Set-Membership Adaptive Equalization and an Updator-Shared Implementation for Multiple Channel Communications Systems

Sridhar Gollamudi, *Student Member, IEEE*, Samir Kapoor,
Shirish Nagaraj, *Student Member, IEEE*, and Yih-Fang Huang, *Fellow, IEEE*

***Abstract*— This paper considers the problems of channel estimation and adaptive equalization in the novel framework of *set-membership parameter estimation.* Channel estimation using a class of set-membership identification algorithms known as optimal bounding ellipsoid (OBE) algorithms and their extension to track time-varying channels are described. Simulation results show that the OBE channel estimators outperform the least-mean-square (LMS) algorithm and perform comparably with the RLS and the Kalman filter. The concept of set-membership equalization is introduced along with the notion of a *feasible equalizer.* Necessary and sufficient conditions are derived for the existence of feasible equalizers in the case of linear equalization for a linear FIR additive noise channel. An adaptive OBE algorithm is shown to provide a set of estimated feasible equalizers. The selective update feature of the OBE algorithms is exploited to devise an updator-shared scheme in a multiple channel environment, referred to as updator-shared parallel adaptive equalization (U-SHAPE). U-SHAPE is shown to reduce hardware complexity significantly. Procedures to compute the minimum number of updating processors required for a specified quality of service are presented.**

I. Introduction

CHANNEL equalization is a common signal processing technique that compensates for channel-induced signal impairment and the resulting intersymbol interference (ISI) in a digital communication system [1], [2]. Insufficient *a priori* information about the channel and, especially in the case of wireless channels, the time-varying nature of the channel response, necessitate adaptive equalization. Linear transversal equalizers and decision feedback equalizers (DFE's) have been used for many decades in conjunction with deterministic or statistical least-squares (LS) algorithms, like the least-mean-square (LMS) algorithm [1], [3], the Kalman filter [4], [5], and the recursive least-squares (RLS) algorithm [6], [7], to adjust the equalizer coefficients. Parameter estimators can either be employed to directly adapt the equalizer taps or to estimate the impulse response of an FIR channel model, which in turn is used to compute an *optimal* equalizer, usually in the sense of MMSE.

Manuscript received October 31, 1996; revised February 19, 1998. This work was supported in part by the National Science Foundation under Grant MIP 97-05173, by Tellabs Research Center, Mishawaka, IN, and by the Center for Applied Mathematics, University of Notre Dame, Notre Dame, IN. The associate editor coordinating the review of this paper and approving it for publication was Dr. Sergios Theodoridis.

The authors are with the Laboratory for Image and Signal Analysis, Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA (e-mail: huang.2@nd.edu).

Publisher Item Identifier S 1053-587X(98)05955-8.

In this paper, channel estimation and equalization problems are addressed in the framework of a novel system-identification paradigm, viz., set-membership identification (SMI). The SMI approach assumes set-theoretic (instantaneous and deterministic) as opposed to statistical information about the model to compute sets of parameter estimates in the parameter space, called membership sets, that are consistent with the model assumptions and observations. Typical set-theoretic assumptions on the noise are those of bounded magnitude or membership in an ellipsoid.

Among many SMI algorithms, the more popular ones are a set of algorithms termed optimal bounding ellipsoid (OBE) algorithms [8]–[11]. The OBE algorithms provide ellipsoids in the parameter space that outer bound the membership set and incorporate *optimization* of the size of the ellipsoids in some meaningful sense. Despite the fundamental difference in their approaches, the OBE recursions are strikingly similar to those of RLS. Moreover, the centers of the ellipsoids in the OBE algorithms are known to be certain weighted LS estimates [11]. The resemblance of the OBE algorithms to the least-squares algorithms, both in terms of the estimates and the implementation of recursions, makes the former an attractive practical estimation tool since LS techniques are used widely, and a large body of work exists on their efficient implementations. Systolic array implementations of the OBE algorithms are discussed in [12].

The OBE algorithms offer a number of advantages over LS techniques. First, they provide an ellipsoidal bound as an "overlay" over the weighted LS estimate. Second, in the context of adaptive tracking of time-varying parameters, the OBE algorithms offer, in a natural way, an indication of loss of tracking. Such information can be invaluable in preventing *runaway* errors by restarting the algorithm when tracking is lost. Third, earlier studies [9], [11], [13] on the OBE algorithms have shown that they exhibit superior tracking and convergence properties compared with the LS algorithms. Fourth, one of the most attractive features of the OBE algorithms, which also forms a basis for the updator-shared scheme to be presented in this paper, is the novel data-dependent weighting approach, which results in a *discerning update* strategy for the parameter estimates. Unlike the LS estimation schemes that require continual updating regardless of the benefit provided by the data, the OBE algorithms can "intelligently" evaluate incoming data for their potential to improve the quality of the estimate and weight the data accordingly. It is common for the algorithms to discard 70–95% of the data while exhibiting performance very similar to, if not better than, the LS estimators. The large idle times of the parameter updators in OBE-trained equalizers are exploited in Section V to devise an updator-shared scheme, which is referred to as updator-shared parallel adaptive equalization (U-SHAPE), for a number of parallel channel estimators/equalizers.

The notion of *set-membership (SM) equalization* is introduced in this paper. An SM equalizer constrains the error (i.e., the Euclidean distance between the transmitted and equalized symbols) to lie within a specified quantity for all data consistent with the assumed set-theoretic assumptions. This is fundamentally different from conventional minimum mean-squares and LS equalizers [1] in the approach to minimizing ISI. The objective of OBE algorithms for SM equalization is to estimate a set of *feasible equalizers* that achieve this deterministic specification on the symbol error.

The discerning update feature of SM equalization is exploited in this paper to reduce the hardware complexity of a multiple-channel communications system. The proposed scheme is applicable to any system that involves independent adaptive filters running simultaneously at the same location, such as a bank of equalizers at the base station of a time-division/frequency-division multiple access cellular system or a receiver employing diversity channels. Infrequent updating facilitates sharing a small number of updating processors among a larger number of channels via U-SHAPE [14]–[16]. Such sharing of processors is not possible with conventional recursive schemes (such as RLS/LMS) since they entail updating at every instant.

Section II gives a summary of the SMI method and a generalized OBE algorithm. Channel estimation based on the SM framework and an extension to time-varying channels are presented in Section III. SM equalization is defined in Section IV, and an OBE solution is proposed. The existence of a feasible equalizer set and its properties are also discussed in this section. The design of an updator sharing scheme via U-SHAPE is described in Section V. Concluding remarks are made in Section VI.

II. Set-Membership Identification and the OBE Algorithm

In the set-membership framework, the linear-in-parameter model

$$y_n = \theta_*^T \mathbf{x}_n + v_n \tag{1}$$

is considered, where θ_* is the underlying complex-valued parameter vector to be estimated, and $\mathbf{x}_n$ is the measurable N-dimensional input vector to the system. Note that the class of systems modeled by (1) includes the ARX model in which $\mathbf{x}_n$ is the regressor vector of the input and output sequences and several nonlinear systems of practical importance. Specifically, $\mathbf{x}_n$ may contain nonlinear functions of past outputs, as is the case in decision feedback equalization. The noise v_n is assumed to be bounded in magnitude with a known bound γ, i.e.,

$$|v_n| \leq \gamma \qquad \forall\, n. \tag{2}$$

It is pertinent to note that system identification with bounded noise has been treated in the LS framework by imposing a dead zone to the update recursions [17]. Moreover, LS estimation under a bounded-disturbance assumption has also been studied in [18].

Combining (1) and (2) yields the *observation-induced set*, which is defined as the set of all parameter vectors that are consistent with the assumed model and the observation at time instant n

$$\mathcal{O}_n = \{\theta \in \mathbf{C}^N\colon\ |y_n - \theta^T \mathbf{x}_n|^2 \leq \gamma^2\} \tag{3}$$

where $\mathbf{C}^N$ is the N-dimensional complex Euclidean space. The *membership set* ψ_n is defined as

$$\psi_n \triangleq \bigcap_{i \leq n} \mathcal{O}_i. \tag{4}$$

Clearly, if θ_* is time invariant, then $\theta_* \in \psi_n$ for all n.

The membership sets as defined in (4) are convex polytopes in the parameter space that are not easily tracked. OBE algorithms seek ellipsoidal outer approximations of the membership sets. The seminal paper in the development of OBE algorithms was by Fogel and Huang [8]. An OBE algorithm featuring a simplified (linear complexity) information checking procedure was developed by Dasgupta and Huang (D–H/OBE algorithm) in [9]. Other contributions to the development of OBE algorithms include the works of Deller *et al.* [10]–[12], [19], Norton *et al.* [20], [21], and Walter *et al.* [22]. To present a general OBE algorithm, we adopt the formulation given by Deller *et al.* [10]. Assume that at time instant $n-1$, the exact membership set ψ_{n-1} is outer bounded by the ellipsoid E_{n-1} described by[1]

$$E_{n-1} = \left\{\theta \in \mathbf{C}^N\colon [\theta - \theta_{n-1}]^H P_{n-1}^{-1} [\theta - \theta_{n-1}] \leq \sigma_{n-1}^2\right\} \tag{5}$$

where

- P_{n-1}^{-1} positive definite matrix that describes the shape, orientation, and size of the E_{n-1};
- σ_{n-1}^2 positive number that, together with P_{n-1}, defines the size of E_{n-1};
- θ_{n-1} geometric center of the ellipsoid.

