Bounded Error Estimation: Set-Theoretic and Least-Squares Formulations

Shirish Nagaraj, Sridhar Gollamudi, Samir Kapoor and Yih-Fang Huang
Department of Electrical Engineering
University of Notre Dame, IN 46556
Email: huang.2@nd.edu *

Abstract

An Optimal Bounding Ellipsoid (OBE) algorithm with asymptotic convergence properties and a selective update capability is developed in this paper. Interesting geometrical insights into the updating rule are found. This algorithm is shown to be robust with respect to certain model violations. Convergence of the algorithm can be established even when the assumption of bounding ellipsoids breaks down due to incorrect initialization. A least-squares formulation is developed next for systems with bounded noise. The recursive estimation procedure of this approach is shown to be identical to that of the proposed OBE algorithm. Fundamental connections between these two approaches are also shown. These results are shown to imply enhanced tracking performance of the algorithm. Simulation results, tested on real channel data, show excellent performance inspite of highly reduced updating. Moreover, the algorithm is shown to render superior tracking capability, as expected.

1 Introduction

The problem of estimation of a linear-in-parameter system, given knowledge of the input and corrupted output observations has been a widely studied problem in Control, Signal Processing and Communication [1]. The classical method of Least-Squares has been extensively studied to address this problem [1]. When there is additional knowledge about the system to be estimated (identified), it is important to include the same in the estimation procedure. Consequently, significant performance gains can be achieved by the explicit use of available information. A common, and realistic assumption which can be made about many such linear systems is that the noise corrupting the observations is bounded instantaneously by an *a priori* known value. Such knowledge leads to a *set* of feasible estimators which are consistent with the given observations and model assumptions. Methodologies exploiting this set theoretic knowledge are conceptually distinct from the philosophy of traditional point-wise estimators (like least-squares, maximum likelihood, etc). These methods have attracted significant attention in the signal processing and control communities over the last decade and in identification literature, they have been termed as Set-Membership Identification (SMI) algorithms ([2, 3, 4]). The objective of such techniques is to maintain consistency with the model assumptions, and are hence, not tied to any specific cost function.

A class of popular SMI algorithms are the group of the so-called Optimal Bounding Ellipsoid (OBE) algorithms ([2, 5]) which are recursive estimation schemes. The OBE algorithms have gained widespread popularity in the research community following the work of Fogel and Huang [5]. The OBE algorithms enjoy a *discerning update* capability, *i.e.*, they use the data selectively for updating the parameter estimates. Traditional least-squares algorithms, on the other hand, lack a simple procedure to exploit possible redundancy in the data [5].

In this paper, we derive a novel OBE algorithm with enhanced performance and robustness to model violations in Section 2. The optimality criterion in this algorithm, for obtaining "tightly" outer bounding ellipsoids, is the same as that used in the algorithm (DH/OBE), by Dasgupta and Huang [2]. This is the so-called σ_i minimization method. The proposed algorithm features a highly selective update policy, while exhibiting excellent performance in terms of mean-square error (MSE). A geometrical insight into the selective updating process is also derived. Convergence of this algorithm is proved under certain conditions.

The robustness results motivate the need for deriving, in Section 3, a least-squares estimator with a bounded noise assumption. It is shown that a certain least-squares like formulation leads to *exactly* the same recursive procedure for the estimate as that of the OBE algorithm developed in Section 2. It is further shown that the problem is amenable to a geometric solution.

This analysis also allows us to provide links between least-squares and OBE estimation paradigms, which were not available in such an explicit and neat manner before. We show that the weights obtained via σ_i minimization are precisely the Lagrangian multipliers for a least-squares cost function. The discerning update capability also arises directly as a consequence of the

*This work was supported, in part, by the Office of Naval Research under contract N00014-95-1-0436, in part by the Center for Applied Math., University of Notre Dame and in part, by the Tellabs Research Center, Mishawaka, IN.

problem statement from the least-squares perspective. As a result, it allows for an optimal choice of weights in an RLS algorithm. Discussions on these aforementioned issues, including tracking performance, are given in Section 4. Simulation results are presented in Section 5. These are for an application to the estimation of a microwave radio communications channel, generated by actual field experiments. A performance comparison of the proposed algorithm is made vis-a'-vis the popular RLS algorithm. It is seen that the algorithm renders excellent performance with highly reduced computational requirements as compared to RLS.

