Adaptive Filtering: Algorithms and Structures

by

Azzedine Zerguine

February 2008

Chapter 1

Adaptive Filtering: Algorithms and Structures

1.1 Introduction

Adaptive systems are playing a vital role in the development of modern telecommunications. Also, adaptive systems proved to be extremely effective in achieving high efficiency, high quality and high reliability of around-the-world ubiquitous telecommunication services.

The role of adaptive systems is wide spread covering almost all aspects of telecommunication engineering, but perhaps most notable in the following context [12] of ensuring high-quality signal transmission over unknown and time varying channels.

Interest in adaptive filters continues to grow as they begin to find practical real-time applications in areas such as echo cancellation [11], channel equalisation [13], noise cancellation [14]-[15] and many other adaptive signal processing applications. This is due mainly to the recent advances in the very large-scale integration (VLSI) technology.

The key to successful adaptive signal processing is understanding the fundamental properties of adaptive algorithms. These properties are stability, speed of convergence, misadjustement errors, robustness to both additive noise and signal conditioning (spectral colouration), numerical complexity, and round-off error analysis of adaptive algorithms. However, some of these properties are often in direct conflict with each other, since consistent fast converging algorithms tend to be in general more complex and numerically sensitive. Also, the performance of any algorithm with respect to any of these criteria is entirely dependent on the choice of the adaptation update function, that is the cost function used in the minimisation process. A compromise must be than reached among these conflicting factors to come up with the appropriate algorithm for the concerned application.

After presenting, in Section 2, the common adaptive system configurations using adaptive filters, Finally, Section 3 will deal with a more explicit development of adaptive filters. Performance evaluation of the resulting algorithms using the properties of the finite impulse response (FIR) adaptive filter are also mentioned.

1.2 Applications of adaptive filters

Adaptive filtering has been successfully applied in such diverse fields as communications, radar, sonar, and biomedical engineering. Although these applications are indeed quite different in nature, nevertheless, they have one basic common feature: an input signal and a desired response are used to compute the error, which is in turn used to control the values of a set of adjustable filter coefficients. However, the main difference among the various applications of adaptive filtering arises in the manner in which the desired response is extracted.

In this context, we may classify an adaptive filter into one of the four following categories:

1.2.1 System identification

In this first application, depicted in Fig. 1.1, the adaptive filter is used to provide a linear model that represents the best fit to the unknown system. Both the adaptive filter and the unknown system are driven by the same input. The error estimate is used to update the filter coefficients of the adaptive filter. After convergence, the adaptive filter output will approximate the output of the unknown system in an optimum sense. Provided that the order of the adaptive filter matches that of the unknown system and the input, $x(n)$, is broad band (flat spectrum) this will be achieved by convergence of adaptive filter coefficients to the same values as the unknown system.

The major practical use of this structure in telecommunications is for echo cancellation [11], [16]-[17]. Typically, the input signal $x(n)$ will be either speech or data.

1.2.2 Inverse modelling

In this second class of applications, the function of the adaptive filter is to provide an inverse model that represents the best fit to the unknown system. Thus, at convergence, the adaptive filter transfer function approximates the inverse of the transfer function of the unknown system. As can be seen from Fig. 1.2, the desired response is a delayed version of the input signal.

The primary use of inverse system modelling is for reducing the effects of intersymbol interference (ISI) in digital receivers. This is achieved through the use of equalisation [13], [18] techniques.

1.2.3 Prediction

In this structure, the function of the adaptive filter is to provide the best prediction of the present value of the input signal from its previous values. The configuration shown in Fig. 1.3 is used for this purpose, where the desired signal, $d(n)$, is the instantaneous value and the input to the adaptive filter is a delayed version of the same signal.

This application is widely used in linear predictive coding (LPC) of speech [19]-[20] and in adaptive differential pulse-code modulation (DPCM) [21]. Another approach to prediction is given in [22].

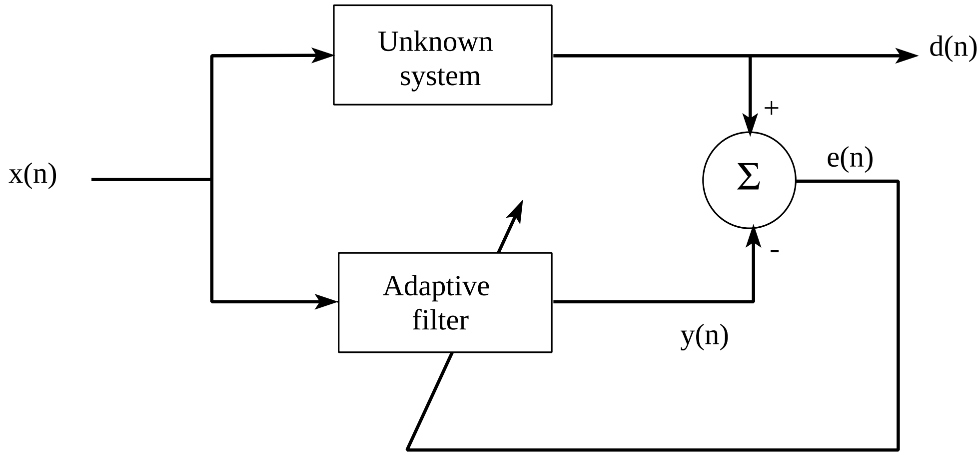

Figure 1.1: Direct system modelling configuration of an adaptive filter.