An ellipsoid E_n that contains the intersection of E_{n-1} and $\mathcal{O}_n$ is given by a linear combination of (3) and (5) [9]

$$\begin{aligned} E_n = \{\theta \in \mathbf{C}^N\colon &\alpha_n [\theta - \theta_{n-1}]^H P_{n-1}^{-1} [\theta - \theta_{n-1}] \\ &+ \beta_n |y_n - \theta^T \mathbf{x}_n|^2 \leq \alpha_n \sigma_{n-1}^2 + \beta_n \gamma^2\} \end{aligned} \tag{6}$$

where $\alpha_n > 0$ and $\beta_n \geq 0$ are parameters to be chosen to optimize a measure of the size of E_n. Different optimality criteria lead to different OBE algorithms. It is straightforward to show that E_n describes an ellipsoid, and

$$E_n = \{\theta \in \mathbf{C}^N\colon [\theta - \theta_n]^H P_n^{-1} [\theta - \theta_n] \leq \sigma_n^2\} \tag{7}$$

[1] Throughout this paper, $\mathbf{x}^*$, $\mathbf{x}^T$, $\mathbf{x}^H$, and $\|\mathbf{x}\|$ denote the complex conjugate, transpose, complex conjugate transpose, and the two-norm, respectively, of the vector $\mathbf{x}$.

where P_n is positive definite, and θ_n is the geometric center of E_n. The D–H/OBE algorithm uses

$$\alpha_n = 1 - \lambda_n; \quad \beta_n = \lambda_n$$

where $0 \le \lambda_n < 1$ and defines λ_n^o, which is the optimal value of λ_n, as the one that minimizes σ_n^2. It has been shown in [9] and [23] that σ_n^2 is an upper bound on a constant multiple of the estimation error, assuming a persistency of excitation condition. An update is not required if the optimal values of α_n and β_n are $\alpha_n^o = 1$, and $\beta_n^o = 0$. The equivalent condition in the case of D–H/OBE is $\lambda_n^o = 0$. The condition that checks if an update is required or not is called the *information evaluation* criterion. D–H/OBE is particularly attractive for the updator sharing scheme (U-SHAPE, see Section V) since its information evaluation criterion is of linear computational complexity. In contrast, most other OBE algorithms require quadratic computational complexity and thereby do not derive any significant computational advantage via updator sharing. However, approximate tests for information evaluation that reduce the computational complexity exist in the literature [11]. The update equations for the D–H/OBE algorithm for real data are given in [9]. By extending those results to complex data, we obtain the update equations at time instant n:
if $\sigma_{n-1}^2 + |\delta_n|^2 \le \gamma^2$, where $\delta_n = y_n - \theta_{n-1}^T \mathbf{x}_n$, then

$$\lambda_n^o = 0$$

and

$$\theta_n = \theta_{n-1}; \quad P_n = P_{n-1}; \quad \sigma_n^2 = \sigma_{n-1}^2 \tag{8}$$

else

$$\begin{aligned} P_n &= \frac{1}{1-\lambda_n}\left[P_{n-1} - \frac{\lambda_n P_{n-1}\mathbf{x}_n^* \mathbf{x}_n^T P_{n-1}}{1-\lambda_n+\lambda_n G_n}\right] \\ \theta_n &= \theta_{n-1} + \lambda_n P_n \mathbf{x}_n^* \delta_n \\ \sigma_n^2 &= (1-\lambda_n)\sigma_{n-1}^2 + \lambda_n \gamma^2 - \frac{\lambda_n(1-\lambda_n)|\delta_n|^2}{1-\lambda_n+\lambda_n G_n} \end{aligned}$$

where it is shown in (9) at the bottom of the page, and where $\lambda_{\max} \in (0, 1)$ is a design parameter. The above algorithm, like all other OBE algorithms, assumes knowledge of a noise bound γ. If reliable noise bounds are not available or if the noise bound is time varying, we can employ bound-tuning strategies like the ones proposed in [20] and [24]. Further, previous studies have shown that the OBE algorithms are quite robust with respect to occasional noise bound violations, although theoretically, the OBE algorithms can diverge because of bound violations. The case of Gaussian noise is considered in the following section.

III. Channel Estimation Using OBE Algorithms

Consider a discrete-time complex baseband equivalent model of a digital communication channel with two-dimensional (2-D) signaling (e.g., M-QAM, M-PSK for any M) comprised of everything from the input of the transmitting filter to the output of the sampler at the receiver. As shown in Fig. 1, the transmitted sequence is denoted by $\mathbf{a} = \{a_n\}_{n=-\infty}^{\infty}$, channel impulse response by $\{c_i\}_{i=-D}^{D}$, additive noise sequence by $\mathbf{v} = \{\nu_n\}_{n=-\infty}^{\infty}$, and channel output sequence by $\mathbf{u} = \{u_n\}_{n=-\infty}^{\infty}$. The equalizer will be considered in the next section. In this section, we describe the application of OBE algorithms to the estimation of $\{c_i\}$. The channel is described by an FIR model with additive noise

$$u_n = \sum_{i=-D}^{D} c_i a_{n-i} + \nu_n. \tag{10}$$

In vector notation, the above relation can be rewritten as

$$u_n = \mathbf{c}^T \phi_n + \nu_n \tag{11}$$

where $\mathbf{c} = [c_{-D} \cdots c_D]^T$, and $\phi_n = [a_{n+D} \cdots a_{n-D}]^T$. Note that an IIR filter formulation also leads to (11), where ϕ contains past outputs. Since (11) is in the same form as the SMI model (1), OBE algorithms find direct application in estimating the channel vector $\mathbf{c}$.

Additive noise in a communications system is usually modeled as a Gaussian process. Bounded noise is theoretically essential for the OBE algorithms, and even a single violation of the bound can potentially make the membership set empty. However, as discussed below, this does not pose a serious practical problem.

Assume that $\{\nu_n\}$ is an ergodic zero-mean Gaussian process with variance σ_ν^2. The probability that the noise violates the assumed bound is $2Q(\gamma/\sigma_\nu)$, where $Q(\cdot)$ is the Q function.[2] For instance, if $\gamma = 3\sigma_\nu$, model violations occur less than 0.3% of the time. It is theoretically possible for the true parameter to fall out of the ellipsoidal estimate due to these violations; however, repeated experiments by several researchers

[2] The Q function is defined as $Q(x) \triangleq \int_x^\infty (1/\sqrt{2\pi})e^{-z^2/2}\,dz$.

$$\begin{aligned} G_n &= \mathbf{x}_n^T P_{n-1} \mathbf{x}_n^* \\ \lambda_n &= \lambda_n^o = \min(\nu_n, \lambda_{\max}) \\ \nu_n &= \begin{cases} \lambda_{\max} & \text{if } \delta_n = 0 \\ \dfrac{(1-\mu_n)}{2} & \text{if } G_n = 1 \\ \dfrac{1}{1-G_n}\left[1 - \sqrt{\dfrac{G_n}{1+\mu_n(G_n-1)}}\right] & \text{if } 1+\mu_n(G_n-1) > 0 \\ \lambda_{\max} & \text{if } 1+\mu_n(G_n-1) \le 0 \end{cases} \\ \mu_n &= \frac{\gamma^2 - \sigma_{n-1}^2}{|\delta_n|^2} \end{aligned} \tag{9}$$

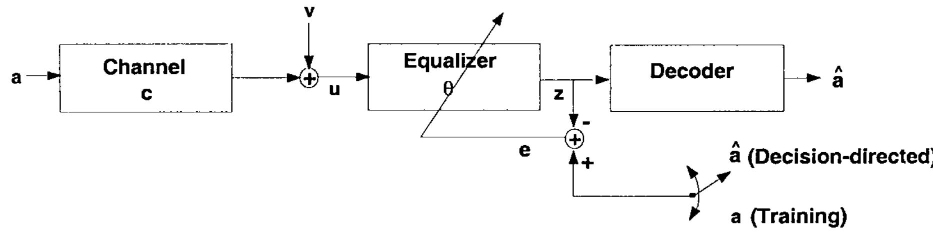


Fig. 1. Discrete-time model of a communication system with equalization.

have shown that the OBE algorithms do not diverge in the presence of infrequent model violations (e.g., when $\gamma = 3\sigma_\nu$). Moreover, infrequent model violations have been shown not to deteriorate the performance of these algorithms significantly in numerous simulations [9], [13], [24]. Such robustness may be attributed to the redundant region around the true membership set provided by ellipsoidal *outer* approximation involved in each updating instant. In the unlikely event that the ellipsoids lose track of the true parameter, an indication of loss of tracking, though not necessarily immediate, is provided by the sign of σ_n^2. Such an indication may be used to perform a simple rescue procedure to recapture the true parameter [13]. Further, once the ellipsoid has converged to the steady state, the corresponding point estimate will produce an output error whose magnitude is at most γ whenever the true noise is smaller than γ (since data consistent with the assumed model do not result in an update in the steady state). In the above example, the steady-state estimate will produce an output error smaller than the noise bound more than 99.7% of the time.

In the case of a time-varying channel, a state-space model can be used to represent the dynamics of the channel vector and the observation model (11) as

$$\mathbf{c}_{n+1} = F_n\mathbf{c}_n + \mathbf{d}_n \tag{12}$$

$$u_n = \phi_n^T\mathbf{c}_n + \nu_n \tag{13}$$

where F_n is a known sequence of matrices, and $\mathbf{d}_n$ is a disturbance known to belong to the ellipsoidal bound

$$\mathcal{D} = \{\mathbf{d}\colon \mathbf{d}^H Q^{-1}\mathbf{d} \le 1\}. \tag{14}$$

In many cases, the only available set-theoretic knowledge on the channel dynamics could be an upper bound γ_d on the magnitude of the channel vector jump at each time. In such cases, we can make F_n the identity matrix (or aI, where $a < 1$) and let $Q = (1/\gamma_d)I$ so that

$$\mathcal{D} = \{\mathbf{d}\colon \|\mathbf{d}\|^2 \le \gamma_d^2\}. \tag{15}$$

To estimate the channel vector using an OBE algorithm, let E_n be a known ellipsoid such that $\mathbf{c}_n \in E_n$. We need to compute the set of all vectors that result from transforming each point in E_n according to (12) for all $\mathbf{d}_n$ satisfying (14). This set, which is given by

$$F_nE_n \oplus \mathcal{D} = \{F_n\mathbf{c} + \mathbf{d}\colon \mathbf{c} \in E_n, \mathbf{d} \in \mathcal{D}\} \tag{16}$$

clearly contains $\mathbf{c}_{n+1}$ and is not an ellipsoid in general. Techniques have been proposed in [25]–[27] to outer bound the above set by an *a priori* (i.e., before using the observation at time $n+1$) ellipsoid $E_{n+1|n}$. The result proposed in [27] that optimizes a positive definite matrix added to P_n is

$$\begin{aligned}\mathbf{c}_{n+1|n} &= F_n\mathbf{c}_n\\ P_{n+1|n} &= P_n + \frac{Q}{\sigma_n^2} + 2\frac{\sqrt{\|F_nP_nF_n^T\|_2\|Q\|_2}}{\sigma_n}I_{2D+1}\\ \sigma_{n+1|n}^2 &= \sigma_n^2\end{aligned} \tag{17}$$

where I_{2D+1} is the identity matrix of size $2D+1$. A summary of techniques for state bounding by ellipsoidal algorithms is found in [28].