2 A Novel OBE Algorithm

The assumed discrete time system model is of the form

$$y_i = x_i^T \theta^* + v_i \tag{1}$$

where $\{y_i\} \in \mathcal{R}$ is the output corrupted by the noise $\{v_i\} \in \mathcal{R}$; $\{x_i\} \in \mathcal{R}^n$ is the known input sequence and $\theta^* \in \mathcal{R}^n$ is the system parameter to be estimated. We assume that the noise satisfies the following boundedness criterion

$$v_i \in \mathcal{V} = \{v \in \mathcal{R} : |v|^2 \le \gamma^2\} \quad \forall i = 1, 2, \dots \tag{2}$$

for a given noise bound $\gamma > 0$.

Equation (2), combined with (1), results in a feasible set of parameters at each instant i which are consistent with the data observed at that time. Obviously, the true parameter lies in this set, which is referred to here as the *observation set* (denoted by $\mathcal{S}_i$), given by

$$\mathcal{S}_i = \{\theta \in \mathcal{R}^n : |y_i - x_i^T \theta|^2 \le \gamma^2\} \tag{3}$$

Given observations $\{y_k, x_k\}_{k=1}^i$ and the fact that $\theta^* \in \mathcal{S}_k$, for all $k = 1, 2, \dots, i$, it follows that $\theta^* \in \Psi_i = \bigcap_{k=1}^i \mathcal{S}_k$. The basic approach of OBE algorithms is to outer bound this *membership set* Ψ_i, by a mathematically tractable ellipsoid $\mathcal{E}_i$ for every i in some optimal manner.

2.1 Recursions and Selective Update Rule

Given an initial ellipsoid $\mathcal{E}_o = \{\theta \in \mathcal{R}^n : (\theta - \hat{\theta}_o)^T P_o^{-1}(\theta - \hat{\theta}_o) \le \sigma_o\}$, with some properly chosen $\hat{\theta}_o, P_o = I$ and $\sigma_o > 0$ such that $\theta^* \in \mathcal{E}_o$, the algorithm establishes a recursive procedure for computing the sequence of ellipsoids $\{\mathcal{E}_i\}$. Specifically, the true parameter θ^* lies in $\mathcal{E}_i$, where $\mathcal{E}_i$ is obtained as an outer bounding ellipsoid of the intersection of $\mathcal{S}_i$ with $\mathcal{E}_{i-1}$. Such an outer bounding ellipsoid is formulated here as a linear combination of $\{\mathcal{E}_i\}$ with $\mathcal{S}_i$. If $\{\mathcal{E}_i\}$ is described by $\{\theta \in \mathcal{R}^n : (\theta - \hat{\theta}_i)^T P_i^{-1}(\theta - \hat{\theta}_i) \le \sigma_i\}$, then the recursions of the algorithm are:

$$\hat{\theta}_i = \hat{\theta}_{i-1} + \lambda_i P_i x_i \delta_i \tag{4}$$

$$P_i^{-1} = P_{i-1}^{-1} + \lambda_i x_i x_i^T \tag{5}$$

$$\sigma_i = \sigma_{i-1} - \eta_i \tag{6}$$

$$\eta_i = \frac{\lambda_i \delta_i^2}{1 + \lambda_i G_i} - \lambda_i \gamma^2 \tag{7}$$

for any $\lambda_i \ge 0$ where the *prediction error* is $\delta_i = y_i - x_i^T \hat{\theta}_{i-1}$ and $G_i = x_i^T P_{i-1} x_i$.

Remark 1: For a properly chosen γ, P_o and $\sigma_o > 0$, we have $\sigma_i > 0$ for all i.

