Tracking Characteristics of an OBE Parameter-Estimation Algorithm

Ashok K. Rao, *Member, IEEE*, and Yih-Fang Huang, *Member, IEEE*

***Abstract*—Recently there seems to have been a resurgence of interest in recursive parameter-bounding algorithms. These algorithms are applicable when the noise is bounded and the bound is known to the user. One of the advantages of such algorithms is that 100% confidence regions (which are optimal in some sense) for the parameter estimates can be obtained at every time instant, rather than asymptotically as in the case of the least squares type algorithms. Another advantage is that these recursive algorithms have the inherent capability of implementing discerning updates, particularly that of allowing no updates of parameter estimates in the recursion. This paper investigates tracking properties of one such algorithm, referred to as the Dasgupta–Huang optimal bounding ellipsoid (DHOBE) algorithm. Conditions that ensure the existence of these 100% confidence regions in the face of small-model parameter variations are derived. For larger parameter variations, it is shown that the existence of the 100% confidence regions is guaranteed asymptotically. A modification is also proposed here to enable the algorithm to track large variations in model parameters. Simulation results show that in general, the modified algorithm has tracking performance comparable, and in some cases superior, to the exponentially weighted recursive least squares algorithm.**

I. Introduction

PERFORMANCE analysis of adaptive filtering is usually done by assuming that the unknown system being modeled is time-invariant. However, in practice, adaptive filters are often used in time-varying environments. It is thus important to investigate the performance of these algorithms, allowing the system-model parameters to vary with time. A considerable amount of attention has been paid to this problem in the adaptive-filtering literature, with analysis of varying amounts of rigor being performed mainly for the least mean square (LMS) and recursive least squares (RLS) algorithms; see, e.g., [1]–[5].

This paper investigates tracking properties of a recursive estimation algorithm, referred to hereafter as the Dasgupta–Huang optimal bounding ellipsoid (DHOBE) algorithm [6]. This algorithm belongs to a class of bounded-error estimation algorithms termed set-membership parameter estimation (SMPE) algorithms [7], [8]. The membership set is a set of parameter estimates compatible with the model of the underlying process, the assumptions on noise, and the observation data. At first glance the DHOBE algorithm appears to be very similar to the RLS algorithm. However, in contrast to the RLS algorithm, which obtains an optimal solution (in the sense of minimum mean-square estimation error) to the underlying problem, the DHOBE algorithm is developed by using a set-theoretic framework, namely, the notion of *optimal bounding ellipsoids* (OBE). This causes the algorithm to behave quite differently from the RLS algorithm in many ways. In addition, the algorithm incorporates a data-dependent forgetting factor that results in a discerning update strategy.

In case of time-varying systems, it is important to ensure that the time-varying true parameters $\{\theta^*(t)\}$ are contained in the bounding ellipsoids $\{E_t\}$ of the DHOBE algorithm. In this paper, such conditions will be derived. It will also be shown that if a jump in the true parameter vector $\theta^*(t)$ causes it to fall outside the bounding ellipsoid, then provided that the jump is not too large the bounding ellipsoids will move toward $\theta^*(t)$ and eventually enclose $\theta^*(t)$ again. A rescue scheme is proposed that will guarantee the existence of bounding ellipsoids in the face of large parameter variations. Some techniques for applying different parameter-bounding algorithms to time-varying systems have been reported by Norton and Mo [9]. One of the techniques suggested for the OBE-type algorithms is to use a fixed scaling factor to inflate the bounding ellipsoid with every new data point. Another technique that can be used if prior knowledge of the parameter increments is available is to vector sum the bounding ellipsoid with the set describing the parameter variation [9]. If the extent of parameter variation is unknown, as is often the case, the first technique will have to use a large scaling factor to cope with possibly large parameter variations and consequently the parameter bounds will be loose. In contrast, the rescue procedure described in this paper can automatically detect and accurately compensate for large parameter jumps.

Simulation results are presented to show that the DHOBE algorithm is able to track slow and abrupt variations in the parameters. The tracking performance, in terms of parameter-estimation error, is comparable to the RLS algorithm with a forgetting factor. Abrupt changes in the parameter can in some cases be tracked better by the DHOBE algorithm than by the RLS algorithm.

Manuscript received September 18, 1990; revised April 10, 1992. This work was supported in part by the National Science Foundation under Grant MIP 87-11174 and in part by the Office of Naval Research under Contract N00014-89-J-1788.

A. K. Rao is with Comsat Laboratories, Clarksburg, MD 20871.

Y.-F. Huang is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556.

IEEE Log Number 9206039.

1053-587X/93$03.00 © 1993 IEEE

II. The DHOBE Algorithm

One of the seminal works in SMPE is that of Fogel and Huang [10]. The algorithm of [10] recursively obtains ellipsoidal outer bounds to the membership set. The model structure considered is the following ARX model:

$$y(t) = \theta^{*T}\Phi(t) + v(t) \tag{2.1}$$

where

$$\theta^* = [a_1 a_2 \cdots a_n b_0 b_1 \cdots b_m]^T$$

is the true parameter vector and

$$\Phi(t) = [y(t-1)y(t-2)\cdots y(t-n) \cdot u(t)u(t-1)\cdots u(t-m)]^T$$

is the measurable regressor vector. The noise $v(t)$ is assumed to be uniformly bounded in magnitude with a known bound γ, i.e.,

$$|v(t)| \le \gamma. \tag{2.2}$$

Assume that at time instant $t-1$, the exact membership set is outer bounded by the ellipsoid E_{t-1} described by

$$E_{t-1} = \{\theta \in \boldsymbol{R}^N: [\theta - \theta(t-1)]^T P^{-1}(t-1) \cdot [\theta - \theta(t-1)] \le \sigma^2(t-1)\} \tag{2.3}$$

where $N = n + m + 1$, $P^{-1}(t-1)$ is a positive-definite matrix, and $\theta(t-1)$ is the center of the ellipsoid. At time instant t, the observation $y(t)$ yields a set S_t, which is a degenerate ellipsoid in the parameter space, namely,

$$S_t = \{\theta \in \boldsymbol{R}^N: [y(t) - \theta^T\Phi(t)]^2 \le \gamma^2\}. \tag{2.4}$$

