Credit: FPG International/ © Orion Press

On the Interaction of Adaptive Filtering, Identification, and Control*

C. RICHARD JOHNSON, JR.

"I was motivated to undertake this study by a belief that there are enough similarities and differences in the philosophies, goals, and analytical techniques found in the two fields to indicate that a concentrated effort to understand these better might lead to a number of ideas for interaction and collaboration among researchers."

—from foreword of Willsky (1979)

My objective in this article is to add evidence for the belief espoused above regarding the two fields of (i) signal processing and (ii) control and estimation. I focus here on the interaction of specialties from (i), i.e., adaptive filtering, and (ii), i.e., adaptive identification and control. For evidence, I use three basic communication systems applications of adaptive filtering to evoke system identification (and adaptive control) style problems amenable to analytical tools popular over (at least) the last decade. The applications raise a number of un(der)resolved issues that challenge existing theory. The connections drawn to adaptive identification and control problems (and relevant analytical tools) suggest approaches to these challenges.

* Based on "Interplay Between Recursive System Identification and Adaptive Filtering in Communication Systems," presented at the 10th IFAC Symposium on System Identification, Copenhagen, Denmark, July 4, 1994.

1053-5888/95/$4.00©1995IEEE

The applications and connections described here are limited to problems I have studied (with the help of others) in some detail. With this admission, I trust the reader will not interpret this article as a thorough survey of advances and issues in the application areas discussed. Another consequence of this idiosyncratic approach to topic selection is that the related issues of mid-1980s adaptive filtering, identification, and control isolated here are not necessarily representative of mid-1990s concerns in the applications discussed, though some are. Nor should the interconnections described here be expected to exhaust the full extent of the interactions of adaptive filtering and adaptive identification and control. Instead, they are intended to bolster the perception of the benefits of cross-fertilization between adaptive filtering and adaptive identification and control.

Overview

Fundamental connections among adaptive filtering, identification, and control have been widely acknowledged since their inception in the late 1960s. However, early on these specializations spawned separate literatures and experts. Explicit attempts at rapprochement emerged in the latter half of the 1970s. This article presents another installment (in a tutorial/survey style) of this continuing saga. For three successful applications of adaptive filtering in communications systems (network echo cancellation, differential pulse code modualtion, and linear channel equalization) that emerged concurrently in the 1960s, mid-1980s application concerns and contemporaneous adaptive parameter estimation theory are linked. Research relevant to some specific issues that emerge from these interconnections is cited.

"Self-adjusting linear filters...are at the heart of high-speed telephone line modems, and have new importance in digital microwave radio, twisted pair (ISDN), undersea, and optical communication systems, and have strongly influenced the broad discipline of linear adaptive systems."

—Gitlin, Hayes and Weinstein (1992)

The distortion, interference, and redundancy removal capabilties of linear filters with adapted parameters are crucial to current communication systems. Three topics dominate the applications chapter in Honig and Messerschmitt (1984) and the more recent overview paper by Murano (1992):

- echo cancellation
- speech coding
- channel equalization

These problems and their adaptive filtering solutions were introduced in the latter half of the 1960s. Within two decades, each application was accorded a major survey paper in the *Proceedings of the IEEE* (Sondhi and Berkeley, 1980; Gibson, 1980; Qureshi, 1985). These three applications dominate this article as well.

The next section presents the equation error formulation of updating the parameters of a tapped delay line, finite impulse response (FIR) model in an identification setting, and its recursive solution via the LMS algorithm. Then, echo cancellation, speech coding, and channel equalization are discussed. The initial task in each of these three application sections is to set the stage with a problem formulation that fits the equation error formulation. Indeed, in each application, this setting and LMS solution are at the heart of decades of actual use and at the core of adaptive filter pedagogy.

Having established this fundamental connection, the objective is to describe some instances of intersection between work on adaptive identification and control and the solutions of these adaptive filtering problems. The instances cited here primarily connect to advances in the theory of adaptive identification and control made during the period of publication of the definitive *Proc. IEEE* surveys cited above, i.e., the first half of the 1980s. As noted earlier, they also draw heavily on the author's own research meanderings since the mid 1980s. The primary, mid-1980s concepts of significance, borrowed here from adaptive identification and control are:

(a) Persistent excitation is important for performance robustness

(b) Lack of persistent excitation is an invitation for parameter estimate drift and bursting in adaptive feedback systems

(c) A general algorithm form for prediction error schemes uses a parameter estimate update that correlates a filtered version of the prediction error and a filtered version of the regressor.

These precepts can all be established through averaging theory (Anderson et al., 1986; Sastry and Bodson, 1989; Solo and Kong, 1995) or similar tools validated by a small stepsize. The slowed convergence of small stepsize algorithms has proven tolerable in a variety of applications, in which a high number (e.g., well over 1000 per estimated parameter) of data samples is processed before the desired filter parameterization varies appreciably. Thus, the tracking requirements are not too strenuous relative to the achievable convergence rate of a small stepsize algorithm such as LMS. This is the situation in a variety of digital filtering tasks in modern communication systems.

Each of the three application sections follow the same pattern. After the stage is set for LMS, the problem is elaborated in some manner that results in a formulation suited to advances made in adaptive identification and control during the early 1980s. For the echo cancellation problem, this is the inclusion in analysis, as in usage, of the feedback path between the output and input of the adaptive filter, which results in parameter drift and bursting behavior. Parameter drift and bursting analysis (and remedy) were active topics in adaptive identification and control in the 1980s, as indicated by points (a) and (b) above.

The elaboration to the speech coding problem is the inclusion of a pole-zero predictor, as in CCITT standard G.721 (established in 1984) for 32 kbps adaptive differential pulse code modulation in 64 kbps trunk lines in the phone system.

This results in an output error formulation—a problem with a long history under numerous guises in the recursive identification literature. However, G.721 uses algorithms quite different from those promoted in the adaptive identification literature for this output error style formulation, as summarized by point (c) above. For example, the predictor parameter adaptation laws in G.721 include sign operators on the update term elements that are almost never included in algorithms analyzed in a systems identification setting, but these operators can be useful in a high data rate communication system application. The presence of these sign operators changes the input signal character needed to provide a persistently exciting regressor, which the recursive identification literature exploited in the 1980s for its enhancement of performance robustness by adaptive identifiers and controllers—point (a) again.

In the channel equalization application, a training signal is needed to create the equation error formulation. In a number of applications of increasing importance, the capacity loss due to a recurrent training signal is too great. Unfortunately, the absence of the training signal, which creates what is termed a blind equalization problem, removes the basis of the prediction error formulation that dominates current recursive identification pedagogy established in the 1980s (Ljung and Söderström, 1983). The multiple local minima in the performance surfaces associated with popular blind equalization algorithms relate to similar multimodal function search problems in adaptive identification and control. Certainly, point (c) above will require alteration. Points (a) and (b) are likely to require re-interpretation. Thus, blind equalization is one example of an adaptive filtering task in a communication system that is quite distinct from problems that dominate the adaptive identification literature. While some analytical concepts and tools of adaptive identification and control remain useful, un(der)resolved adaptive filtering problems in communication systems with substantial practical impact can prove so different that new insights and analytical approaches are required for algorithm development and understanding.

Fundamental Recursive Identifier

A classical system identification configuration is depicted in Fig. 1. The feedback of e with an arrow through the identifier box in Fig. 1 indicates that the prediction error is the performance information used to adjust the identifier parameters. B ($\hat{B}$) indicates the parameterization of the system (identifier).

For a finite impulse response (FIR) model of the system,

$$y(k) = \sum_{i=0}^{N} b_i u(k-i), \tag{1}$$

the identifier uses the same model structure with the same order but (prior to perfect matching convergence) different parameters

$$\hat{y}(k) = \sum_{i=0}^{N} \hat{b}_i u(k-i). \tag{2}$$

Thus, the prediction error is

$$e(k) = y(k) + n(k) - \hat{y}(k) = \tilde{B}^T U(k) + n(k), \tag{3}$$

where

$$\tilde{B}^T = (B - \hat{B})^T = \left[b_0 - \hat{b}_0, b_1 - \hat{b}_1, \ldots, b_N - \hat{b}_N\right] \tag{4}$$

and

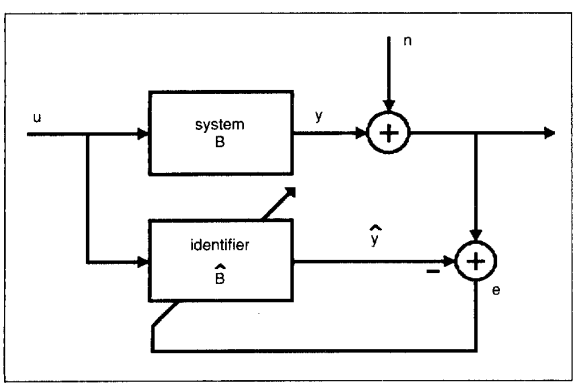
$$U(k) = \left[u(k), u(k-1), \ldots, u(k-N)\right]^T. \tag{5}$$

The prediction error, e is an inner product of the regressor, U and the parameter $\hat{B}$, plus some uncorrelated zero-mean noise, n. This special relationship between e and $\tilde{B}$, U, and n is the foundation of the unbiased equation error formulation (Mendel, 1973). The approximate gradient descent minimization of the average squared prediction error via

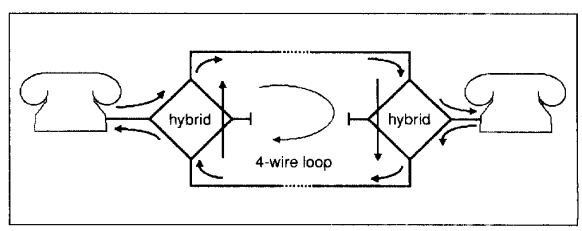
$$\hat{B}(k+1) = \hat{B}(k) + \mu U(k) e(k), \tag{6}$$

where μ is a small positive stepsize, is a viable recursive solution to this system identification problem. In the adaptive filtering literature, Eq. 6 is known as the LMS algorithm.

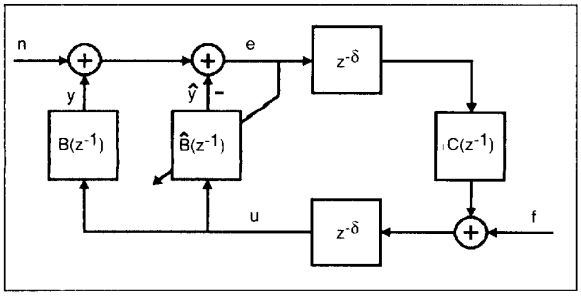
LMS was born (Widrow and Hoff, 1960) during the digitization of the phone system. This enhancement of the voice-band communication system focused attention on the three applications cited earlier: (i) network echo cancellation, (ii) differential pulse code modulation, and (iii) linear equalization, all of which required adaptive solutions for practical success (Murano, 1992). We will now abstract each of these



1. System identification.



2. Telephone system loop.



3. Adaptive echo canceller.

applications so a problem suited to the LMS algorithm emerges.