Remark 2: Each value of λ_i yields a different bounding ellipsoid. To compute λ_i, a measure of optimality which is similar to the one in DH/OBE has been adopted in this algorithm [2]. This is to minimize σ_i with respect to λ_i, under the constraint that $\lambda_i \ge 0$. It is important to note that this measure essentially amounts to maximizing η_i. When the maximum of η_i occurs at $\lambda_i = 0$, it results in no update. This constitutes the *discerning update* capability of the proposed algorithm.

Proposition 1: The optimal value of λ_i, denoted by λ_i^*, is obtained by the following:

$$\lambda_i^* = \arg\max_{\lambda_i} \lambda_i \{\frac{\delta_i^2}{\gamma^2}(\frac{1}{1 + \lambda_i G_i}) - 1\}$$

Moreover,

$$\lambda_i^* = \begin{cases} 0 & \text{if } |\delta_i| \le \gamma \\ \frac{1}{G_i}(\frac{|\delta_i|}{\gamma} - 1) & \text{if } |\delta_i| > \gamma \end{cases} \tag{8}$$

For the next result and much of the analysis later on, we need the following definitions:

Definition 1: The *"bounding hyper planes"* to the observation set, $\mathcal{S}_i$ at time i, denoted by $\mathcal{S}_i^+$ and $\mathcal{S}_i^-$ are given by:

$$\mathcal{S}_i^{\pm} = \{\theta \in \mathcal{R}^n : y_i - x_i^T \theta = \pm\gamma\} \tag{9}$$

Moreover, the *"interior"* of $\mathcal{S}_i$ is given by:

$$\mathcal{S}_i^{in} = \{\theta \in \mathcal{R}^n : |y_i - x_i^T \theta| < \gamma\} \tag{10}$$

It follows from the above construction that we have

$$\mathcal{S}_i = \mathcal{S}_i^+ \cup \mathcal{S}_i^{in} \cup \mathcal{S}_i^-$$

Definition 2: The *"nearest bounding hyper plane"* to a vector, say $\theta \notin \mathcal{S}_i$ at time instant i, denoted by $\mathcal{NB}_i(\theta)$ is:

$$\mathcal{NB}_i(\theta) = \begin{cases} \mathcal{S}_i^+, & \text{if } y_i - x_i^T \theta > \gamma \\ \mathcal{S}_i^-, & \text{if } y_i - x_i^T \theta < -\gamma \end{cases} \tag{11}$$

By the assignment to λ_i (in Proposition 1), we have the following result for the *a posteriori* error, $\epsilon_i = y_i - x_i^T \hat{\theta}_i$ from which we obtain a geometrical interpretation to the update rule.

Proposition 2: The resulting *a posteriori* error magnitude is always less than or equal to the noise bound and is given by:

$$\epsilon_i = \frac{\gamma}{|\delta_i|}\delta_i \quad \text{if } i \in \text{updating instants}$$

$$i.e. \quad |\epsilon_i| = \min(\gamma, |\delta_i|) \quad \forall i$$

Geometrically, the update rule is: *No update is needed if $\hat{\theta}_{i-1} \in \mathcal{S}_i$, else, update $\hat{\theta}_{i-1}$ in the direction of the vector $P_{i-1}x_i$ such that $\hat{\theta}_i \in \mathcal{NB}_i(\hat{\theta}_{i-1})$.*

Remark 3: Deller *et al.* have an algorithm with a similar update rule [3], whose motivation being to achieve

$\mathcal{O}(n)$ complexity of the information evaluation criterion. However, their algorithm uses weights which are computed via a volume or trace minimization procedure once this information evaluation criterion is met and hence belongs to a different class of OBE algorithms altogether.

2.2 Effects of Initial Conditions and Convergence Issues

The choice of the initial condition, *i.e.* σ_o, has been dismissed in existing OBE literature in a somewhat hand waving manner. The motivation for the following analysis is two-fold: (1) It can be shown that various choices of σ_o lead to significantly different performance in DH/OBE [2] and (2) The problem with tracking of time-varying parameters is that of σ_i becoming negative for some i. By a result for the case of model violation, we can effectively explain how this algorithm is more robust to time-varying parameters and consequently offers better performance in such situations.