From (2.1) and (2.2) it is clear that S_t contains the true parameter vector. An ellipsoid E_t, which contains the intersection of E_{t-1} and S_t, is then given by [10]

$$E_t = \{\theta \in \boldsymbol{R}^N: (1-\lambda_t)[\theta - \theta(t-1)]^T P^{-1}(t-1) \cdot [\theta - \theta(t-1)] + \lambda_t[y(t) - \theta^T\Phi(t)]^2\} \le (1-\lambda_t)\,\sigma^2(t-1) + \lambda_t\gamma^2\} \tag{2.5}$$

where λ_t is a positive time-varying updating gain. Note that $(1-\lambda_t)$ can be regarded as a forgetting factor. The formation of the ellipsoid E_t, which contains the intersection of an ellipsoid E_{t-1} and the set S_t, is illustrated by means of a 2-D example in Fig. 1. By performing some algebraic manipulations on (2.5), an expression for E_t can be obtained as

$$E_t = \{\theta \in \boldsymbol{R}^N: [\theta - \theta(t)]^T P^{-1}(t)\,[\theta - \theta(t)] \le \sigma^2(t)\} \tag{2.6}$$

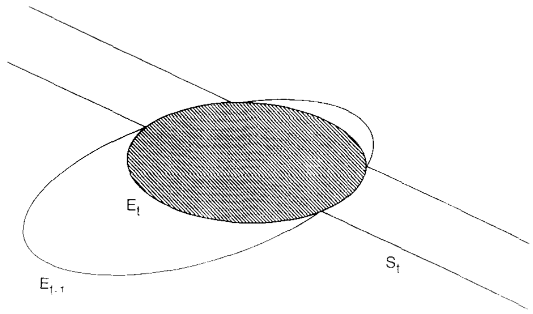


Fig. 1. Formation of the bounding ellipsoid E_t.

where

$$P^{-1}(t) = (1-\lambda_t)P^{-1}(t-1) + \lambda_t\Phi(t)\Phi^T(t) \tag{2.7}$$

$$\sigma^2(t) = (1-\lambda_t)\sigma^2(t-1) + \lambda_t\gamma^2 - \frac{\lambda_t(1-\lambda_t)[y(t) - \Phi^T(t)\theta(t-1)]^2}{1 - \lambda_t + \lambda_t\Phi^T(t)P(t-1)\Phi(t)} \tag{2.8}$$

$$\theta(t) = \theta(t-1) + \lambda_t P(t)\Phi(t) \cdot [y(t) - \Phi^T(t)\theta(t-1)]. \tag{2.9}$$

Using the matrix-inversion lemma in (2.7) yields

$$P(t) = \frac{1}{1-\lambda_t}\left[P(t-1) - \frac{\lambda_t P(t-1)\Phi(t)\Phi^T(t)P(t-1)}{1-\lambda_t + \lambda_t\Phi^T(t)P(t-1)\Phi(t)}\right]. \tag{2.10}$$

Equations (2.6)–(2.9) characterize the update of the bounding ellipsoids. The center $\theta(t)$ of the bounding ellipsoid E_t can be taken to be a point estimate of the parameter vector. Note that different values of λ_t yield different bounding ellipsoids [10]. To ensure convergence, λ_t needs to be chosen to optimize in some sense the bounding ellipsoids and, clearly, different optimization criteria would lead to different OBE algorithms.

In the DHOBE algorithm, the updating gain λ_t is chosen to minimize $\sigma^2(t)$ at every instant t. This has the effect of usually decreasing the size of the ellipsoid from iteration to iteration, though there is no guarantee that the size will be minimized. This choice of λ_t has yielded good results experimentally and in addition has simplified the convergence and tracking analysis of the algorithm. The minimization procedure yields the following updating criterion [6]:

$$\text{if } \sigma^2(t-1) + \delta^2(t) \le \gamma^2 \quad \text{then } \lambda_t = 0 \text{ (i.e., no update)} \tag{2.11}$$

where $\delta(t)$ is the *a priori* prediction error, namely,

$$\delta(t) = y(t) - \Phi^T(t)\theta(t-1). \tag{2.12}$$

Otherwise, if $\sigma^2(t-1) + \delta^2(t) > \gamma^2$, then the optimum value of λ_t is nonzero and can be calculated according to

$$\lambda_t = \min(\alpha, \nu_t)$$

where

$$\nu_t = \begin{cases} \alpha, & \text{if } \delta^2(t) = 0 \quad (2.13.a) \\ \dfrac{1-\beta(t)}{2}, & \text{if } G(t) = 1 \quad (2.13.b) \\ \dfrac{1}{1-G(t)}\left[1 - \sqrt{\dfrac{G(t)}{1+\beta(t)[G(t)-1]}}\right], & \text{if } 1 + \beta(t)[G(t)-1] > 0 \quad (2.13.c) \\ \alpha, & \text{if } 1 + \beta(t)[G(t)-1] \le 0 \quad (2.13.d) \end{cases}$$

and α is a user-chosen upper bound on λ_t satisfying

$$0 < \alpha < 1 \tag{2.14}$$

and

$$G(t) = \Phi^T(t)P(t-1)\Phi(t) \tag{2.15}$$

and

$$\beta(t) = \frac{\gamma^2 - \sigma^2(t-1)}{\delta^2(t)}. \tag{2.16}$$

The initial conditions are chosen to ensure that $\theta^* \in E_0$. A possible choice is

$$P(0) = \mathbf{I}, \quad \theta(t) = 0 \quad \text{and } \sigma^2(0) = 1/\epsilon^2, \quad \text{where } \epsilon \ll 1.$$

Equations (2.8)–(2.16) define recursions of the DHOBE algorithm. In [6], some convergence-type properties such as convergence of the parameter estimates to a ball and boundedness of the prediction error have been shown for time-invariant systems. In [11] and [12], an extension of this algorithm was developed for autoregressive moving average (ARMA) parameter estimation and similar convergence properties have been shown to hold.