Echo Cancellation

"[E]cho mechanisms are mitigated if the hybrid has significant loss between its two four-wire ports. Achieving this large loss unfortunately depends on the knowledge of the two-wire line impedance, which varies significantly over the population of subscriber loops. The four-wire to four-wire frequency dependent loss of the hybrid cannot be depended on to be greater than 6 dB."

—Honig and Messerchmitt (1984)

A typical telephone system loop is depicted in Fig. 2. Messages sent from one end traverse a twisted pair to the local switching station. The twisted pair is one port in a 4-port bridge (hybrid) network. The hybrid is designed so the signals transmitted from the left to the right station in Fig. 2 come out the top of the local hybrid. More importantly, the transmitted signal entering the top of the right hybrid should all go out to the right handset, with none of it emerging from the bottom port of the hybrid. If some signal passes through the hybrid, the potential exists for loop amplification that leads to a phenomenon known as "singing." Thus, the hybrid must be designed to sufficiently inhibit the passthrough gain.

Designing these hybrids requires knowledge of the impedance characteristics at each port. In practice, these values are not known precisely. The variability in telephone channels typically results in a compromise hybrid design achieving a drop of 6 (or more) dB across the hybrid from one half of the 4-wire loop to the other. Note that the signal passing through the hybrid returns to the original transmitter as an echo.

Equation Error Formulation

An adaptive solution to the echo cancellation problem is shown in Fig. 3. Only one hybrid is shown as being adaptive. This adaptive filter, $\hat{B}$, at the "near" end should inhibit echo perceived at the "far" end. The input of the far-end speaker is labeled f. The near-end input is labeled n. The delays, $z^{-\delta}$, represent the transmission delay across each half of the 4 wire loop. The passthrough charateristics of the non-adaptive hybrid at the far end are described by $C(z^{-1})$. The passthrough dynamics of the near-end hybrid are modelled by $B(z^{-1})$. The adaptive filter sits inside the 4-wire loop across the hybrid and attempts to additively cancel y from the output of the hybrid, $y + n$.

The left side of Fig. 3 can be cast as in Eqs. 1-5, and Fig. 1 with exactly the same notation. However, unlike the assumption made regarding n in Eqs. 1-5, due to the feedback connection in Fig. 3, n no longer needs be uncorrelated with the current and N past values of u in U from Eq. 5. This correlation will be slight if δ is sufficiently large compared to the time span of significant temporal correlation in $u(\cdot)$, and f is (substantially) uncorrelated with n.

Becker and Rudin (1966), Sondhi and Presti (1966), and Sondhi (1967) introduced the concept of adaptive echo cancellation using a parameter-adaptive, tapped delay line filter. Within a decade, the cancelling structure of Fig. 3 was described (Widrow, et al., 1975) as one member of a large class of signal processing applications amenable to the LMS filter. Two years later, a tutorial survey of echo cancellation in the telephone network (Weinstein, 1977) acknowledges the LMS algorithm as the standard solution. Several years later, another tutorial survey (Gritton and Lin, 1984) repeats this "fact" that the most common implementation of echo cancellers uses LMS. Single chip adaptive echo cancellers (using an LMS-like algorithm on 128 taps) are commercially available by the end of the 1970s (Duttweiler, 1978).

See Weinstein (1977), Sondhi and Berkley, (1980), Messerschmitt (1984), and Gritton and Lin (1984) for descriptions of various problem driven refinements on this LMS adaptive echo canceller theme. The situation termed double talking, i.e., when f and n are substantial simultaneously, is discussed in all four surveys as a serious practical issue. In terms of the identification problem, in Fig. 1, we would prefer to perform identification only when u is large (and spectrally rich) compared to a therefore negligible n. In Figure 3, if both speakers are talking, then n can overpower y in $y + n$, and the adaptive equalizer can select inappropriate parameters. The fundamental strategy is to inhibit adaptation by setting the stepsize to zero whenever double talking is detected. Since f and n cannot be measured by the hybrid, only u and $y + n$ can, the basic idea is to declare doubletalk when there is less than a drop of 6 dB between u and $y + n$.

"The design of a good double talk detector is difficult and much more of an art than the design of the adaptive filter itself."

—Weinstein (1977)

Bursting in Adaptive Feedback Systems

The adaptive filter in Fig. 3 is rarely studied in the adaptive signal processing literature without the removal of the feedback path connecting the past transmitted (error) signal, e, to the received signal, u. Acknowledging the encasement of the adaptive filter within a feedback loop, as shown in Fig. 3, resonates with the structure of an adaptive feedback controller. In the early 1980s, a hot topic within the field of adaptive control was the issue of operating conditions and algorithm refinements guaranteeing robustness. The challenge was

crystalized by the examples in Rohrs et al. (1985) of non-robust behavior of adaptive control algorithms proven to be globally convergent under more idealized circumstances. One of these examples exhibited parameter drift that erupted into bursting, which was connected to a lack of appropriate excitation (Anderson, 1985; Anderson et al., 1986).

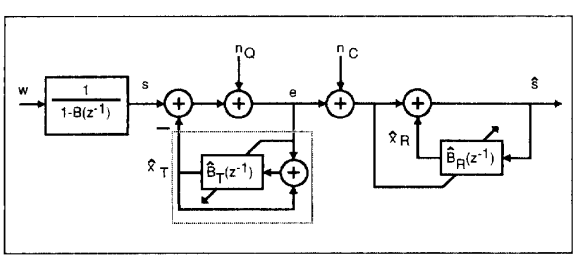
As an instructive example for the echo canceller (Fig. 3), assume $N = 0$, $\delta = 1$, $C(z^{-1}) = \gamma$, and γ, f, and n are positive constants. Thus, $e(k) = (b_0 - \hat{b}_0)u(k) + n(k)$. With $u(k) = f(k-1) + e(k-2)$, and assuming a stable stationary point is reached that zeros e while stabilizing the feedback loop, $\hat{b}_0 = b_0 + (n/f)$. For this value of $\hat{b}_0$, the characteristic equation of the feedback loop in Fig. 3 is $z^2 + (n/f)\gamma$, which is stable as long as $n\gamma/f > 1$. Thus, if $(n/f) > (1/\gamma)$, the assumption of stability of the feedback loop is invalidated. This n-big/f-small scenario is elaborated upon in Sethares et al. (1989). For a start, at $\hat{b}_0 = 0$ with $(n/f) > (1/\gamma)$ and $(n/f) >> b_0$, the prediction error is quickly reduced to near zero and then slowly approaches zero as $\hat{b}_0$ begins to drift toward larger and larger values. Eventually, it results in (temporary) destabilization. This destabilization scenario has implications for a doubletalk detector intended (at the least) to avoid bursting (Ding et al. 1990). An overly zealous doubletalk detector can degrade performance by unnecessarily halting adaptation. A too-relaxed tuning of the doubletalk detector can lead to performance degrading parameter drift/bias, potentially resulting in intermittent catastrophe. With respect to bursting avoidance, the balancing act is to switch off adaptation just upon reaching an application-dependent, dangerous level of n relative to f.

This parameter drift, followed by temporary instability, is the signature of bursting in adaptive echo cancellation and adaptive control. Indeed, simplified versions of the two problems can be reduced to the same system. For this system, bursting is a manifestation of a period doubling bifurcation, which has implications for the burst-avoiding tuning of algorithm modifications such as leakage (Rey et al., 1991). For other than the simplest of problems, describing these connections in order to provide a handle for their exploitation remains a daunting task (Rey et al., 1994).

Speech Coding

"By representing a correlated waveform in terms of appropriate difference samples, or prediction error samples, one can achieve an increased SNR at a given bit rate; or equivalently, a reduction of bit rate for a given requirement of SNR."

—Jayant and Noll (1984)



4. Backward adaptive differential pulse code modulation.

Approximately one fewer bit per sample can be used for every factor of four increase in the signal-to-noise ratio (SNR) improvement, i.e., the variance of the original source divided by the variance of the prediction error to be quantized and transmitted. Typically, with long-term speech or image spectra, the right predictor could offer an increase of up to a factor of 8, or 16, respectively. Unfortunately, in some applications the SNR improvement figure is not directly related to improved speech or image understanding by a human.

Reconstruction of the original source at the receiver from the received version of the transmitted prediction error can be accomplished (Atal and Schroeder, 1970) with the transmission of side information, i.e., the currently appropriate predictor parameters selected at the transmitter to decorrelate the source. Unfortunately, this side information consumes channel capacity in certain applications. Avoiding a side information transmission requirement is one of the advantages of the structure in Fig. 4 (Gibson, 1978; Gibson, 1980).

Equation Error Formulation

In Fig. 4, the source to be encoded and transmitted is s. To isolate a time-series-style identification problem, we will assume (as discussed in Deller et al. (1993)) that s is generated by the autoregressive filtering of a (fictitious, i.e., unmeasurable) white signal w

$$s(k) = w(k) + \sum_{i=1}^{N} b_i s(k-i). \tag{7}$$

The transmitter takes in s and produces the quantized e. The quantization error is indicated by n_Q. In practice, the quantizer is also adapted, but not in a typical (i.e., LMS-like) parameter-adaptive-linear-filter manner (Jayant and Noll, 1984). With our focus here on the predictor parameter adaptation, we will separate the adaptive predictor design from the adaptive quantizer design. This is a common approach. An uncommon approach would be to attempt to simultaneously adapt the quantizer levels and the predictor parameters (Crisafulli et al., 1994).

The backward predictor in the dotted box in Fig. 4 uses e both to force and adapt the predictor parameters, which are the coefficients $\hat{b}_{iT}$ of the polynomial

$$\hat{B}_T(z^{-1}) = \sum_{i=1}^{N} \hat{b}_{iT} z^{-i}. \tag{8}$$

In the absence of channel noise n_C, the reconstruction at the receiver also updates its predictor coefficients with e. For fixed $\hat{b}_{iT}$ and $\hat{b}_{iR}$ the transfer function from s to e is $1 - \hat{B}_T(z^{-}1)$ and from e to $\hat{s}$ is $1/[1 - \hat{B}_R(z^{-}1)]$. When n_Q and

n_C are negligible, $b_i = \hat{b}_{iT}$ will result in e matching w. When n_C is negligible, $\hat{b}_{iT} = \hat{b}_{iR}$ will result in $\hat{s}$ matching $s + n_Q$.