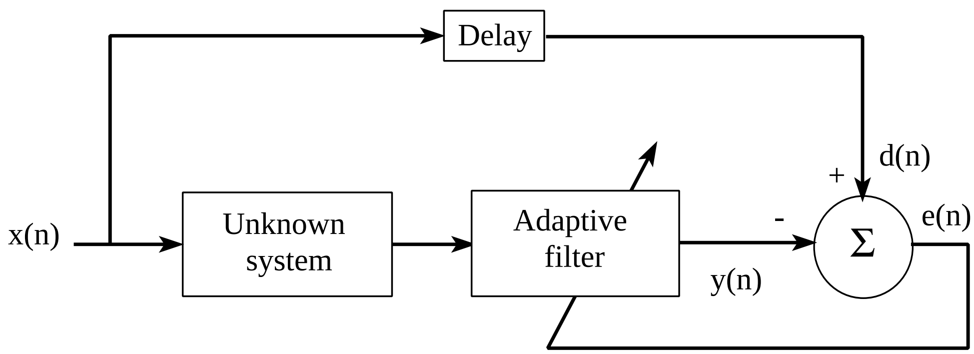

Figure 1.2: Inverse system modelling configuration of an adaptive filter.

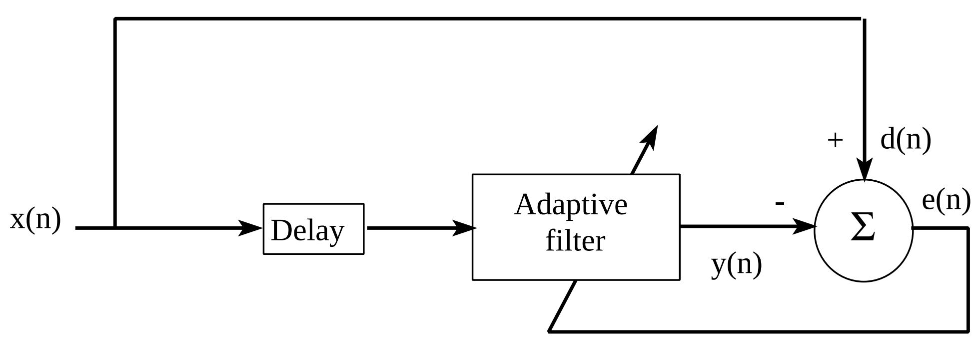

Figure 1.3: Configuration of an adaptive filter as a predictor.

1.2.4 Noise cancellation

In this final class of applications, the adaptive filter is used to cancel unknown interference contained in a primary signal, as Fig. 1.4 depicts it. The primary signal serves as the desired response of the adaptive filter. This type of application is used in adaptive noise cancellation [14]-[15], and adaptive beamforming or adaptive array processing [23].

The principle operation of the adaptive filter in all the four cases is mainly the same, and for the purposes of further development only the case of system identification will be considered. Also, interest in this configuration is related to the type of application we are dealing with in this thesis, namely echo cancellation.

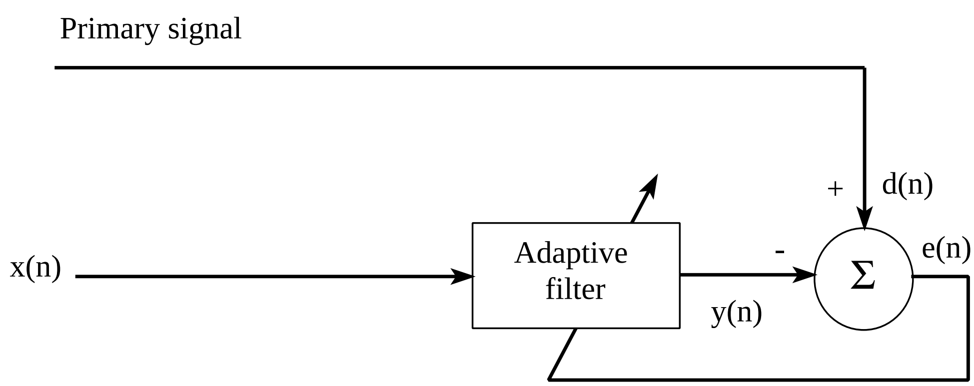

Figure 1.4: Configuration of an adaptive filter as a noise canceller.

1.3 Adaptive filters

Adaptive filters are an important part of signal processing. They are generally defined as filters whose characteristics can be modified to achieve desired objectives and accomplish this modification or adaptation automatically without user intervention. Due to the uncertainty about the input signal characteristics, the designer then uses an adaptive filter which can learn the signal characteristics when first turned on and can later track changes in these characteristics. Adaptive algorithms are responsible for the learning process.

A large number of algorithms for adaptive filters have been prosed. Indeed, adaptive filtering is an example of an optimisation problem and optimisation techniques form an important part of mathematics [24]-[26]. The additional constraint in adaptive filtering is that many of the applications require this optimisation to be performed in real time and so the complexity of the computations must be kept to a minimum.

In what is remaining of this section, different cost functions for adaptive filters are defined with some of the possible structures used in their implementation, and the expression for the optimum FIR filter in the mean square error (MSE) sense is given in terms of autocorrelation and crosscorrelation functions [27].

