

(*i*) The transmitted data C_k assume the values $\pm 1, \pm 3$

(*ii*) The pulses transmitted for the 4 signal values are $h_i(t)$ for $i = \pm 1, \pm 3$

(*iii*) The case of pulse symmetry corresponds to

$h_{+1}(t) = -h_{-1}(t),\ h_{+3}(t) = -h_{-3}(t),\ h_{+1}(t) = h_{+3}(t)/3,\ h_{-1}(t) = h_{-3}(t)/3$

The corresponding transmitted signal is given by

$$x(t) = \sum_k \Big[f_{+1}(C_k)h_{+1}(t-kT) + f_{-1}(C_k)h_{-1}(t-kT) + f_{+3}(C_k)h_{+3}(t-kT) + f_{-3}(C_k)h_{-3}(t-kT)\Big] \tag{2.6}$$

where

$$f_i(C_k) = \begin{cases} 1 & C_k = i \\ 0 & C_k \neq i \end{cases} \tag{2.7}$$

An explicit form of $f_i(C_k)$ is

$$f_i(C_k) = \quad h_0 + h_1 C_k + h_2 C_k^2 + h_3 C_k^3 \ . \tag{2.8}$$

Using Eq. (2.7) and by assigning all possible values of C_k in Eq. (2.8) we can find $h_0,\ h_1,\ h_2$ and h_3 for each $f_i(C_k)$ as given below

$$\begin{aligned} f_{+1}(C_k) &= \frac{9}{16} + \frac{9}{16}C_k - \frac{1}{16}C_k^2 - \frac{1}{16}C_k^3 \\ f_{+3}(C_k) &= -\frac{3}{48} - \frac{1}{48}C_k + \frac{3}{48}C_k^2 + \frac{1}{48}C_k^3 \\ f_{-1}(C_k) &= \frac{9}{16} - \frac{9}{16}C_k - \frac{1}{16}C_k^2 + \frac{1}{16}C_k^3 \\ f_{-3}(C_k) &= -\frac{3}{48} + \frac{1}{48}C_k + \frac{3}{48}C_k^2 - \frac{1}{48}C_k^3 \end{aligned} \tag{2.9}$$

By substituting Eq. (2.9) into Eq. (2.6) we get

$$\begin{aligned} x(t) = \frac{1}{16}\sum_k \Big(& \big[h_{+1}(t-kT) - h_{-1}(t-kT)\big]\big[9C_k - C_k^3\big] \\ & + \frac{1}{3}\big[h_{+3}(t-kT) - h_{-3}(t-kT)\big]\big[C_k^3 - C_k\big] \\ & + \big[h_{+1}(t-kT) + h_{-1}(t-kT)\big]\big[9 - C_k^2\big] \\ & + \frac{1}{3}\big[h_{+3}(t-kT) + h_{-3}(t-kT)\big]\big[3C_k^2 - 3\big] \Big) \end{aligned} \tag{2.10}$$

which shows that nonlinear distortion due to pulse asymmetry includes dc offset pulses, as well as pulses that are proportional to second and third powers of the transmitted data symbol. For equiprobable signal levels the average value of the transmitted signal is

$$\begin{aligned} E[x(t)] = \frac{1}{4}\sum_k \Big(& \big[h_{+1}(t-kT) + h_{-1}(t-kT)\big] \\ & + \big[h_{+3}(t-kT) + h_{-3}(t-kT)\big] \Big) \end{aligned} \tag{2.11}$$

and by the same reasoning as for Eq. (2.4), expanding Eq. (2.11) in a Fourier series,

$$\begin{aligned} E[x(t)] = \frac{1}{4T}\sum_n \Big(& \big[H_{+1}(n\tfrac{2\pi}{T}) + H_{-1}(n\tfrac{2\pi}{T})\big]e^{j\frac{2\pi nt}{T}} \\ & + \big[H_{+3}(n\tfrac{2\pi}{T}) + H_{-3}(n\tfrac{2\pi}{T})\big]e^{j\frac{2\pi nt}{T}} \Big) . \end{aligned} \tag{2.12}$$

Thus in this case (as for binary signals) the transmitted signal contains line components of the bit rate due to the pulse asymmetry. Again taking into account the effect of the transmit filter, the only significant components will be those at the bit rate, which are proportional to the differences of the positive and the negative pulses at that frequency.

2.2.2 Nonlinearities due to the Hybrid Transformer

As mentioned in the introduction the hybrid is used to turn a two-wire transmission facility into an equivalent four-wire connection. Data can then be transmitted

simultaneously in both directions. Hybrids used in telephony are made up of inductive elements in a balanced bridge configuration, and are passive four-port directional couplers.

The ideal hybrid passes the signal to be transmitted from port A to port C with some attenuation, but does not allow any of the signal to enter port B (Fig. 2.4). The received signal passes from port C to port B (again with some attenuation) but no energy of the signal is reflected to port A. This performance is achieved only if the balancing impedance (in port D) is equal to the impedance presented by the two-wire line (assuming a unity ratio hybrid). Because of the different impedances presented by different subscriber lines (due in part to different lengths, bridged taps etc.), an impedance mismatch occurs and some transmitted signal from port A will enter port B.

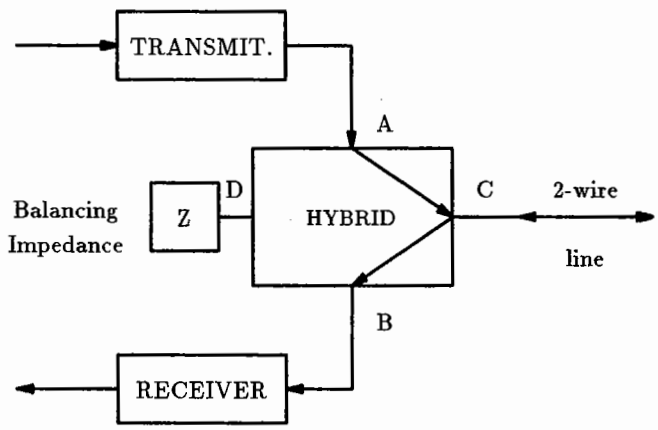


Fig. 2.4 The hybrid as a four-port directional coupler

A typical hybrid circuit shown in more detail is presented in Fig. 2.5 [11]. The purpose of the bridge is to generate a potential at point 1 which is equal to that

of point 2 from the potential divider formed by the transmitter output impedance and the cable impedance. Any remaining interference is the difference between the potential at 1 and 2 which is the return loss of the impedance presented to the cable relative to Z_1.

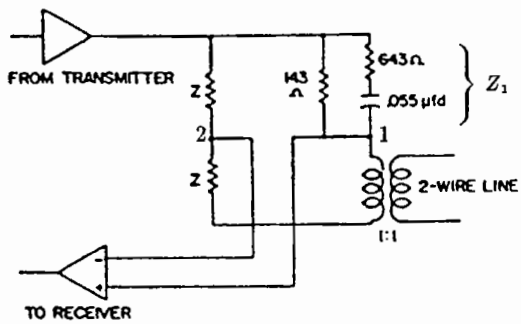


Fig. 2.5 A typical hybrid transformer circuit [11]

It is well known that in a fundamental sense ferromagnetic transformers are all nonlinear. It has proved convenient, because of tradition and mathematical simplicity of linear circuit analysis, to treat the iron-core transformer as if it were a linear circuit element, but in practice only careful design of the transformer allows this assumption to be made. Thus the relationship between flux ϕ and the magnetizing current i_m or the relationship between the flux density B and the magnetizing force H is not linear. The graphical plot (B versus H) exhibits, in addition to saturation, a hysteresis characteristic. Four different curves are shown in Fig. 2.6.

Over a limited range of voltage, frequency and temperature and with careful design, the distortion produced by the inherent nonlinearities does not disturb a

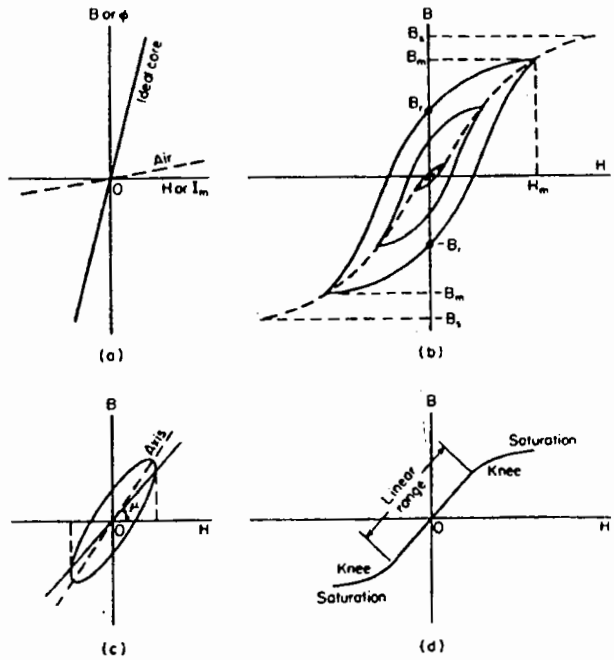


Fig. 2.6 B-H curves. (a) Ideal linear core with high constant slope and no memory B_r. (b) A family of hysteresis curves for soft ferromagnetic core material. B_m is maximum flux density, B_s is saturation flux density. (c) Ellipse approximation of a hysteresis curve, when eddy-current losses dominate. (d) A single-valued curve with nonlinear knee and saturation regions. Hysteresis is neglected. [12]

circuit function. Thus the nonlinearities presented by the hybrid are not as significant as other sources of nonlinearity in the echo path and are usually ignored.

2.2.3 Nonlinearities due to the Data Converters

The data converters are the most important source of nonlinearities particularly where monolithic converters without trimming are to be employed. As described in Section 2.1, the cancellation can be performed digitally, in sampled data form or in

analog form. Depending on the way the cancellation is performed a DAC (Digital to Analog Converter) and/or an ADC (Analog to Digital Converter) is part of the echo path. The nonlinearities present in these converters are described next.

2.2.4 Nonlinearities in Analog to Digital Conversion

The parameters that characterize an ADC are

(1) offset error

(2) gain error

(3) integral nonlinearity

(4) differential nonlinearity

(5) monotonicity

Fig. 2.7 shows the transfer function of a 3-bit ADC. Curve (B) is the transfer function of an ideal converter while curve (A) corresponds to a real converter having integral nonlinearity error. Curve (C) shows a missed code (non-monotonicity) due to excessive differential nonlinearity.

The integral nonlinearity error is defined as the maximum deviation of the transfer curve of the real ADC converter to that of the ideal converter. This error is due to the mismatch of the components. This mismatch could be due to any combination of the following factors: (1) voltage/temperature coefficients of the components (2) sheet resistivity/oxide thickness gradients across the chip (3) random location of the geometrical edges defining each component due to uncertainty in photolithography and/or chip processing. Any of these factors causes a shift in the transition voltages of the transfer curve of the ADC and thus produces nonlinearity. In an ideal ADC

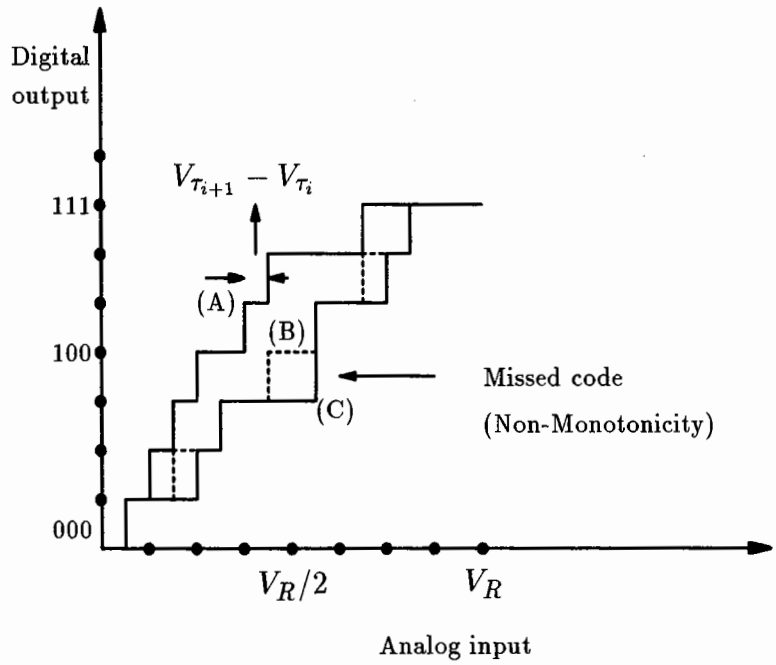


Fig. 2.7 The transfer function of a 3-bit ADC. Curve (B) depicts the ideal case and curves (A) and (C) depict the effect of integral nonlinearity and excessive differential nonlinearity on the ideal transfer function respectively [13]

any two adjacent transition voltages are separated by 1 LSB (least significant bit). The parameter that is usually used to describe the error in this separation is the differential nonlinearity (DNL) defined by the following equation

$$\text{DNL}_i = 1 \text{ LSB} - (V_{\tau_{i+1}} - V_{\tau_i}) \tag{2.13}$$

where DNL_i is the differential nonlinearity at the i_{th} transition and V_{τ_i} is the actual value of the i_{th} transition voltage. For an ideal ADC the difference $V_{\tau_{i+1}} - V_{\tau_i}$ is equal to 1 LSB. Thus, the differential nonlinearity of an ideal converter is zero. For a real ADC however, the component mismatch will cause $V_{\tau_{i+1}} - V_{\tau_i}$ to deviate from

1 LSB and thus will introduce differential nonlinearity.

If the component mismatch is severe, $V_{\tau_{i+1}} - V_{\tau_i}$ can approach zero; i.e. the $(i+1)_{st}$ and i_{th} transitions will coincide, forcing DNL_i to approach 1 LSB. When the DNL_i exceeds 1 LSB, the ADC converter is said to be nonmonotonic (at the i_{th} transition; see curve C in Fig. 2.7). Nonmonotonicity implies that there are one or more missing codes at the output and is major flaw for an ADC. Nonmonotonicity also increases quantization noise.

2.2.5 Nonlinearities in Digital to Analog Conversion

Specific nonlinearities are described in [5] using as an example a DAC to be implemented in MOS technology and using the technique shown in Fig. 2.8.

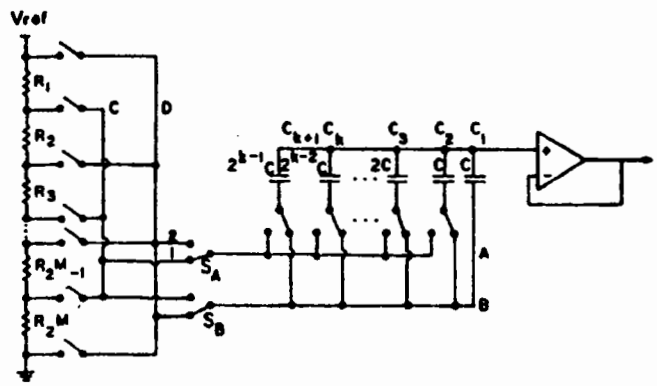


Fig. 2.8 DAC using resistor string and capacitor array (switch control is omitted for simplicity).[5]

Because of the diffusion concentration gradients, voltage coefficient, and photolithographic mismatches the resistors cannot be guaranteed to be equal to within

1 LSB (least significant bit) unless laser trimming is used. Thus in the absence of trimming a nonlinear transfer characteristic results.

Typical DAC transfer functions are shown in Fig. 2.9. In the first case the most important terms in the Volterra series expansion (described in a later section) are the first and third order terms whereas in the second case the first and second order terms predominate.

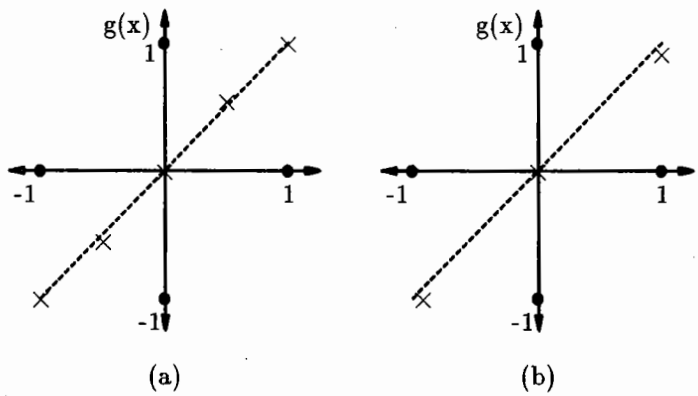


Fig. 2.9 Typical DAC transfer functions
(a) $g(x) = 1.01333x - 0.01333x^3$ (b) $g(x) = x - 0.005|x|$
($\times$ shows a point on the transfer function) [5]

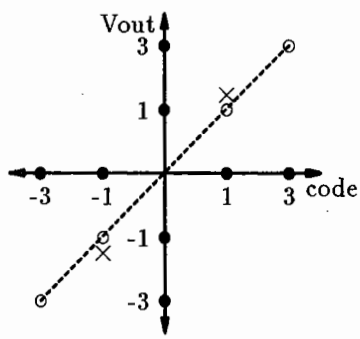


Fig. 2.10 DAC transfer function for 4 level signalling ($\times$ denotes actual transfer values which deviate from linearity)

A typical DAC transfer function for quaternary levels of $\pm 1, \pm 3$ (most probably to be used in an ISDN environment) is shown in Fig. 2.10.

2.3 Nonlinearities in the Communication Channel

In the previous sections we examined the different sources of nonlinearities in the near-end echo path. Nonlinear distortion may also be a characteristic of the communication channel itself. The main concern is the presence of nonlinearities in the far-end data signal. Nonlinearities in the far-end echo are not considered. This is because (1) the far-end echo path includes the communication channel and in general the same kind of nonlinearities will exist in the two cases and (2) nonlinear compensation may not be necessary due to the low degree of cancellation required for the far-end echo.

In the case of digital subscriber loops where the communication channel is the two-wire line, factors which affect the channel characteristics include gauge changes and *bridged taps* (open circuited cable pairs bridged onto the main cable pair) as well as impairments like impulse noise and crosstalk [14]. Nonlinearities may exist due to the transmitted pulse asymmetry, the hybrids, and the ADC in the receiver path, but generally they are not as important as the other kinds of distortion.

The voiceband data modem, while transmitting at a lower speed, encounters many more impairments since the communication channel in this case may include the public switched telephone network. Sources of nonlinear distortion are the ADC and DAC present throughout the telephone network as well as harmonic distortion in amplifiers and repeaters.

The effect of nonlinear distortion in the communication channel is to introduce nonlinear intersymbol interference and to reduce the margin against noise. Individual sources of channel nonlinearity may usually be characterized as memoryless. However, these nonlinear components are embedded in a network where linear filtering operations also take place. Consequently the overall effect of the channel on the input signal is a nonlinearity with memory. The channel may be modelled as a linear dispersion followed by a static nonlinearity which in turn is followed by a second linear dispersion as in Fig. 2.11.

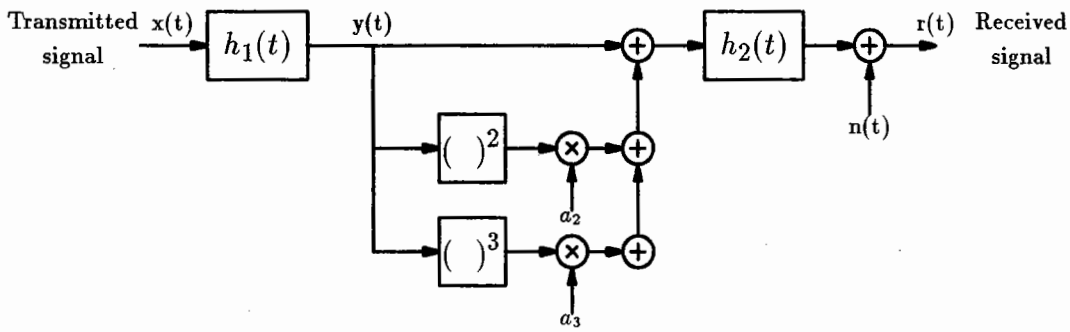


Fig. 2.11 Nonlinear channel model

Filters 1 and 2 have the same bandwidth as the transmitted data signal. The filters may include the input filter at the receiver as well as the linear response of the channel. The quadratic and cubic static nonlinearities with attenuated outputs account for second and third harmonic distortion respectively (usually the most important). The above model can adequately be described by the Volterra series. An additional impairment not shown in Fig. 2.11 is phase jitter. Measurements of the echo parameters including nonlinear distortion were made for voiceband telephone channels in [15] [16] [17].

Nonlinear equalization would improve performance of the most common form of receiver which uses linear equalization and/or decision feedback equalization and thus cannot compensate for nonlinearities. Nonlinear decision feedback equalization is examined in Chapter 4.

2.4 Mathematical Representation of Nonlinear Models

In the previous sections of this chapter, the different configurations for echo cancellation were presented, and the nonlinearities appearing were identified and described. In this section a mathematical representation of the resulting nonlinear models is given.

The Volterra series is perhaps the most used simplification of the Wiener representation of a nonlinear system. In its general form it represents a true subset of the Wiener representation, and is a generalization of a Taylor series expansion with memory. Assuming a finite-memory causal system with input sequence $\{x_k\}$ and output sequence $\{y_k\}$ the relationship between input and output as described by the Volterra series is

$$\begin{aligned} y_k = h^{(0)} + \sum_{i_1=0}^{N-1} h_{i_1}^{(1)} x_{k-i_1} + \sum_{i_1=0}^{N-1}\sum_{i_2=0}^{N-1} h_{i_1 i_2}^{(2)} x_{k-i_1} x_{k-i_2} + \cdots \\ \cdots + \sum_{i_1=0}^{N-1} \cdots \sum_{i_n=0}^{N-1} h_{i_1 \ldots i_n}^{(n)} x_{k-i_1} \cdots x_{k-i_n} \end{aligned} \tag{2.14}$$

where N is the memory length of the system and n is the order of nonlinearity. The factors $h^{(0)}, h_0^{(1)}, h_1^{(1)}, \ldots$ represent the **Volterra coefficients** and describe the system behaviour.

A nonlinear model that adequately describes the Volterra series is the general 3-block system of Fig. 2.12. The problem of system identification for the 3-block system of Fig. 2.12 reduces to identifying the finite set of the Volterra coefficients up to a specified order of nonlinearity n. The Volterra series could adequately describe the nonlinear channel model of Fig. 2.11 in Section 2.3 with $n = 3$ since only 2^{nd} and 3^{rd} order nonlinearities were considered.

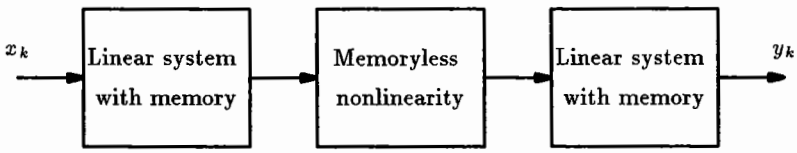


Fig. 2.12 General 3-block nonlinear model

The Volterra series, as described by the general 3-block nonlinear model, may be simplified in three stages which are of direct relevance to the nonlinear models that result from the near-end echo path. The first type of degradation is a memoryless nonlinearity by itself

$$y_k = \sum_{j=0}^{M-1} \alpha_j x_k^j \ .$$

The second degradation occurs when a system is considered which is a memoryless nonlinearity followed by a linear dispersion as in Fig. 2.13 (the Hammerstein model [18]). This system can model the nonlinear distortion due to pulse asymmetry and any linear filtering that follows.

Let us assume that the linear system with memory represents an FIR (finite

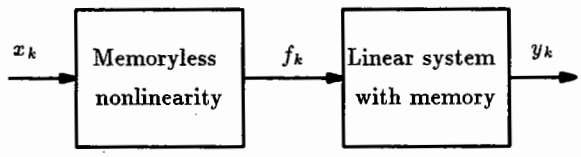


Fig. 2.13 The Hammerstein model

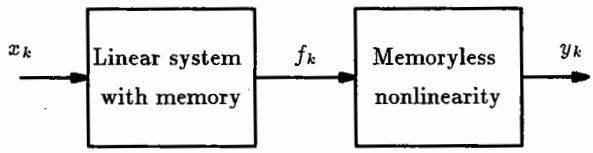


Fig. 2.14 A 2-block nonlinear model (linear-nonlinear)

impulse response) filter with $\{w_i\}$ as the impulse response coefficients so that

$$y_k = \sum_{i=0}^{N-1} w_i f_{k-i} \tag{2.15}$$

and that the memoryless nonlinearity can be parameterized in a polynomial form with parameters $\{\alpha_j\}$ so that

$$f_k = \sum_{j=0}^{M-1} \alpha_j x_k^j \ . \tag{2.16}$$

Then the input/output relation of this model is given by

$$y_k = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} w_i \alpha_j x_{k-i}^j \ . \tag{2.17}$$

Comparing with Eq. (2.14) it can be seen that a relation exists between the $\{h\}, \{w\}$ and $\{\alpha\}$ and that Eq. (2.17) is a special case of Eq. (2.14) where no cross products of the input sequence exist.

The third type of degradation is a linear system with memory followed by a memoryless nonlinearity as shown in Fig. 2.14.

In this case the input/output relationship is given by

$$y_k = \sum_{j=0}^{M-1} \alpha_j f_k^j \tag{2.18}$$

where

$$f_k = \sum_{i=0}^{N-1} w_i x_{k-i} \tag{2.19}$$

the $\{w_i\}$ represent the impulse response coefficients of a linear FIR filter and the $\{\alpha_j\}$ are the nonlinearity parameters. This representation can be used to model the nonlinearity presented by the DAC in the echo canceller path when using analog or sampled data subtraction as in Figures 2.2 (a) and 2.2 (c) respectively.

In the case where digital subtraction is used, both DAC and ADC are present in the near echo path (see Fig. 2.3), and ignoring the nonlinearities of the hybrid transformer, the nonlinear model of Fig. 2.15 emerges.