In the transmitter,

$$\hat{x}_T(k) = \sum_{i=1}^{N} \hat{b}_{iT}(k)\left[\hat{x}_T(k-i) + e(k-i)\right] \tag{9}$$

and

$$e(k) = s(k) - \hat{x}_T(k) + n_Q(k). \tag{10}$$

Rewrite Eq. 10 as $s(k) = e(k) - n_Q(k) + \hat{x}_T(k)$, filter this version of s with B and substitute into Eq. 7 for s in Eq. 10. Also using Eq. 9 for $\hat{x}_T$ in Eq. 10 yields

$$e(k) = \sum_{i=1}^{N}\left[b_i - \hat{b}_{iT}(k)\right]\left[e(k-i) + \hat{x}_T(k-i)\right] + w(k) + n_Q(k) - \sum_{i=1}^{N} b_i n_Q(k-i). \tag{11}$$

The equation error formulation of Fig. 1 and Eq. 3 emerges with $e + \hat{x}_T$, corresponding to u, $s + n_Q$ to $y + n$, $\hat{x}_T$ to $\hat{y}$, $\tilde{b}_i$ to $b_i - \hat{b}_{iT}$, and w plus the FIR filtering of n_Q through $1 - B(z^{-1})$ to n.

The LMS algorithm from Eq. 6, given Eq. 11 for adapting the predictor parameters, is

$$\hat{b}_i(k+1) = \hat{b}_i(k) + \mu e(k)\left[e(k-i) + \hat{x}(k-i)\right]. \tag{12}$$

With reference to Fig. 4, when $n_C(k) = 0$ for all k, Eq. 12 works for both the transmitter $\hat{b}_{iT}$ and the receiver $\hat{b}_{iR}$. To assist resynchronization of the transmitter and receiver predictors when n_C is not zero, Cohn and Melsa (1975) recommend leakage. Leakage refers to the dissipation of the estimate in the continued presence of a zero correction term. This is accomplished in the basic adaptation formula that computes the new estimate as a sum of the old estimate and a correction term by scaling the old estimate with a value less than one. This converts Eq. 12 to

$$\hat{b}_i(k+1) = (1-\lambda)\hat{b}_i(k) + \mu e(k)\left[e(k-i) + \hat{x}(k-i)\right], \tag{13}$$

where the leakage coefficient λ is positive and typically smaller than the stepsize μ. Any difference between $\hat{b}_{iT}$ and $\hat{b}_{iR}$ results in added distortion.

Output Error Formulation with a Pole-Zero Predictor

In the early 1980s, a standard (then labeled CCITT Recommendation G.721) was developed to digitize speech with toll quality (i.e., 8-bit log pulse code modulation (PCM) quality) at 32kb/s with a 4-bit quantizer. For a description of the adapted predictor, see section 6.5 of Jayant and Noll (1984) or Gibson (1984). One decision was to use a predictor structure to match a a pole-zero model for the coloration of n into s. This suggests the need for an adaptive infinite-impulse-response (IIR) filter algorithm (Johnson, 1984).

Adaptive IIR filter algorithms first began to proliferate in the adaptive filtering literature in the mid 1970s. Regalia (1995) provides a deep and broad view of this topic, as a subject of importance in both signal processing and control. During the late 1970s and early 1980s the identification literature was raided for potential algorithms. Hyperstable output error identifiers (Landau, 1976) were appropriated as adaptive IIR filters driven by a measured input (Larimore et al., 1980). Recursive maximum likelihood (RML) algorithms were proposed in Friedlander (1982) for a variety of whitening applications. As described in Johnson (1984), the time-varying regressor filtering algorithm forms of simplified versions of RML appear more suited to adaptive differential pulse code modulation (ADPCM) than the fixed error filtering commonly associated with hyperstable output error identifiers. However, neither matches the chosen standard, which can be interpreted (Treichler et al., 1987) as using a clever LMS-like reformulation and adding leakage and sign operators on the multiplicands in the update correlation term.

"...The theory of adaptive filtering has been vital in the case of echo cancellation technology. In the case of adaptive DPCM, the situation is interesting: a great deal of well-understood LPC theory was not used in the CCITT's ADPCM because the prediction there was not all-pole and not forward adaptive. It is my opinion that the theories of pole-zero filtering and backward adaptation were less mature than the rest of adaptive filtering theory, and yet the time was ripe in 1984 for a somewhat empirical design of the backward adaptive pole-zero predictor for CCITT-ADPCM."
—from letter to author from N. S. Jayant, January 11, 1991

Bonnet et al. (1990) suggest that the suboptimality of the algorithm used in the standard G.721 (versus RML-like schemes) contributes to a degree of robusteness required for the application. This increased robustness includes a self-stabilization property counteracting a bursting phenomenon (Jaidane-Saidane and Macchi, 1988) that is structurally related to the bursting phenomenon of echo cancellation and adaptive control (Rey et al., 1991).

Signed Algorithms

A striking feature of the adaptive algorithms in the CCITT standard G.721 that receives scant attention in the recursive identification and adaptive control literature is the use of sign operators on the prediction error and regressor terms in the update. While the standard uses a signed algorithm in an adaptive IIR (or output error) setting, for perspective we revert to the more basic FIR adaptive filter problem. A recent survey of signed LMS algorithms and their behavior in relationship to their unsigned cousins appears in Sethares (1993).

With reference to LMS in Eq. 6, three immediate variants spring to mind:

- signed-regressor LMS

$$\hat{B}(k+1) = \hat{B}(k) + \mu\,\mathrm{sign}[U(k)]e(k) \tag{14}$$

• signed-error LMS

$$\hat{B}(k+1) = \hat{B}(k) + \mu U(k)\mathrm{sign}[e(k)] \tag{15}$$

• sign-sign LMS

$$\hat{B}(k+1) = \hat{B}(k) + \mu\,\mathrm{sign}[U(k)]\mathrm{sign}[e(k)] \tag{16}$$

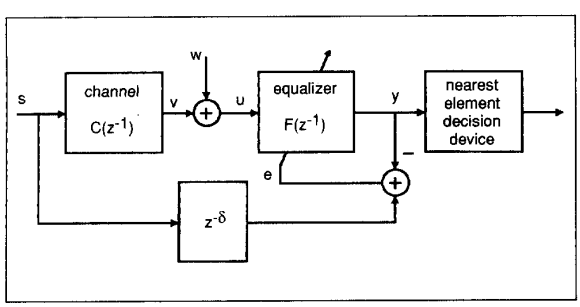
"Overall, it can be said that [sign-sign LMS] is slower than the standard [LMS] algorithm and leads to larger excess output. However, it is very simple; moreover it is robust because of its built-in normalization of its adaptation step, and it can handle nonstationary signals. It is one of the most widely used adaptive filter algorithms."

—Belanger (1987)

Prior to the mid 1980s, almost all examination in the adaptive filtering literature of these signed (or quantized) versions of LMS concerned convergence rate and excess mean squared error performance degradation relative to LMS. The specter of potential ill-convergence by the signed regressor algorithm of Eq. 14 was raised in Claasen and Mecklenbräuker (1981) due to the change in the instantaneous direction vector of the update from U to sign[U]. Bershad (1985) pointed out that some inputs (such as white, zero-mean, and gaussian) could still prove acceptable in an average sense. Claasen and Mecklenbräuker (1986) responded by claiming that other classes of inputs could prove unacceptable, even on average. It turns out that each of these observations is true.

As an identification algorithm where the values of u in U in Eq. 5 could have a strong DC bias, the direction vector sign [U] could always be a vector of plus ones. This would severely constrain the update direction taken by either signed-regressor LMS or sign-sign LMS. An input with a sufficient number of sinusoids for adequate richness (Anderson and Johnson, 1981), and enough positive DC to cause the input samples all to be positive, would result in asymptotic optimality for unsigned LMS from any initialization. If the same signal would cause sign-sign or signed-regressor LMS to misbehave, the applicability of the signed algorithms could be considered limited relative to unsigned LMS.

The first analytical example of such a failure by sign-sign LMS was given in Dasgupta and Johnson (1986). Driven by results from adaptive identification and control from the early 1980s regarding persistent excitation (due in part to its robustifying effect), subsequent studies of signed-regressor (Sethares et al., 1988) and sign-sign (Dasgupta 1990; Dasgupta, 1993) algorithms have helped define the conditions needed to assure global asymptotic (near) optimality in exact matching circumstances. A signed-error version of an output error identifier requires (Garnett et al., 1994) a more stringent operator condition than the strictly positive real (SPR) condition of unsigned output error identifiers (Landau, 1976). While this effort provides some insight into the behavior of signed algorithms, a detailed understanding is not yet available of the foibles of the specific algorithm inside G.721.



5. Adaptive "linear" equalizer with training.

Linear Channel Equalization

"Bandwidth-efficient data transmission over telephone and radio channels is made possible by the use of adaptive equalization to compensate for the time-dispersion introduced by the channel. Spurred by practical applications, a steady research effort over the last two decades has produced a rich body of literature in adaptive equalization and the related more general fields of reception of digital signals, adaptive filtering, and system identification."

—Qureshi (1985)

Twenty years before Qureshi's paper, the seminal paper (Lucky, 1965) appeared with a procedure for design of a tapped-delay-line equalizer to remove the intersymbol interference due to channel distortion in digital communications systems. At that time, the data rate over voice telephone channels was limited to 2400 bits per second due to the onset of overwhelming distortion. Up to that data rate, time-invariant (so-called "compromise") equalizers were designed for a nominal setting within a limited range of expected channel amplitude and delay characteristics. The chapter on equalization in the text (Proakis, 1989) includes figures illustrating templates of this type. Proakis and Miller (1969) relate the tapped-delay-line equalizer, training, and the LMS algorithm.