Since $\mathbf{c}_{n+1}$ belongs to the set in (16), which is, by construction, a subset of $E_{n+1|n}$, the true channel vector $\mathbf{c}_{n+1}$ belongs to the ellipsoid $E_{n+1|n}$. Since $\mathbf{c}_{n+1}$ also belongs to the observation-induced set at time $n+1$, $\mathcal{O}_{n+1}$, the SMI update equations can be used to find the new *a posteriori* ellipsoid E_{n+1} that outer bounds $E_{n+1|n} \cap \mathcal{O}_{n+1}$. A point estimate of the channel vector at time n is taken to be the geometric center $\mathbf{c}_n$ of the *a posteriori* ellipsoid E_n.

Simulation Results: In this section, we examine the performance of an OBE channel estimator *vis-á-vis* those using conventional techniques. For the time-invariant case, RLS (with an exponential weighting factor $\lambda = 0.98$) and LMS (with a step-size $\mu = 0.015$) algorithms are used for comparison, whereas a Kalman filter (KF) is used in the case of time-varying channels. The complex form of the D–H/OBE algorithm, which is described in Section II, is used for the OBE estimator. Extension to time-varying channel estimation is incorporated in the second example.

A randomly generated time-invariant complex-valued channel vector of length five and an output SNR of 15 dB is considered first. The additive noise is a realization of an AWGN process. The noise bound γ is chosen to be $3\sigma_\nu$, where σ_ν is the standard deviation of the noise. The mean-square error performances of D–H/OBE, RLS, and LMS estimators averaged over 1000 independent runs are illustrated in Fig. 2. The plots show comparable performances of the D–H/OBE and RLS algorithms, whereas LMS is much slower in convergence.

A time-varying mobile channel is simulated with the tap coefficients as the outputs of two-pole Butterworth lowpass filters with white noise inputs. The cut-off frequencies of the filters are taken to be the Doppler frequency shift in the mobile communication environment. We simulate the channel with a Doppler frequency of 100 Hz and a symbol rate of 25 000 symbols/s. Fig. 3 shows the tracking behavior of D–H/OBE and the Kalman filter algorithm for an SNR of 10 dB. For illustration, the figure shows just the real part of the second

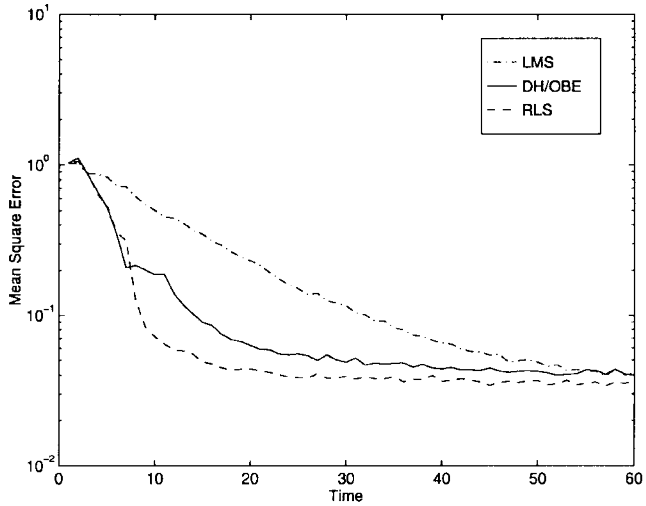


Fig. 2. MSE performance in the estimation of a time-invariant channel.

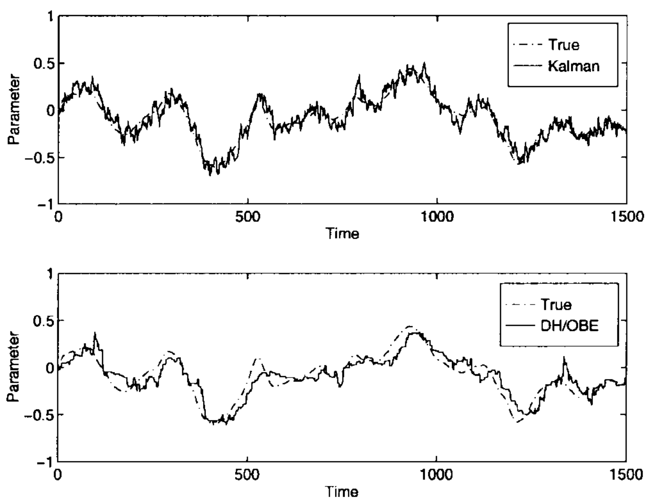


Fig. 3. Tracking of a mobile channel using D–H/OBE and RLS algorithms.

tap coefficient. Table I shows the mean-square prediction errors produced by the D–H/OBE algorithm and the KF for various values of SNR. The D–H/OBE performance is seen to be comparable with that of the KF, which is considered a benchmark for its fast convergence and tracking characteristics [3]. Moreover, as seen from the percentage of updates, this performance is achieved in spite of the infrequent updating by D–H/OBE.

IV. Channel Equalization in the Set-Membership Framework

A. Set-Membership Equalization

A novel concept of adaptive equalization in the SM framework is proposed here. The approach incorporates set-theoretic knowledge of the additive noise, which includes interference and thermal noise. This approach to equalizer design guarantees a specified upper bound on the Euclidean distance between the transmitted and equalized symbols for all inputs belonging to a certain set.

TABLE I
Mean-Square Error in Tracking a Mobile Channel with D–H/OBE and KF Algorithms

SNR (in dB)	0	5	10	15	20
MSE (D-H/OBE)	1.6309	0.5686	0.2455	0.1227	0.0672
MSE (KF)	1.2593	0.4496	0.1707	0.0697	0.0313
% updates (D-H/OBE)	8.76	26.15	36.75	47.19	58.80

Assume a one-dimensional (1-D) or 2-D signal constellation (e.g., BPSK, QPSK, QAM). Consider a linear-in-parameter equalizer with tap weights $\{\theta_i\}_{i=-N}^{N}$, whose output is given by

$$z_n = \theta^T \mathbf{x}_n \tag{18}$$

where $\theta = [\theta_{-N}, \cdots, \theta_N]^T$ is the parameter vector, and $\mathbf{x}_n$ is the input vector to the equalizer. In general, $\mathbf{x}_n$ is a function of the sequence of received samples $\mathbf{u}$ and the past outputs of the decoder $\{\hat{a}_i\}_{i=-\infty}^{n}$, which, in turn, are functions of $\mathbf{a}$ and $\mathbf{v}$. Therefore, the input vector to a linear-in-parameter equalizer is a function of the transmitted symbol sequence and the noise sequence. This category of equalizers includes linear and DFE's, with either fractionally spaced or symbol-spaced taps.

Let the transmitted symbols come from a 2-D constellation $\mathcal{C}$. In addition, assume that we want to design an equalizer that performs according to the desired specification whenever the additive noise ν_n belongs to a set $\mathcal{V}$. No performance requirement is specified for the equalizer when the noise is not from $\mathcal{V}$. For instance, $\mathcal{V}$ could be the set of noise samples of magnitude less than some γ_ν. Let the set of input-noise pairs for which the performance specification is defined be denoted as the *design space*

$$\mathcal{S} = \{\mathbf{a}\colon a_n \in \mathcal{C}\ \forall n\} \times \{\mathbf{v}\colon \nu_n \in \mathcal{V}\ \forall n\}. \tag{19}$$

If the probability that $\nu_n \in \mathcal{V}$ is high, then an equalizer that achieves the desired specification whenever $\nu_n \in \mathcal{V}$ will achieve the specification with a correspondingly high probability.

The output of the equalizer is a function of the input and noise sequences and is parameterized by θ as

$$z_n(\theta, \mathbf{a}, \mathbf{v}) = \theta^T \mathbf{x}_n(\mathbf{a}, \mathbf{v}). \tag{20}$$

The objective of SM equalization is to ensure that the maximum Euclidean distance between the transmitted and equalized symbols is less than a specified value $\gamma > 0$ whenever the input-noise pairs come from the design space, i.e.,

$$|a_n - z_n(\theta, \mathbf{a}, \mathbf{v})|^2 \le \gamma^2, \qquad \text{for all } (\mathbf{a}, \mathbf{v}) \in \mathcal{S}. \tag{21}$$

This is an instantaneous specification on the equalizer performance depending on the input.

If the probability that the input belongs to the design space is p, then this formulation ensures that the probability of meeting the specification is lower bounded by p. In terms of Fig. 1, the SM equalizer constrains the error e to be upper bounded in magnitude by γ for all inputs from $\mathcal{S}$. This is in contrast with the minimum mean-squares and LS equalizers, which are specified to minimize an ensemble or time-average

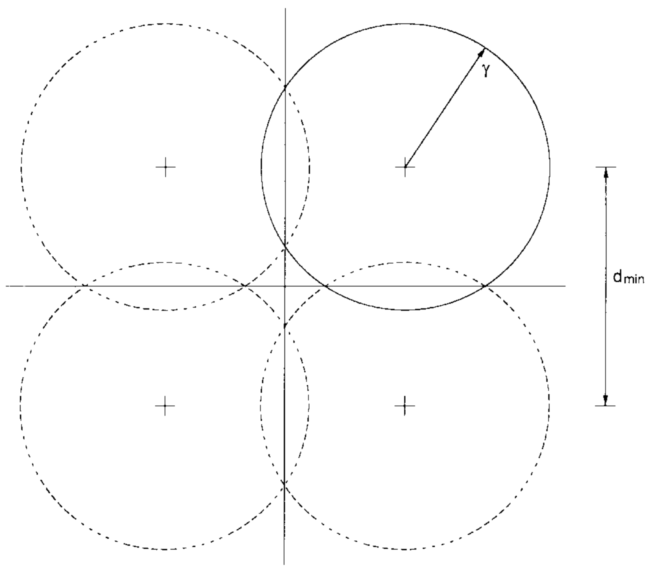


Fig. 4. Set-membership equalization criterion for a QPSK system. The equalized symbol is constrained to lie in a circle of radius γ centered at the transmitted symbol.

of the squared error e^2. Condition (21) requires that the equalized symbol remain in a circle of radius γ centered at the transmitted symbol, as shown in Fig. 4 for QPSK signaling. Note that (21) is independent of the time index n since the inequality must hold for all possible data and noise sequences in $\mathcal{S}$. If the minimum Euclidean distance between any two symbols in the signal constellation $\mathcal{C}$ is $d_{\min}$, then the specification for *error-free equalization* (i.e., no decision errors whenever the inputs are from the design space) is made by choosing $\gamma = d_{\min}/2$.