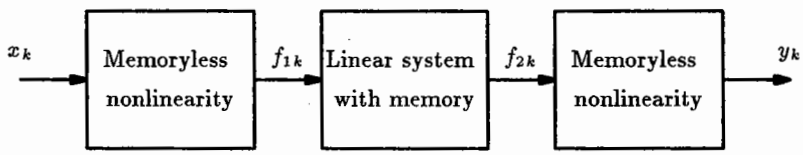


Fig. 2.15 A 3-block nonlinear model (nonlinear-linear-nonlinear)

The input/output relations at the different stages are

$$\begin{aligned} f_{1k} &= \sum_{i=0}^{M-1} \alpha_{1i} x_k^i \\ f_{2k} &= \sum_{j=0}^{N-1} w_j f_{1k-j} = \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} w_j \alpha_{1i} x_{k-j}^i \\ y_k &= \sum_{l=0}^{L-1} \alpha_{2l} f_{2k}^l \end{aligned} \tag{2.20}$$

where the $\{w_j\}$ are the impulse response coefficients of a linear FIR filter representing the linear system with memory, the $\{\alpha_{1i}\}$ are the parameters of the first nonlinearity, and the $\{\alpha_{2l}\}$ are the parameters of the second nonlinearity. An echo canceller that compensates for the nonlinearities presented by this model is described in Chapter 5.

Chapter 3 Table Look-up Structure

The table look-up structure has received considerable attention the last five years in the problem of echo cancellation. It was first introduced by N. Holte and S. Stueflotten [4] in 1981 based on the idea of designing an echo canceller that needs simple signal processing and thus allows high bit rates and/or low power consumption which makes it suitable for digital subscriber lines. Before the table look-up structure is described, a review of the conventional echo canceller structure is made. This review will be useful for reasons of comparison.

3.1 Conventional Echo Cancellation

The typical filter structure in most of the echo cancellation design is the tapped delay line or transversal filter structure shown in Fig. 3.1. Data converters will be omitted for simplicity for the rest of this thesis. The transversal filter is characterized by an impulse response of finite duration. It consists of a set of $N-1$ delay elements (each of which is represented by a unit delay operator z^{-1}) and a corresponding set of N adjustable coefficients (usually called the tap coefficients).

Two basic procedures take place in the adaptive filter

1. The *adaptive process*, which involves the automatic adjustment of the coefficients of the transversal filter in accordance with some algorithm.

2. The *filtering process*, which involves (1) multiplying the tap inputs ($a_k, a_{k-1} \cdots, a_{k-N+1}$) by the corresponding set of coefficients ($h_{0k}, h_{1k} \cdots, h_{N-1k}$) resulting from the adaptive process to produce an estimate of the echo $\hat{e}_k$ and (2) generating a residual signal r_k by subtracting the echo estimate from the received signal s_k. This residual signal is in turn used to actuate the adaptive process.

The signals u_k and n_k represent the far-end signal and additive noise respectively.

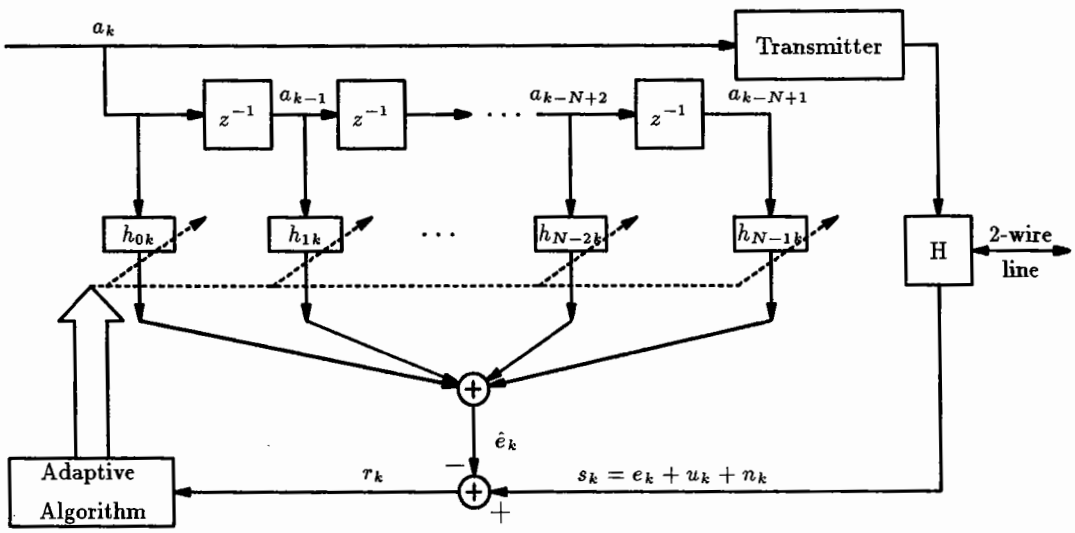


Fig. 3.1 Adaptive transversal filter structure

In the Wiener theory, the minimum mean square error criterion is used to optimize the filter. Specifically, the coefficients of the filter are chosen so as to minimize the mean square value of the estimation error (the estimation error being $e_k - \hat{e}_k$). For the case of stationary inputs, the mean squared error is precisely a second order function of the coefficients of the filter. Thus the dependence of the mean squared error on the coefficients can be viewed as a quadratic surface with a unique minimum. This

surface is usually referred to as the *error performance surface* of the adaptive filter. The adaptive process has the task of continuously seeking the bottom or the minimum of this surface.

3.1.1 Adaptation Process

Various algorithms have been developed for implementing the adaptive process. The most widely used algorithm is known as the least-mean-square (LMS) algorithm (Widrow and Hoff, 1960) or as it is often called the stochastic iteration algorithm (SIA). A significant feature of the stochastic iteration algorithm is its simplicity; it does not require measurements of the pertinent correlation functions, nor does it require matrix inversion. The stochastic iteration algorithm is a modification of the method of steepest descent, a well known technique in optimization theory. Successive corrections to the coefficients are in the direction of the negative of the gradient vector (i.e. in the direction of the steepest descent of the error performance surface) so that iterations eventually approach to the minimum mean squared error, at which point the coefficients approach their optimum values.

The value of the gradient depends on two parameters: (1) the correlation matrix of the tap inputs in the transversal filter and (2) the cross-correlation vector between the received signal and the same tap inputs.

In reality however, exact measurements of the gradient vector are not possible and the gradient vector must be obtained from the available data. According to the stochastic iteration algorithm the instantaneous values of the correlation matrix and the cross-correlation vector are used to derive an estimate for the gradient algorithm.

The adaptation equation for the stochastic iteration algorithm thus reduces to

$$\mathbf{h}_{k+1} = \mathbf{h}_k + 2\ \alpha\ r_k\ \mathbf{a}_k \qquad \text{SIA} \tag{3.1}$$

where α is a stepsize parameter and

$$\mathbf{h}_k = (h_{0k}, h_{1k}, \cdots, h_{N-1k})^T$$

$$\mathbf{a}_k = (a_k, a_{k-1}, \cdots, a_{k-N+1})^T$$

$$r_k = s_k - \hat{e}_k$$

$$= u_k + n_k + e_k - \hat{e}_k$$

(boldface represents a vector and T denotes transposition).

The initial value used for the tap weight vector, $\mathbf{h}_0$, is usually equal to zero. The major limitations of the stochastic iteration algorithm are a relatively slow rate of convergence and a sensitivity to variations in the eigenvalue spread, defined as the ratio of the maximum to minimum eigenvalues of the correlation matrix of the tap inputs.

A modification of the stochastic iteration algorithm is the sign algorithm (SA) where the sign of the residual signal r_k is taken instead of its actual value. The motivation for doing this is simplification of the hardware implementation. It has been shown in [19] that the convergence of the algorithm is not compromised by this simplification, but that the speed of adaptation is reduced. The adaptation equation for the SA is

$$\mathbf{h}_{k+1} = \mathbf{h}_k + 2\ \alpha\ \text{sgn}(r_k)\ \mathbf{a}_k \qquad \text{SA}\ . \tag{3.2}$$

3.1.2 Convergence of the Stochastic Iteration Algorithm

A detailed analysis of the convergence properties of the stochastic iteration algorithm for echo cancellation was done in [20] using data values of ± 1. If R_k^2 is defined as the ratio of the residual echo power to the far-end signal power as given below

$$\begin{aligned} R_k^2 &\stackrel{\triangle}{=} \frac{\varepsilon_k}{\sigma_u^2} \\ &= \frac{E\left[(e_k - \hat{e}_k)^2\right]}{E\left[(u_k)^2\right]} \end{aligned} \tag{3.3}$$

the equation governing the convergence of the echo canceller (starting at $k = 0$) is

$$R_k^2 = \rho^k (R_0^2 - \frac{\alpha N}{1 - \alpha N}) + \frac{\alpha N}{1 - \alpha N} \tag{3.4}$$

where ρ is

$$\rho = (1 - 4\alpha + 4\alpha^2 N) \tag{3.5}$$

The convergence can adequately be described by two parameters R_∞^2 and ν_{20}, where

$$R_\infty^2 = \lim_{k \to \infty} R_k^2 \tag{3.6}$$

and

$$\nu_{20} = \frac{1}{\log_{10}(R_k / R_{k+1})} \quad \text{for} \quad R_k^2 \gg R_\infty^2 \tag{3.7}$$

R_∞^2 characterizes the converged state and ν_{20} represents the number of iterations required for a 20 dB reduction of R_k^2. From Eq. (3.4) R_∞^2 and ν_{20} are given by

$$R_\infty^2 = \frac{\alpha N}{1 - \alpha N} \quad \text{or} \quad R_{\infty,\text{dB}}^2 = 10 \log_{10} \frac{\alpha N}{1 - \alpha N} \tag{3.8}$$

and

$$\nu_{20} = \frac{-2}{\log_{10}(1 - 4\alpha + 4\alpha^2 N)} \quad \text{or} \quad \nu_{20} \approx \frac{1.15}{\alpha} \quad \text{for} \quad \alpha \ll \frac{1}{N} \tag{3.9}$$

Equations (3.8) and (3.9) show the importance of the stepsize α. Its value determines both the rate of convergence and the minimum residual echo that can be obtained.

3.2 Table Look-up Structure

The Table Look-up structure constitutes a different type of an adaptive filter. If the echo impulse response is assumed to be of finite duration, the echo e_k will be a function of only the last N transmitted data symbols a_k. Because the received signal is digital, the echo cancellation needs, in principle, only to be performed at the detection instants. Thus there exists only a finite set (2^N for binary signals) of possible echo samples at the cancellation instants, which can be stored in a digital memory. The correct compensation value is selected by means of an address for a random access memory (RAM), formed by the last N transmitted symbols. Because all possible echo estimates are stored directly in a memory the table look-up method can even compensate for nonlinearities in the echo channel. A two-wire full-duplex baseband digital transmission system using the table look-up method is shown in Fig. 3.2.

3.2.1 Algorithms for the Adaptation Process

Assuming that the two data streams a_k and b_k (b_k are the data transmitted from the far-end transmitter) are uncorrelated, the control signal for the adaptive processor can be taken from the residual signal (Fig. 3.2)

$$r_k = u_k + n_k + e_k - \hat{e}_k \ . \tag{3.10}$$

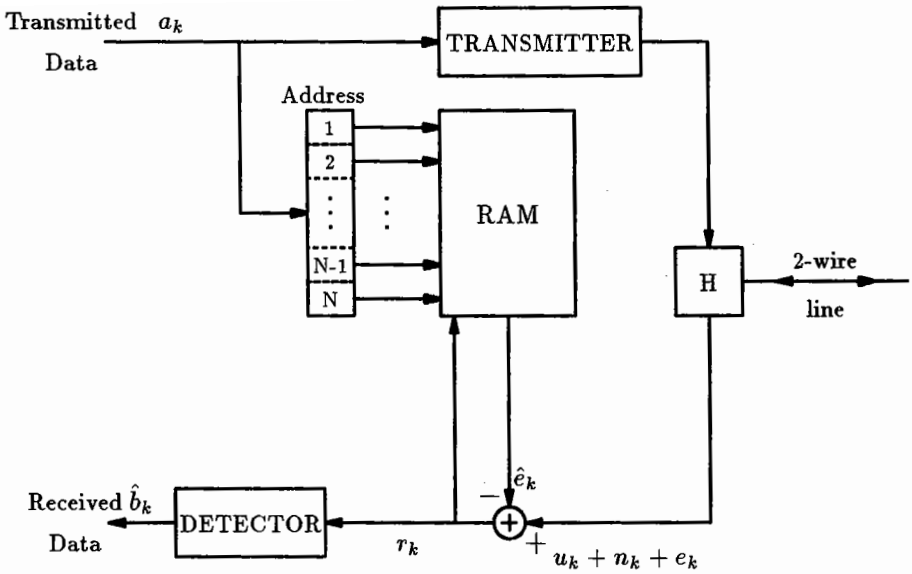


Fig. 3.2 Echo Cancellation using the table look-up method

Each canceled sample r_k contains control information for one corresponding echo estimate $\hat{e}_k$. As with the transversal filter there is a choice between two algorithms for the adaptation process. These are

(1) Stochastic iteration algorithm (SIA)

(2) Sign algorithm (SA)

The update of the echo estimate according to the stochastic iteration algorithm is given by

$$\hat{e}_{k+1} = \hat{e}_k + \alpha\, r_k \qquad \text{SIA} \tag{3.11}$$

where α is the stepsize.

The sign algorithm uses the following update of the echo estimate

$$\hat{e}_{k+1} = \hat{e}_k + \alpha\ \text{sgn}(r_k) \qquad \text{SA} \tag{3.12}$$

where α is the step size of the registers in the memory.

3.2.2 Convergence using the Stochastic Iteration Algorithm

A combination of N binary input data are used as the address for the RAM. The contents of this address is the compensation signal $\hat{e}_k$. If a certain address is read at time k, its content is subsequently updated with a correction term, while the contents of the other addresses remain unchanged. For the stochastic iteration algorithm the correction term is equal to αr_k. As was shown in [7] the convergence is governed by the relation (starting at $k = 0$)

$$R_k^2 = \rho^k R_0^2 + \frac{\alpha^2}{2^N} \frac{1-\rho^k}{1-\rho} \tag{3.13}$$

where

$$\rho = 1 - \frac{\alpha(2-\alpha)}{2^N} \ .$$

For $\alpha \ll 1$, R_∞^2 and ν_{20} are given by

$$R_\infty^2 \approx \frac{\alpha}{2} \quad \text{or} \quad R_{\infty,\text{dB}}^2 \approx 10 \log_{10} \frac{\alpha}{2} \tag{3.14}$$

and

$$\nu_{20} \approx 2.30 \ \frac{2^N}{\alpha} \ . \tag{3.15}$$

3.2.3 Convergence using the Sign Algorithm

As for the stochastic iteration algorithm a combination of N binary input data are used as the address for the RAM. Since in this case we take the sign of r_k instead of the actual value, the correction term for an address is $+\alpha$ or $-\alpha$. The convergence

of the sign algorithm is governed by

$$R_{k+1}^2 = R_k^2 - \frac{2\alpha}{2^N \sigma_u} \Psi(R_k) + \frac{\alpha^2}{2^N \sigma_u^2} \tag{3.16}$$

where

$$\Psi(x) = -2x \int_{-\infty}^{\infty} F_u(-xy\sigma_u)\ y\ p_\epsilon(y)\ dy$$

$F_u(u)$ is the cumulative distribution of u_k and $p_\epsilon(\epsilon)$ is the probability density function (PDF) of the residual echo $\epsilon_k = e_k - \hat{e}_k$. Eq. (3.16) † was derived using results from [7] and [21]. Several cases can be distinguished for different PDF's of the received signal u_k.

Binary Case

For the binary case and by assuming that the residual echo has a gaussian PDF,

$$p_\epsilon(\epsilon) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\epsilon^2}{2}}$$

then

$$\Psi(R_k) = \sqrt{\frac{2}{\pi}} R_k e^{-\frac{1}{2R_k^2}}$$

and Eq. (3.16) becomes

$$R_{k+1}^2 = R_k^2 - \frac{2\alpha}{2^N \sigma_u} \sqrt{\frac{2}{\pi}} R_k e^{-\frac{1}{2R_k^2}} + \frac{\alpha^2}{2^N \sigma_u^2}\ . \tag{3.17}$$

For large values of R_k a simple approximation to Eq. (3.17) can be derived using the fact that

$$\Psi(R_k) \approx \sqrt{\frac{2}{\pi}}\ R_k \qquad \text{for}\ \ R_k \gg 1$$

† The same equation was derived in a recent paper in the IEEE Transactions on Circuits and Systems, March 1987 with title "Multistage RAM: An FIR Filter for Echo Cancellation in a Digital Two-Wire Subscriber Loop", pp. 225-232.

and Eq. (3.17) becomes

$$\begin{aligned} R_{k+1}^2 &= R_k^2 - \frac{2\alpha}{2^N \sigma_u}\sqrt{\frac{2}{\pi}} R_k + \frac{\alpha^2}{2^N \sigma_u^2} \\ &= \left[R_k - \frac{\alpha}{2^N \sigma_u}\sqrt{\frac{2}{\pi}}\right]^2 + \frac{\alpha^2}{2^N \sigma_u^2} - \left[\frac{\alpha}{2^N \sigma_u}\sqrt{\frac{2}{\pi}}\right]^2 \quad \text{for} \quad R_k \gg 1 \end{aligned} \tag{3.18}$$

and for $\alpha \ll 1$

$$R_{k+1} = R_k - \frac{\alpha}{2^N \sigma_u}\sqrt{\frac{2}{\pi}} \qquad \text{for} \quad R_k \gg 1,\ \alpha \ll 1 \tag{3.19.a}$$

$$R_k = R_0 - \frac{\alpha}{2^N \sigma_u} k \sqrt{\frac{2}{\pi}} \qquad \text{for} \quad R_k \gg 1,\ \alpha \ll 1 \tag{3.19.b}$$

For small values of R_k, Eq. (3.17) becomes

$$R_{k+1}^2 = R_k^2 + \frac{\alpha^2}{2^N \sigma_u^2} \qquad \text{for} \quad R_k \ll 1 \tag{3.20.a}$$

$$R_k^2 = R_0^2 + k\frac{\alpha^2}{2^N \sigma_u^2} \qquad \text{for} \quad R_k \ll 1 \tag{3.20.b}$$

Non-Binary Case

As for the binary case the above analysis can be carried out for other cases of u_k. If u_k has a gaussian or a uniform distribution then the convergence properties for large R_k are given from Eq. (3.18) (same as the binary case). For small values of R_k however, and for α small the following approximation can be made for the last two cases

$$\Psi(R_k) \approx \gamma\ R_k^2 \qquad \text{for} \qquad R_k \ll 1 \tag{3.21}$$

where γ is a constant depending on the PDF of u_k. Then Eq. (3.16) becomes

$$R_{k+1}^2 = \rho^k R_0^2 + \frac{\alpha^2}{2^N {\sigma_u}^2}\frac{1-\rho^k}{1-\rho} \quad \text{for} \quad R_k \ll 1 \tag{3.22}$$

where

$$\rho = 1 - \frac{2\alpha\gamma}{2^N \sigma_u}$$

The parameters R_∞^2 and ν_{20} in this case are

$$R_\infty^2 \approx \frac{\alpha}{2\sigma_u \gamma} \tag{3.23}$$

and

$$\nu_{20} \approx 2.30 \frac{2^N \sigma_u}{\alpha\gamma} \tag{3.24}$$

3.2.4 Comparison of the two Algorithms

From the results obtained for the stochastic iteration algorithm and the sign algorithm, we can see that there is a distinct difference between the two. This is due to the fact that the convergence properties of the stochastic iteration algorithm depend only on the stepsize α where for the sign algorithm the convergence properties depend on the PDF of the received signal u_k as well. Thus we are not able to directly compare the two algorithms and we have to distinguish between different cases for the sign algorithm based on the PDF of u_k.

For large values of R_k, the three cases examined for the sign algorithm (binary, gaussian, uniform) behave almost the same as given by Eq. (3.18). For small values of R_k, the uniform and gaussian case behave in a similar manner (Eq. (3.22)) and low values of R_∞ are obtainable. However a different behavior is found for the binary case (Eq. (3.20)).

If the sign algorithm is used for binary signals, R_∞ decreases very slowly with decreasing values of α. The reason for this is that if u_k is binary, then $\mathrm{sgn}[(e_k - \hat{e}_k) + u_k] = \mathrm{sgn}(u_k)$ for almost all k if $e_k - \hat{e}_k$ has become small. Then we expect

the convergence of an echo canceller using the sign algorithm to terminate as soon as the residual error becomes permanently smaller than the amplitude of u_k. However R_∞ still decreases somewhat with decreasing α and this is due to the assumption usually made in the derivation of the convergence equations that the residual error has a gaussian PDF and the tails of this PDF assure that occasionally large values of $e_k - \hat{e}_k$ can still be expected. The speed of convergence is much slower for all cases for the sign algorithm than for the stochastic iteration algorithm if the same value of R_∞ is to be reached.

No matter which algorithm is being used (the stochastic iteration algorithm or the sign algorithm), the table look-up method will need 2^N coefficients rather than N as compared with a linear transversal filter. Thus as it can also be seen from the above results, for large N the required memory becomes very large and the adaptation very slow (an FIR filter with N coefficients is $\frac{2^N}{N}$ times faster). This is the price that has to be paid if compensation of nonlinear distortion is required.

3.2.5 Computer Simulations

In this section computer simulations of the table look-up structure are presented. The near-end and far-end data are uncorrelated ± 1 values. The echo channel is represented by a 3-block nonlinear model (linear– nonlinear–linear) and the communication channel as a 2-block nonlinear model (linear–nonlinear) as shown in Fig. 3.3. The static nonlinearities are third order polynomials ($a_0 + a_1(\cdot) + a_2(\cdot)^2 + a_3(\cdot)^3$) and the ratio of near-end to far-end signal power is 60 dB. The number of bits used to address the RAM (echo canceller) is 6 and all the memory locations of the RAM are

initially set to zero. The coefficients and nonlinearity parameters used for the echo and communication channel are shown in Tables 3.1 and 3.2 respectively.

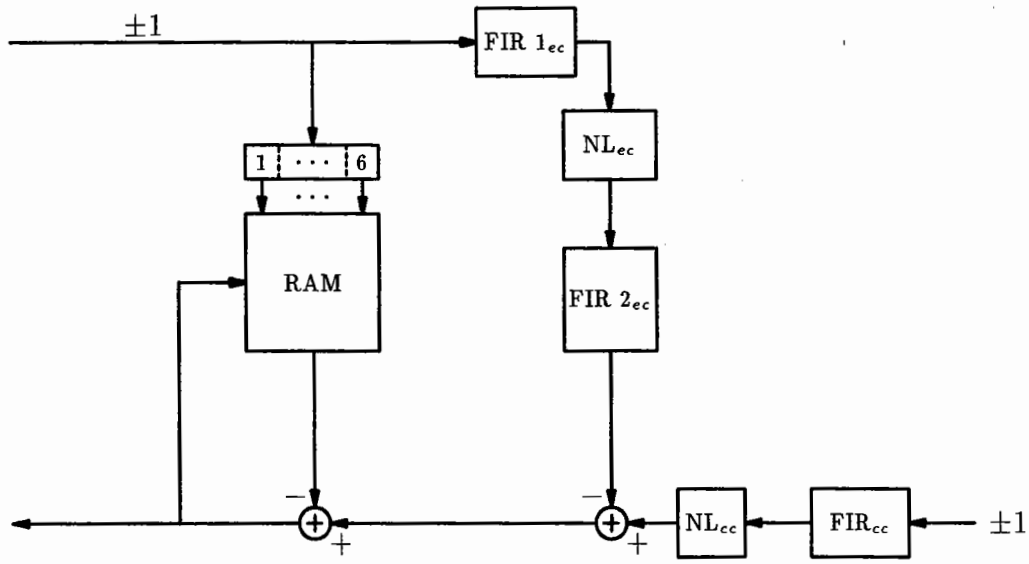


Fig. 3.3 Simulation model for the table look-up structure

NEAR END ECHO CHANNEL										
FIR 1				NL				FIR 2		
h_0	h_1	h_2	h_3	a_0	a_1	a_2	a_3	h_0	h_1	h_2
1.0	0.15	0.07	0.01	0.0	1.0	0.0	0.02	1.0	0.1	0.01

Table 3.1 FIR coefficients and nonlinearity parameters for the near-end echo channel

COMMUNICATION CHANNEL							
FIR 1				NL			
h_0	h_1	h_2	h_3	a_0	a_1	a_2	a_3
1.0	0.1	0.05	0.01	0.0	1.0	0.0	-0.013

Table 3.2 FIR coefficients and nonlinearity parameters for the communication channel

3.2.6 Convergence curves

Fig. 3.4 compares the convergence curves for the sign algorithm obtained from Eq. (3.17) with computer simulations for the binary case and Fig. 3.5 compares the theoretical convergence of the sign algorithm using the exact representation (Eq. (3.17)) and the approximation of Eq. (3.19). Each point in the simulation curves is the sample mean of 100 successive squared errors and normalized by the far-end signal power. The convergence curves show that the simulation results correspond reasonably well with theory. The continued decrease of curve (b) in Fig. 3.5 after iteration 11,000 is expected since this curve approximates the exact representation only for large values of R_k^2.