The linear, tapped-delay-line equalizer in Fig. 5 filters u in an attempt to recover $s(k-\delta)$with

$$y(k) = \sum_{i=0}^{N} f_i(k)u(k-i). \tag{17}$$

Converting y into a member of the source alphabet is done by the nearest element decision device. For example, plus and minus one form a binary source alphabet, for which the sign operator is the nearest element decision device.The source signal, s, is distorted by the channel with transfer function $C(z^{-1})$ into v, and obscured by an additive noise, w before reaching the receiver as

$$u(k) = v(k) + w(k). \tag{18}$$

Equation Error Formulation for an "All-Pole" IIR Channel Model

With a purely autoregressive $C(z^{-1}) = 1/B(z^{-1})$, where $B(z^{-1})$ is a polynomial with a zero delay term so that

$$s(k) = \sum_{i=0}^{N} b_i v(k-i) \quad ,$$

$$v(k) = \frac{s(k)}{b_0} - \sum_{i=1}^{N} \frac{b_i v(k-i)}{b_0}. \tag{19}$$

By prearrangement, during a training period the source entering the channel is known to the receiver. Thus, the prediction error $s(k) - y(k)$ (with δ=0) can be formed. Using Eqs. 17 and 19,

$$\begin{aligned} e(k) = s(k) - y(k) &= \sum_{i=0}^{N} b_i v(k-i) - \sum_{i=1}^{N} f_i(k) u(k-i) \\ &= \sum_{i=0}^{N} (b_i - f_i(k)) u(k-i) - \sum_{i=0}^{N} b_i w(k-i). \end{aligned} \tag{20}$$

With $\tilde{B} = [b_0 - f_0, b_1 - f_1, b_1 - f_2, \ldots, b_N - f_N]^T$ and $U(k)$ defined as in Eq. 5,

$$e(k) = \tilde{B}^T(k) U(k) - \sum_{i=0}^{N} b_i w(k-i). \tag{21}$$

Comparing Fig. 5 (with $C = 1/B$) to Fig. 1 relates F to $\hat{B}$ and $-w$ filtered through B to n. Because the "noise" term $\left(\sum_{i=0}^{N} b_i w(k-i)\right)$ is no longer white, using LMS to minimize $E[e^2]$ will result in a nonzero, asymptotic $\tilde{B}$, i.e., parameter estimate bias.

"Adaptive equalizers, as used in the receivers of data modems operating above 2400 bit/s over telephone channels, are usually realized in the form of a transversal filter with variable tap gains and tap spacing equal to the symbol spacing T. For automatically adjusting the tap gains at the beginning of transmission, and for fine-tuning them later on in an adaptive manner during the entire period of transmission, the least mean-square (LMS) error algorithm has become a standard method."

—Ungerboeck (1976)

"All-Zero" FIR Channel Model

The discrete-time model of symbol (or baud) synchronous behavior of an analog modulation, channel distortion, and demodulation (used in Lucky (1965)) and retained in the modern pedagogy as illustrated by Proakis (1989) uses an FIR model and not the IIR model of Eq. 19. The abstraction in Eq. 19 is limited to stable, minimum-phase channels. But, communication channel impulse responses often exhibit non-minimum-phase characteristics. The possibility exists of using an appropriately parameterized tapped-delay line to approximate the delayed, two-sided inverse of a non-minimum-phase FIR channel. With $C(z^{-1})$ in Fig. 5 a polynomial

$$v(k) = \sum_{i=0}^{P} c_i s(k-i), \tag{22}$$

where $s(k)$ is the source alphabet member of the transmitted sequence at time k and the c_i are the channel impulse response coefficients. The time between transmission of each source sequence member is T seconds. To add channel noise, as illustrated in Fig. 5, we frequently model the received signal u as the sum of v and a (white, zero-mean, gaussian) noise, w (Proakis, 1989)

$$u(k) = \sum_{i=0}^{P} c_i s(k-i) + w(k). \tag{23}$$

A T-spaced tapped delay line for the equalizer also has a finite-impulse-response model

$$y(k) = \sum_{i=0}^{N} f_i u(k-i). \tag{24}$$

The delay time of each delay element in a T-spaced, tapped, delay line is T seconds, with sampling synchronized with the transmission of each new member of the source sequence.

The resulting combined channel-equalizer impulse response coefficient vector

$$H = [h_0, h_1, \ldots, h_{N+P}]^T \tag{25}$$

arising with the convolution of the impulse response vectors of the channel and equalizer can be written as

$$H = \Delta F \tag{26}$$

where

$$\Delta = \begin{bmatrix} c_0 & 0 & \cdots & 0 & 0 \\ c_1 & c_0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ c_{N-1} & c_{N-2} & \cdots & c_0 & 0 \\ c_N & c_{N-1} & \cdots & c_1 & c_0 \\ c_{N+1} & c_N & \cdots & c_2 & c_1 \\ \vdots & \vdots & & \vdots & \vdots \\ c_{P-1} & c_{P-2} & \cdots & c_{P-N} & c_{P-N-1} \\ c_P & c_{P-1} & \cdots & c_{P-N+1} & c_{P-N} \\ 0 & c_P & \cdots & c_{P-N+2} & c_{P-N+1} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & c_P & c_{P-1} \\ 0 & 0 & \cdots & 0 & c_P \end{bmatrix} \tag{27}$$

is the (N + P + 1) by (N + 1) channel impulse response matrix and

$$F = [f_0, f_1, \ldots, f_N]^T. \tag{28}$$

(The description of Δ in Eq. 11 is based on the presumption that $P > N$ because $c_i = 0$ when $i < 0$ or $i > P$. In words, suitable whether $P > N$ or not, Δ is a Toeplitz matrix with the first column containing the $P + 1$ channel parameters c_0 through c_P followed by N zeros and the first row containing c_0 followed by N zeros.) With an objective of a delay of δ units, the desired H is

$$H^* = [0,\ldots,0,1,0,\ldots,0]^T. \tag{29}$$

where the only nonzero element is in the $(\delta + 1)$th entry. Thus, Eq. 26, with H^* from Eq. 29 replacing H, can be solved in a least-squares sense for F, given Δ. With a sufficiently large equalizer length $N + 1$, this solution for F will form an accurate approximation of a truncated, delayed, two-sided inverse of the channel impulse response.

Intersymbol interference reduction short of perfect can prove quite satisfactory. For example, consider the noise-free case with the one in Eq. 29 at the $(\delta + 1)$th location in the desired impulse response. To return the appropriate decision device output, the delayed source $s(k - \delta)$ must equal the quantized version of the equalizer output. With the equalizer output written as the convolution sum of the combined channel-equalizer driven by the source, correct (delayed) source recovery by the decision device requires

$$s(k-\delta) = Q[\sum_{i=0}^{P+N} h_i s(k-i)] \tag{30}$$

for all possible $s(\cdot)$. In Eq. 30, $Q(\cdot)$ signifies quantization to the nearest source alphabet member, and the combined channel-equalizer impulse response coefficients h_i are those achieved with convolution of the channel and the equalizer chosen as the "solution" of the design equation. For a real alphabet with members a minimum of d units apart, Eq. 30 is satisfied if the error between the true source signal and the quantized channel-equalizer output is always less than half the interval between alphabet elements or

$$|s(k-\delta)| - \left|\sum_i h_i s(k-i)\right| < \frac{d}{2} \tag{31}$$

for all possible $s(\cdot)$. As long as the magnitude of the smallest alphabet member(s) is $d/2$, Eq. 31 is satisfied if

$$|h_\delta| > \sum_{i \neq \delta} |h_i| \frac{\max(|s|)}{\min(|s|)} \tag{32}$$

is satisfied, resulting in what is termed an "open-eye" channel-equalizer combination. (See, e.g., Figs. 6 and 7 in Lucky (1965) for examples of closed and open eye pattens.) Certainly $|h_\delta|$ is the term in the combined channel-equalizer impulse response with the largest magnitude. For example, for a binary, discrete alphabet {1, -1}, Eq. 32 reduces to $|h_\sigma| > \sum_{i \neq 0} |h_i|$.

The channel-inverse resolution of Eq. 26 for a sufficiently long f with h replaced by h^* will cause peaks in the equalizer frequency response magnitude at those frequencies where the magnitude of the frequency response of the channel has valleys. Thus, if the channel noise has energy in those frequency bands, this disturbance will be amplified in the equalizer output. This is one of the limitations of the T-spaced linear equalizer structure.

An identification-style problem formulation based on collected response data helps reduce this noise sensitivity. The vector of successive equalizer outputs

$$Y(k) = [y(k), y(k-1), \ldots y(k-M)]^T \tag{33}$$

can be composed as

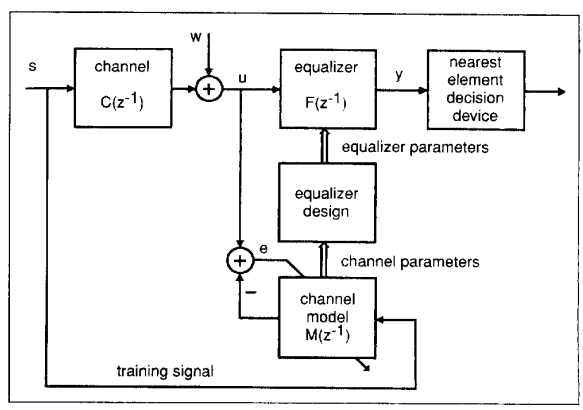
$$Y(k) = G(k)F + [n(k), n(k-1), \ldots, n(k-M)]^T \tag{34}$$

where

$$G(k) = \begin{bmatrix} u(k) & u(k-1) & u(k-2) & \ldots & u(k-N) \\ u(k-1) & u(k-2) & u(k-3) & \ldots & u(k-N-1) \\ \vdots & \vdots & \vdots & & \vdots \\ u(k-M) & u(k-M-1) & u(k-M-2) & \ldots & u(k-M-N) \end{bmatrix}. \tag{35}$$

Effectively, replacing the output vector Y in Eq. 34 by the desired output, i.e., the corresponding delayed source sequence vector, and solving for F via pseudoinversion will minimize the average squared prediction error $s(k-\delta) - y(k)$. Note how this requires knowledge of the transmitted source, at least for the time block over which G is composed. With training, the LMS algorithm of Eq. 6 is a real-time, recursive solution to this problem.

A similar identification problem could be formulated at the receiver for channel model estimation. This estimate could be combined, as in Fig. 6, with your favorite robust (linear) equalizer design. Contrasting this approach with direct adaptation of the equalizer parameters in Fig. 5 is reminiscent of the debate regarding indirect and direct adaptive control begun in the 1970s and reflected in the topical organization of Anderson, et al. (1986), Åström and Wittenmark (1989), and Sastry and Bodson (1989). Because the indirect



6. Simultaneous channel identification and equalization.

approach of channel estimation followed by equalizer redesign is viewed as requiring substantially more computation, the direct approach is commonly chosen to accommodate the data rate given the computational capabilities available at acceptable cost. As computing costs lower, this choice may no longer be automatic.