1.3.1 Cost functions

Before proceeding to discuss any adaptive algorithm, it is necessary to discuss the performance measure (cost function) which is used in adaptive filtering. The adaptive filter has the general form shown in Fig. 1.5, where the FIR filter of order N is considered here. The filter output $y(n)$ is given by

$$\begin{aligned} y(n) &= \sum_{i=0}^{N-1} c_i(n)x(n-i) \\ &= \mathbf{c}^T(n)\mathbf{x}(n), \end{aligned} \tag{1.1}$$

where $\mathbf{x}(n)$ and $\mathbf{c}(n)$ are, respectively, the vector of the last N samples from the time series $x(n)$ and the filter coefficients at sample n, defined as follows:

$$\mathbf{x}^T(n) = [x(n), x(n-1), \cdots, x(n-N+1)]. \tag{1.2}$$

and

$$\mathbf{c}^T(n) = [c_0(n), c_1(n), \cdots, c_{N-1}(n)], \tag{1.3}$$

where T denotes transpose.

In general, adaptive techniques have been classified under two main categories. In one category, the cost function to be optimised in a running sum of squared errors is given by:

$$J(n) = \sum_{j=0}^{n} e^2(j), \tag{1.4}$$

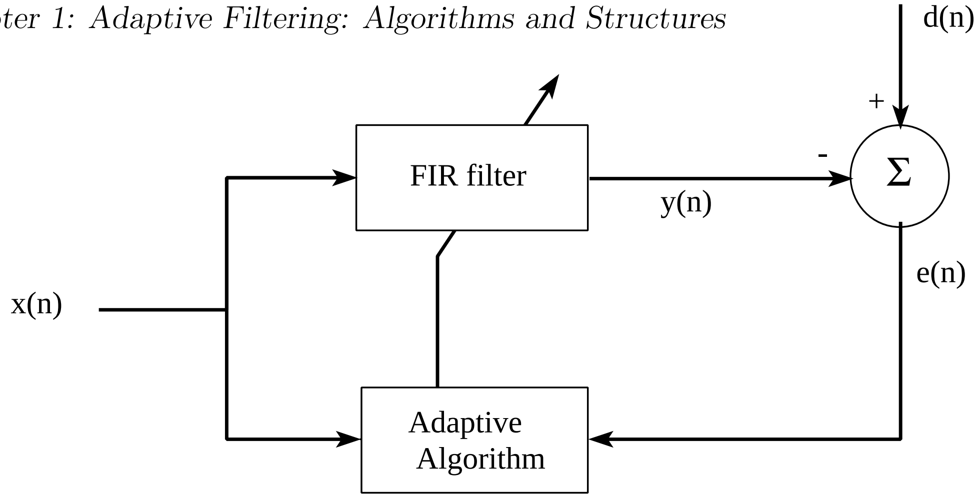

Figure 1.5: General form of an adaptive filter.

where the error $e(n)$ is defined to be the difference value between the desired response $d(n)$ and the output of the adaptive filter $y(n)$, that is,

$$e(n) = d(n) - y(n). \tag{1.5}$$

The approach, defined by (1.4), is based on the method of least squares [2], [28]-[29], which contains the whole class of recursive least squares (RLS) algorithms [7], [16], [30]-[32].

In the other category, the cost function to be optimised is a statistical measure of the squared error, known as the mean squared-error (MSE) [33]. This cost function is given by

$$J(n) = E[e^2(n)], \tag{1.6}$$

where $E[\,]$ denotes statistical expectation. This category contains the whole class of gradient algorithms, which includes the least mean-squared (LMS) algorithm [1], [7], [16].

The two procedures described above for deriving adaptive algorithms differ in some respect on how their respective cost function is chosen. The theory for the Wiener filter is based on statistical concepts, while it is based on the use of time averages for the method of least squares. Also, least squares techniques and stochastic techniques have a number of differences in the way that they perform [34]. Among these differences are on the one hand the much longer time taken for a stochastic gradient algorithm to converge close to the optimum solution and on the other hand the much higher computational complexity in least squares algorithms. Nevertheless, the less computational complexity in the stochastic gradient methods make them much more attractive than their least squares counterparts.

Recently, other minimisation criteria have emerged, in which adaptive structures are derived from minimisation of a class of functions of the form [8]:

$$J^k(n) = E[e^{2k}(n)];\ k \geq 1, \tag{1.7}$$

where k is an integer constant. It is seen from (1.7) that when $k = 1$, the usual MSE criterion is obtained, while the mean fourth-error (MFE) results when $k = 2$.

The cost functions, (1.4) and (1.6), are both convex with a unique minimum point. Accordingly, their use yields a unique solution for the coefficient vector of the FIR filter. Also, the minimisation function, (1.7), is a convex function, and therefore has no local minima. Hence, the use of a gradient based adaptation scheme for the convergence to the minimum can be applied.

Finally, before stating the possible linear structures used in implementing adaptive filters, it is worth mentioning the properties of the cost functions. All the functions presented in this section and others not mentioned in this work should be positive and monotonically increasing [35] for their corresponding algorithms to perform correctly.