The condition to be satisfied while choosing σ_o is:

$$\|\theta^*\|^2 \leq \sigma_o \tag{12}$$

where θ^* is the *true parameter.* In practice, we pick a "large enough" σ_o to start the algorithm. Also, $\hat{\theta}_o = 0$ forms the initial estimate of θ^*.

Remark 4: First, it can be shown that the effect of any arbitrary positive σ_o, such that (12) is satisfied, does not change the sequence of estimates by the algorithm. The above assertion does not hold for DH/OBE.

Now we focus attention to the potentially catastrophic case of a wrong choice of σ_o, such that (12) is not satisfied. In such a scenario, the basic assumption that $\mathcal{E}_o$ is a bounding ellipsoid fails, leading to what we term as a *model violation.* In the following important result, which is a generalization of Remark 4, we show that the recursions given by (4-7) and (8), is robust to these model violations.

Proposition 3: All values of σ_o (not necessarily positive), result in the same sequence of updating instants, weights (λ_i^*) and parameter estimates ($\{\theta_i\}\ \forall i$). Again, this is not true for DH/OBE.

Remark 5: The above results, *i.e.*, the robustness of the algorithm even when OBE assumptions do not hold, motivate the need to answer the following question:

"Does there exist a fundamentally different approach which results in the same estimation strategy?"

The answer to this question is addressed in Section 3 by casting the problem in a least-squares framework. With such an approach, the excellent performance of the algorithm and it's insensitivity to initial conditions can be thus explained.

Convergence issues in OBE algorithms were first addressed rigorously by Dasgupta and Huang [2]. Later, Nayeri, Deller and Liu [6] established convergence results for all known OBE algorithms, along with conditions to guarantee point convergence to the true parameter. The contribution in this work is that we prove asymptotic convergence of the prediction error, cessation of updating and region convergence of the estimator even under model violations. For this, results for the case when σ_0 is chosen to satisfy the model assumption (12) are established. Then, using Proposition 3, it can be shown that all the above results are true even when (12) is not satisfied.

Proposition 4: If G_i is bounded for all i, then **for any** σ_o, the following hold:

- $\limsup_{i\to\infty} |\delta_i| \leq \gamma$
- $\lim_{i\to\infty} \lambda_i = 0$
- $\limsup_{i\to\infty} \|\theta_{i-1} - \theta^*\|^2 \leq (\gamma^2 - \sigma_v^2)/\sigma_x^2$, where σ_v^2 is the noise variance and σ_x^2 is the minimum singular value of R_{xx}, the autocorrelation matrix of the (stationary) input.

We proceed to the alternate approach in the next section.

3 Least-Squares Estimation under Bounded Noise

As the title suggests, our approach is to formulate a least-squares cost function, given the knowledge that the uncertainty is bounded, but otherwise unknown. In the following discussion, we assume that the measurement and input are related by (1) and furthermore, that the noise obeys the bound given by (2).

3.1 Problem Formulation

To develop a recursive mechanism, let us assume that at time $i-1$, we have formulated a certain quadratic cost function given information up to and including time instant $i-1$, that is, $\{y_k, x_k\}_1^{i-1}$. Denote this function as $V_{i-1}(\theta)$, which is of the form:

$$\begin{aligned} V_{i-1}(\theta) &= \{(\theta - \hat{\theta}_o)^T P_o^{-1}(\theta - \hat{\theta}_o) - \sigma_o\} \\ &\quad + \sum_{k=1}^{i-1} \alpha_k \{(y_k - x_k^T\theta)^2 - \gamma^2\} \qquad (13) \\ &= (\theta - \hat{\theta}_{i-1})^T P_{i-1}^{-1}(\theta - \hat{\theta}_{i-1}) - \sigma_{i-1} \qquad (14) \end{aligned}$$

for some appropriately chosen weighting sequence $\alpha_k \geq 0, k = 1, 2, \ldots, i-1$. The choice of the time-varying and (possibly) data-dependent weights will be explained shortly. Note that the additive constants appearing in the cost function do not affect the optimization procedure but only serve to facilitate connections to the OBE methodology. Also, the first term in (13) represents the confidence on the initial guess.