Convergence rate formulas in, for example, Gunnarsson and Ljung (1989) and Egardt, et al. (1992), indicate that the speed of convergence of LMS in frequency domain curve fitting terms tends to be fastest in those frequency bands with the highest input energy. In the indirect approach, if the source is white, then the convergence rate will be uniformly distributed in frequency. In the direct approach, however, the received signal energy is less in those regions where the channel has a null in its frequency response, meaning that the identification of the channel inverse will be slower in those frequency bands where the channel inverse will be largest. Does this difference carry over into the speed with which an open-eye is attained? If so, this could impact the indirect-versus-direct choice.

Adaptation using Decision-Direction or Training

For the zero-forcing minimization strategy of Lucky (1965), Lucky (1966) proposes an adaptive implementation without a training signal for updating the N by 1 vector of equalizer tapped-delay-line coefficients of the form

$$F(k+1) = F(k) + \mu e(k-\delta) Q[Y(k)] \tag{36}$$

where N is even,

$$\delta = N/2, \tag{37}$$

$Q[\cdot]$ signifies quantization to the nearest source alphabet member,

$$e(k) = Q[y(k)] - y(k), \tag{38}$$

and $Y(k)$ is defined in Eq. 33 with M set to N. Several variants on Eq. 36 are compared in (Hirsch and Wolf, 1970) including the sign-sign version of Eq. 36, which is actually promoted more in Lucky (1966) than Eq. 36, due to the implementation simplicity of its sign-sign version. Lucky (1966) suggests a procedure that uses a training signal to help tune the equalizer until y is a suitably accurate replica of s and then switches to (some version of) Eq. 36 for tracking without training. For blind tracking, the use in Eq. 36 of decision direction in the creation of the error of Eq. 38 exploits the robustness of the open-eye condition to small/slow variations in the channel.

"To conserve signal power and bandwidth, the channel monitoring (or system identification) of the adaptive equalizer must be done using only the normal received data signal and without the benefit of added test information."

—Lucky (1966)

A gradient descent solution to a least-squares setup similar to Eqs. 33-35 is proposed in Lucky and Rudin (1966). Least-squares minimization with training and with decision direction are both proposed in Niessen (1967). Least-squares solution using LMS is studied in Proakis and Miller (1969). With training, the LMS algorithm is

$$F(k+1) = F(k) + \mu U(k)(s(k-\delta) - y(k)) \tag{39}$$

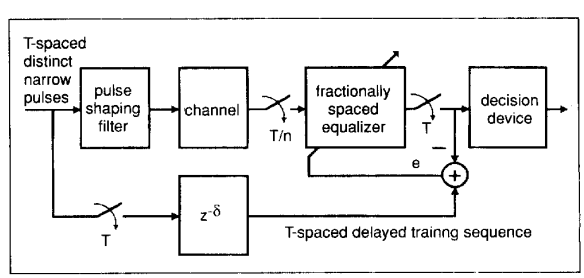
The decision-directed version replaces the source sequence value in Eq. 39, with its current estimate at the output of the memoryless nonlinear decision device

$$F(k+1) = F(k) + \mu U(k)(Q[y(k)] - y(k)). \tag{40}$$

Successful implementations in the late 1960s of trained equalizers using LMS quickly led to performance analyses of excess mean-squared error, convergence rate, and tracking effectiveness, e.g., (Gersho, 1969; Ungerboeck, 1972; Widrow et al., 1976). Such performance analyses of various recursive parameter estimation algorithms under sundry operating scenarios formed the strongest (now traditional) connection between recursive identification and adaptive equalization theories in the 1970s. As a landmark text in terms of pedagogical cohesion in recursive system identification, Ljung and Söderström (1983) captures a large portion of this major effort. The 1970s also saw an explosion in interest in the use of fast Kalman algorithms and lattice realizations to permit the use of fewer data points in the training sequence. As indicated in (Qureshi, 1985), this was a subject in which a number of system identification experts became involved.

Fractional Spacing

Another early pragmatic innovation is the fractionally spaced equalizer, which can be interpreted as sampling the equalizer input at a rate higher than the symbol rate. However, the equalizer output is decimated to the symbol rate. Fractional spacing has been observed to lead to lower minimum mean-squared error in symbol recovery, especially for channels with severe delay distortion near the frequency band edges (Gitlin, Hayes, and Weinstein, 1992). Also, fractionally spaced equalizers can remove the need for a pre-equalizer receive filter. Refer to Figure 7. The pulse shaping filter assists in bandlimiting the frequency content of the signal entering the channel to meet



7. Fractionally-spaced equalizer with training.

constraints on interference with other users in nearby frequency bands. Ideally, a receive filter implemented at the receiver would recover an isolated source pulse as a signal with zero crossings every T seconds except the sample at which this symbol was received.

While a number of papers from the 1970s extol the various virtues of fractionally spaced equalizers, one of their problems studied in (Gitlin et al., 1982) is ambiguity in the stationary points using LMS with training to update—at the symbol rate—the weights on the fractionally spaced equalizer taps. The training decimation relative to the fractionally spaced equalizer tapped delay line sample rate suggests that the fractionally spaced equalizer in Fig. 7 needs to undo the intersymbol interference induced by the pulse shaping filter and channel dynamics only every baud synchronous sample, with the intermediate equalizer outputs taking on any value whatsoever. Thus, various fractionally-spaced equalizer parameterizations could prove equally successful. Gitlin, et al. (1982) suggest using leakage to inhibit the potential for parameter drift and register overflows. Stopping such drift (and subsequent overflow or bursting) with various algorithm modifications and excitation requirements was a prominent component of adaptive identification and control research in the 1980s (Anderson, et al. 1986).

Fractional spacing converts the overspecified design equation $H^* = \Delta F$ from the symbol-spaced case of Eqs. 26-29 into a Diophantine-style equation, commonly associated with algebraic pole placement feedback controller design, as in section 6.7 of Franklin, et al. (1990), or chapter 10 of Åström and Wittenmark (1990). For the moment, imagine that the channel impulse response matrix Δ in Eq. 27 and the equalizer parameter vector F in Eq. 28 are for oversampled versions relative to the source symbol rate. For example, consider the "unit" time delay in Eqs. 23 and 24 to be $T/2$ seconds, where T is the time between transmitted source symbols. This corresponds to $n = 2$ in Fig. 7. At this oversampled rate, every other value in $\{s(k)\}$ is zero. The desired oversampled H^* is [0, ★, 0, ★,...,0, ★, 1, ★, 0,..., ★, 0, ★, 0], where the stars (★) are not of identical value but are instead unconstrained due to the associated zero value of the source signal at those delays.
Of course, the values in the stars' slots will influence the channel noise amplification properties of the equalizer, although they do not effect perfect recovery of a T-spaced source in the absence of channel noise. The design equation $H^* = F$ using an H^* padded with "don't care" slots can be reduced to the set of equations satisfied by removing every other row from Δ in Eq. 27. For $N = P = 3$, this reduced Δ becomes

$$D = \begin{bmatrix} c_0 & 0 & 0 & 0 \\ c_2 & c_1 & c_0 & 0 \\ 0 & c_3 & c_2 & c_1 \\ 0 & 0 & 0 & c_3 \end{bmatrix}. \tag{41}$$

After some rearrangement, with $\delta = 2$,

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} c_0 & 0 & \vdots & 0 & 0 \\ c_2 & c_0 & \vdots & c_1 & 0 \\ 0 & c_2 & \vdots & c_3 & c_1 \\ 0 & 0 & \vdots & 0 & c_3 \end{bmatrix} \begin{bmatrix} f_0 \\ f_2 \\ \cdots\cdots \\ f_1 \\ f_3 \end{bmatrix} \tag{42}$$

is the new design equation. Unlike the design equation $H^* = \Delta F$ from Eqs. 26-29 for a symbol-spaced equalizer, which has no exact solution due to the number of rows of ΔF (and H^*) exceeding the number of rows of F, the fractionally-spaced design equation of Eq. 42 can be solved as long as the square matrix on the right can be inverted.

We can view the comparison of ΔF in Eq. 26-28 to the right side of Eq. 42 with the indicated partition, as a concatenation of two impulse response vector convolutions: (i) one between the T-spaced polynomial of the even coefficients of the oversampled channel, i.e., $C_{\text{even}}(z^{-1}) = c_0 + c_2 z^{-2} + \ldots$, and the "even" portion of the equalizer, i.e., $F_{even}(z^{-1}) = f_0 + f_2 z^{-2} + \ldots$ (where the unit delay is $T/2$ seconds) and (ii) the other between the odd portion of the channel, i.e., $C_{odd}(z^{-1}) = c_1 z^{-1} + c_3 z^{-3} + \ldots$, and the odd portion of the equalizer, i.e., $F_{odd}(z^{-1}) = f_1 z^{-1} + f_3 z^{-3} + \ldots$. This partitioning reveals the Diophantine-style equation format of the reduced design equation in Eq. 42 as

$$C_{even}(z^{-1})F_{even}(z^{-1}) + C_{odd}(z^{-1})F_{odd}(z^{-1}) = z^{-\delta} \tag{43}$$

where δ is even. The polynomial products on the left side of Eq. 43 only have nonzero coefficients for even powers of z^{-1}. If F in this example were chosen with $N > 3$, then the D in Eq. 41 would be altered to have more columns than rows. The design equation then becomes underdetermined and whole subspaces of solutions exist in the equalizer parameter space. This helps interpret the need for leakage discussed above.

Further Blind Linear Equalizer Algorithms

The 1980s saw an explosion of interest in adaptive equalization algorithms that could adapt successfully (e.g., to an eye-opening setting from an initially closed-eye one) without training. As noted above, both training and blind operation were introduced very early (Lucky, 1965; Lucky, 1966). The paradigm used then for blind algorithm development was decision direction, as indicated by the conversion of Eq. 39 to Eq. 40. The design rule for converting an algorithm with training, to one without, is to replace any appearance of the source signal with the output of the decision device. If the channel-equalizer combination is open-eye, this replacement will not alter the algorithm behavior. Indeed, Lucky (1966) established that the decision-directed version of his zero-forcing algorithm would prove asymptotically optimal if initialized at an equalizer parameterization that opens the eye of the channel-equalizer combination.

"Despite the superior performance of the stochastic gradient algorithm, the zero-forcing technique is almost universally

used in microwave digital radio applications. This is certainly due to the simplicity of its implementation."