For a given specification of γ, define the set of all equalizer weight vectors of length $(2N+1)$ that satisfies (21) as the *feasible equalizer set* $\Theta_N(\gamma)$

$$\begin{aligned}\Theta_N(\gamma) &\triangleq \bigcap_{(\mathbf{a},\,\mathbf{v})\in\mathcal{S}} \{\theta \in \mathbf{C}^{2N+1}\colon |a_n - z_n(\theta,\,\mathbf{a},\mathbf{v})|^2 \le \gamma^2\} \\ &= \bigcap_{(\mathbf{a},\mathbf{v})\in\mathcal{S}} \{\theta \in \mathbf{C}^{2N+1}\colon |a_n - \theta^T \mathbf{x}_n(\mathbf{a},\mathbf{v})|^2 \le \gamma^2\}\end{aligned} \tag{22}$$

where $\mathcal{S}$ is the design space, as defined in (19).

The following observations on the feasible equalizer set $\Theta_N(\gamma)$ can be made.

1) $\Theta_N(\gamma)$ is a convex set. This follows from the fact that for each $(\mathbf{a},\,\mathbf{v})$, $\{\theta\colon |a_n - \theta^T\mathbf{x}_n(\mathbf{a},\,\mathbf{v})|^2 \le \gamma^2\}$ describes a convex hyperstrip in the θ space.

2) SM equalization involves estimation of a member of the set $\Theta_N(\gamma)$. Any point in the set yields a valid SM equalizer. Later in this section, we shall derive a procedure to estimate $\Theta_N(\gamma)$ that results in the same recursions as SMI algorithms.

3) $\Theta_N(\gamma)$ may be an empty set for a given N. This means that there exists no SM equalizer as defined above. In general, however, for a given signal constellation, channel, and noise model, it may be possible to find a nonempty $\Theta_N(\gamma)$ for a larger value of N since $\Theta_N(\gamma) \subseteq \Theta_{N+l}$ for all $l > 0$. Theorem 1 below provides a sufficient condition and a necessary condition for $\Theta_N(\gamma)$ to be nonempty in the case of linear equalization of FIR channels.

4) It follows trivially from (22) that $\Theta_N(\gamma) \subseteq \Theta_N(\gamma+\epsilon)$ for any $\epsilon > 0$. Therefore, the possibility of $\Theta_N(\gamma)$ being nonempty increases with γ.

A *linear equalizer* is a linear-in-parameters equalizer where the input vector is simply the regressor vector of the channel outputs. Continuing with the notation from Fig. 1, this implies that

$$\mathbf{x}_n = [u_{n+N}, \cdots, u_{n-N}]^T. \tag{23}$$

When the linear equalizer is considered in conjunction with the FIR channel model (10), we have

$$\mathbf{x}_n = C_N^T \mathbf{s}_n + \mathbf{w}_n \tag{24}$$

where

$$\begin{aligned}\mathbf{s}_n &= [a_{n+K},\,\cdots,\,a_{n-K}]^T;\quad K = D+N \\ \mathbf{w}_n &= [\nu_{n+N},\,\cdots,\,\nu_{n-N}]^T\end{aligned}$$

and the channel convolution matrix C_N is

$$C_N = \begin{pmatrix} c_{-D} & 0 & \cdots & & 0 \\ c_{-D+1} & c_{-D} & & & \\ \vdots & c_{-D+1} & \ddots & & \vdots \\ c_D & & \ddots & \ddots & 0 \\ 0 & \ddots & & \ddots & c_{-D} \\ \vdots & & \ddots & & c_{-D+1} \\ & & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & c_D \end{pmatrix}_{(2K+1)\times(2N+1)}.$$

The output of the equalizer is, therefore, given by

$$z_n = \theta^T(C_N^T \mathbf{s}_n + \mathbf{w}_n). \tag{25}$$

Since the desired output of the equalizer can be expressed as $a_n = \mathbf{e}_0^T \mathbf{s}_n$, where $\mathbf{e}_0$ is the unit vector of dimension $(2K+1)$ with a "1" in the $(K+1)$th position, the SM equalization criterion (21) can be rewritten as

$$|(C_N\theta - \mathbf{e_0})^T\mathbf{s} + \theta^T\mathbf{w}| \le \gamma, \qquad \forall\,(\mathbf{s},\mathbf{w}) \in \mathcal{H} \tag{26}$$

where the design space [space of the signal vector and noise vector pairs $(\mathbf{s},\,\mathbf{w})$ of interest] is denoted

$$\mathcal{H} = \mathcal{C}^{2K+1} \times \mathcal{V}^{2N+1}. \tag{27}$$

Note that the change in notation for the design space from $\mathcal{S}$ is due to the switch from infinite sequences $\mathbf{a}$ and $\mathbf{v}$ to finite dimensional vectors $\mathbf{s}$ and $\mathbf{w}$.

Consequently, the feasible equalizer set $\Theta_N(\gamma)$ is given by

$$\Theta_N(\gamma) = \bigcap_{(\mathbf{s},\mathbf{w})\in\mathcal{H}} \{\theta \in \mathbf{C}^{2N+1}\colon |(C_N\theta - \mathbf{e_0})^T\mathbf{s} + \theta^T\mathbf{w}|^2 \le \gamma^2\}. \tag{28}$$

Conditions for the existence of a nonempty parameter set $\Theta_N(\gamma)$, assuming an FIR channel and a linear equalizer, are stated in the following theorem.

Theorem 1—Existence of the Feasible Equalizer Set: Consider a linear channel model (10) and a linear equalizer as described above. If $\mathcal{V} = \{\nu \in \mathbf{C}\colon |\nu| \le \gamma_\nu\}$ and $\Theta_N(\gamma)$ is as defined in (28), then the following hold.

1) $\Theta_N(\gamma)$ is nonempty if the channel satisfies

$$\gamma_{\overline{a}} \sum_{i=-K}^{K} |\epsilon_i^{(1)}| + \gamma_\nu \sum_{i=-N}^{N} |\theta_i^{(1)}| \le \gamma \tag{29}$$

where $\epsilon^{(1)} = C_N\theta^{(1)} - \mathbf{e_0}$, $\gamma_{\overline{a}}$ is the maximum amplitude in the constellation $\mathcal{C}$, and $\theta^{(1)} = (C_N^H C_N + ((2N+1)\gamma_\nu^2/K\gamma_{\overline{a}}^2)I)^{-1}C_N^H\mathbf{e_0}$. If (29) is satisfied, then $\theta^{(1)} \in \Theta_N(\gamma)$.

2) If $\Theta_N(\gamma)$ is nonempty, then

$$\|\epsilon^{(2)}\|^2\gamma_{\underline{a}}^2 \cos^2\frac{\phi_0}{2} + \|\theta^{(2)}\|^2\gamma_\nu^2 \le \gamma^2 \tag{30}$$

where $\gamma_{\underline{a}}$ is the minimum amplitude in the constellation $\mathcal{C}$, and[3] $\phi_0 \triangleq \max_{a_1\in\mathcal{C}} \min_{a_2\in\mathcal{C}} |\phi(a_1, a_2)|$, $\epsilon^{(2)} = C_N\theta^{(2)} - \mathbf{e_0}$, and $\theta^{(2)} = (C_N^H C_N + (\gamma_\nu^2/\gamma_{\underline{a}}^2 \cos^2(\phi_0/2))I)^{-1}C_N^H\mathbf{e_0}$.

Proof: See Appendix A. ■

Remarks:

1) To gain intuition about the theorem, note that any θ belongs to the feasible set if and only if $J(\theta) \le \gamma$, where $J(\cdot)$ is a cost function as defined in Appendix A [$J(\cdot)$ is the supremum of (54) over $\mathcal{H}$]. In the theorem, $\theta^{(1)}$ is the minimum of a function that upper bounds $J(\cdot)$, and $\theta^{(2)}$ is the minimum of a function that lower bounds $J(\cdot)$.
2) Besides providing conditions for the existence of feasible equalizers, Theorem 1 provides a method to compute an SM equalizer nonrecursively when a channel estimate is available. When (29) is satisfied, $\theta^{(1)}$ is a closed-form solution for an SM equalizer since it belongs to the feasible set.

B. Illustrative Example

Consider a discrete-time additive noise FIR channel with coefficients [2, −1, 1] and a noise bound $\gamma_\nu = 0.2$. The transmitted symbols are QPSK modulated with unit magnitude, i.e., $\gamma_{\underline{a}} = \gamma_{\overline{a}} = 1$, and $\phi_0 = \pi/2$. The square of the LHS of the sufficient condition (29) (dash–dot line), the LHS of the necessary condition (30) (dashed line), and $(d_{\min}/2)^2$ (solid line) are shown in Fig. 5 for equalizer lengths $(2N+1)$. For $\gamma = d_{\min}/2$, the sufficient (necessary) condition is satisfied if the plot corresponding to the sufficient (necessary) condition is below the $(d_{\min}/2)^2$ line. Fig. 5 shows that *error-free* linear equalization is possible for this channel if the number of equalizer taps is seven or more. Simulations with the worst-case input and noise sequences have shown that $\theta^{(1)}$ does indeed produce no bit errors when the sufficient condition is met. The necessary condition is satisfied for linear equalizers of all lengths.

[3]The angle $\phi(a_1, a_2)$ between two complex numbers a_1 and a_2 is the solution of $e^{j\phi} = a_1^* a_2/|a_1||a_2|$, with $\phi \in (-\pi, \pi]$.

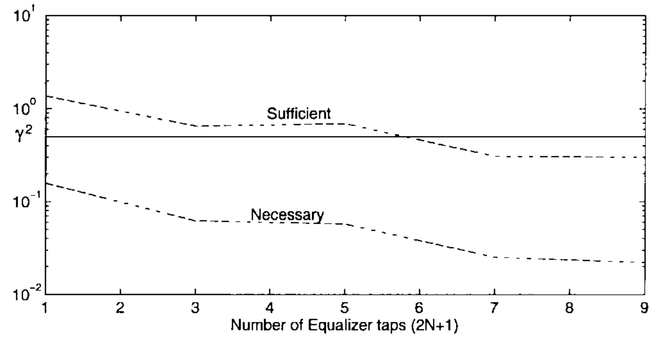


Fig. 5. Necessary and sufficient conditions for the existence of a feasible linear equalizer set for the channel [2, −1, 1] and $\gamma_\nu = 0.2$.