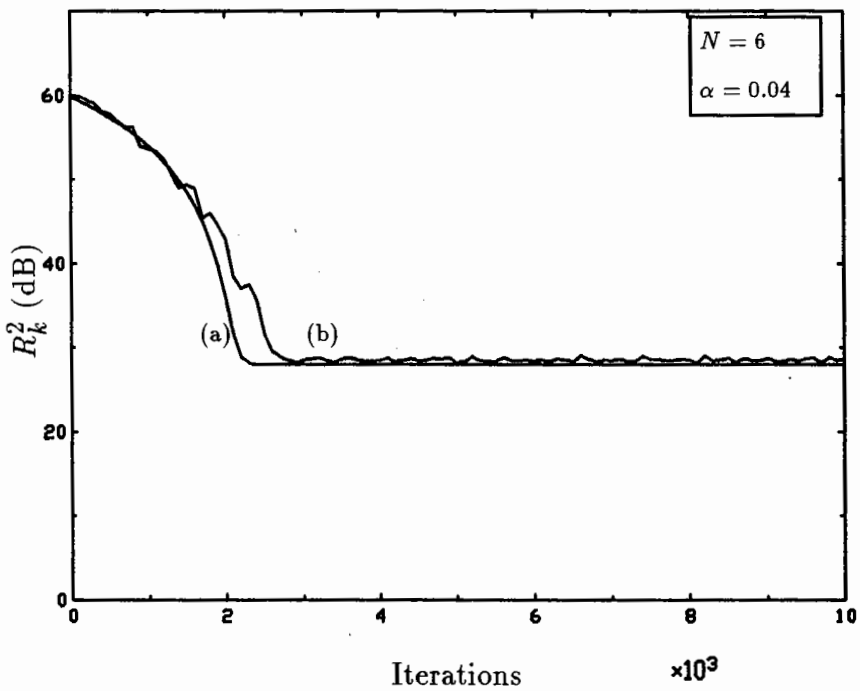


Fig. 3.4 Theory (a) vs Computer simulation (b) for the sign algorithm using the table look-up method

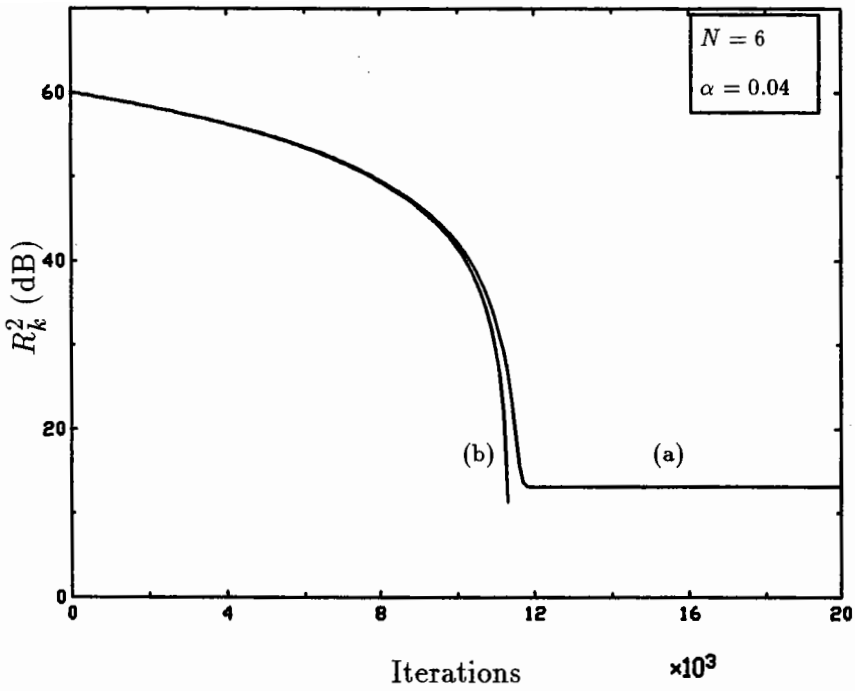


Fig. 3.5 Theoretical convergence of the sign algorithm (a) exact (Eq. (3.17)) (b) approximation (Eq. (3.19))

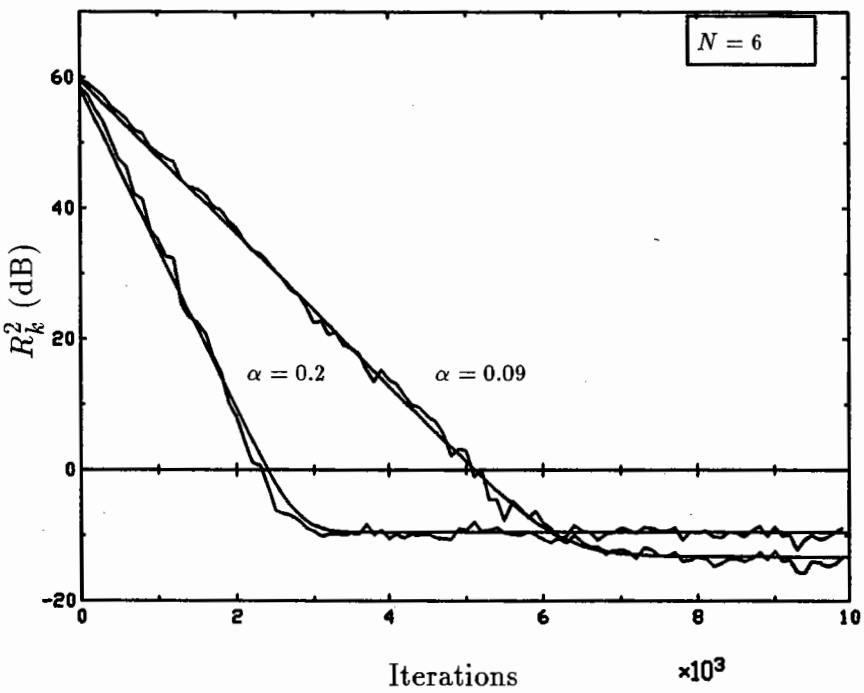


Fig. 3.6 Theory vs Computer simulation for the stochastic iteration algorithm using the table look-up method

Fig. 3.6 compares the convergence curves for the stochastic iteration algorithm according to Eq. (3.13) with computer simulations. The theoretical curve agrees well with the one obtained from computer simulations. $R^2_\infty = -10.0$ dB when the stepsize is 0.2 and -13.47 dB when it is 0.09. ν_{20} is 736 when the stepsize is 0.2 and 1635 when it is 0.09. All the above values are those predicted from Eq. (3.14) and Eq. (3.15).

Chapter 4 The Two-Memory Structure

The presence of the far-end signal in the control signal that is used for the echo canceller adaptation algorithm is a limiting factor in the convergence characteristic. If no far-end signal is present, the cancelled signal should ideally be zero. Because of the disturbing effect of the far-end signal on the adaptation algorithm, conventional echo cancellers (using FIR filters) require a very small stepsize for averaging during full-duplex transmission and hence converge relatively slowly. The same holds if a table look-up structure is used. The aim is to reduce this dependence on the far-end signal. The scheme to be used is called adaptive reference echo cancellation [22].

With adaptive reference echo cancellation, a reference signal which is an estimate of the far-end signal is formed by a reference-former, and is subtracted from the control signal of the echo canceller adaptation algorithm. The reference-former used in [22] was an FIR filter. In this chapter the reference-former uses the table look-up structure.

4.1 Echo Canceller Configuration

In order to understand the problem, the configuration of Fig. 4.1 is first consid-

ered. As it can be seen, the control signal of the echo canceller r_k includes the received far-end data signal u_k. The signal u_k is the output of a transmission path which has as input the far-end data signal b_k (n_k is additive white noise). This transmission path includes transmitter filtering, coding operations and the communication channel. The problem is then the identification of this transmission path which is similar to the problem of echo cancellation. A possible solution is to use a second canceller that will try to cancel the signal u_k from r_k as shown in Fig. 4.2.

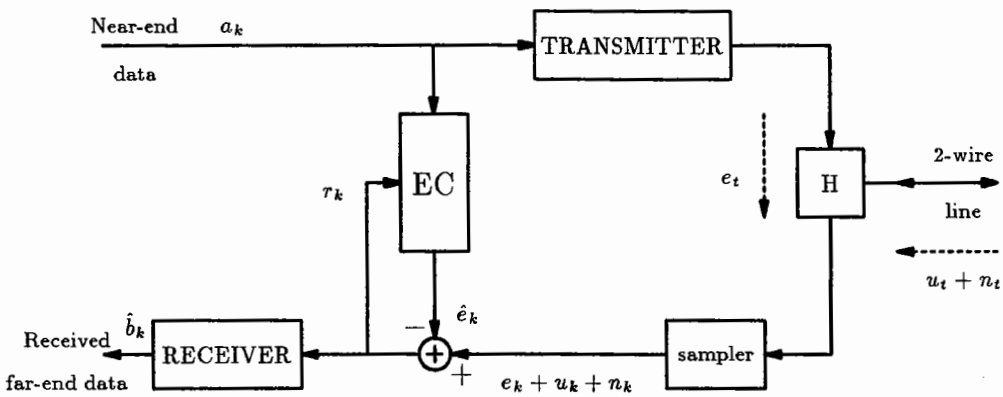


Fig. 4.1 Echo cancellation in the presence of the far-end signal

The purpose of the second memory is to adaptively form a reference signal $\hat{u}_k$ which is an estimate of u_k. The reference signal $\hat{u}_k$ is then subtracted from the signal r_k. The resulting signal r'_k is then used for the adaptation of the echo canceller (RAM1) and of the reference-former (RAM2). As described in Chapter 2, the transmission path through which the far-end data signal passes, may be nonlinear. Thus the table look-up structure for the reference-former will also compensate for any nonlinearities present in the communication channel which distort the received signal, and limit the degree of cancellation that can be reached if an FIR filter is used.

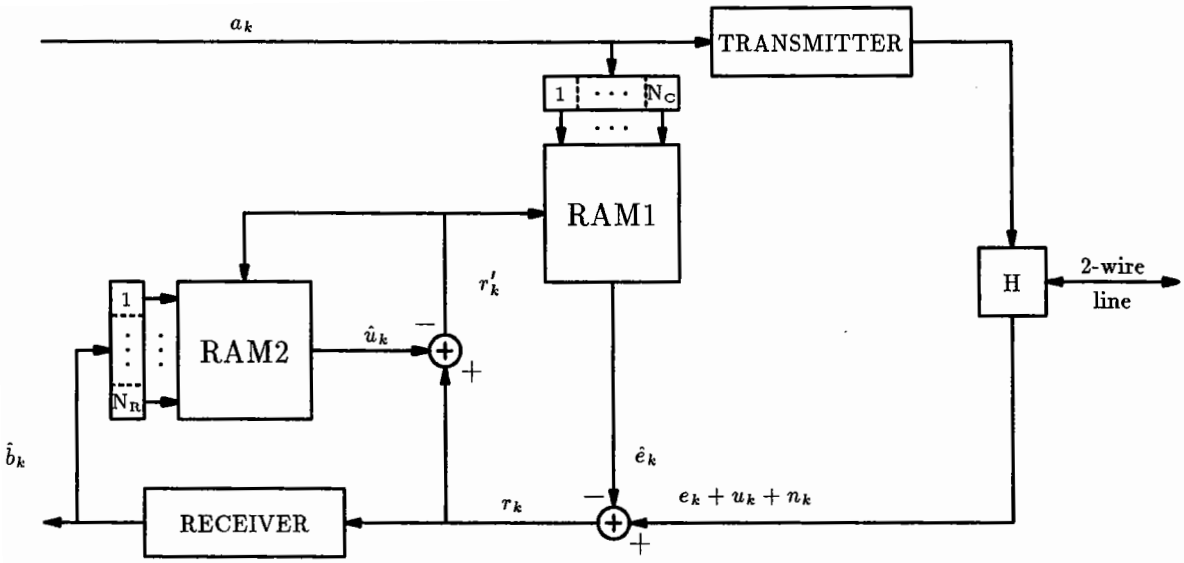


Fig. 4.2 Adaptive Reference echo cancellation using the table look-up structure

The number of bits used to address the echo canceller is chosen so as to span the longest expected echo impulse response, and the number of bits used to address the reference-former is chosen to span the longest expected communication channel impulse response. A delay may be included before subtraction of the reference-former signal $\hat{u}_k$ from the signal r_k because the decisions may be delayed relative to those present in r_k due to receiver filtering, block decoding, etc. (depends on the receiver structure).

The a_k and b_k are again assumed to be uncorrelated and take on the values ± 1. The additive noise n_k results mainly from near-end crosstalk (NEXT) and impulse noise and is assumed to be uncorrelated with the data. The signals e_k and u_k are given from Eq. (4.1) and Eq. (4.2) respectively.

$$e_k = \sum_{i=0}^{\infty} a_{k-i} h_i \tag{4.1}$$

$$u_k = \sum_{i=0}^{\infty} b_{k-i} g_i \tag{4.2}$$

where the h_i are the impulse response coefficients of the near-end echo channel and the g_i are the impulse response coefficients of the communication channel. In case where both channels are nonlinear the above equations take the form of a Volterra series as in Eq. (2.14).

4.2 Adaptation Algorithms

As in Chapter 3 two algorithms that could be used to adapt the two memories are the stochastic iteration algorithm, and the sign algorithm. The adaptation equations are given by

$$\textbf{RAM1} \qquad \hat{e}_{k+1} = \hat{e}_k + \alpha_C \; r'_k \qquad \text{SIA} \tag{4.3}$$

$$\hat{e}_{k+1} = \hat{e}_k + \alpha_C \, \text{sgn}(r'_k) \qquad \text{SA} \tag{4.4}$$

$$\textbf{RAM2} \qquad \hat{u}_{k+1} = \hat{u}_k + \alpha_R r'_k \qquad \text{SIA} \tag{4.5}$$

$$\hat{u}_{k+1} = \hat{u}_k + \alpha_R \, \text{sgn}(r'_k) \qquad \text{SA} \tag{4.6}$$

where α_C and α_R are the stepsizes used for the echo canceller and the reference-former respectively. In most cases the stochastic iteration algorithm is preferred because it converges faster than the sign algorithm.

4.3 Convergence using the Stochastic Iteration Algorithm

The error signal for the echo canceller is $e_k - \hat{e}_k$ and for the reference-former is $u_k - \hat{u}_k$. The control signal for both the echo canceller and the reference-former r'_k

is given by

$$\begin{aligned} r_k' &= r_k - \hat{u}_k \\ &= e_k - \hat{e}_k + u_k - \hat{u}_k + n_k \ . \end{aligned} \tag{4.7}$$

As is shown in Appendix A and by using the following definitions,

$$\begin{aligned} \varepsilon_{C,k} &\triangleq E[(e_k - \hat{e}_k)^2] \\ \varepsilon_{R,k} &\triangleq E[(u_k - \hat{u}_k)^2] \end{aligned} \tag{4.8}$$

the relation governing the convergence of the echo canceller is given by

$$\varepsilon_{C,k+1} = \varepsilon_{C,k}\rho_C + \frac{\alpha_C^2}{2^{N_C}}\varepsilon_{R,k} + \frac{\alpha_C^2}{2^{N_C}}\sigma_n^2 \tag{4.9}$$

and for the reference-former is given by

$$\varepsilon_{R,k+1} = \varepsilon_{R,k}\rho_R + \frac{\alpha_R^2}{2^{N_R}}\varepsilon_{C,k} + \frac{\alpha_R^2}{2^{N_R}}\sigma_n^2 \tag{4.10}$$

where

$$\begin{aligned} \rho_C &= 1 - \frac{\alpha_C(2-\alpha_C)}{2^{N_C}} \\ \rho_R &= 1 - \frac{\alpha_R(2-\alpha_R)}{2^{N_R}} \ . \end{aligned} \tag{4.11}$$

We notice that the convergence of the echo canceller is dependent on the error signal of the reference-former. If the reference-former is not used then $\hat{u}_k$ is zero and Eq. (4.9) reduces to

$$\varepsilon_{C,k+1} = \varepsilon_{C,k}\rho_C + \frac{\alpha_C^2}{2^{N_C}}\sigma_u^2 + \frac{\alpha_C^2}{2^{N_C}}\sigma_n^2 \ . \tag{4.12}$$

If we normalize $\varepsilon_{C,k}$ by σ_u^2 then Eq. (4.12) is similar to Eq. (3.13) which gives the convergence equation using only one memory, with the exception that here additive noise is included. The steady state value for the adaptive reference cancellation scheme is

$$\varepsilon_{C,\infty} = \frac{\alpha_C\sigma_n^2}{2-\alpha_C-\alpha_R} \ . \tag{4.13}$$

From the above equation we see that there would be little penalty by making $\alpha_C = \alpha_R$ since small changes of α_R do not affect the value of $\varepsilon_{C,\infty}$ very much (as long as α_R is greater than zero, the steady state value reached with two memories will be lower than with one memory).

4.4 Decision Feedback Equalization

Decision feedback equalization can be a simple and effective way to mitigate intersymbol interference, especially on subscriber loops with one or more bridged taps where conventional linear adaptive equalization methods may perform poorly. It works on the principle of subtracting the portion of the far-end signal due to previously detected far-end symbols from the received signal.

Decision feedback equalization is performed in most of the cases using an FIR filter. Gerwen *et al.* [7] have realized the DFE (decision feedback equalizer) using a table look-up structure. They have also combined the echo canceller and the DFE using a single memory.

In Section 4.1 a second memory was added to the echo canceller configuration so that it could cancel the far-end data signal from the control signal used in the adaptation of the echo canceller. A third memory could be used as a DFE but this would increase complexity and cost. Thus, in this section the reference-former (RAM2) is exploited and is used also as a DFE with the addition of some simple arithmetic operations.

4.4.1 Relation between the memory contents of the RAM and the Volterra coefficients

The adaptive reference-former was used as a means of identifying the communication channel. If the communication channel is modelled as an FIR filter then this would result in the identification of the tap coefficients which would be the impulse response coefficients of the channel. A subset of the impulse response coefficients cause intersymbol interference (ISI) from previously detected far-end symbols. When a table look-up structure is used the impulse response coefficients are not readily available. The goal is to relate these coefficients with the memory contents of the RAM and use them to cancel intersymbol interference.

Starting from the Volterra series expansion, a new representation of the input/output relationship of a nonlinear channel can be found (Appendix B) in the case of binary signals. This representation has a finite number of terms and is given by

$$f(b_k, b_{k-1}, \ldots, b_{k-N+1}) = g^{(0)} + \sum_{i=0}^{N-1} g_i^{(1)} b_{k-i} + \sum_{i_1 \neq i_2}^{N-1} g_{i_1 i_2}^{(2)} b_{k-i_1} b_{k-i_2} + \cdots$$
$$\cdots + \sum_{i_1 \neq i_2 \cdots \neq i_{N-1}}^{N-1} g_{i_1\ i_2 \ldots i_{N-1}}^{(N-1)} b_{k-i_1}\ b_{k-i_2} \cdots\ b_{k-i_{N-1}} + g^{(N)} b_k b_{k-1} \cdots b_{k-N+1} \tag{4.14}$$

where $f(b_k, b_{k-1}, \ldots, b_{k-N+1})$ represent the contents of one of the 2^N memory locations of the RAM and the g's represent the Volterra coefficients in the new representation (in the case of multilevel transmission, a finite number of terms in the Volterra series expansion is also required). The address of the RAM is specified by $b_k, b_{k-1}, \ldots, b_{k-N+1}$. The relation between the contents of the memory of the RAM and the Volterra coefficients is found by assigning all the combinations of

$b_k, b_{k-1}, \cdots, b_{k-N+1}$ in Eq. (4.14). The matrix resulting from this relation is orthogonal and its inverse can be found very easily. Thus the memory contents of the RAM can be found from the Volterra coefficients and vice-versa. If the communication channel is linear and modelled as an FIR filter then the Volterra coefficients $g_0^{(1)}, \cdots, g_{N-1}^{(1)}$ would be the same as the tap coefficients of the filter and the remaining Volterra coefficients $g^{(0)}, g_{01}^{(2)}, \cdots, g^{(N)}$ would be zero. The number of bits used to address the RAM would still be N.

It should be noted that Eq. (4.14) represents the most general case and that it depends only on N (the memory length) and not on n (the order of nonlinearity). Depending on the value of n, in practical situations, some terms of Eq. (4.14) may be zero and thus fewer terms would be required. Two cases can be distinguished

$$\text{number of terms in Eq. (4.14)} = \begin{cases} 2^N & n \geq N \\ \sum_{M=0}^{n} \binom{N}{M} & n < N \end{cases} \tag{4.15}$$

4.4.2 DFE Configuration

In this section we use the relation between the memory contents of the RAM and the Volterra coefficients in order to determine the DFE configuration. From Eq. (4.14) we can see that the reference-former is trying to identify the Volterra coefficients $g^{(0)}, g_0^{(1)}, \cdots, g^{(N_R)}$ where N_R is the memory length of the communication channel. The DFE would try and cancel all the terms of Eq. (4.14) except $g_0^{(1)} b_k$ and the cross product terms that require knowledge of the current value of b_k. Since these require the current value, they cannot be used.

By knowing the RAM contents at all times, we can find the coefficients $g^{(0)}, g_0^{(1)}, \cdots, g^{(N_R)}$ using Eq. (4.14). Then using previously detected symbols the ISI component can be formed and subtracted from the current received signal before detection. The configuration of the DFE is shown in Fig. 4.3.

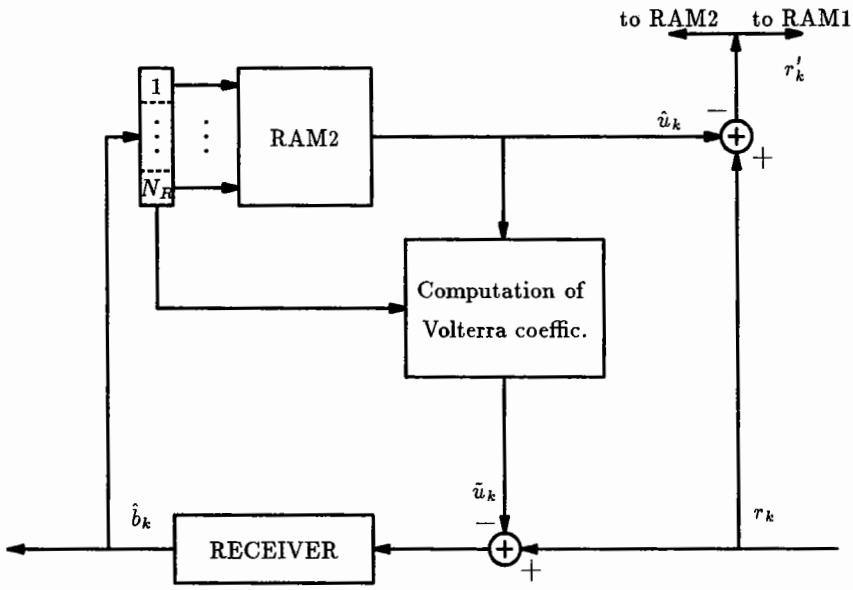


Fig. 4.3 Decision feedback equalizer structure

The figure only shows the DFE part. The rest of the transmission system is the same as in Fig. 4.2. The additional component is a processor that computes a subset of the Volterra coefficients. The processor does not need to perform any multiplications. Only additions and subtractions are required due to the simple relation that exists between the Volterra coefficients and the memory contents of the RAM. In general, the number of Volterra coefficients required is the number of terms that do not contain b_k. This is equal to

$$\sum_{M=1}^{N_R} \binom{N_R - 1}{M - 1} = 2^{N_R - 1} \tag{4.16}$$

This number may be reduced by keeping only those terms that are important in practice. If for example only second and third order nonlinearities are important and $N_R > 3$ then from Eq. (4.15) the number of terms required to form u_k is $\sum_{M=0}^{3} \binom{N_R}{M}$ which is less than 2^{N_R}. For example with $N_R = 8$ only 93 terms are required whereas $2^{N_R} = 256$. Eliminating the terms that contain the current symbol, only 64 coefficients are required to be computed by the processor. The equation that corresponds to Eq. (4.15) for the DFE is

$$\text{number of terms (for DFE)} = \begin{cases} 2^{N_R-1} & n \geq N_R \\ \sum_{M=0}^{n} \binom{N_R-1}{M} & n < N_R \end{cases} \tag{4.17}$$

The computation of the Volterra coefficients can be performed once, after the reference-former has converged, since this will be the time when the coefficients will approximate their correct values. It is also possible to continuously compute the g's after the echo canceller has converged.

4.5 Alternative Solutions

The adaptive reference echo cancellation scheme and the DFE described in the previous sections were premised on the fact that the adaptive reference echo canceller would be based on the table look-up structure. However alternative solutions exist and two of them are suggested here.

4.5.1 Alternative 1

The first alternative is to use a nonlinear filter for the reference-former such as the one suggested in [5] instead of the table look-up structure. The nonlinear filter will

consist of a linear part and a nonlinear part that will form the cross product terms (the number of which will depend on the kind of nonlinearities present in the channel) of the received bits. Then the Volterra coefficients would be available at any time without the need of any extra computations since they will be the tap coefficients of the nonlinear filter. A subset of them could then be used to form the intersymbol interference term.

4.5.2 Alternative 2

The second alternative has as basic components two memories and a nonlinear filter as shown in Fig. 4.4.

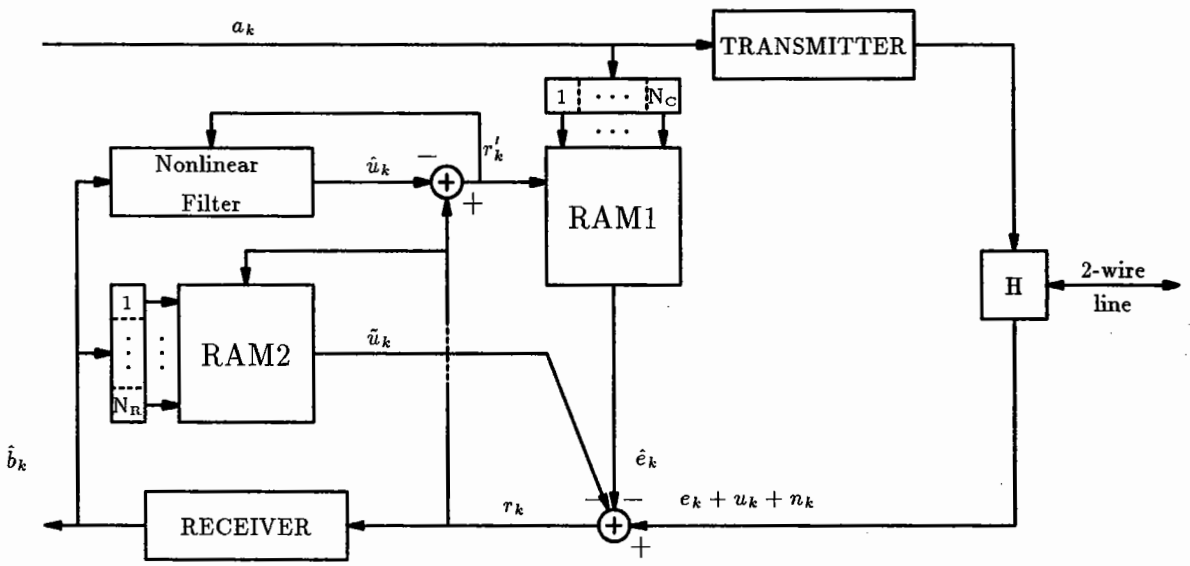


Fig. 4.4 Alternative 2

The first memory (RAM1) is used as an echo canceller, the second memory (RAM2) is used as a DFE, and the nonlinear filter as a reference-former. RAM1 will be addressed by the sequence of bits $(a_k, a_{k-1}, \cdots, a_{k-N_C+1})$ where N_C is as

before, the memory length of the echo channel. RAM2 will be addressed by the sequence of detected bits $(\hat{b}_{k-1}, \hat{b}_{k-2}, \cdots, \hat{b}_{k-N_R+1})$ where N_R is the memory length of the communication channel. RAM2 will form those terms of Eq. (4.14) that do not contain b_k. Thus the nonlinear filter needs to form only those terms that contain b_k as given below

$$\text{number of terms for nonl. filter} = \begin{cases} \sum_{M=0}^{N_R-1} \binom{N_R-1}{M} = 2^{N_R-1} & n \geq N_R \\ \sum_{M=1}^{n} \binom{N_R-1}{M-1} & n < N_R \end{cases} \tag{4.18}$$

For example, if $N_R = 8$ and $n = 3$ the nonlinear filter needs 29 coefficients and RAM2 needs $2^{8-1} = 2^7 = 128$ memory locations. If RAM2 is not used, the nonlinear filter needs 93 coefficients which reduces to the first alternative. The scheme that results without the nonlinear filter is the one suggested in [7].