—Chamberlain, Clayton, Sari, and Vandamme (1986)

The first blind equalizer based on a different strategy (Sato, 1975) appeared a decade after the introduction of decision direction. The primary distinction with the decision-directed LMS algorithm of Eq. 40 is that the reference signal is replaced by a scaled version of the sign of the equalizer output for a multilevel pulse-amplitude-modulated (PAM) source

$$F(k+1) = F(k) + \mu[\gamma \mathrm{sign}(y(k)) - y(k)]U(k) \tag{44}$$

and not the quantized version of the equalizer output emerging from the decision device. For a ±1 alphabet, with $\gamma = 1$, the Sato algorithm of Eq. 44 and the decision-directed LMS algorithm of Eq. 40 are identical. For multilevel PAM, they are different. Using an average cost function stochastic gradient descent style argument, Sato (1975) suggested convergence for a source with infinitely many levels if the initial combined channel-equalizer impulse response coefficients h_i satisfy

$$\frac{\left(\sum_{i=0}^{N+P}|h_i|\right) - \max_i|h_i|}{\max_i|h_i|} < 1. \tag{45}$$

Compare Eq. 45 with the open-eye condition in Eq. 32 for a white zero-mean, binary (±1) source. Simulated examples in Sato (1975) use a center-spike initialization of the equalizer, i.e., $f_\delta(0)$ is nonzero, where δ but all other $f_i(0) = 0$. Global asymptotic optimality is proven by Benveniste, et al. (1980), for a continuously distributed source and an infinite-length equalizer. False minima are noted by Mazo (1980) for a finite-length equalizer and discretely distributed source. Macchi and Eweda (1984) provide the same type of result for the Sato algorithm as described in Lucky (1966) for decision-directed zero-forcing: if the equalizer is initialized with an open-eye channel-equalizer combination, the algorithm will converge to an optimum setting.

A new class of false minima for baud-spaced, finite-length equalizers is decribed in Ding, et al. (1993). Multimodality necessitates clever initialization. The center-spike initialization appears to be the one of choice in the open literature. An approximately matched receive filter is frequently used in practice to initialize a fractionally spaced blind equalizer.

"The most widely tested and used in practice blind equalizer is probably the Godard or constant modulus algorithm"

—Proakis and Nikias (1991)

During the 1980s, a number of algorithms were proposed for blind equalizers that replaced the prediction error in the LMS algorithm with training in Eq. 39 with a memoryless nonlinear function ψ of the equalizer output:

$$F(k+1) = F(k) + \mu U(k)\psi(y(k)). \tag{46}$$

For decision-directed LMS, this error function ψ is simply $\psi_{DD}(y) = Q[y] - y$. For the Sato algorithm, it is $\psi_s(y) = \gamma \mathrm{sign}(y) - y$. A variety of other linear equalizer blind adaptation algorithms of the memoryless error function type, with labels like Bussgang, stop-and-go, and constant modulus to indicate the inspiration of a particular memoryless error function, were proposed in the 1980s. Blind nonlinear equalizers have also been proposed. See the surveys (Johnson, 1991; Proakis and Nikias, 1991; Kennedy, et al., 1992) and the recent collection (Haykin, 1994), especially (Bellini, 1994).

The Godard algorithm (Godard, 1980), or constant modulus algorithm (CMA) (Treichler and Agee, 1983), attempts (in the real case) to minimize the average of $(\rho - y^2)^2$ (where the target modulus ρ is chosen as $(E[s^4(k)])/(E[s^2(k)])$ and has an error function of $\psi_G(y) = y(\rho - y^2)$. In Foschini (1985) and Shalvi and Weinstein (1990), global asymptotic optimality is proven from any intialization of an infinite length equalizer for a discretely valued source. Some attempts have been made using averaging theory to characterize source excitation requirements needed for local stability about a minimum with globally optimal performance (Johnson, et al., 1988). The possibility of false minima for CMA used on a finite-length equalizer is described in Ding, et al. (1991); and similarly in Ding et al. (1994) for a variety of memoryless error function style algorithms. Various performance studies tout CMA as superior to a number of memoryless error function style alternatives (Shynk, et al., 1991; Jablon, 1992). Some circumstances of misbehavior by CMA have also been observed, but remain to be dissected fully (Treichler, 1992).

"The constant modulus (CM) algorithm suggested by John Treichler is currently being investigated to demonstrate its ability to equalize channels carrying constant modulus signals, such as FM and PSK. The simplicity of the algorithm and its intriguing apparent wide range of applications has motivated a search for related algorithms that might be able to solve an even wider range of problems, such as the equalization of channels containing multiple-modulus QAM signals, e.g., CCITT standard V.29."

—from internal company memo by O. L. (Monty) Frost, September 18, 1980, prior to the November issue appearance of Godard (1980)

Just as Sato (1975) suggested a center-spike initialization with only one equalizer tap (the center one) nonzero, so did Godard (1980). The current folklore regarding CMA is that with a white, zero-mean, PAM or quadrature-amplitude-modulated (QAM) source, a sufficiently long equalizer will exhibit asymptotic optimality from an appropriately sized center-spike initialization. Though widely believed, this result has yet to be proven. Unfortunately, source correlation can dismantle this result in circumstances where a white source and a center-spike initialization using the variance matching strategy would lead to asymptotic optimality (Johnson, et al. 1993; LeBlanc, et al., 1994). (The variance-matching-strategy label refers to selection of the center-spike initial magnitude so the initial equalizer output variance

matches that expected from a uniformly distributed source constellation.) Recently, assumptions of exact solvability of Diophantine-style design equations (as in Eqs. 42-43) of a fractionally-spaced equalizer, a white source with equiprobable symbols, and an absence of channel noise have been shown to result in asymptotic optimality by CMA from any initialization (Fijalkow, et al., 1994). This is due to the resulting "egg-carton" type shape of the associated performance surface.

The Advanced Digital High Definition Television (HDTV) system proposed (Hulyalkar, et al., 1993) by the Advanced Television Research Consortium for terrestial simulcast delivery of HDTV utilizes the Godard algorithm for blind start-up of its adaptive channel equalization. After start-up, the system switches to an LMS-style decision-directed adaptive equalizer algorithm for the 32-QAM source. Square-root raised cosine pulse shaping is used. The maximum multipath delay of significance is 35 microseconds. No more than 8 significant multipath echoes are assumed present. On a 6MHz channel, this adaptive equalizer is capable of carrying 20 Mb/s with stringent cochannel interference requirements.

For digital microwave radio applications, Wolff, et al. (1987) describes a $T/2$ spaced equalizer with 64 taps realized for a range of IF input frequencies and bandwidths, source modulation formats (including 64-QAM), and baud rates (up to 40 Mbaud/s). At the 40 Mbaud rate, 64 taps corresponds to an equalizer impulse response duration of 800 nsec (64 $T/2$ spaced taps with $T = 1/(40 \times 10^6)$). The adaptation is decimated to once every 256 symbols. After start-up using CMA, the system switches to LMS-style decision-direction. When successful, the time to the switch is less than 10^4 updates per tap.

Blind linear equalization presents a parameter identification problem outside the prediction error class nominally associated with recursive system identification algorithms. However, connections to the approximate gradient-descent view, which is rampant in adaptive filtering and identification, suggest the utility of analytical techniques from both fields. The growth of wireless applications that could use blind equalizers continues to fuel the search for new algorithms and deeper understanding of old ones that remain in use.

Concluding Remarks

This article has drawn three main threads intertwining mid-1980s concerns in communication system applications of an LMS-adapted tapped-delay-line filter, to mid-1980s adaptive identification and control issues:

(a) echo cancellation → adaptation within a closed-loop → bursting in adaptive control with poor excitation → singing with doubletalk

(b) backward adaptive differential pulse code modulation with a pole-zero predictor → adaptive IIR filtering → output error identification → CCITT standard G.721 → signed adaptive algorithms → alterations in persistent excitation conditions (and associated operator conditions) leading to robust performance relative to unsigned algorithms

(c) linear channel equalization → indirect and direct adaptation → fractionally-spaced realization → pole-placement control design → no training → multimodal performance function interpretation from average update kernel analysis

Since the mid-1980s, the signal processing technology in each of the application areas considered in this article has expanded well beyond the simple LMS adapted tapped-delay-line filter solution that served as the core for each of the threads above. Low bit rate speech coding is now dominated by code excited linear prediction (CELP) methods (Schroeder and Atal, 1985). Echo cancellers now include large numbers of taps, stepsize normalization, and other features due to improved implementation capabilities. Accommodating long delays encountered with acoustic echoes in teleconferencing is a major issue in current echo cancellation work. Rapid variation in channel dynamics in mobile telephony has hastened the development of equalizers that can be satisfactorily tuned from extremely few data points. Nonlinear equalizers using maximum likelihood sequence estimation in the form of the Viterbi algorithm (Forney, 1972) dominate equalizer technology in digital cellular telephone systems. Another nonlinear equalizer structure using decision feedback (Belfiore and Park, 1979) is currently popular for its ability to enhance the performance of a linear equalizer. These applications abound with modes of interaction between adaptive signal processing and system identification and control, that extend well beyond the basic adaptive filter/identifier format exploited in this article.

"From there to here, from here to there, funny things are everywhere."

—*Dr. Suess,*
One Fish, Two Fish, Red Fish, Blue Fish, *1960.*

Acknowledgment

"Just because some of us can read and write and do a little math, that doesn't mean we deserve to conquer the Universe."

—*Kurt Vonnegut,* Hocus Pocus, *1990.*

I would like to acknowledge the debt I owe my various co-authors on adaptive filtering, identification, and control topics for the education they provided me over the last decade (in order of appearance since 1985): W. A. Sethares, D. A. Lawrence, R. R. Bitmead, B. D. O. Anderson, R. L. Kosut, S. Dasgupta, P. V. Kokotovic, L. Praly, B. D. Riedle, C. E. Rohrs, J. D. Mills, J. R. Treichler, M. G. Larimore, I. M. Y. Mareels, M. Gevers, M. A. Poubelle, C. R. Elevitch, G. J. Rey, J. D. Gibson, B. Koo, Z. Ding, G. A. Williamson, J. S. Garnett, R. A. Kennedy, A. M. Baksho, T. L. Fine, S. Crisafulli, R. B. Freund, L. Ljung, C. R. Pollock, V. Wolff, P. C. E. Bennett, J. P. LeBlanc, V. Krishnamurthy, B. Egardt,

G. Kubin, M. R. Frater, C. Canudas de Wit, B. Brogliato, K. Dogancay, A. J. Connolly, A. G. Partanen, I. Fijalkow, F. López de Victoria, and A. Bosser.

I am grateful for funding during the past decade from the National Science Foundation, Tellabs Research Laboratory, and Applied Signal Technology.

"I found it hard, it was hard to find / Oh well, whatever, nevermind."