1.3.2 Structures

A number of different linear structures for adaptive systems have been proposed, which may be subdivided into finite and infinite impulse structures. For the finite impulse response [36]-[38], the transfer function is realised by zeros only, as all the poles of the filter are located at the origin. In the case of the infinite impulse response (IIR) [16], [39] filter, however, both poles and zeros are used to realise the transfer function. Examples of the FIR filter are the linear transversal filter depicted in Fig. 1.6, and the lattice filter [40]-[42]. The structure of the IIR filter is shown in Fig. 1.7. However, difficulties associated with developing adaptive techniques for IIR filter are considerable, because the filter is not unconditionally stable, as it has both poles and zeros in its transfer function. The danger is that the adaptive algorithm will choose a set of coefficients which may place poles outside the unit circle in the z-plane

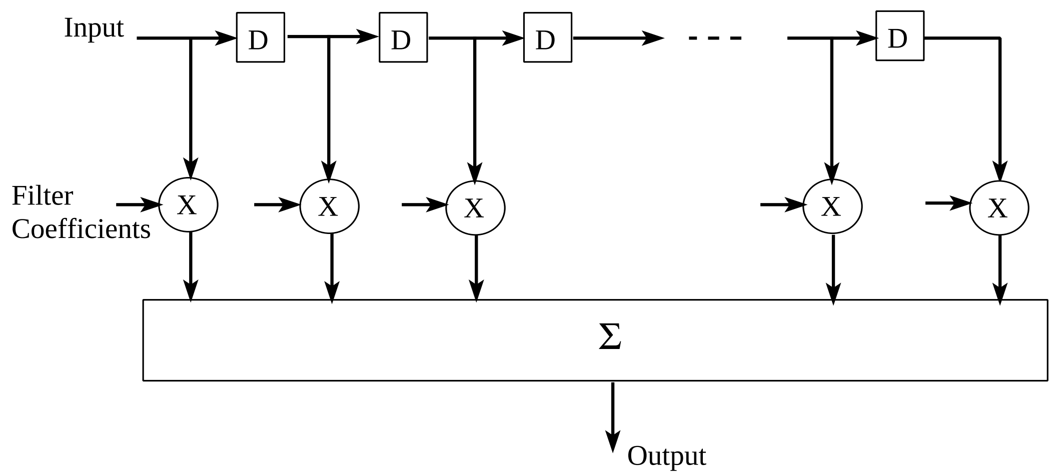

Figure 1.6: Structure of a linear transversal FIR filter.

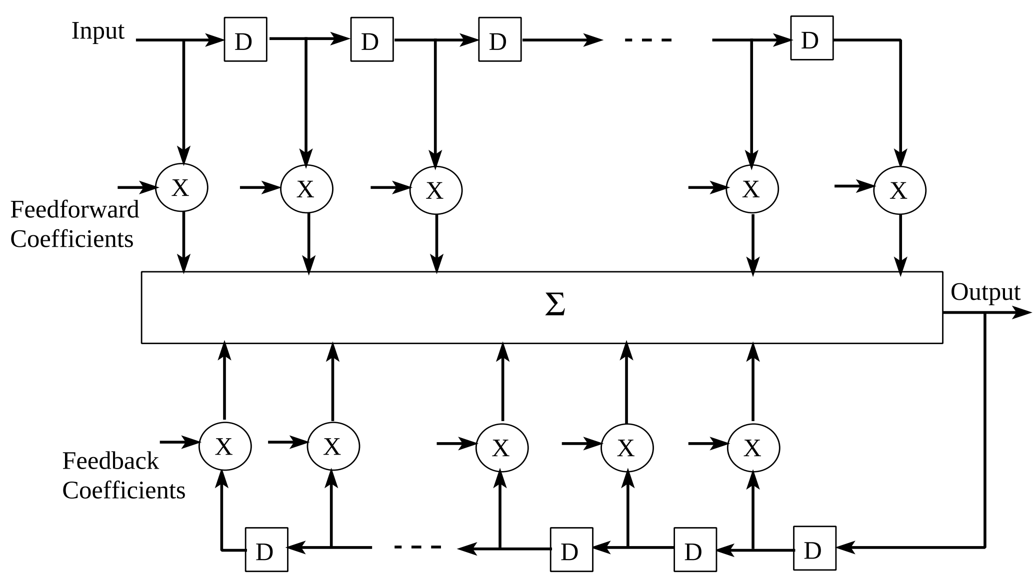

Figure 1.7: Structure of an IIR filter.

and so provoke an unstable response. These difficulties, hence, make the IIR structure less attractive than the well established FIR one.

The work of this course is, therefore, concerned with the linear transversal filter structure and the emphasis is on developing highly efficient algorithms for this well understood and often used structure.

1.3.3 The FIR adaptive filter

Assuming that the input sequence $\{x(n)\}$ and the desired sequence $\{d(n)\}$ are wide sense stationary, the mean-square-error function, equation (1.6), can be more conveniently expressed in terms of the input autocorrelation matrix, $\mathbf{R}$, and the crosscorrelation vector, $\mathbf{p}$, between the desired response and the input components, as follows:

$$J(n) = E[d^2(n)] - 2\mathbf{c}^T(n)\mathbf{p} + \mathbf{c}^T(n)\mathbf{R}\mathbf{c}(n), \tag{1.8}$$

where

$$\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}^T(n)], \tag{1.9}$$

and

$$\mathbf{p} = E[\mathbf{x}(n)d(n)]. \tag{1.10}$$

It is clear from expression (1.8) that the MSE is precisely a quadratic function of the components of the tap coefficients. Thus, the shape associated with this MSE is hyperboloid.

In general, for the linear transversal structure, the surface will be quadratic, when the MSE is used, with a single global minimum. The goal of an adaptive algorithm is to set the filter coefficients so as to obtain an operating point at this minimum, where the filter gives optimum performance.