4.6 Comparison of the Alternative Solutions

The two alternative solutions can be compared with the DFE structure presented in section 4.4 (to be named alternative 0) in terms of the memory and number of Volterra coefficients that are required. If the highest order of nonlinearity is $n = 3$ and a memory location (of the RAM) is assumed equivalent to a coefficient, then we get Tables 4.1 and 4.2 which give the number of coefficients required for each scheme. Fig. 4.5 shows the number of coefficients for different values of N_R for the three cases.

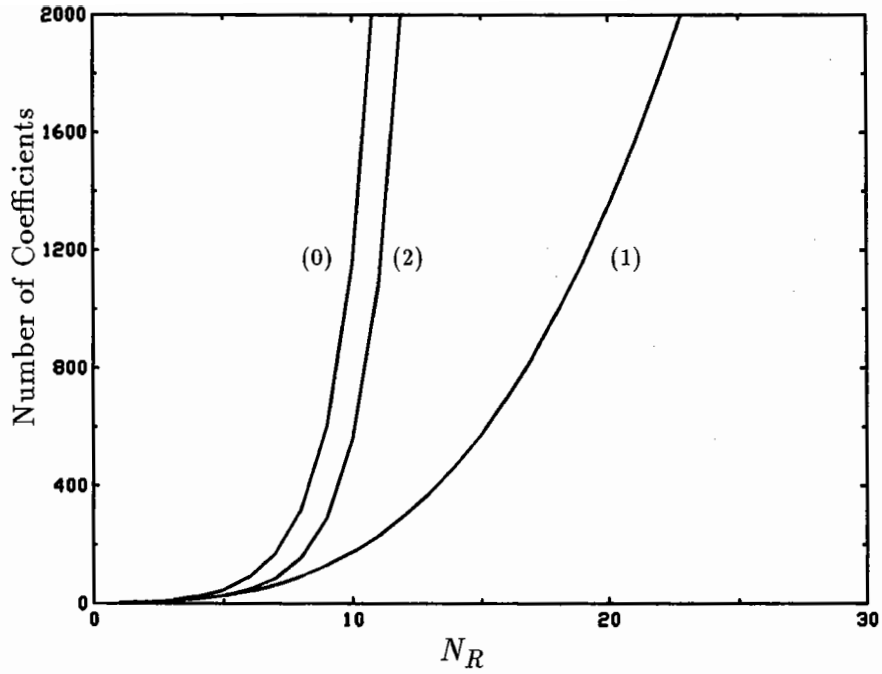


Fig. 4.5 Number of coefficients required for the alternative solutions (0), (1) and (2) as a function of N_R

Number of Coefficients ($n = 3$)	
Alt. 0	$2^{N_R} + \sum_{M=0}^{3} \binom{N_R - 1}{M}$
Alt. 1	$\sum_{M=0}^{3} \binom{N_R - 1}{M}$
Alt. 2	$2^{N_R - 1} + \sum_{M=1}^{3} \binom{N_R - 1}{M - 1}$

Table 4.1 Number of coefficients ($n < N_R$)

Number of Coefficients ($n = 3$)	
Alt. 0	$\frac{3}{2} 2^{N_R}$
Alt. 1	2^{N_R}
Alt. 2	2^{N_R}

Table 4.2 Number of coefficients ($n \geq N_R$)

The choice among the three alternatives will depend on the speed of operation and the complexity of the hardware.

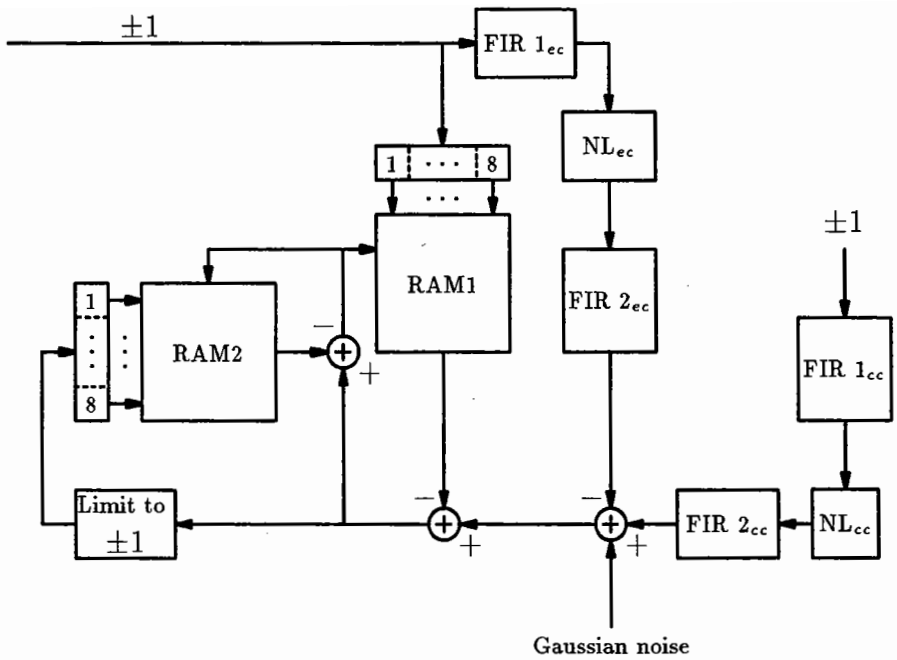
4.7 Computer Simulations for the Two-Memory Structure

In this section computer simulations of the model of Fig. 4.2 are presented and compared with theoretical results. The near-end data and far-end data are uncorrelated ± 1 values. The echo channel consists of a linear part followed by a static nonlinearity which is then followed by a second linear part. This 3-block model is a general nonlinear model which can be described by the Volterra series representation. The communication channel is also simulated as a 3-block nonlinear model. The linear parts are represented by FIR filters. The static nonlinearities are third order polynomials $(a_0 + a_1(\cdot) + a_2(\cdot)^2 + a_3(\cdot)^3)$ so that only second and third order nonlinearities are present.

The ratio of near-end to far-end signal power is 40 dB and the signal-to-noise ratio (SNR) for the communication channel is 20 dB. A Gaussian random variable is used as an additive white noise. The number of bits that address RAM1 (the echo canceller) and RAM2 (the reference-former) is 8. The simulation model is shown in Fig. 4.6 and the filter coefficients and nonlinearity parameters for the near-end echo channel and communication channel are shown in Tables 4.3 and 4.4 respectively.

4.7.1 Convergence Curves

In this section the convergence curves produced by using the model in Fig. 4.6 are plotted. The abscissa for all plots is R_k^2 (in dB) which is the ratio of residual echo



(ec=echo channel, cc=communication channel)

Fig. 4.6 Simulation model for adaptive reference echo cancellation

NEAR END ECHO CHANNEL												
FIR 1						NL				FIR 2		
h_0	h_1	h_2	h_3	h_4	h_5	a_0	a_1	a_2	a_3	h_0	h_1	h_2
1.0	0.23	-0.23	-0.15	-0.06	0.0	0.0	1.0	0.0	0.01	1.0	0.0	-0.2

Table 4.3 FIR filter coefficients and nonlinearity parameters for the near-end echo channel

COMMUNICATION CHANNEL												
FIR 1						NL				FIR 2		
h_0	h_1	h_2	h_3	h_4	h_5	a_0	a_1	a_2	a_3	h_0	h_1	h_2
1.0	-0.25	0.15	-0.1	0.05	-0.01	0.0	1.0	0.0	-0.013	1.0	0.0	-0.125

Table 4.4 FIR filter coefficients and nonlinearity parameters for the communication channel

power to far-end data signal power ($R_k^2 = \frac{\varepsilon_{C,k}}{\sigma_u^2}$) and the ordinate is time in terms of iteration number. The memory contents of RAM1 and RAM2 are initially set to zero. The conditions under which each curve is produced are given within the plots.

Fig. 4.7 compares the theoretical convergence curve using only the echo canceller with the theoretical convergence curve using both the echo canceller and the reference-former (perfect regeneration of the far-end signal is assumed). For curve (a) Eq. (4.12) is used and a stepsize $\alpha_C = 0.2$ and for curve (b) Equations (4.9) and (4.10) are used with $\alpha_C = \alpha_R = 0.2$. In the case where both memories are used, both are activated from the beginning. There are two important things that should be noticed. First, that the steady state value R_∞^2 is almost 20 dB smaller using the reference-former. This also implies that if the same steady state value is to be reached, the two-memory structure will converge faster than the case where only one memory is used (i.e. when only the echo canceller is used). This is shown in Fig. 4.8 where the stepsizes have been chosen to give the same steady state value. Second, that the reference-former has little effect while the echo canceller is still converging as it can be seen from Fig. 4.7 where the two curves coincide up to 6,000 iterations. This is due to the fact that while the echo canceller is converging the level of echo is still much higher than the received signal so that the control signal for the reference-former consists mostly of echo. When the echo canceller converges to the steady state, value the level of echo is reduced below the received signal and the convergence of the reference-former is speeded up.

The steady state value of R_∞^2 using Eq. (4.13) is

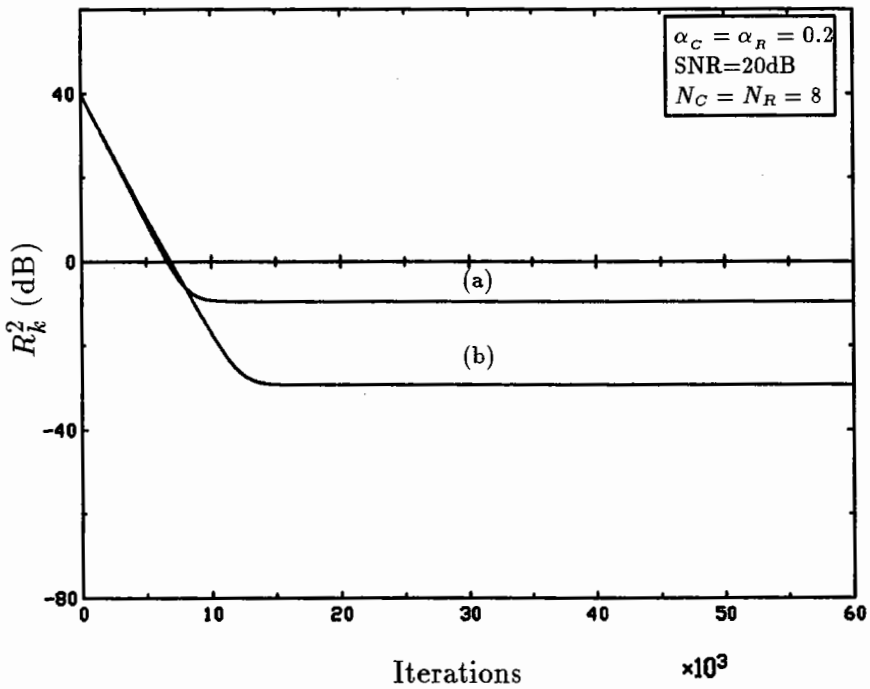


Fig. 4.7 Theoretical convergence curves using (a) only one memory (the echo canceller alone) Eq. (4.12) and (b) using the two-memory structure Eq. (4.9), Eq. (4.10)

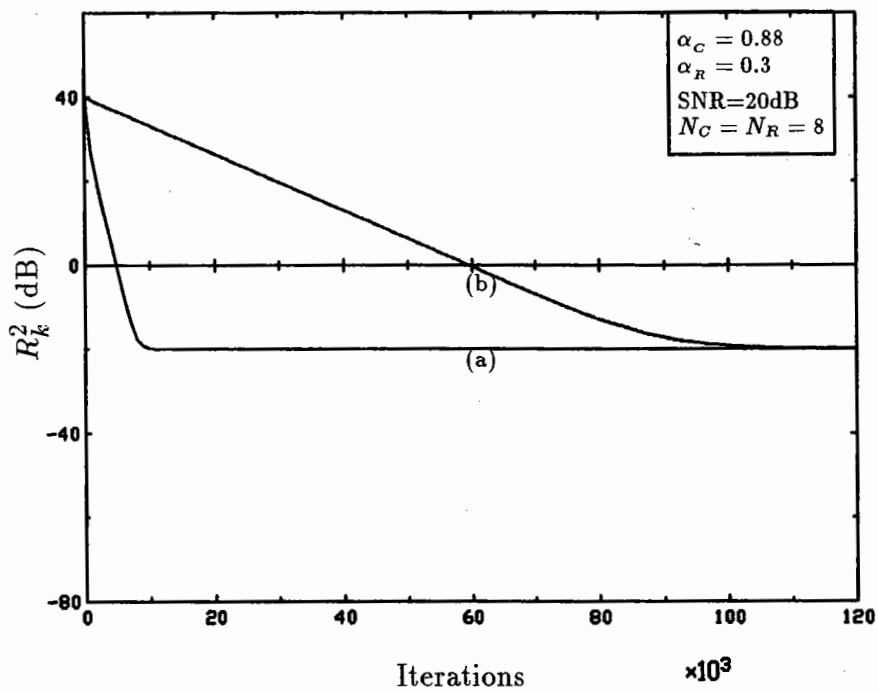


Fig. 4.8 Comparison of the speed of convergence between (a) the two-memory structure and (b) the case when only one memory is used (the echo canceller, with $\alpha_C = 0.02$)

$$R_{\infty}^2 = \frac{\alpha_C \sigma_n^2}{(2 - \alpha_C - \alpha_R)\sigma_u^2} \tag{4.19}$$

and under the above conditions $R_{\infty}^2 = -29.32$ dB for curve (b). From Eq. (4.19) we see that α_R does not affect R_{∞}^2 very much but does affect the speed of convergence of the reference-former as shown in Fig. 4.9. When the stepsize $\alpha_R = 0.08$, $R_{\infty}^2 = -29.32$ dB and when $\alpha_R = 0.05$, $R_{\infty}^2 = -29.71$ dB a difference of just 0.08 dB (not apparent in the figure due to the scale used). If however the stepsize α_C is changed as in Fig. 4.10 the difference in R_{∞}^2 obtained is more apparent. For $\alpha_C = 0.1$, $R_{\infty}^2 = -32.6$ dB and for $\alpha_C = 0.05$, $R_{\infty}^2 = -36.29$ dB.

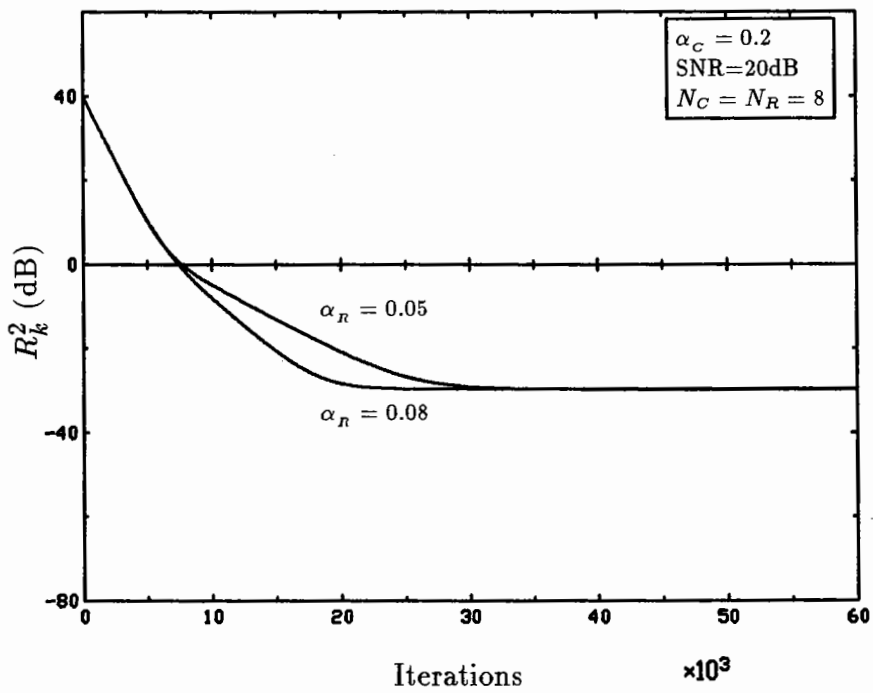


Fig. 4.9 Effect of stepsize used for the reference-former

In all of the above cases (where only theoretical curves were considered) both memories were activated from the beginning and perfect regeneration of the far-

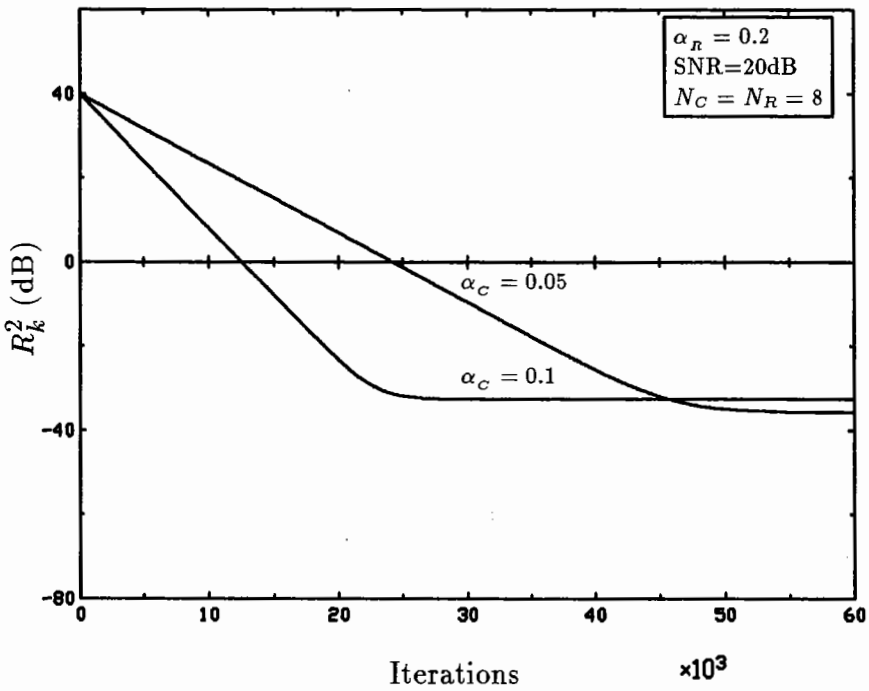


Fig. 4.10 Effect of stepsize used for the echo canceller

end signal was assumed. If in a start-up situation, adaptation of the two memories (starting initially at zero values) is to be achieved without requiring a prearranged "ideal reference" data sequence, it is possible to activate the reference-former at a later stage. This is justified from the fact that without the ideal reference while the echo canceller is converging, most of the decisions used for the reference-former will not be correct. Thus if the reference-former is activated together with the echo canceller it would have no benefit until the echo canceller reached its converged state. In fact it may slow down the overall convergence. However if the reference-former is activated after the echo canceller has converged (when most of the decisions will be correct), it can use a large stepsize and thus converge faster (the steady state value is not affected very much by the large stepsize). The overall convergence time of the modified nonideal-reference algorithm, could be expected to be slower than

an ideal-reference system (an ideal reference system is one where during start-up a prearranged "ideal reference" data sequence is required).

We now compare computer simulations with the theoretical results obtained. In Fig. 4.11 the stepsize for the echo canceller is $\alpha_C = 0.2$ and the stepsize for the reference-former is $\alpha_R = 0.05$. The reference-former is activated at iteration 15,000 (the "noisy" curves represent the computer simulations). The simulation results agree well with the theoretical results. The difference in the two curves at the beginning of the adaptation of the reference-former is due to some initial receiver decision errors. In the case where the stepsize for the echo canceller is low enough so that fewer decision errors occur, then the two curves (theoretical and simulation) are much closer to each other (Fig. 4.12). Each point in the simulation curves is the sample mean of 100 successive squared errors and normalized by the far-end signal power.

If we activate the echo canceller and the reference-former at the same time, the overall convergence time is larger than if the reference-former is activated at a later stage (see Fig. 4.13). We should note that although both methods use the same stepsize for the echo canceller ($\alpha_C = 0.2$) the stepsize for the reference-former in curve (a) (where the reference-former is continually adapting) is much less ($\alpha_R = 0.002$) (hence the apparent slower convergence). This value ($\alpha_R = 0.002$) was chosen so that the two start-up methods achieve the same steady state performance.

The reason α_R is much smaller for curve (a) than it is for curve (b) is because this way the reference-former does not have a significant contribution in the convergence of the echo canceller when the echo signal is large and most of the receiver decisions are not correct. If a large value is used for α_R in curve (a) then when the echo signal

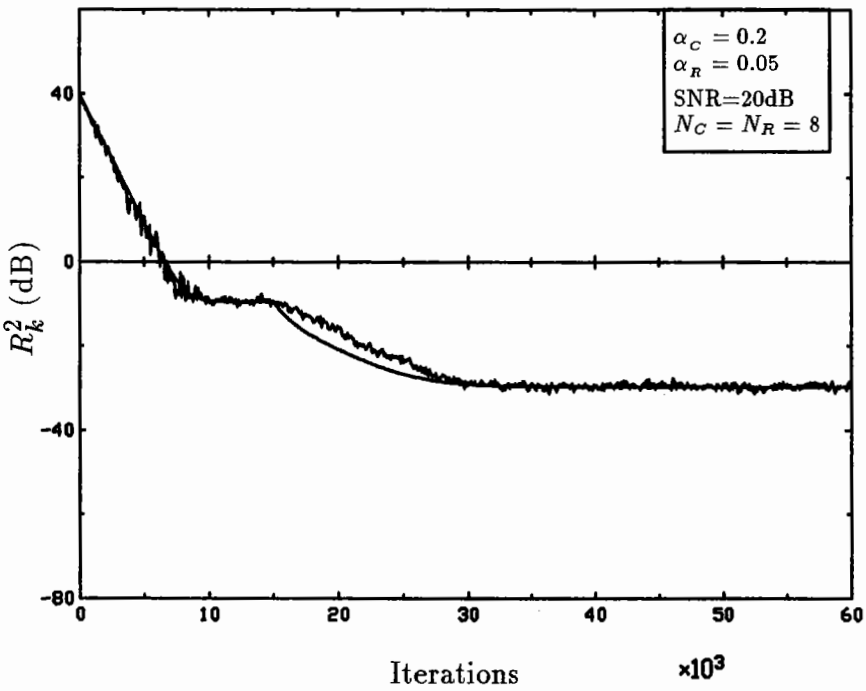


Fig. 4.11 Theory vs Simulation for adaptive reference echo cancellation (reference-former is activated at iteration 15,000)

is still large the reference-former will act as a second echo canceller (because the receiver decisions will be equal to the near-end data symbols with high probability). However the echo estimate from the reference-former is subtracted from the echo canceller control signal and thus the convergence might terminate with a high steady state error. Thus with the high steady state error (which means that the echo signal is still larger than the far-end signal) most of the decisions used for the reference-former would not be correct and no further reduction of the steady state error will be achieved.

In the start-up algorithm of curve (a) the echo canceller tends to converge first. As the residual echo diminishes, more and more receiver decisions are correct, and the reference-former starts to converge. In effect, the two stage start-up procedure is

emulated automatically. Both methods in these simulations achieved the same steady state performance although the initial convergence rate for the two-stage method is faster. For this reason it may be the preferred method, although the second method, with the reference-former continually adapting, may be slightly simpler.

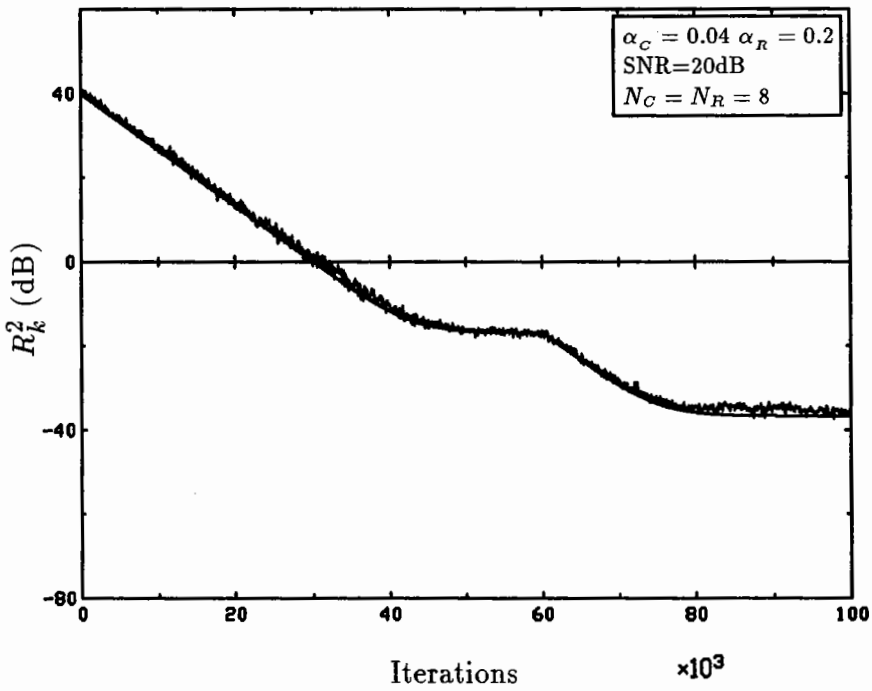


Fig. 4.12 Theory vs Simulation for adaptive reference echo cancellation (the reference-former is activated at iteration 60,000)

We would like now to see the effect of the SNR. In all the above cases, the SNR was 20dB. Fig. 4.14 compares a simulation curve with SNR=20dB with one where SNR=40dB. The decrease in the value of R_∞^2 was expected because of the dependence on the SNR that Eq. (4.19) predicts.