Now, consider the data set at time i, given by the pair $\{y_i, x_i\}$. The information contained by the new data, due to the knowledge of the bounded uncertainty, is characterized by: $(y_i - x_i^T\theta)^2 \leq \gamma^2 \Leftrightarrow \theta \in \mathcal{S}_i$, where θ is any estimate of θ^*. The basic idea of deriving this least-squares estimator is the following:

Given the cost function, $V_{i-1}(\theta)$ and the observation set, $\mathcal{S}_i$, the goal is to find the new estimate, say $\hat{\theta}_i$, *which minimizes* $V_{i-1}(\theta)$ *such that* $\hat{\theta}_i$ *lies in* $\mathcal{S}_i$. In other words, the strategy is to move from $\hat{\theta}_{i-1}$ to a new estimate so as to incur the least increase in the cost function $V_{i-1}(\theta)$, under the constraint that the new estimate has to lie in the observation set $\mathcal{S}_i$.

Remark 6: A very important fact can be noticed immediately, *i.e.*, that when $\hat{\theta}_{i-1} \in \mathcal{S}_i$, then there is no need to move to a new estimate as $\hat{\theta}_{i-1}$ itself is the global minimum of $V_{i-1}(\theta)$. It is pertinent to note that this results in **no update** of the parameter estimate. Moreover, the condition for no update is **exactly** the same as for the OBE, given by Proposition 1! Therefore, without any analysis, we already have established the discerning update rule for this least-squares problem and shown it to coincide with the earlier condition. Note that $\hat{\theta}_{i-1} \notin \mathcal{S}_i$ is equivalent to $|\delta_i| > \gamma$.

For further analysis, we see that this estimation strategy can easily be cast as a constrained optimization problem: At each instant i,

$$\textit{Minimize} \quad V_{i-1}(\theta) \tag{15}$$

$$\textit{Subject to:} \quad (y_i - x_i^T\theta)^2 \leq \gamma^2 \tag{16}$$

The parameter vector minimizing this problem is the new estimate, denoted by $\hat{\theta}_i$.

We first show that the solution to this problem has to lie on the *nearest bounding hyper plane of* $\mathcal{S}_i$, namely, $\mathcal{NB}_i(\hat{\theta}_{i-1})$. The problem constraint then becomes an equality, rather than an inequality as in (16). We solve the above problem by reducing it to an unconstrained optimization problem with the help of the standard Lagrange multipliers technique. We further show the same result to be obtainable by a different geometric approach. Henceforth, it is assumed that $\hat{\theta}_{i-1} \notin \mathcal{S}_i$.

Proposition 5: The new estimate at time i, *i.e.*, $\hat{\theta}_i$ belongs to $\mathcal{NB}_i(\hat{\theta}_{i-1})$.

3.2 The Solution

Given these facts, we now construct an auxiliary function,

$$J_i(\theta, \kappa_i) = V_{i-1}(\theta) + \kappa_i g_i(\theta) \tag{17}$$

$$g_i(\theta) = \{(y_i - x_i^T\theta)^2 - \gamma^2\} \tag{18}$$

where κ_i is the Lagrangian multiplier. Note that setting the gradient of $J_i(\theta, \kappa_i)$ (with respect to θ) to zero, implies that the minimizer is the parameter at which the normal vectors to the surfaces of $V_{i-1}(\theta)$ and $g_i(\theta)$ are collinear. The following solution then follows:

Theorem 1: Given data set $\{y_i, x_i\}$ and the cost function $V_{i-1}(\theta)$ at time i, the solution to the optimization problem posed in (15,16) is given by (4) and the optimal Lagrangian multiplier is $\kappa_i^* = \lambda_i^*$, where the latter is as defined by (8).