III. Analysis of Tracking Characteristics

As mentioned earlier, tracking in the context of OBE algorithms for parameter estimation will mean ensuring that the time-varying true parameter vector is contained in the bounding ellipsoid. The theorems below present conditions under which parameter tracking can be accomplished.

Theorem 1: A sufficient condition for $\theta^*(t) \in E_t$ is

$$(\theta^*(t) - \theta(t-1))^T P^{-1}(t-1)(\theta^*(t) - \theta(t-1)) \le \sigma^2(t-1). \tag{3.1}$$

Proof: If $\theta^*(t) \in E_{t-1}$, then since $\theta^*(t) \in S_t$ and $E_t \supseteq E_{t-1} \cap S_t$, it follows that $\theta^*(t) \in E_t$. And from (2.3), $\theta^*(t) \in E_{t-1}$ is equivalent to (3.1).

Theorem 2: At any time instant t, the true parameter $\theta^*(t) \in E_t$ if and only if

$$(\theta^*(t) - \theta(t-1))^T P^{-1}(t-1)(\theta^*(t) - \theta(t-1)) \le \sigma^2(t-1) + \frac{\lambda_t}{1-\lambda_t}(\gamma^2 - v^2(t)) \tag{3.2}$$

where $v(t)$ is the noise term in (2.1).

Proof: Subtracting $\theta^*(t)$ from both sides of (2.9) yields

$$\theta(t) - \theta^*(t) = \theta(t-1) - \theta^*(t) + \lambda_t P(t)\Phi(t)\delta(t). \tag{3.3}$$

Define the following quadratic function in $\theta^*(t)$

$$V(t) = [\theta(t) - \theta^*(t)]^T P^{-1}(t)[\theta(t) - \theta^*(t)].$$

Using (2.7) and (3.3) it is straightforward though tedious to show that

$$V(t) = (1-\lambda_t)[\theta(t-1) - \theta^*(t)]^T P^{-1}(t-1) \cdot [\theta(t-1) - \theta^*(t)] + \lambda_t v^2(t) - \frac{\lambda_t(1-\lambda_t)\delta^2(t)}{(1-\lambda_t) + \lambda_t G(t)}. \tag{3.4}$$

Using (2.8) in (3.4) yields

$$V(t) - \sigma^2(t) = (1-\lambda_t)[\theta(t-1) - \theta^*(t)]^T P^{-1}(t-1) \cdot [\theta(t-1) - \theta^*(t)] + \lambda_t(v^2(t) - \gamma^2) - (1-\lambda_t)\sigma^2(t-1). \tag{3.5}$$

Since $\theta^*(t) \in E_t$ if and only if $V(t) \le \sigma^2(t)$, thus (3.2) is obtained. ■

It is easy to see from Theorem 1 that if the true parameter $\theta^*(t)$ is constant for all t, then the bounding ellipsoids obtained by the DHOBE algorithm enclose $\theta^*(t)$ at all time instants. This is a property that all well-devised set-membership estimation algorithms should have when applied to estimation of time-invariant parameters. If, on the other hand, $\theta^*(t)$ is time varying, and if at some time instant t_k, $\theta^*(t)$ is found to be out of the bounding ellipsoid E_t, it must not have been included in E_{t-1}. Theorem 2 then demarcates the region in which $\theta^*(t)$ can migrate without loss of tracking. This region is shown in Fig. 2 for a 2-D case. This theorem also shows that by choosing γ^2 to be larger than the actual bound, say γ'^2 on $v^2(t)$, it is possible to increase the tracking capability of the algorithm. The next theorem gives an upper bound on the

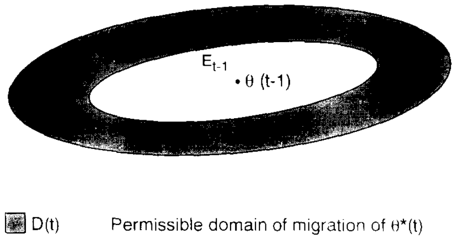


Fig. 2. Region outside E_{t-1} to which $\theta^*(t)$ can belong without loss of tracking.

maximum variation in the parameters for which tracking is guaranteed.

Theorem 3: If $\theta^*(t-1) \in E_{t-1}$ and $\lambda_t \neq 0$, then $\theta^*(t) \in E_t$ if

$$\|\Delta(t)\| \leq \frac{1}{\sqrt{\lambda_{\min}[P^{-1}(t-1)]}} \cdot \left\{ \left[\frac{\lambda_t}{1-\lambda_t} \frac{\lambda_{\min}[P^{-1}(t-1)]}{\lambda_{\max}[P^{-1}(t-1)]} [\gamma^2 - \gamma'^2(t)] + \sigma^2(t-1) \right]^{1/2} - \sqrt{\sigma^2(t-1)} \right\} \tag{3.6}$$

where

$$\Delta(t) = \theta^*(t) - \theta^*(t-1) \tag{3.7}$$

and $\lambda_{\min}$ and $\lambda_{\max}$ denote, respectively, minimum and maximum eigenvalues, and $\|.\|$ denotes the usual Euclidean norm. The quantity γ'^2 is the actual bound on $v^2(t)$, and the threshold γ^2 needed for evaluating the optimal updating gain via (2.11) and (2.16) is chosen to be larger than γ'^2.