The point at the bottom of the performance surface corresponds to the optimal tap coefficients, $\mathbf{c}_{opt}$, or minimum MSE. The gradient method is used to cause the tap coefficients vector to seek the minimum of the performance surface. It is defined as

$$\begin{aligned}\nabla E[e^2(n)] &= \frac{\partial J(n)}{\partial \mathbf{c}(n)} \\ &= [\frac{\partial J(n)}{\partial c_0(n)}, \frac{\partial J(n)}{\partial c_1(n)}, \cdots, \frac{\partial J(n)}{\partial c_{N-1}(n)}]^T \\ &= -2E[e(n)\mathbf{x}(n)] \\ &= 2\mathbf{R}\mathbf{c}(n) - 2\mathbf{p}. \end{aligned} \tag{1.11}$$

To obtain the minimum MSE, the tap-coefficients vector $\mathbf{c}(n)$ is set to its optimal value, $\mathbf{c}_{opt}$, where the gradient is zero, that is,

$$\nabla E[e^2(n)] = \mathbf{R}\mathbf{c}_{opt} - \mathbf{p} = \mathbf{0}. \tag{1.12}$$

Under this condition, the optimum value is given by:

$$\mathbf{c}_{opt} = \mathbf{R}^{-1}\mathbf{p}, \tag{1.13}$$

where this is obtained under the assumption that the autocorrelation matrix $\mathbf{R}$ of the input signal has no nulls in its power spectral density, that is positive definite and hence non singular. Properties of the the autocorrelation matrix $\mathbf{R}$ of the input signal can be found in [7]. The minimum MSE , J_{min}, is hence obtained by substitution of (1.13) in (1.8), that is,

$$J_{min} = E[d^2(n)] - \mathbf{c}_{opt}^T\mathbf{p}. \tag{1.14}$$

The solution for $\mathbf{c}_{opt}$ involves inverting the input autocorrelation matrix $\mathbf{R}$, hence, requiring precise knowledge of the second order statistics of the data, i.e., the autocorrelation matrix and the crosscorrelation vector. Unfortunately, it is the data sequences rather than their second order statistics that are available in practice. Alternatively, an iterative procedure may be used to determine $\mathbf{c}_{opt}$. This is the function of an adaptive FIR filter algorithm which has to find the optimum filter from available data rather than from the second statistics of the data [43]. Thus, an adaptive FIR filter can be defined as an algorithm which operates on the sequences $\{x(n)\}$ and $\{d(n)\}$ to form a time-varying impulse response $\mathbf{c}(n)$ which converges in the mean to $\mathbf{c}_{opt}$ as the number of iterations becomes very large, that is:

$$\lim_{n\to\infty} E[\mathbf{c}(n)] = \mathbf{c}_{opt}. \tag{1.15}$$

1.4 Summary

This chapter concentrated on basic ideas for which both adaptive filters and adaptive algorithms are made up. The issue of adaptive filtering is still and will remain a very active field of research for some considerable time. This is mainly due to the advances in the computing facilities that were not previously available and to the need for such algorithms.

The wide spread use of the least-demanding computing algorithm, i.e., the LMS algorithm, is with no doubt due to its both simplicity and relative performance. The RLS algorithm, for example, gives very fast convergence to the algorithm at the expense of very heavy computational loads, irrespective of the input signal statistics. However, things change when the input signal is white noise, the convergence properties of the LMS algorithm, under certain circumstances, becomes comparable to or the same as those of the RLS algorithm.

Both of these algorithms, and in general all algorithms, operate under different minimisation functions, which are the main reason in their different performances.

Bibliography

[1] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter," *Proc. IEEE, vol.* 64, No. 8, pp. 1151-1162, Aug. 1976.

[2] H. W. Sorenson, "Least-Squares Estimation from Gauss to Kalman," *IEEE Spectrum,* vol. 7, pp. 63-68, July 1970.

[3] D. G. Messerschmitt, "Echo Cancellation in Speech and Data Transmission," *IEEE J. Selected Areas in Comm., vol.* SAC-2, No. 2, pp. 283-297,Mar. 1982.

[4] R. W. Lucky, "Automatic Equalization for Digital Communication," *Bell Syst. Tech. J., vol.*, 44, No. 3, pp 547-588, Apr. 1965.

[5] S. B. Weinstein, "Echo Cancellation on the Telephone Network," *IEEE Commun. Magazine, vol.* 16, No. 11, pp. 9-15, Jan. 1977.

[6] A. Zerguine, C. F. N. Cowan, and M. Bettayeb, "Adaptive Echo Cancellation Using Statistical Shaping," *Twenty-ninth Annual, Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, California* Otc. 29-Nov. 1, 1995.

[7] S. Haykin, *Adaptive Filter Theory,* Prentice-Hall, Englewood Cliffs, NJ, 1991.

[8] E. Walach and B. Widrow, "The Least Mean Fourth (LMF) Adaptive Algorithm and its Family," *IEEE Trans. Inf. Theory, vol.* IT-30, pp. 275-283, Feb. 1984.

[9] D. I. Pazaitis and A. G. Constantinides, "LMS+F algorithm," *Electr. lett., vol.* 31, No. 17, pp. 1423-1424, Aug. 1995.

[10] J. A. Chambers, O. Tanrikulu, and A. G. Constantinides, "Least mean mixed-norm adaptive filtering," *Electr. lett., vol.* 30, No. 19, pp. 1574-1575, Sep. 1994.

[11] M. M. Sondhi, "An Adaptive Echo Canceller," *Bell Syst. Tech. J., vol.*, 46, No. 3, pp 497-511, Mar. 1967.

[12] K. Murano, "Adaptive Signal Processing Applied in Telecommunications," *IFAC Adaptive Systems in Control and Signal Processing* pp. 431-441, 1992.

[13] S. U. H. Qureshi, "Adaptive Equalization," *Proc. IEEE, vol.* 73, No. 9, pp. 1349-1387, Sept. 1985.