Remark: The intent of this simulation example is to illustrate the principle of SM equalization and may not correspond to a realistic channel. It shows clearly that SM equalizers can be designed nonadaptively (without the need for OBE algorithms) with *a priori* knowledge of the channel. In such a case, the resulting SM equalizer will achieve the error-bound specification for all input pairs from $\mathcal{H}$, and therefore, the probability that the output error is less than the specified bound is at least equal to the probability that the input pair belongs to $\mathcal{H}$.

C. A Recursive OBE Solution

We now show that SM recursions can be used to outer bound the *feasible equalizer set* $\Theta_N(\gamma)$. Note that the restriction to linear channel and equalizer models are not required in the sequel. An arbitrary channel and a linear-in-parameter equalizer are assumed. Further, assume that $\mathcal{V} = \{\nu\colon |\nu|^2 \le \gamma_\nu^2\}$ and that all training data come from $\mathcal{H}$. Any data that do not are considered model violations, and they are expected to occur with a small probability (see the discussion on unbounded noise in Section III). Following the notation of SM theory, define the *observation-induced set* at time n

$$\mathcal{O}_n \triangleq \{\theta\colon |a_n - \theta^T\mathbf{x}_n|^2 \le \gamma^2\} \tag{31}$$

where a_n is the transmitted symbol at time n. $\mathcal{O}_n$ is the set of all equalizer weights that equalize the channel output at time n, according to the SM equalization criterion. Analogous to the definition in Section II, we define *membership set* ψ_n as the set of equalizer weights that equalize all the channel outputs until time n

$$\psi_n \triangleq \bigcap_{k\le n} \mathcal{O}_k. \tag{32}$$

If the transmitted symbol sequence and the noise sequence satisfy the assumed set-theoretic model, then it follows from (21) and (31) that $\mathcal{O}_n$ contains $\Theta_N(\gamma)$ for every n, and, consequently,

$$\psi_n \supseteq \Theta_N(\gamma). \tag{33}$$

Since SM techniques estimate the membership set ψ_n with outer bounding approximations, we can use an OBE algorithm to compute ellipsoids E_n such that

$$E_n \supseteq \psi_n \supseteq \Theta_N(\gamma), \qquad \forall n. \tag{34}$$

A few qualitative remarks about the above procedure are given below. They provide intuitive justification for the proposed scheme and are stated without proof.

Remarks:

1) SM equalization does not require the noise to be bounded but only that the data for which the bounded-error specification must be met come from bounded noise. The off-line procedure for the design of an SM equalizer therefore poses no theoretical problems in handling Gaussian noise. However, OBE adaptation requires all the training data to come from the design space, and model violations can potentially result in loss of tracking. Please refer to the discussion following (11) for possible solutions.
2) From (22), (31), and (32), we can see that the feasible equalizer set $\Theta_N(\gamma)$ is an ensemble intersection, and the membership set ψ_n is a time intersection of the same sets over the transmitted symbols and noise. Therefore, if the transmitted symbols and noise samples are "persistently exciting," we can expect the time intersection ψ_n to approach the ensemble intersection $\Theta_N(\gamma)$ in some meaningful sense. This provides a qualitative assurance that the optimal bounding ellipsoids are "tight" since they *tightly* outer bound the membership set.
3) The geometric center of the bounding ellipsoid is taken to be the point estimate for the purpose of equalization. If the bounding ellipsoid is *asymptotically tight*, then the convexity of the desired parameter set makes it likely for the geometric center to asymptotically lie inside the *feasible equalizer set*.
4) The proposed algorithm inherits the discerning update feature of the OBE method and, therefore, allows sharing of the updating processors, as described in the following section.

Simulation Results: To test the performance of the OBE adaptive equalizer, a raised cosine ISI channel [3, p. 414] is considered. The channel impulse response is $\{0.75, 1.0, 0.75\}$. A DFE with nine forward and one feedback taps is used. The OBE equalizer uses $\gamma^2 = 9\sigma_\nu^2$, and RLS uses exponential weighting factor 0.99. Fig. 6 shows the bit error rate (BER) performance of both the equalizers after 500 bits of training.

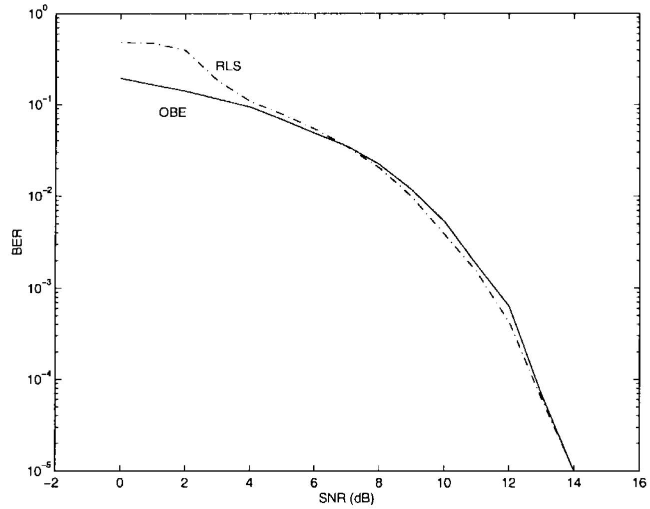


Fig. 6. BER of OBE and RLS equalizers for the channel [0.75, 1.0, 0.75].

V. Updator-Shared Adaptive Parallel Equalization (U-SHAPE)

Consider the problem of simultaneously equalizing M channels. If the probability that more than a certain number of channels request an update at the same time is small, fewer than M updating units (processors) can be shared among the M channels. We shall refer to such a sharing scheme U-SHAPE. Fig. 7 shows a schematic of the proposed system when decision feedback equalizers are employed. Note that all results in this section are valid whenever several independent systems are to be identified using OBE techniques in parallel, including the case of multiple-channel estimation and linear equalization. In the implementation of U-SHAPE, if the number of simultaneous update requests is larger than the number of updating units L, some criterion must be established to decide which L of them to service. A straightforward decision strategy is to choose the L *winning* channels randomly (with equal probability) from the contending channels, but it will be shown in the following that this is not the best approach when the probabilities of update requests on all the channels are unequal. The design problem in U-SHAPE is twofold: to determine the "optimal" number of updating processors required to meet certain specifications and to devise a scheme for contention resolution that would provide uniform quality of service to all users. This section addresses these design issues for U-SHAPE based on the results in [15] and [16].

Assume the channel update requests are stationary, memoryless, independent random processes. At any sample instant, define a binary random variable Q_i to be 1 when the ith channel requests an update and 0 otherwise. Let $p_i \triangleq P_{Q_i}(1)$, which is the probability that the ith channel requests an update. Define another binary random variable U_i to be 1 when an update is performed on the ith channel and 0 otherwise. Let N_Q denote the total number of update requests at any time so that

$$N_Q \triangleq \sum_{i=1}^{M} Q_i. \tag{35}$$

Quality of service provided to the ith user (channel) can be measured in terms of *rejection rate* ρ_i, which is defined as the probability of not updating the channel given an update request, i.e.,

$$\rho_i \triangleq P_{U_i|Q_i}(0|1). \tag{36}$$

A. Solution for Equal Update Request Probabilities

Consider the case of equal probabilities of update requests on all channels, i.e., $p_1 = p_2 = \cdots = p_M = p$. In addition, if $N_Q > L$, assume that $N_Q - L$ update requests are rejected at random with equal probabilities of rejection. This results in $\rho_1 = \rho_2 = \cdots = \rho_M$, and we have the following theorem.

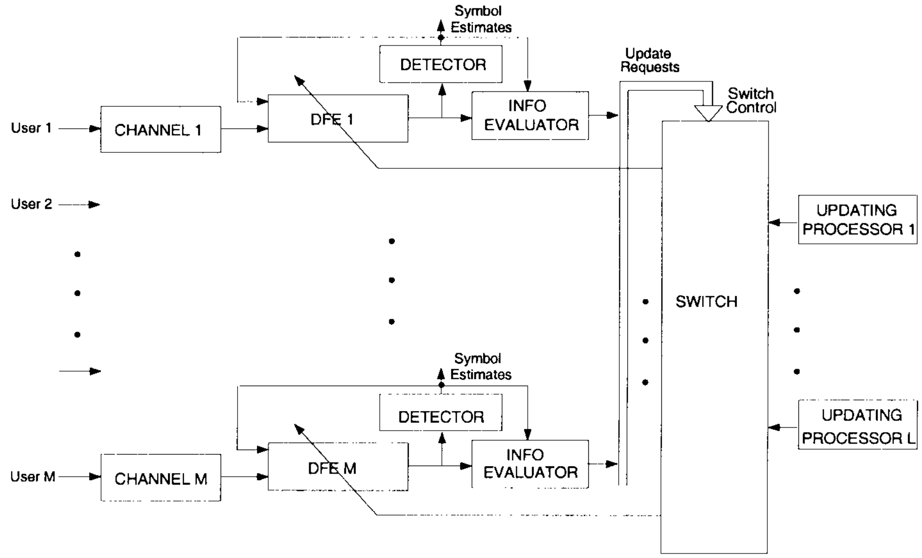


Fig. 7. Updator-shared parallel adaptive equalization (U-SHAPE) scheme using decision feedback equalizers.

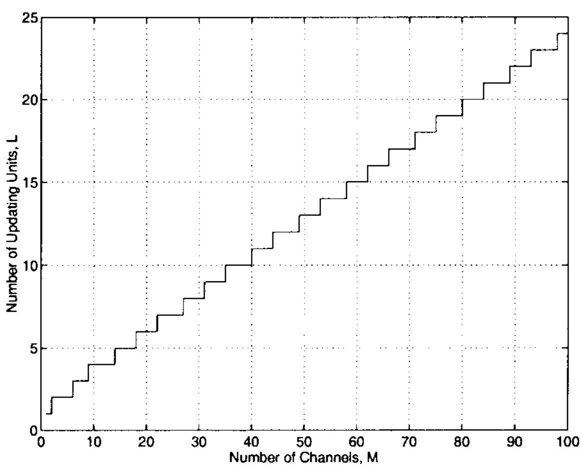


Fig. 8. Number of updators when $p = 0.3$ and $\rho_0 = 0.1$.

Theorem 2: Under the assumptions above, the optimum number of updating processors that results in $\rho_i \le \rho_0$ for every i is the smallest number L of processors that satisfies

$$\rho_i = \sum_{k=L+1}^{M} \frac{k-L}{k} \binom{M-1}{k-1} p^{k-1}(1-p)^{M-k} \le \rho_0 \quad \text{for any } i. \tag{37}$$

Proof: See Appendix B. ■

Since we know that $1 \le L \le M$, the minimum value of L that satisfies (37) can be found by searching in $\{1, 2, \cdots, M\}$. The relation between L and M when $\rho_0 = 0.1$ and $p = 0.3$ is shown in Fig. 8. Simulation results have shown that the variation of ρ_0 from 0 to 0.1 has an insignificant effect on the BER. It is clear that this scheme provides a significant reduction in the number of updating processors required.