Proof: It is straightforward to show that

$$\begin{aligned}[\theta(t-1) - \theta^*(t)]^T P^{-1}(t-1)[\theta(t-1) - \theta^*(t)] &= V(t-1) + \Delta^T(t)\, P^{-1}(t-1)\,\Delta(t) - 2\Delta^T(t) \\ &\quad \cdot P^{-1}(t-1)\,\tilde{\theta}(t-1)\end{aligned} \tag{3.8}$$

where $V(t)$ has been defined previously and

$$\tilde{\theta}(t-1) = \theta(t-1) - \theta^*(t-1).$$

Substituting (3.8) into (3.5) and using the fact that $v^2(t) \leq \gamma'^2$ yield

$$\begin{aligned}V(t) - \sigma^2(t) \leq{}& (1-\lambda_t)[V(t-1) - \sigma^2(t-1)] \\ &+ \lambda_t(\gamma'^2 - \gamma^2) + (1-\lambda_t) \\ &\cdot [\Delta^T(t) P^{-1}(t-1)\Delta(t) - 2\Delta^T(t) \\ &\cdot P^{-1}(t-1)\tilde{\theta}(t-1)].\end{aligned} \tag{3.9}$$

Since $\theta^*(t-1) \in E_{t-1}$, therefore, $V(t-1) \leq \sigma^2(t-1)$, and thus a sufficient condition for $\theta^*(t) \in E_t$ is

$$\Delta^T(t)P^{-1}(t-1)\Delta(t) - 2\Delta^T(t)P^{-1}(t-1)\tilde{\theta}(t-1) \leq \frac{\lambda_t}{1-\lambda_t}(\gamma^2 - \gamma'^2) \tag{3.10}$$

i.e., $\theta^*(t) \in E_t$ if

$$\lambda_{\max}[P^{-1}(t-1)]\,\|\Delta(t)\|^2 + 2\|\Delta(t)\|\,\|\tilde{\theta}(t-1)\|\,\lambda_{\max} \cdot [P^{-1}(t-1)] \leq \frac{\lambda_t}{1-\lambda_t}(\gamma^2 - \gamma'^2). \tag{3.11}$$

Since $V(t-1) \leq \sigma^2(t-1)$, therefore

$$\|\tilde{\theta}(t-1)\|^2 \leq \frac{\sigma^2(t-1)}{\lambda_{\min}[P^{-1}(t-1)]}. \tag{3.12}$$

Substituting (3.12) in (3.11) gives a sufficient condition for $\theta^*(t) \in E_t$ as

$$\lambda_{\max}[P^{-1}(t-1)]\,\|\Delta(t)\|^2 + 2\|\Delta(t)\|\,\sqrt{\sigma^2(t-1)} \cdot \frac{\lambda_{\max}[P^{-1}(t-1)]}{\sqrt{\lambda_{\min}[P^{-1}(t-1)]}} \leq \frac{\lambda_t}{1-\lambda_t}(\gamma^2 - \gamma'^2). \tag{3.13}$$

Solving this quadratic inequality then yields (3.6). ■

It can be seen from (3.6) that if $\lambda_t = 0$, then the difference between γ^2 and γ'^2 cannot be exploited to increase the tracking capability of the algorithm. In this case, $\theta^*(t) \in E_t$ if and only if $\theta^*(t) \in E_{t-1}$. Thus if $\theta^*(t)$ jumps out of E_{t-1} and no updates are performed at future time instants $t+i$, then $\theta^*(t+i) \notin E_{t+i} = E_{t-1}$, and the parameter may never be tracked. However, it can be argued that an update will be performed in a finite interval of time. This is shown heuristically by examining the expression for the magnitude of the prediction error

$$|\delta(t)| = |[\theta^*(t) - \theta(t-1)]^T\Phi(t) + v(t)|.$$

Assume that no updates are performed for a large interval of time, say from time instant t to time instant $t + N_1$. From (2.11) it then follows that

$$|[\theta^*(t+i) - \theta(t-1)]^T\Phi(t+i) + v(t+i)| \leq [\gamma^2 - \sigma^2(t-1)]^{1/2}, \qquad \forall i = 0, 1, \cdots, N_1.$$

If the input and noise sequences are sufficiently rich, then the regressor vector $\Phi(t)$ will span the parameter space in all directions and so $[\theta^*(t+i) - \theta(t-1)]^T\Phi(t+i)$ will not be arbitrarily small for all $i \in [0, N_1]$. If $|v(t+i)|$ is close to its true upper bound γ' for some i in the same interval, and if $\{v(t)\}$ is sufficiently uncorrelated with the input $\{u(t)\}$, then the above inequality will be violated and an update will be performed. It is also clear that to ensure that an update is eventually performed (i.e., violation of the above inequality), the threshold γ^2 should not be chosen much larger than γ'^2.

If the parameter variation is such that (3.2) is violated, then $\theta^*(t) \notin E_t$. The next theorem shows that if $\theta^*(t)$ remains fixed after it jumps out of E_t, and if the jump is not large enough to cause the subsequent ellipsoids E_{t+i} to vanish for $i \geq 0$, then the DHOBE algorithm guarantees that the true parameter will be tracked (enclosed) in finite time.

Theorem 4: Assume that the parameter variation at time instant t causes $\theta^*(t) \notin E_t$. Assume further that:

1) After this variation, the parameter remains constant (i.e., the jump-parameter case).
2) $\sigma^2(t+i) > 0$, for all $i \geq 0$.
3) The algorithm does not stop updating.
4) A lower bound ρ is imposed on all λ_t at all updating instants.

Then there exists an $N_1 > 0$, which depends on the amount of parameter variation and the actual and user-set noise bounds, such that $\theta^*(t) \in E_{t+N_1}$.