[14] B. Widrow, J. R. Glover, Jr., J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong, Jr., and R. C. Goodlin, "Adaptive Noise Cancelling: Principles and Applications," *Proc. IEEE, vol.* 63, No. 12, pp. 1692-1716, Dec. 1975.

[15] C. F. N. Cowan, J. Mavor, and J. W. Arthur, "Noise Cancellation and Inverse Filtering using a Compact High-Performance C.C.D Adaptive Filter," *Electr. lett., vol.* 15, No. 1, pp. 35-37, Jan. 1979.

[16] C. F. N. Cowan and P. M. Grant, *Adaptive Filters*, Prentice-Hall, Englewood Cliffs, NJ, 1985.

[17] M. M. Sondhi and D. A. Berkley, "Silencing Echoes on the Telephone Network," *Proc. IEEE, vol.* 68, No. 8, pp. 948-963, Aug. 1980.

[18] S. U. H. Qureshi, "Adaptive Equalization," *IEEE Commun. Magazine, vol.* 21, No. 2, pp. 9-16, Mar. 1982.

[19] A. H. Gray and J. D. Markel, *Linear Prediction of Speech*, Springer Verlag, Berlin, 1976.

[20] J. Makhoul, "Linear Prediction a Tutorial Review," *Proc. IEEE, vol.* 63, No. 4, pp. 561-580, April 1975.

[21] J. G. Proakis, *Digital Communications* , McGraw-Hill, New York, 1983.

[22] T. Kailath, "A View of three Decades of Linear Filtering Theory ," *IEEE Trans. Inf. Theory, vol.* IT-20, pp. 146-181, March 1974.

[23] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive Antenna Systems Array," *Proc. IEEE, vol.* 55, pp. 2143-2159, Dec. 1967.

[24] G. Leitmann, *Mathematics in Science and Engineering vol 5: Optimization Techniques*, Academic Press, New York, 1962.

[25] G. Leitmann, *Mathematics in Science and Engineering vol 31: Topics in Optimization*, Academic Press, New York, 1967.

[26] D. Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, New York, 1969.

[27] N. Levinson, "The Wiener RMS (Root Mean Square) Error Criterion in Filter Design and Prediction," *J. Math. Phys.,* vol. 25 pp. 261-278, Jan. 1947.

[28] J. M. Mendel, *Discrete Techniques Of Parameter Estimation*, Marcel Dekker, New York, 1973.

[29] T. Kailath, *Lectures on Linear Least-Squares Estimation*, Springer-Verlag, New York, 1981.

[30] D. D. Falconer and L. Ljung, "Application of fast Kalman Estimation to Adaptive Equalization," *IEEE Trans. Commun., vol.* COM-26, pp. 1439-1446, June 1978.

[31] G. Carayannis, D. G. Manolakis, and N. Kalouptsidis, "A Fast Sequential Algorithm for Least-Squares Filtering and Prediction," *IEEE Trans. Acoust., Speech, Signal Processing, vol.* ASSP-31, pp. 1394-1402, Dec. 1983.

[32] J. M. Cioffi and T. Kailath, "Fast, Recursive-Least-Squares Transversal Filters for Adaptive Filtering," *IEEE Trans. Acoust., Speech, Signal Processing, vol.* ASSP-32, pp. 304-337, 1984.

[33] N. Wiener, *The Extrapolation, Interpolation, and Smoothing of stationary Time Series*, John Wiley & Sons, New York, 1962.

[34] C. F. N. Cowan, "Performance Comparison of finite Linear Adaptive Filters," *IEE Proceedings Part F vol.* 134, No. 3, pp. 211-216, June 1987.

[35] B. Mulgrew and C. F. N. Cowan, *Adaptive Filters and Equalizers*, Kluwer Academic Publishers, Norwell, MA, 1988.

[36] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1975.

[37] R. A. Roberts and C. T. Mullis, *Digital Signal Processing*, Addison Wesley, Reading, MA, 1987.

[38] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, Second edition, Macmillan, New York, 1992.

[39] M. G. Larimore, J. R. Treichler, and C. R. Johnson, "SHARF; An Algorithm for Adapting IIR Digital Filters," *IEEE Trans. Acoust., Speech, Signal Processing, vol.* ASSP-28, pp. 1622-1624, Nov. 1976.

[40] E. H. Satorius and S. T. Alexander, "Channel Equalization using Adaptive Lattice Algorithms," *IEEE Trans. Commun., vol.* COM-27, No. 9, pp. 899-905, June 1979.

[41] B. Friedlander, "Lattice Filters for Adaptive Processing," *Proc. IEEE, vol.* 70, pp. 829-867, Aug. 1982.

[42] M. L. Honig and D. G. Messerschmitt, *Adaptive Filter: Structures, Algorithms, and Applications*, Kluwer Academic Publishers, Hingham, MA, 1984.

[43] B. Widrow, "Adaptive Filters," in *Aspects of Network and System Theory* ed. R. E. Kalman and N. Declaris, Holt, Rinehart & Winston, New York, 1971.

[44] G. Ungerboeck, "Theory of the Speed of Convergence in Adaptive Equalizers for Digital Communications," *IBM J. Res. Dev., vol. 16,* pp. 546-552, 1972.

[45] J. E. Mazo, "On the Independence Theory of Equalizer Convergence," *Bell Syst. Tech. Journ.,* No. 58, pp. 963-993, May-June 1979.

[46] I. J. Nagumo and A. Noda, "A Learning Method for System Identification," *IEEE Trans. Automatic Control, vol.* AC-28, pp. 282-287, 1967.