The contention resolution scheme used above is to admit L out of the N_Q update requests with equal probability of admitting (or, equivalently, rejecting) any of the requesting channels. This scheme is referred to here as the *random rejection scheme*. Although such a scheme lends itself to a simple mathematical formulation, it fails when the probabilities of update requests are not the same on all the channels. To illustrate this, consider a two-channel case with one updating processor. Using the random rejection scheme, either channel will be admitted with a probability $1/2$ whenever there is a tie. Then

$$\begin{aligned}\rho_1 =& P_{U_1|Q_1,N_Q}(0|1,\,1)P_{N_Q|Q_1}(1|1)\\ &+ P_{U_1|Q_1,\,N_Q}(0|1,\,2)P_{N_Q|Q_1}(2|1) = \frac{p_2}{2}.\end{aligned} \tag{38}$$

Similarly, $\rho_2 = p_1/2$. Thus, if p_1 is close to 1 and p_2 is close to 0, then channel 1 would be admitted almost every time it requests an update, whereas almost half of channel 2's requests would be rejected. The case of an arbitrary update request distribution is considered in the next subsection.

B. Solution for General Update Request Probabilities

In this section, we remove the constraint that all p_i's are equal. Define a binary random variable R_i for each channel to indicate an update request rejection, i.e.,

$$R_i \triangleq \begin{cases} 1, & \text{if } Q_i = 1 \text{ and } U_i = 0 \\ 0, & \text{otherwise.} \end{cases} \tag{39}$$

Let N_R denote the number of rejected requests at any time. Then

$$N_R = \sum_{i=1}^{M} R_i. \tag{40}$$

In order to remove the dependence of the solution on any particular contention resolution scheme, the performance constraint is modified from a bound on the rejection rates to a bound on the ratio of the average number of rejected channels to the average number of update requests, i.e., define

$$\rho \triangleq \frac{\overline{N}_R}{\overline{N}_Q} \tag{41}$$

where $\overline{N}_R \triangleq E\{N_R\}$ and $\overline{N}_Q \triangleq E\{N_Q\}$, and specify

$$\rho < \rho_0. \tag{42}$$

Note that if there exists a contention resolution scheme that achieves $\rho_1 = \rho_2 = \cdots = \rho_M$, then $\overline{N}_R/\overline{N}_Q = \rho_i$ for every i. Therefore, the above specification reduces to $\rho_i \le \rho_0$, which is identical to the specification in the previous section. Contention resolution schemes are studied in Section V-D. The following theorem provides the solution for the optimum number of updators when the update request probabilities are not (necessarily) equal.

Theorem 3: The optimum number of updating processors required to guarantee (42) is the minimum value of L that satisfies

$$\sum_{k=L+1}^{M} (k-L)P_{N_Q}(k) \le \rho_0 \sum_{i=1}^{M} p_i \tag{43}$$

where

$$P_{N_Q}(k) = \sum_{i_1=1}^{M-k+1} \sum_{i_2=i_1+1}^{M-k+2} \cdots \sum_{i_k=i_{k-1}+1}^{M} \cdot \left(p_{i_1} p_{i_2} \cdots p_{i_k} \prod_{j \in \{1, \cdots, M\} \setminus \{i_1, \cdots, i_k\}} (1-p_j) \right).$$

The rejection ratio is given by

$$\rho = \frac{\sum_{k=L+1}^{M} (k-L)P_{N_Q}(k)}{\sum_{i=1}^{M} p_i}. \tag{44}$$

Proof: See Appendix C. ■

The minimum value of L that satisfies (43) can be determined as before by an off-line search in $\{1, 2, \cdots, M\}$. It can be easily verified that (43) coincides with (37) under the assumption of equal update request probabilities.

C. Design of U-SHAPE Using a Queuing Model

The preceding design procedure assumes time-invariant probabilities of update requests and a fixed number of active users. However, in practice, the number of active users varies. Further, SM equalizers usually require more frequent updating at the beginning of a call. Consequently, the probability of update request on any channel is large at the beginning of a call and then decays to a steady-state value. Since the number of updating processors is fixed, these variations result in a stochastic rejection rate. Computation of the number of updating processors L must be altered to satisfy a statistical specification on the rejection rate. Instead of an upper bound on the rejection rate, the design criterion can be a confidence interval, i.e., for given ρ_0 and ϵ, the design specification is

$$P[\rho > \rho_0] < \epsilon. \tag{45}$$

Computation of $P[\rho > \rho_0]$ as a function of L requires a model for the time variations of the number of active users M and the p_i's. We employ a queuing model for the telephone network and model the decay of p_i's by an exponential function. The telephone network is modeled as an m-server loss queuing system [29]. In this model, the customer population is assumed to be infinite, and the number of available telephone channels (i.e., the number of servers) is denoted by N. The average rate of incoming calls is λ calls per second, and the average service rate for any call is μ calls per second. Any call that arrives when all channels are busy is rejected. The call interarrival time and the service time for each call are exponentially distributed with parameters λ and μ, respectively. Note that λ and μ, as used here, are different from the notation in the discussion on OBE algorithms in Section II. For this system, the probability of M active users is given by

$$P_M(m) = \frac{\left(\frac{\lambda}{\mu}\right)^m \Big/ m!}{\sum_{k=0}^{N} \left(\frac{\lambda}{\mu}\right)^k \Big/ k!} \qquad m = 0, 1, \cdots, N. \tag{46}$$

The average number of active users is

$$\overline{M} = E\{M\} = \frac{\lambda}{\mu}(1 - P_M(N)) \approx \frac{\lambda}{\mu}. \tag{47}$$

If the number of updating processors L is computed using $M = M_0$ and $\rho = \rho_0$ in (37) or (43), then, neglecting the effect of variations in p_i's, we get

$$P[\rho > \rho_0] = \sum_{m=M_0+1}^{N} P_M(m) \tag{48}$$

and

$$E\{\rho\} = \sum_{m=0}^{N} \rho(m) P_M(m) \tag{49}$$

where $\rho(m)$ is the rejection rate when the number of active users is m.

Variation in the probabilities of update requests is modeled as an exponential decay with an initial value of $p_{init} = 1$ to a steady-state probability p_{ss}, i.e.,

$$p_i(t) = (p_{init} - p_{ss})e^{-t/\tau} + p_{ss} \qquad i = 0, 1, \cdots, N$$

where time t is normalized by the symbol interval, and τ is in number of symbols.

A worst-case estimate of the set of parameters for the queuing system and the exponential model for p_i's is

$$\begin{aligned} &N = 100 \;\; \lambda = 1/180 \text{ s}^{-1} \;\; \frac{\lambda}{\mu} = 50 \;\; \rho_o = 0.1 \\ &p_{init} = 1.0 \;\; p_{ss} = 0.3 \;\; \tau = 100. \end{aligned} \tag{50}$$

Observation of the transient behavior of the D–H/OBE algorithm in nonstationary environments has shown that the above choice of τ and p_{ss} represents a worst-case scenario. Even if we assume a symbol rate as small as 2400 symbols/s, the average interarrival time for incoming calls is approximately equal to 2.4×10^5 symbols. Since τ is insignificant compared with this number, we can expect the variation in p_i's to have little impact on the effective rejection rate. This conjecture is supported by simulation results with the system parameters as in (50), and $M_0 = \overline{M} = 50$. From (37), we get $L = 15$. Using

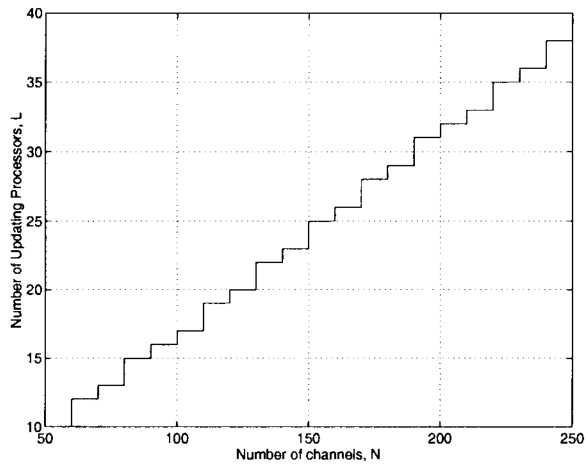


Fig. 9. Number of updators when $p = 0.3$ and $\rho_0 = 0.1$, using the queuing model.

(49), $E\{\rho\} = 0.0941$, whereas the observed $\rho = 0.0964$. Since (49) is derived without considering the effect of variations in p_i's, it is clear that the exponential decay of the update request probabilities does not significantly affect the overall rejection rate. Therefore, L can be computed assuming steady-state update request probabilities without loss of accuracy.

Given ρ_0 and ϵ in (45), the number of updating processors L is computed as follows.

1) Using (48), find M_0, which is the minimum value of M that satisfies (45).
2) Compute L using $\rho = \rho_0$ and $M = M_0$ in (37) or (43).

Example: For $\epsilon = 0.1$, $\rho_0 = 0.1$, and the above system parameters, the design procedure yields $M_0 = 59$ and $L = 17$. The simulation included 50 runs, each for more than 4×10^8 symbols, with a symbol rate of 2400 symbols/s. Each run was initialized with the ending network status (number of active calls, call-lengths, etc.) of the previous run. The average rejection rate was observed to be

$$E\{\rho\}_{obs} = 0.054 \ \ (E\{\rho\}_{theory} = 0.046).$$

The fraction of times that $\rho > \rho_0 = 0.1$ was

$$P[\rho > \rho_0]_{obs} = 0.109 \ \ (P[\rho > \rho_0]_{theory} = 0.100).$$

Simulations show very good agreement with the theoretical predictions made after neglecting the effect of varying update-request probabilities. Moreover, simulation results have shown that small variations in the rejection rate do not adversely affect the final BER of the equalizers. This example shows that a system with 100 channels needs just 17 updating processors to perform with negligible deterioration.

Fig. 9 shows the relation between the number of updating processors L and the number of channels N when the expected number of active calls $\overline{M} = N/2$, with all other parameters as above. The plot clearly shows that the above design procedure results in a significant reduction in hardware complexity.