—*Kurt Cobain,* Smells Like Teen Spirit, *1991.*

C. Richard Johnson Jr. is currently a Professor of Electrical Engineering and a member of the Graduate Field of Applied Mathematics at Cornell University, Ithaca, NY. His research in adaptive parameter estimation theory with applications in digital control and signal processing (for which he was elected a Fellow of the IEEE in 1989) has been supported by the National Science Foundation, the Engineering Foundation, the National Aeronautics and Space Administration, Tellabs Research Laboratory, MOOG Technology Center, and Applied Signal Technology. His current research interest is blind adaptive fractionally-spaced linear equalization for intersymbol interference removal from received QAM sources.

References

1. Anderson, B. D. O. (1985). "Adaptive Systems, Lack of Persistency of Excitation and Bursting Phenomena." *Automatica*, 21, 247-258.

2. Anderson, B. D. O., R. R. Bitmead, C. R. Johnson, Jr., P. V. Kokotovic, R. L. Kosut, et al., (1986). *Stability of Adaptive Systems*. MIT Press, Cambridge.

3. Anderson, B. D. O. and C. R. Johnson, Jr. (1982). "Exponential Convergence of Adaptive Identification and Control Algorithms," *Automatica, 18, 1-13.*

4. Åström, K. J. and B. Wittenmark (1989). *Adaptive Control*, Addison Wesley, Reading.

5. Åström, K. J. and B. Wittenmark (1990). *Computer-Controlled Systems: Theory and Design*, Addison Wesley, Reading.

6. Atal, B. S. and M. R. Schroeder (1970). "Adaptive Predictive Coding of Speech Signals." *Bell Syst. Tech. J.*, 49, 1973-1986.

7. Becker, F. K. and H. R. Rudin (1966). "Application of Automatic Transversal Filters to the Problem of Echo Suppression." *Bell Syst. Tech. J.*, 45, 1847-1850.

8. Belfiore, C. A. and J. H. Park (1979). "Decision Feedback Equalization." *Proc. IEEE*, 67, 1143-1156.

9. Bellanger, M. G. (1987). *Adaptive Digital Filters and Signal Analysis*. Marcel-Dekker, New York.

10. Bellini, S. (1994). "Bussgang Techniques for Blind Deconvolution and Equalization." *Blind Deconvolution*, ed. S. Haykin, Prentice-Hall, Englewood Cliffs, 8-59.

11. Benveniste, A., M. Goursat, G. Ruget (1980). "Robust Identification of a Nonminimum Phase System: Blind Adjustment of a Linear Equalizer in Data Communications." *IEEE Trans. on Automatic Control*, AC-25, 385-399.

12. Bershad, N. J. (1985). "Comments on 'Comparison of the Convergence of Two Algorithms for Adaptive FIR Digital Filters.' " *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-33, 1604-1606.

13. Bonnet, M., O. Macchi, and M. Jaidane-Saidane (1990). "Theoretical Analysis of the ADPCM CCITT Algorithm." *IEEE Trans. on Comm.*, 38, 847-858.

14. Chamberlain, J. K., F. M. Clayton, H. Sari, and P. Vandamme (1986). "Receiver Techniques for Microwave Digital Radio." *IEEE Communications Magazine*, 24, 11:43-54.

15. Claasen, T. A. C. M. and W. F. G. Mecklenbräuker (1981). "Comparison of the Convergence of Two Algorithms for Adaptive FIR Digital Filters." *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-29, 670-678.

16. Claasen, T. A. C. M. and W. F. G. Mecklenbräuker (1986). Authors' Reply to "Comments on 'Comparison of the Convergence of Two Algorithms for Adaptive FIR Digital Filters.'" *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-34, 202-203.

17. Cohn, D. L. and J. L. Melsa (1975). "The Residual Encoder - An Improved ADPCM System for Speech Digitization." *IEEE Trans. on Communications*, COM-23, 935-941.

18. Crisafulli, S., G. J. Rey, C. R. Johnson, Jr., and R. A. Kennedy (1994). "A Coupled Approach to ADPCM Adaptation." *IEEE Trans. on Speech and Audio Processing*, 2, 90-93.

19. Dasgupta S. and C. R. Johnson, Jr. (1986). "Some Comments on the Behavior of Sign-Sign Adaptive Identifiers." *Systems and Control Letters*, 7, 75-82.

20. Dasgupta, S., C. R. Johnson, Jr., and A. M. Baksho (1990). "Sign-sign LMS Convergence with Independent Inputs," *IEEE Trans. on Information Theory*, 36, 197-201.

21. Dasgupta, S., C. R. Johnson, Jr., and A. M. Baksho (1993). "Characterizing Persistent Excitation for the Sign-Sign Equation Error Identifier." *Automatica*, 29, 1473-1489.

22. Deller, J. R., Jr., J. G. Proakis, and J. H. L. Hansen (1993). *Discrete-Time Processing of Speech Signals*, MacMillan, New York, chapter 5.

23. Ding, Z., C. R. Johnson, Jr., and W. A. Sethares (1990). "Frequency-dependent Bursting in Adaptive Echo Cancellation and its Prevention using Double-talk Detectors." *Int. Journal of Adaptive Control and Signal Processing*, 4, 219-236.

24. Ding, Z., R. A. Kennedy, B. D. O. Anderson, and C. R. Johnson, Jr. (1991). "Ill-Convergence of Godard Blind Equalizers in Data Communication Systems." *IEEE Trans. on Communications*, 39, 1313-1327.

25. Ding, Z., R. A. Kennedy, B. D. O. Anderson, and C. R. Johnson, Jr. (1993). "Convergence of Sato Blind Algorithm and Generalizations Under Practical Constraints." *IEEE Trans. on Information Theory*, 39, 129-144.

26. Ding, Z., C. R. Johnson, Jr., and R. A. Kennedy (1994). "Global Convergence Issues with Linear Blind Adaptive Equalizers." *Blind Deconvolution*, ed. S. Haykin, Prentice-Hall, Englewood Cliffs, 60-120.

27. Duttweiler, D. L. (1978). "A Twelve-Channel Digital Echo Canceller." *IEEE Trans. on Communications*, COM-26, 647-653.

28. Egardt, B., C. R. Johnson, Jr., L. Ljung, and G. A. Williamson (1992). "Adaptive Systems Performance in the Frequency Domain." *Preprints 4th IFAC Symposium on Adaptive Systems in Control and Signal Processing*, 33-39.

29. Fijalkow, I., F. López de Victoria, and C. R. Johnson, Jr. (1994). "Adaptive Fractionally Spaced Blind Equalization." *Proc. 6th IEEE DSP Workshop*, 257-260.

30. Forney, G. D. (1972). "Maximum Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference." *IEEE Trans. on Information Theory*, IT-18, 363-378.

31. Foschini, G. J. (1985). "Equalization Without Altering or Detecting Data." *AT&T Technical J.*, 64, 1885-1911.

32. Franklin, G. F., J. D. Powell, and M. L. Workman (1990). *Digital Control of Dynamic Systems*, 2nd ed., Addison Wesley, Reading.

33. Friedlander, B. (1982). "System Identification Techniques for Adaptive Signal Processing." *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-30, 240-246.

34. Garnett, J., S. Dasgupta, and C. R. Johnson, Jr. (1994). "Convergence of the Signed Output Error Adaptive Identifier." *IEEE Trans on Automatic Control*, 39, 1387-1399.

35. Gersho, A. (1969). "Adaptive Equalization of Highly Dispersive Channels for Data Transmission." *Bell Syst. Tech. J.*, 48, 55-70.

36. Gibson, J. D. (1978). "Sequentially Adaptive Backward Prediction in ADPCM Speech Coders." *IEEE Trans. on Communications*, COM-26, 145-150.

37. Gibson, J. D. (1980). "Adaptive Prediction in Speech Differential Encoding Systems." *Proc. IEEE*, 68, 488-525.

38. Gibson, J. D. (1984). "Adaptive Prediction for Speech Encoding." *IEEE ASSP Magazine*, 1, 12-26.

39. Gitlin, R. D., J. F. Hayes, and S. B. Weinstein (1992). *Data Communication Principles*, Plenum Press, New York.

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

41. Godard, D. N. (1980). "Self-recovering Equalization and Carrier Tracking in Two-dimensional Data Communication Systems." *IEEE Trans. on Comm.*, COM-28, 1867-1875.

42. Gritton, C. W. K. and D. W. Lin (1984), "Echo Cancellation Algorithms." *IEEE ASSP Magazine*, 1, 30-38.

43. Gunnarsson S. and L. Ljung (1989). "Frequency Domain Tracking Characteristics of Adaptive Algorithms. *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 37, 1072-1089.

44. Haykin, S. (ed.) (1994). *Blind Deconvolution*, Prentice-Hall, Englewood Cliffs.

45. Hirsch, D. and W. J. Wolf (1970). "A Simple Adaptive Equalizer for Efficient Data Transmission." *IEEE Trans. on Commun. Technol.*, COM-18, 5-12.

46. Honig, M. L. and D. G. Messerschmitt (1984). *Adaptive Filters: Structures, Algorithms, and Applications*, Kluwer, Boston.

47. Hulyalkar, S. N., Y-S. Ho, K. S. Challapali, D. A. Bryan, C. Basile, H. White, N.D. Wilson, and B. Bhatt (1993). "Advanced Digital HDTV Transmission System for Terrestial Video Simulcasting." *IEEE Journal on Selected Areas in Communications*, 11, 119-126.

48. Jablon, N. K. (1992). "Joint Blind Equalization, Carrier Recovery, and Timing Recovery for High-Order QAM Signal Constellations." *IEEE Trans. on Signal Processing*, 40, 1383-1398. 49. Jaidane-Saidane, M. and O. Macchi, "Quasi-Periodic Self-Stabilization of Adaptive ARMA Predictors." *Int. J. of Adaptive Control and Signal Processing*, 2, 1-31.

50. Jayant, N. S. and P. Noll (1984). *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, chapter 6, Prentice-Hall, Englewood Cliffs.

51. Johnson, C. R., Jr. (1984). "Adaptive IIR Filtering: Current Results and Open Issues." *IEEE Trans. on Information Theory*, IT-30, 237-250.

52. Johnson, C. R., Jr. (1991). "Admissibility in Blind Adaptive Channel Equalization." *IEEE Control Systems Magazine*, 11, 3-15.

53. Johnson, C. R., Jr., S. Dasgupta, and W. A. Sethares (1988). "Averaging Analysis of Local Stability of a Real Constant Modulus Algorithm Adaptive Filter." *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 36, 900-910.

54. Johnson, C. R., Jr., J. P. LeBlanc, and V. Krishnamurthy (1993). "Godard Blind Equalizer Misbehavior with Correlated Sources: Two Examples." *Journal Marocain d'Automatique, d'Informatique et de Traitement du Signal*, 2, 1-39.