Remark 7: The first thing to notice, of course, is that the recursion for the parameter estimate has turned out to be exactly the same as in the case of the OBE algorithm. What is further appealing is that the update rule, along with the assignment to the time-varying weights is the same as before. We have the additional

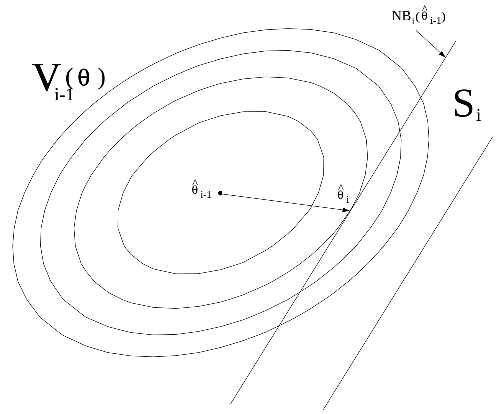

Figure 1: Contour Plot of $V_{i-1}(\theta)$ along with Geometric Representation of Updating Process

understanding that the weights are the Lagrangian multiplier to the constrained optimization problem.

Remark 8: It was not clear, using OBE arguments, as to the significance of the particular direction $P_{i-1}x_i$ to move from the earlier estimate, $\hat{\theta}_{i-1}$, in Proposition 2. It now becomes obvious that this strategy is essentially a compromise between reaching the observation set at each instant, while not increasing the extra cost incurred on the earlier error surface, $V_{i-1}(\theta)$.

A novel geometric solution to the same problem is possible by linearly transforming the parameter space so that the contours of $V_{i-1}(\theta)$ become speheroids instead of ellipsoids. By appropriately transforming $\mathcal{S}_i$ also, the solution is obtained as the projection of $\hat{\theta}_{i-1}$ on to $\mathcal{NB}_i(\hat{\theta}_{i-1})$, in the transformed space. The recursions then turn out to be the same as before. The updating process is depicted in Figure (1).

Next, we need a method for construction of $V_i(\theta)$, given knowledge of $V_{i-1}(\theta)$ and the data pair, $\{y_i, x_i\}$. The property that should be possessed by this function is that it is quadratic and of the same form as $V_{i-1}(\theta)$. Also, $\hat{\theta}_i$ should be the global minimizer of this function.

Effectively, the question is regarding the choice of the weight, $\alpha_i \geq 0$, as in (13). Therefore,

$$\begin{aligned} V_i(\theta) &= V_{i-1}(\theta) + \alpha_i\{(y_i - x_i^T\theta)^2 - \gamma^2\} \quad (19) \\ &= (\theta - \hat{\theta}_i)^T P_i^{-1}(\theta - \hat{\theta}_i) - \sigma_i \end{aligned}$$

with $\hat{\theta}_i$ given by Theorem 1. It turns out that the only assignment to α_i, which is consistent with the above formulation, is given by κ_i^* from Theorem 1, which, in turn, is the same as λ_i^*.

Proposition 6: The recursions for $V_i(\theta)$, defined via P_i and σ_i, with $\hat{\theta}_i$ given by Theorem 1, are given by (5-7) and $\alpha_i = \lambda_i^*$.

To summarize, we have formulated a least-squares cost function at each instant and obtained a solution constrained upon the need for the estimate to lie in the observation set at that instant. We have solved this problem by the Lagrange multiplier technique and also shown that the solution can be obtained by a geometrical argument. Further, we have shown a recursive method for constructing the cost function.

One of the important points that arises out of this formulation is that *for the first time, the least-squares criterion has been used to rigorously achieve "selective updating" in a recursive formulation without the use of any ellipsoidal optimality criterion.* Additional knowledge assumed of the model, in the form of a known bound on the noise process, has allowed an intelligent method to pick the weighting sequence in a recursive least-squares algorithm. To date, the choice of such weights was subject to heuristic arguments, based on experimentally observed performance. Also, by the analysis presented in this section, it now becomes amply clear as to the (least-squares) optimality criterion of the OBE algorithm under consideration.

The next section discusses some other interesting aspects of this approach, namely, explicit connections to the meaning of σ_i minimization in this Optimal Bounding Ellipsoid algorithm. In passing, we also discuss the connections to DH/OBE, which is the other algorithm in this class.

4 Discussions

In this section, we first discuss the case when the parameter is time-varying. Also, connections between σ_i minimization and the least-squares solution are explored.