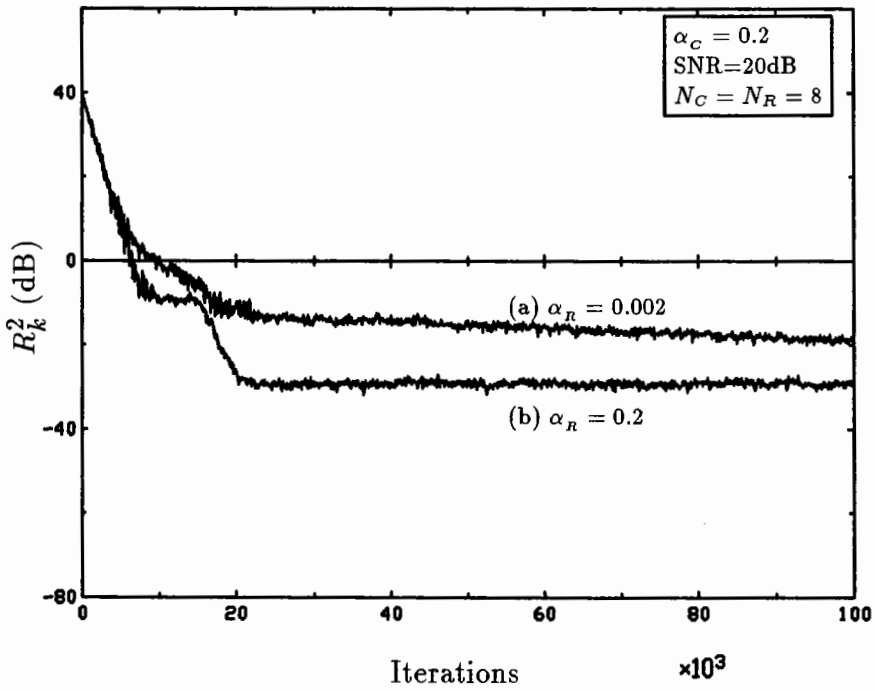


Fig. 4.13 Comparison of two start-up methods
(a) reference-former is continually adapting
(b) reference-former is activated at iteration 15,000

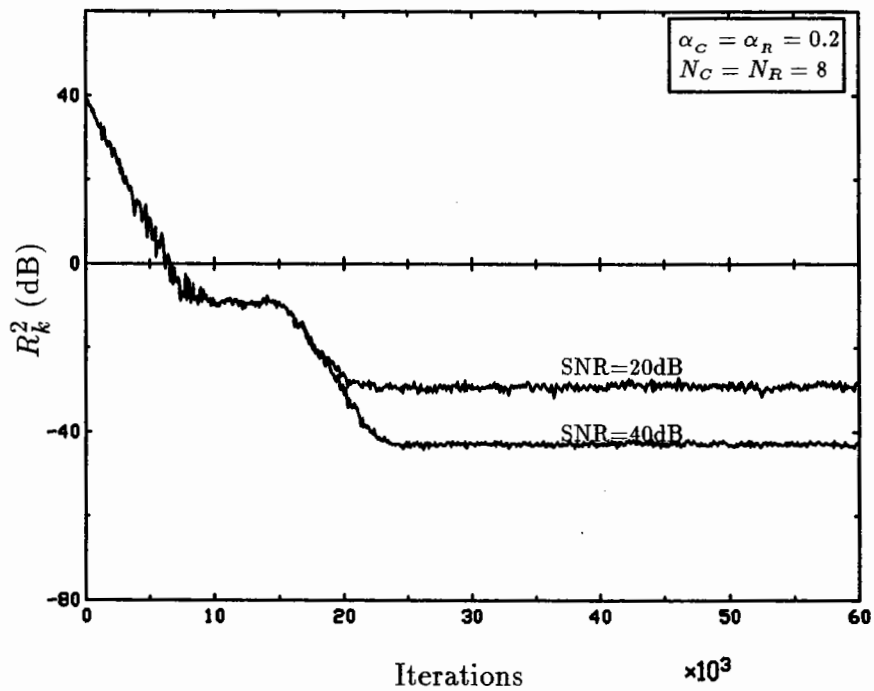


Fig. 4.14 Effect of SNR to the convergence characteristic

4.8 Computer Simulations for the DFE

In this section computer simulations of the decision feedback equalizer are presented. The simulation model is given in Fig. 4.15. All three models for the communication channel as shown in Fig. 4.16 will be used.

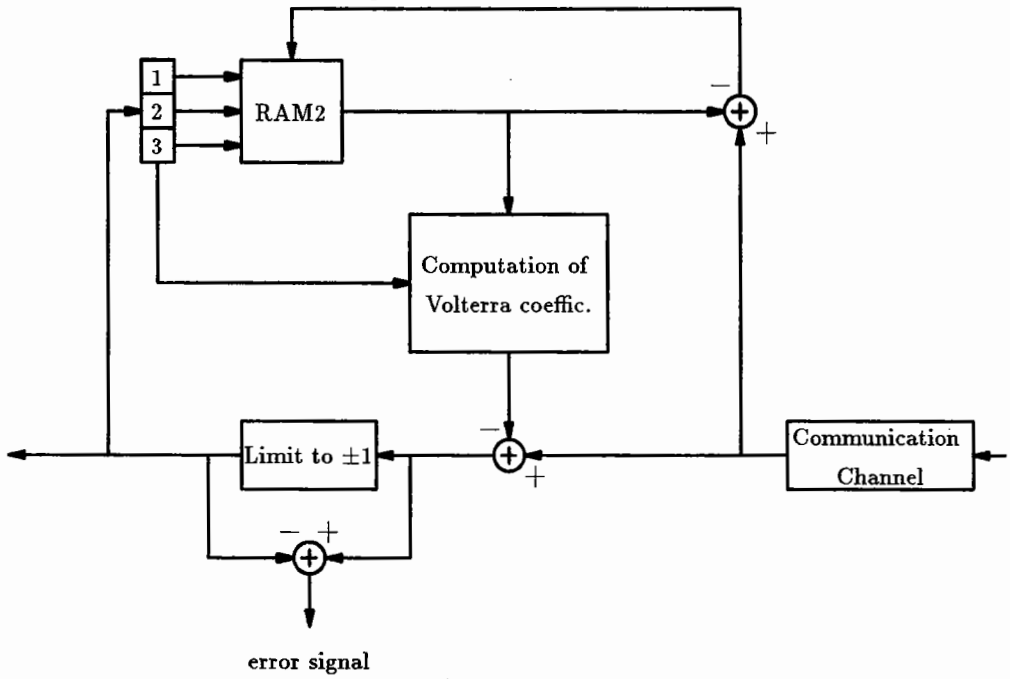


Fig. 4.15 Simulation model for the decision feedback equalizer

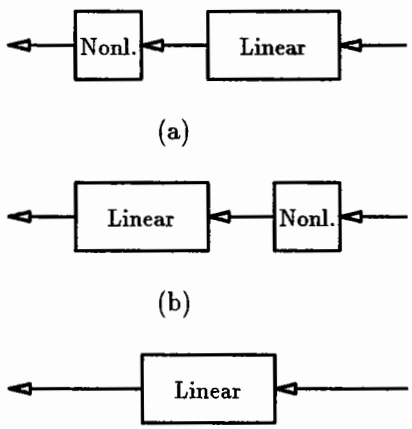


Fig. 4.16 Simulation model for the communication channel

COMMUNICATION CHANNEL						
Linear			Nonlinear			
h_0	h_1	h_2	a_0	a_1	a_2	a_3
1.0	0.6	0.02	0.0	1.0	0.0	0.01

Table 4.5 FIR coefficients and nonlinearity parameters for the communication channel

The number of bits that address RAM2 is $N_R = 3$. The input data to the communications channel are uncorrelated ± 1 values. The linear section is in both cases modelled as an FIR filter with 3 tap coefficients. The nonlinear part is in both cases a static nonlinearity represented by the polynomial $a_0 + a_1(\cdot) + a_2(\cdot)^2 + a_3(\cdot)^3$. The square of the error signal (E_k) averaged over 100 samples is plotted against the number of iterations.

4.8.1 Convergence Curves

Fig. 4.17 displays three curves corresponding to cases (a), (b) and (c) of Fig. 4.16. The stepsize for RAM2 is $\alpha_R = 0.02$. The relation between the Volterra coefficients and the contents of the memory was found using Eq. (4.14). In this case only $g_1^{(1)}$ and $g_2^{(1)}$ are computed by the processor and used to form $\tilde{u}_k$ so that no cross products of the detected symbols and no cross products of the detected with the present symbol are present.

The difference between the linear case (case (c)) and the nonlinear cases (cases (b) and (c)) is due to the fact that in the linear case the processor has all the information required to cancel ISI ($g_1^{(1)}$ and $g_2^{(1)}$) whereas in the nonlinear cases the cross product

terms are missing. The difference between case (a) and case (b) is because in case (b) only $g^{(0)}$ is missing whereas in case (a) $g^{(0)}, g_{01}^{(2)}, g_{02}^{(2)}, g_{12}^{(2)}$ and $g^{(3)}$ are missing.

In Fig. 4.18 $g^{(0)}$ and $g_{12}^{(2)}$ are included in the computation for all cases. We can see that case (b) and (c) behave the same since all the terms required to compute intersymbol interference are included (in case (b) only powers of the present and past symbols are present. No cross products exist). In case (a) however there is a difference because the cross product terms that include b_k (the present symbol) are not computed. The result in Fig. 4.18 is better than in Fig. 4.17 (for (a)) where $g^{(0)}$ and $g_{12}^{(2)}$ are not included.

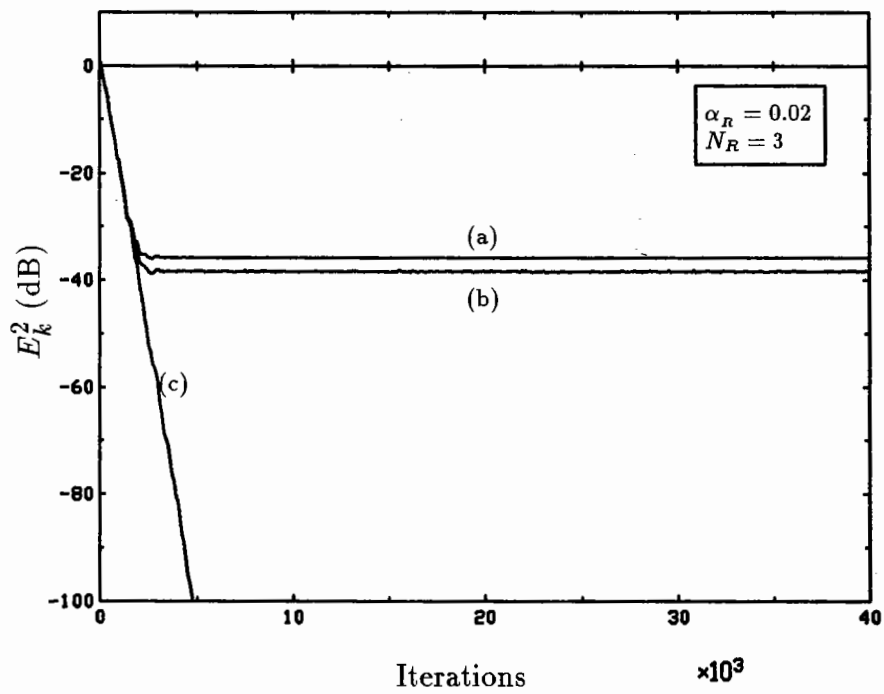


Fig. 4.17 DFE error signal with some Volterra coefficients missing

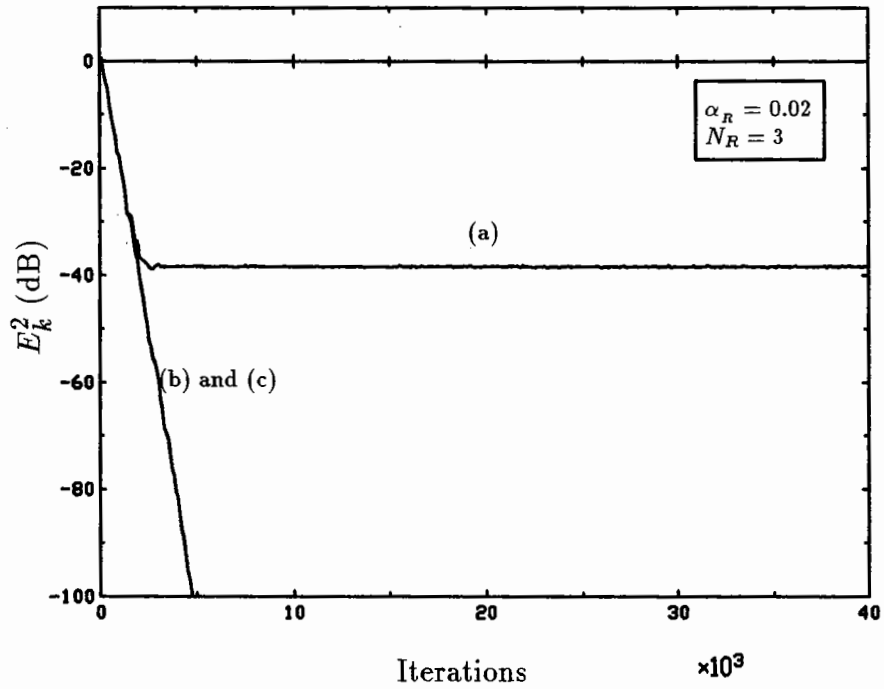


Fig. 4.18 DFE error signal when all possible Volterra coefficients are computed

Chapter 5 The Combined Structure

In Chapter 4, a modification of the table look-up structure was made by adding a second memory that could form an estimate of the received far-end signal and subtract it from the echo canceller control signal. This modification improved performance by increasing the speed of convergence of the echo canceller. In this chapter a different modification is being made by adding a linear FIR filter in parallel to the echo canceller. As in Chapter 4 the goal is to increase the speed of convergence of the echo canceller.

The motivation for the addition of the FIR filter to the table look-up structure comes from the fact that the convergence time of the first is much less than that of the second (it is actually $\frac{2^N}{N}$ greater for the table look-up structure if the RAM is addressed by N bits and N coefficients are used for the filter). Thus the FIR filter can compensate for the linear part of the echo much faster than the table look-up structure would. Then the table look-up structure can compensate for the nonlinear part. If the truncated echo impulse response is N bits long then the FIR filter should have N coefficients and the RAM should be addressed by N bits. The echo canceller configuration is described next.

5.1 Echo Canceller Configuration

The new structure for the echo canceller is shown in Fig. 5.1.

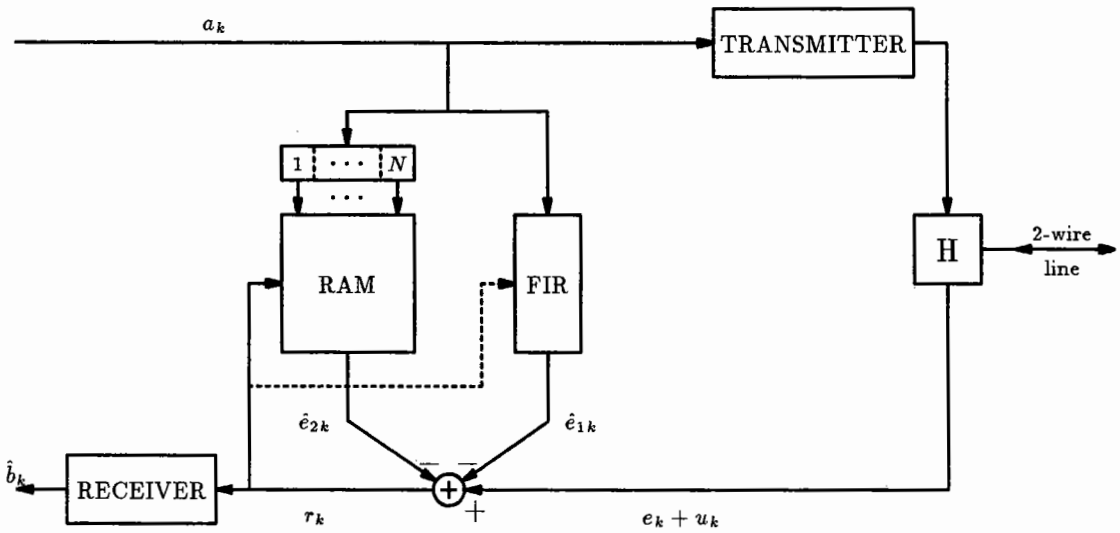


Fig. 5.1 The Combined Structure

The a_k are the near-end transmitted data and are assumed to be uncorrelated with b_k which are the far-end transmitted data. The signal u_k is the received far-end data signal and e_k is the echo signal.

The echo estimate from the FIR filter $\hat{e}_{1k}$ is given by

$$\hat{e}_{1k} = \sum_{i=0}^{N-1} \hat{h}_{ik}\, a_{k-i} \tag{5.1}$$

where the $\hat{h}_{ik}$ are the tap coefficients of the FIR filter. The echo estimate from the RAM, $\hat{e}_{2k}$ is equal to the contents of the RAM addressed by the sequence $(a_k, a_{k-1}, a_{k-2}, \ldots, a_{k-N+1})$ and is given by

$$\hat{e}_{2k} = f(a_k, a_{k-1}, a_{k-2}, \ldots, a_{k-N+1}) \tag{5.2}$$

where $f(\cdot)$ represents the contents of the memory and is a nonlinear function of the transmitted bits. The total echo estimate is the sum of the echo estimates given by

the FIR filter and the RAM as given below

$$\begin{aligned}\hat{e}_k &= \hat{e}_{1k} + \hat{e}_{2k} \\ &= \sum_{i=0}^{N-1} \hat{h}_{ik} a_{k-i} + f(a_k, a_{k-1}, a_{k-2}, \ldots, a_{k-N+1}) \ .\end{aligned} \tag{5.3}$$

The nonlinear function $f(\cdot)$ can be represented by the Volterra series expansion. If we define the vector $\hat{\mathbf{g}}_k$ as the estimate of the Volterra coefficient vector, then the function $f(\cdot)$ can be obtained in matrix form as

$$f(a_k, a_{k-1}, a_{k-2}, \ldots, a_{k-N+1}) = \hat{\mathbf{g}}_k^T \mathbf{a}_k^{(\mathrm{NL})} \tag{5.4}$$

where

$$\hat{\mathbf{g}}_k \triangleq \begin{bmatrix} \hat{g}_k^{(0)} \\ \hat{g}_{0k}^{(1)} \\ \hat{g}_{1k}^{(1)} \\ \vdots \\ \hat{g}_{N-1k}^{(1)} \\ \hat{g}_{01k}^{(2)} \\ \hat{g}_{02k}^{(2)} \\ \vdots \\ \hat{g}_k^{(N)} \end{bmatrix} \qquad \mathbf{a}_k^{(\mathrm{NL})} \triangleq \begin{bmatrix} 1 \\ a_k \\ a_{k-1} \\ \vdots \\ a_{k-N+1} \\ a_k a_{k-1} \\ a_k a_{k-2} \\ \vdots \\ a_k a_{k-1} \cdots a_{k-N+1} \end{bmatrix} .$$

The dimension of both vectors is 2^N, since as was shown in Chapter 4, the truncated Volterra series has a finite number of Volterra coefficients, in the case of binary signals equal to 2^N. If we also define the vectors $\hat{\mathbf{h}}_k$ and $\mathbf{a}_k$ as

$$\hat{\mathbf{h}}_k \triangleq \begin{bmatrix} \hat{h}_0 \\ \hat{h}_{1k} \\ \vdots \\ \hat{h}_{N-1k} \end{bmatrix} \qquad \mathbf{a}_k \triangleq \begin{bmatrix} a_k \\ a_{k-1} \\ \vdots \\ a_{k-N+1} \end{bmatrix} .$$

Eq. (5.3) can be expressed in matrix form as

$$\hat{e}_k = \hat{\mathbf{h}}_k^T \mathbf{a}_k + \hat{\mathbf{g}}_k^T \mathbf{a}_k^{(\mathrm{NL})} \ . \tag{5.5}$$

The control signal used for the adaptation of both the FIR filter and the RAM, r_k is given by

$$\begin{aligned} r_k =& s_k - \hat{e}_k \\ =& e_k + u_k - \hat{e}_k \end{aligned} \tag{5.6}$$

and in matrix form

$$r_k = \left[s_k - (\hat{\mathbf{h}}_k^T \mathbf{a}_k + \hat{\mathbf{g}}_k^T \mathbf{a}_k^{(\mathrm{NL})}) \right] . \tag{5.7}$$

We assume a two-stage start-up method in which while the FIR filter converges, the RAM is not activated. Thus during this time,

$$r_k = \left[s_k - (\hat{\mathbf{h}}_k^T \mathbf{a}_k) \right] \tag{5.8}$$

and the minimum mean square error ($E[(e_k - \hat{e}_{1k})^2]$) can be reached by using the stochastic iteration algorithm that uses the following adaptation equation for the tap coefficients of the FIR filter

$$\hat{\mathbf{h}}_{k+1} = \hat{\mathbf{h}}_k + 2\ \alpha_F\ r_k\ \mathbf{a}_k \tag{5.9}$$

where α_F is the stepsize (F for FIR). When the FIR filter reaches its steady state there is no more improvement in the error performance. Thus $\hat{\mathbf{h}}_k$ is no longer adapted and the adaptation of the RAM begins ($\hat{\mathbf{h}}_k$ can be continuously adapted in the case of a nonstationary channel). The control signal r_k is now given by

$$r_k = \left[(s_k - \hat{e}_{1k}) - (\hat{\mathbf{g}}_k^T \mathbf{a}_k^{(\mathrm{NL})}) \right] . \tag{5.10}$$

As it was shown in [5] a linear update algorithm (like the SIA) can still be used to adapt the nonlinear coefficient vector $\hat{\mathbf{g}}_k$ (because as it was found, the expected value of the squared error is a quadratic function of the elements of $\hat{\mathbf{g}}_k$). However since a

RAM is used, the adaptation is automatically done by addressing the RAM by the vector $\mathbf{a}_k$ and updating the contents of the RAM by an appropriate factor. Thus we get

$$\hat{e}_{2k+1} = \hat{e}_{2k} + \alpha_T \; r_k \tag{5.11}$$

where α_T is the stepsize of the registers in the memory (T for table look-up). The steady state value reached will be that governed by the convergence equation for the table look-up structure (Eq. (3.13)) and consequently by α_T.

5.2 Convergence of the FIR filter

In Chapter 3 the equations governing the convergence of an FIR filter and of the table look-up structure were presented for the stochastic iteration algorithm. The derivation of Eq. (3.4) for the FIR filter assumed a linear echo channel. In this chapter we assume that the echo channel is nonlinear so that Eq. (3.4) no longer holds. The equation governing the convergence of the table look-up structure is still valid because its derivation assumed a general channel. In Appendix C we derive the convergence equation for a linear FIR filter under the constraint of a nonlinear echo channel. This is given by

$$R_k^2 = \rho^k \left[R_0^2 - \frac{\alpha_F N + G/\sigma_u^2}{(1 - \alpha_F N)} \right] + \frac{\alpha_F N + G/\sigma_u^2}{(1 - \alpha_F N)} \tag{5.12}$$

where R_k^2 is as in previous chapters the ratio of the residual echo power to the far-end signal power, and

$$\rho \stackrel{\triangle}{=} (1 - 4\alpha_F + 4\alpha_F^2 N) \; .$$

G is a nonlinear factor representing the sum of the squares of the nonlinear terms of the Volterra coefficient vector. The nonlinear terms represent that part of the Volterra coefficient vector which the FIR filter cannot identify. The steady state value of R_k^2, R_∞^2 is given by

$$R_\infty^2 = \frac{\alpha_F N + G/\sigma_u^2}{(1 - \alpha_F N)} \ . \tag{5.13}$$

If we compare Eq. (5.13) with what was found for the case of a linear echo channel (see Table 5.1 below) we find that there is a difference of $(G/\sigma_u^2)/(1 - \alpha_F N)$.

It can be seen that for small values of the stepsize α_F, not much improvement is expected in the degree of cancellation and the nonlinear factor G is more dominant. However when small nonlinearities are present in the channel, a smaller stepsize will have more effect in the steady state value of the mean square error. Of course if no nonlinearities are present the nonlinear factor G will be zero. We now summarize the new results found in this section in Table 5.1 where we also present the results for the linear channel case.

Nonlinear Echo Channel	Linear Echo Channel
$R_k^2 = \rho^k[R_0^2 - \frac{\alpha_F N + G/\sigma_u^2}{(1 - \alpha_F N)}] + \frac{\alpha_F N + G/\sigma_u^2}{(1 - \alpha_F N)}$	$R_k^2 = \rho^k[R_0^2 - \frac{\alpha_F N + G/\sigma_u^2}{(1 - \alpha_F N)}] + \frac{\alpha_F N}{(1 - \alpha_F N)}$
$R_\infty^2 = \frac{\alpha_F N + G/\sigma_u^2}{(1 - \alpha_F N)}$	$R_\infty^2 = \frac{\alpha_F N}{(1 - \alpha_F N)}$
$\nu_{20} = \frac{-2}{\log_{10} \rho}$	$\nu_{20} = \frac{-2}{\log_{10} \rho}$

Table 5.1 Convergence of FIR filter using the Stochastic Iteration Algorithm

5.3 Computer Simulations

The structure of Fig. 5.1 is simulated using the simulation model shown in Fig. 5.2. The number of bits used to address the RAM is $N = 8$ and the FIR filter consists of 8 coefficients. The input data to the echo channel and the communication channel are uncorrelated ± 1 values. The ratio of near-end to far-end signal power is 40 dB. The values used for the linear and nonlinear sections of the echo and communication channels are the same as the ones used in Chapter 4 for adaptive reference echo cancellation.

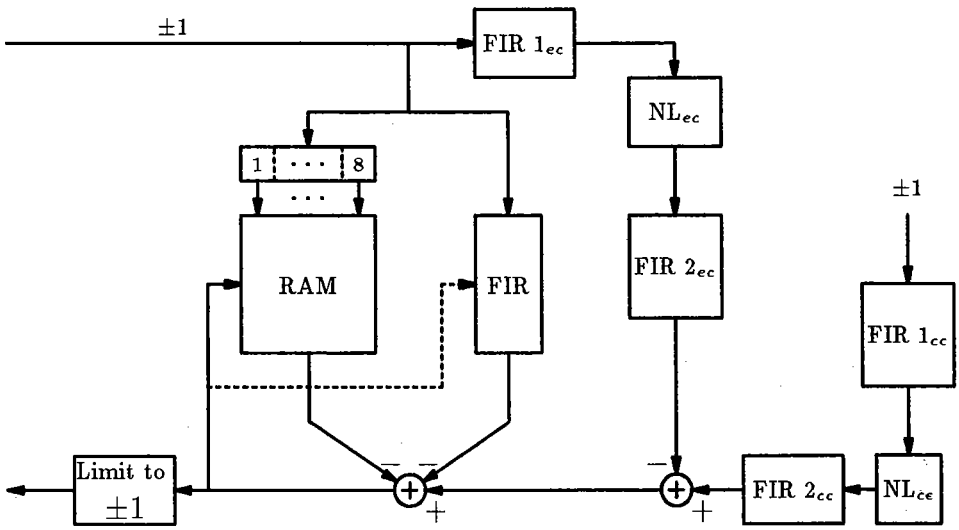


Fig. 5.2 Simulation model of the combined structure

5.3.1 Convergence Curves

In this section the residual echo power to far-end data signal power (R_k^2) is plotted in dB against the number of iterations. Fig. 5.3 shows the convergence curve for using a stepsize of $\alpha_F = 0.005$ for the FIR and a stepsize $\alpha_T = 0.09$ for the RAM. The

RAM is activated at iteration 3000 and at that time the coefficients of the FIR filter are 'frozen'. The amount of cancellation achieved is predicted from Eq. (3.13) ($R^2_\infty = -13.46$ dB).