Proof: Since $\theta^*(t) \notin E_t$, define

$$\eta = [\theta(t) - \theta^*(t)]P^{-1}(t)[\theta(t) - \theta^*(t)] - \sigma^2(t) > 0. \tag{3.14}$$

Assumption (1) will imply that $\Delta(t + N_1) = \Delta(t+1) = 0$ for arbitrary positive N_1. Substituting in (3.9) and iterating from $t + N_1$ to $t + 1$ yields

$$\begin{aligned} V(t+N_1) &- \sigma^2(t+N_1) \\ &= \eta \prod_{i=t+1}^{t+N_1} (1-\lambda_i) + \sum_{i=t+1}^{t+N_1} q_{i,t+N_1}[\gamma'^2(t) - \gamma^2] \end{aligned} \tag{3.15}$$

where $q_{i,t}$ is defined as

$$q_{it} = \begin{cases} \lambda_i \prod_{j=i+1}^{t} (1-\lambda_j), & \text{if } i < t \\ \lambda_t, & \text{if } i = t. \end{cases}$$

Assumption (3) will ensure that some of the λ_{t+i}, $i \geq 0$, will be nonzero. This ensures that the first term on the right-hand side of (3.15) will tend to zero. Since the second term on the right-hand side of (3.15) is negative, the difference $V(t + N_1) - \sigma^2(t + N_1)$ will tend to zero as N_1 increases. Thus there exists an N_1 such that

$$V(t+N_1) - \sigma^2(t+N_1) \leq 0 \tag{3.16}$$

thereby ensuring that $\theta^*(t) \in E_{t+N_1}$. ■

IV. A Rescue Procedure

In many cases when the parameter jump is large or if the ellipsoid has shrunk to a very small size, the intersection of E_{t-1} and S_t can be void. This situation is illustrated in Fig. 3. In such cases, $\sigma^2(t)$ will become negative, thus indicating that a bounding ellipsoid could not be constructed. To circumvent such a failure of the algorithm, a rescue procedure is proposed. If at any time instant t, $\sigma^2(t)$ becomes negative, then $\sigma^2(t-1)$ is increased by an appropriate amount, thereby increasing the size of E_{t-1} so that the intersection of S_t and this enlarged E_{t-1} will no longer be void. As such, an ellipsoid E_t will be constructed. Alternatively, γ^2 could be increased to permit a non-null intersection. However, the former procedure is preferable because it causes $\theta(t)$ to migrate towards $\theta^*(t)$, thereby reducing the parameter-estimation

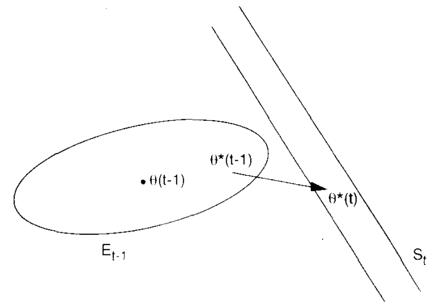


Fig. 3. A case in which a jump in the parameter causes the intersection of E_{t-1} and S_t to be void.

error. The rescue procedure is similar to the covariance-resetting technique used in RLS algorithms to cope with time-varying systems [13]. However, in the RLS case, a jump in the parameters has to be detected by some other means before the covariance matrix can be reset, whereas for the DHOBE algorithm, $\sigma^2(t)$ becoming negative is an automatic indicator of a jump. The amount of increase in $\sigma^2(t-1)$ required to make $\sigma^2(t)$ positive in such a case is now calculated.

Recall that the optimal updating gain λ_t is the one that minimizes $\sigma^2(t)$. The minimum occurs either at a stationary point of $\sigma^2(t)$ or at one of the boundaries $\lambda_t = 0$ and $\lambda_t = \alpha$. Since it is assumed that a failure occurs when $\sigma^2(t-1) > 0$ and $\sigma^2(t) \leq 0$, an update, therefore, has to occur at t and so $\lambda_t \neq 0$. The case that the minimum occurs at a stationary point, which is strictly inside the interval $[0, \alpha]$, and the case that the minimum occurs at $\lambda_t = \alpha$ are considered separately.

Case 1:

$$\left.\frac{d\sigma^2(t)}{d\lambda_t}\right|_{\lambda_t = \nu_t} = 0, \qquad \text{and } 0 < \nu_t < \alpha.$$

From (2.13) it is clear that this case occurs if and only if $1 + \beta(t)[G(t) - 1] > 0$ and $\nu_t < \alpha$. Setting the derivative of $\sigma^2(t)$ in (2.8) to zero yields

$$\gamma^2 - \sigma^2(t-1) - \frac{1-\lambda_t}{1-\lambda_t+\lambda_t G(t)}\delta^2(t) + \frac{\lambda_t G(t)}{(1-\lambda_t+\lambda_t G(t))^2}\delta^2(t) = 0.$$

Substituting $\sigma^2(t-1)$ from above into (2.8) yields

$$\sigma^2(t) + \frac{(1-\lambda_t)^2}{(1-\lambda_t+\lambda_t G(t))^2}\delta^2(t) = \gamma^2. \tag{4.1}$$

Thus, $\sigma^2(t)$ is negative if and only if

$$|\delta(t)| > \frac{1-\lambda_t+\lambda_t G(t)}{1-\lambda_t}\gamma. \tag{4.2}$$

On substituting for λ_t from (2.13b) and (2.13c), (4.2) can be expressed, respectively, as

$$|\delta(t)| > \frac{G(t) - 1}{\sqrt{G(t)[1 + \beta(t)(G(t) - 1)]} - 1} \gamma, \quad \text{if } G(t) \neq 1$$

$$|\delta(t)| > \frac{2\gamma}{1 + \beta(t)}, \quad \text{if } G(t) = 1. \tag{4.3}$$