One of the singular problem encountered with tracking via OBE algorithms is the case when the intersection of $\mathcal{S}_i$ with $\mathcal{E}_{i-1}$ becomes void at some time instant i. At such instants, σ_i becomes negative and the OBE algorithms have to resort to certain "rescue procedures" in order that tracking is restored [4]. However, as have been shown in the earlier section, σ_i becoming negative does not affect the updating recursions (4-7), although the OBE interpretation breaks down. In the absence of rescue procedures, other OBE estimators lose track of the parameter while such is not the case with this algorithm, which explains good tracking characteristics observed in practice. From the least-squares perspective, this fact is even more apparent. The only objective in that formulation is to optimally "move" to the observation set at the present time instant. Hence, intuitively, we see that the estimate is going to be "close" to the time-varying true parameter at all times.

However, if we wish the estimator to operate according to OBE principles, a scheme using ideas from set-membership state estimation can be implemented for tracking, thus giving a concrete method to circumvent the need for rescue mechanisms. For more details on the state estimation via set-membership principles, the reader is referred to [7].

It can be shown that the process of obtaining an extremum of $J_i(\cdot)$ with respect to κ_i, (17), in order to find the Lagrange multiplier, is the same as finding the extremum of σ_i with respect to κ_i. Moreover, with κ_i^* denoting the value of κ_i which achieves this extremum, it turns out that

$$\frac{d^2\sigma_i}{d\kappa_i^2} > 0 \quad \text{for} \quad \kappa_i = \kappa_i^*$$

which implies that κ_i^* is in fact a minimizer of σ_i.

By the above analysis, we have shown how the problem of finding the Lagrange multiplier is the same as finding the minimum of σ_i over all $\kappa_i \geq 0$. Also, the geometrical significance of σ_i is that it defines the depth of the minimum point of $V_i(\theta)$ below the θ-plane.

5 Simulation Results

We present simulation results for the following three cases:

- Identification of an 8^{th} order FIR filter under uniformly distributed noise.
- Estimation of a *real microwave radio communications channel* with truncated Gaussian noise.
- Tracking of a fast time varying 2^{nd} order FIR filter.

Note: This digital microwave radio channel was extracted from processing of field measurements of received signals.

We omit minute details of the simulation for the sake of brevity. The results are shown in Figures (2), (3) and (4). For all cases considered, the Signal to Noise Ratio (SNR) was 15 dB and the input was taken to be from a binary PAM alphabet.

We notice the similar performance of the proposed algorithm as compared to RLS for case 1. In fact, one of the simulation results showed that an order of magnitude improvement on the *mean square parameter error* was obtained over RLS. This translates to similar performance in the *mean-square prediction error*, as can be seen from Figure (1). Importantly, the number of updates was around 30 out of the total of 1000 samples. It is pertinent to note the excellent performance in spite of the sparse number of inputs used by the algorithm for updating.

For case 2, we observe that the rate of convergence of OBE is slower than RLS, though the final MSE is comparable. This hints to the known fact that OBE algorithms perform best under uniform noise. The selective updating feature may be one of the main reasons for the slower convergence.

In case 3, it can be seen that the OBE algorithm outperforms RLS (with an exponential forgetting factor of 0.9) in terms of tracking, while using only around 25 data points for updating, out of 100. The time variations in the parameter are introduced by having random, but bounded jumps at every 15 samples. This superior tracking capability is an outcome of the "intelligent" choice of weights as also the selective updating mechanism.

6 Conclusions

This paper presented a novel OBE algorithm along with a performance analysis. Results on the robustness of the algorithm to model violations as well as convergence under such circumstances were established. Further, a least-squares like criterion was used to derive the same algorithm. The analysis presented allowed for fundamental connections between OBE and least-squares estimation methods. The alternative formulation also shed light on the observed robustness to model violations and superior tracking properties of the algorithm. Simulation results for various cases, including that of estimation of a real microwave channel were presented. The excellent tracking behavior observed indicates much promise for the algorithm developed.