D. Contention Resolution Schemes

As we have seen earlier, the *random rejection scheme,* in which all requesting channels have equal probability of being rejected, is suboptimal, except in the special case when the probabilities of update requests are equal on all the channels. In the general case, we need to determine which L of the requesting channels require updates more urgently than the others. This would result in a *priority rejection scheme* in which some measure of priority would be used to determine the priority of each channel so that a higher priority channel is serviced in case of a tie.

In the D–H/OBE algorithm, the quantity $\Delta_k \triangleq \delta_k^2 + \sigma_{k-1}^2 - \gamma^2$ is used as a measure of the information content in the received data set. In particular, an update is requested if and only if $\Delta_k > 0$. Therefore, Δ_k is an intuitively satisfying measure for the *priority rejection scheme*.

The rejection ratios of all the channels can be made equal if the priority of each channel is measured in terms of estimates of their rejection rates. To do this, define *rejection fraction* $\tilde{\rho}_i$ of the ith channel as the fraction of its update requests that have been rejected until that time, i.e., $\tilde{\rho}_i(t) \triangleq \langle R_i \rangle_t / \langle Q_i \rangle_t$, where $\langle \cdot \rangle_t$ denotes time average until time t. Assuming ergodicity, $\tilde{\rho}_i(t)$ equals the rejection ratio in the limit as $t \to \infty$. A channel with a higher rejection fraction would win in a tie to bring the rejection fractions closer. Estimates of $\langle R_i \rangle_t$ and $\langle Q_i \rangle_t$ can be obtained as

$$r_i(t) = \zeta r_i(t-1) + (1-\zeta) R_i(t) \tag{51}$$

$$q_i(t) = \zeta q_i(t-1) + (1-\zeta) Q_i(t) \tag{52}$$

where $0 < \zeta < 1$ is a forgetting factor. Then, $\tilde{\rho}(t) \approx r_i(t)/q_i(t)$. Therefore, each channel can be prioritized by the fraction $\tilde{\rho}(t)$. Additional time-dependent terms could be added in the equations to assign higher priority to equalizers that have started operating more recently.

VI. Summary and Conclusions

The excellent convergence and tracking characteristics of SM algorithms and the computational advantage provided by their discerning update feature have inspired the development of SM equalization in this paper. The close resemblance of the OBE update equations to those of LS estimators make them attractive practical tools. SM algorithms and schemes that utilize their selective updating feature can be very useful in applications to adaptive filtering in communications, where the requirement is for high performance algorithms with low computational complexity.

SMI methodology has been applied to channel identification, showing comparable convergence performance and better tracking performance *vis-á-vis* the LS techniques at a fraction of the computational burden. The notion of SM equalization has been introduced, and the issues of the existence and design of SM equalizers have been addressed. OBE algorithms have been shown to be attractive tools for adaptive SM equalization. An updator-shared implementation of parallel OBE equalizers/channel estimators called U-SHAPE has been proposed to exploit the selective updating feature of OBE algorithms. Solutions have been derived for the minimization of the number of updating processors subject to a specified bounds on quality of service to all users. U-SHAPE has been shown to offer a large reduction in the number of updating processors. For instance, a 100-channel system only needs 17 updating processors to operate with negligible deterioration in performance. Different strategies have been suggested to handle update contentions.

APPENDIX A
PROOF OF THEOREM 1

Proof: From the definition of the feasible equalizer set (22), a necessary and sufficient condition for $\Theta_N(\gamma)$ to be nonempty is

$$J^o \triangleq \inf_{\theta} \sup_{(\mathbf{s},\mathbf{w})\in\mathcal{H}} J(\theta, \mathbf{s}, \mathbf{w}) \le \gamma \tag{53}$$

where

$$J(\theta, \mathbf{s}, \mathbf{w}) \triangleq |(C_N\theta - \mathbf{e_0})^T\mathbf{s} + \theta^T\mathbf{w}|. \tag{54}$$

The filtered noise magnitude $|\theta^T\mathbf{w}|$ is maximized when $|w_i| = \gamma_\nu \ \forall i$ and $\angle(w_i) = \phi - \angle(\theta_i), \ \forall i$, for any angle ϕ. Given any θ and s, $J(\theta, \mathbf{s}, \mathbf{w})$ is maximized when ϕ is chosen such that $\angle(C_N\theta - \mathbf{e_0})^T\mathbf{s} = \angle(\theta^T\mathbf{w})$. Therefore, in (53)

$$\begin{aligned}\sup_{(\mathbf{s},\mathbf{w})\in\mathcal{H}} & J(\theta, \mathbf{s}, \mathbf{w}) \\ &= \sup_{\mathbf{s}\in\mathcal{C}^{2K+1}} |(C_N\theta - \mathbf{e_0})^T\mathbf{s}| + (|\theta_{-N}| + \cdots + |\theta_N|)\gamma_\nu. \end{aligned}\tag{55}$$

Defining $\gamma_{\overline{a}} \triangleq \max_{a\in\mathcal{C}} |a|$, we also have

$$|(C_N\theta - \mathbf{e_0})^T\mathbf{s}| \le (|\epsilon_{-K}| + \cdots + |\epsilon_K|)\gamma_{\overline{a}} \tag{56}$$

where $\epsilon = C_N\theta - \mathbf{e_0}$. Combining this with (53) and (55) yields

$$J^o \le \inf_\theta \{(|\epsilon_{-K}| + \cdots + |\epsilon_K|)\gamma_{\overline{a}} + (|\theta_{-N}| + \cdots + |\theta_N|)\gamma_\nu\}. \tag{57}$$

Therefore, $\Theta_N(\gamma)$ is nonempty if there exists a $\theta^{(1)}$ such that

$$(|\epsilon_{-K}^{(1)}| + \cdots + |\epsilon_K^{(1)}|)\gamma_{\overline{a}} + (|\theta_{-N}^{(1)}| + \cdots + |\theta_N^{(1)}|)\gamma_\nu \le \gamma \tag{58}$$

where $\epsilon^{(1)} = C_N\theta^{(1)} - \mathbf{e_0}$.

Since

$$\begin{aligned} J^2(\theta, \mathbf{s}, \mathbf{w}) &\le 2(|(C_N\theta - \mathbf{e_0})^T\mathbf{s}|^2 + |\theta^T\mathbf{w}|^2) \\ &\le 2(\|C_N\theta - \mathbf{e_0}\|^2 K\gamma_{\overline{a}}^2 + \|\theta\|^2(2N+1)\gamma_\nu^2) \end{aligned}\tag{59}$$

choose $\theta^{(1)}$ to be the LS solution of the RHS of the above inequality to obtain the sufficient condition (29).

For a necessary condition, we need to lower bound $\sup_{(\mathbf{s},\mathbf{w})\in\mathcal{H}} J(\theta, \mathbf{s}, \mathbf{w})$. Letting $\epsilon = C_N\theta - \mathbf{e_0}$, we have

$$|(C_N\theta - \mathbf{e_0})^T\mathbf{s}| = \left|\sum_{i=-K}^{K} \epsilon_i s_i\right|. \tag{60}$$

Choose each a_i to minimize the angular distance of each term in the RHS above to the real axis, i.e., let

$$a_i = \arg\min_{a\in\mathcal{C}} |\phi(a, \epsilon_i^*)|. \tag{61}$$

Defining

$$\gamma_{\underline{a}} \triangleq \min_{a\in\mathcal{C}} |a| \tag{62}$$

$$\phi_0 \triangleq \max_{a_1\in\mathcal{C}} \min_{a_2\in\mathcal{C}} |\phi(a_1, a_2)| \tag{63}$$

the choice (61) results in

$$|(C_N\theta - \mathbf{e_0})^T\mathbf{s}| \ge (|\epsilon_{-K}| + \cdots + |\epsilon_K|)\gamma_{\underline{a}} \cos\frac{\phi_0}{2}. \tag{64}$$

Equations (53), (55), and (64) imply

$$\begin{aligned} J^o &\ge \inf_\theta (|\epsilon_{-K}| + \cdots + |\epsilon_K|)\gamma_{\underline{a}} \cos\frac{\phi_0}{2} \\ &\quad + (|\theta_{-N}^{(2)}| + \cdots + |\theta_N^{(2)}|)\gamma_\nu \\ &\ge \inf_\theta \|\epsilon\|\gamma_{\underline{a}} \cos\frac{\phi_0}{2} + \|\theta\|\gamma_\nu \\ &\ge \inf_\theta \sqrt{\|\epsilon\|^2\gamma_{\underline{a}}^2 \cos^2\frac{\phi_0}{2} + \|\theta\|^2\gamma_\nu^2} \\ &= \sqrt{\|\epsilon^{(2)}\|^2\gamma_{\underline{a}}^2 \cos^2\frac{\phi_0}{2} + \|\theta^{(2)}\|^2\gamma_\nu^2} \end{aligned}\tag{65}$$

where $\epsilon^{(2)} = C_N\theta^{(2)} - \mathbf{e_0}$, and $\theta^{(2)} = (C_N^H C_N + (\gamma_\nu^2/\gamma_{\underline{a}}^2 \cos^2(\phi_0/2))I)^{-1}C_N^H\mathbf{e_0}$. The necessary condition (30) follows from the above. ■

APPENDIX B
PROOF OF THEOREM 2

Proof: The probability mass function of N_Q, which is denoted by $P_{N_Q}(k)$, follows a binomial distribution given by

$$P_{N_Q}(k) = \binom{M}{k} p^k (1-p)^{M-k} \tag{66}$$

and

$$P_{N_Q|Q_i}(k|1) = \binom{M-1}{k-1} p^{k-1}(1-p)^{M-k}. \tag{67}$$

Therefore,

$$\begin{aligned} \rho_i &= \sum_{k=L+1}^{M} P_{U_i|Q_i, N_Q}(0|1, k) P_{N_Q|Q_i}(k|1) \\ &= \sum_{k=L+1}^{M} P_{U_i|Q_i, N_Q}(0|1, k) \binom{M-1}{k-1} p^{k-1}(1-p)^{M-k}. \end{aligned}\tag{68}$$

If $k - L$ requests are rejected at random with equal probability, the probability that the ith channel is one of the rejected channels is

$$P_{U_i|Q_i, N_Q}(0|1, k) = \frac{k-L}{k}. \tag{69}$$

Putting (69) in (68) and bounding ρ_i's by a specified ρ_0, the desired result follows. ■

APPENDIX C
PROOF OF THEOREM 3

Proof: The probability mass function of R_i is given by

$$P_{R_i}(1) = P_{R_i|Q_i}(1|0)P_{Q_i}(0) + P_{R_i|Q_i}(1|1)P_{Q_i}(1) = \rho_i p_i. \tag{70}$$

Therefore, the expected number of rejections is

$$E\{N_R\} = \sum_{i=1}^{M} E\{R_i\} = \sum_{i=1}^{M} P_{R_i}(1) \tag{71}$$

or

$$\overline{N}_R = \sum_{i=1}^{M} \rho_i p_i. \tag{72}$$

From (35), we also have

$$\overline{N}_Q \triangleq E\{N_Q\} = \sum_{i=1}^{M} E\{Q_i\} \tag{73}$$

or

$$\overline{N}_Q = \sum_{i=1}^{M} p_i. \tag{74}$$

Now, putting (74) in (43), we obtain

$$\overline{N}_R \leq \rho \sum_{i=1}^{M} p_i. \tag{75}$$

Since $N_R = N_Q - L$ if $N_Q > L$ and 0 otherwise, the expected number of rejected requests can be expressed in terms of $P_{N_Q}(k)$ as

$$\overline{N}_R = \sum_{k=L+1}^{M} (k - L) P_{N_Q}(k). \tag{76}$$

The required inequality follows from (75) and (76). ■

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers for their careful review and extremely valuable suggestions. They also wish to express their heart-felt appreciation to Prof. J. R. Deller for numerous valuable comments.