Using the definition of $\beta(t)$ from (2.16) in (4.3) and manipulating terms yields a necessary and sufficient condition for $\sigma^2(t)$ to be negative in terms of $\sigma^2(t-1)$

$$\sigma^2(t-1) < \frac{1}{G(t) - 1}\left[\delta^2(t) + \gamma^2[G(t) - 1] - \frac{[\gamma[G(t) - 1] + |\delta(t)|]^2}{G(t)}\right] = K_1, \quad \text{if } G(t) \neq 1$$

and

$$\sigma^2(t-1) < \delta^2(t) + \gamma^2 - 2\gamma|\delta(t)| = K_1, \quad \text{if } G(t) = 1.$$

Note that the last inequality was obtained because $\nu_t = (1 - \beta(t))/2 < 1$; hence, $1 + \beta(t) > 0$. Thus, if the calculated value of $\sigma^2(t)$ is negative, the rescue procedure will replace $\sigma^2(t-1)$ by $K_1 + \zeta$, where ζ is a positive constant, thereby increasing the size of E_{t-1}. The optimum updating gain will then be recalculated, and the resulting value will be used to calculate $\sigma^2(t)$, $\theta(t)$, and $P(t)$. Our simulation studies have shown that using a value of $\zeta = 1$ yields satisfactory results.

Case 2: $\lambda_t = \alpha$

In this case, from (2.8), $\sigma^2(t)$ is negative if and only if

$$\delta^2(t) \geq [1 - \alpha + \alpha G(t)]\left[\frac{\sigma^2(t-1)}{\alpha} + \frac{\gamma^2}{1-\alpha}\right].$$

Thus, $\sigma^2(t)$ is negative if and only if

$$\sigma^2(t-1) < \alpha\left[\frac{\delta^2(t)}{1 - \alpha + \alpha G(t)} - \frac{\gamma^2}{1-\alpha}\right] = K_2.$$

In this case, $\sigma^2(t-1)$ would be replaced by $K_2 + \zeta$ and the value of the updating gain would be recalculated and used to calculate $\sigma^2(t)$, $\theta(t)$, and $P(t)$.

V. Simulation Examples

The tracking properties of the DHOBE algorithm are studied for an ARX(1,1) model

$$y(t) = ay(t-1) + bu(t) + v(t)$$

The nominal values for the parameters were $a = -0.5$ and $b = 1.0$. The noise sequence $\{v(t)\}$ and the input sequence $\{u(t)\}$ were both generated by a pseudorandom-number generator with a uniform distribution in $[-1, 1]$. This corresponds to a signal-to-noise ratio (SNR) of 0 dB. For the DHOBE algorithm, we chose $\alpha = 0.2$, $\gamma^2 = 1.0$, and $\sigma^2(0) = 100$. In all the examples shown here, the parameter estimates are taken to be the centers of the optimal bounding ellipsoids. The parameters were varied as follows:

Case 1: Slow Variation in the Parameter Vector

The parameters a and b were varied by 1% for every 10 samples, starting from the first sample, and the output data $\{y(t)\}$ were generated for $t = 1, 2, \cdots, 1000$. It was then observed that the bounding ellipsoids created by the DHOBE algorithm contain the true parameter at all time instants. The final parameter-estimation error was 7.0×10^{-3}. The parameter estimates, i.e, the centers of the OBE, are plotted against the true parameters in Fig. 4. From the figure it is clear that the DHOBE algorithm tracks slow time variations in the parameters quite well.

Case 2: Slow Variation in the Parameter Vector from $t = 500$

The parameters a and b were varied by 1% for every 10 samples, starting from the five-hundredth sample. The final parameter-estimation error was 3.0×10^{-3}. All the bounding ellipsoids were seen to contain the true parameter. The parameter estimates are plotted against the true parameters in Fig. 5. The figure shows that the algorithm can track slow time variations in the parameters even after it has "converged."

Case 3: Jump in the MA Parameter at $t = 500$

The parameter b was changed by 100% at the five-hundredth sample, and a was kept constant at its nominal value at all times. Several runs of the DHOBE algorithm were performed with different input and noise sequences. It was observed that the true parameter vector was out of the bounding ellipsoid at $t = 500$ and would be recaptured by the bounding ellipsoid after some number of samples (usually less than 50), thus verifying the claims made in Theorem 4. It was also observed that the jump causes the resulting bounding ellipsoids to have smaller sizes. Intuitively, a jump at time t causes the set S_i, $i \geq t$, to have a smaller intersection with E_{i-1}, and so the ellipsoid that bounds the intersection is also smaller. In one particular run, the parameter was recaptured at $t = 530$, and the final parameter estimation error at $t = 1000$ was 1.3×10^{-4}. The parameter estimates (the centers of the bounding ellipsoids) are plotted against the true parameters in Fig. 6. Fig. 7 shows the parameter estimates obtained for this run by applying the RLS algorithm with a forgetting factor $\lambda(t) = 0.9$ and $\lambda(t) = 0.99$. Observe that the RLS parameter estimates are extremely jumpy when $\lambda(t) = 0.9$, probably because the forgetting factor is not large enough to average out the noise. Fig. 8 shows the estimates when the variable forgetting factor proposed by Fortescue and Kershenbaum [13] is incorporated into the RLS algorithm. This variable forgetting factor $\lambda(t)$ is a

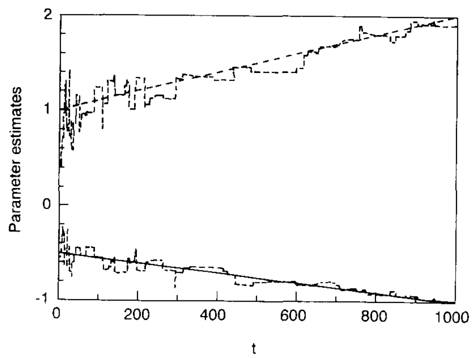


Fig. 4. DHOBE parameter estimates for the case of slow variation in the true parameter from $t = 1$.