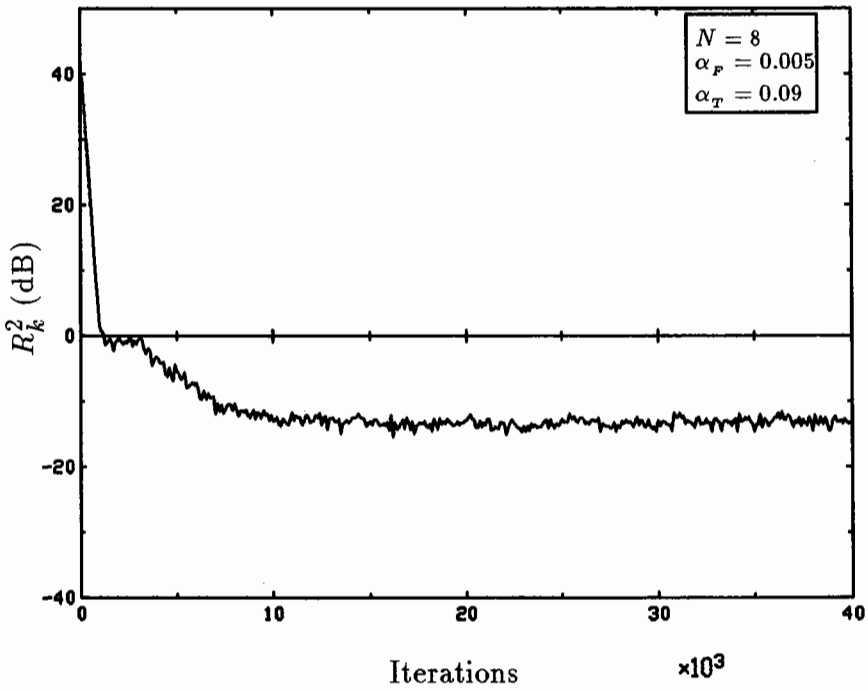


Fig. 5.3 Computer simulation for the combined structure

To see the effect of the FIR filter Fig. 5.4 shows the convergence curve of Fig. 5.3 and the convergence of using only the RAM. The stepsize for the RAM is again $\alpha_T = 0.09$. The increase in the speed of convergence is considerable. With the FIR filter, steady state is reached at approximately 12,000 iterations whereas without the FIR filter steady state is reached at 22,000 iterations.

It is also interesting to see the capabilities of the RAM to cancel nonlinearities as compared to a linear FIR filter. Fig. 5.5 compares the convergence curves of using only an FIR filter with the combined structure.

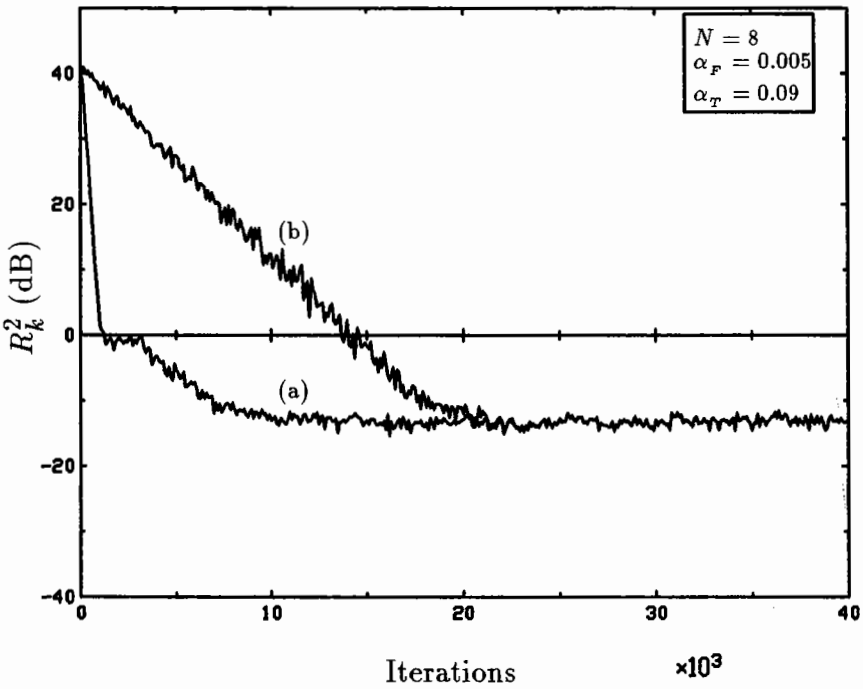


Fig. 5.4 Combined structure (a) vs Table look-up structure (b)

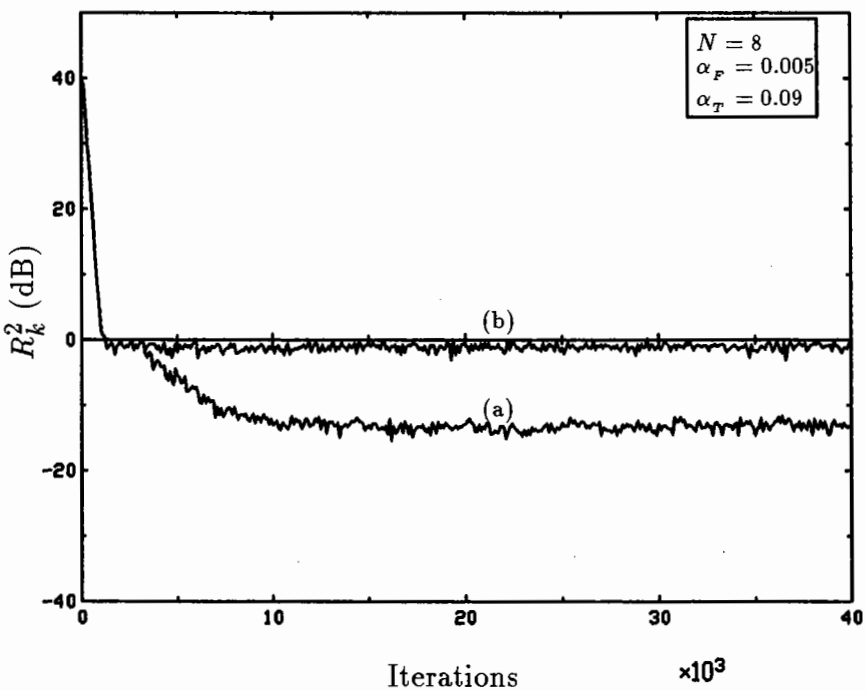


Fig. 5.5 Combined structure (a) vs FIR filter (b)

A simple example is now presented in order to compare the theoretical convergence of the FIR filter corresponding to Eq. (5.12) with computer simulations. We model the nonlinear echo channel as a cascade of an FIR filter with 2 coefficients, and a nonlinear section with 4 parameters (Fig. 5.6). The values used are shown in Table 5.2.

ECHO CHANNEL					
FIR filter		Nonlinearity			
w_0	w_1	β_0	β_1	β_2	β_3
1.0	0.2	0.0	1.0	0.01	0.01

Table 5.2 FIR filter coefficients and nonlinearity parameters used for the echo channel

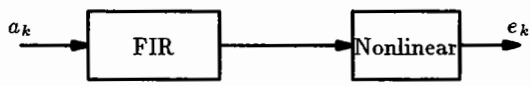


Fig. 5.6 Nonlinear echo channel

The communication channel remains the same as the previous simulations. The Volterra coefficients of the echo channel are $2^N = 2^2 = 4$ in number and were computed to be

$$g^{(0)} = \beta_0 + \beta_2 w_0^2 + \beta_2 w_1^2 = 0.0104$$

$$g_0^{(1)} = \beta_1 w_0 + \beta_3 w_0^3 + 3 w_0 w_1^2 \beta_3 = 1.0112$$

$$g_1^{(1)} = \beta_1 w_1 + \beta_3 w_1^3 + 3 w_0^2 w_1 \beta_3 = 0.2061$$

$$g^{(2)} = 2\beta_2 w_0 w_1 = 0.0040 \ .$$

The coefficients $g_0^{(1)}$ and $g_1^{(1)}$ represent the linear part of the channel and the coefficients $g^{(0)}$ and $g^{(2)}$ the nonlinear part. In this special case, the vectors $\mathbf{g}_1^{(L)}$ and

$\mathbf{g}_2^{(\mathrm{NL})}$ (defined as the linear and nonlinear Volterra coefficient vectors respectively) take the following form

$$\mathbf{g}_1^{(\mathrm{L})} = \begin{bmatrix} 1.0112 \\ 0.2061 \end{bmatrix} \qquad \mathbf{g}_2^{(\mathrm{NL})} = \begin{bmatrix} 0.0104 \\ 0.0040 \end{bmatrix}$$

We can now calculate the nonlinear factor G so that we can use Eq. (5.12). The value of G is found to be equal to 1.241×10^{-4}. Fig. 5.7 shows the convergence curve using Eq. (5.12) with a stepsize of $\alpha_F = 0.001$ and the convergence curve obtained from computer simulations. The steady state value of R_k^2 is that predicted by Eq. (5.13) ($R^2_{\infty,\mathrm{dB}} = 0.65$ dB). If the stepsize is changed to $\alpha_F = 0.0007$ there is no apparent improvement in the value of R^2_∞.

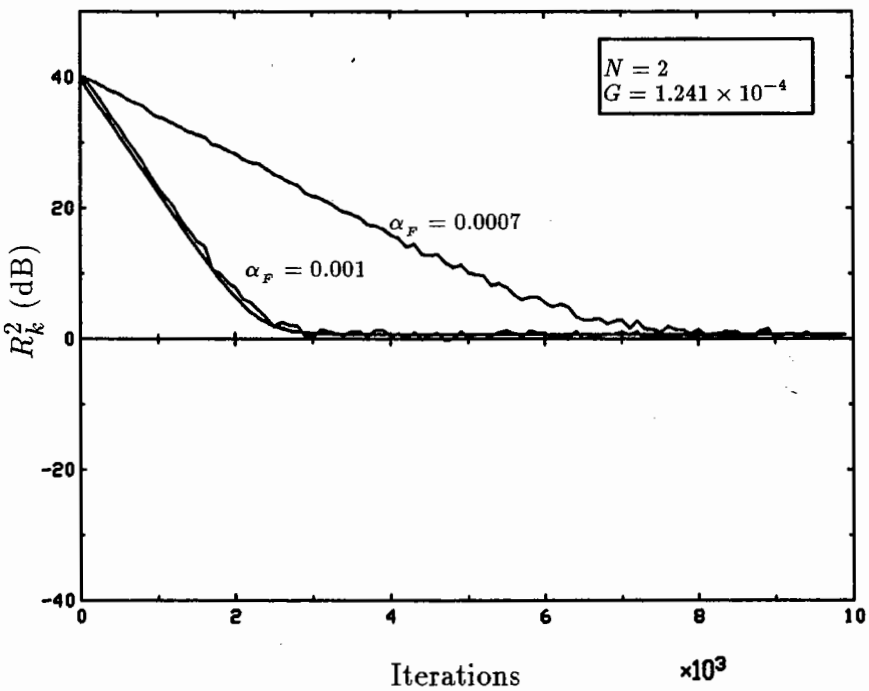


Fig. 5.7 Theoretical convergence curve vs Computer simulations for a linear FIR filter operating under a nonlinear channel

However if we change the nonlinear parameters β_2 and β_3 to 0.001, the change

in R^2_∞ is much larger. This shows the effect of the nonlinearities in the overall convergence of the FIR filter ($R^2_{\infty,\mathrm{dB}} = -18.65$ dB) (see Fig. 5.8).

We now compare the theoretical convergence curve of the combined structure using Eq. (5.12) for the FIR and Eq. (3.13) for the table look-up structure with computer simulations. The RAM is addressed by two bits (Fig. 5.9).

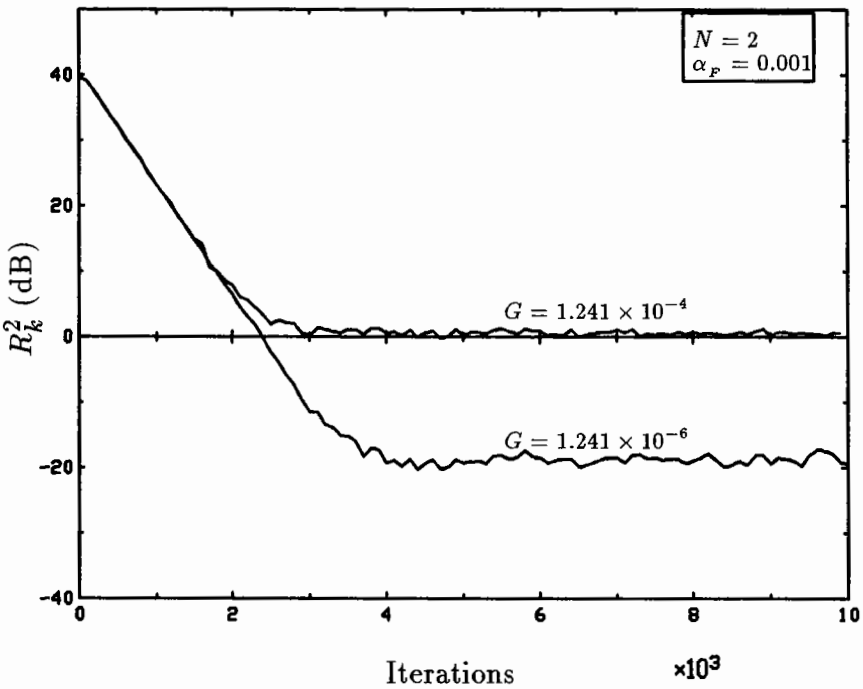


Fig. 5.8 Effect of the nonlinear factor G in the convergence of a linear FIR filter

The two-stage start-up method considered so far is similar to the one used in Chapter 4 for the adaptive reference echo cancellation scheme. A second alternative is to activate the table look-up structure and the FIR filter at the same time. Fig. 5.10 shows computer simulations for the two start-up methods.

In the second method the FIR filter converges fast up to a point after which the convergence is slowed down. At this point the echo estimate from the RAM is having

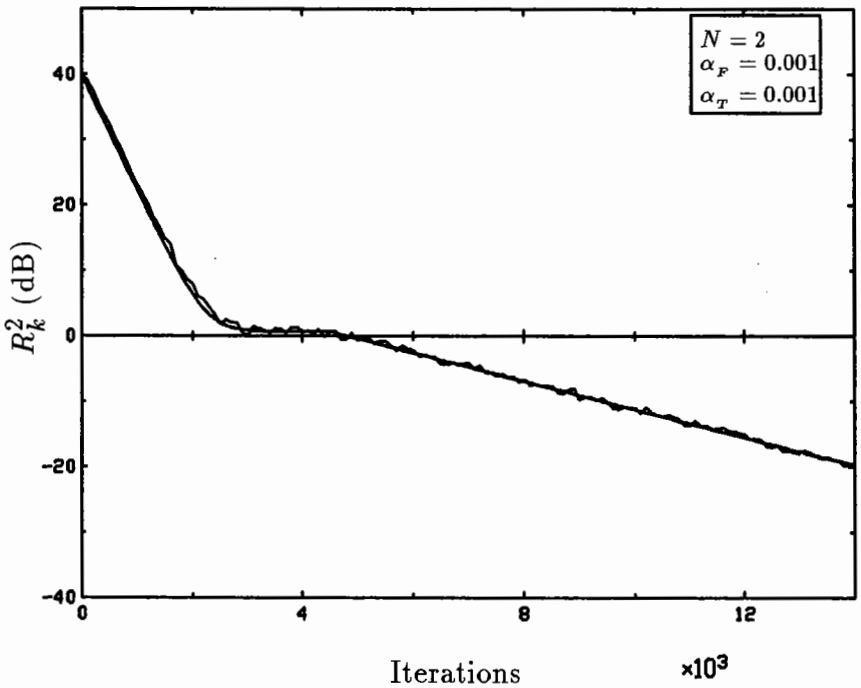


Fig. 5.9 Theoretical convergence of combined structure vs computer simulation

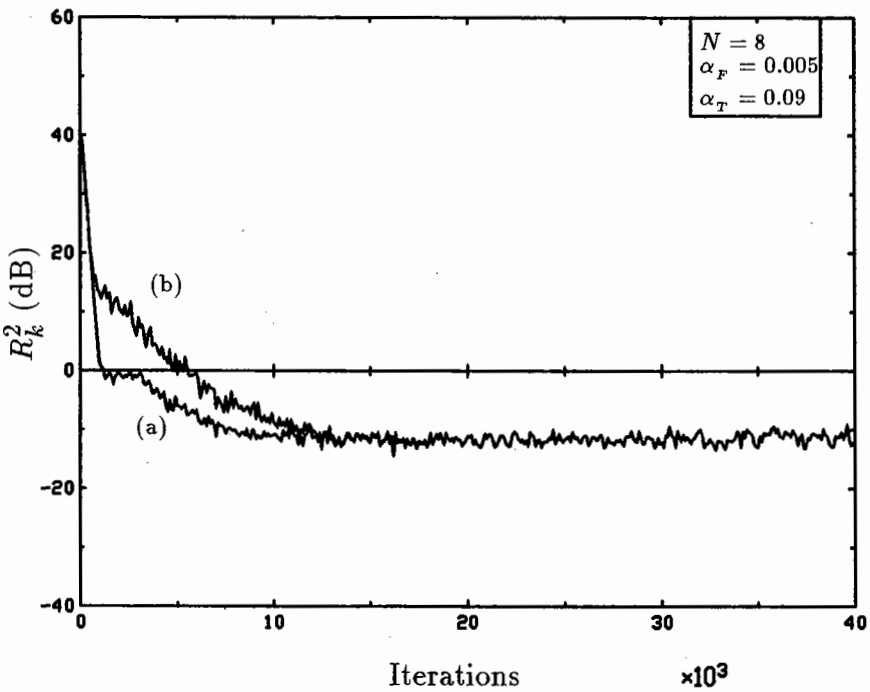


Fig. 5.10 Comparison of two start-up methods: (a) the RAM is activated at iteration 3000 (b) the RAM is continually adapting

a significant contribution to the residual signal and changes in the filter coefficients are smaller. Thus the table look-up structure takes over compensating for the remaining linear, and the nonlinear part of the echo channel at a much slower rate. In effect the two-stage start-up procedure is emulated automatically. The two-stage method might be preferred because typically it is desirable to compensate for the linear part of the echo channel as fast as possible.

Similar results to the ones found in this chapter could be obtained if an echo canceller configuration is considered in which the FIR filter and RAM use separate control signals, i.e, the filter control signal to be the difference between the echo and the filter echo estimate, and the RAM control signal to be the difference between the filter control signal and the RAM echo estimate.

Chapter 6 Compensators for Nonlinear Channels

The nonlinear filters described in the previous Chapters were based on the table look-up structure. As mentioned in the introduction, another alternative is to use a nonlinear filter that approximates the nonlinear echo replica by a Volterra-like expansion. This can be viewed as adding extra nonlinear taps to a linear transversal filter. Both the table look-up structure and the Volterra series nonlinear filter (as it will be called in the sequel) can compensate for general nonlinearities. In practical situations however, nonlinear channels may be more specific. Thus, specific nonlinear filters or compensators could be used instead of the general ones. In this chapter one such nonlinear compensator is examined, that is modelled in the same way as the channel that is trying to compensate for. The model consists of the cascade of a nonlinear part, a linear part and a second nonlinear part. Before the new echo canceller configuration is described, a review of the Volterra series nonlinear filter and two of the more specific compensators is given.

6.1 The Volterra Series Nonlinear Filter

Assuming a finite-memory causal system with input sequence $\{a_k\}$ and output sequence $\{y_k\}$, the relationship between input and output can be generally expressed as

$$y_k = f(a_k, a_{k-1}, \cdots, a_{k-N+1}) \tag{6.1}$$

where N is the memory length of the system and $f(\cdot)$ is given by the truncated Volterra series (Eq. 4.14)). The system described by Eq. (6.1) can be implemented by a finite-length transversal filter, a number of multipliers, and a summing bus as shown in Fig. 6.1.

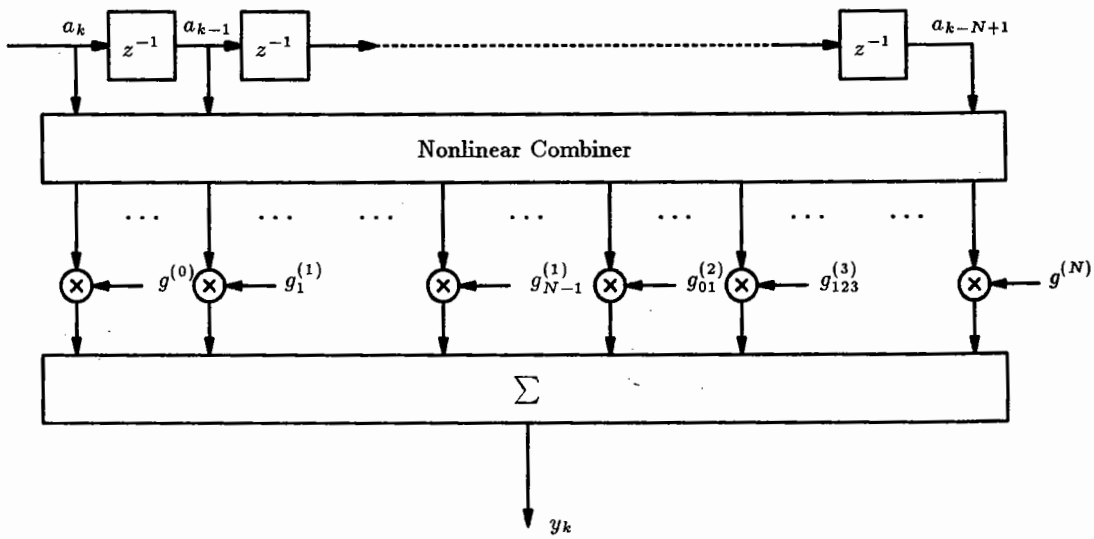


Fig. 6.1 The Volterra series nonlinear filter

From the above figure we can see that the difference between the nonlinear filter and the conventional linear filter is that in the nonlinear case the addition of $2^N - N$ nonlinear taps has been made (assuming again binary signals). It was suggested in [5] that in many practical cases the nonlinear function $f(\cdot)$ departs only slightly from

a linear function so that only a few of the extra nonlinear terms need to be retained. It is in this property that the usefulness of the above representation resides. If the full complement of terms is required, then it becomes equivalent to the table look-up structure.

6.1.1 Adaptation Algorithm

The output of the nonlinear filter can be written in matrix form as

$$y_k = \mathbf{g}_k^T \mathbf{a}_k^{(\mathrm{NL})} \tag{6.2}$$

where $\mathbf{g}_k$ and $\mathbf{a}_k^{(\mathrm{NL})}$ are defined as

$$\mathbf{a}_k^{(\mathrm{NL})} \triangleq \begin{bmatrix} 1 \\ a_k \\ a_{k-1} \\ \vdots \\ a_k a_{k-1} \\ a_k a_{k-2} \\ \vdots \\ a_k a_{k-1} \cdots a_{k-N+1} \end{bmatrix} \qquad \mathbf{g}_k \triangleq \begin{bmatrix} g_k^{(0)} \\ g_{0k}^{(1)} \\ g_{1k}^{(1)} \\ \vdots \\ g_{01k}^{(2)} \\ g_{02k}^{(2)} \\ \vdots \\ g_k^{(N)} \end{bmatrix}$$

and $\mathbf{g}_k$ represents the tap coefficient vector of the nonlinear filter (T denotes transposition).

The adaptation equation for the vector $\mathbf{g}_k$ can be found by minimizing the mean square error between the output of the nonlinear filter y_k and a desired response d

$$\begin{aligned} E[\epsilon_k^2] = E[(d - y_k)^2] &= E[d^2 - 2d(\mathbf{a}_k^{(\mathrm{NL})T} \mathbf{g}_k) + (\mathbf{g}_k^T \mathbf{a}_k^{(\mathrm{NL})})(\mathbf{a}_k^{(\mathrm{NL})T} \mathbf{g}_k)] \\ &= E[d^2] - 2\, \mathbf{p}^T \mathbf{g}_k + \mathbf{g}_k^T \mathbf{R}\, \mathbf{g}_k \end{aligned} \tag{6.3}$$

where

$$\begin{aligned} \mathbf{p} &= E[d\, \mathbf{a}_k^{(\mathrm{NL})}] \\ \mathbf{R} &= E[\mathbf{a}_k^{(\mathrm{NL})}\, \mathbf{a}_k^{(\mathrm{NL})T}] \end{aligned} \tag{6.4}$$

This shows that the expected value of the squared error is a quadratic function of the elements of $\mathbf{g}_k$ and for this reason, the stochastic iteration algorithm can be used to adapt $\mathbf{g}_k$

$$\mathbf{g}_{k+1} = \mathbf{g}_k + 2\ \alpha\ \epsilon_k\ \mathbf{a}_k^{(\mathrm{NL})}\ . \tag{6.5}$$

6.2 Compensators for Specific Nonlinear Channels

The Volterra series nonlinear filter is general as to the kind of nonlinearities that it can compensate for. Other models have been considered in the literature which compensate for specific nonlinear channels. Fig. 6.2(a) shows the model of a compensator used in [23] to compensate for the memoryless nonlinearity present in the DAC in the canceller path (the Wiener scheme). The model consists of the cascade of a linear filter and a memoryless nonlinearity. Fig. 6.2(b) shows the model used in [18] (the Hammerstein model). The Hammerstein model consists of a memoryless nonlinearity preceding a linear filter.