REFERENCES

[1] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.

[2] ———, "Adaptive equalization for TDMA digital mobile radio," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 333–341, May 1991.

[3] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[4] R. E. Lawrence and H. Kaufman, "The Kalman filter for the equalization of a digital communication channel," *IEEE Trans. Commun. Technol.*, vol. CT-19, pp. 1137–1141, 1971.

[5] S. McLaughlin, "Adaptive equalization via Kalman filtering techniques," in *Proc. Inst. Elect. Eng.*, vol. 138, no. 4, pt. F, pp. 388–396, Aug. 1991.

[6] D. D. Falconer, M. Abdulrahman, N. W. K. Lo, B. R. Anderson, and A. U. H. Sheikh, "Advances in equalization and diversity for portable wireless systems," *Digital Signal Process.*, vol. 3, pp. 148–162, 1993.

[7] N. W. K. Lo *et al.*, "Adaptive equalization for co-channel interference in a multipath fading environment," *IEEE Trans. Commun. Syst.*, vol. 43, pp. 1441–1453, 1995.

[8] E. Fogel and Y. F. Huang, "On the value of information in system identification—Bounded noise case," *Automatica*, vol. 18, pp. 229–238, 1982.

[9] S. Dasgupta and Y. F. Huang, "Asymptotically convergent modified recursive least-squares with data dependent updating and forgetting factor for systems with bounded noise," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 383–392, 1987.

[10] J. R. Deller, Jr., M. Nayeri, and M. S. Liu, "Unifying the landmark developments in OBE processing," *Int. J. Automat. Contr. Signal Process.*, vol. 8, pp. 43–60, 1994.

[11] J. R. Deller, Jr., M. Nayeri, and S. F. Odeh, "Least-square identification with error bounds for real-time signal processing and control," *Proc. IEEE*, vol. 81, pp. 813–849, June 1993.

[12] J. R. Deller, Jr. and S. F. Odeh, "Adaptive set-membership identification in $\mathcal{O}(m)$ time for linear-in-parameters models," *IEEE Trans. Signal Processing*, vol. 41, pp. 1906–1924, May 1993.

[13] A. K. Rao and Y. F. Huang, "Tracking characteristics of an OBE parameter estimation algorithm," *IEEE Trans. Signal Processing*, vol. 41, pp. 1140–1148, Mar. 1993.

[14] Y. F. Huang and S. Gollamudi, "Set-membership identification for adaptive equalization," in *Proc. 38th Midwest Symp. Circuits Syst.*, Rio de Janeiro, Brazil, Aug. 13–16, 1995, pp. 819–822.

[15] S. Gollamudi and Y. F. Huang, "Updator-shared adaptive parallel equalization (U-SHAPE) using set-membership identification," in *Proc. IEEE Intl. Conf. Circuits Syst.*, Atlanta, GA, May 12–15, 1996, pp. 13–16.

[16] S. Gollamudi, Y. F. Huang, S. Kapoor, and S. Nagaraj, "Set-membership identification for multi-channel adaptive equalization," in *Proc. Symp. Modeling, Anal., Simulation, CESA IMACS Multiconf.*, July 1996, pp. 831–836.

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

[18] E. W. Bai, K. M. Nagpal, and R. Tempo, "Bounded-error parameter estimation: Noise models and recursive algorithms," *Automatica*, vol. 32, pp. 3–15, July 1996.

[19] J. R. Deller, Jr., "Set-membership identification in digital signal processing," *IEEE Acoust., Speech, Signal Processing Mag.*, vol. 6, pp. 4–22, Oct. 1989.

[20] D. Maksarov and J. P. Norton, "Tuning of noise bounds in state bounding," in *Proc. Symp. Modeling, Anal. Simulation, CESA IMACS Multiconf.*, July 1996.

[21] J. P. Norton, Ed., Special issues on "Bounded-error methods in system identification," *Int. J. Automat. Contr. Signal Process.*, vol. 8, Jan./Feb. 1994; vol. 9, Jan./Feb. 1995.

[22] E. Walter and H. Piet-Lahanier, "Estimation of parameter bounds from bounded-error data: A survey," *Math. Comput. Simul.*, vol. 32, pp. 449–468, 1990.

[23] S. Kapoor, S. Gollamudi, S. Nagaraj, and Y. F. Huang, "Tracking of time-varying parameters using optimal bounding ellipsoid algorithms," in *Proc. 34th Annu. Allerton Conf. Commun., Contr., Comput.*, Univ. Illinois, Urbana-Champaign, Oct. 2–4, 1996.

[24] M. Nayeri, T. M. Lin, and J. R. Deller, Jr., "Blind error-bound estimation for OBE identification," in *Proc. 34th Annu. Allerton Conf. Commun., Contr., Comput.*, Univ. Illinois, Urbana-Champaign, Oct. 2–4, 1996.

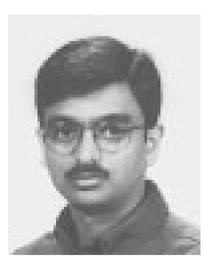
[25] F. C. Schweppe, "Recursive state estimation: Unknown but bounded errors and system inputs," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 22–28, Feb. 1968.

[26] F. L. Chernous'ko, "Optimal guaranteed estimates of indeterminacies with the aid of ellipsoids," *Eng. Cybern.*, vol. 18, pp. 1–9, 1980.

[27] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y. F. Huang, "Set-membership state estimation with optimal bounding ellipsoids," in *Proc. Int. Symp. Inform. Theory Applicat.*, Victoria, B.C., Canada, Sept. 17–20, 1996, pp. 262–265.

[28] D. G. Maksarov and J. P. Norton, "State bounding with ellipsoidal set description of the uncertainty," *Int. J. Contr.*, vol. 65, pp. 847–866, 1996.

[29] L. Kleinrock, *Queuing Systems, Vol. I: Theory.* New York: Wiley, 1975.



Sridhar Gollamudi (S'97) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Bombay, in May 1994 and the M.S.E.E. degree from University of Notre Dame, Notre Dame, IN, in January 1996. He is currently pursuing the Ph.D. degree at the Department of Electrical Engineering, University of Notre Dame.

He was awarded the Center for Applied Mathematics Fellowship at the University of Notre Dame for the academic year 1996 to 1997. From May to December 1997, he was an Intern at Lucent Technologies-Bell Laboratories, Whippany, NJ, where he worked on interference suppression for wireless CDMA systems. His research interests are in the areas of statistical and adaptive signal processing, communication theory, multiuser communications, and antenna array processing.

Mr. Gollamudi a member of Eta Kappa Nu.

Samir Kapoor received the B.Tech. degree from the Indian Institute of Technology, Bombay, in 1992, the M.S. degree from Washington State University, Pullman, in 1994, and the Ph.D. degree from the University of Notre Dame, Notre Dame, IN, in 1998, all in electrical engineering.

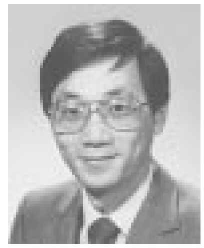
From 1994 to 1998, he was employed as a Research Engineer with Tellabs Research Center, Mishawaka, IN. He is currently a Research Staff Member at NEC C&C Research Laboratories, Princeton, NJ. His research interests include array processing, adaptive signal processing, and digital communications.

Dr. Kapoor is a member of Tau Beta Pi and Eta Kappa Nu.

Shirish Nagaraj (S'97) received the B.Tech. degree from the Indian Institute of Technology, Bombay, in May 1995 and the M.S. degree from the University of Notre Dame, Notre Dame, IN, in May 1997, both in electrical engineering. He is currently pursuing the Ph.D. degree at the University of Notre Dame.

His research interests include adaptive signal processing, statistical estimation and detection theory, set-membership filtering, and multiuser communications.

Mr. Nagaraj received the Michael J. Birck Fellowship for the period 1996 to 1997 and a graduate fellowship from the Center for Applied Mathematics, University of Notre Dame, for the period 1997 to 1998.



Yih-Fang Huang (F'95) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, R.O.C., in June 1976, the M.S.E.E. degree from the University of Notre Dame, Notre Dame, IN, in January 1980, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in October 1982.

Since August 1982, he has been on the faculty at the University of Notre Dame, where he is currently Professor and Chair of Electrical Engineering. In Spring 1993, he received the Toshiba Fellowship and was the Toshiba Visiting Professor at Waseda University, Tokyo, Japan, in the Department of Electrical Engineering. His research interests are in the general areas of statistical communications and signal processing. He has also worked on image source coding and, more recently, joint source-channel coding.

Dr. Huang served as Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS from 1989 to 1991 and for *Express Letters* for the same journal from 1992 to 1993. From 1995 to 1996, he was Chairman for the Digital Signal Processing Technical Committee of the IEEE Circuits and Systems Society. He was a Co-Chairman for Workshops/Short Courses for the 1997 International Symposium on Circuits and Systems. He serves as Regional Editor of America for the *Journal of Circuits, Systems, and Computers.* He also served as Vice President of Publications for the IEEE Circuits and Systems Society from 1997 to 1998.