55. Kennedy, R. A., K. Yamazaki, Z. Ding, S. Verdu, and S. Vembu (1992). "Approaches to Blind Equalization: Open Problems. *Preprints 4th IFAC Int. Symp. on Adaptive Systems in Control and Signal Processing*, 627-632.

56. Landau, I. D. (1976). "Unbiased Recursive Identification Using Model Reference Adaptive Techniques." *IEEE Trans. on Automatic Control*, AC-21, 194-202.

57. Larimore, M. G., J. R. Treichler, and C. R. Johnson, Jr. (1980). "SHARF: An Algorithm for Adapting IIR Digital Filters." *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-28, 428-440.

58. LeBlanc, J. P., K. Dogancay, R. A. Kennedy, and C. R. Johnson, Jr. (1994). "Effects of Input Data Correlation on the Convergence of CMA." *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing.*

59. Ljung, L. and T. Söderström (1983). *Theory and Practice of Recursive Identification*. MIT Press, Cambridge.

60. Lucky, R. W. (1965). "Automatic Equalization for Digital Communication." *Bell System Technical Journal*, 44, 547-588.

61. Lucky, R. W. (1966). "Techniques for Adaptive Equalization of Digital Communication Systems." *Bell Syst. Tech. J.*, 45 255-286.

62. Lucky, R. W. and H. Rudin (1967). "An Automatic Equalizer for General-Purpose Communication Channels." *Bell Syst. Tech. J.*, 46, 2179-2208.

63. Mendel, J. M. (1973). *Discrete Techniques of Parameter Estimation: The Equation Error Formulation*. Marcel-Dekker, New York.

64. Macchi, O. and E. Eweda (1984). "Convergence Analysis of Self-Adaptive Equalizers." *IEEE Trans. on Info. Theory*, 46, 161-176.

65. Mazo, J. E. (1980). "Analysis of Decision-directed Equalizer Convergence." *Bell Syst. Tech. J.*, 59, 1857-1876.

66. Messerschmitt, D. G. (1984). "Echo Cancellation in Speech and Data Transmission." *IEEE J. Selected Areas in Comm.*, SAC-2, 283-297.

67. Murano, K. (1992) "Adaptive Signal Processing Applied in Telecommunications." *Preprints 4th IFAC Int. Symp. on Adap. Syst. in Control and Signal Proc.*, pp. 269-279.

68. Niessen, C. W. (1967). "Automatic Channel Equalization Algorithm." *Proc. IEEE*, 55, 689.

69. Proakis, J. G. (1989). *Digital Communications*. McGraw-Hill.

70. Proakis, J. G. and J. H. Miller (1969). "An Adaptive Receiver for Digital Signaling Through Channels with Intersymbol Interference." *IEEE Trans. on Information Theory*, IT-15, 484-497.

71. Proakis, J. G. and C. L. Nikias (1991). "Blind Equalization." *SPIE Vol. 1565 Adaptive Signal Processing*, 76-87.

72. Qureshi, S. U. H. (1985) "Adaptive Equalization." *Proc. IEEE*, 73, 1349-1387.

73. Regalia, P. A., (1995). *Adaptive IIR Filtering in Signal Processing and Control*, Marcel-Dekker, New York.

74. Rey, G. J., R. R. Bitmead, and C. R. Johnson, Jr. (1991). "The Dynamics of Bursting in Simple Adaptive Feedback Systems with Leakage." *IEEE Trans. on Circuits and Systems*, 38, 476-488.

75. Rey, G. J., R. R. Bitmead, and C. R. Johnson, Jr. (1994). "Qualitative Classification of the Dynamics of a Class of Adaptive Feedback Systems." *IEEE Trans. on Circuits and Systems - I:Fundamental Theory and Applications*, 41, 529-538.

76. Rohrs, C. E., L. Valavani, M. Athans,and G. Stein (1985). "Robustness of Continuous-Time Adaptive Control Algorithms in the Presence of Unmodelled Dynamics." *IEEE Trans. on Auto. Control*, AC-30, 881-889.

77. Sastry, S. S. and M. Bodson (1989). *Adaptive Control: Stability, Convergence, and Robustness*, Prentice-Hall, Englewood Cliffs.

78. Sato, Y. (1975). "A Method of Self-recovering Equalization for Multilevel Amplitude Modulation." *IEEE Trans. on Comm.*, COM-23, 679-682.

79. Schroeder, M. R. and B. S. Atal (1985). "Code-Excited Linear Prediction (CELP): High-Quality Speech at Very Low Bit Rates." *Proc. IEEE ICASSP*, 937-940.

80. Sethares, W. A. (1993). "The LMS Family." *Efficient System Identification and Signal Processing Algorithms*, ed. N. Kalouptsidis and S. Theodoridis, Springer-Verlag, Berlin.

81. Sethares, W. A., I. M. Y. Mareels, B. D. O. Anderson, C. R. Johnson, Jr.,and R. R. Bitmead (1988). "Excitation Conditions for Signed-regressor Least Mean Squares Adaptation." *IEEE Trans. on Circuits and Systems*, 35, 613-624.

82. Sethares W. A., C. R. Johnson, Jr., and C. E. Rohrs (1989). "Bursting in Adaptive Hybrids." *IEEE Trans. on Communications*, 37, 791-799.

83. Shalvi, O. and E. Weinstein (1990). "New Criteria for Blind Deconvolution of Non-minimum Phase Systems (Channels)." *IEEE Trans. on Info. Theory*, IT-36, 312-321.

84. Shynk, J. J., R. P. Gooch, G. Krishnamurthy, and C. K. Chan (1991). "A

Comparative Performance Study of Several Blind Equalization Algorithms." *SPIE Vol. 1565 Adaptive Signal Processing*, 102-117.

85. Solo, V. and X. Kong (1995). *Adaptive Signal Processing Algorithms: Stability and Performance*, Prentice-Hall, Englewood Cliffs.

86. Sondhi, M. M. (1967). "An Adaptive Echo Canceller." *Bell System Technical Journal*, 46, 487-511.

87. Sondhi M. M. and D. A. Berkley (1980). "Silencing Echoes on the Telephone Network." *Proc. IEEE*, 68, 948-963.

88. Sondhi, M. M. and A. J. Presti (1966). "A Self-Adaptive Echo Canceller." *Bell System Technical Journal*, 45, 1851-1854.

89. Treichler, J. R. (1992). "Applications of Blind Equalization Techniques to Voiceband and RF Modems." *Preprints 4th IFAC Int. Symp. on Adaptive Systems in Control and Signal Processing*, 705-713.

90. Treichler, J. R. and B. G. Agee (1983). "A New Approach to Multipath Correctionof Constant Modulus Signals." *IEEE Trans. on Acoustics, Speech, and Signal Processing*, ASSP-31, 349-472.

91. Treichler, J. R., C. R. Johnson, , Jr., and M. G. Larimore (1987). *Theory and Design of Adaptive Filters*. Wiley-Interscience, New York.

92. Ungerboeck, G. (1972). "Theory on the Speed of Convergence in Adaptive Equalizers for Digital Communication." *IBM J. Res. Develop.*, 16, 546-555.

93. Ungerboeck, G. (1976). "Fractional Tap-spacing Equalizer and Consequences for Clock Recovery in Data Modems." *IEEE Trans. on Communications*, COM-24, 856-864.

94. Weinstein, S. B. (1977). "Echo Cancellation in the Telephone Network." *IEEE Communications Society Magazine*, 9-15.

95. Widrow, B. and M. E. Hoff, Jr. (1960). "Adaptive Switching Circuits." *IRE WESCON Conv. Record*, pt. 4, pp. 96-104.

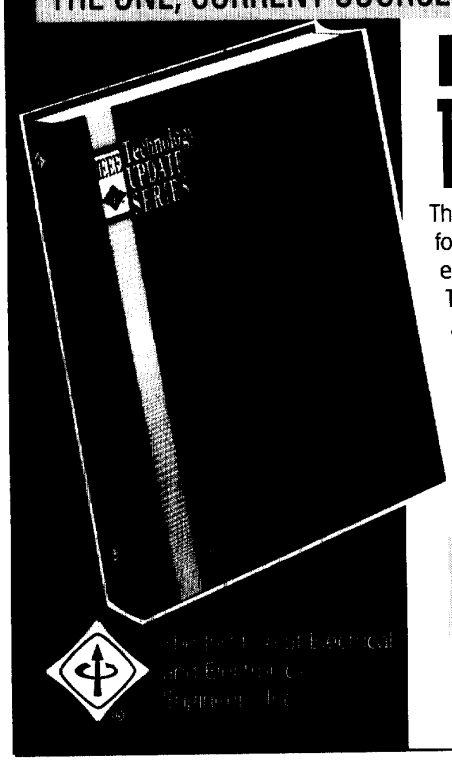
96. Widrow, B., J. R. Glover, Jr., J. M. McCool, J. Kaunitz, C. S. Williams, et al., (1975). "Adaptive Noise Cancelling: Principles and Applications." *Proc. IEEE*, 63, 1692-1716.

97. Widrow, B., J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr. (1976). "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter." *Proc. IEEE*, 64, 1151-1162.

98. Willsky, A. S. (1979). *Digital Signal Processing and Control and Estimation Theory: Points of Tangency, Areas of Intersection, and Parallel Directions*, MIT Press, Cambridge.

99. Wolff, V., R. Gooch, and J. R. Treichler (1987). "Specification and Development of an Equalizer/Demodulator for Wideband Digital Microwave Radio Signals." *WESCON Conf. Record*, paper 6.3.

THE ONE, CURRENT SOURCE FOR THE NEWEST APPLICATIONS IN FUZZY LOGIC!



FUZZY LOGIC
TECHNOLOGY & APPLICATIONS

This 1994 volume details the recent advances in the design, applications, and foundations of fuzzy logic. This volume is the *key* source where a professional entering or working in the field of fuzzy logic can find the *best* practical papers. This probabilistic discipline now provides real-world solutions in process control and manufacturing, navigation, robotics and automation, and control systems.

Fuzzy Logic contains selected technical design and applications papers on fuzzy logic handpicked by acknowledged experts in the field. This material was chosen from recent IEEE conference records with special introductory material by leading figures in the field, including a preface by Lofti Zadeh, the father of fuzzy logic. *This one source covers it all!*

IEEE Member Price: $48.00 List Price: $59.95
IEEE Order No. 94CR1016-PDL ISBN: 0-7803-1383-6
1994 Hardcover 470 pp

FOR FASTEST SERVICE, CALL TOLL-FREE (IN USA)
1-800-678-IEEE
FAX 1-908-981-9667; PHONE 1-908-981-0060