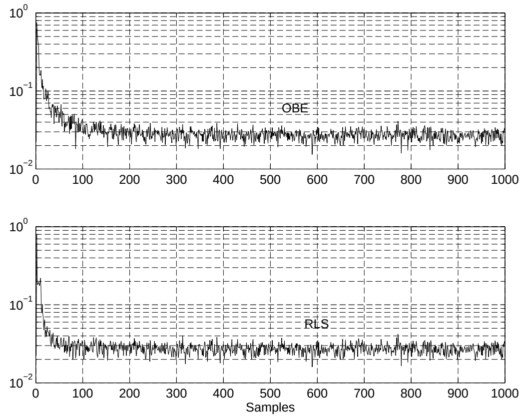

Figure 2: Estimation of an 8 Tap FIR filter with Uniform Noise

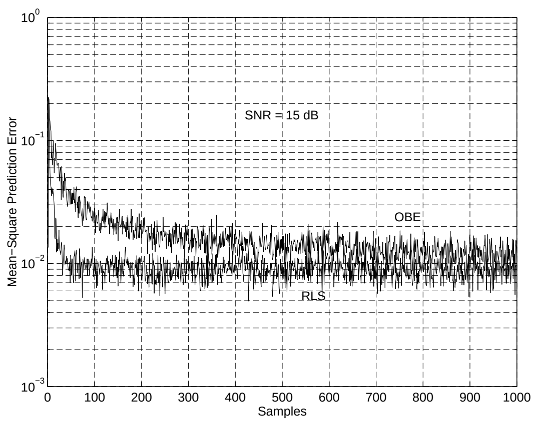

Figure 3: Estimation of a Microwave Radio Channel

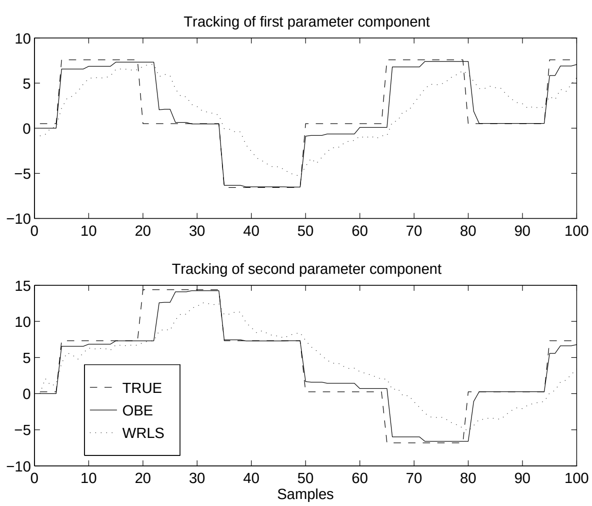

Figure 4: Tracking performance comparison

7 Acknowledgments

The authors wish to thank Prof. J. R. Deller at Michigan State University, East Lansing, MI, for useful and insightful discussions. The data used for the simulation study in case 2 was obtained from the website: *http://backhoe.ee.cornell.edu/BERG/*. We are grateful to the BERG group at Cornell and Applied Signal Technology, CA, for making the data available on the internet.

References

[1] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 1996.

[2] 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 Transactions on Information Theory*, Vol. 33, No. 3, pp. 383-392, May 1987.

[3] J. R. Deller, M. Nayeri and M. S. Liu, "Unifying the landmark developments in OBE processing," *Intl. Journal of Automatic Control and Signal Processing*, Vol. 8, pp. 43-60, Jan.-Feb. 1994.

[4] A. K. Rao and Y. F. Huang, "Tracking Characteristics of an OBE Parameter Estimation Algorithm," *IEEE Transactions on Signal Processing*, Vol. 41, No. 3, pp. 1140-1148, March 1993.

[5] 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, March 1982.

[6] M. Nayeri, M. S. Liu and J. R. Deller, "Do Interpretable Optimal Bounding Ellipsoid Algorithms Converge?: Part I and II," *IFAC Symposium on System Identification*, Copenhagen, 1994.

[7] S. Gollamudi, S. Nagaraj, S. Kapoor and Y.F. Huang, "Set Membership State Estimation with Optimal Bounding Ellipsoids," *Proc. Intl. Symp. on Information Theory and its App.*, pp. 262-265, Victoria, B.C., Canada, September 17-20, 1996.