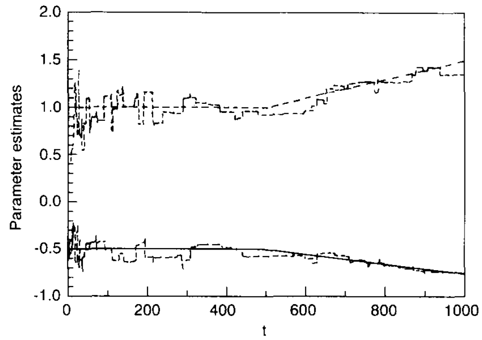


Fig. 5. DHOBE parameter estimates for the case of slow variation in the true parameter from $t = 500$.

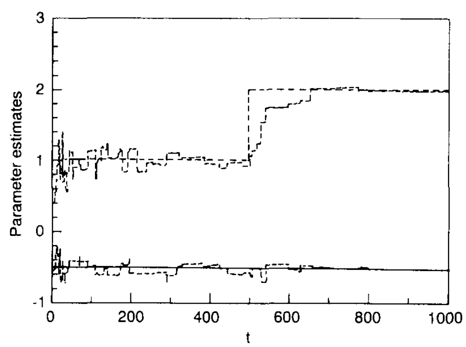


Fig. 6. DHOBE parameter estimates for the case of a jump in the MA parameter at $t = 500$.

function of the prediction error and is given by

$$\lambda(t) = 1 - \alpha' \frac{\delta^2(t)}{1 + G(t)}.$$

A value of $\alpha' = 0.01$ was used because it yields steady-state tracking error of about the same magnitude as does the DHOBE algorithm. From these figures, it is evident that the DHOBE algorithm can track jumps in the parameters at least as well as the exponentially weighted RLS algorithm.

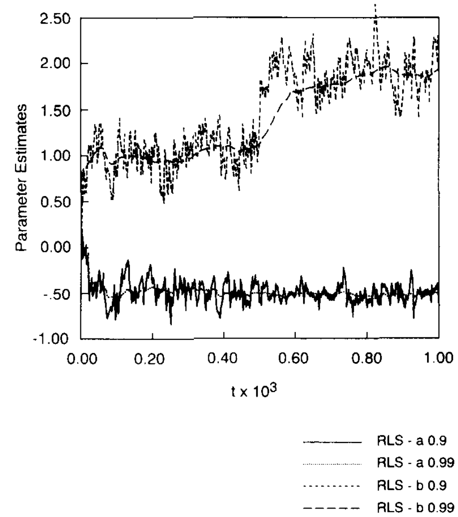


Fig. 7. RLS (with $\lambda(t) = 0.9$ and $\lambda(t) = 0.99$) parameter estimates for the jump-parameter case.

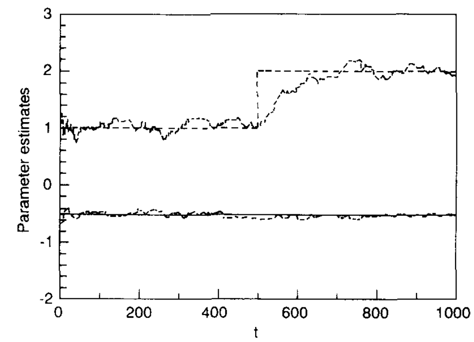


Fig. 8. RLS (with variable forgetting factor) parameter estimates for the jump-parameter case.

The effect of varying γ^2 was also studied. A value of $\gamma^2 = 2$ was taken. In this case, the true parameter did not jump out of the bounding ellipsoid at $t = 500$. The parameter estimates are identical to those in Fig. 6. But the ellipsoids are larger, as expected.

For a different run, i.e., with a different input and noise sequence, the jump at $t = 500$ caused $\sigma^2(t)$ to become negative. The rescue procedure was then used and yielded remarkable results. The true parameter was captured immediately at $t = 501$. The final parameter-estimation error was 2.4×10^{-4}. Fig. 9 shows that the parameters are tracked extremely rapidly in this case.

Tracking Performance in Gaussian Noise

It is well known that least squares algorithms are optimal in the constant-parameter case for Gaussian-distributed noise. It is thus interesting to compare the tracking abilities of the DHOBE and RLS algorithms in Gaussian noise. The same ARX model was used with the noise sequence $v(t)$ now being generated as zero-mean white

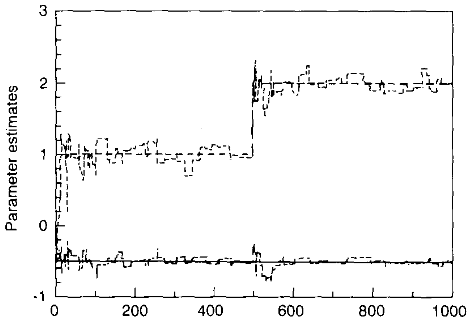


Fig. 9. DHOBE parameter estimates when the rescue procedure is activated in the jump-parameter case.

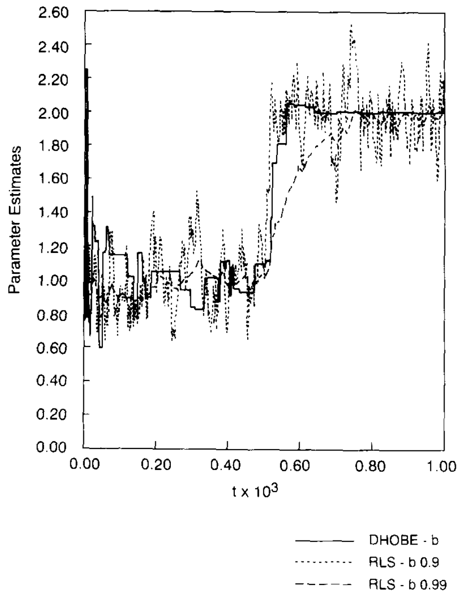


Fig. 10. Tracking performance of DHOBE and RLS algorithms for Gaussian noise.