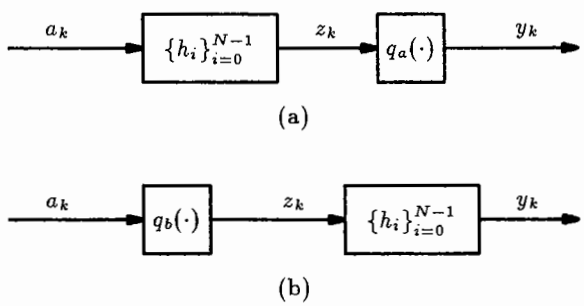


Fig. 6.2 Compensators for specific nonlinear channels (a) The Wiener scheme (b) The Hammerstein model

In both cases the $\{h_i\}$ are the tap coefficients of a linear filter and the nonlinearity $q(\cdot)$ is parameterized in a polynomial form with parameters $\{\alpha_j\}$ as

$$\begin{aligned} q_a(z_k) &= \sum_{j=0}^{M-1} \alpha_j z_k^j \\ q_b(a_k) &= \sum_{j=0}^{M-1} \alpha_j a_k^j \end{aligned} \tag{6.6}$$

The above models can compensate for channels that have the same representation as the one in each model. The coefficients $\{h_i\}$ and the parameters $\{\alpha_j\}$ are adapted separately in each case until the minimum mean square error of the output y_k and a desired response d is reached.

Another model that can be specific but more general than the above two models is the one that consists of the cascade of a linear part, a static nonlinear part, and a second linear part (all being adaptive). This structure is attractive because it can adequately represent channels that are described by the Volterra series. However, computer simulations showed that adapting such a configuration leads to local minima and suboptimal solutions.

6.3 The New Nonlinear Compensator

A nonlinear model that might result in the near-end echo channel (as explained in Section 2.5) is the one shown in Fig. 6.3 (Fig. 2.15 reproduced here for convenience). This is a cascade of a memoryless nonlinearity, a linear system with memory, and a second memoryless nonlinearity.

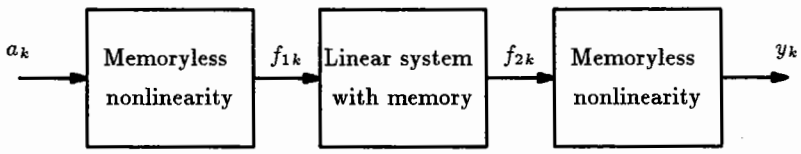


Fig. 6.3 A specific nonlinear channel

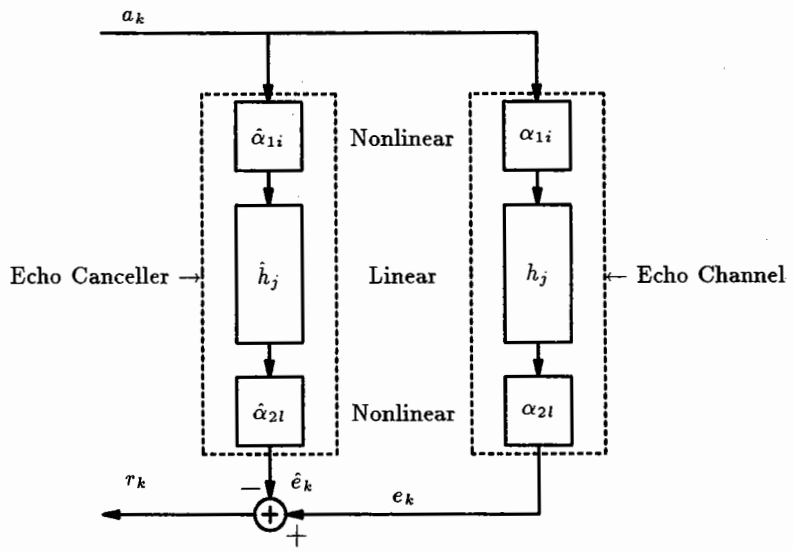


Fig. 6.4 The new nonlinear compensator

The input/output relations at the different stages are

$$
\begin{aligned}
f_{1k} &= \sum_{i=0}^{M-1} \alpha_{1i} a_k^i \\
f_{2k} &= \sum_{j=0}^{N-1} h_j f_{1k-j} = \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} h_j \alpha_{1i} a_{k-j}^i \\
y_k &= \sum_{l=0}^{L-1} \alpha_{2l} f_{2k}^l
\end{aligned}
\tag{6.7}
$$

where the $\{h_j\}$ are the impulse response coefficients of the linear FIR filter representing the linear system with memory, the $\{\alpha_{1i}\}$ are the parameters of the first nonlinearity, and the $\{\alpha_{2l}\}$ are the parameters of the second nonlinearity. Let us assume that the same scheme is used for the compensator (echo canceller) as shown in Fig. 6.4.

The $\{\hat{h}_j\}$,$\{\hat{\alpha}_{1i}\}$, and $\{\hat{\alpha}_{2l}\}$ represent estimates of $\{h_j\}$, $\{\alpha_{1i}\}$, and $\{\alpha_{2l}\}$ respectively and obey the same relations given by Eq. (6.7).

Our objective is to minimize the mean square error ε_k

$$\varepsilon_k = E[(e_k - \hat{e}_k)^2] \tag{6.8}$$

A gradient search algorithm can be used where the instantaneous squared error is taken as an estimate of the true square error. Then the gradient of the instantaneous squared error is found with respect to $\{\hat{h}_j\}$, $\{\hat{\alpha}_{1i}\}$, and $\{\hat{\alpha}_{2l}\}$. It is easy to verify that the following adaptation equations result

$$\hat{\alpha}_{1i\ k+1} = \hat{\alpha}_{1i\ k} + 2\left[\sum_{l=0}^{L-1} l\ \hat{\alpha}_{2lk}\ \hat{f}_{2k}^{l-1} \sum_{j=0}^{N-1} \hat{h}_{jk}\ a_{k-j}^{i}\right] r_k\ \mu_i \tag{6.9.a}$$

$$i = 0, \cdots, M-1$$

$$\hat{h}_{j\ k+1} = \hat{h}_{j\ k} + 2\hat{f}_{1\ k-j}\left[\sum_{l=0}^{L-1} l\ \hat{\alpha}_{2lk}\ \hat{f}_{2k}^{l-1}\right] r_k\ \nu_j \tag{6.9.b}$$

$$j = 0, \cdots, N-1$$

$$\hat{\alpha}_{2l\ k+1} = \hat{\alpha}_{2l\ k} + 2\ \hat{f}_{2k}^{l}\ r_k\ \xi_l \tag{6.9.c}$$

$$l = 0, \cdots, L-1$$

where μ, ν, and ξ represent the stepsizes. The summation $\sum_{l=0}^{L-1} l\ \hat{\alpha}_{2lk}\ \hat{f}_{2k}^{l-1}$ is the derivative of the instantaneous estimate of $\hat{e}_k$ as given in Eq. (6.7) (with y_k replaced by $\hat{e}_k$).

A general study of convergence is highly complicated because $\hat{f}_{1k}$ and $\hat{f}_{2k}$ are not stationary (as the coefficients of the linear filter and the parameters of the nonlinearities are time varying). Furthermore the interrelations between the three sets of parameters (first nonlinear , linear, second nonlinear) are nonlinear. In order to gain some insight for the convergence, the following assumption is made.

The parameters of the static nonlinearities and the coefficients of the linear filter converge in stages so that when the parameters of the first nonlinearity begin to adapt the linear filter has already reached its optimal value, and when the parameters of the second nonlinearity begin to adapt both the first nonlinearity and the linear filter have reached their optimum values. From the adaptive point of view, this can be seen as the result of selecting a much lower convergence speed for the nonlinear parts of the system.

Under the above assumption $\hat{\alpha}_{1i}$ and $\hat{\alpha}_{2l}$ can remain at their initial values while the linear filter is adapting so that

$$\hat{\alpha}_{11} = 1 \quad , \qquad \hat{\alpha}_{1i} = 0 \quad i \neq 1 \quad \text{at } k = 0$$

$$\hat{\alpha}_{21} = 1 \quad , \qquad \hat{\alpha}_{2l} = 0 \quad l \neq 1 \quad \text{at } k = 0$$

The advantage of using the nonlinear structure described in this section over a general nonlinear filter, is that the former requires fewer coefficients than the latter. For example, if the Volterra series nonlinear filter was used to compensate for a nonlinear channel as in Fig. 6.3, 2^N coefficients would be required. If the new nonlinear structure is used $N + M + L$ coefficients would be necessary, where typically $M + L \leq 8$.

6.4 Computer Simulations

The structure of Fig. 6.4 is simulated using the simulation model shown in Fig. 6.5. The input data to the echo channel and the communication channel are uncorrelated ± 1 values. The ratio of near-end to far-end signal power is 40 dB. The nonlinear parts of the echo canceller and of the echo channel are represented by polynomials of 3^{rd} order so that only second and third order nonlinearities are present. The FIR filter consists of 10 coefficients. The communication channel consists only of a linear part, represented by an FIR filter of 6 coefficients. The values of the tap coefficients

and of the nonlinearity parameters are shown in Tables 6.1 and 6.2. Initially the filter coefficients for the echo canceller and all the nonlinearity parameters are 0 except $\hat{\alpha}_{11}$ and $\hat{\alpha}_{21}$ (the linear terms) which are equal to 1.

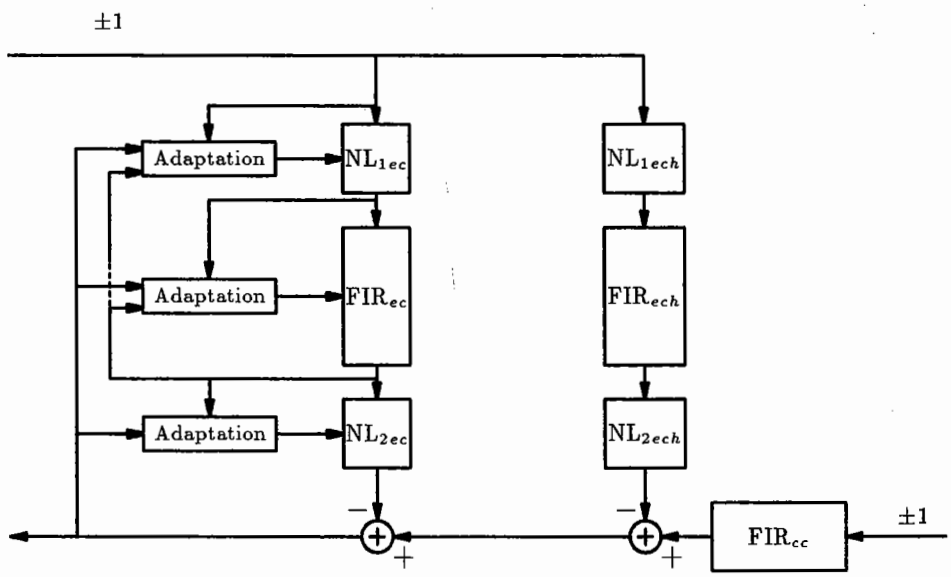


Fig. 6.5 Simulation model for the nonlinear compensator

ECHO CHANNEL									
FIR									
h_0	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9
1.0	0.25	0.11	0.05	0.03	0.02	0.015	0.01	0.009	0.008
NL_1					NL_2				
α_{10}		α_{11}	α_{12}	α_{13}	α_{20}	α_{21}	α_{22}	α_{23}	
0.0		1.0	0.0	0.01	0.0	1.0	0.0	0.01	

Table 6.1 FIR filter coefficients and nonlinearity parameters used for the echo channel

COMMUNICATION CHANNEL					
FIR					
h_0	h_1	h_2	h_3	h_4	h_5
1.0	0.25	0.015	0.01	0.005	0.001

Table 6.2 FIR filter coefficients used for the communication channel

6.4.1 Convergence Curves

The residual echo power to far-end data signal power R_k^2 is plotted against the number of iterations. Eq. (6.9) is used to adapt the coefficients of the FIR filter and the nonlinearity parameters (with $\mu_i = \mu$, $\nu_j = \nu$, $\xi_l = \xi$). Fig. 6.6 shows the convergence curve obtained by using $\mu = 0.009$, $\nu = 0.05$ and $\xi = 0.009$.

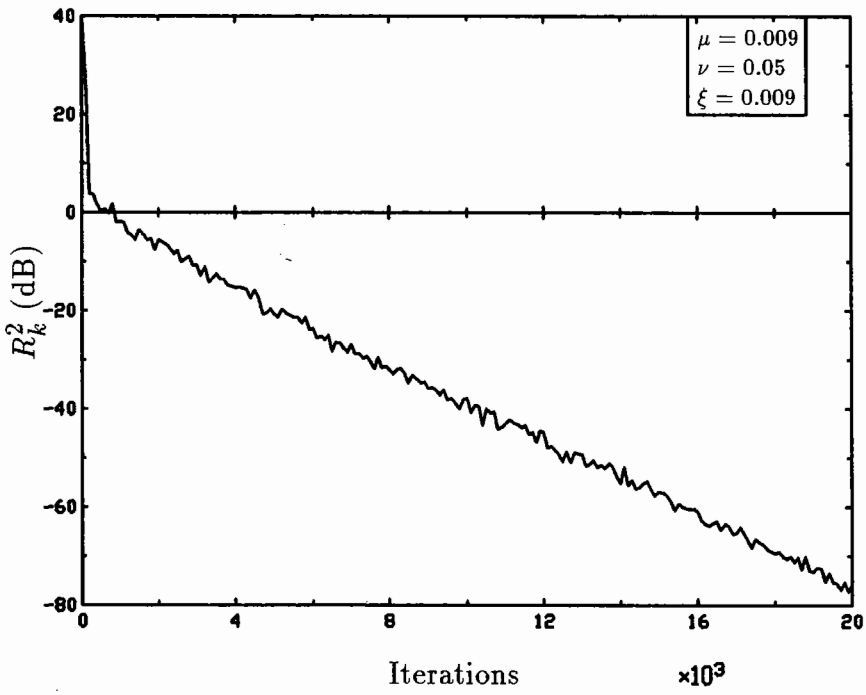


Fig. 6.6 Convergence curve for nonlinear echo canceller (computer simulation)

There are two phases in the convergence curve. The first phase is the time during which the FIR filter converges and is characterized by the sharp decrease of R_k^2 from 40 dB to about 4 dB. The second phase is the time during which the nonlinear parameters start to have significant values so that they can influence the error and eventually bring it to its minimum value and is characterized by the slow decrease of R_k^2 after about 300 iterations.

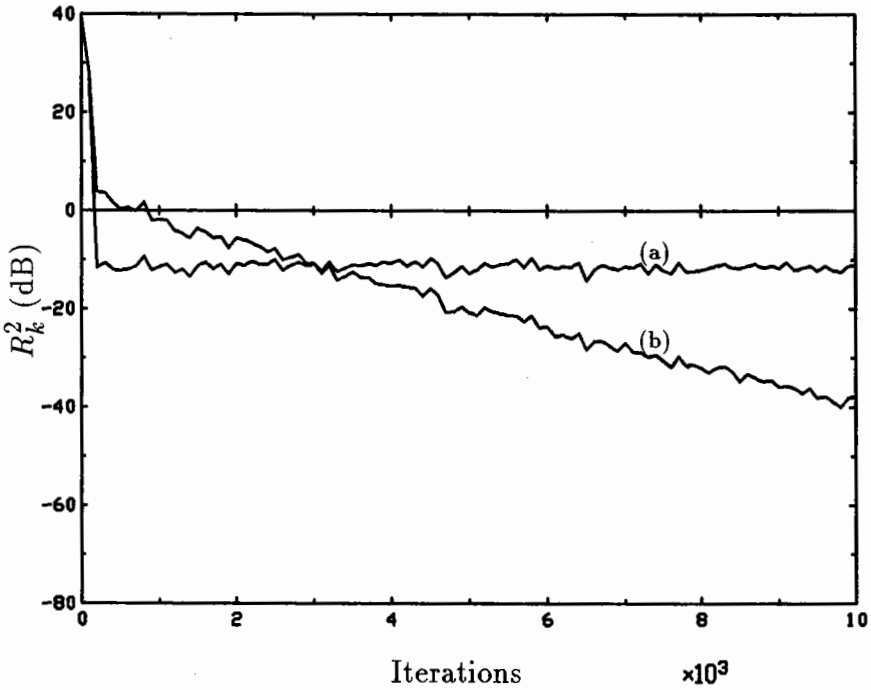


Fig. 6.7 Comparison of linear vs nonlinear echo canceller
(a) Linear ($\nu = 0.05$)
(b) Nonlinear ($\mu = 0.009, \nu = 0.05, \xi = 0.009$)

For purposes of comparison, the evolution of R_k^2 for a linear canceller is depicted in Fig. 6.7 together with the one corresponding to the nonlinear structure. The convergence curve for the linear canceller is implemented by making the stepsizes used for the nonlinear parts μ and ξ equal to zero. Thus no nonlinear parts are

present in the echo canceller structure, and the adaptation equation for the FIR filter coefficients (6.9.b) reduces to the stochastic iteration algorithm. The stepsize for the FIR is $\nu = 0.05$. The difference in the degree of cancellation achieved by the nonlinear structure as compared to a linear structure is considerable.

To test the ability of the echo canceller to adapt to changes in the echo channel and to see if it converges to the same minimum point, one of the coefficients of the FIR filter of the echo channel is changed. Fig. 6.8 shows the convergence curve obtained when the coefficient h_1 is changed from 0.25 to 0.2. The change is made at iteration 30,000 and as it can be seen the canceller converges to the same minimum point.

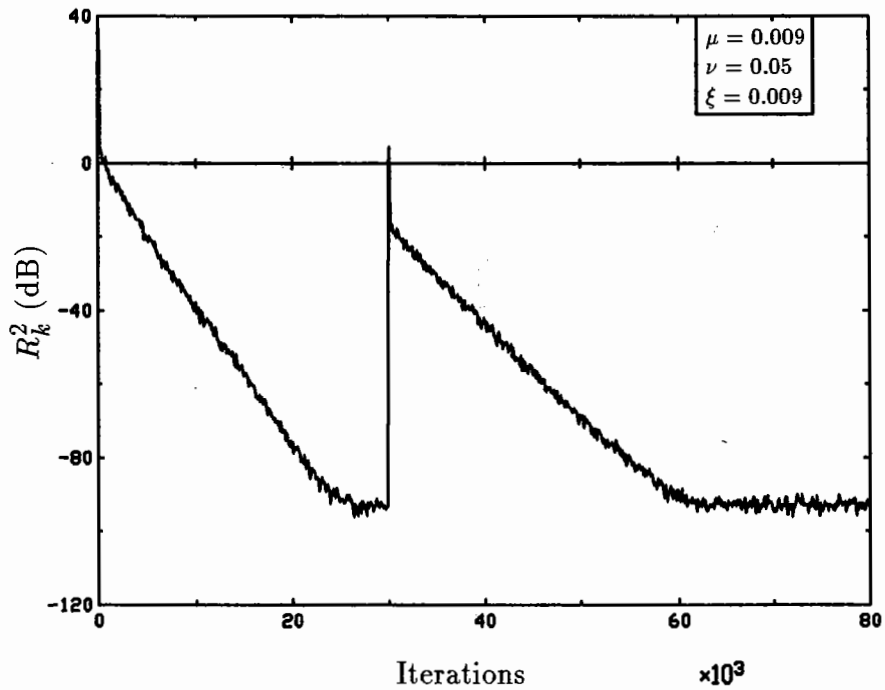


Fig. 6.8 Test of ability of echo canceller to adapt to changes in the echo channel

Chapter 7 Summary and Conclusions

7.1 Summary

This thesis has dealt with nonlinear adaptive filtering for echo cancellation and decision feedback equalization. Four models have been presented, three for echo cancellation and one for combined echo cancellation and decision feedback equalization. For the latter case, two more alternatives were suggested.

Chapter 2 described the sources of nonlinear distortion. The most important sources are the data converters in the echo path and in the implementation of the echo canceller itself, the transmitted pulse asymmetry, and the saturation of the transformers. The exact placement of these sources depends on the application (digital subscriber loops or voiceband data modems) and on the implementation of the modem. In the case of voiceband data modems, however, the communication channel has more severe nonlinearities than the communication channel in the digital subscriber loops. The Volterra series expansion, being general in form, was found to adequately describe most of the configurations encountered in practice.

Chapter 3 examined the table look-up structure and its convergence characteristics for the stochastic iteration algorithm and the sign algorithm. Although the sign algorithm is simpler to implement, it results in a much slower convergence as compared to the stochastic iteration algorithm. Also the sign algorithm is highly dependent on the probability density function of the received far-end signal. However even with the stochastic iteration algorithm, the table look-up structure has a much slower convergence time as compared to a linear transversal filter (though the latter cannot compensate for nonlinearities).

The adaptive reference echo cancellation scheme in Chapter 4 used a table look-up structure for both the echo canceller and the reference-former. This scheme has a much faster convergence than an echo cancellation scheme with only one memory if the same steady state value is to be reached. It can also reach much lower steady state values with the same convergence times as the one memory echo cancellation scheme. New expressions for the convergence properties were presented. Two simple start-up procedures, not involving transmission of an ideal reference or any other special action on the part of the transmitting modem, have been demonstrated by simulation. Both methods achieve the same steady state performance in the simulations, although the initial convergence for the two-stage method is faster. For this reason it may be the preferred method, although the second method with the reference-former continually adapting, may be slightly simpler. The associated convergence times, although somewhat longer than those of the ideal-reference start-up, are still much less than that of the one memory scheme.

The adaptive reference-former was further exploited as a decision feedback equal-

izer. Only simple arithmetic operations were needed due to the simple relation that was found to exist between the memory contents of the RAM and the Volterra coefficients. The two other alternatives suggested were also based on the idea of combining adaptive reference echo cancellation with decision feedback equalization. The alternative incorporating the reference-former and the echo canceller in a single nonlinear filter requires the least number of coefficients. The choice among the three alternatives will depend on the speed of operation, the complexity of the hardware, and the need to deal with nonlinearities.

The combination of a linear FIR filter with the table look-up structure presented in Chapter 5 is a different approach in reducing the convergence time of the latter. From the simulations presented we have seen that the FIR filter is able to compensate for the linear part of the echo channel with its characteristic fast convergence time. Later in the process the table look-up structure takes over, reducing the mean square error to a lower value by compensating for the nonlinear part. Two start-up methods were also demonstrated in the simulations. The two-stage method is preferable due to the need in practice for initial fast compensation of the linear part of the echo channel. An expression for the convergence of a linear FIR filter under a nonlinear echo channel constraint was also presented.

The nonlinear compensator in Chapter 6 (nonlinear-linear-nonlinear) was the only model not based on the table look-up structure. Its advantage over general nonlinear filters is its simple structure and the low number of coefficients and nonlinearity parameters required (2^N for a general nonlinear filter, $N + M + L$ for the nonlinear compensator where typically $M + L \leq 8$). The simulation results demonstrate that

it is well suited for the compensation of a specific nonlinear channel. The FIR filter can compensate for the linear part of the echo channel with a fast convergence time and the nonlinear parts can compensate, at a lower rate, for the nonlinear part. The adaptation procedure however is more complicated than the simple stochastic iteration algorithm. Also the adaptive procedure might find local minima.

7.2 Conclusions and Suggestions for Future Research

In almost all the simulations (except in Chapter 6) the most general echo and communication channels were considered so that the nonlinear adaptive filters can compensate for general nonlinearities in both digital subscriber loops and voiceband data modems applications. In the latter case the baseband configurations should be extended to the passband with appropriate modulation techniques.

In general, the degree of cancellation required differs (1) for the reference-former (2) for the decision feedback equalizer and (3) for the echo canceller. (1) requires the least amount of cancellation and (3) the most. Taking this into account some reductions in complexity can be considered in the general models presented. For example, the decision feedback equalizer configuration presented in Chapter 4 (alternative 0) may need only the computation of $N-1$ coefficients (instead of 2^{N-1} in the worst case) in the case of digital subscriber loops where nonlinear equalization is not so important (in contrast with voiceband data modems). Also the reference-former can use less memory than the echo canceller.

If nonlinear echo cancellation is to be avoided then some tradeoffs exist: (1) the required echo canceller accuracy can be minimized by use of a selective termination

for the hybrid and (2) the data converters can be designed with sufficient linearity to allow a linear echo canceller.

We conclude that the choice of the optimum structure for any particular application will highly depend upon such factors as

- required number of taps,
- necessity to deal with nonlinearities,
- maximum allowable convergence time,
- maximum allowable power consumption, and
- available IC (integrated circuit) technology.

The specifications of the ISDN (requiring few taps, high dynamic range, and low power consumption) make the table look-up structures serious candidates in this application.

Other issues for further discussion could be (1) the combination of the DFE structures presented with forward equalization and (2) the problem of timing jitter. Timing jitter results in small variations in the sampling rate, which in turn cause slow fluctuations in the sampled impulse response seen by the echo canceller (also the impulse response seen by the reference-former). The faster convergence properties of the two-memory structure and the combined structure may allow it to track and compensate for this time variation more effectively.

Appendix A. Convergence of the Two-Memory Structure

In this appendix we derive the convergence properties of the echo canceller that has the two-memory structure (Section 4.1) and that uses the stochastic iteration algorithm.