[47] R. D. Gitlin, H. C. Meadors, and S. B. Weinstein, "The Tap-Leakage Algorithm: An Algorithm for the Stable Operation of a Digitally Implemented, Fractionally Spaced Adaptive Equalizer," *Bell Syst. Tech. J., vol.* 61, pp. 1817-1839, Oct. 1982.

[48] R. W. Harris, D. M. Chabries, and F. A. Bishop "A Variable Step (VS) Adaptive Filter Algorithm," *IEEE Trans. Acoust., Speech, Signal Processing, vol.* ASSP-34, pp. 309-316, Feb. 1986.

[49] G. A. Clark, S. K. Mitra, and S. R. Parker, "Block Implementations of Adaptive Digital Filters," *IEEE Trans. Circuits and Systems, vol.* CAS-28, pp. 584-592, Jun. 1981.

[50] M. L. Honig, "Echo Cancellation of Voiceband Data Signals Using Recursive Least Squares and Stochastic Gradient Algorithms," *IEEE Trans. Commun., vol.* COM-33, No. 1, pp. 65-73, Jan. 1985.

[51] C. W. K. Gritton and D. W. Lin, "Echo Cancellation Algorithms," *IEEE ASSP Magazine*, pp. 30-37, April 1984.

[52] L. Guidoux and B. Peuch, "Binary Passband Echo Canceller in a 4800 Bit/s Two-Wire Duplex Modem," *IEEE J. Selected Areas in Comm., vol.* SAC-2, No. 5, pp. 711-721, Sept. 1984.

[53] A. Lewis, "Adaptive Filtering - Applications in Telephony," *BT Technol. J., vol.* 10, No. 1, pp. 49-63, Dec. 1992.

[54] P. T. Brady and G. K. Helder, "Echo Suppressor Design in Telephone Communications," *Bell Syst. Tech. J., vol.* 42, pp. 2893-2917, 1963.

[55] CCITT Recommendation G. 164, "Echo Suppressors."

[56] D. G. Messerschmitt, "An Electronic Hybrid with Adaptive Balancing for Telephony," *IEEE Trans. Commun., vol.* COM-28, No. 8, pp. 1399-1407, Aug. 1980.

[57] B. E. Dotter, Jr., A. De La Plaza, D. A. Hodges, and D. G. Messerschmitt, "Implementation of an Adaptive Balancing Hybrid," *IEEE Trans. Commun., vol.* COM-28, No. 8, pp. 1408-1416, Aug. 1980.

[58] K. Murano, S. Unagami, and F. Amano, "Echo Cancellation and Applications," *IEEE Commun. Magazine, vol.* 28, No. 1, pp. 49-55, Jan. 1990.

[59] O. Agazzi, D. A Hodges, and D. G. Messerschmitt, "Nonlinear Echo Cancellation of Data Signals," *IEEE Trans. Commun., vol.* COM-30, No. 11, pp. 2421-2433, Nov. 1982.

[60] M. J. Smith, C. F. N. Cowan, and P. F. Adams, "Nonlinear Echo Cancellers Based on Transposed Distributed Arithmetic," *IEEE Trans. Cir. and Syst., vol.* CAS-35, No. 1, pp. 6-18, Jan. 1988.

[61] K. Yamazaki, S. Aly, and D. Falconer, "Convergence Behavior of a Jointly Adaptive Transversal and Memory-Based Echo Canceller," *IEE Proceedings-F, vol.* 138, No. 4, pp. 361-370, Aug. 1991.

[62] 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 Trans. Acoust., Speech, Signal Processing, vol.* ASSP-27, No. 6, pp. 768-781, Dec. 1979.

[63] S. B. Weinstein, "A Passband Data-Driven Echo Canceller for Full-Duplex Transmission on Two-Wire Circuits," *IEEE Trans. Commun., vol.* COM-25, No. 7, pp. 654-666, July 1977.

[64] N. Holte and S. Stueflotten,"A New Digital Echo Canceler for Two-Wire Subscriber Lines," *IEEE Trans. Commun., vol.* COM-29, No. 11, pp. 1573-1581, Nov. 1981.

[65] V. G. Koll and S. B. Weinstein, "Simultaneous Two-Way Data Transmission over A Two-Wire Circuit," *IEEE Trans. Commun., vol.* COM-21, No. 2, pp. 143-147, Feb. 1973.

[66] K. H. Mueller, "A New Digital Echo Canceler for Two-Wire Full-Duplex Data Transmission," *IEEE Trans. Commun., vol.* COM-24, No. 9, pp. 956-962, Sep. 1976.

[67] J. M. Cioffi and J. J. Werner, "Effects of Biases on Digitally Implemented Data-Driven Echo Cancelers," *AT&T Techn. J., vol.* 64, No. 1, pp. 115-138, Jan. 1985.

[68] P. F. Adams, "Adaptive Filters in Telecommunications," ed. C. F. N. Cowan and P. M. Grant, Prentice-Hall, Englewood Cliffs, NJ, 1985.

[69] O. Agazzi, D. A Hodges, and D. G. Messerschmitt, "Large-Scale Integration of Hybrid-Method Digital Subscriber Loops," *IEEE Trans. Commun., vol.* COM-30, No. 9, pp. 2095-2108, Sep. 1982.

[70] K. Shenoi, *Digital Signal Processing in Telecommunications*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[71] H. Fan and W. K. Jenkins, "An Investigation of an Adaptive IIR Echo Canceler: Advantages and Problems," *IEEE Trans. Acoust., Speech, Signal Processing, vol.* ASSP-36, No. 12, pp. 1819-1834, Dec. 1988.