Gaussian noise with variance 0.25, which corresponds to an SNR of 1.25 dB. To satisfy the bounded-noise assumption, $v(t)$ was truncated to the range $[-1, 1]$, resulting in a slightly larger SNR. The parameter b was changed by 100% at the five-hundredth sample, and a was kept constant at its nominal value at all times. Several runs of the DHOBE algorithm were performed with different noise sequences. As in the uniform-noise case, it was found that in a few runs the rescue procedure was activated, consequently causing extremely rapid acquisition of the parameter. In most of the runs, the true parameter was acquired by the bounding ellipsoid without requiring rescue. The acquisition usually happened in less than twenty samples after the change occurred. Fig. 10 compares the tracking performance of the RLS algorithm (with $\lambda(t) = 0.9$ and $\lambda(t) = 0.99$) to the DHOBE algorithm for a run in which the rescue procedure was *not* activated. The curves shown are plots of estimates of parameter b by both algorithms. It is seen that RLS with $\lambda(t) = 0.9$ seems to track a little faster than the DHOBE algorithm. However, the steady-state RLS estimates are extremely jerky. The tracking performance of RLS with $\lambda(t) = 0.99$ is definitely inferior to that of the DHOBE algorithm; however, its steady-state performance prior to the jump is superior. Another point of note is that the DHOBE estimates become much less jerky after the jump on account of the decrease in the size of the ellipsoids.

VI. Conclusion

The tracking properties of a recursive set-membership parameter estimation algorithm, the DHOBE algorithm, have been investigated. Some sufficient and other necessary conditions that ensure parameter tracking have been derived. A modification of the DHOBE algorithm is proposed to improve its tracking capability for larger parameter variations. Simulation results show that the tracking performance of the DHOBE algorithm is comparable to that of the exponentially weighted RLS algorithm. In some cases of large parameter jumps, the automatic activation of a rescue procedure causes the parameters to be tracked extremely rapidly.

References

[1] B. Widrow *et al.*, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, no. 8, pp. 1151–1162, Aug. 1976.

[2] A. Benveniste and G. Ruget, "A measure of the tracking capability of recursive stochastic algorithms with constant gains," *IEEE Trans. Automat. Contr.*, vol. AC-27, no. 3, pp. 639–649, June 1982.

[3] E. Eleftheriou and D. D. Falconer, "Tracking properties and steady-state performance of RLS adaptive filter algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 5, pp. 1097–1109, Oct. 1986.

[4] B. D. Rao and R. Peng "Tracking characteristics of the constrained IIR adaptive notch filter," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 9, pp. 1466–1479, Sept. 1988.

[5] S. Gunnarsson and L. Ljung, "Frequency domain tracking characteristics of adaptive algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 1072–1089, July 1989.

[6] 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, no. 3, pp. 383–392, May 1987.

[7] J. Deller, "Set membership identification and signal processing," *IEEE ASSP Mag.*, vol. 6, no. 4, pp. 4–20, Oct. 1989.

[8] E. Walter and H. Piet-Lahanier, "Estimation of parameter bounds from bounded-error data: A survey," *Math. Comput. Simulation*, vol. 32, no. 5-6, pp. 449–465, Dec. 1990.

[9] J. P. Norton and S. H. Mo, "Parameter bounding for time-varying systems," *Math. Comput. Simulation*, vol. 32, no. 5-6, pp. 527–534, Dec. 1990.

[10] E. Fogel and Y. F. Huang, "On the value of information in system identification-bounded noise case," *Automatica*, vol. 18, no. 2, pp. 229–238, Mar. 1982.

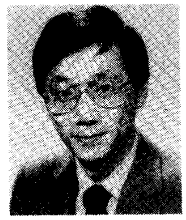
[11] A. K. Rao, "Membership-set parameter estimation via optimal bounding ellipsoids," Ph.D. dissertation, Dept. Elec. Comput. Eng., Univ. of Notre Dame, Notre Dame, IN, Jan. 1990.

[12] A. K. Rao, Y. F. Huang, and S. Dasgupta, "ARMA parameter estimation using a novel recursive estimation algorithm with selective updating," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 3, pp. 447–457, Mar. 1990.

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

Ashok K. Rao (S'84–M'89) was born in New Delhi, India, on January 12, 1962. He received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Delhi, India, in 1984 and M.S.E.E. and Ph.D. degrees in electrical engineering from the University of Notre Dame, Notre Dame, IN, in 1987 and 1989, respectively.

In 1989, he joined Comsat Laboratories, Clarksburg, MD, as a Member of the Technical Staff in the Communication Technology Division. He has been actively involved in the development of algorithms and hardware for the compression of conventional and high-definition television. His research interests are in the areas of image coding and adaptive filtering.



Yih-Fang Huang (S'80–M'82) received the B.S.E.E. degree from National Taiwan University, Taipei, Taiwan, in 1976, the M.S.E.E. degree from the University of Notre Dame, Notre Dame, IN, in 1979, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1982.

In 1982, he joined the faculty of the University of Notre Dame, where he is currently Professor of Electrical Engineering. In Spring 1993, he was a Visiting Professor in the Department of Electrical Engineering, Waseda University, Tokyo, Japan. His research interests are in statistical signal processing and artificial neural networks.

Dr. Huang is currently Associate Editor of Express Letters for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS. He also served as Associate Editor of Neural Networks and Signal Processing for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS in 1989–1991.