A.1 Convergence for the Echo Canceller (RAM1)

The adaptation equation for the echo canceller is given by

$$\hat{e}_{k+1} = \hat{e}_k + \alpha_C r'_k \tag{A.1}$$

where $\hat{e}_k$ represents the output of RAM1 at time k. For a particular address i ($i = 0, 1, 2, \cdots, N_C - 1$) the contents of RAM1 are represented by $\tilde{e}_{ik}$. If at time k, $i = l$ then $\hat{e}_k = \tilde{e}_{lk}$ and

$$\tilde{e}_{i\ k+1} = \begin{cases} \tilde{e}_{ik} & i \neq l \\ \tilde{e}_{ik} + \alpha_C r'_k & i = l \end{cases} \tag{A.2}$$

We now make the following definitions

$$\epsilon_{1,k} \stackrel{\triangle}{=} e_k - \hat{e}_k = \tilde{\epsilon}_{1,lk} \tag{A.3}$$

$$\epsilon_{2,k} \stackrel{\triangle}{=} u_k - \hat{u}_k = \tilde{\epsilon}_{2,mk} \tag{A.4}$$

where $\hat{u}_k$ represents the output of RAM2 at time k and for a particular address j ($j = 0, 1, 2, \cdots, N_R - 1$) the contents of RAM2 are $\tilde{u}_{jk}$. If at time k, $j = m$ then $\hat{u}_k = \tilde{u}_{mk}$. The error signal for each address i of RAM1 at time k is $\tilde{\epsilon}_{1,ik}$ and the error signal for RAM2 is $\epsilon_{2,k}$. The control signal used for the adaptation of both

memories is given by

$$\begin{aligned} r'_k &= r_k - \hat{u}_k \\ &= e_k - \hat{e}_k + u_k - \hat{u}_k + n_k \\ &= \tilde{\epsilon}_{1,lk} + \tilde{\epsilon}_{2,mk} + n_k \end{aligned} \tag{A.5}$$

Substituting into Eq. (A.2) we get

$$\tilde{e}_{i\ k+1} = \begin{cases} \tilde{e}_{ik} & i \neq l \\ \tilde{e}_{ik} + \alpha_C \left[\tilde{\epsilon}_{1,ik} + \tilde{\epsilon}_{2,jk} + n_k\right] & i = l \end{cases} \tag{A.6}$$

Subtracting both sides of Eq. (A.6) from e_k, noting that for a stationary channel $e_{k+1} = e_k$, and using Eq. (A.3) and Eq. (A.4) we get

$$\tilde{\epsilon}_{1,ik+1} = \begin{cases} \tilde{\epsilon}_{1,ik} & i \neq l \\ \tilde{\epsilon}_{1,ik}(1-\alpha_C) - \alpha_C \tilde{\epsilon}_{2,jk} - \alpha_C n_k & i = l \end{cases} \tag{A.7}$$

Define now

$$\begin{aligned} \varepsilon_{1,ik} &\stackrel{\triangle}{=} E[\tilde{\epsilon}^2_{1,ik}] \quad i = 0, 1, 2, \cdots, N_C - 1 \\ \varepsilon_{2,k} &\stackrel{\triangle}{=} E[\epsilon^2_{2,k}] \end{aligned} \tag{A.8}$$

By squaring both sides of Eq. (A.7) and by taking expectations we get

$$\begin{aligned} \varepsilon_{1,i\ k+1} = P(i \neq l)\varepsilon_{1,ik} + P(i = l)\Big[&(1-\alpha_C)^2 E[\tilde{\epsilon}^2_{1,ik}] \\ &+ \alpha_C^2 E[\epsilon^2_{2,k}] + \alpha_C^2 E[n_k^2] + 2\alpha_C E[\epsilon_{2,k} n_k] \\ &- 2(1-\alpha_C)\alpha_C \left(E[\tilde{\epsilon}_{1,ik}\epsilon_{2,k}] + E[\tilde{\epsilon}_{1,ik} n_k]\right)\Big] \end{aligned} \tag{A.9}$$

where $P(\cdot)$ denotes probability. For equiprobable addresses $P(i = l) = \frac{1}{2^{N_C}}$ and by assuming that the near-end data and far-end data are uncorrelated and that the additive noise n_k is also assumed to be uncorrelated with the data, Eq. (A.9) reduces to

$$\varepsilon_{1,i\ k+1} = \varepsilon_{1,ik}\ \rho_C + \frac{\alpha_C^2}{2^{N_C}}\varepsilon_{2,k} + \frac{\alpha_C^2}{2^{N_C}}\sigma_n^2 \tag{A.10}$$

for $i = 0, 1, 2, \cdots, N_C - 1$, where

$$\rho_C = 1 - \frac{\alpha_C(2 - \alpha_C)}{2^{N_C}}$$

What we really want to find is $E[\epsilon_{1,k}^2] = E[(e_k - \hat{e}_k)^2]$. This is given by

$$\begin{aligned} E[\epsilon_{1,k}^2] &= \sum_{i=0}^{N_C-1} E[\epsilon_{1,k}^2/l = i)]P(l = i) \\ &= \sum_{i=0}^{N_C-1} E[\tilde{\epsilon}_{1,ik}^2]P(l = i) \\ &= \sum_{i=0}^{N_C-1} \varepsilon_{1,ik}P(l = i) \end{aligned} \tag{A.11}$$

If $\varepsilon_{1,k} \stackrel{\triangle}{=} E\left[\epsilon_{1,k}^2\right]$ from Eq. (A.11) and Eq. (A.10) we get

$$\varepsilon_{1,k+1} = \varepsilon_{1,k}\ \rho_C + \frac{\alpha_C^2}{2^{N_C}}\varepsilon_{2,k} + \frac{\alpha_C^2}{2^{N_C}}\sigma_n^2 \tag{A.12}$$

which gives the convergence equation for the echo canceller.

A.2 Convergence for the Reference-former (RAM2)

The equations governing the convergence of the reference-former can be derived in exactly the same way as for the echo canceller. Thus by interchanging $\varepsilon_{1,k}$ with $\varepsilon_{2,k}$, α_C with α_R, and N_C with N_R in Eq. (A.12) we get the following equation for the convergence of the reference-former

$$\varepsilon_{2,k+1} = \varepsilon_{2,k}\ \rho_R + \frac{\alpha_R^2}{2^{N_R}}\varepsilon_{1,k} + \frac{\alpha_R^2}{2^{N_R}}\sigma_n^2 \tag{A.13}$$

where

$$\rho_R = 1 - \frac{\alpha_R(2 - \alpha_R)}{2^{N_R}}$$

A.3 The Steady State Mean Square Error

We wish now to find the steady state mean square error $\varepsilon_{1,\infty}$. Define a two dimensional state vector,

$$\mathbf{A}_k \triangleq \begin{bmatrix} \varepsilon_{1,k} \\ \varepsilon_{2,k} \end{bmatrix}$$

Then equations (A.12) and (A.13) can be expressed in a matrix form as in Eq. (A.14).

$$\mathbf{A}_{k+1} = \mathbf{B}\mathbf{A}_k + \mathbf{C} \tag{A.14}$$

or

$$\begin{bmatrix} \varepsilon_{1,k+1} \\ \varepsilon_{2,k+1} \end{bmatrix} = \begin{bmatrix} \rho_C & \frac{\alpha_C^2}{2^{N_C}} \\ \frac{\alpha_R^2}{2^{N_R}} & \rho_R \end{bmatrix} \begin{bmatrix} \varepsilon_{1,k} \\ \varepsilon_{2,k} \end{bmatrix} + \begin{bmatrix} \frac{\alpha_C^2}{2^{N_C}}\sigma_n^2 \\ \frac{\alpha_R^2}{2^{N_R}}\sigma_n^2 \end{bmatrix}$$

If Eq. (A.14) is iterated we get

$$\mathbf{A}_k = \mathbf{B}^k\mathbf{A}_0 + \sum_{i=0}^{k-1} \mathbf{B}^i\mathbf{C} \tag{A.15}$$

It should be noted that $\mathbf{B}$ is independent of the echo channel and communication channel impulse responses.

The steady state value is given by

$$\lim_{k\to\infty} \mathbf{A}_k = (\mathbf{I} - \mathbf{B})^{-1}\mathbf{C} \tag{A.16}$$

($\mathbf{I}$ is the identity matrix). From Eq. (A.16) we can find $\varepsilon_{1,\infty}$ and this is equal to

$$\varepsilon_{1,\infty} = \frac{\alpha_C\sigma_n^2}{2 - \alpha_C - \alpha_R}$$

The transient term $\mathbf{B}^k\mathbf{A}_0$ involves the k^{th} powers of the eigenvalues $\{\lambda\}$ of $\mathbf{B}$. It will converge to zero if all eigenvalues are within the unit circle. Its rate of convergence is governed by the largest eigenvalue.

Appendix B. Relation between the Memory Contents of the RAM and the Volterra Coefficients

In this appendix we find a new representation for the input/output relationship of a nonlinear channel in the case of binary signals. This representation will lead to the relation between the memory contents of the RAM and the Volterra coefficients.

Let us assume that the nonlinear channel can be represented by the truncated Volterra series expansion as follows

$$u_k = h^{(0)} + \sum_{i_1=0}^{N-1} h_{i_1}^{(1)} b_{k-i_1} + \sum_{i_1=0}^{N-1}\sum_{i_2=0}^{N-1} h_{i_1 i_2}^{(2)} b_{k-i_1} b_{k-i_2} + \cdots + \cdots + \sum_{i_1=0}^{N-1} \cdots \sum_{i_n=0}^{N-1} h_{i_1 \ldots i_n}^{(n)} b_{k-i_1} \cdots b_{k-i_n} \,. \tag{B.1}$$

The factors $h^{(0)}, h_0^{(1)}, h_1^{(1)}, \ldots$ represent the Volterra coefficients, b_k are the input data symbols to the channel, and the u_k are the output data symbols. In the case of binary signals a finite number of coefficients is required to represent the nonlinear signal u_k (this is also true for any case where there is a finite number of possible transmitted signals). This is demonstrated by an example and then the results are generalized.

B.1 Example

For the case where $N = 2$ and $n = 3$ Eq. (B.1) reduces to

$$\begin{aligned} u_k = {} & h^{(0)} + h_0^{(1)} b_k + h_1^{(1)} b_{k-1} \\ & + h_{00}^{(2)} b_k b_k + h_{01}^{(2)} b_k b_{k-1} + h_{11}^{(2)} b_{k-1} b_{k-1} + h_{10}^{(2)} b_{k-1} b_k \\ & + h_{000}^{(3)} b_k b_k b_k + h_{001}^{(3)} b_k b_k b_{k-1} + h_{010}^{(3)} b_k b_{k-1} b_k + h_{011}^{(3)} b_k b_{k-1} b_{k-1} \\ & + h_{100}^{(3)} b_{k-1} b_k b_k + h_{101}^{(3)} b_{k-1} b_k b_{k-1} + h_{110}^{(3)} b_{k-1} b_{k-1} b_k + h_{111}^{(3)} b_{k-1} b_{k-1} b_{k-1} \,. \end{aligned} \tag{B.2}$$

It is now convenient to use the fact that the Volterra coefficients are symmetric functions of their indices. Thus the complexity of Eq. (B.2) is reduced since we group together terms differing only by a permutation of their indices (for example $b_k b_{k-1}$ and $b_{k-1} b_k$). Then Eq. (B.2) reduces to

$$\begin{aligned} u_k = {} & h^{(0)} + h_0^{(1)} b_k + h_1^{(1)} b_{k-1} \\ & + h_{00}^{(2)} b_k^2 + h_{01}^{(2)} b_k b_{k-1} + h_{11}^{(2)} b_{k-1}^2 \\ & + h_{000}^{(3)} b_k^3 + h_{001}^{(3)} b_k^2 b_{k-1} + h_{011}^{(3)} b_k b_{k-1}^2 + h_{111}^{(3)} b_{k-1}^3 \, . \end{aligned} \tag{B.3}$$

For $b_k = \pm 1$ the even powers of b_k and b_{k-1} are $+1$ and thus their Volterra coefficients can be combined with $h^{(0)}$ to form the constant term. Also the odd powers will take the same values as b_k and b_{k-1} and their Volterra coefficients can be combined with those of b_k and b_{k-1}. Thus we get Eq. (B.4) where the combined h coefficients result to a new coefficient g (the letter g will from now on be used to represent the Volterra coefficients in the new representation)

$$u_k = g^{(0)} + g_0^{(1)} b_k + g_1^{(1)} b_{k-1} + g^{(2)} b_k b_{k-1} \, . \tag{B.4}$$

If $N = 3$ and $n = 3$, u_k is given by

$$\begin{aligned} u_k = g^{(0)} + g_0^{(1)} b_k + g_1^{(1)} b_{k-1} + g_2^{(1)} b_{k-2} + g_{01}^{(2)} b_k b_{k-1} + g_{02}^{(2)} b_k b_{k-2} \\ + g_{12}^{(2)} b_{k-1} b_{k-2} + g^{(3)} b_k b_{k-1} b_{k-2} \, . \end{aligned} \tag{B.5}$$

In the case where $N = 3$ and $n > 3$ the expression for u_k is still the same as in Eq. (B.5).

B.2 General case

The above results can be generalized to the N^{th} case

$$\begin{aligned} u_k &= f(b_k, b_{k-1}, \ldots, b_{k-N+1}) \\ &= g^{(0)} + \sum_{i=0}^{N-1} g_i^{(1)} b_{k-i} + \sum_{i_1 \neq i_2}^{N-1} g_{i_1 i_2}^{(2)} b_{k-i_1} b_{k-i_2} + \cdots \\ &\cdots + \sum_{i_1 \neq i_2 \cdots \neq i_{N-1}}^{N-1} g_{i_1\ i_2 \ldots i_{N-1}}^{(N-1)} b_{k-i_1}\ b_{k-i_2} \cdots\ b_{k-i_{N-1}} + g^{(N)} b_k b_{k-1} \cdots b_{k-N+1} \end{aligned} \tag{B.6}$$

The total number of terms can be obtained by observing the number of combinations of N bits taken M at a time ($M = 0, 1, \cdots, N$). Thus the total number of terms in Eq. (B.6) is

$$\sum_{M=0}^{N} \binom{N}{M} = \sum_{M=0}^{N} \frac{N!}{M!(N-M)!} = 2^N \tag{B.7}$$

which is finite ($2^2 = 4$ for Eq. (B.4) and $2^3 = 8$ for Eq. (B.5)). It should be noted that Eq. (B.6) represents the most general case and that its depended only on N (the memory length) and not on n (the order of nonlinearity). Depending on the value of n, in practical situations, some terms of Eq. (B.6) may be zero and thus less terms would be required. Two cases can be distinguished

$$\text{number of terms in Eq. (B.6)} = \begin{cases} 2^N & n \geq N \\ \sum_{M=0}^{n} \binom{N}{M} & n < N \end{cases} \tag{B.8}$$

If the table look-up structure is used to form the value of u_k then 2^N memory locations are needed which will be addressed by N bits. In the case where $b_k = \pm 1$ the relation between the contents of the memory of the RAM and the Volterra coefficients is found by assigning all the combinations for b_k, b_{k-1} and b_{k-2} in Eq. (B.5) (or in

the general case by assigning all combinations of $b_k, b_{k-1}, \cdots, b_{k-N+1}$ in Eq. (B.6)) which in matrix form is

$$\begin{bmatrix} f(-1,-1,-1) \\ f(-1,-1,+1) \\ f(-1,+1,-1) \\ f(-1,+1,+1) \\ f(+1,-1,-1) \\ f(+1,-1,+1) \\ f(+1,+1,-1) \\ f(+1,+1,+1) \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} g^{(0)} \\ g_0^{(1)} \\ g_1^{(1)} \\ g_2^{(1)} \\ g_{01}^{(2)} \\ g_{02}^{(2)} \\ g_{12}^{(2)} \\ g^{(3)} \end{bmatrix} \tag{B.9}$$

where $f(b_k, b_{k-1}, b_{k-2})$ represent the contents of one of the 8 memory locations of the RAM. The address is specified by b_k, b_{k-1}, and b_{k-2}. This matrix is orthogonal and its inverse can be found very easily (actually the inverse is its transpose with all the terms multiplied by 0.125).

From the above analysis is concluded that the memory contents of the RAM can be found from the Volterra coefficients and vice-versa.

Appendix C. Convergence of a linear FIR filter under a Nonlinear Echo Channel Constraint

In this appendix we derive the equation governing the convergence of a linear FIR filter under a nonlinear echo channel constraint.

First we make the following definitions

$$\mathbf{a}_{1k} \stackrel{\triangle}{=} (a_k, a_{k-1}, \cdots, a_{k-N+1})^T$$

$$\mathbf{a}_{2k} \stackrel{\triangle}{=} (a_k a_{k-1}, a_k a_{k-2}, \cdots, a_k a_{k-1} \cdots a_{k-N+1})^T$$

$$\mathbf{g}_1 \stackrel{\triangle}{=} (g_0^{(1)}, g_1^{(1)}, \cdots, g_{N-1}^{(1)})^T$$

$$\mathbf{g}_2 \stackrel{\triangle}{=} (g^{(0)}, g_{01}^{(2)}, g_{02}^{(2)}, \cdots, g^{(N)})^T$$

$$\mathbf{h}_k \stackrel{\triangle}{=} (h_{0k}, h_{1k}, \cdots, h_{N-1k})^T$$

The dimensions of the above vectors are

$\mathbf{a}_{1k} =< N >$ linear transmitted data vector

$\mathbf{a}_{2k} =< 2^N - N >$ nonlinear transmitted data vector

$\mathbf{g_1} =< N >$ linear Volterra coefficient vector

$\mathbf{g_2} =< 2^N - N >$ nonlinear Volterra coefficient vector

$\mathbf{h}_k =< N >$ FIR filter coefficient vector

We assume that the stochastic iteration algorithm is used for the adaptation of the filter coefficients which uses the following equation

$$\mathbf{h}_{k+1} = \mathbf{h}_k + 2\ \alpha_F\ r_k\ \mathbf{a}_{1k} \tag{C.1}$$

where

$$r_k = e_k - \hat{e}_k + u_k$$

and α_F is the stepsize.

Using the above definitions the error signal is equal to

$$\begin{aligned} e_k - \hat{e}_k =& \mathbf{a}_{1k}^T \mathbf{g_1} + \mathbf{a}_{2k}^T \mathbf{g_2} - \mathbf{a}_{1k}^T \mathbf{h}_k \\ =& \mathbf{a}_{1k}^T (\mathbf{g_1} - \mathbf{h}_k) + \mathbf{a}_{2k}^T \mathbf{g_2} \ . \end{aligned} \tag{C.2}$$

By denoting ε_k as the mean square error we get

$$\begin{aligned} \varepsilon_k =& E\Big[(e_k - \hat{e}_k)^2\Big] \\ =& E\Big[\big[\mathbf{a}_{1k}^T (\mathbf{g_1} - \mathbf{h}_k) + \mathbf{a}_{2k}^T \mathbf{g_2}\big]^2\Big] \ . \end{aligned} \tag{C.3}$$

After some vector algebra and by recalling the assumptions about the statistical independence of the data symbols $\{a_k\}$ and about the statistical independence of $\mathbf{a}_{1k}$ and $\mathbf{h}_k$ Eq. (C.3) reduces to

$$\begin{aligned} \varepsilon_k =& E\Big[(\mathbf{g_1} - \mathbf{h}_k)^T (\mathbf{g_1} - \mathbf{h}_k)\Big] + E\Big[\mathbf{g_2}^T \mathbf{g_2}\Big] \\ =& \varepsilon_{1k} + \varepsilon_{2k} \end{aligned} \tag{C.4}$$

where

$$\varepsilon_{1k} = E\Big[(\mathbf{g_1} - \mathbf{h}_k)^T (\mathbf{g_1} - \mathbf{h}_k)\Big]$$

$$\varepsilon_{2k} = E\Big[\mathbf{g_2}^T \mathbf{g_2}\Big] \ .$$

From Eq. (C.1) and Eq. (C.2) we get

$$\mathbf{h}_k = \mathbf{h}_{k-1} + 2\alpha_F \mathbf{a}_{1k-1}\Big[\mathbf{a}_{1k-1}^T (\mathbf{g_1} - \mathbf{h}_{k-1}) + \mathbf{a}_{2k-1}^T \mathbf{g_2}\Big] + 2\alpha_F \mathbf{a}_{1k-1} u_{k-1} \ . \tag{C.5}$$

Substituting Eq. (C.5) into Eq. (C.4) we get the following expression for the mean square error

$$\begin{aligned} \varepsilon_k =& (1 - 4\alpha_F + 4\alpha_F^2 N)\Big[E\Big[(\mathbf{g_1} - \mathbf{h}_{k-1})^T (\mathbf{g_1} - \mathbf{h}_{k-1})\Big] + E\Big[\mathbf{g_2}^T \mathbf{g_2}\Big]\Big] \\ &+ 4\alpha_F^2 N \sigma_u^2 + 4\alpha_F E\Big[\mathbf{g_2}^T \mathbf{g_2}\Big] \ . \end{aligned} \tag{C.6}$$

From Equations (C.4), (C.6) and by noting that $E\Big[\mathbf{g_2}^T \mathbf{g_2}\Big] = \mathbf{g_2}^T \mathbf{g_2}$ we get the following iterative equation for the mean square error

$$\varepsilon_k = (1 - 4\alpha_F + 4\alpha_F^2 N)\varepsilon_{k-1} + 4\alpha_F^2 N \sigma_u^2 + 4\alpha_F (\mathbf{g_2}^T \mathbf{g_2}) \ . \tag{C.7}$$

From the above equation we can also get the expression for ε_{1k} which represents the mean square error for the linear part of the channel only.

$$\varepsilon_{1k} = (1 - 4\alpha_F + 4\alpha_F^2 N)\varepsilon_{1k-1} + 4\alpha_F^2 N(\sigma_u^2 + \mathbf{g_2}^T\mathbf{g_2}) \; . \tag{C.8}$$

By defining R_k^2 as the ratio of the residual echo power to the far-end signal power,

$$R_k^2 \triangleq \frac{\varepsilon_k}{\sigma_u^2} \tag{C.9}$$

and G as

$$G \triangleq \left(\mathbf{g_2}^T\mathbf{g_2}\right) \tag{C.10}$$

from Eq. (C.7) we get

$$R_k^2 = \rho^k \left[R_0^2 - \frac{\alpha_F N + G/\sigma_u^2}{(1 - \alpha_F N)}\right] + \frac{\alpha_F N + G/\sigma_u^2}{(1 - \alpha_F N)} \tag{C.11}$$

where

$$\rho \triangleq (1 - 4\alpha_F + 4\alpha_F^2 N) \; .$$

Eq. (C.11) represents the convergence equation for a linear FIR filter under a nonlinear echo channel constraint.

References

1. D.G. Messerschmitt, "Echo cancellation in speech and data transmission", *IEEE Journal on Selected Areas in Communications*, vol. SAC-2, No. 2, pp. 283–297, March, 1984.
2. A.B. Clark and R.C. Mothes, "Echo suppressors for long telephone circuits", *Proc. AIEE*, vol. 44, pp. 481–490, April, 1925.
3. M.M. Sondhi, "An adaptive echo canceller", *Bell Syst. Tech. J.*, vol. 46, No. 3, pp. 497–511, March, 1967.
4. N. Holte and S. Stueflotten, "A new digital echo canceller for two-wire subscriber lines", *IEEE Transactions on Communications*, vol. COM-29, No. 11, pp. 1573–1581, November, 1981.
5. O. Agazzi, D.G. Messerschmitt and D.A. Hodges, "Nonlinear echo cancellation of data signals", *IEEE Transactions on Communications*, vol. COM-30, No. 11, pp. 2421–2433, November, 1982.
6. B.X.C. Maitre, M.Y. Levy, S.R. Surie, "The multi memory echo canceller for two wire digital transmission", *IEEE Global Telecommunications Conference*, Atlanta, Georgia, pp. 196–202, November, 1984.
7. P.V. Gerwen, N.A.M. Verhoeckx and T.A.C.M. Claasen, "Design considerations for a 144 kbits/s digital transmission unit for the local telephone network", *IEEE Journal on Selected Areas in Communications*, vol. SAC-2, No. 2, pp. 314–323, March, 1984.
8. C.F.N. Cowan, M.J. Smith and P.F. Adams, "A nonlinear adaptive echo canceller for a digital transmission system", *IASTED, Int. Sym. on Applied Sig. Proc. and Dig. Filtering*, Paris, France, pp. 19–21, June, 1985.
9. M.J. Coker and D.N. Simkins, "A nonlinear adaptive noise canceller", *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, Denver, Colorado, vol. 2, pp. 470–473, April, 1980.
10. O. Agazzi, D.A. Hodges and D.G. Messerschmitt, "Large-scale integration of hybrid-method digital subscriber loops", *IEEE Transactions on Communications*, vol. COM-30, No. 9, pp. 2095–2108, September, 1982.
11. D.D. Falconer, "Timing jitter effects on digital subscriber loop echo cancellers: Part II- considerations for squaring loop timing recovery", *IEEE Transactions on Communications*, vol. COM-33, No. 8, pp. 833–838, August, 1985.
12. N.R. Grossner, *Transformers for Electronic Circuits*, McGraw-Hill Book Company, 1967.
13. B. Fotouhi and D.A. Hodges, "High-resolution A/D conversion in MOS/LSI", *IEEE Journal of Solid-State Circuits*, vol. SC-14, No. 6, pp. 920–926, December, 1979.
14. S.V. Ahmed, P.P. Bohn, and N.L. Gottfried, "A tutorial on two-wire digital transmission in the loop plant", *IEEE Transactions on Communications*, vol. COM-29, No. 9, pp. 1554–1564, November, 1981.

15. P.H. Wittke, S.R. Penstone, and R.T. Keightley, "Measurements of echo parameters pertinent to high-speed full-duplex data transmission on the telephone circuits", *IEEE Journal on Selected Areas in Communications*, vol. SAC-2, No. 2, pp. 703–710, September, 1984.

16. F.P. Duffy and T.W. Thatcher, Jr, "Analog transmission performance on the switched telecommunications network", *Bell Syst. Tech. J.*, No. 4, pp. 1311–1347, April, 1971.

17. R.R Anderson and D.D. Falconer, "Modem evaluation on real channels using computer simulation", *Proc. National Telecomm. Conf.*, San Diego, pp. 877–883, December, 1974.

18. J.C. Stapleton and S.C. Bass, "Adaptive noise cancelling for a class of nonlinear, dynamic reference channels", *IEEE Transactions on Circuits and Systems*, vol. CAS-32, No. 2, pp. 143–150, February, 1985.

19. D.L. Duttweiler, "Adaptive filter performance with nonlinearities in the correlation multiplier", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-30, pp. 578–586, August, 1982.

20. N.A.M. Verhoeckx, H.C. van Den Elzen, F.A.M. Snijders and P.J. van Gerwen, "Digital echo cancellation for baseband data transmission", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-27, No. 6, pp. 768–781, December, 1979.

21. T.A.C.M. Claasen and W.F.G. Mecklenbrauker, "Comparison of the convergence of two algorithms for adaptive FIR digital filters", *IEEE Transactions on Circuits and Systems*, vol. CAS-28, No. 6, pp. 510–518, June, 1981.

22. D.D. Falconer, "Adaptive reference echo cancellation", *IEEE Transactions on Communications*, vol. COM-30, No. 9, pp. 2083–2094, September, 1982.

23. J.R. Casar-Corredera, M. Carcia-Otero and A.R. Figueiras-Vidal, "Data echo nonlinear cancellation", *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, Tampa, Florida, vol. 3, pp. 1245–1248, March, 1985.