[72] E. R. Ferrara, Jr., "Frequency Domain Adaptive Filtering," ed. C. F. N. Cowan and P. M. Grant, Prentice-Hall, Englewood Cliffs, NJ, 1985.

[73] F. Ling, "Echo Cancellation," ed. N. Kalouptsidis and S. Theodoridis, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[74] D. L. Duttweiler and Y. S. Chen, "A Single Chip VLSI Echo Canceller," *Bell Syst. Tech. J., vol.,* 59, No. 2, pp. 149-161, Feb. 1980.

[75] J. J. Shynk, "Adaptive IIR Filtering," *IEEE ASSP Magazine, vol.* 28, No. 1, pp. 4-21, Apr. 1989.

[76] C. R. Johnson, Jr., "Adaptive IIR Filtering: Current Results and Open Issues," *IEEE Trans. Inform. Theory, vol.* IT-30, No. 2, pp. 237-250, Mar. 1984.

[77] G. Long, D. Shwed, and D. D. Falconer, "Study of a Pole-Zero Adaptive Echo Canceller ," *IEEE Trans. Circuits Systems, vol.* CAS-34, No. 7, pp. 765-769, Jul. 1987.

[78] T. A. C. M. Classen and W. F. G. Mecklenbrauker, "Comparison of the Convergence of two Algorithms for the Adaptive FIR Digital Filters," *IEEE Trans. Acoust., Speech, Signal Processing, vol.* ASSP-29, pp. 670-678, June 1981.

[79] H. Sari, "Performance Evaluation of three Adaptive Equalization Algorithms," *in Proc. ICASSP'82* , pp. 1385-1388.

[80] C. J. Macleod, E. Ciapala, and Z. J. Jelonek, "Quantisation in non-recursive Equalisers for Data Transmission," *PROC. IEE, Vol 122, No. 10,* pp. 1105-1110, Oct. 1975.

[81] P. Xue and B. Liu, "Adaptive Equalizer Using Finite-Bit Power-of-Two Quantizer," *IEEE Trans. Acoust., Speech, Signal Processing, vol.* ASSP-34, pp. 1603-1611, December 1986.

[82] R. E. Kalman , "A New Approach to Linear Filtering and Prediction Problems," *Trans. ASME J. Basic Eng.,* vol. 82D, pp. 35-45, March 1960.

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

[84] D. N. Godard, "Channel Equalization using a Kalman Filter for Fast Data Transmission," *IBM J. Res. Dev., vol.* 18, pp. 267-273, 1976.

[85] M. Morf, A. Vieira, and D. T. Lee, "Ladder Forms for Identification and Speech Processing," *Proc. IEEE CDC'77,* pp. 1074-1078.

[86] M. Morf and D. T. Lee, "Recursive Least Squares Ladder Forms for Fast Parameter Tracking," *Proc. IEEE CDC'78,* pp. 1362-1367.

[87] F. Ling and J. G. Proakis, "Numerical Accuracy and Stability: Two Problems of Adaptive Estimation Algorithms Caused by Round-Off Error," *Proc. IEEE ICASSP'84,* pp. 30.3.1-30.3.4.

[88] F. Ling, D. Manolakis, and J. G. Proakis, "New Forms of LS Lattice Algorithms and an Analysis of their Round-Off Error Characteristics," *Proc. IEEE ICASSP'85,* pp. 1739-1742.

[89] I. K. Proudler, J. G. McWhirter, and T. J. Shepherd, *QRD-Based Lattice Filter Algorithms*, SPIE, San Diego, Calif., 1989.

[90] P. A. Regalia and M. G. Bellanger, "On the Duality between Fast QR Methods and Lattice Methods in Least Squares Adaptive Filtering," *IEEE Trans. Acoust., Speech, Signal Processing, vol.* ASSP-39, pp. 879-891, 1991.

[91] G. C. Goodwin and R. L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press, New York, 1977.

[92] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1985.

[93] O. Macchi and E. Eweda, "Second-Order Convergence Analysis of Stochastic Adaptive Linear Filtering," *IEEE Trans. Automat.Contr., vol.* AC-28, pp. 76-85, Jan. 1983.

[94] R. R. Bitmead, "Convergence in Distribution of LMS-Type Adaptive Parameter Estimates," *IEEE Trans. Automat.Contr., vol.* AC-28, pp. 54-60, Jan. 1983.

[95] W. A. Gardner, "Learning Characteristics of Stochastic-Gradient-Descent Algorithms: A general Study, Analysis and Critique," *Signal Processing, vol.* 6, pp. 113-133, 1984.

[96] C. L. Nikias and A. P. Petropulu, *High-Order Spectra Analysis: A Nonlinear Signal Processing Framework*, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[97] A. Papoulis, *Probability, Random Variables, and Stochastic Processes, Second Ed.*, McGraw-Hill Publishing Co., New York, 1984.

[98] A. Zerguine, C. F. N. Cowan, and M. Bettayeb, "Adaptive Echo Cancellation Using the L_2 Norm of both Near-End and Far-End Errors," *Proceedings of the First Communications Conference*, pp. 212-216, Muscat, Oman, Mar. 1996.

[99] D. C. Shin and C. L. Nikias, "Adaptive Interference Canler for Narrowband and Wideband Interferences Using Higher Order Statistics," *IEEE Trans. Signal Processing, vol.* SP-42, pp. 2715-2728, Oct. 1994.